# Viroid-Protein-Komplexe und deren Bedeutung für Replikation und Transport von Viroiden

Inaugural - Dissertation

zur

Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf vorgelegt von

# **Oliver Bannach**

aus Essen

Düsseldorf

2006

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf

Erster Berichterstatter: Prof. Dr. Detlev Riesner Zweiter Berichterstatter: Prof. Dr. Dieter Willbold Dritte Berichterstatterin: Prof. Dr. Sieglinde Ott

Tag der mündlichen Prüfung: 21. 12. 2006

# Danksagung

Mein ganz besonderer Dank gebührt Herrn Professor Detlev Riesner für die ausgezeichneten Arbeitsmöglichkeiten an seinem Institut und die mir gewährte wissenschaftliche Freiheit.

Für seine große Begeisterung für meine Arbeit und vieles mehr danke ich meinem Freund und Mentor Axel.

Vielen Dank sage ich auch Gerhard für seine geduldige Hilfsbereitschaft. Insbesondere beim Zusammenschreiben der Arbeit war er eine ganz große Hilfe.

Danke Elfi und Goldi und Ilka und Eva für einige richtig schöne Urlaube.

Elke danke ich herzlich für die nette Gesellschaft in Raum 27 und ihre tolle Unterstützung in der Schlussphase dieser Arbeit.

Ich danke vielmals Luitgard, Bernd und Schmitzi, die mir mit ihrem großen Erfahrungsschatz oft weitergeholfen haben.

Lieben Dank auch den Wagners, ganz vornan Reini, für die vielen Köstlichkeiten.

Bei allen Angehörigen des Instituts, auch den vielen Ehemaligen, möchte ich mich für die schöne Zeit bedanken.

Für ihre Liebe danke ich Silke. Ohne ihre Unterstützung hätte ich es nicht geschafft.

Schließlich danke ich meiner Mutter für alles was sie mir gab. Sie wird immer bei mir sein.

# Inhaltsverzeichnis

1.	Einl	eitung	1
	1.1	Das Viroid	2
		1.1.1 Die Entdeckung der kleinsten bekannten Parasiten	2
		1.1.2 Der Habitus	3
		1.1.3 Der Replikationszyklus	4
		1.1.4 Die funktionellen Strukturmotive für die Replikation	5
		1.1.5 Der Transport <i>in planta</i>	7
		1.1.6 Die pathogene Wirkung	8
	1.2	Die potentiellen Wechselwirkungspartner	10
		1.2.1 Das Tata Binding Protein	10
		1.2.2 Der Transkriptionsfaktor IIIA und das ribosomale Protein L5 $\dots$	10
	1.3	Aufgabenstellung	12
2.	Mat	erial	15
	2.1	Bezugsquellen	15
	2.2	Chemikalien und Lösungen	15
	2.3	Plasmide	19
3.	Met	hoden	21
	3.1	Arbeit mit Protein	21
		3.1.1 Extraktion von Gesamtprotein aus Pflanzenmaterial	21
		3.1.2 In vitro-Translationsexperimente	21
		3.1.3 Expression rekombinanter Proteine mit dem pTYB-System	22
		3.1.4 Aufreinigung rekombinater Proteine mit Intein-Fusionstag	22
		3.1.5 Expression rekombinanter Proteine mit dem pGEX-System	22

	3.1.6	Aufreinigung der GST-Fusionsproteine	23
	3.1.7	Aufkonzentrierung von Proteinlösungen	24
3.2	Arbeit	mit Nukleinsäure	25
	3.2.1	Extraktion mit Phenol und Chloroform	25
	3.2.2	Präzipitation von Nukleinsäure	25
	3.2.3	Quantifizierung von Nukleinsäuren	26
	3.2.4	Polymerase-Kettenreaktion (PCR)	27
	3.2.5	PCR mit $\alpha^{32}$ P -dCTP	29
	3.2.6	Sequenzierung von DNA	29
	3.2.7	In vitro-Transkription	29
	3.2.8	In vitro-Transkription mit $\alpha^{32}$ P -UTP	31
	3.2.9	In vitro-Synthese von mRNA	31
	3.2.10	In vitro-Synthese von infektiöser Viroid-RNA	31
	3.2.11	Isolation von mRNA	32
	3.2.12	Herstellung einer cDNA-Bank	32
	3.2.13	Infektion von Tomatenpflanzen	32
	3.2.14	Extraktion von Nukleinsäure aus Pflanzenmaterial	33
	3.2.15	Nachweis von Viroid-RNA in Pflanzenmaterial	33
3.3	Mikro	biologische Arbeiten	34
	3.3.1	Anzucht und Aufbewahrung von Bakterienkulturen	34
	3.3.2	Herstellung von kompetenten Zellen zur Elektrotransformation	34
	3.3.3	Elektrotransformation	35
	3.3.4	Single Colony PCR	35
	3.3.5	Mini-Präparation von Plasmiden	36
	3.3.6	Midi-Präparation von Plasmiden	37
	3.3.7	Maxi-Präparation von Plasmiden	37
	3.3.8	Restriktionsendonuklease-Hydrolyse	38
	3.3.9	Dephosphorylierung von Vektor-DNA	38
	3.3.10	Ligase-Reaktionen	39
	3.3.11	Klonierung in pGEM-T	39
3.4	Gelele	ktrophoresen	41
	3.4.1	Agarosegele	41

		3.4.2	Denaturierende Polyacrylamidgele	41
		3.4.3	Verzögerungsgelelektrophorese	41
		3.4.4	Bidirektionelle Gelelektrophorese	43
		3.4.5	Gelelutionen	43
		3.4.6	Silberfärbung von Polyacrylamid-Gelen	44
		3.4.7	Autoradiographie von Polyacrylamid-Gelen	44
		3.4.8	SDS-Polyacrylamidgelelektrophorese	44
		3.4.9	Färbung mit Coomassie-Blau	46
		3.4.10	Semi-Dry-Blot (Western-Blot)	46
		3.4.11	Immunologischer Proteinnachweis	47
4.	Erge	bnisse		49
	4.1	Infekt	iosität von punktmutierten Viroiden	49
		4.1.1	Synthese und Inokulation	50
		4.1.2	Wachstumskinetik der PSTVd-Mutanten	51
		4.1.3	Genetische Stabilität der PSTVd-Mutanten	52
	4.2	Analy	se von Viroid-Protein-Wechselwirkungen	54
		4.2.1	Sequenzauswahl der potentiellen Wechselwirkungspartner	54
		4.2.2	Identifikation von TFIIIA und L5 aus <i>L. esculentum</i>	55
		4.2.3	Expression rekombinanter Proteine in <i>E. coli</i>	59
		4.2.4	Beschreibung des Expressionssystems	59
		4.2.5	Klonierung von <i>St</i> TBP in pTYB	60
		4.2.6	Klonierung von AtTFIIIA und AtL5 in pTYB	60
		4.2.7	Überexpression und Aufreinigung mit dem pTYB-System	61
		4.2.8	Klonierung von TFIIIA, L5 und TBP in pGEX	63
		4.2.9	Überexpression und Aufreinigung mit dem pGEX-System	64
		4.2.10	Zusammenfassung zur rekombinaten Proteinexpression	68
		4.2.11	Nukleinsäurebindung der rekombinanten Proteine	69
		4.2.12	Analyse der Nukleinsäurebindeeigenschaften von TFIIIA	73
		4.2.13	Analyse der Nukleinsäurebindeeigenschaften des ribosomalen	
			Proteins L5	80

	4.3	Alterr	nativ gespleißte TFIIIA-mRNA: Ein konserviertes Phänomen $in$	
		planta	l	
		4.3.1	Sequenzanalyse des alternativen Produktes	
		4.3.2	Vergleich mit anderen pflanzlichen TFIIIA-Sequenzen	
		4.3.3	In vitro-Translation der alternativ gespleißten TFIIIA-mRNA .	
		4.3.4	Bindestudien mit rekombinantem kTFIIIA	
		4.3.5	In vivo-Nachweis von kTFIIIA mittels Immunoblot	
		4.3.6	Alternativ gespleißte TFIIIA-mRNA: Zusammenfassung der	
			Ergebnisse	
5.	Disk	cussion		
	5.1	Die St	artstelle der PSTVd-Transkription	
		5.1.1	Struktur der Startdomäne	
		5.1.2	Reversionsmechanismus mutierter Viroid-Nukleotide	
	5.2	Viroid	l-Protein-Komplexe	
		5.2.1	Methodische Gesichtspunkte	
		5.2.2	Nukleinsäurebindeeigenschaften von <i>Le</i> TFIIIA	1
		5.2.3	Die Fehlleitung von Wirtsproteinen durch das Viroid	1
	5.3	Die A	lternative Spleißform der TFIIIA-mRNA	1
	5.4	Ausbl	ick	1
ß	7.00	mmonf	a company	1
υ.	<b>z</b> iusa	annneill	assung	1
	Lite	raturve	rzeichnis	1

# Abbildungsverzeichnis

1.1	Struktur von PSTVd.	3
1.2	Asymmetrischer <i>rolling-circle</i> -Mechanismus der <i>Pospiviroidae</i>	6
1.3	Loop E-TFIIIA-Wechselwirkung.	11
2.1	Vektorkarte, pTYB1 (NEB).	19
2.2	Vektorkarte, pGEX-4T-3 (Amersham).	20
4.1	Position der Mutationen im linken terminalen Loop von PSTVd. $\ldots$	50
4.2	Nachweis der G359-Nachkommenschaft, 28 Tage $post\ infectum.$	51
4.3	Infektiosität der Mutante G359.	52
4.4	Infektiosität der Mutante G1.	53
4.5	Identifikation eines $Le$ TFIIIA-ORF	56
4.6	Alignment von TFIIIA-Sequenzen pflanzlichen und tierischen Ursprungs.	57
4.7	Alignment von Sequenzen des ribosomalen Proteins L5 aus pflanzlichen	
	und tierischen Organismen.	58
4.8	Aufreinigung der Proteine $At$ TFIIIA, $At$ L5 und $St$ TBP mit	
	C-terminalem Fusions tag über eine Affinitätschromatographie	62
4.9	Rekombinante Über expression der GST-Fusionsproteine in $E.~coli$ am	
	Beispiel von $GST$ - $LeL5$ .	66
4.10	Einfluss von Sarkosyl auf die Reinigung von GST- <i>Le</i> L5.	67
4.11	Konzentrationsbestimmung des gereinigten GST-LeTFIIIA nach Dialyse.	68
4.12	In die Gelverzögerungsexperimente eingesetzte Nukleinsäurespezies.	71
4.13	EMSA zur Analyse der Bindung von TFIIIA an 5S-DNA unter	
	verschiedenen Pufferbedingungen.	72
4.14	EMSA zur Bestimmung des apparenten $K_{\rm d}$ -Wertes für die Bindung von	
	TFIIIA an 5S-DNA.	73

4.15	EMSA zur Bestimmung des Anteils an bindeaktivem TFIIIA. $\ .\ .$ .	75
4.16	Bindung von TFIIIA an 5S-DNA und an unspezifische DNA ohne	
	$MgSO_4$ im Gel.	76
4.17	EMSA zur Bestimmung des apparenten $K_{\rm d}$ -Wertes für die Bindung von	
	TFIIIA an RNA.	77
4.18	Bindung von TFIIIA an unspezifische, einzelsträngige RNA.	78
4.19	Bindung von TFIIIA an 5S-RNA und dsRNA. $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$	79
4.20	Bindung von L5 an RNA.	81
4.21	TFIIIA, Amplifikation von cDNA.	82
4.22	Intron/Exon-Struktur des TFIIIA-Gens (Zugriffsnr. DQ882180).	83
4.23	In vitro-Translation von TFIIIA-mRNA <sup>alt</sup> .	86
4.24	In vitro-Translation von TFIIIA-mRNA und TFIIIA-mRNA $^{\rm alt}.$	87
4.25	Rekombinante Expression von GST-kTFIIIA und dessen	
	Nukleinsäurebindeeigenschaften.	88
4.26	Nachweis von TFIIIA und kTFIIIA in vivo mit Antiserum 24	89
4.27	Nachweis von TFIIIA und kTFIIIA <i>in vivo</i> mit Antiserum 23	90
<b>F</b> 1	NMD Staultur des UUCC Tetrologne	0.4
<b>J.1</b>	NMR-Struktur des UUCG-Tetraloops.	94
5.2	Modell zum Reversionsmechanismus von PSTVd-Mutationen.	97
5.3	Modell für den intrazellulären Transport von PSTVd.	101

# Tabellenverzeichnis

1.1	Übersicht über die bekannten zellulären Wechselwirkungspartner des	
	Viroids.	9
2.1	DNA-Oligonukleotide.	18
4.1	Übersicht über die Molekulargewichte der rekombinanten Proteine	61
4.2	Dissoziationskonstanten für die Nukleinsäurebindung von TFIIIA	80
4.3	TFIIIA, Zusammenfassung der Datenbankzugriffsnummern.	84

 $,, Post\ experimentum\ ante\ experimentum\ est.``$ 

- frei nach Sepp Herberger

# 1

# Einleitung

Zu Beginn der fünfziger Jahre klärten Watson und Crick die Doppelhelixstrukur der Desoxyribonukleinsäure (DNA) auf und ebneten den Weg für weitere Pionierleistungen auf dem neuen Feld der Molekularbiologie. Der genetische Code wurde entschlüsselt und der Fluss der genetischen Information von der DNA zum Protein als zentrales Dogma der Biologie etabliert.

Die Rolle von Ribonukleinsäure (RNA) im biologischen System wurde indes lange Zeit unterschätzt. So wurde *transfer*-RNA (tRNA) als Transportmolekül, *messenger*-RNA (mRNA) als passiver Informationsträger und ribosomale RNA (rRNA) lediglich als Gerüstsubstanz für Proteine verstanden.

Erst Anfang der 80er Jahre begann man die faszinierende Struktur- und Funktionsvielfalt der RNA zu erhellen. Angefangen mit der Entdeckung selbstspaltender RNAs, den Ribozymen, bis hin zur Identifikation winziger, regulatorischer *micro*-RNAs erschlossen sich der Wissenschaft ganz neue Einblicke in die molekularen Prozesse der Zelle.

RNA ähnelt chemisch durchaus der DNA und vermag ebenso wie diese genetische Information zu tragen; *in puncto* Struktur- und Funktionsdiversität hingegen erinnert RNA eher an die Mannigfaltigkeit der Proteine. So verwundert es nicht, dass RNA auch proteintypische, enzymatische Aufgaben übernehmen kann: RNA katalysiert die Bildung der Peptidbindung während der ribosomalen Proteinbiosynthese und ist maßgeblich an der Prozessierung von Vorläufer-mRNA in den Splicosomen beteiligt.

Mit diesem Wissen um die pluripotenten RNA-Funktionen wurde ferner die Hypothese formuliert, wonach zu Beginn des Lebens auf der Erde selbstreplizierende RNA- Moleküle existierten, aus denen sich durch molekulare Evolution komplexere Systeme entwickelten. Ein rezentes Fossil dieser urtümlichen RNA-Welt könnte eine überaus besondere RNA-Spezies verkörpern: Das Viroid.

# 1.1 Das Viroid

# 1.1.1 Die Entdeckung der kleinsten bekannten Parasiten

In den 60er Jahren versuchte Theodor Diener den Erreger einer schweren Pflanzenkrankheit, namentlich der Kartoffelspindelknollensucht, zu identifizieren. Interessanterweise konnte die Infektiosität von Extrakten aus erkrankten Pflanzen nicht durch Behandlung mit Phenol oder Protease, wohl aber durch Ribonuklease aufgehoben werden. Erreger bakteriellen oder viralen Ursprungs schieden somit aus. Auf Grundlage dieses und nachfolgenden Ergebnissen formulierte Diener, dass es sich bei dem infektiösen Agens um eine nackte RNA mit sehr niedrigem Molekulargewicht handelte. Er nannte diesen neuen Erregertypus Viroid, also virusähnlich (zur Übersicht: Diener, 2003; Riesner & Gross, 1985).

Es stellte sich heraus, dass sich die Ähnlichkeit zu Viren lediglich in der Symptomausprägung manifestiert: Kranke Pflanzen sind von kleinem Wuchs durch gestauchte Internodien. Zudem können Blattmissbildungen (Epinastien), Chlorosen oder gar Nekrosen beobachtet werden. Von Viroiden befallen werden Kulturpflanzen, darunter viele Nachtschattengewächse, Citrusgewächse, Chrysanthemen, Hopfen, Avocado, Gurke, Apfel oder Wein. Bemerkenswerterweise hängt das Ausmaß der pathogenen Effekte gleichermaßen von der Varietät der infizierten Pflanzenart als auch von der genauen Viroidsequenz ab. Natürlich infizierte Wildtyp-Pflanzen reagieren für gewöhnlich asymptomatisch und können so als natürliches Reservoir dienen. Daher liegt die Vermutung nahe, dass Viroide in der Pflanzenwelt stark verbreitet sein könnten. Bis *dato* sind annähernd 40 Viroid-Spezies mit einer Vielzahl von Sequenzvarianten charakterisiert worden. Diese werden in zwei Familien eingeordnet, die *Avsunviroidae* und die *Pospiviroidae*, von denen letztere mit dem Potato Spindle Tuber Viroid (PSTVd) den prominentesten Vertreter und auch das Untersuchungsobjekt der vorliegenden Arbeit stellen.



Abbildung 1.1: Struktur von PSTVd. Die fünf funktionell relevanten Domänen sind mit  $T_L$  und  $T_R$  (linker und rechter Terminus), VM (Virulenz-modulierende Region), CCR (zentrale konservierte Region) und VR (variable Region) gekennzeichnet. Basenpaarungen sind durch Punkte markiert. Sequenzen, die alternative Strukturen (Hairpin I und II) bilden können, sind hellgrau dargestellt.

# 1.1.2 Der Habitus

Hinsichtlich ihrer molekularen Physiologie lassen sich die Viroide sehr deutlich von den Viren abgrenzen. Viroide sind 246 bis 401 nt lange RNA-Einzelstränge, die zirkulär geschlossen sind. Das Fehlen offener Enden verleiht der Viroid-RNA eine gewisse Resistenz gegen RNasen. Im Gegensatz zu Viren kodiert das Viroidgenom aber weder ein Protein, noch ist es in einer proteinösen Hülle verpackt. Daraus folgt wiederum, dass das Viroid hinsichtlich seiner Vermehrung und seines Transports durch die Pflanze vollständig von Faktoren seines Wirts abhängig ist.

Begünstigt durch ein hohes Maß an Selbstkomplementarität nimmt die PSTVd-RNA unter nativen Bedingungen eine stäbchenförmige, unverzweigte Konformation ein (siehe Abb. 1.1). Dabei wechseln durchgängig basengepaarte und ungepaarte Bereiche (sogenannte *loops*) ab. Sequenzhomologien innerhalb der *Pospiviroidae* erlauben eine Unterteilung des Moleküls in verschiedene Domänen (Keese & Symons, 1985). Die zentrale konservierte Region (CCR) ist durch ihren GC-Reichtum thermodynamisch sehr stabil. Sie ist an einer Vielzahl von biologischen Funktionen des Viroids beteiligt. Rechts schließt sich die wenig konservierte, variable Region (VR) an, während linksseitig die Virulenzmodulierende Region (VM) liegt. Punktmutationen im letzt genannten Abschnitt führen zu den stammspezifischen Symptomatiken. Begrenzt wird das Viroid von zwei terminalen *loop*-Domänen (T<sub>L</sub> und T<sub>R</sub>).

Verzweigungen oder übergeordnete Tertiärstrukturen konnten für das PSTVd ausgeschlossen werden. Dennoch scheinen einige ungepaarte Abschnitte keinesfalls unstrukturierter Natur zu sein. Vielmehr existieren Sequenzabschnitte aus nicht-kanonischen Basenpaaren, die komplexe Strukturen ausbilden können. Ein derartiges Motiv repräsentiert beispielsweise der zentral gelegene Loop E, der große Ähnlichkeit mit einem Loop der 5 S-RNA oder dem Sarcin/Ricin-Loop in der 23 S-RNA besitzt.

Es sind wahrscheinlich solche lokalen Strukturelemente, vermittels derer das Viroid mit seiner molekularen Umwelt zu kommunizieren vermag. Die Sekundärstruktur des Viroids ist also Teil dessen genetischer Information. Die primäre Sequenz scheint im Gegensatz zu Protein-kodierenden RNAs eine eher untergeordnete Rolle zu spielen. So werden Punktmutationen durch kompensatorische Mutationen, die helikale Strukturelemente wiederherstellen, toleriert (Loss *et al.*, 1991).

Darüber hinaus kann das Viroid alternative, sogenannte metastabile Strukuren ausbilden. Diese sind nicht Bestandteil der nativen Stäbchenform, sondern stellen Übergangskonformationen bei der Denaturierung des Moleküls dar. *In vivo* werden solche Strukturen bei der sequentiellen Faltung des naszierenden Transkriptes gebildet. Metastabile RNA-Strukturen befinden sich – kinetisch begünstigt – in Zuständen lokaler Energieminima, sie bedürfen Aktivierungsenergie um sich in die thermodynamisch optimale Konformation umzulagern.

Auf diese Weise speichert das Viroid mehrere sekundäre Strukturinformationen in einer einzigen primären Sequenz. Ein ähnliches Phänomen war bislang nur bei einigen genomischen DNA-Sequenzen bekannt, die in überlappenden Leserastern bzw. über ihren Gegenstrang für zwei Proteine kodieren. Die funktionelle Bedeutung der verschiedenen Strukturmotive des Viroids wird in den folgenden Kapiteln beleuchtet.

#### 1.1.3 Der Replikationszyklus

Die Pospiviroidae werden von der wirtseigenen RNA Polymerase II (Pol II) im Zellkern transkribiert. Das erstaunliche Ergebnis, dass eine DNA-abhängige Polymerase ein RNAtemplate akzeptieren kann, fanden Mühlbach & Sänger (1979) in Replikationsstudien mit dem Pol II-spezifischen Hemmstoff  $\alpha$ -Amanitin. Auch die in Chloroplasten lokalisierten Avsunviroidae werden von einem DNA-abhängigen Enzym transkribiert, der plastidären RNA Polymerase. In vitro gelang gar eine Transkription des Peach Latent Mosaic Viroid (PLMVd) durch E. coli-Polymerase, für die man eine Homologie zu der Chloroplasten-Polymerase nahelegt (Pelchat et al., 2001). Die Replikation des PSTVd folgt einem asymmetrischen *rolling-circle*-Mechanismus (Branch & Robertson, 1981, 1984; Feldstein *et al.*, 1998). Dabei umläuft die Polymerase mehrmals das zirkuläre Molekül und schreibt es dabei in einen multimeren Strang entgegengesetzter Polarität um (siehe Abb. 1.2). Dieser *per definitionem* (-)-Strang dient wiederum als *template* für die Synthese eines multimeren (+)-Stranges, welcher schließlich in Monomere exakter Einheitslänge geschnitten und zu reifen Zirkeln ligiert wird. Dieser in Analogie zur Maturierung von Vorläufer-mRNA als Prozessierung bezeichnete Schritt wird von zwei unterschiedlichen Wirtsenzymen katalysiert, deren genaue Identitäten bisher nicht geklärt werden konnten (Klümper, 2002).

Nach einem grundsätzlich anderen Mechanismus werden Viroide der Familie der Avsunviroidae repliziert (zur Übersicht: Daròs et al., 2006). Da hier zirkuläre Replikationsintermediate zu beobachten sind, spricht man von einem symmetrischen rolling-circle-Mechanismus (Daròs et al., 1994; Hutchins et al., 1986). Ferner bilden Avsunviroidae Ribozymstrukturen aus und können autokatalytisch prozessieren, d.h. ohne Beteiligung von Wirtsfaktoren. Für das chloroplastidäre Protein PARBP33 konnte aber gezeigt werden, dass es als RNA-Chaparone mit dem Avocado Sun Blotch Viroid (ASBVd) wechselwirkt und so dessen Selbstprozessierung stimuliert (Daròs & Flores, 2002).

## 1.1.4 Die funktionellen Strukturmotive für die Replikation

## Transkription des (+)-Stranges

Die Transkription des (+)-strängigen Zirkels beginnt am linken terminalen *loop*. Dies konnte aus Primer-Extension-Analysen an aufgereinigten Replikationsintermediaten abgeleitet werden (Kolonko, 2003). Die Existenz einer einheitlichen Startstelle der Transkription legt die Vermutung nahe, dass diese von einem Promotorelement auf dem Viroid festgelegt wird. Schon in früheren Arbeiten zur Replikation von PSTVd wurden GC-reiche Abschnitte (sogenannte GC-Boxen) als potentielle Promotoren diskutiert (Heinrich, 1991; Loss *et al.*, 1991; Qu *et al.*, 1993). Für PLMVd konnte ein Promotorelement in Form einer RNA-Hairpinstruktur charakterisiert werden. Dieses scheint essentiell für die spezifische Transkriptionsinitiation durch eine DNA-abhängige Polymerase (Pelchat *et al.*, 2002).



Abbildung 1.2: Asymmetrischer rolling-circle-Mechanismus der Pospiviroidae. Das per definitionem (+)-strängige PSTVd-Molekül wird von der wirtseigenen RNA-Polymerase II (Pol II) in einen multimeren (-)-Strang transkribiert. Dieser dient wiederum als template für die Synthese eines multimeren (+)-Stranges. Durch Prozessierung und Ligation, vermittelt von Wirtsfaktoren (RNase und Ligase), werden schließlich zirkuläre PSTVd-Moleküle maturiert.

### Transkription des (-)-Stranges

Bei der thermischen Denaturierung der PSTVd-RNA rearrangieren sich bestimmte Sequenzbereiche und formen drei charkteristische Haarnadelstrukturen, genannt HP(hairpin)I-III. Diese metastabilen Strukturelemente könnten auch *in vivo* kotranskriptionell bei der Synthese des (-)-Stranges gebildet werden (Loss *et al.*, 1991). Durch gerichtete Mutagenese des HPII wurde diesem ferner eine funktionelle Bedeutung bei der Replikation zugesprochen: Mutationen im Kernbereich des HPII revertierten zur Wildtypsequenz bzw. wurden durch den Einbau kompensatorischer Mutationen ausgeglichen. Mutationen in der Peripherie des HPII hingegen erwiesen sich als genetisch stabil (Loss *et al.*, 1991; Qu *et al.*, 1993). Die Existenz des HPII konnte sowohl *in silico* durch Simulation kinetischer Faltungswege als auch experimentell mit HPII-spezifischen Sonden belegt werden (Repsilber *et al.*, 1999; Schmitz & Steger, 1996; Schröder & Riesner, 2002).

Der GC-Gehalt der HPII-Helix erinnert an eukaryotische GC-Boxen, die als Promotoren von konstitutiv exprimierten Genen dienen. Ferner zeigt die Helixgeometrie von kurzen, GC-reichen DNA-Oligonukleotiden eine RNA-typische A-Form (Heinemann *et al.*, 1987). Diese Tatsache, dass sich DNA und RNA strukturell ähneln können, sollte dem Viroid die Mimikry eines DNA-Promotors erleichtern. Dadurch könnten eukaryotische Faktoren für die Transkriptionsinitiation auf dem (-)-Strang rekrutiert werden.

#### **Prozessierung und Ligation**

Bei der Prozessierung des multimeren (+)-Stranges spielen ebenfalls metastabile Strukturen und deren Erkennung durch Wirtsproteine eine essentielle Rolle. So konnte gezeigt werden, dass ausschließlich eine in der CCR verzweigte, metastabile Struktur nukleolytisch geschnitten wird. Als Erkennungsmotiv für die Wirtsnuklease wird ein Tetraloop vom Typ GNRA diskutiert. Die thermodynamisch optimale Konformation der CCR stellt schließlich das Substrat für die Ligationsaktivität dar (Baumstark & Riesner, 1995a; Baumstark *et al.*, 1997). Eine Minimal-RNA, die lediglich aus der postulierten Prozessierungsdomäne von PSTVd besteht, wird in einem Kernextrakt aus Kartoffelzellen ebenfalls korrekt geschnitten und zu Zirkeln ligiert (Schrader *et al.*, 2003). Das Viroid komprimiert demnach – wie oben erwähnt – die Information für die Ausbildung zweier funktionelle Strukturmuster in einer einzigen Primärsequenz.

### 1.1.5 Der Transport in planta

Das Viroid gelangt über Verletzungen an der Pflanzenoberfläche in die Wirtszelle. Die Infektion verläuft in der Regel systemisch, d.h. es werden nahezu alle Pflanzenteile befallen. Dabei ist der Weg durch die Pflanze keineswegs als passiver Diffusionsprozess zu verstehen. Es wird vielmehr ein spezifischer, von zellulären Faktoren vermittelter Transportprozess angenommen (zur Übersicht: Ding *et al.*, 2005).

Zunächst muss das Viroid zum Ort seiner Replikation, dem Zellkern, gelangen. Mit großer Wahrscheinlichkeit wird das Viroid in einem Ribonukleoprotein-Komplex (RNP) spezifisch in den Kern geleitet. Offenbar besitzen die *Pospiviroidae* die dafür notwendigen strukturellen Signale. Andere RNA-Fragmente ähnlicher Größe oder die chloroplastenständigen *Avsunviroidae* werden indes nicht in den Kern importiert (Woo *et al.*, 1999). Als Vermittler des intrazellulären Transports diskutierten Martínez de Alba *et al.* (2003) ein Viroid-bindendes Protein namens VirP1. Dieses Protein besitzt ein Kernlokalisierungssignal (NLS) und bindet PSTVd an einen asymmetrischen internen *loop*, dem sogenannten RY-Motif. Mutationen in diesem Abschnitt zerstören die Infektiosität und bestätigen somit dessen biologische Relevanz.

Auf subzellulärer Ebene existieren weitere sehr spezifische Transportprozesse. So akkumuliert PSTVd im Nukleolus zu einem beachtlichen Titer von  $10^4$  bis  $10^5$  Molekülen, wobei allerdings nur jeder vierte bis fünfte Nukleolus befallen ist (Harders *et al.*, 1989; Schumacher *et al.*, 1983b). Aus Bindestudien von Viroid an rekombinantes Nukleolin, dem abundantesten Protein im Nukleolus, folgerte Schmitz (2003), dass 99% der Viroid-RNA mit Nukleolin komplexiert ist. Durch diese Art der Maskierung einer schädlichen RNA könnten sich asymptomatische Pflanzen der pathogenen Wirkung entziehen. Funktionell bedeutsam für den intrazellulären Transport könnten überdies Komplexe aus PSTVd und 5 S-RNA sein. Eine Bindung des PSVTd an diese kleine ribosomale RNA konnte *in vivo* nachgewiesen und *in vitro* rekonstituiert werden (Aschermann, 2003).

Weite Strecken innerhalb der Pflanze kann das Viroid über das Phloem zurücklegen. Dabei folgt es dem Strom der Photoassimilate von den photosynthetisch aktiven Quellorganen hin zu den jungen Trieben an der Blattspitze (Palukaitis, 1987; Stark-Lorenzen *et al.*, 1997). Es handelt sich aber auch hierbei keinewegs um einen passiven Transport. Viemehr wird er von Wirtsfaktoren aktiv vermittelt. *In vitro* und *in vivo* wurde ein Ribonukleoprotein-Komplex zwischen dem Hop Stunt Viroid (HSVd) und dem Phloem Protein 2 (CsPP2) aus Gurke identifiziert (Gómez & Pallás, 2001, 2004; Owens *et al.*, 2001). Dieses Protein wechselwirkt ferner mit den Plasmodesmata und kann so den interzellulären Transport des Viroids vermitteln.

## 1.1.6 Die pathogene Wirkung

Die molekularen Grundlagen des Pathogenitätsmechanismus von PSTVd sind weitgehend unverstanden. Es ist unklar, warum manche Varietäten ein und derselben Wirtsart bei Viroid-Befall deutliche Krankheitssymptome entwickeln, während andere Varietäten völlig symptomlos reagieren. Ebenso weiß man nicht, warum einzelne Mutationen in der VM-Region des Viroids die Symptomausprägung drastisch modulieren. Da die Viroid-RNA nicht kodogen ist, scheiden Pathogenitätsfaktoren als Verursacher der Krankheitssymptome aus. Vielmehr wird die schädigende Wirkung darauf zurückgeführt, dass durch Wechselwirkung mit dem Viroid essentielle Wirtsfaktoren entzogen werden. Diese können gleichsam proteinöse Faktoren und RNA sein. Es wurde diskutiert, dass durch Basen-

9

Tabelle 1.1: Übersicht über die bekannten zellulären Wechselwirkungspartner des Viroids. Auflistung von Proteinen, denen verschiedene Funktionen bei Replikation und Transport des Viroids zugesprochen wurden. Als einzig bekannte interagierende Nukleinsäurespezies wurde die 5 S-RNA beschrieben.

Faktor	System	Funktion	Zitat	Kap.
Polymerase II	PSTVd	Transkription	Mühlbach & Sänger (1979)	1.1.3
plastidäre RNA-Polymerase	ASBVd	Transkription	Navarro et al. (2000)	1.1.3
PARBP 33, PARBP 35	ASBVd	Prozessierung	Daròs & Flores (2002)	1.1.3
CsPP2	HSVd	Transport	Gómez & Pallás (2004)	1.1.5
VirP1	PSVTd	Transport	Gozmanova et al. (2003)	1.1.5
43 kDa Histon-Homolog	PSTVd	unbekannt	Klaff <i>et al.</i> (1989)	5.2.1
Proteinkinase PKR	PSTVd	Pathogenese	Hiddinga et al. (1988)	5.2.1
5S ribosomale RNA	PSTVd	Transport	Aschermann (2003)	1.1.5

paar-Interaktionen rRNA-Reifung, mRNA-Spleissen oder der Einbau der 7 S-RNA in das Signal Recognition Particle (SRP) beeinträchtigt sein könnte (Diener, 2001).

Neuste Modelle zur Pathogenität von PSTVd basieren auf pflanzlichen RNA-Interference-Mechanismen (Papaefthimiou et al., 2001; Schmitz, 2003). Dabei wird doppelsträngige RNA, wie sie etwa im Lebenszyklus mancher Viren auftritt, von einer Wirtsnuklease namens Dicer zu kleinen RNAs prozessiert, den sogenannten small interfering RNAs (siRNAs). Diese siRNAs binden an die Virus-RNA und determinieren so deren Degradation. Neben dieser Schutzfunktion kann die Pflanze mit einem ähnlichen Mechanismus auch die eigene Genaktivität regulieren. Genomisch kodierte *micro*-RNA (miRNA) wird aus einem Vorläufermolekül gebildet, das eine unverzweigte Haarnadelstruktur aufweist. Dieser Hairpin wird von Dicer prozessiert, wobei eine miRNA entlassen wird. Diese miRNA kann mit komplementärer mRNA hybridisieren und deren ribosomale Translation inhibieren. Bemerkenswerterweise zeigen Sequenzabschnitte aus der Pathogenitätsdomäne von PSTVd besonders deutliche Homologien zu einigen pflanzlichen mRNA-Sequenzen. Da die endständigen Domänen des Viroids den pflanzlichen miRNA-Vorläufermolekülen ähneln, könnte das Viroid gemäß Modellvorstellung von Dicer akzeptiert und geschnitten werden. Die auf diese Weise freigesetzte siRNA mit PSTVd-Sequenz könnte aufgrund ihrer Komplementarität die Degradation von Wirts-mRNA einleiten oder deren Translation hemmen. Somit wäre der Wirtsmetabolismus empfindlich gestört und dadurch der pathogene Effekt der Viroid-RNA zu erklären.

# 1.2 Die potentiellen Wechselwirkungspartner

Seit Beginn der Viroidforschung stellt sich die Frage nach den Wirtsfaktoren, die mit dem Viroid wechselwirken. Wie aus den vorangegangen Kapiteln ersichtlich, kennt man die Identität und Funktion erst weniger solcher Faktoren (siehe Tab. 1.1). Des Weiteren gibt es eine Reihe von Wirtsfaktoren, für die eine Funktion für Replikation und Transport des Viroids postuliert werden kann. Im Folgenden werden drei dieser potentiellen Wechselwirkungspartner vorgestellt.

### 1.2.1 Das Tata Binding Protein

Tata Binding Protein (TBP) ist die DNA-Bindedomäne des Transkriptionsfaktors IID (TFIID). TBP bindet an TATA-Boxen eukaryotischer PolII-Promotoren und liefert so das Fundament für den Aufbau des Initiationskomplexes aus weiteren Faktoren. Dieser determiniert schließlich den exakten Beginn der Transkription.

Auch die Pol II-vermittelte Transkription der Viroid-RNA beginnt an einer definierten Stelle, dem linken terminalen *loop* (Kolonko, 2003). Demnach liegt es nahe, dass die Viroid-RNA ein Strukturmotiv besitzt, welches als Pol II-Promotor wirken könnte. Tatsächlich konnte bereits gezeigt werden, dass ein RNA-Konstrukt, welches eine Domäne nahe des linken Terminus von PSTVd umfasst, an humanes TBP bindet (Gande, 2002). Die Dissoziationskonstante dieser Wechselwirkung wurde im submikromolaren Bereich abgeschätzt. Daher scheint eine artifizielle, unspezifische Wechselwirkung unwahrscheinlich, aber für eine detaillierte Hypothese wäre eine Analyse eines homologen Systems nötig.

## 1.2.2 Der Transkriptionsfaktor IIIA und das ribosomale Protein L5

Die Funktion des Transkriptionsfaktors IIIA (TFIIIA) besteht unter anderem darin, die 5S-RNA-Synthese zu regulieren (Engelke *et al.*, 1980; Pelham & Brown, 1980). Dieses Protein besteht aus neun Zink-Finger Bindemotiven vom Typ Cys<sub>2</sub>His<sub>2</sub>, mit denen es die interne Kontrollregion (ICR) des 5S-DNA-Promotors binden kann. Erstaunlicherweise bindet TFIIIA auch das Transkript, die 5S-RNA. Auf Basis dieses Befundes wurde eine Autoregulation der 5S-RNA-Transkription in Form eines negativen Rückkopplungsmechanismus postuliert. Darüber hinaus ist TFIIIA an intrazellulären RNA-Transportprozessen beteiligt. Im Zytosol von *Xenopus*-Oozyten findet man Spei-



**Abbildung 1.3: Loop E-TFIIIA-Wechselwirkung.** Röntgenkristallstruktur des Loop E der 5 S-RNA im Komplex mit Zink Finger 4 von TFIIIA aus *Xenopus laevis*. RNA-Nukleotide sind in violett und blau, Peptidketten in gelb dargestellt. Wasserstoffbrücken sind rot gekennzeichnet, Basenpaar-*stacking* grün. Das exponierte Nukleotid 75G wechselwirkt über Wasserstoffbrücken mit den Aminosäuren Asp120, His119 und Lys118 (Röntgenkristallstruktur aus Lu *et al.* (2003)).

cherpartikel in Form von 7 S-RNPs, die sich aus TFIIIA und 5 S-RNA zusammensetzen (Guddat *et al.*, 1990; Pieler & Rudt, 1997). Ebenso in den nukleozytoplasmatischen Transport involviert ist das ribosomale Protein L5. Es aggregiert mit der 5 S-RNA zu einem 5 S-RNP, wird so in den Nukleolus geleitet und dort in die große Untereinheit des Ribosoms eingebaut (Michael & Dreyfuss, 1996). 3D-Strukturdaten zeigten, dass die TFIIIA/5 S-RNA- und die TFIIIA/DNA-Wechselwirkungen einem grundsätzlich anderen Bindungsmodus unterliegen (Lu *et al.*, 2003; Vitolo *et al.*, 2004). Die B-förmige Geometrie der DNA erlaubt eine spezifische Promotorerkennung durch TFIIIA, wobei dessen Zink-Finger in der großen Furche der DNA mit den Basen wechselwirken. Dagegen wird die 5 S-RNA über ihre Tertiärstruktur erkannt. So bilden sich beispielsweise zwischen der Spitze des vierten Zink-Fingers des TFIIIA und dem Loop E der 5 S-RNA Wasserstoffbrücken und *stackinq*-Wechselwirkungen aus (siehe Abb. 1.3).

Die CCR der *Pospiviroidae* zeigt signifikante Homologien zur 5 S-RNA aus *Solanum tuberosum*. Neben dem Loop E ähneln sich auch die Sequenzen der angrenzenden Helices von 5 S-RNA und PSTVd-CCR (Schrader *et al.*, 2003). Legt man einen gemeinsamen evolutionären Ursprung zu Grunde, so ist es durchaus vorstellbar, dass beide RNAs denselben intrazellulären Transportmechanismus nutzen.

Ferner ist es denkbar, dass die Wechselwirkung des Viroids mit einem Transportprotein durch die 5 S-RNA vermittelt wird. Somit könnte das Viroid in einem ternären Verbund – etwa gemeinsam mit der 5 S-RNA und dem ribosomalen Protein L5 – innerhalb der Zelle passagiert werden.

# 1.3 Aufgabenstellung

Viroide sind nicht kodogen und müssen daher mit Proteinen der Wirtspflanze interagieren, um ihre Population erfolgreich aufrechtzuhalten. Dies betrifft alle oben genannten Aspekte ihrer biologischen Funktionen: die Replikation, die Prozessierung und den Transport *in planta*. Durch molekulare Mimikry gelingt es der Viroid-RNA Wirtsfaktoren für den eigenen Bedarf fehlzuleiten. Der Entzug essentieller Faktoren durch Bindung an Viroid-RNA könnte ferner den pathogenen Effekt erklären.

Trotz intensiver Forschung sind bislang nur wenige Wechselwirkungspartner mit empirischen Methoden identifiziert und funktionell charakterisiert worden (siehe Tab. 1.1). In der vorliegenden Arbeit wird ein komplementärer hypothetischer Ansatz verfolgt. Basierend auf Strukturanalysen des Viroids wurden zuvor potentielle Wechselwirkungspartner diskutiert und in funktionelle Modelle integriert. Ziel dieser Arbeit war es, einige dieser Faktoren rekombinant herzustellen und hinsichtlich ihrer Interaktion mit PSTVd zu analysieren. Mit Kenntnis der Bindungsaffinität der rekombinanten Proteine an das Viroid sollte es möglich sein, die gegenwärtigen Modellsysteme zu den Bereichen Transport und zur Replikation von PSTVd zu überprüfen oder zu verwerfen.

Für das TATA-Binding-Protein wird eine wesentliche Rolle bei der Transkription des Viroids vermutet. Die Transkriptionsfaktoren IIIA (TFIIIA) und das ribosomale Protein L5 könnten in den intrazellulären Transport des Viroids involviert sein. Für die Analyse der Wechselwirkung ist es wünschenswert, die Sequenzen der untersuchten Proteine aus einer Wirtspflanze des Viroids zu kennen. Da TFIIIA bislang lediglich aus *Arabidopsis thaliana* charakterisiert wurde, musste zunächst eine entsprechende Sequenz aus einem Nachtschattengewächs gefunden werden.

Der Ergebnisteil beginnt indes mit einer anderen Fragestellung zum Thema Replikation von PSTVd. Zum Zeitpunkt des Beginns dieser Arbeit wurde die Startstelle der Transkription im linken terminalen *loop* des PSTVd-Moleküls bestimmt (Kolonko, 2003). Um auf die genaue Identität des Startnukleotids schließen zu können, sollten im Rahmen der vorliegenden Arbeit einzelne Nukleotide im linken terminalen *loop* ausgetauscht werden und die Infektiosität der so mutierten Viroide *in planta* charakterisiert werden.

2

# Material

# 2.1 Bezugsquellen

Sämtliche Chemikalien und Lösungsmittel wurden mit dem Reinheitsgrad *pro analysi* verwendet. Enzyme wurden überwiegend von den Firmen Boehringer (Mannheim), Gene-Craft (Lüdinghausen), New England Biolabs (Ipswich, USA), Fermentas (Burlington, Canada) und Promega (Wisconsin, USA) bezogen.

# 2.2 Chemikalien und Lösungen

Alle Puffer und Lösungen wurden mit hochreinem Milli-Q-Wasser (Millipore GmbH, Neu Isenburg) angesetzt und wenn möglich 20 min bei 120 °C autoklaviert. Nicht autoklavierbare Lösungen wurden durch  $0,2 \ \mu$ m Membranfilter (FP 030/3, Schleicher und Schuell, Dassel) sterilfiltriert. Prozentangaben erfolgen, falls nicht anders vermerkt, in weight per volume (w/v). Phenol wurde mit 100 mM Tris/HCl auf pH 8.0 äquilibriert und mit einem Volumen Chloroform versetzt (Phenol/Chloroform). 30%ige Acrylamid-Stammlösungen wurden hergestellt durch Lösen von Acrylamid und N,N'-Methylenbisacrylamid in den Verhältnissen 30:1 oder 43:1. Die Stammlösungen wurden mit 5 g/l Amberlite Ionentauscher-Partikel für 30 min gerührt und anschließend filtriert. Die am häufigsten verwendeten Puffer finden sich nachfolgend aufgelistet, alle übrigen sind unter den entsprechenden Methoden verzeichnet.

1  imes TE	$10 \mathrm{~mM}$	Tris/HCl, pH 8.0
	$1 \mathrm{mM}$	EDTA
$10  imes  ext{TBE}$	890 mM	Tris
	$890 \mathrm{~mM}$	Borsäure
	$25 \mathrm{~mM}$	EDTA
	pH 8.3	
10  imes NaTA	1 M	Natriumacetat
	400  mM	Tris/Essigsäure, pH 8.4
$50  imes \mathbf{TAE}$	2 M	Tris/HCl, pH 8.3
	1 M	Essigsäure
	$20 \mathrm{~mM}$	EDTA
$1 \times TNE$	$100 \mathrm{~mM}$	Tris/HCl, pH 7.5
	$100 \mathrm{~mM}$	Natriumchlorid
	$10 \mathrm{~mM}$	EDTA
70 % Formamid-Auftragspuffer	70~%	Formamid
	30~%	$1 \times \text{TE-Puffer}$
	$0{,}05~\%$	Bromphenolblau
	$0{,}05~\%$	Xylencyanol
$1 \times PBS$	$137 \mathrm{~mM}$	Natriumchlorid
	$2{,}7~\mathrm{mM}$	Kaliumchlorid
	$10 \mathrm{~mM}$	Dinatriumhydrogenphosphat
	1,8 mM	Kaliumdihydrogenphosphat
	pH 8.4	

Zur Anzucht von Bakterien wurden folgende Medien verwendet:

LB-Medium	10 g/l Bacto-Trypton
	5  g/l Bacto-Yeast-Extrakt
	10 g/l NaCl
	pH 7.4 mit NaOH
YT-Medium	10 g/l Bacto-Trypton
	5 g/l Bacto-Yeast-Extrakt
	5 g/l NaCl
	pH 7.4 mit NaOH
LB-Agar	wie LB-Medium, zusätzlich $15{\rm g/l}$ Agar
TB-Medium	12 g/l Bacto-Trypton
	24 g/l Bacto-Yeast-Extract
	$17 \text{ mM} \text{ KH}_2 \text{PO}_4$
	$72 \text{ mM}  \text{K}_2\text{HPO}_4 \ge 3 \text{ H}_2\text{O}$
	0,4 % Glycerin (v/v)
	pH 8.4
SOB-Medium	20 g/l Bacto-Trypton
	5 g/l Bacto-Yeast-Extract
	10 mM NaCl
	2,5  mM KCl
	$10 \text{ mM} \text{MgCl}_2$
	$10 \text{ mM} \text{MgSO}_4$
	pH 6.8
SOC-Medium	wie SOB-Medium, zus. 20 mM Glucose

Als selektionierendes Antibiotikum wurde Ampicilin (100 $\mu{\rm g}/{\rm ml})$ und/oder Chloramphenicol (34 $\mu{\rm g}/{\rm ml})$ eingesetzt.

**Tabelle 2.1: DNA-Oligonukleotide.** Alle in der vorliegenden Arbeit eingesetzten Oligonukleotide wurden von MWG Biotech (Martinsried) bezogen. Zusätzlich wurde eine Qualitätskontrolle mittels Massenspektrometrie in Auftrag gegeben.

Name	Sequenz
5SDNA1	GGATGCGATCATACCAGC
At5SDNA2	CCGAAAAAGTAGTTAAAGGTC
Le5SDNA2	TTAAACACAACCTTCCTCCC
G359+	CCGCAGTTGGTTCCGCGGAACTAAACTCGTG
G359-	CACGAGTTTAGTTCCGCGGAACCAACTGCGG
G1+	CCGCAGTTGGTTCCTGGGAACTAAACTCGTG
G1-	CACGAGTTTAGTTCCCAGGAACCAACTGCGG
XXLT7	TAATACGACTCACTATAGGTGCGGCGTCTTG
XXLPolyA	T(30)GATTATAAGCACAACATTACTTATATCTC
TFIIIA1C	GGTGGTGTCGACATGGCGGAAGAAGCTAAAG
TFIIIA2CQ	GCAAGTTTCGTGTTCTTCTGC
L51C	GGTGGTGGTGTCGACATGGTGTTTGTGAAGTCC
L52CQ	CTCTTCATCGTCCTCATC
pGEX5'	GGGCTGGCAAGCCACGTTTGGTG
pGEX3'	CCGGGAGCTGCATGTGTCAGAGG
TBPC2	GGTGGTTGCTCTTCCGCATTGCTGATTCTTCCTGAACTC
TBPC1	GGTGGTGGTGCTAGCATGGCAGATCAGGGATTAGAG
prTBPpGX4+	TATCATAGGATCCATGGCAGATCAGGGATTAGAG
prTBPpGX4-	ATTCTATGTCGACTCATTGCTGATTCTTCCTGAACTC
prL5LEpGX4+	TATCATAGGATCCATGGCATTCATCAAAGTCCAG
prL5LEpGX4-	ATTCTATGTCGACTTTTATAATTCTAGTGTCTTCACTC
pr3ApGX4+	TATCATAGGATCCATGGCGGAAGAAGCTAAAG
prL5ATpGX4-	ATTCTATGAATTCGATGGTGTTTGTGAAGTCCAC
prL5ATpGX4-	ATTCTATGTCGACCTGATTTACTCTTCATCGTCC
kurz3A-	ATTCTATGTCGACTCAAAATATCTCAATCTCGAGTG
lang3A+	TATCATAGGATCCATGCAAGAGAGGGCCATTC
AT3Avoll+	GTGCGGCGTCTTGATGGAGG
AT3Avoll-	GATTATAAGCACAACATTACTTATATCTCAAACGTTGGG
M13uni	TGTAAAACGACGGCCAGT
M13rev	TGTAAAACGACGGCCAGT



Abbildung 2.1: Vektorkarte, pTYB1 (NEB).

# 2.3 Plasmide

Die Vektoren pTYB1, pTYB2 und pTYB11 (NEB) wurden für die Expression rekombinanter Fusionsproteine mit einem selbstspleißenden Affinitätstag verwendet (siehe Abb. 2.1). pGEX-4T-3 (Amersham) dient zur rekombinanten Expression von GST-Fusionsproteinen (siehe Abb. 2.2). Das Plasmid pGEM-T (Promega) erlaubt die einfache Klonierung von PCR-Produkten mit 3'-Adenosinüberhang. Die Sequenzen können unter Kontrolle eines T7- oder SP6-Promotors in RNA überführt werden. Der sogenannte Feldsteinvektor pRZ6-2 erhält eine monomere PSTVd (*intermediate* DI)-Sequenz, die über *Bam*HI kloniert wurde (Owens *et al.*, 1995). Die Sequenz wird flankiert von Ribozymsequenzen, die bei *in vitro*-Transkription ein lineares PSTVd-Molekül exakter Einheitslänge prozessieren.



Abbildung 2.2: Vektorkarte, pGEX-4T-3 (Amersham).

3

# Methoden

# 3.1 Arbeit mit Protein

## 3.1.1 Extraktion von Gesamtprotein aus Pflanzenmaterial

Zur Herstellung eines pflanzlichen Proteinextraktes wurden 1-2 g Blätter von Lycopersicon esculentum in flüssigem Stickstoff zu einem feinen Pulver gemörsert. Das Blattmaterial wurde direkt in einem Volumen  $2 \times$  Laemmli-Auftragspuffer aufgenommen (siehe Abschnitt 3.4.8) und 5 min bei 94 °C erhitzt. Unlösliches Material wurde durch Zentrifugation sedimentiert, und der lösliche Überstand auf einer SDS-PAGE analysiert (siehe Abschnitt 3.4.8).

## 3.1.2 In vitro-Translationsexperimente

Mit dem zellfreien Weizenkeimextrakt (Wheat Germ Extract Plus, Promega) können Proteine auf Basis von exogen zugefügter mRNA translatiert werden. Der Extrakt enthält alle Komponenten, die für eine Proteinsynthese essentiell sind: Aminosäuren, Ribosomen, tRNA und Translationsfaktoren. Ferner wurde endogene mRNA durch *Micrococcus*-Nuklease entfernt. Die *in vitro*-Translation erfolgte nach Angaben des Herstellers. Zur Detektion *de-novo*-synthetisierter Proteine wurde dem Reaktionsansatz <sup>35</sup>S-markiertes L-Methionin zugefügt. Die Translationsprodukte wurden auf einer SDS-PAGE aufgetrennt (siehe Abschnitt 3.4.8). Das Gel wurde über Nacht auf einem Gel-Trockner unter Vakuum bei 80 °C getrocknet und 4 h auf einem Phosphoimager exponiert.

## 3.1.3 Expression rekombinanter Proteine mit dem pTYB-System

Mit Einzelkolonien wurden 5 ml LB-Medium, supplementiert mit 50  $\mu$ g/ml Ampicilin und 40  $\mu$ g/ml Chloramphenicol, angeimpft. Diese Vorkulturen wurden über Nacht bei 37 °C auf einem Rundschüttler inkubiert und am nächsten Tag auf 500 ml Hauptkultur übertragen. Nach Erreichen einer optischen Dichte von 0,6 OD wurde die T7-kontrollierte Expression durch Zugabe von IPTG zu 0,4 mM induziert. Standardmäßig wurden die Hauptkulturen bei 15 °C über Nacht geschüttelt. Die Zellen wurden bei 4 °C und 5.000 g im JA10-Rotor abzentrifugiert. Die Bakterien wurden in je 20 ml 175 mM NaCl resuspendiert, in 12 ml-Greiner-Gefäßen aliquotiert und erneut zentrifugiert. Die gewaschenen Zellpellets wurden bis zur Aufreinigung bei -20 °C gelagert.

### 3.1.4 Aufreinigung rekombinater Proteine mit Intein-Fusionstag

Falls nicht anders erwähnt wurden alle Arbeitsschritte bei 4 °C durchgeführt. Bei der Reinigung von rekombinanten TFIIIA enthielten alle verwendeten Lösungen 100  $\mu$ M ZnCl<sub>2</sub>. Ein Zellpellet wurde in 10 ml Säulenpuffer (20 mM Tris/HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA, 0,1 % TritonX-100) resuspendiert und im Eiswasserbad sonifiziert. Unlösliches Material wurde bei 17.000 g im JA17-Rotor abzentrifugiert. Eine Chromatographiesäule (1 cm Durchmesser) wurde mit 4 ml 50 %igen Chitinpartikeln (NEB) beschickt und mit 40 ml Säulenpuffer äquilibriert. Das geklärte Lysat wurde langsam auf die Matrix geladen und anschließend mit 80 ml Säulenpuffer, enthaltend 1 M bis 2 M NaCl, gewaschen. Die Abspaltung des Zielproteins vom Affinitätstag wurde durch 10 ml Säulenpuffer, enthaltend 50 mM DTT, induziert. Die Säule wurde sodann verschlossen und bei 4 °C über Nacht inkubiert. Am nächsten Morgen wurde das rekombinante Protein mit 1 ml Säulenpuffer (+50 mM DTT) eluiert. Bei allen Arbeitsschritten wurden Aliquots entnommen und auf einer denaturierenden SDS-PAGE analysiert (siehe Abschnitt 3.4.8). Dabei wurden die Proben mit Auftragspuffer ohne reduzierendes Agens versetzt, um eine Inteinspaltung nicht nachträglich zu induzieren.

## 3.1.5 Expression rekombinanter Proteine mit dem pGEX-System

Die pGEX-Expressionsvektoren wurden gemäß Herstellerangaben mittels Hitzeschock in den *E. coli*-Stamm Rosetta $2^{\text{TM}}$ (Novagen) transformiert. Mit Einzelkolonien wurden 5 ml
LB-Medium, supplementiert mit 50  $\mu$ g/ml Ampicilin und 40  $\mu$ g/ml Chloramphenicol, angeimpft und über Nacht im temperierbaren Schüttler bei 30 °C inkubiert. 500  $\mu$ l dieser Vorkultur wurden am nächsten Morgen auf 50 ml LB<sup>amp/cam</sup> übertragen und bei 37 °C bis zum Erreichen von 0,5 bis 1 OD inkubiert. Sodann wurde die Expression durch IPTG-Zugabe zu 0,1 mM induziert. Zur Analyse des Wachstumsverhaltens der Expressionskulturen wurden stündlich nach Induktion OD-Messungen durchgeführt. Für die Analyse des Expressionslevels auf einer SDS-PAGE wurden 0,1 OD Zellen pelletiert und in Auftragspuffer resuspendiert (siehe Abschnitt 3.4.8). Dauer und Temperatur der Expression wurden variiert (72 h bei 4 °C, 14 h bei 15 °C, 8 h bei 24 °C oder 2-4 h bei 37 °C). Zur Beendigung der Expression wurden die Bakterien auf Eis abgekühlt und bei 5.000 g und 4 °C sedimentiert. Das Zellpellet wurde in eiskaltem PBS-Puffer resuspendiert und aliquotiert. Nach erneuter Zentrifugation und Abnahme des Überstandes konnten die Pellets bei -20 °C dauerhaft gelagert werden.

#### 3.1.6 Aufreinigung der GST-Fusionsproteine

Die affinitätschromatographische Aufreinigung der unlöslich exprimierten GST-Fusionsproteine GST-*Le*L5 und GST-*Le*TFIIIA gelang mittels des ionischen Detergenz n-Lauroylsarkosin (Sarkosyl). Dazu wurde ein von Frangioni & Neel (1993) publiziertes Protokoll in der vorliegenden Arbeit modifiziert. Alle Schritte finden bei 4 °C oder auf Eis statt, solange nicht anders erwähnt. Bei der Reinigung von GST-*Le*TFIIIA enthielten alle verwendeten Lösungen zusätzlich 100  $\mu$ M ZnCl<sub>2</sub>.

- Ein 125 ml Zellpellet resuspendieren in 6,25 ml STE (10 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA)
- 2. Zugabe von n-Lauroylsarkosin (Sarkosyl) zu $1,5\,\%$
- 3. Zugabe von Lysozym zu 100  $\mu$ g/ $\mu$ l, PMSF zu 0,1 mM und DTT zu 5 mM
- 4. Sonifizieren im 50 ml Falconröhrchen, 40 W, 0.9 s Puls,  $3\times30\,{\rm s}$
- 5. Zentrifugation bei 30.000 g, 30 min, Abnahme des geklärten Überstandes
- 6. Zugabe von TritonX-100 zu $2\,\%$

- Zugabe von 50 %iger Glutathion Sepharose<sup>TM</sup>4B Suspension (Amersham), äquilibriert nach Herstelleranweisung
- 8. 30 min langsames Mischen bei Raumtemperatur (head-over-tail rotating)
- 9. Transfer der Sepharosepartikel auf eine Chromatographiesäule (Durchmesser 1 cm)
- 10. Waschen mit insgesamt 50 Volumen Waschpuffer (20 mM Tris/HCl pH 8.0, 1 M NaCl, 5 mM  $\beta$ -Mercaptoethanol)
- 11. Transfer der Sepharosepartikel in ein Eppendorfgefäß
- GST-LeL5: Elution mit einem Volumen EBL5 (50 mM Tris/HCl pH 8.0, 10 mM reduziertes L-Glutathion, 0,1 % TritonX-100) bzw. GST-LeTFIIIA: Elution mit einem Volumen EB3A (50 mM Tris/HCl pH 8.8, 10 mM reduziertes L-Glutathion, 0,1 % TritonX-100, 1 mM β-Mercaptoethanol, 100 μM ZnCl<sub>2</sub>)
- 13. 30 min langsames Mischen bei Raumtemperatur (head-over-tail rotating)
- 14. Sepharosepartikel bei 500 g sedimentieren, 1 min
- 15. Überstand abnehmen, zu 10 % Glycerin versetzen und bei -70 °C lagern
- 16. Elution ggf. wiederholen.

#### 3.1.7 Aufkonzentrierung von Proteinlösungen

#### Acetonfällung

Diese denaturierende Fällungsmethode wurde verwendet, um Proteinlösungen vor der Analyse in einer SDS-PAGE aufzukonzentrieren. Dazu wurden die Proben mit 5 Volumen -20 °C kaltem Aceton versetzt und 10 min bei -20 °C inkubiert. Das Präzipitat wurde durch fünfminütige Zentrifugation in einer Tischzentrifuge bei 20.000 g, 4 °C sedimentiert und nach Abnahme des Überstandes kurz luftgetrocknet. Abschließend wurde das Pellet in Auftragspuffer aufgenommen.

#### Ultrafiltration

Zur nativen Aufkonzentrierung von Proteinlösungen kamen Centricon<sup>TM</sup>- und Microcon<sup>TM</sup>-Röhrchen (Millipore) mit geeignetem Ausschlussvolumen zum Einsatz. Die Handhabung erfolgte gemäß Anweisung des Herstellers. Mit diesem Ultrafiltrationssystem konnten auch Pufferwechsel durchgeführt werden. Dazu wurden die Proteinlösungen zweimalig mit 10 Volumen Puffer versetzt und aufkonzentriert.

# 3.2 Arbeit mit Nukleinsäure

#### 3.2.1 Extraktion mit Phenol und Chloroform

Nukleinsäure wurde gereinigt durch Extraktion mit Phenol und Chloroform. Die Lösungen wurden mit einem Volumen eines 1:1-Gemisches aus äquilibrierten Phenol und Chloroform versetzt und sodann ausgiebig gemischt. Die Phasentrennung wurde durch fünfminütige Zentrifugation bei 25.000 g beschleunigt. Die wässrige Phase wurde mit einem Volumen Chloroform versetzt, gemischt und erneut wie oben beschrieben zentrifugiert. Bei Nukleinsäurepräparationen aus Blattgewebe wurde ein zusätzlicher Extraktionsschritt mit reinem Phenol vorangestellt.

#### 3.2.2 Präzipitation von Nukleinsäure

#### Fällung mit Ethanol und Natriumacetat

Nukleinsäure kann durch Fällung einfach aufkonzentriert oder umgepuffert werden. Dazu wurde die Lösung auf 300 mM Natriumacetat eingestellt und mit einem dreifachen Volumen an 96 %igen Ethanol versetzt. Nach zehnminütiger Inkubation auf Eis wurde die Nukleinsäure durch 30-minütige Zentrifugation bei 25.000 g und 4 °C sedimentiert. Der Überstand wurde sodann dekantiert und das Pellet durch Zugabe von 70 %igem Ethanol salzfrei gewaschen. Nach kurzer Zentrifugation wurde der Überstand abgenommen, das Pellet 5 min luftgetrocknet und schließlich in einem geeigneten Puffer resuspendiert.

#### Fällung mit Ethanol und Ammoniumacetat

Radioaktiv markierte Nukleinsäure wurde mit Ethanol und Ammoniumacetat gefällt, da hier nicht-inkorporierte radioaktive Triphosphate im Überstand verbleiben. Zu der Nukleinsäurelösung wurde ein halbes Volumen 7,5 M Ammoniumacetat und 3 Volumen 96 %iges Ethanol gegeben. Nach halbstündiger Inkubation auf Eis wurde das Präzipitat durch 30-minütige Zentrifugation bei 25.000 g, 4 °C pelletiert. Der Überstand wurde dekantiert, das Pellet getrocknet und in einem geeigneten Puffer resuspendiert.

#### Fällung mit Polyethylenglykol und Natriumchlorid

Mittels Fällung mit  $PEG_{6000}$  und NaCl können Nukleinsäuren selektiv, entsprechend ihrer Größe gefällt werden (Schmitz & Riesner, 2006). Die zu präzipierende Nukleinsäurelösung wurde auf 6-18 %  $PEG_{6000}$  und 500 mM NaCl eingestellt und bei Raumtemperatur für 30 min inkubiert. Das Präzipitat wurde durch anschließende Zentrifugation (20-60 min, 18 °C, 30.000 g) gewonnen. Salzreste wurden durch Zugabe von 70 % igem Ethanol gelöst. Nach kurzer Zentrifugation bei Raumtemperatur wurde der Überstand abgenommen, das Pellet 5 min getrocknet und in einem geeigneten Puffer resuspendiert.

#### 3.2.3 Quantifizierung von Nukleinsäuren

Zur Ermittlung der Konzentration sowie zur Überprüfung des Reinheitsgrades von Nukleinsäuren wurden UV-Absorptionsspektren von 340 nm bis 210 nm in einem Spektralphotometer (Modell Spectophotometer DU 640; Beckman, USA) aufgenommen. Die Konzentration wurde aus der Absorption bei einer Wellenlänge von 260 nm nach Sambrook *et al.* (1989) wie folgt berechntet:

DNA	$1 \mathrm{A_{260}} = 50 \mu\mathrm{g/m^2}$
RNA	$1 \mathrm{A_{260}} = 40 \mu\mathrm{g/m^2}$
DNA-Oligonukleotide	$1  A_{260} = 33  \mu g/m_{\star}^2$

Die Messungen wurden in 50  $\mu$ l-fassende Quarzküvetten durchgeführt. Es wurde ferner darauf geachtet, dass die Absorption zwischen 0,1 A<sub>260</sub> und 1 A<sub>260</sub> lag, damit ein linearer Zusammenhang zwischen Absorption und Konzentration besteht. Bei sehr geringen Nukleinsäuremengen wurde die Konzentrationsbestimmung mittels einer  $\sqrt{10}$ -Verdünnungsreihe im Vergleich mit quantifizierten Standards alternativ in einer denaturierenden PAGE durchgeführt (siehe Abschnitt 3.4.2). Die Quantifizierung von <sup>32</sup>Pmarkierter Nukleinsäure erfolgte durch Messung der Cerenkov-Strahlung in einem Szintillationszähler (LS 5000 TD, Beckman) ohne Szintillationsflüssigkeit.

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

Polymerase-Kettenreaktionen (PCR) dienen der selektiven Amplifikation von DNA (Mullis *et al.*, 1986). Hauptanwendungen der PCR in der vorliegenden Arbeit waren:

- Genspezifische DNA-Amplifikation ausgehend von cDNA
- Sequenzbestimmung von Viroidvarianten
- Herstellung von Templates für in vitro-Transkriptionen

Für diese Zwecke erforderlich ist eine 3'-Exonuklease-Aktivität (*proof reading*) der verwendeten Polymerase zur fehlerfreien DNA-Vervielfältigung.

#### PCR-Reaktionsansatz mit Pfx-Polymerase

$1 \times$	Pfx-Polymerase-Puffer
$1 \times$	Enhancer-Lösung
$1 \mathrm{mM}$	$\mathrm{MgSO}_4$
$0,3 \mathrm{~mM}$	dNTP-Mix
$0{,}3~\mu\mathrm{M}$	Primer A
$0{,}3~\mu\mathrm{M}$	Primer B
0,02 u/µl	Pfx-Polymerase
1-20 ng	Template
$ad~50~\mu l$	H <sub>2</sub> O

Die PCR-Proben wurden mit einem Tropfen Paraffinöl überschichtet, um Evaporation während der Reaktion zu vermeiden. Der Thermozykler (Modell Varius V45, Landgraf, Langenhagen) wurde wie folgt programmiert:

#### PCR-Zyklen, Pfx-Polymerase

1 Zyklus	$30\mathrm{s}$	$94^{\circ}\mathrm{C}$
35 Zyklen	$15\mathrm{s}$	$94^{\circ}\mathrm{C}$
	$30\mathrm{s}$	$50-68^{\circ}\mathrm{C}$
	$60\mathrm{s}$	$68^{\circ}\mathrm{C}$

1 Zyklus 600 s 72 °C

Nach Beendigung des letzten Zyklus wurden die Reaktionsansätze auf 4°C abgekühlt. Die PCR-Produkte wurden mit dem QIAquick PCR Purification Kit (QIAGEN, Hilden) nach Anweisung des Herstellers aufgereinigt. Üblicherweise wurde 1/10 des Eluats auf einem Agarosegel analysiert.

Im Verlaufe der vorliegenden Arbeit wurde in den PCR-Reaktionen anstelle der Pfx-Polymerase die Phusion-Polymerase (Finnzymes) eingesetzt. Bei diesem neuartigen Enzym ist eine dsDNA-Bindedomäne fusioniert an eine Polymerase mit *proof reading*-Aktivität, was deren Prozessivität signifikant erhöht. Somit konnten deutlich höhere Produktausbeuten bei kürzeren Extensionszeiten erreicht werden.

#### PCR-Reaktionsansatz, Phusion-Polymerase

$1 \times$	HF-Buffer
$0,2 \mathrm{~mM}$	dNTP-Mix
$0{,}5~\mu\mathrm{M}$	Primer A
$0,5~\mu{ m M}$	Primer B
0,02 u/µl	Phusion-Polymerase
1-20 ng	Template
$ad~50~\mu l$	$H_2O$

#### **PCR-Zyklen**

1 Zyklus	$30\mathrm{s}$	$98^{\circ}\mathrm{C}$
35 Zyklen	$10\mathrm{s}$	$98^{\circ}\mathrm{C}$
	$20\mathrm{s}$	$50-72^{\circ}\mathrm{C}$
	$20\mathrm{s}/1\mathrm{kb}$	$72^{\circ}\mathrm{C}$
1 Zyklus	$600\mathrm{s}$	72 °C

Die PCR-Proben wurden mit einem Tropfen Paraffinöl überschichtet, um Evaporation während der Reaktion zu vermeiden. Die  $T_m$ -Werte der verwendeten Primer wurden nach der *nearest-neighborhood*-Methode ermittelt (Chester & Marshak, 1993). Zu diesem Zweck bietet die Firma Finnzymes einen Webdienst an (http://www.finnzymes. fi/Java/tm\_determination.htm). Die annealing-Temperatur wurde bei 3 °C oberhalb des niedrigsten  $T_m$ -Werts der beiden Primer angesetzt. Bei den Klonierungsexperimenten war es notwendig, genspezifische Primer mit einem für Restriktionsenzyme zugänglichen, 5'-seitigen Sequenzabschnitt zu versehen, um das PCR-Produkt in Vektor DNA zu klonieren. In diesem Falle wurde der  $T_m$ -Wert nur für den genkomplementären Sequenzbereich errechnet und die annealing-Temperatur für die ersten fünf Zyklen entsprechend herabgesetzt. Nach Beendigung der finalen Extension wurden die Reaktionsansätze auf 4 °C abgekühlt. Die PCR-Produkte wurden wie oben beschrieben gereinigt und analysiert.

## 3.2.5 PCR mit $\alpha^{32}$ P -dCTP

Radioaktiv markierte DNA wurde in einer PCR durch Einbau von  $\alpha^{32}$ P -dCTP synthetisiert. Die eigesetzten Komponenten wurden identisch gewählt wie in den oben beschriebenen PCR-Reaktionsansätzen mit Phusion-Polymerase. Anstelle des dNTP-Mixes wurden je 0,2 mM dATP, dGTP und dTTP, 0,05 mM dCTP und 3  $\mu$ Ci/50  $\mu$ l  $\alpha^{32}$ P -dCTP eingesetzt. Die PCR-Produkte wurden mit dem QIAquick PCR Purification Kit (QIAGEN, Hilden) nach Anweisung des Herstellers aufgereinigt. Ein Aliquot des Eluats wurde im Szintillationszähler vermessen (siehe Abschnitt 3.2.3).

#### 3.2.6 Sequenzierung von DNA

Sequenzierungen wurden wahlweise bei den Firmen Seqlab (Göttingen) oder Genterprise (Mainz) in Auftrag gegeben. Dazu wurden Plasmid- und Primerkonzentration den jeweiligen Erfordernissen angepasst. Die Elektropherogramme wurden mit dem Programm CHROMAS, Version 1.45 ausgewertet. Die Sequenzen wurden mit dem BLAST-Algorithmus verifiziert.

#### 3.2.7 In vitro-Transkription

RNA wurde *in vitro* mit der RNA-Polymerase des Phagen T7 erzeugt (Milligan *et al.*, 1987). Die im hiesigen Institut präparierte T7-Polymerase wurde mir freundlicherweise von Herrn Bernd Esters zur Verfügung gestellt. Als Template für die Transkription diente wahlweise linearisiertes Plasmid oder PCR-Produkt mit integriertem T7-Promotor. Je

nach Anwendungszweck wurden unterschiedliche Mengen an RNA benötigt; daher wurden zwei verschiedene Protokolle verwendet.

#### IVT im kleinen Maßstab

50  nM	Template-DNA
$1 \times$	IVT A
$1 \times$	IVT B
$1 \mathrm{~mM}$	rNTP-Mix (je 4 mM ATP, CTP, GTP, UTP)
$1 \mathrm{mM}$	rNTP-Mix (je 4 mM ATP, CTP, GTP, UTP)
$15 \text{ U}/\mu \text{l}$	T7-Polymerase
$1 \mathrm{U}/\mu\mathrm{l}$	$\mathrm{RNAsin}^{\mathrm{I\!R}}$
ad 10 µl	H <sub>2</sub> O

#### IVT im großen Maßstab

$100 \ \mu { m g/ml}$	Plasmid
$1 \times$	$\operatorname{RiboMAX}^{\textcircled{R}}$ -Reaktionspuffer
$4 \mathrm{mM}$	rNTP-Mix (je 4 mM ATP, CTP, GTP, UTP)
3000  U/ml	T7-Polymerase
$1000~{\rm U/ml}$	RNAsin®
$5 \mathrm{U/ml}$	Pyrophophatase
$ad 100 \ \mu l$	H <sub>2</sub> O

Die Transkription erfolgte für 3h bzw. 4h bei  $37 \,^{\circ}$ C unter erneuter Zugabe von  $15 \,\text{U}/\mu$ l bzw.  $3000 \,\text{U/ml}$  T7-Polymerase nach der Hälfte der Inkubationszeit. Anschließend wurde das DNA-Template durch eine Behandlung mit DNaseI (Ambion) nach Angaben des Herstellers entfernt. Die Mg<sup>2+</sup>-Ionen wurden mit EDTA vollständig komplexiert und die RNA mit Phenol/Chloroform und Chloroform extrahiert. Die Transkripte wurden schließlich mit PEG<sub>6000</sub> und NaCl gefällt und die Pellets mit 70 %igem Ethanol gewaschen.

# 3.2.8 In vitro-Transkription mit $\alpha^{32}$ P -UTP

Die Herstellung von radioaktiv markierten Transkripten erfolgte in Anlehnung an das unter Abschnitt 3.2.7 beschriebene IVT-Protokoll im kleinen Maßstab. Dabei wurde die UTP-Konzentration auf 250  $\mu$ M abgesenkt und dem Reaktionsansatz zusätzlich 25  $\mu$ Ci  $\alpha^{32}$ P-UTP hinzugefügt. Die Reaktionsdauer wurde auf 45 min beschränkt und das DNA-Template nicht entfernt. Die markierte RNA wurde mit Ammoniumacetat und Ethanol gefällt, in Formamid-Auftragspuffer gelöst und aus einem denaturierenden Gel eluiert (siehe Abschnitt 3.4.5). Ein Aliquot wurde im Szintillationszähler vermessen (siehe Abschnitt 3.2.3). Typischerweise lag die Ausbeute bei 1-5 × 10<sup>6</sup> cpm. Vor dem Einsatz in eine Verzögerungsgelelektrophorese wurde die markierte Nukleinsäure mit Tris/HCl, pH 8.0 auf 5.000 cpm/ $\mu$ l verdünnt.

#### 3.2.9 In vitro-Synthese von mRNA

Messenger RNA wurde von PCR-Produkten *in vitro* transkribiert, deren kodogener Strang am 3'-Ende 30 Adenosine aufweist. Somit wird bei der Transkription ein charakteristischer PolyA-Tail generiert. Ferner wurde dem Transkriptionsansatz 5'7-Methyl Guanosin Nukleotid (Ribo m<sup>7</sup>G Cap Analog, Promega) im Verhältnis 10:1 zu GTP hinzugefügt. Somit tragen im statistischen Mittel 90 % aller Transkripte eine physiologische 5'-ständige Cap-Struktur. Die Transkriptionsreaktion erfolgte nach dem RiboMAX<sup>®</sup>-Protokoll (siehe Abschnitt 3.2.7).

#### 3.2.10 In vitro-Synthese von infektiöser Viroid-RNA

Als Template für die *in vitro*-Synthese von infektiösen Viroid-Molekülen diente das Plasmid pRZ6-2 (Owens *et al.*, 1995). In diesem pRZ18R-Derivat ist eine monomere PSTVd *IntermediateDI* Sequenz kloniert. Die Viroidsequenz wird flankiert von Ribozymkassetten, die die T7-Transkripte zu exakter PSTVd-Einheitslänge prozessieren. Die Transkription erfolgte nach dem RiboMAX<sup>®</sup>-Protokoll, wobei die MgCl<sub>2</sub>-Konzentration um 20 mM erhöht wurde, um die Ribozym-Aktivität zu steigern. Die Volllängen-PSTVd-Moleküle wurden in einem denaturierenden PAA-Gel von unvollständig prozessierten Transkripten separiert und eluiert. PSTVd-Moleküle mit einzelnen Basenaustauschen wurden von punktmutierten pRZ6-2-Plasmiden synthetisiert. Die gerichtete Mutagenese wurde mit dem Site Directed Mutagenesis Kit<sup>TM</sup>(Stratagene) nach Anweisung des Herstellers durchgeführt. Dabei wurden die Primer G359+, G359-, G1+ und G1- verwendet.

#### 3.2.11 Isolation von mRNA

Polyadenylierte mRNA wurde mit dem Oligotex<sup>TM</sup>mRNA Mini Kit (QIAGEN, Hilden) nach Angaben des Herstellers isoliert. Es wurden 200-400  $\mu$ g Gesamt-RNA eingesetzt. Die typische Ausbeute an mRNA entsprach ungefähr 1-3 % der vorgelegten Gesamt-RNA.

#### 3.2.12 Herstellung einer cDNA-Bank

Für die reverse Transkription von mRNA in cDNA wurden 1  $\mu$ g mRNA mit 2  $\mu$ l polydT-Nukleotid (100  $\mu$ M) und 2  $\mu$ l dNTP-Mix (10 mM) in einem Gesamtvolumen von 26  $\mu$ l versetzt. Die Probe wurde 5 min auf 65 °C erhitzt, auf Eis abgekühlt und das Kondensat kurz abzentrifugiert. Nach Zugabe von 8  $\mu$ l First-Strand-Buffer (Invitrogen) und 4  $\mu$ l 0,1 M DTT wurde die Lösung auf 42 °C kurz vorgewärmt und anschließend die Polymerisation mit Zugabe von 2  $\mu$ l Superscript II reverse Transkriptase (Invitrogen) gestartet. Nach einer Inkubation von 50 min bei 42 °C wurde das Enzym 15 min bei 70 °C inaktiviert und das RNA-Template mit 1  $\mu$ l 4 $u/\mu$ l RNase bei 37 °C entfernt. Die so produzierte cDNA konnte ohne weitere Aufreinigung direkt in eine PCR für eine genspezifischen DNA-Amplifikation eingesetzt werden.

#### 3.2.13 Infektion von Tomatenpflanzen

Die Effekte der Punktmutationen auf die Replikation von PSTVd wurde in Infektionsstudien von Tomatenpflanzen (*Lycopersicum esculentum*) der Handelssorte Rutgers analysiert. Tomaten dieser Varietät reagieren mit ausgeprägter Symptomatik auf Viroidbefall. Die Pflanzen wuchsen in einem beleuchteten, auf 30 °C temperierten Gewächshaus bei einer relativen Luftfeuchtigkeit von 60 %. Zwei Wochen nach ihrer Aussaat in Torfkultursubstrat (Typ TKS2) wurden die Pflanzen vereinzelt und konnten nach weiteren zwei Wochen mit PSTVd-Transkripten infiziert werden. Dazu wurde ein junges Keimblatt mit Siliziumcarbid (Carborundum<sup>TM</sup>, Schleifmittelwerke GmbH, Düsseldorf) bestreut, 5 ng Viroid-Transkript (siehe Abschnitt 3.2.10) in  $10 \,\mu$ l H<sub>2</sub>O aufgetragen und mit einer Pipettenspitze vorsichtig eingerieben. Als Negativkontrollen wurden Scheininfektionen mit Wasser durchgeführt. Zu verschiedenen Zeitpunkten *post infectum* wurden Blattproben nahe der Pflanzenspitze genommen und der Viroidgehalt analysiert.

Nicht infiziertes Blattmaterial wurde zur Herstellung pflanzlicher Proteinextrakte verwendet (siehe Abschnitt 3.1.1).

#### 3.2.14 Extraktion von Nukleinsäure aus Pflanzenmaterial

Der erste Schritt bei den Klonierungsexperimenten war die Extraktion von Gesamtnukleinsäure aus Pflanzenmaterial. Aus diesem Nukleinsäure-Pool wurde mRNA isoliert und diese wiederum in cDNA kopiert. Gesamtnukleinsäure wurde extrahiert aus:

- Filtrat einer flüssigen Zellkultur von Solanum tuberosum der Linie HH258
- einzelnen Blättern von Lycopersicon esculentum
- ganzen Pflanzen von Arabidopsis thaliana

Das Pflanzenmaterial wurde zunächst abgewogen und anschließend unter flüssigem Stickstoff zu einem feinen Pulver gemörsert. Die Aufarbeitung des Gewebes wurde mit dem RNeasy Plant Total RNA Kit<sup>TM</sup>(QIAGEN, Hilden) gemäß der Protokollvorschrift durchgeführt. Wenn notwendig wurden die Proben durch eine Natriumacetat/Ethanol-Fällung aufkonzentriert.

#### 3.2.15 Nachweis von Viroid-RNA in Pflanzenmaterial

Für den Nachweis von Viroid RNA wurden Lycopersicon esculentum Blätter nahe der Pflanzenspitze abgeschnitten und sofort auf Eis transferiert. Das Blattmaterial wurde gewogen, schockgefroren und in einem 1,5 ml Reaktionsgefäß mit einem Feinspatel gemörsert. Das feine Pulver wurde mit 2 Volumen TNE<sup>++</sup>-Puffer (TNE-Puffer, enthaltend 1 % SDS und 10 mM  $\beta$ -Mercaptoethanol) und 2 Volumen Phenol versetzt und anschließend gründlich gemischt. Die Phasentrennung wurde durch kurze Zentrifugation beschleunigt. Die wässrige Phase wurde mit zwei Volumen Phenol/Chloroform und anschließend mit zwei Volumen Chloroform erneut extrahiert. Der Überstand wurde mit 500 mM NaCl und 6 % bzw. 15 % PEG<sub>6000</sub> fraktioniert gefällt. Das Präzipitat wurde abzentrifugiert und in einem geeigneten Puffer resuspendiert.

# 3.3 Mikrobiologische Arbeiten

Mikrobiologische Arbeiten wurden unter sterilen Bedingungen durchgeführt. Wenn möglich, wurden alle verwendeten Lösungen, Puffer und Medien bei 121 °C für 30 min autoklaviert. Temperaturinstabile Lösungen wurden durch 0,2 mm Celluloseacetat-Membranen sterilfiltriert. Glasgeräte wurden mindestens 4 h bei 210 °C gelagert.

Alle Arbeiten mit bakteriellem Material wurden gemäß den gesetzlichen Vorschriften für gentechnisch veränderte Organismen (GVO) durchgeführt. Mikrobiologischer Abfall wurde stets autoklaviert, Kontaminationen wurden mit 2% iger Incidinlösung<sup>®</sup> beseitigt.

#### 3.3.1 Anzucht und Aufbewahrung von Bakterienkulturen

Um Bakterienzellen zu vermehren wurden mittels einer sterilen Platinöse einige Kristalle einer tiefgefrorenen Stammkultur abgekratzt und auf einer LB-Agarplatte ausgestrichen. Diese wurde bis zum Wachstum von Einzelkolonien üblicherweise über Nacht bei 37 °C inkubiert. Agarplatten mit Einzelkolonien wurden bis zu vier Wochen bei 4 °C kurzfristig gelagert. Eine einzelne Kolonie wurde mit einem Zahnstocher oder einer Platinöse in 3-10 ml LB-Flüssigmedium überführt und bei 37 °C und 220 rpm geschüttelt. Nach Erreichen der stationären Phase wurde die Bakteriensuspension für eine Protein- oder Plasmidpräparation aufgearbeitet.

Für eine Protein- oder Plasmidpräparation im großen Maßstab wurden bis zu 800 ml Flüssigmedium mit 1/100 Volumen einer Vorkultur angeimpft.

Zur längerfristigen Lagerung von Bakterienstämmen wurden glycerinhaltige Stammkulturen angelegt. Dazu wurden wie oben beschrieben Flüssigkulturen angesetzt. Nach Erreichen der logarithmischen Wachstumsphase, erkennbar durch deutliche Schlierenbildung, wurde die Kultur mit einem Volumen SOB, enthaltend 40 % Glycerin, versetzt und auf Eis abgekühlt. Die Stammkultur wurde bei -70 °C gelagert.

#### 3.3.2 Herstellung von kompetenten Zellen zur Elektrotransformation

Zunächst wurden von einer bei -70 °C gelagerten Stammkultur von DH5 $\alpha$ -Zellen einige Kristalle mittels einer Platinöse auf einer LB<sup>amp</sup>-Platte ausgestrichen und diese über Nacht bei 37 °C inkubiert. Mit einer einzelnen Kolonie 10 ml wurde LB-Flüssigmedium angeimpft und diese bei 37 °C über Nacht bei 220 rpm geschüttelt. Die Vorkultur wurde auf 11 LB-

Medium umgesetzt und bis zu einer optischen Dichte von 0.8 OD bei  $37 \,^{\circ}$ C und  $220 \,$ rpm inkubiert. Hierauf wurden die Zellen rasch auf Eis abgekühlt, bei  $4 \,^{\circ}$ C,  $3750 \,$ g,  $15 \,$ min pelletiert und folgender Waschprozedur unterzogen:

- 1. Resuspendieren in 11 eiskaltem, sterilen H<sub>2</sub>O. Pelletieren bei 4 °C, 3750 g, 15 min
- 2. Resuspendieren in 0,51 eiskaltem, sterilen H<sub>2</sub>O. Pelletieren bei 4 °C, 3750 g, 15 min
- 3. Resuspendieren in 0,51 eiskaltem, sterilen H<sub>2</sub>O. Pelletieren bei 4 °C, 3750 g, 15 min
- 4. Resuspendieren in 20 ml eiskaltem, sterilen H<sub>2</sub>O. Pelletieren bei 4 °C, 3750 g, 15 min

Schließlich wurde das Pellet in 2 ml sterilfiltriertem, eiskalten 10 % Glycerin aufgenommen und in 100  $\mu$ l und 200  $\mu$ l Aliquots bei -70 °C gelagert.

#### 3.3.3 Elektrotransformation

Plasmide wurden mittels Elektroporation in kompetente Zellen transformiert. Die Reaktion wurde in einem ECM 600 Electro Cell Manipulator (BTX, San Diego, USA) durchgeführt, da hier besonders hohe Transformantenzahlen erreicht wurden. Sofern nicht anders erwähnt, wurden alle Arbeitsschritte mit eisgekühlten Materialien durchgeführt.

Es wurden 5  $\mu$ l Ligationsansatz (salzfrei in H<sub>2</sub>O) mit 45  $\mu$ l kompetenten Zellen vorsichtig gemischt und in eine Gene-Pulser-Küvette überführt (0,2 cm Elektrodenabstand, BIO-RAD). Nach einminütiger Inkubation auf Eis wurde die automatische Ladungs- und Impulssequenz ausgelöst. Dabei wurden folgende Einstellungen verwendet: R: R5 (129 ohm), S: 2,5 kV, T: 7,4 ms, C: keine Einstellung.

Unmittelbar nach Beendigung des Impulses wurden  $950 \,\mu$ l raumtemperiertes SOC-Medium in die Küvette gegeben und die Bakteriensuspension ausgiebig gemischt. Diese wurde schließlich in einem Eppendorfgefäß für exakt 1 h bei 37 °C, 220 rpm geschüttelt. 1- $100 \,\mu$ l dieses Transformationsansatzes wurde auf LB-Agarplatten aufgebracht, die selektives Antibiotikum enthielten.

#### 3.3.4 Single Colony PCR

Mitunter war es notwendig eine sehr große Anzahl von *E. coli*-Kolonien auf den Gehalt von *insert*-DNA zu überprüfen. Aufgrund der besseren Handhabung wurde hier eine *single*-

colony-PCR (scPCR) der "konventionellen" Plasmid-Restriktionsanalyse vorgezogen. Für eine scPCR wurde eine Bakterienkolonie mit einer 200  $\mu$ l Pipettenspitze aufgenommen, auf eine sogenannte Masterplatte kopiert und schließlich in 100  $\mu$ l Aqua dest. überführt. Nach gründlichem Durchmischen wurden 5  $\mu$ l dieser Suspension in eine PCR-Reaktion eingesetzt. Verwendet wurde die Taq-Polymerase (Promega) sowie standardisierte Primer, die das Insert flankieren (z.B. M13-Primer). Die Durchführung der Reaktion erfolgte nach Empfehlung des Herstellers. Nach Beendigung der PCR wurden 5  $\mu$ l Reaktionslösung ohne Aufreinigung auf einem Agarosegel analysiert.

#### 3.3.5 Mini-Präparation von Plasmiden

Für die Präparation von Plasmid-DNA im kleinen Maßstab  $(1-10 \,\mu\text{g})$  wurde ein von Schmitz & Riesner (2006) etabliertes Protokoll verwendet. Hierbei handelt es sich um eine modifizierte Version der von Sambrook *et al.* (1989) publizierten Versuchsvorschrift, bei der die Plasmind-DNA nicht phenolisiert werden muss.

Mit einzelnen Bakterienkolonien wurden 3 oder 5 ml LB-Medium inklusive selektivem Antibiotikum angeimpft und über Nacht bei 37 °C, 220 rpm geschüttelt. 2 ml dieser Kultur wurden am nächsten Tag bei maximaler Drehzahl in einer Tischzentrifuge pelletiert und in 300  $\mu$ l Puffer P1 (50 mM Tris/HCl, pH 8.0, 10 mM EDTA, 100 ng/ $\mu$ l RNaseA) resuspendiert. Die alkalische Lyse wurde durch Zugabe von 300  $\mu$ l Puffer P2 (200 mM NaOH, 1 % SDS) und vorsichtigem Invertieren des Reaktionsgefäßes gestartet. Nach fünfminütiger Inkubation bei Raumtemperatur wurde die Reaktionslösung durch Zugabe von 300  $\mu$ l Puffer P3 (3M Kaliumacetat, pH 5.5) neutralisiert. Zelltrümmer und genomische DNA wurden abzentrifugiert bei 20.000 g, 4 °C, 10 min. Der klare Überstand, in dem sich die Plasmid-DNA befindet, wurde vorsichtig abgenommen und auf 500 mM NaCl und 8 % PEG<sub>6000</sub> eingestellt. Nach einer halbstündigen Inkubation auf Eis wurde das Präzipitat durch Zentrifugation gewonnen, mit 1 ml 70 % EtOH gewaschen und das salzfreie Pellet in 20  $\mu$ l Tris/HCl pH 8.0 gelöst. Die Plasmid Konzentration wurde spektralphotometrisch bestimmt.

#### 3.3.6 Midi-Präparation von Plasmiden

Plasmid-DNA wurde im mittleren Maßstab  $(50-100 \,\mu g)$  mit dem Plasmid Midi Kit oder dem HiSpeed Plasmid Midi Kit (QIAGEN, Hilden) nach Vorschrift des Herstellers präpariert.

#### 3.3.7 Maxi-Präparation von Plasmiden

Solange nicht anders erwähnt fanden bei der Plasmidpräparation im großen Maßstab (1-2 mg) alle Arbeitsschritte bei 4 °C bzw. auf Eis statt. Es wurden 500 ml LB<sup>amp</sup>-Medium mit 5 ml einer Vorkultur angeimpft und über Nacht bei 37 °C bei 220 rpm geschüttelt. Die Bakterienzellen wurden 5 min bei 5.000 g pelletiert und in 20 ml Lösung I (25 mM Tris/HCl, pH 8.0, 50 mM Glucose, 10 mM EDTA) resuspendiert. Die alkalische Lyse der Zellen erfolgte durch Zugabe frisch angesetzter Lösung II (0,2 N NaOH, 10% SDS (w/v)) und vorsichtigem Mischen. Nach einer Inkubation auf Eis für 10 min wurde die Reaktionslösung mit 30 ml frisch angesetzter, eiskalter Lösung III (3 M Kaliumacetat, 2 M Eisessig) neutralisiert und 20 min inkubiert. Die unlöslichen Bestandteile wurden 5 min bei 15.000 g abzentrifugiert. Der Überstand wurde über einen Faltenfilter gegeben und mit einem halben Volumen 30 % PEG<sub>6000</sub>, 1,5 M NaCl 1 h gefällt. Die Nukleinsäure wurde durch 30-minütige Zentrifugation bei 15.000 g sedimentiert, das Pellet in 10 ml TE-Puffer gelöst und in ein 50 ml Greiner Reaktiongefäß überführt. Um RNA zu entfernen wurde die Lösung auf 50  $\mu$ g/ml RNaseA eingestellt und für 15 min bei 37 °C inkubiert. Es folgte eine Proteinase-K Behandlung (50  $\mu$ g/ml) mit 1 % SDS für 15 min bei 37 °C. Die Plasmid-DNA wurde mit reinem Phenol, zweimal mit Phenol/Cloroform (1:1) und mit Chloroform extrahiert. Die Phasen wurden jeweils durch Zentrifugation bei 5000 g getrennt. Die Plasmid-DNA in der wässrigen Phase wurde mit 100 mM NaCl und 10 ml Isopropanol gefällt. Nach 30-minütiger Inkubation wurde die Probe 30 min bei 5.000 g zentrifugiert. Das Pellet wurde in TE-Puffer resuspendiert, mit 10% PEG<sub>6000</sub>, 500 mM NaCl gefällt und bei 14.000 g zentrifugiert. Die Plasmid-DNA wurde mit 70 % igem Ethanol entsalzt, getrocknet und in 500  $\mu$ l TE-Puffer aufgenommen. Die Qualität der Präparation wurde sowohl spektralphotometrisch als auch über ein Agarose-Gel bestimmt.

#### 3.3.8 Restriktionsendonuklease-Hydrolyse

Restriktionsendonuklease-Reaktionen wurden in der Regel nach Anweisung des Herstellers durchgeführt. Für eine möglichst vollständige Verdauung wurde die Inkubationszeit verlängert und mehr Enzym in die Reaktion eingesetzt. Bei einem *screening* einer großen Anzahl von Klonen wurde den Reaktionsansätzen 1 mM Spermidin hinzugefügt, welches den Bedarf an Enzym auf 1/10 verringert. Nach Beendigung der Inkubation wurde die DNA phenolisiert und mit PEG<sub>6000</sub> und NaCl gefällt.

#### 3.3.9 Dephosphorylierung von Vektor-DNA

Religationsreaktionen von linearisierter Vektor-DNA sind kinetisch favorisiert, da sie Reaktionen 1. Ordnung darstellen. Um unerwünschte Religation von Vektor-DNA zu unterdrücken, wurden die endständigen Phosphatgruppen der Vektor-DNA enzymatisch hydrolysiert. Hierfür wurde eine alkalische Phosphatase aus Kalbsthymus (CIP, Roche, Mannheim) verwendet.

#### Dephosphorylierungsansatz

Der Reaktionsansatz wurde 30 min bei 37 °C inkubiert. Die Phosphatase wurde durch Zugabe von 5 mM EDTA und zehnminütiger Erhitzung auf 78 °C deaktiviert. Es folgte eine Phenol/Chloroform-Extraktion sowie eine Fällung mit 10 % PEG<sub>6000</sub> und 500 mM NaCl. Da die anschließende Ligationsreaktion äußerst salzsensitiv ist, wurde das Pellet mit 70 %igem Ethanol zweimal gewaschen und in einem geeigneten Volumen 10 mM Tris/HCl, pH 8.0, resuspendiert. Die Effizienz der Dephosphorylierung wurde in Ligations- und Transformationsreaktionen überprüft.

#### 3.3.10 Ligase-Reaktionen

Die Ligation eines linearisierten Plasmides mit Insert-DNA erfolgte enzymatisch mit der DNA-Ligase des Bakteriophagen T4. Diese katalysiert die Bildung von Phosphodiesterbindungen zwischen benachbarten 5'Phosphat- und 3'Hydroxyl-Resten.

#### Ligationsansatz

50  ng	Plasmid (linearisiert, dephosphoryliert)
x ng	Insert-DNA (dreifacher molarer Überschuss)
$1 \times$	T4 DNA-Ligase-Puffer
$1 \ \mu l$	T4 DNA-Ligase
$1 \mathrm{mM}$	ATP
$ad 10 \mu l$	H <sub>2</sub> O

Die Lösung wurde in einem wasserdichten Reaktionsgefäß mit Schraubdeckel angesetzt und in einem auf 12 °C temperierten Wasserbad für 1 h inkubiert. Wurden glatte Enden ligiert, erfolgte die Inkubation bei 4 °C über Nacht. Nachdem die Inkubation beendet war, wurden die Proben mit 8 % PEG<sub>6000</sub> und 500 mM NaCl gefällt. Das Pellet wurde zweimal mit 70 % Ethanol gewaschen und in 5  $\mu$ l H<sub>2</sub>O gelöst. Schließlich wurden 2,5  $\mu$ l in eine Transformationsreaktion eingesetzt (siehe Abschnitt 3.3.3).

#### 3.3.11 Klonierung in pGEM-T

Zur exakten Charakterisierung der Sequenz von PCR-Produkten war es notwendig, diese in den pGEM-T-Vektor (Promega) zu subklonieren. Dieser lineare Vektor weist überhängende Tymin-Nukleotide an den 5'-Enden auf. Somit können PCR-Produkte, die mit Taq-Polymerase erzeugt wurden, auf einfache Art und Weise kloniert werden, da diese 3'-überhängende Adenosine aufweisen. Da im Rahmen dieser Arbeit überwiegend mit *Proof-Reading-*Polymerasen gearbeitet wurde, die glatte Enden produzieren, mussten zunächst diese PCR-Produkte mit einem 3'-seitigen Adenosin versehen werden (*Atailing*).

A-tailing-Ansatz			
	25  nM	PCR-Produkt	
	$0{,}2~\mathrm{mM}$	dATP	
	$1,5~\mathrm{mM}$	MgCl2	
	$1 \times$	Taq-Polymerase-Puffer	
	$1 \mu l$	Taq-Polymerase	
-0	$nd 10 \mu l$	H <sub>2</sub> O	

Der Ansatz wurde für 30 min bei 70 °C inkubiert. 1  $\mu$ l dieses Ansatzes wurde direkt ohne weitere Aufreinigung in eine Ligationsreaktion mit dem Vektor pGEM-T eingesetzt.

#### Ligationsansatz mit pGEM-T

50  ng	pGEM-T
$1 \mu l$	DNA aus $A$ -tailing-Reaktion
$1 \times$	T4 Ligationspuffer
$1 \mu l$	T4 DNA-Ligase
$ad  10  \mu l$	$H_2O$

Parallel wurden Kontrollen ohne Insert und mit Testinsert durchgeführt. Der Ligationsansatz wurde bei 4 °C über Nacht inkubiert. Es wurden 25  $\mu$ l Tricin/KOH (25 mM, pH 8.4) hinzugefügt und die Ligase bei 65 °C für 15 min deaktiviert. Das Ligationsprodukt wurde mit 8 % PEG<sub>6000</sub> und 500 mM NaCl präzipitiert, zweimal mit 70 % Ethanol gewaschen und in 5  $\mu$ l Wasser resuspendiert. Es wurden 2,5  $\mu$ l in einer Transformationsreaktion eingesetzt (siehe Abschnitt 3.3.3). Der eingesetzte Vektor ermöglicht eine Überprüfung auf Insert mittels Blau/Weiß-Selektion. Hierzu wurden je 100  $\mu$ l IPTG (200 mM) und X-Gal (2 %) auf LB<sup>amp</sup>-Platten ausgestrichen und diese 30 min getrocknet. Es wurden 10  $\mu$ l und 100  $\mu$ l des Transformationsansatzes ausplattiert.

# 3.4 Gelelektrophoresen

#### **3.4.1** Agarosegele

DNA-Fragmente größer als 500 bp wurden in Agarosegelen aufgetrennt. Die Elektrophorese in Agarosegelen wurde durchgeführt in einer Minigel-Apparatur (Modell Mini Sub DNA Cell, BIO-RAD) oder in einer Flachbettapparatur, die die hiesige mechanische Werkstatt hergestellt hatte. Die 1-2 %igen Agarosegele sowie der Laufpuffer enthielten  $1 \times \text{TAE}$  und  $1 \,\mu\text{g/ml}$  Ethidiumbromid. Die Elektrophorese verlief bei 8-10 V/cm für 1-2 h. Abschließend wurde die Nukleinsäure auf einem UV-Transluminator (UVT 2035; Herolab, St.Leon) bei 302 nm visualisiert und das Gel mit einer Sofortbildkamera der Firma Polaroid (Modell MP-4 Land Camera; Massachusetts, USA) dokumentiert.

## 3.4.2 Denaturierende Polyacrylamidgele

Kleinere Nukleinsäure bis 500 bp wurde in denaturierenden Polyacrylamidgelen analysiert. Dabei erwärmt der hohe Stromfluss das nur 0,5 mm dicke Gel auf 60 °C; in Kombination mit einer Harnstoffkonzentration von 8 M werden somit denaturierende Bedingungen geschaffen, unter denen sich sämtliche Nukleinsäurestrukturen auflösen. Als Gelund Laufpuffer wurde  $0.5 \times$  TBE verwendet. Die Nukleinsäureproben wurden mit Formamid-Auftragspuffer zu einer Formamid-Konzentraion von mindestens 50 % versetzt. Anschließend wurden die Proben in einem Heatblock bei 94 °C für 3 min denaturiert und in einem Ethanol/Wasser-Gemisch von -20 °C rapide abgekühlt. Vor dem Auftragen der Proben wurde das Gel durch Anlegen einer Spannung von 1000 V auf 60 °C vorgewärmt. Der Gellauf wurde beendet, sobald der Bromphenolblau-Farbstoff in das untere Pufferbad übergetreten war. Die Nukleinsäure wurde anschließend durch eine im hiesigen Institut etablierte Silberfärbung sichtbar gemacht (siehe Abschnitt 3.4.6).

#### 3.4.3 Verzögerungsgelelektrophorese

Protein-Nukleinsäure-Wechselwirkungen können mit der Verzögerungsgelelektrophorese analysiert werden. Dabei macht man sich zu Nutze, dass an Protein gebundene Nukleinsäure retardiert wird. Anhand des Verhältnisses von gebundener zu nicht gebundener Nukleinsäure lässt sich die Affinität der Bindung quantifizieren. Als Gelapparatur diente das Modell 22 der Firma BIO-RAD, in der zwei vertikale Gele (175 mm x 120 mm x 1,5 mm) Platz finden. Nach Gießen des Gels und Beendigung der Polymerisation wurden die Gele samt Apparatur bei 4°C für mindestens 2h vorgekühlt.

#### Verzögerungsgellösung

$5 \ \%$	Acrylamid/Bisacrylamid (43:1)
0,2 $\times$	Na-TA Puffer
$0{,}5~\mathrm{mM}$	$MgSO_4$ (optional)
$0{,}1~\%$	TEMED
$0{,}03~\%$	APS

#### $1 \times$ Bindepuffer (BP)

- 20 mM Tris/HCl, pH 8.0
- 100 mM KCl
  - $1 \text{ mM} \text{ MgCl}_2$
- $100 \ \mu M \quad ZnCl_2 \text{ (optional)}$
- 0,004 % Bromphenolblau
- 0,008 %  $\beta$ -Mercaptoethanol
  - 3 % Glycerin

Die in  $1 \times$  Bindepuffer gelagerten Proteine wurden auf Eis aufgetaut. Eventuell vorhandene bindeinaktive Proteinaggregate wurden durch Zentrifugation (32.000 g, 4°C, 30 min) entfernt. Protein-Konzentrationsreihen wurden erstellt durch Verdünnung in  $1,25 \times$  BP. Durch Zugabe von  $1 \mu$ l markierter Nukleinsäure (5.000 cpm) zu  $4 \mu$ l Proteinlösung wurde eine Endkonzentration von  $1 \times$  BP erreicht. Die Proben wurden durch Schnippen vermischt und bei Raumtemperatur für 30 min inkubiert. Nach Abkühlung auf Eis wurden die Bindeansätze schließlich auf das vorgekühlte Gel aufgetragen. Die Elektrophorese verlief bei 4°C, 20 V/cm für 90 min. Nach Beendigung des Gellaufs wurde das Gel in 10% Ethanol, 0,5% Essigsäure 10 min fixiert und anschließend in 10% Glycerin für 10 min geschwenkt. Zur besseren Handhabung wurde das Gel über Nacht getrocknet und am folgenden Tag für 2h auf einem Phosphoimagerscreen exponiert.

#### 3.4.4 Bidirektionelle Gelelektrophorese

Zum Viroidnachweis aus Pflanzenmaterial kam eine bidirektionelle Gelelektrophorese zum Einsatz. Aufgrund seiner strukturellen Charakteristika kann das Viroid in diesem Gelsystem von allen anderen Nukleinsäurespezies separiert werden. Zunächst erfolgt unter nativen Bedingungen eine Trennung des stäbchenförmigen Viroids von größeren genomischen Nukleinsäuren (1. Dimension). Nach Einstellen von denaturierenden Bedingungen durch Temperaturerhöhung wurden sodann die Pufferbäder umgepolt (2. Dimension). Der Übergang in eine offenkettige, ringförmige Struktur retardiert das Viroid-Molekül drastisch verglichen zu kleinen Nukleinsäuren.

Als Gelbedingungen wurden 5% Polyacrylamid (43:1 Vernetzungsgrad), 4M Harnstoff sowie  $0.4 \times$  TBE als Gel- und Laufpuffer gewählt. In der ersten Dimension lief das Gel spannungsgesteuert (800 V) 75 min bei 15 °C. Das Gel wurde auf 40 °C erwärmt und nach Umpolung der Pufferbäder die "Umkehr"-Elektrophorese gestartet. Die Laufzeit in der zweiten Dimension betrug 45 min bei 800 V.

#### 3.4.5 Gelelutionen

Uniforme Nukleinsäurespezies wurden über eine Gelaufreinigung erhalten. Viroid-Transkripte und <sup>32</sup>P-markierte Nukleinsäure wurden in einem 1 mm dicken denaturierenden PAA-Gel aufgetrennt (siehe Abschnitt 3.4.2). Die korrespondierenden Banden wurden durch Ethidiumbromidfärbung oder Autoradiographie identifiziert und mit einem sterilen Skalpell ausgeschnitten. Die Gel-Stücke wurden in 300 µl Elutionspuffer (300 mM NaOAc, 1 mM EDTA, 1% SDS) mindestens 4 h geschüttelt. Viroid-Transkripte wurden einer Extraktion mit Phenol/Chloroform und Chloroform unterzogen. Anschließend wurde die Nukleinsäure aus dem Überstand durch Zugabe von Ethanol gefällt und abzentrifugiert. Das Pellet wurde in 10 mM Tris/HCl, pH 8.0 gelöst. Aus Agarosegelen wurde Nukleinsäure mit dem QIAquick Gel Extraktion Kit<sup>TM</sup>(QIAGEN, Hilden) gemäß den Herstelleranweisungen eluiert. Die Qualität aller Nukleinsäureeluate wurde in einem geeigneten Gelsystem und spektralphotometrisch untersucht.

#### 3.4.6 Silberfärbung von Polyacrylamid-Gelen

Mittels Silberfärbung kann in Polyacrylamid-Gelen weniger als 1 ng Nukleinsäure nachgewiesen werden. Das Protokoll wurde ursprünglich für Proteingele entwickelt (Sammons *et al.*, 1981), später für Nukleinsäuren modifiziert (Schumacher *et al.*, 1983a). Die zu färbenden Gele wurden in den nachfolgend aufgelisteten Lösungen inkubiert:

Fixierung	10~% Ethanol	$15\mathrm{min}$
	0,5~%Essigsäure	
Färbung	0,19 % Silberni trat	$30\mathrm{min}$
Waschen	aqua dest.	$3 \times 15 \mathrm{s}$
Entwicklung	1,5~%NaOH	$15\mathrm{min}$
	0,008 % NaBH <sub>4</sub>	
	$0,\!15~\%$ Formaldehyd	
Neutralisierung	0,75 % Na HCO $_3$	$10\mathrm{min}$

#### 3.4.7 Autoradiographie von Polyacrylamid-Gelen

Aufgrund der erhöhten Bandenschärfe und der besseren Handhabung wurden radioaktive PAA-Gele vor der Exposition getrocknet. Dazu wurden Protein-Gele in 50 % Methanol, 7 % Essigsäure und anschließend in 7 % Glyerin, 7 % Methanol, 7 % Essigsäure für jeweils 10 min fixiert. Nukleinsäure-Gele wurden in 10 % Ethanol, 0,5 % Essigsäure 10 min inkubiert. Es folgte jeweils eine Behandlung mit 10 % Glycerin für 10 min, um ein Brechen des Gels während des Trocknens zu verhindern. Die Gele wurden auf Whatman 3MM Papier aufgezogen und auf einem Geltrockner (Modell 1125 B, BIO-RAD) mit angeschlossener Membranpumpe für 3 h bei 80 °C getrocknet. Zum Nachweis <sup>32</sup>P-markierter Nukleinsäure und <sup>35</sup>S-markierten Proteinen wurden die getrockneten PAA-Gele auf Röntgenfilm (Amersham Hyperfilm und Kodak Xomat AR) mit Verstärkerfolie (Lightning Plus, Dupont) exponiert. Alternativ wurden die Gele auf Phosphoimager-Screens und die Signale mit einem Bioimager (FAS 3000, Fuji) detektiert.

#### 3.4.8 SDS-Polyacrylamidgelelektrophorese

Mit dem von Laemmli *et al.* (1970) entwickelten diskontinuierlichen Gelsystem können Proteine gemäß ihres Molekulargewichts aufgetrennt werden. Die regelmäßige Anlagerung von SDS an das Peptidrückgrat des entfalteten Proteins erzeugt ein konstantes Masse/Ladungs-Verhältnis. Die Elektrophorese wurde in einer kühlbaren Gelapparatur der Firma Hoefer Scientific Instruments durchgeführt. Das temperierbare Pufferbad des Anodenraumes kühlt gleichsam das vertikal positionierte Gel. Somit kann ein hoher Stromfluss eingestellt und eine kurze Gellaufzeit erreicht werden. Zur Herstellung vertikaler Plattengele (150 mm x 160 mm x 1,1 mm) wurden 20 ml Trenngellösung eingefüllt und mit wassergesättigtem Isobutanol überschichtet. Nachdem die Polymerisation beendet war, wurde das Isobutanol entfernt und dreimal mit Wasser gespült. Abschließend wurden 5 ml Sammelgellösung auf das Trenngel geschichtet. Als Molekulargewichtsstandard wurde wahlweise Mark12<sup>TM</sup>(Invitrogen) oder Rainbow<sup>TM</sup>(Amersham Pharmacia Biotech) eingesetzt. Die Elektrophorese lief bei 35 V/cm,  $10 \,^{\circ}\text{C}$  für 90 min.

#### Trenngellösung

$375~\mathrm{mM}$	Tris/HCl, pH 8.8
10-14 %	Acrylamid/Bisacrylamid (30:0,8)
$0{,}1~\%$	SDS
$0{,}1~\%$	TEMED
$0{,}03~\%$	APS

#### Sammelgellösung

$125~\mathrm{mM}$	Tris/HCl, pH 6.8
3-5 %	Acrylamid/Bisacrylamid (30:0,8)
$0{,}1~\%$	SDS
$0{,}1~\%$	TEMED
$0{,}03~\%$	APS

#### $\mathbf{2} imes \mathbf{SDS-Auftragspuffer}$

125  mM	Tris/HCl, pH 6.8
5 %	SDS
30~%	Glycerin
5~%	$\beta$ -Mercaptoethanol
0,02~%	Bromphenolblau

#### Laufpuffer nach Laemmli

25  mM	Tris
$0{,}1~\%$	SDS
$190 \mathrm{mM}$	Glycin

#### 3.4.9 Färbung mit Coomassie-Blau

Das Sammelgel wurde mit einem Kleenextuch vorsichtig entfernt und das Trenngel für 10 min in 0,1% (w/v) Coomassie Blue R250, 45% Methanol und 10% Essigsäure geschwenkt. Die Entfärbung in 10% Methanol und 10% Essigsäure dauerte mehrere Stunden, konnte aber durch Zugabe von Zellstoff als Adsorbens beschleunigt werden. Als einfache und nichttoxische Alternative erwies sich, das Gel in H<sub>2</sub>O für 15 min in einem handelsüblichen Mikrowellenherd zu kochen.

#### 3.4.10 Semi-Dry-Blot (Western-Blot)

Mittels Western-Blot können Proteine auf Membranen aus Polyvinyldifluorid (PVDF) übertragen werden. Dazu wurde das Trenngel, Chromatographie-Papier (3MM Chr, Whatman) und eine Ethanol- benetzte PVDF-Membran ( $45 \,\mu$ m, Millipore GmbH, Neu Isenburg) in Laemmli-Laufpuffer ohne SDS geschwenkt. Die Komponenten wurden in einer Semi-Dry Electrophoretic Transfer Cell (BIO-RAD, Hercules, USA) in folgender Reihenfolge geschichtet: Kathode, drei Lagen Chromatographiepapier, Trenngel, PVDF-Membran, drei Lagen Chromatographiepapier, Anode. Der Transfer wurde für 1 h bei 1,5 mA/cm<sup>2</sup> und maximal 25 V durchgeführt.

#### 3.4.11 Immunologischer Proteinnachweis

Zunächst wurde die PVDF-Membran in TBST (10 mM Tris/HCl, pH 8,0, 150 mM NaCl, 0,01 % Tween 20) und 5 % (w/v) Milchpulver 1 h inkubiert. In diesem als *blocking* bezeichneten Schritt werden freie Proteinbindeplätze belegt. Anschließend wurde der Lösung Antiserum immunisierter Kaninchen in einer Konzentration von 1:50 bis 1:400 hinzugefügt und über Nacht bei 4 °C sanft geschwenkt. Ungebundene Antikörper wurden durch ausgiebiges Waschen mit TBST entfernt. Es folgte eine Behandlung mit dem Zweitantikörper GAR (Peroxidase gekoppelter Ziege-anti-Maus-Antikörper) in einer Verdünnung von 1:5000 in TBST. Nach erneutem Waschen in TBST wurden die Proteine mittels Chemolumineszenz detektiert. Hierbei kam das ECL-Plus<sup>TM</sup>Detektions Reagenz (Amersham Bioscience, Buckinghamshire, UK) zum Einsatz. Die Detektion erfolgte nach Angaben des Herstellers.

# 4

# Ergebnisse

Da Viroide nicht für Proteine kodieren, müssen sie mit Wirtsfaktoren interagieren um in der Pflanze repliziert und transportiert zu werden. Basierend auf strukturellen Erkenntnissen kann für einige Wirtsproteine eine funktionelle Beziehung zur Viroid-RNA postuliert werden. Das zentrale Thema dieser Arbeit bilden Experimente, in denen solche hypothetischen Wechselwirkungspartner von PSTVd synthetisiert wurden, um ihr Bindungsverhalten an die Viroid-RNA zu untersuchen. Anhand der ermittelten Daten sollten vorhandene Modellsysteme zu Replikation und Transport von PSTVd überprüft werden.

Abschließend wird der eher zufällige Befund weiter analysiert, dass die mRNA eines überaus prominenten Transkriptionsfaktors, namentlich TFIIIA, alternativ gespleißt wird. Dieses Phänomen, das in höheren Pflanzen offenbar konserviert ist, konnte in der vorliegenden Arbeit experimentell behandelt werden.

Zu Beginn des Ergebniskapitels werden zunächst Mutagenesestudien präsentiert, die methodisch in diese Arbeit passen und die kurzfristig wichtig wurden um die genaue Startstelle der Viroidreplikation zu charakterisieren.

# 4.1 Infektiosität von punktmutierten Viroiden

Im ersten Schritt des Replikationszyklus von PSTVd wird der (+)-strängige Zirkel in einen multimeren (-)-Strang übersetzt. In einer vorangehenden Arbeit wurde die Startstelle der Transkription bestimmt (Kolonko, 2003). Dabei wurde einem transkriptionsaktiven Kernextrakt, präpariert aus *Solanum tuberosum*-Zellkulturen, exogenes *template* in Form von

 $1G \stackrel{G}{=} C^{G} GAACUAA^{A_{10}}$  $359G \stackrel{G}{=} U_{C} CUUGGUU_{G_{350}}$ 

Abbildung 4.1: Position der Mutationen im linken terminalen Loop von PSTVd. Für die Analyse der Infektiosität *in planta* wurden zwei punkmutierte PSTVd-Moleküle synthetisiert. Dabei wurden die potentiellen Startstellen U359 bzw. C1 durch G359 bzw. G1 ersetzt.

(+)-strängigen PSTVd-Zirkeln hinzugefügt. *De-novo*-synthetisierter (-)-Strang wurde affinitätschromatographisch aufgereinigt und dessen 5'-Ende mittels *primer extension*-Analyse bestimmt. Als Startpunkt dieser Synthese des (-)-Strangs wurde der linke terminale Loop des nativen PSTVd-Moleküls identifiziert(siehe Abb. 4.1). Aufgrund der niedrigen Auflösung des Sequenziergels war es hingegen nicht möglich, das *primer extension*-Signal einem der Nukleotide U359 oder C1 im linken terminalen Loop zuzuordnen. Zur näheren Charakterisierung des Startnukleotids wurden in der vorliegenden Arbeit gerichtete Mutagenese-Studien durchgeführt. Dabei sollten die potentiellen Startnukleotide U359 und C1 jeweils durch ein Guanin ersetzt werden und die Infektiosität dieser Mutanten *in planta* analysiert werden. Es wurden Wachstumskinetiken der Viroidpopulationen aufgenommen und die genetische Stabilität der Mutationen durch Sequenzanalyse bestimmt.

#### 4.1.1 Synthese und Inokulation

Als Inokulum wurden *in vitro*-Transkripte mit mutierter PSTVd-Sequenz verwendet, die von punktmutierten pRZ6-2-Plasmiden, sogenannte Feldsteinvektoren (Owens *et al.*, 1995), gewonnen wurden. In diesem Vektorsystem wird eine über *Bam*HI klonierte PSTVd-Sequenz flankiert von speziell modifizierten Versionen des *hammerhead*- bzw. *paperclip*-Ribozyms. Bei *in vitro*-Transkription mit erhöhter MgCl<sub>2</sub>-Konzentration spalten sich diese Ribozymsequenzen autokatalytisch ab. Dabei entlassen sie die lineare PSTVd-Sequenz exakter Einheitslänge, deren endständige Nukleotide G<sub>88</sub> bzw. G<sub>87</sub> ein 5'-Hydroxyl bzw. ein 2', 3'-Cyclophosphat aufweisen. Die besondere Beschaffenheit der freien Enden ermöglicht *in planta* eine Ligation zum reifen, zirkulären Viroid. Daher sind diese synthetischen, linearen Transkripte gleich infektiös wie natürliche, zirkuläre PSTVd-Moleküle (Feldstein *et al.*, 1998).

Jeweils vier 14 Tage alte Tomatenpflanzen wurden mit 5 ng Wildtyp-PSTVd (*inter-mediate* DI) und den Mutanten G359 und G1 mechanisch inokuliert. Als Nullkontrolle wurden vier weitere Pflanzen mit Phosphatpuffer scheininfiziert. Blattproben wurden



Abbildung 4.2: Nachweis der G359-Nachkommenschaft, 28 Tage post infectum. Bidirektionelles, 5% iges PAA-Gel (30:1), 0,4 x TBE, 4 M Urea, silbergefärbt. Aufgetragen sind jeweils 2 und 10  $\mu$ g Gesamtnukleinsäure von der scheininfizierten Pflanze (Null) und den mit G359 infizierten Pflanzen 9-12 (P9-P12), von der mit Wildtyp-PSTVd (*Intermediate*DI) infizierten Pflanze (WT) sind 1, 3 und 10  $\mu$ g Gesamtnukleinsäure aufgetragen. Das viroidspezifische Signal ist mit einem Stern (\*) gekennzeichnet.

21, 25, 28 und 35 Tage *post infectum* (d.p.i.) in Nähe der Pflanzenspitze entnommen und deren Viroidgehalt in einem bidirektionellen Gel semi-quantitativ bestimmt (siehe Abb. 4.2). Durch Vergleich der Bandenintensität von Mutante und Wildtyp wurde die relative Akkumulation der Nachkommenschaft ermittelt (relative Akkumulation = <u>Bandenintensität Mutante</u>).

#### 4.1.2 Wachstumskinetik der PSTVd-Mutanten

Der Titer des Wildtyps erreichte 21 d.p.i. einen Maximalwert. Durch Bandenvergleich mit einem spektralphotometrisch quantifiziertem Standard wurde die Viroidkonzentration auf 0,7 ng/mg Blattmaterial abgeschätzt (Daten nicht gezeigt).

Die Viroidakkumulation in den mit Mutanten inokulierten Pflanzen ist in Abb. 4.3A und 4.4A dargestellt. Es zeigte sich, dass bereits drei Wochen d.p.i. in der mit G359 infizierten Pflanze 12 ein viroidspezifisches Signal zu detektieren war. Dies ist ein starker



Abbildung 4.3: Infektiosität der Mutante G359. A) Wachstumskinetik. Gezeigt ist die relative Akkumulation der G359-Nachkommenschaft, bestimmt durch Vergleich der Bandenintensität von Mutante und Wildtyp in einem silbergefärbten, bidirektionellen Gel. Die mit einem Kreis markierten Proben wurden einer Sequenzanalyse unterzogen. B) Sequenz der G359-Nachkommenschaft in Pflanze 12, 21 d.p.i, Elektropherogramm. Die Sequenzanalyse zeigt, dass nach drei Wochen in Pflanze 12 die Mutation G359 zum Wildtyp C359 revertierte. Denselben Befund lieferten alle sequenzierten Produkte der G359-Nachkommenschaft.

Hinweis auf eine sehr frühe Reversion, da in den übrigen drei Pflanzen (9, 10 und 11) kein Viroid nachgewiesen werden werden konnte.

Dahingegen zeigte die Analyse 25 d.p.i, dass schon drei von vier Pflanzen die Mutante G1 replizieren – wenngleich mit niedrigerer Effizienz als den Wildtyp –, was auf eine genetisch stabile Mutation schließen lässt. Dafür spricht weiterhin, dass diese Mutante auch 35 Tage *post infectum* nicht den maximalen Titer des Wildtyps erreichte (siehe Abb.). Im Gegensatz dazu wiesen nach 35 Tagen die G359-inokulierten Pflanzen 9 und 12 ein Viroid-Signal auf, welches der Wildtyp-Konzentration entspricht. Hier könnte eine Reversion von G359 zum Wiltyp-U359 stattgefunden haben. Über den gesamten Beobachtungszeitraum konnten keine viroidspezifischen Signale in den scheininfizierten Pflanzen detektiert werden.

#### 4.1.3 Genetische Stabilität der PSTVd-Mutanten

Zur Analyse der genetischen Stabilität der Punktmutationen musste die Sequenz der Viroidspezies in den verschiedenen Pflanzenextrakten bestimmt werden. Aus den RNA-



Abbildung 4.4: Infektiosität der Mutante G1. A) Wachstumskinetik. Gezeigt ist die relative Akkumulation der G1-Nachkommenschaft, bestimmt durch Vergleich der Bandenintensität von Mutante und Wildtyp in einem silbergefärbten, bidirektionellen Gel. Die mit einem Kreis markierten Proben wurden einer Sequenzanalyse unterzogen. B) Sequenz der G1-Nachkommenschaft in Pflanze 5, 21 d.p.i, Elektropherogramm. 21 Tage *post infectum* ist mittels RT-PCR die mutante Sequenz nachzuweisen. C) 35 d.p.i. wurde in Pflanze 5 ein schwaches Wildtypsignal (C1) beobachtet, welches gleichwohl vom Signal des mutierten G1 überlagert wird. In allen übrigen Nukleinsäureproben konnte auschließlich die G1-Sequenz gefunden werden.

Extrakten, die in den Abbildungen 4.3A und 4.4A mit Punkten markiert sind, wurde per RT-PCR Viroid-cDNAs hergestellt, welche anschließend bei der Firma Seqlab (Göttingen) sequenziert wurden. Die Sequenzierungen zeigten jeweils ein einheitliches Bild: Die Mutation G359 revertierte ausnahmslos zum Wildtypnukleotid U359. Die Mutation G1 erwies sich in allen untersuchten Proben als genetisch stabil. Lediglich in einer einzigen Blattprobe war 35 d.p.i. ein schwaches Wildtypsignal zu detektieren. Es ist anzunehmen, dass in dieser Pflanze die Wildtyppopulation die G1-Population sehr rasch überwachsen wird.

Offenbar lastet ein sehr viel stärkerer Reversionsdruck auf G359 als auf G1, so dass man – vor dem Hintergrund der biochemischen Daten – der Position U359 eine Funktion beim Transkriptionsstart zuschreiben kann. Neuste Erkenntnisse über die Struktur der Startdomäne werden im Rahmen der Diskussion erörtert. Darüber hinaus wird ein mechanistisches Model zur Reversion mutierter Nukleotide entwickelt (siehe Kapitel 5.1.2).

# 4.2 Analyse von Viroid-Protein-Wechselwirkungen

Für die Interaktion von Viroiden mit Wirtsproteinen wurden bislang empirische Methoden angewendet und dadurch die Identität einiger Wechselwirkungspartner aufgeklärt.

In der vorliegenden Arbeit wurde ein hypothetischer Ansatz verfolgt. Es sollten einige Proteine, für die basierend auf strukturellen Erkenntnissen eine Wechselwirkung mit dem Viroid angenommen werden kann, rekombinant exprimiert und ihre Bindung an das Viroid analysiert werden. Das Tata Binding Protein wurde ausgewählt, da es potentiell bei der Pol II-vermittelten Transkription des Viroids fungiert. Es wird vermutet, dass die Viroid-RNA Promotorstrukturen präsentiert, über die eine korrekte Initiation der Transkription gewährleistet wird. Der Transkriptionsfaktor IIIA und das ribosomale Protein L5 stehen in direktem Zusammenhang mit dem intrazellulären 5 S-RNA-Transport. Da PSTVd 5Sähnliche Strukturmotive besitzt und überdies eine direkte Wechselwirkung mit der 5 S-RNA nachgewiesen wurde, könnte sich auch das Viroid dieser Faktoren für seine zelluläre Lokalisation bedienen.

#### 4.2.1 Sequenzauswahl der potentiellen Wechselwirkungspartner

#### **Tata Binding Protein**

Die DNA-Bindekomponente des Transkriptionsfaktors IID, das sogenannte Tata Binding Protein (TBP), ist ebenso wie das korrespondierende Tata-Motiv der DNA in der belebten Natur stark konserviert. Dies vereinfacht die Suche nach entsprechenden Gensequenzen, und es ist somit wenig verwunderlich, dass TBP bereits für eine Vielzahl von Organismen beschrieben und annotiert wurde. Unter anderem wurde TBP aus *Solanum tuberosum* charakterisiert (Holdsworth *et al.*, 1992) und die vollständige Gensequenz in der Datenbank abgelegt (Zugriffsnr. X62494). Da die Kartoffel aus der Familie der Nachtschattengewächse eine natürliche Wirtspflanze des PSTVd repräsentiert, wurde dieses Protein für die Klonierung aus cDNA und die rekombinante Expression in *E. coli* ausgewählt. Das Tata Binding Protein aus Kartoffel wird im Folgenden mit *St*TBP abgekürzt.

#### Trankriptionsfaktor IIIA und ribosomales Protein L5

Zu Beginn der vorliegenden Arbeit kannte man lediglich eine einzige pflanzliche Sequenz des Transkriptionsfaktors IIIA (TFIIIA). Diese wurde gemeinsam mit der Sequenz des ribosomalen Proteins L5 aus Arabidopsis thaliana bestimmt (Mathieu et al., 2003). Im Folgenden werden beide Proteine als AtTFIIIA bzw. AtL5 bezeichnet. Obwohl Arabidopsis keine natürliche Wirtspflanze des Viroids ist, erschien es dennoch sinnvoll beide Sequenzen für die rekombinante Expression zu klonieren: Arabidopsis thaliana verfügt über die enzymatische Maschinerie, um PSTVd zu replizieren. Eine systemische Infektion scheitert, da der Transport des Viroids über weite Strecken in Arabidopsis nicht zu funktionieren scheint (Daròs & Flores, 2004). Die Arabidopsis-Sequenzen wurden gewissermaßen in Reserve vorgehalten, falls es zu Schwierigkeiten bei der im Folgenden beschriebenen Identifikation oder Präparation der TFIIIA- und L5-Sequenzen aus der Wirtspflanze Tomate gekommen wäre.

#### 4.2.2 Identifikation von TFIIIA und L5 aus L. esculentum

Um von Ergebnissen aus Wechselwirkungsstudien *in vitro* auf die Situation *in vivo* schließen zu können, ist es vorteilhaft, rekombinante Proteine aus einer Wirtspflanze zu verwenden. Daher wurde versucht, die Sequenzen von TFIIIA und L5 aus dem natürlichen Wirt *Lycopersicon esculentum* zu ermitteln. Analog der für die anderen Proteine eingeführten Terminologie werden Transkriptionsfaktor IIIA und ribosomales Protein L5 aus *Lycopersicon esculentum* nachfolgend kurz als *Le*TFIIIA und *Le*L5 bezeichnet.

Ein ClustalW-Alignment (Thompson *et al.*, 1994) von bekannten TFIIIA-Proteinsequenzen von Frosch, Mensch und Zebrafisch ergab eine recht schwache Konservierung von lediglich 50%. Vergleicht man die bisher einzig bekannte pflanzliche TFIIIA-Sequenz aus Arabidopsis thaliana mit den tierischen Sequenzen, so zeigt sich eine Übereinstimmung von gar nur 20%. Die Ähnlichkeit beschränkt sich weitestgehend auf die neun Zink-Finger-DNA-Bindemotive, die in allen bekannten TFIIIA-Sequenzen zu finden sind. Eine Ausnahme bilden TFIIIA-Sequenzen einiger Hefearten, die aus zehn Zink-Fingern bestehen (Schulman & Setzer, 2002). Augrund der schwachen Konservierung schien eine Identifikation mittels Hybridisierungsverfahren wenig aussichtsreich. Wegen der großen Zahl annotierter EST-Sequenzen (*expressed sequence tags*) aus Tomate war es erfolgsversprechend die LeTFIIIA-Sequenz durch Datenbankanalyse zu bestimmen. Dazu wurde das Programm BLAST (*Basic Local Alignment Search Tool*; Altschul *et al.*, 1990), ein *online*-Dienst des NCBI eingesetzt. Dieser Algorithmus ermittelt Regionen, die lokale Ähnlichkeiten zwischen Sequenzen besitzen. Dabei vergleicht es vorgegebene Nukleinsäure- oder Pro-



Abbildung 4.5: Identifikation eines LeTFIIIA-ORF. Mit der annotierten TFIIIA-Proteinsequenz aus Arabidopsis thaliana (AAO73339) wurde nach EST-Sequenzen aus Lycopersicon esculentum gesucht, die für homologe Proteine kodieren. Anhand der gefundenen Datenbankeinträge AW930406, BI926029 und BG129248 war es möglich, ein durchgängiges Leseraster bis zum achten Zink-Finger-Motif zu konstruieren. Das fehlende 3'-Ende konnte experimentell in einer 3'-RACE-Reaktion ermittelt werden. Prozentangaben: Sequenzübereinstimmung auf Proteinebene. Die Zahlen 1-9 repräsentieren die Zink-Finger-Motive des TFIIIA.

teinsequenzen mit Sequenzen aus Datenbanken und bewertet die statistische Signifikanz von Übereinstimmungen.

Zunächst wurde versucht aus den bereits annotierten EST-Sequenzen von Lycopersicon esculentum ein möglichst vollständiges Leseraster (ORF, für engl. open reading frame) zu identifizieren, welches für neun Zink-Finger-DNA-Bindemotive kodiert. Eine tblastn-Suchanfrage mit der AtTFIIIA-Proteinsequenz lieferte lediglich unvollständige mRNA-Sequenzen, die auf Proteinebene aber signifikante Homologien aufwiesen [siehe Abb. 4.5: AW930406 (52 % Übereinstimmung), BI926029 (44 %), BG129248 (48 %)].

Darüber hinaus überlappten sich die Sequenzen, so dass eine Zusammenfügung zu einem durchgängigen Leseraster ohne Stopp-Codon möglich war. Auf diese Weise konnte eine *Le*TFIIIA-Sequenz vom Start-Methionin bis hin zum achten Zink-Finger-Motiv nahegelegt werden. Lediglich der C-Terminus des Proteins konnte nicht durch korrespondierende mRNA-Sequenzen abgedeckt werden. Dies könnte darin begründet sein, dass die entsprechende Sequenz noch nicht in der EST-Datenbank abgelegt wurde, oder diese aufgrund einer sehr schwachen Konservierung des C-Terminus schlichtweg bei der BLAST-Suche nicht gefunden werden konnte. Das fehlende 3'-Ende konnte schließlich in Kooperation mit Dr. Axel Schmitz in einer 3'-RACE-Reaktion (*rapid ampflification of cDNA ends*) ausgehend von einer *Lycopersicon esculentum*-cDNA-Bank amplifiziert und sequenziert werden.

L.esculentum	0		-MGEV	/GEE
A.thaliana	0		MAE-EA	AKVD
O.sativa D.sativa	0	MG	SVELGAE	SERD
D.rerio	0			
A.Idevis M. muggulug	0			. — — — — т. 🗔
M.Musculus H.ganieng	0		BGLGGAG	
n.sapiens	0	MRSSGRDAGRCLVIARAF-GSVFASR-EGSAGSRGFGARFFARVSARGSAFG	FGLGGAG	ALL
L.esculentum	7	KREVVFRDIRRYYCEECGLCRSKK	SLISS	OS
A.thaliana	8	VKTSAKKDIRNYLCOYCGISRSKN	JYLITKHI	0s
O.sativa	12	VACSVVRSK	GLIRAHV	<b>LEH</b>
D.rerio	0	GETETCSYPDCHAYYNRE	WKLQAHI	скн
X.laevis	0	WGEKALPVVYKRYICSFADCGAAYNKN	JWKLQAHI	гскн
M.musculus	38	PRVSVAEAVSSLTIADAFVGACEGPAPPRPALPSRFICSFPDCSASYNKA	WKLDAHL	гскн
H.sapiens	60	PPAVVAESVSSLTIADAFIAAGESSAPTPPRPALPRRFICSFPDCSANYSK	WKLDAHI	гскн
		2		
L.esculentum	41	HQDVIESLKDDATENAKGSKMNICEECGATFQKPAHLKQHLQSHSL	ERPFVCHI	DDC
A.thaliana	42	HQMELEEERDDEACEVDEESSSNHT-CQECGAEFKKPAHLKQHMQSHSLE	ERSFICYV	DDC
O.sativa	55	HKDEVDDLDDYLGRGGGETCKEMDHDCKVCGASFKKPAHLRQHMQSHSL	RPFSCHV	DGC
D.rerio	41	TGERPYKCKYKKCSKSFCTKHHLSRHVLTHTGE	SKPYRCME	DGC
X.laevis	37	TGEKPFPCKEEGCEKGFTSLHHLTRHSLTHTGE	KNFTCDS	BDGC
M. musculus	100		K P F V C A L	NGC NGC
n.sapiens	122	3 4	<b>KPFVCAA</b>	IN GC
L.esculentum	98	QSSYRRKDHLSRHLLL-HQGKLFECPVDGCKSTFSFQGNMTRHVKEMH		
A.thaliana	101	AASYRRKDHLNRHLLT-HKGKL TKCPKENCKSEFSVQGNVGRHVKKYHSN	IDNRDKDN	ITGL
0.sativa	115	PFSYSRKDHLNRHLLT-HQGKL TACPMEGCNRKFTIKGNIQRHVQEMH		
D.rerio	84	KEGFTTNSNLQKHISRIHRQETKQYICTFEGCGKAFKKNNQLKTHE-CTH		
X.laevis	80	DLRFTTKANMKKHFNRFHNIKICVYVCHFENCGKAFKKHNQLKVHQ-FSH		
M.musculus	141	NQKFNTKSNLKKHIERKHGNPQKQYLCSYEGCKKAFKKHQQLRTHQ-CQH		
H.sapiens	165	DQKFNTKSNLKKHFERKHENQQKQYICSFEDCKKTFKKHQQLKIHQ-CQN		
L.esculentum	145		PKEYVCS	EAG
	160	GDGDKDNTCKGDDDKEKSGSGGCEKENEGNGGSGKDNNGN-DSQPAEC-SMG	JOKOVVCK	EIG
	1 ( )	"Daabaha"	WWRRTON	1 10 10 11
O.sativa	162	KDGSPCES	KKEFICF	EEN
O.sativa D.rerio	162 133		-KKEFICF Q-LPFLCT	PEEN QEG
O.sativa D.rerio X.laevis	162 133 129		-KKEFICF 2-LPFLCI 20LPYECF	EEN QEG HEG
O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 190 214		- KKE FICF Q-LPFLCT QQLPYECF SEPLFRCT	EEN QEG HEG HEG
O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 190 214		-KKEFICE 2-LPFLCT 2QLPYECE 3EPLFRCT 3EPLFKCT	PEEN QEG HEG HEG QEG
D.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum	162 133 129 190 214 165		- KKEFICE QLPFLCT QLPFECE SEPLFRCT IEPLFKCT IQ-HIDCE	PEEN QEG HEG HEG QEG
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana	162 133 129 190 214 165 221	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH	- KKEFICE 2 - LPFLCI 2 QLPYECE 5 EPLFRCI NEPLFKCI 1 Q - HIDCE 1 Q - HINCE	PEEN QEG HEG QEG QEG
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa	162 133 129 190 214 165 221 180	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH	- KKEFICE 2 - LPFLCI 2 QLPYECE 5 EPLFRCI NEPLFKCI 1 Q - HIDCE 1 Q - HINCE 1 R - HVVCI	EEN QEG HEG HEG IQEG ICG ICG
O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio	162 133 129 190 214 165 221 180 144	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH	- KKEFICE - LPFLCT QLPYECE SEPLFRCT NEPLFKCT NC-HIDCE NC-HINCE NC-HVVCT IVRVQCT	PEEN QEG HEG QEG I CG I CG I CG V CG QCQ
O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis	162 133 129 190 214 165 221 180 144 141	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKVFKYASKLRKHEDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKTFKYASKLQKHESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYPCKKDDSCSFVGKTWTLYLKHVAECH	KKEFICF D-LPFLCT QQLPYECE SEPLFRCT VEPLFKCT Q-HIDCE Q-HIDCE R-HVVCT VVCT VVCC VQLAVCT	PEEN QEG HEG QEG ICG ICG ICG QCQ QCQ
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus	162 133 129 190 214 165 221 180 144 141 202	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKTFKYASKLQKHESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCKKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHAGYLCQKGCSFMGKTWTELLKHMREAH	KKEFICF OLPFLCT SEPLFRCT VEPLFKCT QOLHIDCE R-HVVCT VCDLAVCT KEDITCN	PEEN QEG HEG QEG ICG ICG VCG VCG VCN
A.thaiian O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 190 214 165 221 180 144 141 202 226	5 CGKVFKYASKLRKHEDSHVKLETVE CGKAFKYPSOLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSOLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH CGKTFKYASKLQKHESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTBMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTBMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCKKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFMGKTWTELLKHMREAH CGKHFASPSKLKRHAKAHEGYVCQK-GCSFVAKTWTELLKHVRETH	KKEFIC - LPFLCT SEPLFRCT NEPLFKCT Q-HIDCE Q-HIDCE R-HVVCI HVRVQCI QDLAVCI KEDITCN KEEILCE	PEEN QEG HEG HEG ICG ICG VCG QCQ VCN VCQ VCR
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum	162 133 129 190 214 165 221 180 144 141 202 226 225	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTBMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCCET-EGCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHAGYCCQK-GCSFWGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYUCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEG	KKEFIC - LPFLCT 20LPFCT SEPLFRCT NEPLFKCT 10-HINCE IR-HVVCI IR-HVVCI IVRVQCI IVRVQCI IVRVQCI IVRELICE SKPFSCS	
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana	162 133 129 190 214 165 221 180 144 141 202 226 225 281	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH CGKTFKYASKLQKHESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CGKHFASPSRLKRHEKVHAGYLCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVAKTWTELLKHMREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVAKTWTELLKHWRETH 7 7 1 TK-QLKKNIKRHLRTHEGPISERTKCEFQDCLHTFSSRSNLAQHIKAVHVC	KKEFICF - LPFLCT 20LPFCT SEPLFRCT VEPLFKCT 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HIDCE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40-HICE 40	PEEN QEG HEGG HEGG ICGG VCG VCG VCC VCR VCR
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa	162 133 129 190 214 165 221 180 144 141 202 226 225 281 240	5 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKYFTNEECLKSHIRSCH CGRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYLCQK-GCSFVAKNWTELLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVAKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVAKTWTELLKHVRETH 7 7 8 TK-QLKKNIKRHLRTHEEGPISERIKCEYEGCSSTFSKASNLQKHMKAVHDU TK-QLKKNFKRHQRMHEGSCVTERVRCHLKDCKLSFSKKSNLDKHVKAVHEC	KKEFICF - LPFLCT 20LPFCT SEPLFRCT IEPLFKCT IQ-HIDCE IQ-HIDCE IQ-HINCE IX-HVCT IXEDITCN KEEILCE SKPFSCS IRPFVCC IRPFVCC	PEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHEEGG PHE
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio	162 133 129 190 214 165 221 180 144 141 202 226 225 281 240 199	5       6         CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH       TO         6       5         6       6         CGKAFKYPSOLOKHODSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH       CGKTFKYASKLCKHESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH         CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH       CGKTFKYASKLQKHESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH         CGKTFSDPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH       CDKRFSLPSRLKRHEKVHAGYLCQK-GCSFVGKTWTLYLKHVAECH         CGKHFASPSRLKRHEKVHAGYLCQK-GCSFVGKTWTELLKHMREAH       CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTELLKHMREAH         CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVAKTWTELLKHVRETH       7         7       8         TK-QLKKNIKRHLRTHEGPISERTKCEFQDCLHTFSSRSNLAQHIKAVHOC       8         K-HLKKNIKRHLRTHEGSVTERVRCHLKDCKLSFSKASNLQKHMKAVHDC       8         KTFRDSWFLKQHQHVHSEERLV-FHCPRDGCTRSYTTAFNLQSHILSFHED       8	KEFICF CLPFLCT CLPFLCT CLPFLCT SEPLFRCT IEPLFKCT IQ-HIDCE IQ-HIDCE IR-HVVCT IUVVCT IUVVCT IUVVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT IUVCT	PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PERSON PE
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis	162 133 129 190 214 165 221 180 144 141 202 226 225 281 240 199 198	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTBMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYECKKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHAGYECKKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHAGYECQK-GCSFVAKNWTBLKHVRETH CGKHFASPSRLKRHGKVHEGYECQK-GCSFVGKTWTELLKHWREAH CGKHFASPSRLKRHGKVHEGYECQK-GCSFVGKTWTELLKHVRETH 7 8 TK-QLKKNIKRHLRTHEGFISER TK-QLKKNIKRHLRTHEGPISER KCEFQDCLHTFSSRSNLAOHIKAVHOT TK-QLKKNFKRHQRMHEGSCVTERVRCHLKDCKLSFSKASNLQKHMKAVHDT KTFRDSWFLKQHQHVHSEERLV-FHCPRDGCTRSYTTAFNLQSFHE	KEFICF - LPFLCT QLPFECT SEPLFRCT NEPLFKCT Q-HIDCE IQ-HIDCE IQ-HIDCE IR-HVVCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI ILVRVQCI	PEEGGG HHEEGG HHEEGG IICCGQCNQC ICCCQNQC SFSPAG HHA
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus	162 133 129 214 165 221 180 222 144 141 202 226 225 281 240 199 257	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYECKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHAGYECKDDSCSFVGKTWTELKHWREAH CGKHFASPSRLKRHEKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHEKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHEKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHEKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFWGKTWTELKHVKAVHC KTFFOSWFLKQHQHYSEERLVFHCPRDGCTRSYTTAFNLQSHILSFHEE RKFRHKDYLKQHMKTHAPERDVYRCPRQGCGRTYTTVFNLQSHILSFHEE	KEFICF LPFLCT LPFLCT SEPLFRCT VEPLFRCT VEPLFKCT VEPLFKCT VCL VCL VCL VCL VCL VCL VCL VCL	PEEGGG CGGGQC CCCCQC CCCCCC CCCCCCC CCCCCCC CCCCCCC
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 214 165 221 180 225 225 281 240 198 257 281	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCCKKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFVGKTWTELKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFVGKTWTELKHVRETH 7 7 8 TK-QLKKNIKRHLRTHEGGSCVTERVCHLKCCKLSFSKKSNLQKHWKAVHC KTFRDSWFLKQHQHVHSEERLVFHCPRDGCTRSYTTAFNLGSHILSFHEE RKFRHKDYLKQHMKTHAPERDVYRCPRQGCGRTYTTVFNLQSHILSFHEE KTFKRKDYLKQHMKTHAPERDVCRCPREGCGRTYTTVFNLQSHILSFHEE	KEFIC LPFLCT QLPYECF SEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VEPLFRCT VE	E E E E E E E E E E E E E E E E E E E
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 214 165 221 180 222 144 141 202 226 225 281 240 199 257 281 286	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFMGKTWTELLKHNREAH CGKHFASPSRLKRHGKVHEGYCQK-GCSFMGKTWTELLKHNREAH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFMGKTWTELLKHVRETH 7 7 7 8 TK-QLKKNIKRHLRTHEGPISERIKCEVEGCSSTFSKASNLQKHKAVHDI TK-QLKKNIKRHLRTHDEDSSPGEIKCEVEGCSSTFSKASNLQKHKAVHDI TK-QLKKNFKRHQRMHEGSCVTERVRCHLKDCKLSFSKKSNLDKHVKAVHEG KFFRHKDYLRDHQKHEKETV-YLCPRDGCTRSYTTAFNLRSHILSFHEE KFFRKRDYLKQHMKTHAPERDV-YRCPRQGCGRTYTTVFNLQSHILSFHEE KTFKRKDYLKQHMKTHAPERDV-CRCPREEVDFORLSPECOP	KKEFICF LPFLCT QLPFECT SEPLFRCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT SKPFSCS DIRPFVCC SKPFSCS DIRPFVCC SKPFVCE SRPFVCE SRPFVCE SRPFVCE SRPFVCE	
A.thaliana D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana	162 133 129 190 214 165 221 180 141 202 226 225 281 240 199 198 257 281 286	5 6 CGKVFKYASKLRKHEDSHVKLETVE CGKAFKYPSQLQKHQDSHVKLETVE CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKEHIESCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTN-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTN-KKVH CDKRFSLPSRLKRHEKVHAGYCCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVGKTWTELLKHVRETH 7 TK-QLKKNIKRHLRTHEGSPISERIKCEFQDCLHTFSSRSNLAQHIKAVHVC SK-HLKKNIKRHLRTHEGSVTERVRCHLKDCKLSFSKKSNLQKHMKAVHDI TK-QLKKNFKRHQRMHEGSCVTERVRCHLKDCKLSFSKKSNLDKHVKAVHEC RKFRHKDYLRQHQVHSEERLVFHCPRDGCTRSYTTAFNLQSHILSFHEC RKFRHKDYLRQHQKTHEKERTV-YLCPRQGCGRTYTTVFNLQSHILSFHEC SMFKRRDYLKQHMKTHAPERDVYCRPQGCGRTYTTVFNLQSHILSFHEC 9 CGMKFAFKHVRDKHEQSGCHVYTPGDFEEVDEQFLSRPRGGRKREPTLFE	KKEFICF CLPFLCT CLPFLCT CLPFLCT CLPFCT REPLFKCT CLPLFKCT CLPLFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT C	PEERCE QECGQ HEGGGGGGGG SICCGQ VCCQQ VCCQQ VCCQQ VCCQQ VCCQQ VCCQ SFSGGG HAGG HAGG HAGG
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio	162 133 129 190 214 165 221 180 141 202 226 225 281 240 199 1987 281 286 3401	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGKRFSDPSRLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVE CDKRFSLPSRLKRHEKVHEGYSCET-EGCSFVAKNWTEMTNH-KKVE CGKHFASPSRLKRHEKVHEGYCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSKLKRHAKAHEGYVCQK-GCSFVGKTWTELLKHMREAH CGKHFASPSKLKRHAKAHEGVVCQK-GCSFVAKTWTELLKHMREAH CGKHFASPSKLKRHAKAHEGYCQK-GCSFVAKTWTELLKHVRETH 7 7 8 TK-QLKKNIKRHLRTHEGSPISERIKCEFQDCLHTFSSRSNLAQHIKAVHVC KTFRDSWFLKQHQHVHSEERLVFHCPRDGCTRSYTTAFNLQSHILSFHEC KFRHKDYLRDHQKTHEKERTVYLCPRDGCDRSYTTAFNLQSHILSFHEC RKFRHKDYLKQHMKTHAPERDVYRCPRQGCGRTYTTVFNLQSHILSFHEE 8 TGGMKFAFKHVRDKHEQSGCHVYTPGDFEEVDEQFLSRPRGGRKREPTLFE CGMKFAFKHVRDKHENSGYHVYTCGDFVETDEDFTSRPRGGLKRKQVTAF	KKEFICF CLPFLCT CLPFLCT CLPFLCT CLPFCT KEPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLPFKCT CLFFKCT CLFFKCT CLFFKCT CLPFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFKCT CLFFFKCT CLFFFKCT CLFFFKCT CLFFFKCT CLFFFKCT CLFFFKCT CLFFFKCT CLFFFFKCT CLFFFFFF CLFFFFFF CLFFFFFFF CLFFFFFFFFFF	PEERCE QEGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio	162 133 129 214 165 2214 165 221 180 144 141 2026 225 2810 198 257 281 286 342 3019	5 6 CGKVFKYASKLRKHEDSHVKLETVE AMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSOLOKHODSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHINKSCH CGRRFSORGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVI CDKRFSLPSRLKRHEKVHAGYCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHEGYLCQK-GCSFVGKTWTELLKHMREAH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTELLKHMREAH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTELLKHMREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVGKTWTELLKHVRETH 7 TK-QLKKNIKRHLRTHEGPISERIKCEFQDCLHTFSSRSNLAOHIKAVHVC SK-HLKKNIKRHLRTHEGSVTERVRCHLKDCKLSFSKKSNLOKHMKAVHEO KTFRDSWFLKQHQHVHSEERLV-FHCPRDGCTRSYTTAFNLQSHILSFHEE RKFRHKDYLRDHQKTHEKERTV-YLCPRDGCDRSYTTAFNLQSHILSFHEE RMFKRRDYLKQHMKTHAPERDV-YRCPRQGCGRTYTTVFNLQSHILSFHEE KTFRLSVLKQHMKTHAPERDV-YCCRCPREGCGRTYTVFNLQSHILSFHEE CGMKFAFKHVRDKHEQSGCHVYTPGDFEEVDEQFLSRPRGCKKREPTLFE CGMKFAFKHVRDKHENSGYHVYTCGDFVETDEDFTSRPRGCLKRKQVTAF CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGERPRQAGGRKKAIPVG	KEFICF CLPFLCT CLPFLCT CLPFLCT CLPFLCT SEPLFRCT IPLFKCT IQ-HINCE IRPFVCC CLPFVCC IRPFVCC CLRFFVCC CRPFVCC CRPFVCC CRPFVCC CRPFVCC SSIMRKRI ESLMRKRI SSLMRKRV KTROTKT	PEERCE PHEECG CCCCCQN VCCQN VCCQN SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SFFSC SF
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 1299 214 165 2210 144 141 2026 225 2810 198 257 281 2409 198 257 281 2401 258 342 3019 258	5 6 CGKVFKYASKLRKHEDSHVKLETVE AMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGKRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYECKKDDSCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHEKVHAGYLCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTELLKHMREAH CGKHFASPSKLKRHAKAHEGYLCQK-GCSFVGKTWTELLKHVRETH 7 TK-QLKKNIKRHLRTHEGPISER TK-QLKKNIKRHLRTHEGPISER TK-QLKKNIKRHLRTHEGPISER KTFRDSWFLKQHQHVHSEERLV-FHCPRDGCTRSYTTAFNLQSHILSFHEG RKFRHKDYLKQHMKTHAPERDV-YLCPRDGCDRSYTTAFNLQSHILSFHEG RKFRHKDYLKQHMKTHAPERDV-YLCPRDGCDRSYTTAFNLRSHIQSFHEE RMFKRRDYLKQHMKTHAPERDV-YCCDFVETDEDFTSRPR-GCRKREPTLFE CGMKFAFKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GCLKRKQVTAF CGKSFSYKHVRDNHEKSSAHVYVQANFEELDGG	KEFICF CLPFLCT CLPFLCT CLPFLCT CLPFLCT SEPLFRCT IC - HIDCE IC - HIDCE	PEENCE PHEGG PHEGEG SICCGQCQQ VCCQQVCQQVCQVCQ VCCQCQVCQCQVCQC VCCQCQVCQCQUCQCQU
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 214 165 2210 144 141 202 226 225 281 2409 198 257 281 259 2581 259 259 2577	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKAFTNLECLKSHIRSCH CGKRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTELLKHWREAH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTLLKHVAECH CGKHFASPSRLKRHGKVHEGYLCQK-GCSFVGKTWTELLKHWREAH CGKHFASPSKLKRHAKAHEGYVCQK-GCSFVGKTWTELLKHVRETH 7 7 8 TK-QLKKNIKRHLRTHEGFUSERIKCEFQDCLHTFSSRSNLAOHIKAVHVC SK-HLKKNIKRHLRTHEGFUSERIKCEVEGCSSTFSKASNLQKHMKAVHDI TK-QLKKNFKRHQRMHEGSCVTERVRCHLKDCKLSFKKSNLDKHVKAVHDC KTFRDSWFLKQHQHVHSEERLV-FHCPRDGCDRSYTTAFNLRSHILSFHEF RKFRHKDYLKQHMKTHAPERDV-YRCPRQGCGRTYTTVFNLQSHILSFHEF 8 CGMKFAFKHVRDKHEQSGCHVYTPGDFEEVDEQFLSRPR-GGRKREPTLFH CGMRFAYKHVRNKHENSGYHVYTCGDFVETDEDFTSRPR-GGLKRKQVTAF CGKSFSYKHVRNHESSAHVYVQANFEEIDGERPRQAGGRKRKAIPVE CGKAFSMKQSLQRHGV-VHDPEKKQMKPRFKRSLASRLSGYKSF	KKEFICF LPFLCT LPFLCT LPFLCT SEPLFRCT LPLFKCT LPLFKCT LPLFKCT LPLFKCT LCF LCF LCF LCF LCF LCF LCF LCF	PEENCE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 214 165 221 184 141 202 226 225 281 240 198 257 281 259 2581 259 2587 341	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEECLKSHIRSCH CGKRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTLLKHVAECH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSKLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSKLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSKLKRHQKHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCCQK-GCSFVGKTWTELKHVKAVHEC 7 7 8 TK-QLKKNIKRHLRTHEGGSCVTERVRCHLKDCKLSFSKASNLQKHMKAVHDI KTFRDSWFLKQHQHWESEERLVFHCPRDGCTRSYTTAFNLQSHILSFHEE RMFKRRDYLKQHMKTHAPERDVYRCPRQGCGRTYTTVFNLQSHILSFHEE 8 FFKRKDYLKQHMKTHAPERDVGRCPREGCGRTYTTVFNLQSHILSFHEE 8 CGMKFASFKHVRDKHEQSGCHVYTPGDFEEVDEQFLSRPRGGLKRKQVTAF CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGFRPRKSLASRLSGYKSF CGKAFSMKQSLQRHGVVHDPEKKMKLKVRAPRFKRSLASRLSGYKSF CGKCFAMKKSLERHSVVHDPEKKKKKKK-SREKRSLASHLSGYIFF CGKTFAMKQSLTRHAVVHDPDKKKMKLKVKK-SREKRSLASHLSGYIFF	KKEFICF LPFLCT LPFLCT LPFLCT SEPLFRCT LPLFRCT LPLFRCT LPLFRCT LPLFRCT LPLFRCT LPLFRCT LCF LCF LCF LCF LCF LCF LCF LCF	PEENCE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE PHEGE
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 129 214 165 221 184 141 202 226 225 281 240 198 257 281 2581 259 2581 259 2587 341	5 6 CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH CGKAFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYSCET-EGCSFVAKNWTEMTNH-KKVH CDKRFSLPSRLKRHEKVHAGYCCQK-GCSFVGKTWTLYLKHVAECH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTLLKHVAECH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSRLKRHGKVHEGYCCQK-GCSFVGKTWTELKHWREAH CGKHFASPSKLKRHAKAHEGYCCQK-GCSFVGKTWTELKHVREAH CGKHFASPSKLKRHQKHEGYCCQK-GCSFVGKTWTELKHVREAH CGKHFASPSKLKRHAKAHEGYCCQK-GCSFVGKTWTELKHVREAH CGKHFASPSKLKRHAKAHEGYCCQK-GCSFVGKTWTELKHVREAH CGKHFASPSKLKRHAKAHEGYCCQCGSTFSKASNLQKHWKAVHDT TK-QLKKNIKRHLRTHEGSCVTERVRCHLKDCKLSFSKKSNLDKHVKAVHEC KTFRDSWFLKQHQHWSEERLVFHCPRDGCTRSYTTAFNLQSHILSFHEE RMFKRRDYLKQHMKTHAPERDVYCCPRQGCGRTYTTVFNLQSHILSFHEE RMFKRRDYLKQHMKTHAPERDVYRCPRQGCGRTYTTVFNLQSHILSFHEE SCGMKFASFKHVRDKHEQSGCHVYTPGDFEEVDEQFLSRPRGGLKRKQVTAH CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGERPRQAGGRKRKAIPVE CGKAFSMKQSLQRHGVVHDPEKKQMKPRPKRSLASRLSGYKSF CGKCFAMKKSLERHSVVHDPEKKKKKKCPRPKRSLASRLSGYKSF CGKTFAMKQSLTRHAVVHDPDKKKMKLKVRAPRERSLASRLSGYFFF	KKEFICF LPFLCT LPFLCT LPFLCT SEPLFRCT LPLFRCT LPLFRCT LPLFRCT LQ-HIDCE LR-HVCC LR-HVCC LR-HVCC LRFFVCC SKFFSCS LRFFVCC SKFFSCS LRFFVCC SKFFSCS LRFFVCE SKPFVCE SKRFFVCE SSRFFVCE SSLMRKRI KRKQEPC KSKEKNA KSKEKNA	PEENCE PHEGE HEGEG SICGGEG VCGQ VCCQ VCCQ VCCQ VCCQ VCCQ VCCQ VCC
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana	162 133 129 214 165 221 184 141 202 226 225 281 240 199 257 281 257 281 257 281 257 281 257 281 257 2581 257 346 346 346 346 346 346 346 346 346 346	5       6         CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH       7         CGKAFKYPSOLOKHODSHVKLDSVEAFCS-EPGCMKYFTNEECLKSHIRSCH       7         CGKAFKYPSOLOKHODSHVKLDSVEAFCS-EPGCMKKFTNEECLKSHIRSCH       7         CGKRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH         CDKRFSLPSRLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH         CGKHFASPSRLKRHEKVHAGYDCKKDDSCSFVAKNWTELLKHWRACH         CGKHFASPSRLKRHEKVHAGYDCKKDDSCSFVAKNWTELLKHWRATH         CGKHFASPSRLKRHEKVHAGYDCKKDDSCSFVAKNWTELLKHWRACH         CGKHFASPSRLKRHEKVHAGYDCKKDDSCSFVAKNWTELLKHWRATH         CGKHFASPSRLKRHKAKHEGYDCQK-GCSFVAKTWTELLKHWRATH         CGKHFASPSKLKRHAKAHEGYVCQKGCSFVAKTWTELLKHVRETH         7         8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       9         7       9         7       9	KEFICF LPFLCT LPFLCT LPFLCT SEPLFRCT LFPLFRCT LPLFRCT LPLFRCT LQ-HIDCE LQ-HIDCE LR-HVCC LR-HVCC LRFFVCC SKPFSCS LRFFVCC SKPFSCS LRFFVCC SSRFFVCE SSRFFVCE SSRFFVCE SSRFFVCE SSLMRKRI KRKQEPC KRKQC SLMRKRI KRKQCC LESLARKEN KRKQCC LESLARKEN KRKQCC LESLARKEN KRKQCC LESLARKEN KRKQCC LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARKEN LESLARK	PEERGE PHEEGE HEEGE CUCCQQ VCCQQ VCCQQ VCCQ VCCQ VCCQ VCCQ
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana	162 133 129 1904 165 2210 144 141 202 226 225 2810 1998 257 28 2587 2587 341 3462 2587 34622 2682	5       6         CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHTESCH       70         CGKVFKYASKLRKHEDSHVKLDSVEAFCS-EPGCMKKFTNEECLKSHTSCH       70         CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKKFTNEECLKSHTSCH       70         CGKTFKYASKLQKHEESHVKLDYSEVICC-EPGCMKAFTNLECLKAHNKSCH       70         CDKRFSLPSRLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH       70         CGKHFASPSRLKRHEKVHAGYSCET-EGCSFVAKNWTELKHWREAH       70         CGKHFASPSRLKRHEKVHAGYDCKKDDSCSFVGKTWTELLKHWREAH       70         CGKHFASPSRLKRHEKVHAGYUCQK-GCSFVAKTWTELLKHVRETH       70         7       8         7       7         8       7         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       8         7       9         7       8         7	KKEFICF LPFLCT QLPYECF SEPLFRCT VEPLFKCT Q-HIDCE Q-HIDCE Q-HIDCE IVRVQCI VCU VCU KEDITCN KEEILCE SKPFSCS QRPFVCE SSRPFVCE SSRPFVCE SSRPFVCE SSRPFVCE SSLMRKRI KRKQEPI KKRQGQ KRKQGQ SKREKNA	PEEEEG PHEEGG CUCCCQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQQU CUCCQU CUCCQQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUCCQU CUC
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio	162 133 129 1904 165 2210 1402 226 22810 1998 2281 2409 1987 286 23401 257 258 340 258 3401 258 3401 34620 312 340 3120 3120 3120 3120 3120 3120 3120 312	5       6         CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH       75         CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKHFTNEKCLKEHIESCH       75         CGKRFSQRGKLKRHESHVKLDSVEAFCS-EPGCMKAFTNLECLKAHNKSCH       76         CGRRFSQRGKLKRHEKVHKGYSCET-EGCSFVAKNWTEMTNH-KKVH       78         CGKHFASPSRLKRHEKVHKGYECKDDSCSFVGKWTELYLKHVAECH       78         CGKHFASPSRLKRHEKVHEGYECKDDSCSFVGKTWTELLKHWREAH       78         CGKHFASPSRLKRHEKVHEGYECKDDSCSFVGKTWTELLKHVRETH       78         TK-QLKKNIKRHLRTHEGPISER       78         CKFRHKNIKRHLRTHEGPISER       78         CKFRHKKNIKRHLRTHEGPISER       78         CKFRHKUPLQHUYSEERLVYUCQK-GCSSTFSKASNLQKHMKAVHDI       78         TK-QLKKNIKRHLRTHEGPISER       79         CKFRHKUPLQHUYSEERLVYUCQKSISSTSKASNLQKHMKAVHDI         TK-QLKKNIKRHLRTHESSYNCHLVCKLSFSKKSNLDKHVKAVHEG         KFRHKUPLKQHMKTHAPERDVYLCPNGCCRSYTTAFNLQSHILSFHEG         GKKFRHKHVDKHKUPRHESSYNVYTPGDFEEVDEGFLSRPR-GGRKREPTLFE         SCGMKFAFKHVRDKHENSGYHVYTCGDFVETDEDFTSRPR-GGLKRKQTAH         CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGEPRPRQAGGRKKKÄIPVE         CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGEPRPRKSLASCLTGYIPF         CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGEPRPRKSLASCLTGYIPF         CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGEPRPKRSLASCLTGYIPF         CGKSFSYKHVRDNHEKSSAHVYVQANFEEIDGEPRPKRSLASCLTGYIPF	KKEFICF CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPLFKCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLPFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLFFLCT CLF	PEEGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	162 133 1299 214 165 2280 144 12026 2281 2499 198 257 281 2598 2581 2409 19872 2842 3019 2583 340 2598 314 3402 314 3402 314	5         6         7         6         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7 <td< td=""><td>KKEFICF OLPFLCT OLPFLCT OLPFLCT SEPLFRCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT</td><td>PEEGG GCGCQU PHEEGG GCCCQCQU VVCQQR BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSG BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSG BFFSGGC BFFSG BFFSGGC BFFSGGC BFFSGGC BFFSG BFFSG BFFSGGC BFFSG BFFSGC BFFSGC BFFSG BFFSGC BFFSGC BFFSG BFFSGC BFFSGC BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFS BFFS</td></td<>	KKEFICF OLPFLCT OLPFLCT OLPFLCT SEPLFRCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT VEFT	PEEGG GCGCQU PHEEGG GCCCQCQU VVCQQR BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGGC BFFSGGGC BFFSGGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSG BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSGGC BFFSG BFFSGGC BFFSG BFFSGGC BFFSGGC BFFSGGC BFFSG BFFSG BFFSGGC BFFSG BFFSGC BFFSGC BFFSG BFFSGC BFFSGC BFFSG BFFSGC BFFSGC BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFSG BFFS BFFS
A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens L.esculentum A.thaliana O.sativa D.rerio X.laevis M.musculus H.sapiens	$\begin{array}{c} 1 \ 6 \ 2 \\ 1 \ 3 \ 3 \\ 1 \ 2 \ 9 \\ 1 \ 2 \\ 1 \ 6 \ 5 \\ 1 \ 2 \ 2 \\ 1 \ 4 \\ 1 \ 4 \\ 1 \ 2 \\ 2 \ 2 \\ 2 \ 2 \\ 2 \ 4 \\ 1 \ 9 \\ 2 \\ 2 \\ 2 \\ 2 \\ 1 \ 9 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ 1$	5       6         CGKVFKYASKLRKHEDSHVKLETVEAMCL-EPGCMKHFTNEKCLKEHIESCH       7         CGKAFKYPSQLQKHQDSHVKLDSVEAFCS-EPGCMKAFTNLECLKAHNKSCH       7         CGKTFKYASKLQKHESHVKLDSVEAFCS-EPGCMKAFTNLECLKAHNKSCH       7         CGRTFSVASKLQKHESHVKLDYSEVTCC-EPGCMKAFTNLECLKAHNKSCH       7         CGKTFSLPSRLKRHEKVHAGYSCET-EGCSFVGKTWTLYLKHVAECH       7         CGKHFASPSKLKRHEKVHAGYSCET-EGCSFVGKTWTLELKHMREAH       7         CGKHFASPSKLKRHGKVHEGYLCQK-GCSFVGKTWTLELKHVRETH       7         7       8         7       7         7       7         7       7         7       8         7       7         7       8         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7         7       7 <td>KKEFIC KKEFIC LPFLCT SEPLFRCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPFVCE VEPFVCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE V</td> <td>PEEEGG GCCCQN PHEEGG GCCCCQN VCCQQN VCCQQN VCCQQN PEESS PEHAGG PEESS PEAG PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEES</td>	KKEFIC KKEFIC LPFLCT SEPLFRCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPLFKCT VEPFVCE VEPFVCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFCE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE VEFFFE V	PEEEGG GCCCQN PHEEGG GCCCCQN VCCQQN VCCQQN VCCQQN PEESS PEHAGG PEESS PEAG PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEESS PEES

Abbildung 4.6: Alignment von TFIIIA-Sequenzen pflanzlichen und tierischen Ursprungs. Das Alignment wurde mit dem MAFFT-Algorithmus (Katoh *et al.*, 2002) erstellt und mit dem Programm BioEdit manuell überarbeitet. Zu über 50 % identische Aminosäuren sind dunkelgrau hinterlegt, ähnliche Aminosäuren sind hellgrau hinterlegt. Die Zink-Finger-Bindemotive sind umrandet und von 1–9 durchnummeriert.



Abbildung 4.7: Alignment von Sequenzen des ribosomalen Proteins L5 aus pflanzlichen und tierischen Organismen. Das Alignment wurde mit dem MAFFT-Algorithmus (Katoh *et al.*, 2002) erstellt. Zu über 50 % identische Aminosäuren sind dunkelgrau, ähnliche Aminosäuren sind hellgrau hinterlegt.

Mit Kenntnis des 3'-Bereiches der *Le*TFIIIA-mRNA war es nun möglich, das vollständige Leseraster und somit die komplette Proteinsequenz inklusive der neun Zink-Finger zu postulieren. Experimentelle Evidenz lieferte letztlich eine PCR-Amplifikation der TFIIIA-cDNA mit genspezifischen Primern, die den kompletten ORF flankieren. Das PCR-Produkt der erwarteten Länge wurde in den pGEM-T-Vektor ligiert und sequenziert. Die *Le*TFIIIA-cDNA-Sequenz stimmte abgesehen von einem einzelnen Nukleotidaustausch vollständig mit der auf Basis der Datenbankeinträge geforderten Sequenz überein. Die erstmalig für Tomate bestimmte TFIIIA-Sequenz wurde unter der Zugriffsnr. DQ882179 in der Sequenzdatenbank des NCBI abgelegt.

Überraschenderweise konnte in der PCR-Reaktion mit TFIIIA-spezifischen Primern ein zweites Produkt detektiert werden, was auf eine alternativ gespleißte mRNA-Variante
schließen lässt. Dieser eher zufällige Befund wurde in dieser Arbeit weiter verfolgt und wird in einem eigenen Kapitel (siehe 4.3) behandelt.

Ein Sequenzvergleich von LeTFIIIA mit einigen bekannten TFIIIA-Sequenzen zeigt Abb. 4.6. Die Ähnlichkeit zwischen den Sequenzen manifestiert sich vornehmlich in den Zink-koordinierenden Aminosäuren C und H sowie den hydrophoben Resten innerhalb der Zink-Finger-Domänen (F, I, L). Diese sind für die korrekte Faltung des Proteins essentiell. Auffällig ist eine *spacer*-Region in den pflanzlichen Sequenzen zwischen den Zink-Fingern 1 und 2 von ca. 20 zusätzlichen Aminosäuren. Der 66 Aminosäuren umfassende *spacer* zwischen Zink-Finger 5 und 6 ist dagegen auf die Arabidopsis-Sequenz beschränkt. Verglichen mit dem tierischen System fehlt dem LeTFIIIA ein Kernlokalisierungssignal (NLS), ein Kernexportsignal (NES) sowie ein Transkriptionsaktivierungssignal (TAS), was gleichwohl auch für AtTFIIIA berichtet wurde (Mathieu *et al.*, 2003).

Die Identifikation von *LeL5* gestaltete sich unterdessen leichter. Die L5-Sequenz erweist sich als konserviert, selbst unter sehr entfernt verwandten Arten (siehe Abb. 4.7). So zeigen die L5-Sequenzen aus entfernt verwandten Arten wie *Arabidopsis* und Mensch eine über 50 %ige Übereinstimmung, lediglich der C-terminale Bereich des Proteins divergiert stark. Die gute Konservierung führte dazu, dass sich die L5-Sequenz aus *Lycopersicon esculentum* durch automatische Annotationsverfahren erschloss. In der EST-Datenbank finden sich eine Reihe von Sequenzen aus Tomate, die als putatives ribosomales Protein L5 bezeichnet sind. Ein vollständiger ORF beinhaltet die Sequenz mit der Zugriffsnr. BT014015. Diese Sequenz sollte schließlich in dieser Arbeit für die rekombinante Expression des Proteins *LeL5* kloniert werden.

#### 4.2.3 Expression rekombinanter Proteine in E. coli

#### 4.2.4 Beschreibung des Expressionssystems

Erste Versuche einer Überexpression der rekombinanten Proteine wurden mit dem pTYB-Vektorsystem der Firma NEB durchgeführt. Die Expression wird hierbei von einem T7-Promotor unter Kontrolle eines Lac-Repressors vermittelt. An das klonierte Zielprotein wird bei Translation ein 55 kDa-Protein fusioniert, welches eine Inteindomäne und eine Chitinbindedomäne umfasst (Intein*tag*). Dies erlaubt die einfache Aufreinigung über eine Affinitätsmatrix. Durch reduzierende Bedingungen wird ein Selbstspleißen der Inteindomäne induziert. Somit ist es möglich, in nur einem chromatographischen Schritt das Zielprotein in exakter Länge von dem Intein*tag* zu trennen, ohne Proteasen verwenden zu müssen. Je nach Vektor ist der Intein*tag* C-terminal (pTYB1 und pTYB2) oder Nterminal positioniert (pTYB11 und pTYB12).

#### 4.2.5 Klonierung von StTBP in pTYB

Die Sequenz von StTBP wurde von Kartoffel-cDNA per PCR amplifiziert, in pGEM-T subkloniert und durch Sequenzierung verifiziert. Die Primer wurden dahingehend entworfen, dass sie an ihrem 5'-Ende geeignete Restriktionsschnittstellen trugen und somit eine gerichtete Klonierung möglich war. Die StTBP-Subklone und die Expressionsvektoren pTYB1 und pTYB11 wurden mit NheI/SapI oder NdeI/SapI behandelt. Als vermeintliches Hindernis stellte sich die Tatsache dar, dass die NdeI-Erkennungssequenz auch innerhalb der StTBP-Sequenz enthalten ist. Dieses Problem konnte erfolgreich umgangen werden, indem sowohl die im Restriktionsansatz eingesetzte NdeI-Konzentration als auch die Inkubationsdauer limitiert wurde. Durch diesen unvollständigen Restriktionsverdau entstand somit u.a. korrekt prozessierte insert-DNA, die nach Isolation in den pTYB11-Vektor ligiert werden konnte.

#### 4.2.6 Klonierung von AtTFIIIA und AtL5 in pTYB

Zu Beginn der vorliegenden Arbeit war die Sequenz von TFIIIA und L5 aus der Wirtspflanze Tomate noch unbekannt. Für erste Expressionsversuche wurden diese Proteine daher aus Arabidopsis thaliana kloniert. Die Sequenzen von AtTFIIIA und AtL5 wurden von Arabidopsis-cDNA gewonnen und zunächst in pGEM-T subkloniert. Dabei wurden mit den TFIIIA-spezifischen Primern ebenfalls wie von Tomaten-cDNA zwei verschiedene Produkte gebildet (vgl. Abschnitt 4.2.2), was auf ein alternatives Spleißen der TFIIIAmRNA schließen lässt. Dieses beachtenswerte Phänomen wird in einem späteren Kapitel behandelt (siehe Kap. 4.3).

Für die N-terminale Fusion des Affinitäts*tags* sollten beide Sequenzen in pTYB11 kloniert werden. Um nach Selbstspaltung des *tags* ein Protein von nativer, exakter Sequenz zu entlassen war eine Klonierung über die *Sap*I-Schnittstelle des Vektors erforderlich. Da diese jedoch sowohl in der AtTFIIIA- als auch in der AtL5-Sequenz enthalten ist, mussten

Protein	ohne Fusions $tag$	mit 55 kDa-Intein $tag$	mit 26 kDa-GST $tag$
<i>At</i> TFIIIA	46,6	112,6	73,1
AtL5	34,4	90,4	61,3
StTBP	22,3	78,3	_
LeTFIIIA	42,9	_	69,1
LeL5	34,5	_	60,9

**Tabelle 4.1: Übersicht über die Molekulargewichte der rekombinanten Proteine.** Alle Angaben in kDa.

nach dem SapI-Verdau des jeweiligen Subklons zunächst die *insert*-Fragmente isoliert und reassoziiert werden. Dies erwies sich als unproblematisch, da das Enzym SapI außerhalb der Erkennungssequenz schneidet. Somit konnten die zweifach SapI-geschnittenen *inserts* in definierter Orientierung wieder zusammengefügt werden.

Zur C-terminalen Fusion des Intein*tags* wurden die AtTFIIIA- und AtL5-Sequenzen via *Sal*I und *Sma*I in pTYB2 ligiert. Eine Übersicht über die Molekulargewichte der klonierten Konstrukte findet sich in Tab. 4.1.

#### 4.2.7 Überexpression und Aufreinigung mit dem pTYB-System

Die Expressionsvektoren wurden in den *E. coli*-Stamm BL21 CodonPlus(DE3)RIL (Stratagene) transformiert. Dieser Expressionsstamm eignet sich besonders für die Expression eukaryotischer Proteine, da er mit zusätzlichen tRNAs für in *E. coli* seltene Codons supplementiert ist.

Die pTYB-Expressionsklone wurden wie im Methodenkapitel (siehe 3.1.3) beschrieben angezogen. Die rekombinanten Proteine wurden gemäß Abschnitt 3.1.4 aufgereinigt. Im Coomassie-gefärbten Gel ist bei allen drei klonierten Proteinen (AtTFIIIA, AtL5 und StTBP) mit C-terminalen Fusionstag kein Signal infolge einer Überexpression zu detektieren (Daten nicht gezeigt). Eine Variation der IPTG-Konzentration und der Expressionstemperatur konnte keine Verbesserung erzielen. Dennoch konnte Protein an die chitinöse Matrix gebunden und anschließend eluiert werden (siehe Abb. 4.8). Im Falle von AtTFIIIA erkennt man in der Spur BV das 113 kDa-Vorläuferprotein, welches an die Affinitätsmatrix gebunden ist. Durch Zugabe von DTT lässt sich das 47 kDa-TFIIIA abspalten (siehe Spuren Eluat1 und 2), während der 55 kDa-Intein*tag* an den Chitinbeads verbleibt (Spur BN).



Abbildung 4.8: Aufreinigung der Proteine AtTFIIIA, AtL5 und StTBP mit C-terminalem Fusions tag über eine Affinitätschromatographie. SDS-PAGE, 10 % PAA (30:0,8). BV: Chithinbeads vor Inteinspaltung, BN: Chitinbeads nach Inteinspaltung mittels DTT. E1: Eluat 1, gewonnen nach 24-stündiger Inkubation im Elutionspuffer, E2: Eluat 2, nach 48 Stunden.

Die Qualität der Eluate indes war nicht zufriedenstellend. Aufgrund von Verunreinigungen mit endogenem *E. coli*-Protein zeigt das Gel neben dem erwarteten Produkt ein sehr heterogenes Bandenmuster. Im Falle des *St*TBP ist bei erwarteter Länge von 22 kDa gar kein Signal vorhanden. Ein ausgiebigeres und stringenteres Waschen der Matrix mit erhöhter NaCl-Konzentration von 1,5 M bis 2 M hatte keinen Einfluss auf die Reinheit des Eluats. Darüber hinaus ließen sich die Eluate von *At*TFIIIA und *At*L5 nicht durch Ultrafiltration aufkonzentrieren, da die Proteine nach autokatalytische Abspaltung vom Intein*tag* offensichlich zur Aggregation tendierten. Es gelang nicht mit erhöhter NaCl-Konzentration von 1 M im Elutionspuffer und 1 % Triton X die Proteine stabil in Lösung zu halten.

Das StTBP mit N-terminalem Fusionstag zeigte als einziges Protein ein starkes Signal nach der Induktion (Daten nicht gezeigt). Gleichwohl entsprach das apparente Molekulargewicht von 55 kDa nicht dem erwarteten Wert von 78,3 kDa. Da das Produkt mit großer Effizienz an die Chitin-Matrix assozierte, nicht aber unter reduzierenden Bedingungen zu spalten war, lag der Verdacht nahe, dass hier lediglich der N-terminale 55 kDa-Inteintagohne das Zielprotein StTBP überexprimiert wurde. Letztendliche Gewissheit über die Identität des Signals ergab eine Maldi-ES-Analyse des trypsinbehandelten 55 kDa-Proteins: Es konnten lediglich die Massen der Intein*tag*-Fragmente, nicht aber Massen von TBP-Fragmenten gemessen werden. Daher ist anzunehmen, dass *St*TBP toxisch für *E. coli* ist. Aufgrund der hohen Assoziationskonstante würde selbst ein nur basal exprimiertes TBP Tata-ähnliche DNA-Abschnitte okkupieren, wodurch die intrinsische Genregulation empfindlich gestört würde. Somit kann offenbar die korrekte Sequenz des *St*TBP-Expressionsvektors dem selektiven Druck nicht standhalten. Diese Beobachtung steht allerdings im Gegensatz zu Arbeiten, in denen erfolgreich TBP – auch aus *Solanum tuberosum* – in *E. coli* exprimiert werden konnte (Holdsworth *et al.*, 1992).

Aufgrund der unbefriedigenden Ergebnisse dieser ersten Expressionsexperimente mit den pTYB-Vektoren wurde auf ein anderes Expressionssystem ausgewichen. Mit den pGEX Expressionsvektoren (Amersham) werden Proteine mit Glutathion S-Transferase (GST) fusioniert. Dies erlaubt die Isolation des rekombinanten Proteins über Glutathion-Sepharose mit über 90 %iger Reinheit. Es handelt sich hierbei um ein weit verbreitetes, etabliertes Expressionssystem, welches sowohl gute Überexpression als auch gute Löslichkeit der rekombinanten Proteine verspricht.

#### 4.2.8 Klonierung von TFIIIA, L5 und TBP in pGEX

Die Sequenzen von AtTFIIIA, LeTFIIIA, StTBP und LeL5 wurden von cDNA bzw. Subklonen mittels PCR amplifiziert und über BamHI und SalI in den Expressionsvektor pGEX-4T-3 kloniert. LeL5 wurde via EcoRI und SalI ligiert. In dem verwendeten Vektorsystem vom Typ pGEX-4T sind GST- und Zielproteinsequenz durch eine Thrombin-Schnittstelle verbunden. Dies würde die Abspaltung des Affinitätstags nach der chromatographischen Aufreinigung erlauben, worauf aber in dieser Arbeit aus folgenden Gründen verzichtet wurde: 1) Reinigungsversuche mit dem Intein-System deuteten bereits daraufhin, dass die exprimierten Proteine nach Abspaltung des tags zur Aggregation tendieren (siehe Abschnitt 4.2.7). 2) Mathieu *et al.* (2003) konnten GST-AtTFIIIA und GST-AtL5 erfolgreich überexprimieren und zeigen, dass der GST-Fusionstag die biologische Aktivität nicht negativ beeinflusst.

Die gewonnenen Klone wurden einer Restriktionskontrolle unterzogen und anschließend sequenziert. Mit Ausnahme von pGXStTBP hatten alle Konstrukte die gewünschte Sequenz. Dies ist verwunderlich, da auch die StTBP-insert-DNA durch PCR mit einer proof reading-Polymerase amplifiziert wurde. Als template diente darüber hinaus ein Subklon (pGEMStTBP) mit definierter, 100% ig korrekter Sequenz. Bei den Sequenzfehlern handelte es sich nicht um sogenannte stille Mutationen, d.h. die Nukleotidaustausche hatten allesamt Aminosäuremutationen zur Folge, die sich zudem in den stark konservierten Bereichen des TBP befanden. Dies stellte die Funktionalität der mutierten Proteinsequenzen in Frage. Es wurde daher nun versucht, die *insert*-DNA von dem sequenzierten Subklon in den Expressionsvektor zu klonieren. Durch diesen PCR-freien Schritt können Mutationen in der *insert*-DNA ausgeschlossen werden. Eine Sequenzanalyse von sechs Klonen bestätigte die Fehlerfreiheit der klonierten TBP-DNA. Allerdings war bei allen geprüften Klonen die insert-DNA im Vektor fehlerhaft positioniert. Durch geschwächte Spezifität (sogenannte star-Aktivität) des Restriktionsenzyms BamHI an einer BamHIähnlichen (GGATCT) Erkennungssequenz upstream der MCS (multi cloning site) wurden überhängende Enden generiert, die eine falsche Ligation des *inserts* ermöglichten. Die auf diese Weise entstandenen Klone waren unbrauchbar, da durch Deletion das offene Leseraster zerstört wurde. Bemerkenswerterweise wurden mit derselben zweifach restringierten Vektor-DNA die übrigen *inserts* (TFIIIA und L5) erfolgreich kloniert. Daher liegt die Annahme eines TBP-spezifischen Problems nahe. Die Ergebnisse der StTBP-Klonierung lassen nur einen Schluss zu: StTBP ist toxisch für E. coli. Um ihr Überleben zu sichern müssen Zellen für eine sehr kleine Population an fehlerhaft ligierter oder mutierter TBPinsert-DNA selektieren. Klone, die das korrekte Konstrukt tragen, sind lethal. Bemerkenswerterweise wurde humanes TBP, welches TBP aus Kartoffel zu 81% identisch ist, als GST-Fusionprotein erfolgreich in *E. coli* exprimiert (Sommer *et al.*, 1994).

Aufgrund der negativen Ergebnisse bei der Klonierung und Expression von StTBP mit zwei verschiedenen Vektorsystemen fasste der Autor schließlich den Entschluss, die Arbeit an diesem Protein einzustellen. Ein Umstieg auf ein sekretorisches, eukaryotisches Expressionssystem kam aus Zeitgründen nicht in Frage.

#### 4.2.9 Überexpression und Aufreinigung mit dem pGEX-System

Die Expression der Vektoren pGXLeL5 und pGXLeTFIIIA erfolgte in dem BL21-Derivat Rosetta2<sup>TM</sup> (Novagen). Dieser *E. coli*-Stamm trägt zusätzliche tRNAs für in *E. coli* seltene Arginin-, Leucin- und Isoleucin-Codons. In Vorversuchen, in denen verschiedene

Transformanten auf Uberexpression getestet wurden, zeigte sich, dass nahezu alle untersuchten Klone das jeweilige rekombinante Protein produzierten. Exemplarisch zeigt Abb. 4.9 starke Signale von GST-*Le*L5 in Zellaufschlüssen nach IPTG-Induktion.

Zur Analyse der Löslichkeit der überexprimierten Proteine wurden die Zellen in PBS mit 100  $\mu$ g/ $\mu$ l Lysozym sonifiziert und unlösliche Bestandteile abzentrifugiert. Aliquots der löslichen und unlöslichen Fraktion wurde auf einer SDS-Page analysiert. Hierbei wurde evident, dass alle Proteine unlöslich exprimiert wurden; sie lagen aggregiert in *inclusion bodies* vor (siehe Abb. 4.10, Spur PS ohne Sarkosyl). Ein unvollständiger Aufschluss der Zellen konnte durch lichtmikroskopische Kontrolle ausgeschlossen werden. Daher wurde versucht, die Expressionsbedingungen so zu optimieren, dass der Anteil an löslichem, rekombinantem Protein nennenswert zunimmt. Folgende Parameter wurden hierzu variiert:

- IPTG-Induction bei höherer Zelldichte (1,0 OD statt 0,4 OD)
- Reduktion der Expressionstemperatur (20°C, 15°C, 4°C)
- Verkürzung der Expressionsdauer (30 min, 60 min)
- Zugabe von 2% Ethanol zum Wachstumsmedium

Mit keinem der genannten Optimierungsversuche gelang es aber die Löslichkeit der rekombinanten Proteine signifikant zu erhöhen. Lediglich mittels Expression von GST-LeL5 bei Raumtemperatur konnte ein schwaches Signal in der löslichen Fraktion verzeichnet werden (Daten nicht gezeigt). Die Konzentration war indes zu gering für eine affinitätschromatographische Aufreinigung in größerem Maßstab. Schließlich wurde mit BLR ein anderer Expressionsstamm getestet, der nicht auf die Expression eukaryotischer Proteine optimiert ist. Eine Reduzierung des Expressionslevels könnte der Bildung von inclusion bodies entgegenwirken. Außerdem konnten mit diesem *E. coli*-Stamm Mathieu *et al.* (2003) erfolgreich lösliches rekombinantes GST-AtTFIIIA und GST-AtL5 produzieren. Obwohl somit Expressions- und Extraktionsbedingungen exakt den veröffentlichten Methoden entsprachen, gelang es in der vorliegenden Arbeit nicht, die GST-Fusionsproteine in löslicher Form aufzureinigen.

Eine vollständige Denaturierung mit stark chaotropen Reagenzien wie Harnstoff oder Guanidiniumhydrochlorid wurde zunächst nicht erwogen, da sich eine Renaturierung einer entfalteten Polypeptidkette zu einem biologisch aktiven Protein als äußerst diffizil



Abbildung 4.9: Rekombinante Überexpression der GST-Fusionsproteine in *E. coli* am Beispiel von GST-*LeL5.* SDS-PAGE, 10 % PAA (30:0,8). Ein Aliquot von 0,1 OD *E. coli*-Zellen (Rosetta2<sup>TM</sup>) wurde vor Induktion (v) bzw. 2h und 4h nach Induktion mit IPTG entnommen. M: Mark12<sup>TM</sup> MW-Standard mit Proteinen der Größen 116, 97, 66, 55, 36, 31, 22 und 14 kDa. Alle drei untersuchten *LeL5*-Klone zeigen sehr gute Überexpression des rekombinanten Proteins GST-*LeL5*.

gestalten kann. Zwar existieren in der Literatur einige Renaturierungsprotokolle für rekombinantes TFIIIA aus *Xenopus laevis* (Del Río & Setzer, 1991). Da die Sequenzen von pflanzlichem und tierischem TFIIIA aber nur zu 20% übereinstimmen, schien es wenig erfolgsversprechend, dieses Protokoll für die vorliegende Arbeit zu adaptieren.

In einem nächsten Optimierungsschritt in Richtung verbesserter Löslichkeit wurden daher die Extraktionsbedingungen modifiziert. Zunächst wurden dem Extraktionspuffer milde Detergenzien zugefügt. Durch Zugabe von 1 % Triton X-100 konnte die Löslichkeit für beide GST-Fusionsproteine nicht signifikant verbessert werden.

Durch Behandlung mit N-Lauroylsarkosin (Sarkosyl), ein ionisches Detergenz, welches in niedriger Konzentration nicht denaturierend wirkt, lassen sich unlöslich exprimierte GST-Fusionsproteine solubilisieren und über eine Affinitätschromatographie aufreini-



Abbildung 4.10: Einfluss von Sarkosyl auf die Reinigung von GST-LeL5. SDS-PAGE, 10 % PAA (30:0,8). CL: cleared lysate, PS: post sonicate, FT: flow through, W: wash, E: Eluat, M: Mark12<sup>TM</sup> MW-Standard mit Proteinen der Größen 116, 97, 66, 55, 36 und 31 kDa. Ohne Sarkosyl im Extraktionspuffer gelingt es nicht, GST-LeL5 in Lösung zu bringen (starkes Signal in PS). Durch Zugabe von 1,5 % Sarkosyl wird GST-LeL5 solubilisiert und kann chromatographisch sauber aufgereinigt werden (siehe Spur E).

gen (Frangioni & Neel, 1993). Auch im Falle der hier exprimierten GST-*Le*TFIIIA und GST-*Le*L5 gelang es endlich, die Produkte durch Zugabe von 1,5 % Sarkosyl zu lösen und zu reinigen (siehe Abb. 4.10).

Die GST-Fusionsproteine wurden über Glutathion-Sepharose<sup>TM</sup>4B (Amersham) aufgereinigt. Dazu wurde ein spezielles *batch-column-batch*-Verfahren entwickelt. Um der langsamen Assoziationskinetik der GST-Fusionsproteine an die Glutathionmatrix Rechnung zu tragen, wurde die Bindung in einem *batch*-Verfahren durchgeführt. Dies ermöglicht eine lange Inkubationszeit. Anschließend war ein ausgiebiges Waschen erforderlich, um endogene *E. coli*-Proteine, insbesondere RNasen, restlos zu entfernen. Es erwies sich gleichsam als effektiv und praktikabel, die Sepharose-Partikel in eine Chromatographiesäule zu übertragen, und diese per Gravitationsfluss extensiv zu waschen. Die Elution erfolgte wiederum in einem Eppendorf-Gefäß, da die Dissoziation der GST-Fusionsproteine in Anwesenheit von reduziertem Glutathion sehr langsam verläuft (Herstellerangabe, Amersham). Eine detailierte Dokumentation der optimierten Aufreinigung der GST-Fusionsproteine findet sich im Methodenteil unter Abschnitt 3.1.6.



Abbildung 4.11: Konzentrationsbestimmung des gereinigten GST-*Le*TFIIIA nach Dialyse. Spuren 1–5) BSA-Konzentrationreihe (2, 1, 0,5 0,25, 0,125  $\mu$ g), Spuren 6–10) GST-*Le*TFIIIA-Eluat, dialysiert (10, 5, 2,5, 1,25, 0,63  $\mu$ l), Spur M) Mark12<sup>TM</sup> MW-Standard mit Proteinen der Größen 116, 97, 66, 55, 36, 31 und 22 kDa. Durch Bandenvergleich mit dem BSA-Standard wurde eine TFIIIA-Konzentration von 100 ng/ $\mu$ l bestimmt, was bei einem Molekulargewicht von 69 kDa einer molaren Konzentration von 1,45  $\mu$ M entspricht. Die Reinheit des dialysierten Eluats beträgt über 90%.

Die Reinheit der Eluate lag deutlich über 90 % (siehe Abb. 4.11). Die Proteine wurden gegen  $1 \times$  Bindungspuffer dialysiert. Eine TFIIIA-Konzentration von  $1,45 \,\mu$ M wurde in einem Coomassie-Gel durch Vergleich mit einem definierten Standard (BSA) ermittelt.

#### 4.2.10 Zusammenfassung zur rekombinaten Proteinexpression

Mit dem pTYB-Vektorsystem gelang es in der vorliegenden Arbeit nicht, die Proteine AtTFIIIA, AtL5 und StTBP zu exprimieren und aufzureinigen. Nach Abspaltung des matrixgebundenen Fusions*tags* wurde TFIIIA unlöslich. L5 konnte nur sehr unsauber und

in sehr niedriger Ausbeute chromatographisch gereinigt werden. StTBP ließ sich aufgrund seiner Toxizität sowohl in pTYB als auch in pGEX nicht exprimieren.

Datenbankanalytisch und experimentell gelang es, die TFIIIA-Sequenz aus der PSTVd-Wirtspflanze Lycopersicon esculentum zu identifizieren. Nach TFIIIA aus Arabidopsis thaliana (Mathieu et al., 2003) ist dies erst der zweite experimentell verifizierte Datenbankeintrag einer pflanzlichen TFIIIA-Sequenz.

Mit dem pGEX-Vektorsystem konnten die Proteine TFIIIA und L5 aus Lycopersicon esculentum erfolgreich, wenn auch unlöslich überexprimiert werden. Eine Teillyse gelang mit Hilfe des ionischen Detergenz Sarkosyl. Somit konnten LeTFIIIA und LeL5 als GST-Fusionsproteine mittels modifizierter Affinitätschromatographie sauber präpariert werden.

#### 4.2.11 Nukleinsäurebindung der rekombinanten Proteine

#### Methodik

Um die erfolgreich exprimierten, rekombinanten Proteine *Le*TFIIIA und *Le*L5 hinsichtlich ihrer Nukleinsäurebindeaktivität zu anaylsieren wurde die Methode der Gelverzögerung angewandt (EMSA, von engl.: *electrophoretic mobility shift assay*). Das Prinzip dieser Anwendung beruht darauf, dass Nukleinsäure, die im Komplex mit einem Protein gebunden ist, langsamer in der nativen Gelmatrix als freie Nukleinsäure migriert. Aus dem Verhältnis von freier zu gebundender Nukleinsäure lässt sich nach dem Massenwirkungsgesetz die Dissoziationskonstante berechnen:

$$K_d = \frac{P_f \cdot N_f}{C}$$

- mit  $K_d$  : Dissoziationskonstante,
  - $P_f$ : freie Proteinkonzentration,
  - $N_f$ : freie Nukleinsäurekonzentration,
  - C : Komplexkonzentration.

Im Experiment wird Nukleinsäure mit steigenden Proteinkonzentrationen titriert. Ermittelt wird dann ein Konzentrationsverhältnis, bei dem die Hälfte der eingesetzten Nukleinsäure  $N_0$  komplexiert vorliegt  $(N_0/2 = N_f = C)$ . Dann gilt:

$$K_d = P_f;$$

Wenn ferner gewährleistet ist, dass die eingesetzte Nukleinsäurekonzentration  $N_0$  viel kleiner ist als die eingesetzte Proteinkonzentration  $P_0$ , dann ist:

$$K_d = P_0;$$

Die Proteinkonzentration bei Halbsättigung entspricht also der Dissoziationskonstanten der Reaktion im Gleichgewicht. Experimentell kann der Punkt der Halbsättigung durch Bandenvergleich mit einer Referenzspur, in der genau die Hälfte an Nukleinsäure ohne Protein aufgetragen wird, identifiziert werden.

#### **Eingesetzte Proteine**

In den im Folgenden beschriebenen Gelverzögerungsexperminenten wurde die Nukleinsäurebindung von rekombinantem GST-*Le*TFIIIA und GST-*Le*L5 analysiert. Die Sequenz dieser Proteine stammt aus *Lycopersicon esculentum*, einer Wirtspflanze von PSTVd. Weiterhin tragen die Faktoren einen N-terminalen GST-*Tag*, der aber die Bindungseigenschaften der *Arabidopis*-Homologen nicht beeinflusst (Mathieu *et al.*, 2003). Die beiden verwendeten Proteine werden im Folgenden der Einfachheit halber als TFIIIA und L5 bezeichnet. Vor dem Einsatz im Experiment wurden die gegen Bindepuffer dialysierten Proteine 30 min bei 30.000 g abzentrifugiert, um inaktive Aggregate zu entfernen.

#### Eingesetzte DNA

Als Positivkontrolle wurde mit dem internen Promotor des 5 S-Gens die natürliche DNA-Bindestelle von TFIIIA ausgewählt (siehe Abb. 4.12A). Es wurde ein 330 bp großes Fragment aus Tomatenzellkern-DNA ampflifiziert (Basen 1-330 der DNA-Sequenz mit der Zugriffsnr. X55697), welches die Promotorelemente A-Box, IE (*intermediate element*) und C-Box enthält. Unspezifische DNA einer Länge von 330 bp wurde von dem T7-Transkriptionsplasmid pRH701 gewonnen, das mir freundlicherweise von Herrn Bernd Esters zur Verfügung gestellt wurde. Das Fragment wurde auch als Negativkontrolle in EMSAs mit dem RNA-bindenden ribosomalen Protein L5 eingesetzt.

#### Eingesetzte RNA

Alle eingesetzten RNAs wurden als *run off*-Transkripte von PCR-Produkten mit T7-Promotor synthetisiert. Die 5 S-RNA (siehe Abb. 4.12B), repräsentiert den natürlichen RNA-



#### A) 5S-DNA



Abbildung 4.12: In die Gelverzögerungsexperimente eingesetzte Nukleinsäurespezies. A) Der Transkriptionsfaktor IIIA bindet an die Promotorelemente A-Box, IE und C-Box, welche innerhalb der kodierenden Sequenz positioniert sind. Der Transkriptionsstart der 5 S-RNA-Synthese ist mit +1 gekennzeichnet. B) 5 S-RNA als Positiv- und C) 7 S-RNA als Negativkontrolle. Alle Sequenzen stammen aus Lycopersicon esculentum.

Bindepartner von TFIIIA und L5. Ihre Sequenz aus *Lycopersicon esculentum* ist unter der Zugriffsnr. X06842 annotiert. Von vordringlichem Interesse war die Bindung der Proteine an PSTVd. Hier wurden an Position 146 linearisierte Transkripte von Einheitslänge eingesetzt. Als unspezifische Bindepartner wurden weiterhin untersucht:

- 7 S-RNA aus Lycopersicon esculentum (Zugriffsnr. Z46744, siehe Abb. 4.12C)
- dsRNA, durchgängig doppelsträngige RNA, 174 bp mit PSTVd-Sequenz
- rr1-RNA, 341 nt-Fragment aus 5'ETS-Region der 18S-Vorläufer-rRNA (Schmitz, 2003).
- TBP-mRNA, 140 nt-Transkript von PstI-linearisiertem pGEMTBP-Subklon

Alle Nukleinsäuren wurden mit  ${}^{32}$ P wie in Abschnitt 3.2.8 beschrieben radioaktiv markiert. Vor dem Einsatz in das Bindungsexperiment wurden die RNAs bei > 90°C



Abbildung 4.13: EMSA zur Analyse der Bindung von TFIIIA an 5S-DNA unter verschiedenen Pufferbedingungen. Autoradiographie einer nativen PAGE, 5% PAA, 0,2 x NaTA, 0,5 mM MgSO<sub>4</sub>. In verschiedenen Bindepuffern wurden 5.000 cpm 5S-DNA mit steigenden TFIIIA-Konzentrationen versetzt (0, 54, 109, 218, 435, 870 nM). Spuren 1–6) NaOAc pH 5,2, Spuren 7–12) Tris/HCl pH 8,8, 100 mM NaCl, Spuren 13–18) Tris/HCl pH 8,0, 100 mM KCl.

vollständig denaturiert und auf Eis abgekühlt. Bei der Präparation der Nukleinsäure wurden ausschließlich EDTA-freie Reagenzien verwendet. Die Bindung von Zn<sup>2+</sup> durch EDTA ist äußerst stark ( $K_d=10^{-16}$  M; Nyborg & Peersen, 2004). Reste von EDTA im Bindepuffer würden dem Metalloprotein TFIIIA die Zn<sup>2+</sup>-Ionen entziehen, da diese im Protein deutlich schwächer gebunden werden ( $K_d=10^{-10}$  M). Die koordinative Bindung von Zn<sup>2+</sup> in den Zink-Finger-Motiven des TFIIIA ist essentiell für dessen strukturelle Integrität und somit notwendig für dessen Bindeaktivität.

#### Wahl des Bindepuffers

In einem ersten Gelverzögerungsexperiment mit TFIIIA wurde die DNA-Bindung in verschiedenen Reaktionspuffern analysiert. Der pH-Wert der Puffer lag dabei stets weit oberhalb oder unterhalb des pI-Wertes von 7,2, um ein Ausfallen des Proteins zu verhindern. Zur Unterdrückung unspezifischer, elektrostatischer Wechselwirkungen enthielten zwei Puffer 100 mM NaCl bzw. KCl. Ferner wurden reduzierende Bedingungen gewählt, um eine Oxidation der strukturrelevanten Cysteine zu unterbinden.

Da der Bindepuffer zusätzlich Glycerin und Bromphenolblau enthielt, war es möglich, die Proben nach Inkubation direkt auf das Gel aufzutragen. Eine zusätzliche Verdünnung



Abbildung 4.14: EMSA zur Bestimmung des apparenten  $K_d$ -Wertes für die Bindung von TFIIIA an 5 S-DNA. Autoradiographie einer nativen PAGE, 5% PAA, 0,2 x NaTA, 0,5 mM MgSO<sub>4</sub>. Autoradiographie. 5.000 cpm 5 S-DNA wurden mit steigenden TFIIIA-Konzentrationen versetzt (3,5, 7, 14, 28, 55 nM). Die Spuren 1 und 7 enthalten 2.500 cpm 5 S-DNA ohne Protein, Spuren 1– 6) Tris/HCl pH 8,8, 100 mM NaCl, Spuren 7–12) Tris/HCl pH 8,0, 100 mM KCl. Halbsättigung kann bei einer Proteinkonzentration von 7 nM beobachtet werden.

mit Auftragspuffer und die damit einhergehende Störung des thermodynamischen Gleichgewichts wurde somit umgangen. Binde- und Elektrophoresepuffer waren stets EDTAfrei. GST-TFIIIA konnte nicht gegen Detergenzien-freien Bindepuffer dialysiert werden. Ein Pufferwechsel während einer Ultrafiltration führte zu einem kompletten Verlust des Proteins. Auch in einer langsamen Dialyse gegen Puffer ohne Triton X-100 fiel das Protein aus. Zur Förderung der Solubilität enthielt der Bindepuffer demnach 0,1 % Triton X-100.

#### 4.2.12 Analyse der Nukleinsäurebindeeigenschaften von TFIIIA

#### Bindung von TFIIIA an 5S-DNA

Ein Retardierungsgel mit TFIIIA zeigt Abb. 4.13. Mit steigender Proteinkonzentration geht eine Abnahme der freien 5S-DNA einher. Eine diskrete Komplexbande mit geringerer Mobilität wird sichtbar. Überdies lassen sich in den Tris-haltigen Proben zusätzliche Banden detektieren, die auf eine Mehrfachbeladung der DNA mit Protein hinweisen. Ein diffuser Schmier, insbesondere in den Spuren 8–10 und 14–16, offenbart, dass Komplexe auch während der Elektrophorese irreversibel zerfielen. Dies stellt jedoch bei der Auswertung kein Problem dar, da der  $K_{\rm d}$ -Wert anhand der Abnahme der freien DNA bestimmt werden kann.

Halbsättigung konnte in diesem Experiment nicht beobachtet werden, da die niedrigste TFIIIA-Konzentration von 54 nM offensichlich noch oberhalb des  $K_d$  liegt. Indes zeigt sich, dass die Bindung in den leicht basischen Tris-Puffern effektiver ist, als im sauren Acetat-Puffer. Eine Beteiligung des GST-Fusions*tags* an der DNA-Bindung kann ausgeschlossen werden. In Bindeanalysen mit GST-AtTFIIIA konnte gezeigt werden, dass GST alleine nicht zur DNA-Bindung befähigt ist (Mathieu *et al.*, 2003). Auch in der vorliegenden Arbeit konnte keine DNA-Bindung für ein anderes GST-Fusionsprotein gemessen werden (siehe Abb. 4.25).

Zur Ermittlung des apparenten  $K_d$  wurden in einem Folgeexperiment die Proteinkonzentrationen gesenkt, um Halbbindung zu erreichen (siehe Abb. 4.14). In den Spuren 3 und 9 liegt die 5S-DNA zu 50% ungebunden vor, was aus Vergleich mit den Referenzspuren 1 und 7 ersichtlich ist. Daraus ergibt sich ein apparenter  $K_d$  von 7 nM, was auf eine sehr affine Bindung schließen lässt. Obwohl offenbar mehrere Bindeplätze auf der eingesetzten DNA existieren, ist es dennoch zulässig die Dissoziationskonstante über Halbsättigung – also unter der Annahme nur eines Bindungsplatzes – zu bestimmen. Dies wird näher in Abschnitt 4.2.12 erläutert.

In den nächsten Verzögerungsgelen mit TFIIIA wurde KCl-haltiger Trispuffer verwendet, da KCl zum einen von größerer physiologischer Relevanz ist als NaCl und zum anderen die Bindungsdaten besser mit Literaturwerten verglichen werden können. In den veröffentlichten Bindeexperimenten mit TFIIIA wird überwiegend KCl eingesetzt (Mathieu *et al.*, 2003; Schulman & Setzer, 2002).

#### Bestimmung des Anteils an bindeaktivem TFIIIA

Die für die TFIIIA/5 S-DNA-Bindung gemessenen Dissoziationskonstanten lagen in den ersten Experimenten bei 7 nM und damit um ungefähr eine Größenordnung oberhalb des Literaturwertes für AtTFIIIA (0,33 nM; Mathieu *et al.*, 2003). Eine mögliche Ursache könnte ein nennenswerter Anteil fehlgefalteten Proteins darstellen, welcher nicht zur DNA-Bindung beitragen kann. Zur Quantifizierung der bindeaktiven Protein-Konzentration ([P]<sub>aktiv</sub>) wurde wie in Carey & Smale (2002) beschrieben vorgegangen: Die TFIIIA-Konzentration wird im Bindungsansatz auf 350 nM erhöht. Sie liegt somit um einen Fak-



Abbildung 4.15: EMSA zur Bestimmung des Anteils an bindeaktivem TFIIIA. Autoradiographie einer nativen PAGE, 5% PAA,  $0,2 \times NaTA$ ,  $0,5 \text{ mM} MgSO_4$ . 5.000 cpm 5 S-DNA wurden mit 350 nM TFIIIA versetzt. Die Proben wurden mit steigender Konzentration an unmarkierter 5 S-DNA ([5S DNA]<sub>komp</sub>) titriert, Spuren 2–11) 47, 66, 93, 131, 185, 263, 372, 525, 741, 1050 nM unmarkierte 5 S-DNA als Kompetitor. Spur 1 enthält 2.500 cpm 5 S-DNA ohne Protein und Kompetitor-DNA. Halbsättigung ist in Spur 10 zu beobachten. Folglich ist das rekombinante TFIIIA zu 100% bindeaktiv (siehe Text).

tor 50 höher als der apparente  $K_d$ -Wert von 7 nM. Dies hat zur Folge, dass die markierte 5 S-DNA näherungsweise vollständig gebunden wird. Sodann wird unmarkierte 5 S-DNA ([5S DNA]<sub>komp</sub>) in steigender Konzentration hinzugefügt, die ebenfalls komplexiert wird, solange ihre Konzentration unterhalb [P]<sub>aktiv</sub> liegt. Übersteigt [5S DNA]<sub>komp</sub> nun [P]<sub>aktiv</sub>, so konkurrieren markierte und unmarkierte 5 S-DNA um die Bindung an das Protein, wodurch markierte DNA freigesetzt wird. Bei dem Konzentrationsverhältnis, an dem die Hälfte der markierten DNA aus der Bindung verdrängt wird, ist [5S DNA]<sub>komp</sub> genau doppelt so hoch wie [P]<sub>aktiv</sub>. Mit Kenntnis von [P]<sub>aktiv</sub> kann schließlich das apparente  $K_d$ korrigiert werden:

$$K_{\rm d} = K_{\rm apparent} \cdot \frac{[P]_{\rm aktiv}}{[P]_{\rm gesamt}}$$

Das Experiment zur Bestimmung des Anteils an bindeaktiven TFIIIA zeigt Abb. 4.15. Halbsättigung kann in Probe 10 beobachtet werden, in der 750 nM [5S DNA]<sub>komp</sub> eingesetzt wurden. Dies entspricht in etwa dem Doppelten der eingesetzten TFIIIA-Konzentration. Daraus kann gefolgert werden, dass die TFIIIA-Präparation 100 % bindeaktives Protein lieferte.



Abbildung 4.16: Bindung von TFIIIA an 5 S-DNA und an unspezifische DNA ohne MgSO<sub>4</sub> im Gel. Native PAGE, 5 % PAA,  $0.2 \times$  NaTA, Autoradiographie. 5.000 cpm 5 S-DNA (Spuren 1–9) und unspezifische DNA (Spuren 10–18) wurden mit 0, 0.2, 0.4, 0.9, 1.8, 3.5, 7, 14, 28 nM TFIIIA versetzt. Die Spuren 1 und 10 enthalten jeweils 2.500 cpm DNA. Halbsättigung zwischen Spur 3 und 4 bzw. zwischen Spuren 17 und 18.

#### Optimierung der DNA-Bindung

Die bislang gezeigten Verzögerungsgele enthielten  $0.5 \,\mathrm{mM}$  MgSO<sub>4</sub>. Diese annähernd physiologische Mg<sup>2+</sup>-Konzentration im Gel findet allerdings in der Literatur kaum Anwendung. Daher wurde in der vorliegenden Arbeit nun versucht, in einem Mg<sup>2+</sup>-freien Gelsystem eine noch stärkere Bindung zu erzielen. Zusätzlich sollte die Bindung von unspezifischer DNA untersucht werden. Abb. 4.16 zeigt ein Mg<sup>2+</sup>-freies Verzögerungsgel. Es kann eine hochaffine Bindung von TFIIIA an 5S-DNA beobachtet werden. Halbbindung tritt bei sehr niedriger Proteinkonzentration auf (Spur 3). Der ermittelte  $K_d$ -Wert von  $0.6 \,\mathrm{nM}$ liegt in derselben Größenordnung wie der Literaturwert von  $0.33 \,\mathrm{nM}$  für die AtTFIIIA/5 S-DNA-Bindung (Mathieu *et al.*, 2003). Hingegen wird unspezifische DNA mit einem apparenten  $K_d$ -Wert von 20 nM signifikant schlechter gebunden. Da ferner die unspezifische DNA ca. sechs Bindeplätze für TFIIIA aufweist, die alle gleich affin gebunden werden, ist der tatsächliche  $K_d$ -Wert ungefähr doppelt so groß (Steger & Riesner, 2006) und liegt hier somit bei 40 nM. Signale oberhalb der freien unspezifischen DNA stellen keine Retardierungseffekte dar, sondern resultieren aus Verunreinigungen bei der DNA-Präparation (siehe Spur 10, ohne Protein).



Abbildung 4.17: EMSA zur Bestimmung des apparenten  $K_d$ -Wertes für die Bindung von TFIIIA an RNA. Autoradiographie einer nativen PAGE, 5 % PAA, 0,2 x NaTA. Spuren 1–6) 5 S-RNA, Spuren 7–12) PSTVd, Spuren 13–18) 7 S-RNA. Die RNAs wurden mit 0, 1,7, 3,5, 7, 14 und 27 nM TFIIIA titriert. Je 5.000 cpm RNA, Spuren ohne Protein (1,7 und 13) enthalten 2.500 cpm RNA.

Obwohl die hier eingesetzte 5 S-DNA multiple Protein-Bindestellen umfasst, wurde der  $K_d$ -Wert nur für die TFIIIA-Promotor-Wechselwirkung bestimmt. Dies ist legitim unter der Voraussetzung, dass die unspezifischen Bindestellen deutlich schlechter gebunden werden. Unter der ungefähren Annahme, dass auf der eingesetzten 5 S-DNA (330 bp) ein spezifischer und fünf weitere unspezifische Bindungsplätze existieren, welche aber apparent 30fach schwächer komplexiert werden, ist die  $K_d$ -Bestimmung über Halbsättigung hinreichend genau. Aus dem Vergleich der Bindung von spezifischer und unspezifischer DNA in Abb. 4.16 ist deutlich zu erkennen, dass im Konzentrationsbereich der spezifischen Bindung noch keine unspezifische Bindung zu sehen ist.

#### Bindung von TFIIIA an RNA

Zunächst wurden die RNA-Bindeeigenschaften von TFIIIA mit 5S-RNA, PSTVd und 7S-RNA untersucht. Abb. 4.17 zeigt ein Verzögerungsgel mit RNA. Mit steigender Proteinkonzentration korrespondiert eine Abnahme der freien RNA. Dieser Effekt ist nicht auf RNase-Kontamination der Proteinpräparation zurückzuführen, da unterhalb der freien



Abbildung 4.18: Bindung von TFIIIA an unspezifische, einzelsträngige RNA. Native PAGE, 5% PAA, 0,2x NaTA, Autoradiographie. Spuren 1–6) 5S-RNA, Spuren 7–12) TBP-mRNA, Spuren 13–18) PSTVd. Die RNAs wurden mit 0, 3,5, 7, 14, 27, 54 nM TFIIIA titriert. Je 5.000 cpm RNA, Spuren ohne Protein (1,7 und 13) enthalten 2.500 cpm RNA.

RNA keine Abbruchfragmente zu detektieren sind. Allerdings wurde in diesem Experiment evident, dass alle eingesetzten einzelsträngigen RNA-Spezies, darunter auch die 7S-RNA, gleichermaßen sehr affin gebunden werden. Aus dem Vergleich mit den Referenzspuren, die kein Protein enthalten, kann man ein  $K_{\rm d}$  von 20 nM abschätzen. Diese Dissoziationskonstante befindet sich in derselben Größenordnung wie die von Mathieu et al. (2003) für GST-AtTFIIIA gemessene. Allerdings konnte diese Arbeitsgruppe keine Bindung an ein unspezifisches mRNA-Fragment beobachten. In einem weiteren Verzögerungsgel wurde die TFIIIA-Bindung an ein 140 nt-langes Fragment der StTBP-mRNA (siehe Abb. 4.18) sowie an rr1, einem Teilstück einer 18 S-Vorläufer-RNA (Schmitz, 2003) analysiert. Auch diese RNAs werden gleich wie die 5S-RNA und PSTVd bei sehr niedriger TFIIIA-Konzentration gebunden  $(K_d = 15 \text{ nM})$ . Um die Spezifität der RNA-Bindung zu erhöhen wurde in einem Folgeexperiment die KCl-Konzentration im Bindeansatz auf 250 mM erhöht und  $MgSO_4$  zur Gellösung (0,5 mM) zugefügt. Hier konnten ebenfalls keine Unterschiede in den RNA-Affinitäten beobachtet werden (Daten nicht gezeigt). Schließlich war die Bindung von einzelsträngiger RNA in Anwesenheit von  $12,5 \text{ ng}/\mu l$  Poly dIdC und  $125 \text{ ng}/\mu l$  BSA um einen Faktor 20 schwächer (Daten nicht gezeigt). Einzig doppelsträngie RNA, die



Abbildung 4.19: Bindung von TFIIIA an 5S-RNA und dsRNA. Native PAGE, 5% PAA, 0,2 x NaTA, Autoradiographie. Spur 1) Je 2.500 cpm 5S-RNA und dsRNA Spuren 2-9) jeweils 5.000 cpm 5S-RNA und dsRNA simultan titriert mit 7, 14, 27, 54, 110, 220, 440, 880 nM TFIIIA Spur 10) 2.500 cpm dsRNA, Spur 11-14) 5.000 cpm dsRNA titriert mit 110, 220, 440, 880 nM TFIIIA.

durch äquimolare Hybridisierung von (+)- und (-)-strängigen PSTVd-Transkripten gewonnen wurde, wurde deutlich schlechter gebunden. Diesen Befund zeigt ein Retardierungsexperiment, in dem 5 S-RNA und dsRNA simultan mit steigenden TFIIIA-Konzentrationen versetzt wurden (siehe Abb. 4.19).

#### Zusammenfassung der Nukleinsäurebindeeigenschaften von TFIIIA

Das rekombinante GST-TFIIIA aus Tomate zeigt eine außerordentlich affine Bindung an seinen natürlichen Promotor. Es konnte ein  $K_d$ -Wert von 0,6 nM gemessen werden, welcher sehr gut mit den Literaturdaten übereinstimmt (siehe Tab. 4.2). Dagegen wird unspezifische DNA um einen Faktor 60 deutlich schlechter gebunden. Auch 5 S-RNA wird an TFIIIA gebunden; allerdings ist die Affinität geringer und liegt im Bereich der unspezifischen DNA-Bindung ( $K_d = 20 \text{ nM}$ ). Für die RNA-Bindung divergieren die Literaturwerte von pflanzlichem TFIIIA ( $K_d = 30 \text{ nM}$ ; Mathieu *et al.*, 2003) und tierischem TFIIIA ( $K_d = 1 \text{ nM}$ ; Romaniuk, 1985). Bemerkenswerterweise wird die Viroid-RNA gleich affin wie 5 S-RNA gebunden. Gleichwohl muss festgehalten werden, dass alle hier untersuchten

Nukleinsäure	Lycopersicon esculentum <sup>a</sup>	Arabidopsis thaliana <sup>b</sup>	S. cerevisae <sup>c</sup>	S. pombe <sup>d</sup>	Xenopus laevis <sup>e,f</sup>
5S DNA	$0,6\mathrm{nM}$	$0,3\mathrm{nM}$	$0,1\mathrm{nM}$	$0,2\mathrm{nM}$	$1,9\mathrm{nM}$
Unspez. DNA	40 nM	_	_	_	$\approx 190$ nM
5S RNA	20 nM	30 nM	_	_	1  nM
Unspez. RNA	20 nM	> 140 nM	_	_	> 100 nM
PSTVd	20 nM	_	_	_	_
dsRNA	440 nM	_	_	_	_

Tabelle 4.2: Dissoziationskonstanten für die Nukleinsäurebindung von TFIIIA

<sup>a</sup>vorliegende Arbeit

- <sup>b</sup>Mathieu *et al.* (2003)
- <sup>c</sup>Rowland & Segall (1996)
- <sup>d</sup>Schulman & Setzer (2002) <sup>e</sup>Romaniuk (1985)
- <sup>f</sup>Romaniuk (1990)

(1000)

einzelsträngigen RNA mit demselben  $K_d$ -Wert gebunden werden. Lediglich die Affinität von TFIIIA an doppelsträngige RNA ist um einen Faktor 22 geringer.

## 4.2.13 Analyse der Nukleinsäurebindeeigenschaften des ribosomalen Proteins L5

Die 5 S-RNA ist ein zellulärer Bindungspartner des ribosomalen Proteins L5. Beide werden im einem 5 S-RNP-Komplex in den Nucleolus transportiert, um gemeinsam in die große ribosomale Untereinheit inkorporiert zu werden (Michael & Dreyfuss, 1996; Rudt & Pieler, 1996). Für die Bindungsstärke von rekombinantem L5 aus *Xenopus laevis* an die 5 S-RNA wird in der Literatur ein  $K_d$  von 2 nM angegeben (Scripture & Huber, 1995). In der vorliegenden Arbeit wurde die Bindung von L5 an 5 S-RNA, PSTVd und unspezifische TBP-mRNA gemessen (siehe Abb. 4.20). Die Pufferbedingungen wurden entsprechend den TFIIIA-Bindeexperimenten gewählt. Es wurde lediglich der pH des Bindepuffers auf 8,0 gesenkt und auf Zn<sup>2+</sup> verzichtet. Für keine der untersuchten RNAs konnte eine spezifische Bindung ermittelt werden. Die Dissoziationskonstanten wurden auf ca. 300 nM abgeschätzt und lagen somit deutlich, um mehr als zwei Größenordnungen, über den veröffentlichen Zahlen. Doppelsträngige DNA, welche nicht als natürliches Bindeziel von L5 beschrieben ist, wird im micromolaren Bereich unspezifisch gebunden. Eine Variation



Abbildung 4.20: Bindung von L5 an RNA. Native PAGE, 5% PAA, 0,2x NaTA, Autoradiographie. Spuren 1–6) 5S-RNA, Spuren 7–12) TBP-mRNA, Spuren 13–18) PSTVd. Die RNAs wurden mit 0, 60, 180, 600, und 1700 nM L5 titriert. Je 5.000 cpm RNA, Spuren ohne Protein (1,7 und 13) enthalten 2.500 cpm RNA.

einiger experimenteller Parameter, wie Salzkonzentration und pH-Wert des Bindepuffers, brachte keine Verbesserung der beobachteten Affinitäten. Daher ist zu vermuten, dass das unlöslich exprimierte Protein L5 nicht in aktiver Form unter teildenaturierenden Bedingungen aufgereinigt werden konnte. Dies könnte auf Fehlfaltungen in dessen RNA-Bindedomäne zurückzuführen sein.

# 4.3 Alternativ gespleißte TFIIIA-mRNA: Ein konserviertes Phänomen *in planta*

Das letzte Kapitel des Ergebnisteils widmet sich einem Befund, der bei der Klonierung der TFIIIA-Sequenzen eher zufällig beobachtet wurde. Für die rekombinante Expression von TFIIIA musste zunächst die TFIIIA-Sequenz aus einer cDNA-Bank aus *Lycopersicon esculentum* amplifiziert werden. In PCR-Reaktionen mit Primern, die den ORF der TFIIIA-mRNA flankieren, konnten interessanterweise zwei Produkte detektiert werden. Davon entsprach eines der erwarteten Länge von 1119 nt. Das andere Produkt erwies sich als ca. 100 nt länger und wurde sogar in größerer Menge detektiert (siehe Abb. 4.21, Spur Le). Das Auftreten eines alternativen Produktes deutet auf die Existenz einer alternativ gespleißten TFIIIA-mRNA hin. Da es sich bei TFIIIA um ein äußerst wichtiges, multifunktionelles Protein handelt, wurde dieses initiale Ergebnis in der vorliegenden Arbeit weiter verfolgt. TFIIIA aus *Xenopus laevis* war der erste eukaryotische, genspezifische Transkriptionsfaktor, der isoliert und eingehend charakterisiert wurde (Engelke *et al.*, 1980). Nichtsdestoweniger wurde alternatives Spleißen für TFIIIA bislang nicht in der Literatur beschrieben.

#### 4.3.1 Sequenzanalyse des alternativen Produktes

Zur Sequenzanalyse wurden beide Banden getrennt geleluiert und die aufgereinigte DNA in pGEM-T subkloniert. Mittels der Standardoligonukleotide M13fwd und M13rev wurden beide Polaritäten der *insert*-DNA sequenziert. Die kürzere TFIIIA-mRNA zeigte, wie oben beschrieben, die erwartete Sequenz mit einem ununterbrochenen ORF, der für neun Zink-Finger-Bindemotive kodiert (Zugriffsnr. DQ882179). Die längere TFIIIA-mRNA dagegen weist ein zusätzliches Exon auf, welches ein vorzeitiges Stopp-Codon enthält (siehe Abb. 4.22). Durch konventionelle Translation dieser alternativ gespleißten mRNA würde somit eine verkürzte Form von TFIIIA gebildet, welche bei einem Molekulargewicht von 10,7 kDa nur aus den ersten beiden Zink-Finger-Domänen bestehen würde. Für die alter-



Abbildung 4.21: TFIIIA, Amplifikation von cDNA. 1% ige Agarosegele,  $1 \times \text{TAE}$ , mit Ethidiumbromid gefärbt. Spur At) PCR-Produkt von cDNA aus Arabidopsis thaliana, Spur M1) Lambda-DNA/HindIII, 9419, 6559, 4261, 2322, 2023 bp, Spur Le) PCR-Produkt von cDNA aus Lycopersicon esculentum, Spur M2) kB-Leiter 10000, 8000, 6000, 5000, 4000, 3000, 2500, 2000, 1500, 1000, 750, 500 bp. Die beiden Isoformen der TFIIIA-mRNA sind durch Doppelpfeile gekennzeichnet.

#### A) TFIIIA-mRNA



Abbildung 4.22: Intron/Exon-Struktur des TFIIIA-Gens (Zugriffsnr. DQ882180). Exons sind durch nummerierte graue Boxen, Introns durch Striche gekennzeichnet. A) Die konstitutiv gespleißte TFIIIA-mRNA kodiert für die TFIIIA-Vollänge mit neun Zink-Finger-Motiven (Zugriffsnr. DQ882180). B) Alternativ gespleißte TFIIIA-mRNA mit zusätzlichem Exon 3, welches ein vorzeitiges UGA-Stopp-Codon enthält (Zugriffsnr. DQ882179).

nativ gespleißte TFIIIA-mRNA wurde ein eigener Datenbankeintrag unter der Zugriffsnr. DQ882180 angelegt; sie wird im Folgenden als TFIIIA-mRNA<sup>alt</sup> bezeichnet. Die Sequenz des TFIIIA-Gens aus *Lycopersicon esculentum* findet sich unter der Zugriffsnr. DQ882178.

#### 4.3.2 Vergleich mit anderen pflanzlichen TFIIIA-Sequenzen

Interessanterweise konnte oben beschriebenes Phänomen auch bei der Klonierung der TFIIIA-Sequenz aus Arabidopsis thaliana beobachtet werden. Hier wurden ebenfalls zwei mRNAs amplifiziert (siehe Abb. 4.21, Spur At). Die Sequenzen der beiden Produkte korrespondieren mit den Datenbankeinträgen AY186610 und AY054225. Die TFIIIA-mRNA<sup>alt</sup> aus Arabidopsis ist in der Datenbank aber lediglich als "kodierend für ein putatives Zink-Finger-Protein" annotiert.

Des Weiteren konnte durch BLAST-Suche in der EST-Datenbank ein TFIIIA-Homolog für *Oryza sativa* gefunden werden (Zugriffsnr. AK099528). Diese Sequenz aus Reis zeigt sogar deutlich bessere Übereinstimmung mit der *Arabidopsis*- und Tomaten-TFIIIA-Sequenz als der automatisch annotierte Eintrag Nr. BAA90519. Eine weitere *in silico*-Ana-

Spezies	TFIIIA-mRNA	$\rm TFIIIA\text{-}mRNA^{alt}$	TFIIIA-Gen
$Ly copersicon \ esculent um$	DQ882180	DQ882179	DQ882178
Arabidopsis thaliana	AY186610	AY054225	AC069273
Oryza sativa	AK099528	AK101631	AC093088
Homo sapiens	D32257	CR614643	AL137059

Tabelle 4.3: TFIIIA, Zusammenfassung der Datenbankzugriffsnummern.

lyse ergab ferner, dass auch in Reis eine alternative TFIIIA-mRNA existiert (AK101631). Somit kann für die nur entfernt verwandten Familien der dikotylen *Brassicaceae* und *Solanaceae* sowie der monokotylen *Poaceae* ein gemeinsamer Mechanismus zum alternativen Spleißen der TFIIIA-mRNA nahegelegt werden. In allen drei Familien findet sich eine alternativ gespleißte TFIIIA-mRNA, die für eine auf zwei Zink-Finger-Motive verkürzte TFIIIA-Variante kodiert. Dieses Genprodukt wird im Folgenden als kTFIIIA bezeichnet. Schließlich sei erwähnt, dass sogar menschliche TFIIIA-mRNA in einer alternativ gespleißten Form vorliegt (CR614643). Die verkürzte Variante weist allerdings vier statt im pflanzlichen System zwei Zink-Finger auf.

Für mRNAs, die aufgrund von Fehlern bei der Transkription oder durch alternative Spleißprozesse ein vorzeitiges Stopp-Codon (PTC, von engl.: premature termination codon) enthalten, ist als ein Abbaumechanismus der sogenannte nonsense mediated decay (NMD) beschrieben (zur Übersicht: Conti & Izaurralde, 2005; Lejeune & Maquat, 2005). Auf diese Weise kann sich die Zelle vor abberanten Transkripten schützen, die für trunkierte und potentiell schädliche Proteine kodieren. Für den beschriebenen Fall der TFIIIA-mRNA<sup>alt</sup> scheint dieser Mechanismus indes nicht zu greifen, da diese sogar in höherer Konzentration als die konstitutiv gespleißte Form vorliegt.

Dieses ganz offensichtlich konservierte Phänomen, dass die TFIIIA-mRNA in einer alternativ gespleißten Isoform auftritt, wurde im Folgenden experimentell genauer untersucht. Dabei sollten Antworten auf die folgenden Fragen gefunden werden:

- Wie wird die TFIIIA-mRNA<sup>alt</sup> in vitro translatiert?
- Besitzt rekombinantes kTFIIIA Nuklensäurebindeaktivität?
- Lässt sich kTFIIIA in vivo nachweisen?

#### 4.3.3 In vitro-Translation der alternativ gespleißten TFIIIA-mRNA

Zunächst wurde die Frage gestellt, welches Genprodukt von der alternativ gespleißten TFIIIA-mRNA gebildet wird. Dazu bietet sich der *wheat germ extract plus*<sup>TM</sup>(Promega) als ein zellfreies in vitro-Translationssystem an. Gemäß konventioneller Vorstellung würde diese mRNA ausgehend von ihrem 5'-Cap gescannt werden. Bei Erkennung eines AUG-Startcodons würde die Translation initiert und im Exon 3, welches das Stopp-Codon enthält, terminiert werden. Somit würde am Beispiel von Arabidopsis thaliana ein 14,1 kDa-Protein übersetzt, welches zwei Zink-Finger-Motive enthält. Alle drei beschriebenen pflanzlichen TFIIIA-mRNA<sup>alt</sup> tragen ein solches Stopp-Codon in Exon 3. Die Exons 4-7 bilden ein intaktes, ununterbrochenes Leseraster, welches den C-Terminus von TFIIIA kodiert. Dessen Expression scheint aus klassischer Sichtweise nicht möglich. Andererseits wurden Mechanismen beschrieben, nach denen RNA unabhängig vom 5'-Cap translatiert werden kann. So vermögen einige virale RNAs Strukturen auszubilden, sogenannte internal ribosome entry sites (IRES), mit denen sie direkt an die kleine ribosomale Untereinheit binden und translatiert werden können. Darüber hinaus gibt es erste Hinweise auf die Existenz von zellulären IRES-Elementen, die gleichwohl äußerst kontrovers diskutiert werden (Jackson, 2005; Kozak, 2005). Mit einer in vitro-Translation könnte gezeigt werden, ob eine unkonventionelle Expression der TFIIIA-mRNA<sup>alt</sup> möglich ist.

Für diese Analyse wurde die TFIIIA-mRNA und die TFIIIA-mRNA<sup>alt</sup> aus Arabidopsis thaliana ausgewählt, da diese vollständig, inklusive ihrer 3'- und 5'-untranslatierten Regionen, sequenziert und annotiert wurden (Zugriffsnr. AY054225). Auf dieser Grundlage wurden Oligonukleotide entworfen, welche die komplette mRNA flankieren. Der 5'-Primer war mit einer T7-Promotorsequenz ausgestattet, um das PCR-Produkt mittels *in vitro*-Transkription in RNA überführen zu können. Der 3'-Primer wurde um 30 Thymidin-Nukleotide verlängert, um einen natürlichen Poly(A)-*tail* zu generieren. Durch Zugabe eines 7-Methylguanosin-Cap-Analogs zur IVT-Reaktion wurden letzlich RNAs erzeugt, die identisch zu den natürlichen, eukaryotischen mRNAs waren (siehe auch 3.2.9). Auf diese Weise wurden beide AtTFIIIA-mRNA-Varianten synthetisiert.

In einem ersten Translationsexperiment wurden steigende Konzentrationen an TFIIIA-mRNA<sup>alt</sup> in den Weizenkeimextrakt gegeben. Die *de novo* synthetisierten Proteine sind <sup>35</sup>S markiert und können nach Auftrennung in einer SDS-PAGE mittels Autoradiographie sichtbar gemacht werden. Korrelierend mit der Menge zugegebener mRNA



Abbildung 4.23: In vitro-Translation von TFIIIA-mRNA<sup>alt</sup>. De novo-synthetisierte Proteine sind <sup>35</sup>S-markiert. Autoradiographie einer 12 % igen SDS-PAGE. Spur 1) 0,5  $\mu$ g Standard-BMV-RNA, translatiert in 94-, 35-, 20-kDa-Proteine, Spur 2–6) 8, 4, 2, 1, 0,5  $\mu$ g TFIIIA-mRNA<sup>alt</sup>, Spur 7) Nullkontrolle ohne Zugabe von exogener mRNA.

erscheint ein Signal zunehmender Intensität auf Höhe von ca. 15 kDa (siehe Abb. 4.23). Diese Beobachtung kann in Einklang mit einer konventionellen Translation der TFIIIAmRNA<sup>alt</sup> gebracht werden, weil nur der N-Terminus exprimiert wird. Des Weiteren tritt ein Signal bei etwas höherem Molekulargewicht in Erscheinung. Dies könnte dadurch zu erklären sein, dass sich das erste AUG-Startkodon in einem ungünstigen Sequenzkontext befindet und daher die Translation bevorzugt am zweiten AUG initiert wird. Stärkere Signale im unteren Drittel des Gels sind auf freies oder unspezifisch verknüpftes <sup>35</sup>S-L-Methionin zurückzuführen. Ein sehr schwaches Signal im Bereich von ca. 30 kDa scheint artifizieller Natur zu sein, da es sowohl unabhängig von der mRNA-Konzentration auftritt als auch in der Nullkontrolle sichtbar ist.

In einem weiteren Experiment wurde zusätzlich das Translationsmuster der "normal" gespleißten TFIIIA-mRNA analysiert. Außerdem wurden die eingesetzten mRNAs einer Strukturbehandlung unterzogen. Nach vollständiger Hitzedenaturierung wurden die Proben wahlweise schnell auf Eis oder langsam auf Raumtemperatur abgekühlt. Dadurch stellt sich eine kinetisch kontrollierte oder die thermodynamisch optimale Struktur ein. Man erkennt in Abb. 4.24, dass auch die konstitutiv gespleißte mRNA konventionell trans-



Abbildung 4.24: In vitro-Translation von TFIIIA-mRNA und TFIIIA-mRNA<sup>alt</sup>. De-novo-synthetisierte Proteine sind <sup>35</sup>S-markiert. Autoradiographie einer 12 %igen SDS-PAGE. Vor Zugabe zum Weizenkeimextrakt wurde die RNA struktureingestellt Spuren SN) "snap cool", Spuren SL) "slow cool"), BMV) Standard-BMV-RNA, translatiert in 94-, 35-, 20 kDa-Proteine, Spuren TFIIIA-mRNA<sup>alt</sup>) mit 5'-Cap und PolyA-tail bzw. nur mit PolyA-tail. Spuren TFIIIA-mRNA) mit 5'-Cap und PolyA-tail bzw. nur mit PolyA-tail, Spur Null) Nullkontrolle ohne Zugabe von exogener mRNA.

latiert wird. Auffallend ist, dass alle kinetisch gefalteten mRNAs effizienter translatiert werden als die jeweiligen thermodynamisch optimalen Strukturen. Überraschenderweise wird die TFIIIA-mRNA<sup>alt</sup> mit einem 5'-Cap nur sehr schwach exprimiert. In den folgenden Experimenten sollten Hinweise auf Nukleinsäurebindeeigenschaften und die *in vivo*-Existenz der TFIIIA-mRNA<sup>alt</sup> gefunden werden.

#### 4.3.4 Bindestudien mit rekombinantem kTFIIIA

Das alternative Genprodukt kTFIIIA besitzt bei einer Größe von 14,1 kDa zwei vollständige Zink-Finger-Motive. Es kann also durchaus davon ausgegangen werden, dass es sich hierbei um ein funktionelles, intaktes Protein handelt. Um nähere Hinweise auf die Funktion von kTFIIIA zu erhalten, wurde die Nukleinsäurebindeaktivität von rekombi-



Abbildung 4.25: Rekombinante Expression von GST-kTFIIIA und dessen Nukleinsäurebindeeigenschaften. A) SDS-PAGE, 10 % PAA (30:0,8), gefärbt mit Coomassie-Brilliant-Blue. Spur M) Mark12<sup>TM</sup> MW-Standard mit Proteinen der Größen 200, 116, 97, 66, 55, 36, 31, 22 und 14 kDa, Spur TC) *total cell*, Spur CL) *cleared lysate*, Spur FT) *flow through*, Spur E1) Eluat 1, Spur E2) Eluat 2. B) Autoradiographie eines 5 %igen PAA-Gels (43:1), 0,2 x NaTAE. Spur 0) 2.500 cpm 5 S-DNA ohne Protein, Spuren 1–7) 5.000 cpm 5 S-DNA titriert mit steigenden Konzentrationen an GST-kTFIIIA (14, 27, 54, 109, 218, 435, 870 nM).

nantem kTFIIIA in einem Gelverzögerungsexperiment analysiert. Dazu wurde die Sequenz von kTFIIIA aus Lycopersicon esculentum kloniert und ein GST-kTFIIIA-Fusionsprotein in *E. coli* synthetisiert. Auch dieses rekombinante Protein erwies sich als unlöslich. Die Aufreinigung erfolgte daher mittels Sarkosyl, wie bereits für LeTFIIIA beschrieben (siehe Methoden, Abschnitt 3.1.6). Auf diese Weise gelang es einen kleinen Anteil an GST-kT-FIIIA in eine lösliche Form zu überführen (siehe Abb. 4.25A, Spur CL). Aus dem Lysat konnte hochreines, lösliches Eluat gewonnen werden (siehe Abb. 4.25A, Spuren E1 und E2), das gegen Bindepuffer in 1,5 micromolarer Konzentration dialysiert werden konnte. Eine weitere Aufkonzentrierung per Ultrafiltration schlug indes fehl, da GST-kTFIIIA oberhalb einer Konzentration von ca. 1,5  $\mu$ M aggregierte und ausfiel. Als Bindepartner für GST-kTFIIIA wurden 5S-DNA und 5S-RNA getestet. Wie aus Abb. 4.25B ersichtlich, lässt sich bis zu einer GST-kTFIIIA-Konzentration von 870 nM keine 5S-DNA-Bindung detektieren. Auch konnte keine Bindung an 5S-RNA festgestellt werden (Daten nicht gezeigt).



Abbildung 4.26: Nachweis von TFIIIA und kTFIIIA *in vivo* mit Antiserum 24. A) SDS-PA-GE, 12 % PAA (30:0,8), gefärbt mit Coomassie-Brilliant-Blue. Spur 1–3) Gesamtproteinextrakt aus *Lycopersicon esculentum* (PE), Verdünnungsreihe. B) Westernblot mit Anti-TFIIIA-Serum, Kaninchen 24. Spuren 4–6) Proteinextrakt (PE), Konzentrationsreihe wie in A, Spuren 7–10) Antigen GST-TFIIIA (AG), 0,2, 2, 20, 200 ng *inclusion bodies*.

#### 4.3.5 In vivo-Nachweis von kTFIIIA mittels Immunoblot

Die Fragestellung, ob die verkürzte TFIIIA-Variante *in vivo* existiert, wurde mit einem immunologischen Nachweisverfahren behandelt. Polyklonale Antikörper gegen GST-*Le*TFIIIA wurden durch Immunisierung zweier Kaninchen erhalten. Zur Präparation des Antigens wurde GST-*Le*TFIIIA rekombinat unlöslich überexprimiert. Die *inclusion bodies* wurden durch Sonifikation und Zentrifugation vorgereinigt und aus einer präparativen SDS-PAGE in großem Maßstab (ca. 2 mg) geleluiert. Die Gelstücke wurden gründlich gewässert, um toxische SDS-Reste zu entfernen. Die Immunisierung wurde bei der Firma Biotrend (Köln) in Auftrag gegeben.

Ein Gesamtproteinextrakt wurde aus Tomatenblättern, wie im Methodenteil unter 3.1.1 beschrieben, präpariert und durch eine SDS-PAGE aufgetrennt (siehe Abb. 4.26A). Per Westernblot wurde versucht, sowohl die TFIIIA-Volllänge (42,9 kDa) als auch die alternative Variante kTFIIIA (10,7 kDa) nachzuweisen. Für das Detektionsverfahren mittels Chemolumineszenz musste zunächst eine optimale Erstantikörper-Konzentration ermittelt



Abbildung 4.27: Nachweis von TFIIIA und kTFIIIA *in vivo* mit Antiserum 23. A) Westernblot mit Anti-TFIIIA-Serum, Kaninchen 23. Spur 1) Antigen GST-TFIIIA (AG), 20 ng *inclusion bodies*, Spuren 2–4) Proteinextrakt aus Lycopersicon esculentum (PE), Verdünnungsreihe wie in Abb. 4.26A, B) Westernblot mit Präimmunserum, Kaninchen 23. Spuren 5–7) Proteinextrakt aus Lycopersicon esculentum (PE), Verdünnungsreihe wie in Abb. 4.26A.

werden. Eine 1:200 Verdünnung des Anti-TFIIIA-Serums zeigte gute Ergebnisse bezüglich hoher Sensitivität und hohem Signal/Rausch-Verhältnis (Abb. 4.26B). Die Nachweisgrenze für das GST-TFIIIA-Antigen lag bei etwa 2 ng. In Spur 6 mit der höchsten Gesamtprotein-Konzentration zeigt sich ein sehr schwaches, diffuses Bandenmuster im Bereich von ca. 25–45 kDa. Signale, die mit den erwarteten Molekulargewichten von 10,7 bzw. 42,9 kDa korrespondieren, konnten indes nicht detektiert werden. Banden unterhalb des 69 kDa-Antigens resultieren aus Kreuzreaktionen mit endogenem *E. coli*-Protein.

In einem weiteren Experiment wurde für den TFIIIA-Nachweis das Antiserum des zweiten immunisierten Kaninchens getestet (siehe Abb. 4.27). Hierbei lag die Nachweisgrenze für das Antigen um etwa eine Größenordnung schlechter bei 20 ng. In den Spuren mit Gesamtprotein konnte eine intensive Doppelbande bei etwa 35 kDa beobachtet werden. Dieses Signal trat allerdings auch in einem Detektionsversuch mit Präimmunserum auf (siehe Abb. 4.27B). Daher liegt der Verdacht nahe, dass das Serum dieses Individuums mit pflanzlichem Protein unspezifisch kreuzreagiert.

## 4.3.6 Alternativ gespleißte TFIIIA-mRNA: Zusammenfassung der Ergebnisse

Es konnte experimentell gezeigt werden, dass in Lycopersicon esculentum und Arabidopsis thaliana das TFIIIA-Gen alternativ gespleißt wird. Der Mechanismus scheint in der Pflanzenwelt konserviert zu sein, wie ein *in silico*-Vergleich mit Oryza sativa zeigt. Die TFIIIA-mRNA<sup>alt</sup> enthält ein zusätzliches Exon mit vorzeitigem Stopp-Codon. Sie wird in der Pflanze offenbar nicht durch NMD degradiert. In einem zellfreien Weizenkeimextrakt wurde die TFIIIA-mRNA<sup>alt</sup> erwartungsgemäß in ein 15 kDa-Protein (kTFIIIA) translatiert. Die Translationseffizienz dieser mRNA mit 5'-ständigem 7-Methylguanosin-Cap war geringer als ohne 5'-Modifikation.

Für rekombinantes GST-kTFIIIA konnte auch bei hoher Konzentration von 870 nM keine Nukleinsäurebindeaktivität nachgewiesen werden. Ferner war kTFIIIA wie auch TFIIIA mittels Westernblot nicht *in vivo* nachzuweisen.

# 5

# Diskussion

### 5.1 Die Startstelle der PSTVd-Transkription

Das Viroid wird von der wirtseigenen Pol II repliziert. In einer vorausgehenden Arbeit konnte gezeigt werden, dass an Position U359 oder C1 die Transkription des (+)strängigen Zirkels beginnt (Kolonko, 2003). Auf einem Startnukleotid der Transkription sollte *in planta* ein starker Selektionsdruck lasten. Demzufolge sollten Mutationen dieses Nukleotids eine hohe Reversionsrate aufweisen. In der vorliegenden Arbeit wurde daher die Infektiosität von zwei PSTVd-Sequenzen analysiert, die an Position 359 bzw. 1 punktmutiert waren. Es konnte gezeigt werden, dass über einen Zeitraum von fünf Wochen die Mutation G1 genetisch stabil war, während G359 zu U359 revertierte.

#### 5.1.1 Struktur der Startdomäne

Der linke, terminale Loop von PSTVd mit der Sequenz CUCG zeigt einige Charakteristika eines extrastabilen Tetraloops vom Typ UNCG (Scholtysik, 2006; Varani *et al.*, 1991). Das Denaturierungsprofil dieses Motivs lässt auf eine hohe thermodynamische Stabilität schließen. Mittels NMR-Spektroskopie konnten ferner UUCG-typische Resonanzen gemessen werden. Anhand der 3D-Konformation eines UNCG-Tetraloops, gezeigt in Abb. 5.1, lässt sich der Einfluss der Mutationen G359 und G1 auf die Tetraloopstruktur abschätzen.

Das Nukleotid an Position 359 befindet sich nicht im strukturellen Kontext des Tetraloopmotivs, da es mit Position N in der Konsensussequenz UNCG korrespondiert. Somit



Abbildung 5.1: NMR-Struktur des UUCG-Tetraloops. Der UUCG-Tetraloop und der CUCG-Loop des Viroids weisen ein identisches Resonanzprofil auf (Scholtysik, 2006; Varani *et al.*, 1991). Die Ribosen der Nukleotide sind nummeriert analog ihrer Position im Viroid. Stabilisierende *stacking*-Wechselwirkungen existieren zwischen den Basen U358 und C1 sowie zwischen G2 und G3. Hingegen trägt das Uracil an Position 359 nicht zu den stabilisierenden Wechselwirkungen in der Tetraloopstruktur bei.

hat die Mutation G359 keinen Einfluss auf die Struktur des Viroids, sondern es handelt sich folglich um eine reine Sequenzänderung. Auf diesem Nukleotid lastet aber offenbar ein starker selektiver Druck. Eine Mutation nach Guanosin an dieser Stelle wird nicht toleriert, sie revertiert zum Retardierungseffekte Uracil des Wildtyps.

Dagegen destabilisiert ein Austausch von C1 durch G1 den Tetraloop signifikant und zerstört somit dessen strukturelle Integrität (Dr. M. Schmitz, mündliche Mitteilung). Bemerkenswerterweise erwies sich diese Mutation *in planta* als genetisch stabil. Gleichwohl zeigt die G1-Population eine langsamere Wachstumskinetik, verglichen mit der Wildtyppopulation. Dies könnte darauf zurückzuführen sein, dass durch den Verlust der Tetraloopstruktur die Transkription nur suboptimal initiiert wird. Das Startnukleotid U359 ist nur in einer korrekten Tetraloopumgebung besonders exponiert und somit für eine Wechselwirkung mit Wirtsfaktoren gut zugänglich. Daher ist es durchaus wahrscheinlich, dass G1 in Folgegenerationen zum Wildtyp C1 revertiert und somit die Tetraloopstruktur wiederhergestellt wird. Tatsächlich konnte in einer G1-infizierten Pflanze eine solche Reversion in einem Frühstadium beobachtet werden. Es ist zu erwarten, dass sich diese Wildtyp-Population in der Pflanze etabliert und aufgrund ihrer größeren *fitness* die G1-Nachkommenschaft überwächst.

Auch für ASBVd und PLMVd, zwei Viroiden aus der Gruppe der Avsunviroidae, wurden Uracil-Nukleotide, die sich in endständigen *loops* befinden, als Startpunkte der
Transkription identifiziert (Navarro & Flores, 2000; Pelchat *et al.*, 2002). Aufgrund dieser Übereinstimmung einen gemeinsamen evolutionären Ursprung nahezulegen scheint indes fragwürdig. Da die Transkription der *Avsunviroidae* von einer plastidären Polymerase vermittelt wird, gehorcht sie wahrscheinlich einem gänzlich anderen Initiationsmechanismus.

Die anhand der Mutagenesestudien erzielten Ergebnisse stehen in gutem Einklang mit den biochemischen und strukturellen Befunden von Kolonko (2003) und Scholtysik (2006). Dennoch kann nicht *a priori* ausgeschlossen werden, dass die Mutationen andere Aspekte des Infektionsmechanismus von PSTVd als die Transkription beeinflussen. Für den systemischen Befall der Wirtspflanze müssen letztlich sämtliche funktionellen Domänen des Viroids intakt sein. Der Ausfall einer einzelnen Funktion des Viroids kann somit zu einem kompletten Verlust der Infektiosität führen. Daher ist es auch vorstellbar, dass durch Mutagenese des linken Loops die Transportfähigkeit des Viroids gestört wurde. Dies kann sowohl den intra- als auch den interzellulären Transport betreffen, die beide mit großer Wahrscheinlichkeit auf RNA-/Proteinwechselwirkungen zurückzuführen sind.

Darüber hinaus könnten Mutationen die *template*-Aktivität des (-)-Stranges beeinflussen. Ein selektiver Druck lastet auch auf solchen Nukleotiden, die auf (-)-Strangebene Strukturen ausbilden, um Wirtsfaktoren zu rekrutieren. Beispielsweise kann auch die komplementäre Sequenz des linken Terminus (CGAG) einen extrastabilen Tetraloop formen. Dieser ist vom Typ YNMG (Y=Pyrimidin, M=C oder A). Die Mutationen G359 und G1 zerstören allerdings nicht die Tetraloopstruktur des (-)-Stranges, da beide Sequenzmotive, CGCG und CCAG, dem Konsensus YNMG entsprechen. Dennoch könnte auch für den Transkriptionsstart auf dem (-)-Strang die exakte Sequenz eine entscheidende Rolle spielen. Da auch auf (-)-Strangebene ein Promotorelement in Form des metastabilen Hairpin II existiert (Loss *et al.*, 1991; Qu *et al.*, 1993), ist anzunehmen, dass auch hier ein definierter Transkriptionsbeginn festgelegt ist. Ein mutiertes Startnukleotid der (-)-Strangtranskription könnte demnach auch dazu führen, dass die Infektiosität geschwächt wird.

Indes scheint es äußerst unwahrscheinlich, dass Mutationen im linken Terminus mit der Prozessierung des Viroids interferieren, da die Erkennungssignale für die nukleolytische Aktivität innerhalb der zentralen konservierten Region (CCR) liegen (Baumstark & Riesner, 1995b; Baumstark *et al.*, 1997). So werden Minimalviroide, die lediglich aus der CCR bestehen, *in vitro* prozessiert und zu sogenannten Minizirkeln ligiert (Schrader *et al.*, 2003). Am Beispiel der Prozessierung und Ligation wird ferner offensichtlich, dass das Viroid mit nur einer Sequenz zwei unterschiedliche Funktionen erfüllen kann: Durch Ausbildung von Alternativstrukturen wird es wahlweise geschnitten oder ligiert.

Dies verdeutlicht ein inhärentes Problem von Infektionsstudien mit mutiertem PSTVd: Allein auf Basis von Wachstums- und Reversionskinetiken von mutierten PSTVd-Molekülen ist es nicht möglich, Struktur-/Funktionsbeziehungen herzuleiten. Dazu bedarf es komplementärer biochemischer und struktureller Analysen.

Als eine zusätzliche Methode, die hier dargestellten Ergebnisse zu untermauern, bietet sich eine Transkriptionsanalyse *in vitro* an. Im hiesigen Institut wurde ein transkriptionsaktiver Kernextrakt aus Kartoffelzellen der Linie HH258 etabliert (Fels *et al.*, 2001; Kolonko *et al.*, 2006; Nagel-Steger, 1990). Mit diesem System lassen sich (+)- und (-)-Strangtranskription getrennt analysieren. Nach Zugabe eines exogenen Viroid-*templates* kann man *de novo*-synthetisierten (-)-Strang *per* RT-PCR nachweisen. Sofern kein (-)strängiges Intermediat ausgehend von Mutante G359 transkribiert würde, wäre bestätigt, dass die Position 359 essentiell für die Transkription des (+)-Stranges ist. Somit wäre letztlich auch die Ursache des starken selektiven Druckes auf der Mutation G359 *in planta* eindeutig identifiziert.

Nach derzeitiger Kenntnislage, basierend auf molekularbiologischen, biochemischen und strukturellen Daten, lässt sich bezüglich der Transkriptionsinitiation von PSTVd zusammenfassend postulieren:

- Die Transkription des zirkulären PSTVd beginnt an Position U359.
- Notwendig für die Initiation ist die Sequenzidentität des Nukleotids U359.
- Bei intakter Tetraloopstruktur des linken Terminus wird die Transkription effektiver initiiert.

#### 5.1.2 Reversionsmechanismus mutierter Viroid-Nukleotide

Mutierte Viroidmoleküle sind in der Pflanze selektivem Druck ausgesetzt. Mutierte Nukleotide, die biologische Funktionen des Viroids negativ beeinflussen, revertieren mit großer Wahrscheinlichkeit zum Wildtypnukleotid, sofern eine geringe Replikation überhaupt noch möglich ist. Dagegen können sich Mutationen als genetisch stabil erweisen, wenn funktionsrelvante Strukturen des Viroids weitestgehend intakt bleiben.



Abbildung 5.2: Modell zum Reversionsmechanismus von PSTVd-Mutationen. Eine Mutation des Startnukleotids (X) führt dazu, dass das Viroid nicht mehr von Pol II als *template* akzeptiert wird. Stattdessen wird das Viroid durch andere zelluläre Polymerasen (Pol I, Pol III oder RdRP) basal transkribiert. Die übrigen Schritte im Replikationszyklus werden nicht durch die Mutation gestört. Durch die natürliche Fehlerrate der RNA-Polymerasen kann sowohl bei der Transkription des (+)-Stranges als auch des (-)-Stranges die Wildtypsequenz wiederhergestellt werden. Eine Reversion (R!) hat zur Folge, dass das Viroid mit hoher Rate transkribiert wird und sich der Wildtyp gegen die Mutante *in planta* durchsetzt.

Mutationen, die beispielsweise für den Transkriptionsstart essentiell sind, können bewirken, dass das Viroid nicht mehr von Pol II kopiert wird. Trotzdem kann die Population der Viroidmutanten überleben, wenn sie auf basaler Ebene von anderen zellulären RNA-Polymerasen vermehrt wird (siehe Abb. 5.2).

In vitro konnten bei der Transkription von mutierten Viroid-*templates* Signale detektiert werden, die sich als unabhängig von dem Pol II-Inhibitor  $\alpha$ -Amanitin erwiesen (Hu, 2000; Moors, 2004). Dies spricht für eine Beteiligung anderer RNA-Polymerasen an der basalen Viroidtranskription. Unabdingbar für das frequente Auftreten von Reversionen ist ferner eine gewisse Ungenauigkeit der RNA-Polymerase. Im Gegensatz zu DNA-Polymerasen zeigen RNA-Polymerasen eine recht hohe Fehlerrate. Diese liegt für die Pol II bei  $1, 4 \times 10^{-3}$  (Nesser *et al.*, 2006). Durch Einbau eines "falschen" Nukleotids durch die RNA-Polymerasen kann während der (+)- oder der (-)-Strang-Transkription die Wildtypsequenz wiederhergestellt werden. Die Wildtypsequenz besitzt die höhere *fitness*, weil sie effizienter transkribiert wird. Dies hat schließlich zur Folge, dass die Wildtyppopulation die Mutantenpopulation sehr rasch überwächst.

## 5.2 Viroid-Protein-Komplexe

#### 5.2.1 Methodische Gesichtspunkte

Protein/Viroid-Wechselwirkungen *in vivo* konnten erst in wenigen Fällen identifiziert werden. Mittels UV-Crosslinking gelang es, rekonstituierte Viroid/Protein-Komplexe nachzuweisen. Es konnte ein 43 kDa-Protein aufgereinigt werden, für das man wegen seiner ausgeprägten Basizität eine Ähnlichkeit zu Histonen nahelegte (Klaff *et al.*, 1989). Aus Phloem-Exsudat konnten Viroid-Protein-Komplexe spezifisch immunopräzipitiert werden (Gómez & Pallás, 2004). So erhielt man starke Hinweise darauf, dass das Phloem Protein 2 (CsPP2) an der systemischen Verbreitung des Hop Stunt Viroid (HSVd) beteiligt ist. Des Weiteren wurde *in vivo* durch *screening* einer cDNA-Expressionsbank das Protein VirP1 als ein Interaktionspartner des Viroids ermittelt (Maniataki *et al.*, 2003). Die Wechselwirkung wurde in einem *yeast-three-hybrid*-Systems bestätigt (Martínez de Alba *et al.*, 2003). Das Protein VirP1 wurde mit dem intrazellulären Transport von PSTVd in Verbindung gebracht.

Ein indirekter Nachweis einer Viroid/Protein-Wechselwirkung konnte im Falle der Pol II geführt werden. Die Replikation des Viroids erwies sich als sensititiv gegenüber dem Pol II-Inhibitor  $\alpha$ -Amanitin (Mühlbach & Sänger, 1979). Auch wurde eine Wechselwirkung des Viroids mit der Proteinkinase PKR nahegelegt. Diese ist in Proteinextrakten aus Viroid-infizierten Pflanzen verstärkt phosphoryliert (Hiddinga *et al.*, 1988). Allerdings konnte dieser Befund nicht reproduziert werden. Darüber hinaus wurde die Viroid-RNA *in vitro* nur sehr schwach von aufgereinigter PKR gebunden (Schmitz, 2003).

In der vorliegenden Arbeit wurden Viroid-Proteinwechselwirkungen mit einer gänzlich anderen Herangehensweise behandelt. Anstelle eines *in vivo*-Nachweises mit einer der oben aufgeführten Methoden wurden hier Modellproteine, für die basierend auf strukturellen Erkenntnissen eine Funktion bei Replikation oder Transport des Viroids postuliert wurde, rekombinant synthetisiert und auf ihre Viroid-Bindungseigenschaften untersucht. Für die Analyse solcher RNA/Protein-Wechselwirkungen hält die Molekularbiologie ein breites Methodenspektrum bereit. Dabei lässt sich auf dieselben biochemischen Methoden zurückgreifen, mit denen erfolgreich DNA/Protein-Wechselwirkungen charakterisiert wurden. Darunter fallen die Verzögerungsgelelektrophorese, der Filterbindungstest und die *footprint*-Analyse. Um die Dissoziationskonstante für die Protein/Nukleinsäure-Wechselwirkungen zu bestimmen, wurden in der vorliegenden Arbeit Gelverzögerungsexperimente durchgeführt. Dabei wurde auf den im hiesigen Institut etablierten Nachweis durch Silberfärbung zugunsten eines autoradiographischen Nachweises verzichtet. Da sich radioaktiv markierte Nukleinsäure in subnanomolaren Konzentrationen nachweisen lässt, können somit auch sehr niedrige Dissoziationskonstanten gemessen werden. Desweiteren sollten definierte Komplexbanden sichtbar gemacht werden können, was aufgrund der Protein-Cofärbung im Silbergel schwierig ist.

Im Vergleich zu den "klassischen" Filterbindungstests bietet die Verzögerungsgelelektrophorese eine Reihe von Vorteilen. TFIIIA/Nukleinsäure-Komplexe haben eine Halbwertszeit von wenigen Minuten (Setzer *et al.*, 1996). Es besteht also die Gefahr, dass beim Waschen der Filtermembran der Komplex zerfällt und demnach eine zu schwache Bindung angenommen wird. Dagegen kann die Verzögerungsgelelektrophorese auch einer schnellen Dissoziationskinetik Rechnung tragen, da freie und gebundene Nukleinsäure beim Eintritt in die Gelmatrix sehr schnell separiert werden. Zerfallen die Komplexe während der elektrophoretischen Auftrennung, lässt sich die Konzentration der Reaktanden über die freie Nukleinsäure quantifizieren. Bei hinreichend stabilen Wechselwirkungen lassen sich distinkte Komplexbanden nachweisen, die interessante Hinweise auf den Beladungszustand der Nukleinsäure geben können. Um die Bindeplatzzahl in einem Filterbindungstest zu ermitteln, müssten zusätzliche Sättigungstitrationen mit hohem Bedarf an wertvollem Probenmaterial durchgeführt werden. Ferner lassen sich Degradationen der Nukleinsäure anhand der sehr sensitiven Autoradiographien detektieren. Somit kann leicht beurteilt werden, ob die Proteinpräparationen eventuell mit Nukleasen kontaminiert sind.

Gleichgewichte von freier und komplexierter Nukleinsäure lassen sich auch mittels analytischer Ultrazentrifugation (AUZ) studieren. Sehr affine Bindungen lassen sich mit dieser Methode gleichwohl nicht quantifizieren, da man bei sehr geringer Konzentration der eingesetzten Komponenten an die Nachweisgrenze der Absorptionsoptik stößt. Um starke Wechselwirkungen mittels AUZ nachzuweisen, muss die Zentrifuge mit einer Fluoreszenzoptik ausgestattet sein. Mit einem Gerät dieser Bauart konnten Goodman *et al.*  (1984)eine Dissoziationskonstante von 53 nM für die Bindung der PolII an das Viroid ermitteln.

#### 5.2.2 Nukleinsäurebindeeigenschaften von LeTFIIIA

In der vorliegenden Arbeit wurde für die Bindung von *Le*TFIIIA an seine Promotor-DNA eine hochspezifische Bindung im subnanomolaren Bereich beobachtet. Dagegen wurde unspezifische DNA deutlich schlechter gebunden. Das Protein wurde also in aktiver Form aus der rekombinanten Expression und Aufreinigung gewonnen. Da die RNA-Bindung von *Le*TFIIIA unter den identischen experimentellen Bedingungen gemessen wurde, ist es unklar, warum sich die Bindungsaffinitäten zu 5 S-RNA und zu anderen einzelsträngigen RNAs nicht unterschieden. Lediglich durchgängig doppelsträngige RNA, welche freilich nur bedingt physiologisch relevant ist, wird sehr viel schlechter gebunden.

Für die RNA-Bindung von rekombinantem TFIIIA aus *Xenopus laevis* wurden in der Literatur Dissoziationskonstanten in derselben Größenordnung wie für die TFIIIA/DNA-Bindung beschrieben (Romaniuk, 1985; Setzer *et al.*, 1996). Dies ist verwunderlich, tragen doch nur die Zink-Finger 4–6 zur RNA-Bindung bei. Der zusätzliche Energiebeitrag durch Bindung der übrigen Zink-Finger sollte sich eigentlich in einer affineren DNA-Wechselwirkung widerspiegeln.

Da sich die ermittelten Bindungsaffinitäten von *Le*TFIIIA an verschiedene RNAs nicht unterschieden, muss die Frage offen bleiben, ob TFIIIA eine funktionelle Rolle beim Transport des Viroids einnimmt. Das folgende Modell zur intrazellulären Passage von PSTVd basiert daher im wesentlichen auf der bemerkenswerten Ähnlichkeit von 5 S-RNA und PSTVd. Vermittels der ähnlichen Struktursignale sollte sich PSTVd die intrazellulären Transportwege der 5 S-RNA zu Nutze machen können.

#### Transport der 5S-RNA

Der intrazellulären Transport der 5 S-RNA in *Xenopus*-Oozyten wird von den Proteinen TFIIIA und L5 vermittelt (Pieler & Rudt, 1997). Die 5 S-RNA verlässt komplexiert mit TFIIIA den Zellkern, um im Cytoplasma als 7 S-Ribonucleoprotein (RNP) gespeichert zu werden. Auch das ribosomale Protein L5 bindet an die 5 S-RNA und bildet einen 5 S-RNP-Komplex (Rudt & Pieler, 1996). Einem regulatorischen Modell zufolge führt eine steigende L5-Produktion dazu, dass die 5 S-RNA aus dem Komplex mit TFIIIA



Abbildung 5.3: Modell für den intrazellulären Transport von PSTVd. Das Viroid verbreitet sich systemisch von Zelle zu Zelle über die Plasmodesmata. Dem Modell zufolge würde PSTVd im Cytoplasma ein Komplex mit TFIIIA eingehen, um so in den Nukleus zu gelangen. Dort wird es repliziert und zu reifen Zirkeln prozessiert. Ein Teil des zirkulären PSTVd kann nun von einem unbekannten Faktor X aktiv aus dem Zellkern transportiert und eine benachbarte Zelle befallen. Alternativ wird zirkuläres PSTVd im Nukleus von TFIIIA gebunden und in den Nukleolus geleitet, wo es akkumuliert.

verdrängt wird. Gemeinsam mit L5 wird sie dann als 5 S-RNP in den Kern transportiert und gelangt schließlich in den Nukleolus, wo sie in die große ribosomale Untereinheit eingebaut wird. Im Gegensatz zu *Xenopus*-Oozyten gibt es in somatischen Zellen von Mensch und *Arabidopsis* keine Hinweise auf Speichpartikel in Form von RNPs. Hingegen konnte bei *Arabidopsis* beobachtet werden, dass GFP-TFIIIA-Fusionsproteine bevorzugt im Nukleolus akkumulieren (Mathieu *et al.*, 2003). Die Autoren glauben daher, dass in der Pflanze Komplexe aus 5 S-RNA und TFIIIA als 7 S-RNP im Nukleolus gelagert werden.

#### Modell zum Transport von PSTVd

Das Viroid gelangt über ein spezifisches Transportmotiv über die Plasmodesmata in das Cytoplasma (siehe Abb. 5.3). Dort wird es von freiem TFIIIA komplexiert und in den Zellkern geschleust. Hier findet die Replikation inklusive der Prozessierung des Viroids statt. Innerhalb des Nukleus scheinen keine aktiven Transportprozesse zu existieren (Lewis & Tollervey, 2000; Phair & Misteli, 2000). Die intranukleäre Lokalisation von Faktoren wird über die transiente Bindung an funktionsgekoppelte Komponenten sichergestellt. Die Annahme, dass PSTVd im Nukleolus komplexiert mit TFIIIA vorliegt, wäre konsistent mit der Beobachtung, dass beide Partner im Nukleolus akkumulieren (Harders *et al.*, 1989; Mathieu *et al.*, 2003; Qi & Ding, 2003). Die Tatsache, dass Qi & Ding (2003) ferner kein (-)-stränges PSTVd im Nukleolus detektieren konnten, könnte so verstanden werden, dass der (-)-Strang kein Loop E-Motif bilden kann und daher nicht für eine Bindung an TFIIIA in Frage käme.

Alternativ könnten die multimeren (+)-Stränge, wie von Harders *et al.* (1989) diskutiert, im Nukleolus zu Monomeren geschnitten und ligiert werden. In diesem Zellorganell werden auch unreife rRNAs und tRNAs prozessiert (Lewis & Tollervey, 2000).

Um die nächste Zelle zu infizieren, müssen einige Viroidmoleküle den Kern verlassen. Dieser aktive Transport könnte vielleicht durch das ribosomale Protein L5 bewerkstelligt werden. Eine mögliche Beteiligung des ribosomalen Proteins L5 am intrazellulären Transport von PSTVd wurde bereits in anderen Arbeiten erwogen (Aschermann, 2003; Steitz *et al.*, 1988).

Ein weiteres Szenario für den intrazellulären Transport von PSTVd wäre eine direkte Beteiligung der 5 S-RNA. Aschermann (2003) konnte Komplexe aus PSTVd und 5 S-RNA nachweisen, eine genauere Eingrenzung der 5 S/PSTVd-Kontaktstelle ist Gegenstand aktueller Untersuchungen. Die 5 S-RNA könnte als ein Adaptermolekül verstanden werden, welches eine Kopplung von PSTVd an das Transportprotein TFIIIA oder L5 vermittelt. Somit würde sich das Viroid gewissermaßen als "blinder Passagier" der Transportmittel der 5 S-RNA bedienen, um seine korrekte subzelluläre Lokalisation sicherzustellen.

#### 5.2.3 Die Fehlleitung von Wirtsproteinen durch das Viroid

#### Molekulare Mimikry

Im Zusammenhang mit Viroiden wird oft von "molekularen Parasiten" gesprochen, da ihr Überleben an den Wirtsmetabolismus gekoppelt ist. Auch findet oftmals – in Analogie zum Parasitismus in höheren Eukaryoten – der Begriff einer "molekularen Mimikry" Verwendung: Das Viroid sollte wirtsähnliche Strukturen imitieren können, um Wirtsfaktoren zu binden und deren Funktionen zu usurpieren. Eine Mimikry kann jedoch immer nur als eine Reaktion auf einen äußeren Stimulus verstanden werden. Sie ist eine zeitlich verzögerte, sukzessive Anpassung des Phänotyps. Das Modell einer molekularen Mimikry berücksichtigt jedoch nicht die Möglichkeit, dass bestimmte Viroid- und RNA-Strukturen des Wirts aus Koevolution hervorgegangen sein können, und demnach *per definitionem* homolog sind. Der Begriff "Mimikry" trifft demnach nur für analoge Erscheinungsmerkmale zu, die sich gemäß der Darwinschen Theorie von Mutation und Selektion ausbildeten.

Vor dem Hintergrund der in Kapitel 5.1.2 dargelegten hohen Mutationsrate scheint es durchaus nachvollziehbar, dass sich das Viroid im Laufe der Evolution pflanzlichen RNAs anglich, um sich beispielsweise deren Transportwege zu bedienen. Doch wie vermag das Viroid ein DNA-abhängiges Enzym für seine Replikation fehlzuleiten? Auf den ersten Blick scheint dies nicht möglich zu sein, liegt doch typischerweise die DNA in einer Bförmigen Helixgeometrie vor, welche die RNA aufgrund von sterischen Einschränkungen durch ihre 2'OH-Gruppe gar nicht einnehmen kann. Jedoch existieren für die DNA strukturelle Alternativen zur B-Form. So konnte gezeigt werden, dass kurze GC-reiche DNA-Doppelstränge in der RNA-typischen A-Helix vorliegen können (Heinemann et al., 1987). Diese Tatsache, dass GC-reiche DNA der Struktur einer RNA ähnelt, führte zu der Annahme, dass DNA-bindende Proteine durchgängig basengepaarte, GC-reiche Abschnitte auf der Viroid RNA erkennen können; darunter fallen die sogenannten GC-Boxen auf dem (+)-Strang oder der Hairpin II auf dem (-)-Strang (Fels et al., 2001; Heinrich, 1991; Loss et al., 1991; Qu et al., 1993). Diese Hypothese muss jedoch im Folgenden einer Reevaluierung standhalten, da unterdessen neuere strukturelle Daten zu multifunktionellen Proteinen vorliegen, welche sowohl DNA als auch RNA binden können.

#### Transkriptionsfaktor IIIA: Nukleinsäurebindungseigenschaften

Der Prototyp eines DNA- und RNA-bindenden Proteins ist der Transkriptionsfaktor IIIA (TFIIIA), welcher auch in der vorliegenden Arbeit als möglicher Bindungspartner des Viroids untersucht wurde. Für seine natürliche, recht GC-reiche DNA-Promotorsequenz wurde anfänglich eine A-förmige Helixgeometrie nahegelegt (Rhodes & Klug, 1986). Diese Vermutung konnte kurze Zeit später durch röntgenkristallographische Analysen gestützt werden (McCall *et al.*, 1986). Somit glaubte man zu verstehen, wie TFIIIA gleichsam DNA und RNA erkennt: Die natürlichen DNA- und RNA-Bindepartner von TFIIIA ähneln sich

trotz ihrer chemischen Unterschiede auf struktureller Ebene. Also könnten sowohl DNA als auch RNA nach demselben Mechanismus gebunden werden. Jedoch zeigten Aboul-ela *et al.* (1988), dass die gemessene A-Form des Promotorfragments artifizieller Natur war. In Lösung lag dasselbe Fragment in der DNA-typischen B-Form vor.

Einsicht in den genauen Nukleinsäurebindungsmodus von TFIIIA erbrachten erst hochaufgelöste Strukturen von TFIIIA/DNA- (Nolte *et al.*, 1998) und TFIIIA/5S-RNA-Komplexen (Lu *et al.*, 2003; Vitolo *et al.*, 2004). Ein Vergleich der Strukturen offenbart, dass TFIIIA mechanistisch gänzlich anders an DNA als an RNA bindet. Im Falle der DNA-Bindung dringen die Zink-Finger-Bindemotive in die große Furche ein und können dort basenspezifische Wechselwirkungen etablieren. Auch können die Zinkfinger entlang des Phosphatrückgrats positioniert werden. Dort dienen sie lediglich dazu, die notwendigen Abstände zu gewährleisten. Dagegen wird die 5S-RNA über ihre Tertiärstruktur spezifisch komplexiert. Für die Erkennung essentiell sind die hochstrukturierten Motive Loop A und Loop E, welche mit den Zink-Finger-"Spitzen" 4 und 6 interagieren.

#### DNA- und RNA-bindende Transkriptionsfaktoren

Neben TFIIIA gibt es weitere Transkriptionsfakoren, die befähigt sind, sowohl DNA als auch RNA zu binden. Darunter fallen TRA-1, STAT-1, *bicoid* und der überaus prominente Faktor p53 (zur Übersicht: Cassiday & Maher, 2002). Ferner existieren Arbeiten, in denen mittels *in vitro*-Selektionsverfahren RNA-Aptamere syntetisiert wurden, die hochspezifisch an Transkriptionsfaktoren binden. Für einen Komplex eines RNA-Aptamers mit dem p50-Homodimer von NF- $\kappa$ B wurde die Struktur durch Röntgenkristallographie aufgeklärt (Huang *et al.*, 2005). Die RNA nimmt eine haarnadelförmige Tetraloop-Struktur an, die zentral ungepaarte Basen präsentiert. Es konnte gezeigt werden, dass das Aptamer verglichen mit der DNA dieselbe Bindungsschnittstelle im Protein rekrutiert. Obwohl die DNA in der typischen B-Form und das Aptamer in einer komplexen Tertiärstruktur gebunden werden, ähneln sich Ladungsmuster und H-Brückenkonformation an der Wechselwirkungsstelle im Protein weitgehend.

Auch für das Tata Binding Protein (TBP) konnten bindungsaktive RNA-Aptamere synthetisiert werden (Fan *et al.*, 2004). Diese konkurrieren mit der natürlichen DNA-Tata-Sequenz um die Bindung an das TBP. Im hiesigen Institut konnte Gande (2002) eine recht affine Bindung von humanem TBP an Viroid-RNA messen. Aus Materialmangel konnte dieses Ergebnis indes nicht reproduziert und mit den notwendigen Kontrollversuchen belegt werden. Da sich in der vorliegenden Arbeit das TBP aus Kartoffel als toxisch für *E. coli* erwies, konnten keine weiteren Daten zu den RNA-Bindeeigenschaften des TBP erhalten werden. Somit bleibt auch die Frage nach den Promotorelementen auf der Viroid-RNA, die das Nukleotid U359 als Startpunkt der Transkription festlegen, ungeklärt. Als mögliche Promotorelemente auf dem Viroid wurden bislang GC-reiche, basengepaarte Abschnitte diskutiert. Jedoch gibt es in der Literatur keine Beispiele, in denen gezeigt wird, dass ein Protein dieselbe Helixgeometrie auf RNA- und DNA-Ebene erkennen kann. Basierend auf den bisher bekannten Bindungsmechanismen von DNA- und RNA-bindenden Proteinen scheinen die folgenden zwei Szenarien für die Rekrutierung DNA-bindender Proteine auf die Viroid-RNA plausibler:

- Das Viroid offeriert Strukturen, mit denen es die drei-dimensionale Anordnung von Ladungen und von Wasserstoffbrückendonoren- und akzeptoren einer DNA im Proteinkomplex mimikriert.
- Das Viroid interagiert über seine hochstrukturierten Bereiche, wie z.B. den Loop E, mit dem Protein. Dabei belegt es aber andere Wechselwirkungspositionen im Protein verglichen zu DNA.

### 5.3 Die Alternative Spleißform der TFIIIA-mRNA

In der vorliegenden Arbeit konnte gezeigt werden, dass die TFIIIA-mRNA in einer alternativ gespleißten Variante (TFIIIA-mRNA<sup>alt</sup>) existiert. Dies wurde experimentell für die Organismen *Arabidopsis thaliana* und *Lycopersicon esculentum* und *in silico* für *Oryza sativa* belegt. Der Mechanismus dieses alternativen Spleißens scheint in allen drei untersuchten Spezies konserviert zu sein: Die TFIIIA-mRNA<sup>alt</sup> enthält ein optionales Exon 3 mit einem vorzeitigem UGA-Stopp-Codon, welches sich auf Proteinebene unmittelbar hinter dem zweiten Zink-Finger-Motiv befindet. Bei der "normal"-gepleißten Variante, welche für die TFIIIA-Volllänge kodiert, wird dieses Exon 3 bei der Prozessierung der prä-mRNA entfernt.

Aufgrund dieser signifikanten Übereinstimmung selbst zwischen Mono- und Dikotylen wurde eine funktionelle Bedeutung der TFIIIA-mRNA<sup>alt</sup> postuliert. Durch *in vitro*-Translation konnte gezeigt werden, dass die TFIIIA-mRNA<sup>alt</sup> in eine verkürzte TFIIIA- Isoform (kTFIIIA), bestehend aus den ersten beiden Zink-Fingern, übersetzt wird. Unklar ist die Beobachtung, dass die TFIIIA-mRNA<sup>alt</sup> mit 5'-Cap sehr viel schlechter als ohne Modifikation translatiert wurde. Dies könnte als Hinweis gedeutet werden, dass im Weizenkeimextrakt die Translation von natürlichen mRNAs mit einem vorzeitigem Stopp-Codon inhibiert wird.

Rekombinantes kTFIIIA zeigte indes keine Nukleinsäurebindung im mikromolaren Konzentrationsbereich. *In vivo* konnte die verkürzte Variante, ebenso wie die Volllänge nicht *per* Westernblot detektiert werden. Dies könnte darin begründet sein, dass die TFIIIA-Konzentration in der Zelle zu gering ist. Im Gegensatz zu *Xenopus*-Oozyten, die 10<sup>12</sup> TFIIIA-Kopien beherbergen (Ginsberg *et al.*, 1984), ist der TFIIIA-Gehalt in somatischen Zellen eher gering. In *A. castellanii* und HeLa-Zellen wurden lediglich 170-400 Moleküle gezählt (Moorefield & Roeder, 1994; Polakowski & Paule, 2002).

Durch alternatives Spleißen vermag die Zelle ihre Genexpression auf posttranskriptioneller Ebene zu steuern (zur Übersicht: Stamm *et al.*, 2005). Ein Beispiel repräsentiert das zirkadian regulierte, RNA-bindende Protein AtGRP7 aus *Arabidopsis thaliana* (Staiger *et al.*, 2003). Durch Wechselwirkung mit der eigenen prä-mRNA aktiviert AtGRP7 eine kryptische Spleißstelle, wodurch die mRNA alternativ prozessiert wird. Die alternativ gepleißte AtGRP7-mRNA wird nicht exprimiert und schnell degradiert. Dieses Ergebnis steht im guten Einklang mit dem Befund, dass AtGRP7 seine eigene Expression über einen negativen Rückkopplungsmechanismus reguliert (Heintzen *et al.*, 1997).

Die eukaryotische Zelle ist in der Lage, gespleißte mRNA einer Qualitätskontrolle zu unterziehen. In einem als nonsense mediated decay (NMD) bezeichneten Prozess werden abnormale messenger, die ein vorzeitiges Stopp-Codon enthalten, enzymatisch abgebaut (zur Übersicht: Conti & Izaurralde, 2005). Es gibt starke Hinweise auf analoge Degradationsmechanismen in planta (Petracek et al., 2000; van Hoof & Green, 1996). Die Natur der meisten beteiligten Komponenten ist bislang aber weitgehend unbekannt. Kürzlich konnten für die NMD-involvierten UPF-Proteine die entsprechenden Homologe aus Arabidopsis beschrieben werden (Arciga-Reyes et al., 2006; Hori & Watanabe, 2005). Die in der vorliegenden Arbeit beobachtete Abundanz der TFIIIA-mRNA<sup>alt</sup> zieht die Frage nach sich, warum diese aufgrund ihres frühzeitigen Stopp-Codons nicht der zellulären Degradationsmaschinerie zugeführt wird. Möglicherweise ist TFIIIA-mRNA<sup>alt</sup> vorwiegend im Nukleus lokalisiert. Sie wäre in diesem Kompartiment vor dem cytoplasmatischen Abbau via NMD geschützt und ließe sich folglich mittels RT-PCR detektieren. Dennoch besteht die Möglichkeit, dass TFIIIA-mRNA<sup>alt</sup> nach Transport in das Cytosol degradiert wird und nicht translatiert wird. Tatsächlich weisen neuste Ergebnisse darauf hin, dass TFIIIAmRNA<sup>alt</sup> offenbar vom NMD-Mechanismus betroffen ist. In *Arbabidopsis*-Mutanten, die NMD-defizient sind, liegt TFIIIA-mRNA<sup>alt</sup> höher konzentriert vor als in Wildtyppflanzen (Yoine *et al.*, 2006).

Ein Vergleich von EST- und cDNA-Sequenzen mit den Genomen von Arabidopsis und Reis zeigt, dass ungefähr ein Fünftel aller pflanzlichen Gene alternativ gespleißt wird (Wang & Brendel, 2006). Da hiervon wiederum ein Viertel nach einem konservierten Mechanismus prozessiert wird, lässt sich nahelegen, dass diese mRNA-Varianten funktionell bedeutend sind.

Welche Rolle die TFIIIA-mRNA<sup>alt</sup> im biologischen System spielt, muss in künftigen Studien eruiert werden. Denkbar ist eine Funktion bei der posttranskriptionellen Genregulation, aber auch ihre kodogenen Eigenschaften dürfen nicht außer Acht gelassen werden. Aufgrund ihrer signifikanten Konservierung innerhalb der Pflanzenwelt scheint es indes äußerst unwahrscheinlich, dass es sich bei der TFIIIA-mRNA<sup>alt</sup> um einen bloßen Fehler der Natur handelt.

## 5.4 Ausblick

Mit dem Viroid steht der Wissenschaft seit über 25 Jahren ein ausgezeichnetes Modellsystem zur Verfügung, das es erlaubt, die unterschiedlichsten biologischen Thematiken zu behandeln. Wie kann z. B. RNA von einer DNA-abhängigen Polymerase transkribiert werden? Oder wie wird RNA in der Pflanze transportiert? Was sind die strukturellen Voraussetzungen von RNA-*interference*? Die Beantwortung dieser Fragen ist von großem Allgemeininteresse, das weit über die eigentliche Aufklärung der Funktionsweise des Viroids hinausgeht.

Die vorliegende Arbeit begann mit der Frage nach den Wechselwirkungspartnern des Viroids und führte schließlich ganz unerwartet zu neuen Erkenntnissen bezüglich einer Fragestellung von großer Bedeutung: Ist alternatives Spleißen bei eukaryotischen Transkriptionsfaktoren ein allgemeines Regulationsprinzip? Es gelang die Identifikation und nachfolgende Charakterisierung des Tomaten-Homologs von TFIIIA – eines der wichtigsten Transkriptionsfaktoren aller eukaryotischen Organismen. So konnte gezeigt werden, dass ein genau definierter, alternativer Spleißmechanismus der mRNA dieses Proteins innerhalb der höheren Pflanzen konserviert ist und damit ganz offenbar eine wichtige Funktion erfüllt. Auch für menschliches TFIIIA konnte eine alternativ gespleißte mRNA in der Datenbank gefunden werden. Mithilfe weiterer bioinformatischer Analysen sollte geklärt werden können, ob alternative Spleißmechanismen von TFIIIA auch im tierischen System konserviert sind. Von großer Wichtigkeit wird es sein, die funktionelle Bedeutung der alternativ gespleißten TFIIIA-Variante aufzuklären. Hierbei gilt es insbesondere die Frage zu klären, ob diese ribosomal translatiert wird oder aufgrund ihres vorzeitigen Stopp-Codons lediglich als Intermediat innerhalb des zellulären Degradationsmechanismus zu verstehen ist. Somit könnten sich ganz neue Einblicke in das regulatorische Netzwerk der 5 S-RNA-Synthese erschließen.

# 6

# Zusammenfassung

Viroide sind nicht-kodogene RNA-Moleküle, die in geeigneten Wirtspflanzen parasitieren. Vermittels definierter Sekundärstrukturen wechselwirkt die Viroid-RNA mit Wirtsfaktoren um in der Pflanze repliziert und transportiert zu werden. Auf struktureller Basis kann für einige Wirtsproteine eine Interaktion mit dem Viroid postuliert werden. In der vorliegenden Arbeit sollten solche hypothetischen Wechselwirkungspartner rekombinant hergestellt und ihre Bindung an das Potato Spindle Tuber Viroid (PSTVd) charakterisiert werden. Anhand der ermittelten Daten sollten Modellsysteme zu Replikation und Transport von PSTVd überprüft werden.

In einer vorausgehenden Arbeit wurde gezeigt, dass die Transkription am linken terminalen Loop der nativen Viroid-Sekundärstruktur (Position 1C oder 359U) beginnt. (Kolonko *et al.*, 2003). In der vorliegenden Arbeit konnten Infektiositätsstudien mit punktmutierten PSTVd-Molekülen zeigen, dass die Mutation G1 genetisch stabil ist, während G359 zu U359 revertiert. Nach NMR-Analysen ist das Nukleotid U359 für die Struktur des linken Tetraloops ohne Bedeutung (Scholtysik, 2006). Daher muss also nicht etwa ein Strukturmotiv, sondern die reine Sequenz essentiell für den Transkriptionsstart sein.

Da die Transkription von PSTVd durch die Polymerase II an einer definierten Position beginnt, liegt es nahe, dass diese von einem Promotor auf der Viroid-RNA determiniert wird. Das Viroid könnte DNA-Strukturen imitieren und so den basalen Faktor TATA-Binding-Protein (TBP) für die korrekte Transkriptionsinitiation rekrutieren. Das stark konservierte TBP konnte bereits aus einigen Organismen als GST-Fusionsprotein rekombinant hergestellt werden. In der vorliegenden Arbeit wurde bemerkenswerterweise beobachtet, dass TBP aus der Wirtspflanze *Solanum tuberosum* aufgrund seiner Toxizität für eine rekombinante Expression in *E. coli* nicht zugänglich war.

Der Transkriptionsfaktor IIIA (TFIIIA) und das ribosomale Protein L5 regulieren die 5S-RNA-Synthese und stehen in direktem Zusammenhang mit dem intrazellulären 5S-RNA-Transport. Da PSTVd über 5S-ähnliche Strukturmotive verfügt, könnte es sich dieser Proteine für seine eigene zelluläre Lokalisation bedienen. Die Sequenz von TFIIIA aus der Wirtspflanze *Lycopersicon esculentum* wurde in dieser Arbeit aufgeklärt, und es wurde ein Aufreinigungsverfahren für die in *E. coli* unlöslich exprimierten Proteine GST-TFIIIA und GST-L5 etabliert. GST-L5 zeigte nach Reinigung lediglich unspezifische RNA-Bindung. Für GST-TFIIIA konnte eine hochspezifische Bindung an 5S-DNA beobachtet werden. 5S-RNA, PSTVd und andere einzelsträngige RNA wurde um mehr als eine Größenordnung schwächer gebunden. Es wurde ein Modell formuliert, wonach PSTVd über sein Loop E-Motiv mit TFIIIA wechselwirkt und somit im Nukleolus akkumulieren kann.

Bei der Klonierung der TFIIIA-Sequenzen wurde sowohl von *Lycopersicon*- als auch von *Arabidopsis*-cDNA eine alternativ gespleißte TFIIIA-mRNA amplifiziert. Durch Sequenzanalyse und *in-vitro*-Translation konnte gezeigt werden, dass dieses optionale Genprodukt nur aus den ersten beiden Zink-Finger-Bindemotiven von TFIIIA besteht. Diese verkürzte TFIIIA-Variante ließ sich per Westernblot-Analyse nicht *in vivo* nachweisen und zeigte in rekombinanter Form keine Nukleinsäurebindeaktivität. Aufgrund seiner Konservierung in höheren Pflanzen wurde dem alternativen Spleißmechanismus des TFIIIA-Gens eine regulatorische Funktion zugesprochen.

# Literaturverzeichnis

- Aboul-ela, F., Varani, G., Walker, G.T. & Tinoco, I. (1988). The TFIIIA recognition fragment d(GGATGGGAG)·d(CTCCCATCC) is B-form in solution. Nucleic Acids Res., 16(8), 3559–3572. 104
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol., 215(3), 403–410. 55
- Arciga-Reyes, L., Wootton, L., Kieffer, M. & Davies, B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. *Plant J.*, 47(3), 480–489. 106
- Aschermann, K. (2003). Identifikation eines zellulären RNA-Wechselwirkungspartners von PSTVd in vitro und in vivo. Dissertation, Heinrich-Heine-Universität Düsseldorf. 8, 9, 102
- Baumstark, T. & Riesner, D. (1995). Only one of four possible secondary structures of the central conserved region of potato spindle tuber viroid is a substrate for processing in a potato nuclear extract. *Nucleic Acids Res.*, 23, 4246–4254. 7
- Baumstark, T. & Riesner, D. (1995). Only one of four possible secondary structures of the central conserved region of potato spindle tuber viroid is a substrate for processing in a potato nuclear extract. *Nucleic Acids Res.*, 23, 4246–4254. 95
- Baumstark, T., Schröder, A.R. & Riesner, D. (1997). Viroid processing: switch from cleavage to ligation is driven by a change from a tetraloop to a loop E conformation. *EMBO J.*, 16, 599–610. 7, 95
- Branch, A. D. & Robertson, H. D. (1981). Longer-than-unit length viroid minus strands are present in RNA from infected plants. *Proc. Nat. Acad. Sci. U.S.A.*, 78, 6381–6385. 5
- Branch, A. D. & Robertson, H. D. (1984). A replication cycle for viroids and other small infectious RNA's. Science, 223, 450–455. 5

- Carey, M. & Smale, S.T., Hrsg. (2002). Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques. Cold Spring Harbor Laboratory Press, New York. 74
- Cassiday, L.A. & Maher, L.J. (2002). Having it both ways: transcription factors that bind DNA and RNA. *Nucleic Acids Res.*, **30**(19), 4118–4126. **104**
- Chester, N. & Marshak, D.R. (1993). Dimethyl sulfoxide-mediated primer  $T_m$  reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Anal. Biochem.*, **209**(2), 284–290. 28
- Conti, E. & Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.*, **17**(3), 316–325. 84, 106
- Daròs, J.A., Elena, S.F. & Flores, R. (2006). Viroids: an Ariadne's thread into the RNA labyrinth. *EMBO Rep.*, 7(6), 593–598. 5
- Daròs, J.A. & Flores, R. (2002). A chloroplast protein binds a viroid RNA in vivo and facilitates its hammerhead-mediated self-cleavage. *EMBO J.*, **21**(4), 749–759. 5, 9
- Daròs, J.A. & Flores, R. (2004). Arabidopsis thaliana has the enzymatic machinery for replicating representative viroid species of the family Pospiviroidae. Proc. Nat. Acad. Sci. U.S.A., 101(17), 6792–6797. 55
- Daròs, J. A., Marcos, J. F., Hernández, C. & Flores, R. (1994). Replication of avocado sunblotch viroid: evidence for a symmetric pathway with two rolling circles and hammerhead ribozyme processing. *Proc. Nat. Acad. Sci. U.S.A.*, **91**, 12813–12817. 5
- Del Río, S. & Setzer, D.R. (1991). High yield purification of active transcription factor IIIA expressed in *E. coli. Nucleic Acids Res.*, **19**(22), 6197–6203. 66
- Diener, TO. (2001). The viroid: biological oddity or evolutionary fossil? Adv. Virus. Res., 57, 137–184. 9
- Diener, T.O. (2003). Discovering viroids-a personal perspective. Nat. Rev. Microbiol., 1(1), 75–80. 2
- Ding, B., Itaya, A. & Zhong, X. (2005). Viroid trafficking: a small RNA makes a big move. Curr. Opin. Plant Biol., 8(6), 606–612. 7
- Engelke, D.R., Ng, S.Y., Shastry, B.S. & Roeder, R.G. (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*, 19(3), 717–728. 10, 82

- Fan, X., Shi, H., Adelman, K. & Lis, J.T. (2004). Probing TBP interactions in transcription initiation and reinitiation with RNA aptamers that act in distinct modes. *Proc. Nat. Acad. Sci. U.S.A.*, **101**(18), 6934–6939. **104**
- Feldstein, P. A., Hu, Y. & Owens, R. A. (1998). Precisely full length, circularizable, complementary RNA: an infectious form of potato spindle tuber viroid. *Proc. Nat. Acad. Sci. U.S.A.*, **95**, 6560–6565. **5**, **50**
- Fels, A, Hu, K & Riesner, D (2001). Transcription of potato spindle tuber viroid by RNA polymerase II starts predominantly at two specific sites. *Nucleic Acids Res*, 29(22), 4589–4597. 96, 103
- Frangioni, J.V. & Neel, B.G. (1993). Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal. Biochem., 210(1), 179–187. 23, 67
- Gande, R. (2002). Strukturelle Analyse der Promotor-ähnlichen Elemente im Kartoffelspindelknollensuchtviroid. Diplomarbeit, Heinrich-Heine-Universität Düsseldorf. 10, 104
- Ginsberg, A.M., King, B.O. & Roeder, R.G. (1984). Xenopus 5S gene transcription factor, TFIIIA: characterization of a cDNA clone and measurement of RNA levels throughout development. Cell, 39(3 Pt 2), 479–489. 106
- Gómez, G. & Pallás, V. (2001). Identification of an in vitro ribonucleoprotein complex between a viroid RNA and a phloem protein from cucumber plants. *Mol. Plant. Microbe Interact.*, 14(7), 910–913. 8
- Gómez, G. & Pallás, V. (2004). A long-distance translocatable phloem protein from cucumber forms a ribonucleoprotein complex in vivo with Hop stunt viroid RNA. J. Virol., 78(18), 10104–10110. 8, 9, 98
- Goodman, T.C., Nagel, L., Rappold, W., Klotz, G. & Riesner, D. (1984). Viroid replication: equilibrium association constant and comparative activity measurements for the viroid-polymerase interactions. *Nucleic Acids Res.*, 12, 6231–6246. 99
- Gozmanova, M., Denti, M.A., Minkov, I.N., Tsagris, M. & Tabler, M. (2003). Characterization of the RNA motif responsible for the specific interaction of potato spindle tuber viroid RNA (PSTVd) and the tomato protein Virp1. Nucleic Acids Res., 31(19), 5534–5543. 9

- Guddat, U., Bakken, A.H. & Pieler, T. (1990). Protein-mediated nuclear export of RNA:
  5 S rRNA containing small RNPs in *Xenopus* oocytes. *Cell*, 60(4), 619–628. 11
- Harders, J., Lukács, N., Robert-Nicoud, M., Jovin, T.M. & Riesner, D. (1989). Imaging of viroids in nuclei from tomato leaf tissue by *in situ* hybridization and confocal laser scanning microscopy. *EMBO J.*, 8, 3941–3949. 8, 102
- Heinemann, U., Lauble, H., Frank, R. & Blöcker, H. (1987). Crystal structure analysis of an A-DNA fragment at 1.8 Åresolution: d(GCCCGGGC). Nucleic Acids Res., 15, 9531–9550. 7, 103
- Heinrich, C. (1991). Basenpaaraustausche mit gerichteter Mutagenese in der konservierten Region der Haarnadelstruktur II des Potato Spindle Tuber Viroid (PSTVd). Diplomarbeit, Heinrich-Heine-Universität Düsseldorf. 5, 103
- Heintzen, C., Nater, M., Apel, K. & Staiger, D. (1997). AtGRP7, a nuclear
  RNA-binding protein as a component of a circadian-regulated negative feedback
  loop in Arabidopsis thaliana. Proc. Nat. Acad. Sci. U.S.A., 94(16), 8515–8520. 106
- Hiddinga, H. J., Crum, C. J., Hu, J. & Roth, D. A. (1988). Viroid-induced phosphorylation of a host protein related to a dsRNA-dependent protein kinase. *Science*, 241, 451–453. 9, 98
- Holdsworth, M.J., Grierson, C., Schuch, W. & Bevan, M. (1992). DNA-binding properties of cloned TATA-binding protein from potato tubers. *Plant. Mol. Biol.*, 19(3), 455–464. 54, 63
- Hori, K. & Watanabe, Y. (2005). UPF3 suppresses aberrant spliced mRNA in Arabidopsis. *Plant J.*, 43(4), 530–540. 106
- Hu, K. (2000). Wechselwirkung beider Startstellen der Transkription von mutierten zirkulären Viroid-Molekülen. Dissertation, Heinrich-Heine-Universität Düsseldorf.
   97
- Huang, D.B., Phelps, C.B., Fusco, A.J. & Ghosh, G. (2005). Crystal structure of a free kappaB DNA: insights into DNA recognition by transcription factor NF-kappaB. J. Mol. Biol., 346(1), 147–160. 104
- Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. (1986). Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res.*, 14, 3627–3640. 5

- Jackson, R.J. (2005). Alternative mechanisms of initiating translation of mammalian mRNAs. Biochem. Soc. Trans., 33(Pt 6), 1231–1241. 85
- Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, **30**(14), 3059–3066. **57**, 58
- Keese, P. & Symons, R. H. (1985). Domains in viroids: Evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. Proc. Nat. Acad. Sci. U.S.A., 82, 4582–4586. 3
- Klaff, P., Gruner, R., Hecker, R., Sättler, A., G., Thissen & D., Riesner (1989). Reconstituted and cellular viroid-protein complexes. J. gen. Virol., 70, 2257–2270. 9, 98
- Klümper, S. (2002). Prozessierung des Kartoffel-Spindelknollensucht-Viroids (PSTVd): Charakterisierung der beteiligten Enzyme der Wirtspflanze. Dissertation, Heinrich-Heine-Universität Düsseldorf. 5
- Kolonko, N. (2003). Isolation und Bestimmung des 5'-Endes der
  (-)-Repklikationsintermediären des potato spindle tuber viroids (PSTVd).
  Dissertation, Heinrich-Heine-Universität Düsseldorf. 5, 10, 12, 49, 93, 95
- Kolonko, N., Bannach, O., Aschermann, K., Hu, K.H., Moors, M., Schmitz, M.,
  Steger, G. & Riesner, D. (2006). Transcription of potato spindle tuber viroid by
  RNA polymerase II starts in the left terminal loop. *Virology*, 347(2), 392–404. 96
- Kozak, M. (2005). A second look at cellular mRNA sequences said to function as internal ribosome entry sites. Nucleic Acids Res., 33(20), 6593–6602. 85
- Laemmli, U.K., Beguin, F. & Gujer-Kellenberger, G. (1970). A factor preventing the major head protein of bacteriophage T4 from random aggregation. J. Mol. Biol., 47(1), 69–85. 44
- Lejeune, F. & Maquat, L.E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr. Opin. Cell Biol.*, 17(3), 309–315. 84
- Lewis, J.D. & Tollervey, D. (2000). Like attracts like: getting RNA processing together in the nucleus. Science, 288(5470), 1385–1389. 102
- Loss, P., Schmitz, M., Steger, G. & Riesner, D. (1991). Formation of a thermodynamically metastable structure containing hairpin II is critical for

infectivity of potato spindle tuber viroid RNA. *EMBO J.*, **10**, 719–727. **4**, **5**, **6**, **95**, **103** 

- Lu, D., Searles, M.A. & Klug, A. (2003). Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature*, **426**(6962), 96–100. **11**, 104
- Maniataki, E., Martinez de Alba, A.E., Sägesser, R., Tabler, M. & Tsagris, M. (2003). Viroid RNA systemic spread may depend on the interaction of a 71-nucleotide bulged hairpin with the host protein VirP1. RNA, 9(3), 346–354. 98
- Martínez de Alba, A.E., Sägesser, R., Tabler, M. & Tsagris, M. (2003). A bromodomain-containing protein from tomato specifically binds potato spindle tuber viroid RNA *in vitro* and *in vivo*. J. Virol., **77**(17), 9685–9694. 7, 98
- Mathieu, O., Jasencakova, Z., Vaillant, I., Gendrel, A.V., Colot, V., Schubert, I. & Tourmente, S. (2003). Changes in 5 S rDNA chromatin organization and transcription during heterochromatin establishment in *Arabidopsis. Plant Cell*, 15(12), 2929–2939. 55, 59, 63, 65, 69, 70, 74, 76, 78, 79, 80, 101, 102
- McCall, M., Brown, T., Hunter, W.N. & Kennard, O. (1986). The crystal structure of d(GGATGGGAG): an essential part of the binding site for transcription factor IIIA. *Nature*, **322**(6080), 661–664. 103
- Michael, W.M. & Dreyfuss, G. (1996). Distinct domains in ribosomal protein L5 mediate 5 S rRNA binding and nucleolar localization. J. Biol. Chem., 271(19), 11571–11574. 11, 80
- Milligan, J.F., Groebe, D.R., Witherell, G.W. & Uhlenbeck, O.C. (1987).
  Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, 15(21), 8783–8798. 29
- Moorefield, B. & Roeder, R.G. (1994). Purification and characterization of human transcription factor IIIA. J. Biol. Chem., 269(33), 20857–20865. 106
- Moors, M. (2004). Mutagenesestudien zur Funktion potentieller Startstellen der Viroidreplikation. Diplomarbeit, Heinrich-Heine-Universität Düsseldorf. 97
- Mühlbach, H.P. & Sänger, H.L. (1979). Viroid replication is inhibited by  $\alpha$ -amanitin. Nature, **278**, 185–188. **4**, **9**, **98**
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.*, **51 Pt 1**, 263–273. 27

- Nagel-Steger, L. (1990). Untersuchungen zur Replikation des Potato Spindle Tuber Viroids durch die DNA-abhängige RNA-Polymerase II aus Wirtsgewebe. Dissertation, Heinrich-Heine-Universität Düsseldorf. 96
- Navarro, J.A. & Flores, R. (2000). Characterization of the initiation sites of both polarity strands of a viroid RNA reveals a motif conserved in sequenz and structure. *EMBO J.*, **19**, 2662–2670. **95**
- Navarro, J.A., Vera, A. & Flores, R. (2000). A chloroplastic RNA polymerase resistant to tagetitoxin is involved in replication of avocado sunblotch viroid. *Virology*, 268(1), 218–225. 9
- Nesser, N.K., Peterson, D.O. & Hawley, D.K. (2006). RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. Proc. Nat. Acad. Sci. U.S.A., 103(9), 3268–3273. 97
- Nolte, R.T., Conlin, R.M., Harrison, S.C. & Brown, R.S. (1998). Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. *Proc. Nat. Acad. Sci. U.S.A.*, **95**(6), 2938–2943. 104
- Nyborg, J.K. & Peersen, O.B. (2004). That zincing feeling: the effects of EDTA on the behaviour of zinc-binding transcriptional regulators. *Biochem. J.*, **381**(Pt 3), e3–e4. 72
- Owens, R.A., Blackburn, M. & Ding, B. (2001). Possible involvement of the phloem lectin in long-distance viroid movement. Mol. Plant. Microbe. Interact., 14(7), 905–909. 8
- Owens, R.A., Chen, W., Hu, Y. & Hsu, Y.H. (1995). Suppression of potato spindle tuber viroid replication and symptom expression by mutations which stabilize the pathogenicity domain. *Virology*, **208**, 554–564. **19**, **31**, **50**
- Palukaitis, P. (1987). Potato spindle tuber viroid: Investigations of the long-distance, intra-plant transport route. Virology, 158, 101–107. 8
- Papaefthimiou, I., Hamilton, A., Denti, M., Baulcombe, D., Tsagris, M. & Tabler, M. (2001). Replicating potato spindle tuber viroid RNA is accompanied by short RNA fragments that are characteristic of post-transcriptional gene silencing. *Nucleic Acids Res.*, **29**(11), 2395–2400. 9

- Pelchat, M., Coté, F. & Perreault, J.P. (2001). Study of the polymerization step of the rolling circle replication of peach latent mosaic viroid. Arch. Virol., 146(9), 1753–1763. 4
- Pelchat, M., Grenier, C. & Perreault, J.P. (2002). Characterization of a viroid-derived RNA promoter for the DNA-dependent RNA polymerase from Escherichia coli. *Biochemistry*, 41(20), 6561–6571. 5, 95
- Pelham, H.R. & Brown, D.D. (1980). A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc. Nat. Acad. Sci. U.S.A., 77(7), 4170–4174. 10
- Petracek, M.E., Nuygen, T., Thompson, W.F. & Dickey, L.F. (2000). Premature termination codons destabilize ferredoxin-1 mRNA when ferredoxin-1 is translated. *Plant J.*, **21**(6), 563–569. 106
- Phair, R.D. & Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. Nature, 404(6778), 604–609. 102
- Pieler, T. & Rudt, F. (1997). Nucleocytoplasmic transport of 5S ribosomal RNA. Semin. Cell Dev. Biol., 8(1), 79–82. 11, 100
- Polakowski, N. & Paule, M.R. (2002). Purification and characterization of transcription factor IIIA from Acanthamoeba castellanii. Nucleic Acids Res., 30(9), 1977–1984. 106
- Qi, Y. & Ding, B. (2003). Differential subnuclear localization of RNA strands of opposite polarity derived from an autonomously replicating viroid. *Plant Cell*, 15(11), 2566–2577. 102
- Qu, F., Heinrich, C., Loss, P., Steger, G., Tien, P. & Riesner, D. (1993). Multiple pathways of reversion in viroids for conservation of structural elements. *EMBO J.*, 12, 2129–2139. 5, 6, 95, 103
- Repsilber, D., Wiese, U., Rachen, M., Schröder, A.R., Riesner, D. & Steger, G. (1999).
  Formation of metastable RNA structures by sequential folding during transcription: Time-resolved structural analysis of potato spindle tuber viroid (-)-stranded RNA by temperature-gradient gel electrophoresis. *RNA*, 5, 574–584.
- Rhodes, D. & Klug, A. (1986). An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA. *Cell*, 46(1), 123–132. 103

Riesner, D. & Gross, H. J. (1985). Viroids. Ann. Rev. Biochem, 54, 531–564. 2

- Romaniuk, P.J. (1985). Characterization of the RNA binding properties of transcription factor IIIA of Xenopus laevis oocytes. Nucleic Acids Res., 13(14), 5369–5387. 79, 80, 100
- Romaniuk, P.J. (1990). Characterization of the equilibrium binding of Xenopus transcription factor IIIA to the 5 S RNA gene. J. Biol. Chem., 265(29), 17593–17600. 80
- Rowland, O. & Segall, J. (1996). Interaction of wild-type and truncated forms of transcription factor IIIA from *Saccharomyces cerevisiae* with the 5 S RNA gene. J. Biol. Chem., 271(20), 12103–12110. 80
- Rudt, F. & Pieler, T. (1996). Cytoplasmic retention and nuclear import of 5S ribosomal RNA containing RNPs. *EMBO J.*, **15**(6), 1383–1391. **80**, 100
- Sambrook, J., Fritsch, E. F. & Maniatis, T., Hrsg. (1989). Molecular Cloning. Cold Spring Harbor Laboratory Press. 26, 36
- Sammons, D.W., Adams, L.D. & Nishizawa, E.E. (1981). Ultrasensitive silver-based colour staining of polypeptids in polyacrylamid gels. *Electrophoresis*, 2, 135–141. 44
- Schmitz, A. (2003). Untersuchungen zum Pathogenitätsmechanismus von Viroid-RNA. Dissertation, Heinrich-Heine-Universität Düsseldorf. 8, 9, 71, 78, 98
- Schmitz, A. & Riesner, D. (2006). Purification of nucleic acids by selective precipitation with polyethylene glycol 6000. Anal. Biochem., 354(2), 311–313. 26, 36
- Schmitz, M. & Steger, G. (1996). Description of RNA folding by "simulated annealing". J. Mol. Biol., 255, 254–266. 6
- Scholtysik, R. (2006). Struktur der terminalen loops des Viroids PSTVd. Diplomarbeit, Heinrich-Heine-Universität Düsseldorf. 93, 94, 95
- Schrader, O, Baumstark, T & Riesner, D (2003). A mini-RNA containing the tetraloop, wobble-pair and loop E motifs of the central conserved region of potato spindle tuber viroid is processed into a minicircle. *Nucleic Acids Res.*, **31**(3), 988–998. 7, 11, 95
- Schröder, A.R. & Riesner, D. (2002). Detection and analysis of hairpin II, an essential metastable structural element in viroid replication intermediates. *Nucleic Acids Res.*, **30**(15), 3349–3359. 6

- Schulman, D.B. & Setzer, D.R. (2002). Identification and characterization of transcription factor IIIA from *Schizosaccharomyces pombe*. Nucleic Acids Res., **30**(13), 2772–2781. 55, 74, 80
- Schumacher, J, Randles, J.W. & Riesner, D. (1983). A two-dimensional electrophoretic technique for the detection of circular viroids and virusoids. Anal. Biochem., 135, 288–295. 44
- Schumacher, J., Sänger, H.L. & Riesner, D. (1983). Subcellular localization of viroids in highly purified nuclei from tomato leaf tissue. *EMBO J.*, 2, 1549–1555. 8
- Scripture, J.B. & Huber, P.W. (1995). Analysis of the binding of *Xenopus* ribosomal protein L5 to oocyte 5 S rRNA. The major determinants of recognition are located in helix III-loop C. J. Biol. Chem., 270(45), 27358–27365. 80
- Setzer, D.R., Menezes, S.R., Del Rio, S., Hung, VS. & Subramanyan, G. (1996).
  Functional interactions between the zinc fingers of *Xenopus* transcription factor
  IIIA during 5S rRNA binding. *RNA*, 2(12), 1254–1269. 99, 100
- Sommer, M.H., Scully, A.L. & Spector, D.H. (1994). Transactivation by the human cytomegalovirus IE2 86-kilodalton protein requires a domain that binds to both the TATA box-binding protein and the retinoblastoma protein. J. Virol., 68(10), 6223–6231. 64
- Staiger, D., Zecca, L., Wieczorek Kirk, D.A., Apel, K. & Eckstein, L. (2003). The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. *Plant J.*, **33**(2), 361–371. 106
- Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. & Soreq, H. (2005). Function of alternative splicing. *Gene*, **344**, 1–20. **106**
- Stark-Lorenzen, P., Guitton, M.-C., Werner, R. & Mühlbach, H.-P. (1997). Detection and tissue distribution of potato spindle tuber viroid in infected tomato plants by tissue print hybridization. Arch. Virol., 142, 1289–1296. 8
- Steger, G. & Riesner, D. (2006). A-Modul in Molekularer Biophysik Hydrodynamik. Praktikumsskript, Heinrich-Heine-Universität Düsseldorf. 76
- Steitz, J.A., Berg, C., Hendrick, J.P., La Branche-Chabot, H., Metspalu, A., Rinke, J. & Yario, T. (1988). A 5S rRNA/L5 complex is a precursor to ribosome assembly in mammalian cells. J. Cell Biol., 106(3), 545–556. 102

- Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22, 4673–4680, http://bimas.dcrt.nih.gov/clustalw/clustalw.html. 55
- van Hoof, A. & Green, P.J. (1996). Premature nonsense codons decrease the stability of phytohemagglutinin mRNA in a position-dependent manner. *Plant J.*, **10**(3), 415–424. 106
- Varani, G., Cheong, C. & Tinoco, I. (1991). Structure of an unusually stable RNA hairpin. *Biochemistry*, **30**(13), 3280–3289. **93**, 94
- Vitolo, J.M., Yang, Z., Basavappa, R. & Hayes, J.J. (2004). Structural features of transcription factor IIIA bound to a nucleosome in solution. *Mol.*. Cell Biol., 24(2), 697–707. 11, 104
- Wang, B.B. & Brendel, V. (2006). Genomewide comparative analysis of alternative splicing in plants. Proc. Natl. Acad. Sci. U.S.A., 103(18), 7175–7180. 107
- Woo, Y.M., Itaya, A., Owens, R.A., Tang, L., Hammond, R.W., Chou, H.C., Lai, M.M. & Ding, B. (1999). Characterization of nuclear import of potato spindle tuber viroid RNA in permeabilized protoplasts. *Plant. J.*, **17**(6), 627–635. 7
- Yoine, M., Ohto, M.A., Onai, K., Mita, S. & Nakamura, K. (2006). The lba1 mutation of UPF1 RNA helicase involved in nonsense-mediated mRNA decay causes pleiotropic phenotypic changes and altered sugar signalling in Arabidopsis. *Plant* J., 47(1), 49–62. 107

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 10. November 2006

(Oliver Bannach)