

**Der Ursprung der IF-Multigenfamilie bei Chordaten.  
Neue Daten und Analysen zur  
Evolution zytoplasmatischer Intermediärfilament-Proteine.**

Von der Gemeinsamen Naturwissenschaftlichen Fakultät  
der Technischen Universität Carolo-Wilhelmina  
zu Braunschweig  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr. rer. nat.)

genehmigte  
Dissertation

von Alexander Zimek  
aus Essen

1. Referentin oder Referent:	<u>Prof. Dr. J. Wehland</u>
2. Referentin oder Referent:	<u>Prof. Dr. K. Weber</u>
eingereicht am:	<u>14. August 2002</u>
mündliche Prüfung (Disputation) am:	<u>29. Oktober 2002</u> <u>2002</u>



### Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Gemeinsamen Naturwissenschaftlichen Fakultät, vertreten durch die Mentorin oder den Mentor/die Betreuerin oder den Betreuer der Arbeit, in folgenden Beiträgen vorab veröffentlicht.

Publikationen:

Riemer, D., Wang, J., Zimek, A., Swalla, B.J. und Weber, K. (2000). Tunicates have unusual nuclear lamins with a large deletion in the carboxyterminal tail domain. *Gene* **255**, 317-325.

Wang, J., Karabinos, A., Zimek, A., Meyer, M., Riemer, D., Hudson, C., Lemaire, P. und Weber, K. (2002). Cytoplasmic intermediate filament protein expression in tunicate development; a specific marker for the test cells. *Eur. J. Cell Biol.* **81**, 302-311.

Zimek, A. und Weber, K. (2002). The gene for a cytoplasmic intermediate filament (IF) protein of the hemichordate *Saccoglossus kowalevskii*; definition of the unique features of chordate IF proteins. *Gene* **288**, 187-193.

# Inhalt

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Material und Methoden</b>	<b>9</b>
2.1	Materialien	9
2.1.1	Chemikalien	9
2.1.2	Wachstumsmedien und Zusätze	9
2.1.3	Enzyme und Proteine	10
2.1.4	DNA-Oligonukleotide	10
2.1.5	Verbrauchsmaterial	10
2.1.6	Geräte	11
2.1.7	Wasser	11
2.1.8	Organismen	12
2.1.9	Standard- und Pufferlösungen	12
2.1.10	Plasmide, Bakterienstämme und $\lambda$ -Phagen	15
2.1.10.1	Plasmide	15
2.1.10.2	Bakterienstämme	15
2.1.10.3	$\lambda$ -Phagen	15
2.2	Protein- und Biochemische Methoden	16
2.2.1	Proteinfällung	16
2.2.2	Antikörper (AK)	16
2.2.2.1	Erstellung und Reinigung polyklonaler Peptidantikörper	16
2.2.2.2	Kopplung des Peptids an BSA und Sepharose	16
2.2.2.3	Affinitätsreinigung polyklonaler Antikörper	18
2.2.3	Diskontinuierliche SDS-PAGE	18
2.2.4	Dot-Blot	19
2.2.5	Western-Blot Analyse	20
2.2.6	Mikrosequenzierung von Proteinen und Peptiden	20
2.2.7	Expression rekombinanter Proteine	21
2.2.7.1	Induktion	21
2.2.7.2	Inclusionbody Präparation	21
2.2.7.3	Harnstoff-FPLC	22
2.3	Molekularbiologische Methoden	23
2.3.1	Konzentrationsbestimmung von Nukleinsäuren	23
2.3.2	RNA-Isolierung	24
2.3.2.1	Gewinnung von gesamt-RNA aus Tiergeweben	24
2.3.2.2	Anreicherung von poly A <sup>+</sup> RNA aus Gesamt-RNA	24
2.3.3	Arbeiten mit $\lambda$ -cDNA-Banken	25
2.3.3.1	Erstellung von $\lambda$ -cDNA-Banken	25
2.3.3.2	$\lambda$ -kompetente <i>E. coli</i> Zellen und Plattierung von $\lambda$ -cDNA-Banken	25
2.3.3.3	Isolierung von IF-Kandidaten aus $\lambda$ -cDNA-Banken mit monoklonalen Antikörpern	26
2.3.3.4	cDNA-Screen von $\lambda$ -cDNA-Genbanken	27
2.3.4	Polymerase-Kettenreaktion (PCR)	28
2.3.4.1	Genomische PCR	29
2.3.4.2	Kolonie-PCR	29
2.3.4.3	Mutagenese-PCR	30
2.3.4.4	Halbspezifische PCR-Reaktion	30
2.3.4.5	PCR auf $\lambda$ -Phagen	30
2.3.4.6	Synthese von DNA-Sonden mit DIG-dUTP	30
2.3.5	Extraktion von genomischer DNA	31
2.3.6	Agarosegelelektrophorese	31

2.3.7	DNA-Extraktion aus Agarosegelen	32
2.3.8	Fällung von Nukleinsäuren	32
2.3.9	Restriktionsverdau	33
2.3.10	Ligation von DNA-Molekülen	33
2.3.11	Transformation von <i>E. coli</i> Zellen	33
2.3.12	DNA-Isolierung	34
2.3.12.1	Analytische Plasmidisolierung (Mini-Präparation)	34
2.3.12.2	Präparative Plasmidisolierung (Midi-Präparation)	34
2.4	Datenverarbeitung	35
<b>3</b>	<b>Ergebnisse</b>	<b>37</b>
3.1	Isolierung von IF-cDNA-Klonen bei ausgewählten Invertebraten	37
3.1.1	Arbeiten mit $\lambda$ -cDNA-Expressionsbanken	38
3.1.1.1	Erstellung von $\lambda$ -cDNA-Expressionsbanken	38
3.1.1.2	Plattierung von $\lambda$ -cDNA-Expressionsbanken	39
3.1.1.3	Isolierung von IF-cDNAs mit dem IFA-Antikörper	40
3.1.1.4	Klonierung der isolierten cDNA-Fragmente mittels PCR	41
3.2	Zwei neue IF-Klone bei <i>Ciona intestinalis</i>	42
3.2.1	Charakterisierung von <i>C. intestinalis</i> Lamin L2	42
3.2.2	Räumlich-zeitliche Analyse der Expression von Lamin L2	43
3.2.3	Charakterisierung von <i>Ciona</i> IF A2	44
3.2.4	Weitere Untersuchungen der Tunikaten-IF-Proteine	44
3.3	Das erste zytoplasmatische IF-Protein des Hemichordaten <i>S. kowalevskii</i>	46
3.3.1	Sacco IF1 head-Domäne	46
3.3.2	Die Organisation der rod-Domäne	47
3.3.3	Die Intronverteilung des Sacco IF1 Gens	49
3.3.3.1	Sacco IF1 ist das erste deuterostomische IF <sup>zyt</sup> -Protein mit einem Laminhomologiesegment in der tail-Domäne	50
3.3.4	Rekombinante Expression des Sacco IF1-Proteins	51
3.3.5	Erstellung eines polyklonalen anti-Sacco IF1 Peptidantikörpers	52
3.4	Klonierung eines zytoplasmatischen IF-Proteins von <i>Priapulus caudatus</i>	53
3.5	Die Phylogenie der Intermediärfilamentproteine	54
<b>4</b>	<b>Diskussion</b>	<b>57</b>
4.1	Charakterisierung eines zytoplasmatischen Hemichordaten-IF-Proteins	58
4.2	Die kurze coil 1b-Variante entstammt den ursprünglichen Chordaten	59
4.3	Lamine: Der Grundtyp und seine Varianten	60
4.4	Die Evolution der zytoplasmatischen IF-Proteine	63
4.5	Wie viele IF-Proteine braucht ein Tier ?	65
4.6	Kein Königsweg zur Identifizierung von IF-Proteinen	66
<b>5</b>	<b>Zusammenfassung</b>	<b>68</b>
<b>6</b>	<b>Referenzen</b>	<b>70</b>
<b>7</b>	<b>Anhang</b>	<b>75</b>

## Abkürzungen

Amp.	Ampicillin
APS	Ammoniumperoxidsulfat
ATP	Adenosintriphosphat
bp	Basenpaar(e)
β-ME	β-Mercaptoethanol
BSA	Rinderserumalbumin
cDNA	zur mRNA komplementäre DNA
°C	Grad Celsius
Da	Dalton, Molekulargewichtseinheit
ddH <sub>2</sub> O	autoklaviertes entmineralisiertes Wasser
DEPC	Diethylpyrocarbonat
DMSO	Dimethylsulfoxid
dNTP	Desoxyribonukleosidphosphat
DNA	Desoxyribonukleinsäure
EDTA	Ethylendiamintetraacetat
EST	Expressed Sequence Tags (Sequenzprojekte über willkürlich klonierte cDNA-Fragmente)
EtOH	Ethanol
FPLC	Fast Protein Liquid Chromatography
g	Zentrifugalkraft: 1g=9,81m/s <sup>2</sup>
H <sub>2</sub> O	entmineralisiertes Wasser
IF	Intermediärfilamente
IPTG	Isopropylthiogalaktosid
LHS	Laminhomologiesegment
M	Konzentration in Mol/Liter
mRNA	messenger-(Boten-) Ribonukleinsäure
nm	Nanometer (10 <sup>-9</sup> m)
NLS	Kernlokalisierungssignal
OD <sub>600</sub>	optische Dichte, Wellenlänge in nm
PAGE	Polyacrylamid-Gelelektrophorese
PBS	Phosphatgepufferte Saline-Lösung
PCR	Polymerase Kettenreaktion
pfu	Plauebildende Einheit, Anzahl der infektiösen Phagenpatrikel
PEG	Polyethylenglykol
Primer	Oligonukleotide für verschiedene <i>in vitro</i> -Verfahren

rDNA	ribosomale Desoxyribonukleinsäure
RNA	Ribonukleinsäure
RT	Raumtemperatur (25 °C)
RT-PCR	Reverse Transkriptions-PCR
Sarcosyl	Natriumlauroylsarcosin
SDS	Natriumdodecylsulfat
ssss-DNA	einzelsträngige Lachssperma-DNA
TEMED	N,N,N',N'-Tetramethylethyldiamin
Tet.	Tetrazyklin
Tris	Tris-(hydroxymethyl)-aminomethan
Tween 20	Polyoxyethylen-Sorbitan Monolaurat
U	Unit (Einheit der Enzymaktivität)
Upm	Umdrehungen pro Minute
UTR	Nichttranslatierbarer Bereich der DNA-Sequenz
UV	Ultraviolett
v/v	Volumen pro Volumen
w/v	Gewicht pro Volumen
w/w	Gewicht pro Gewicht
XGal	5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galaktopyranosid
zyt	Kurzform für zytoplasmatisch, hier im speziellen für IF-Proteine

Alle weiteren Abkürzungen entsprechen den Vorschlägen der

IUPAC (International Union of Pure and Applied Chemistry)

und

IUBMB (International Union of Biochemistry and Molecular Biology).

Nukleotidsequenzen werden in 5'-3'-Richtung aufgeführt. Weiterhin wurden die anerkannten SI-Einheiten verwendet.





## 1 Einleitung

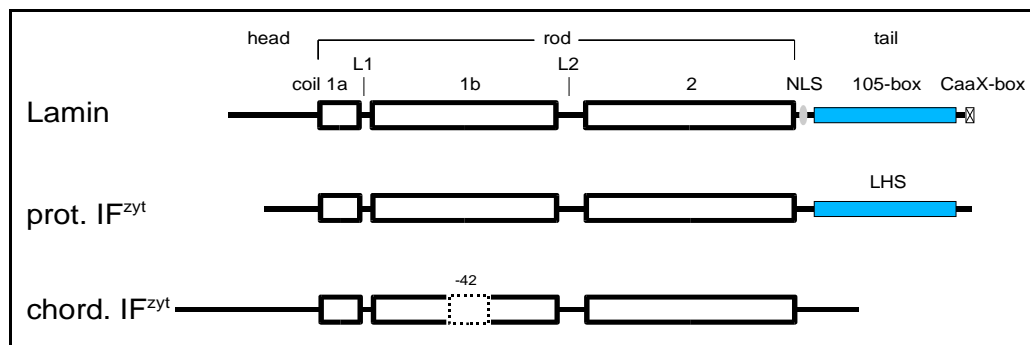
Die Zelle ist die kleinste Einheit der belebten Natur, die in der Lage ist, alle Grundfunktionen der Organismen zu erfüllen. Dazu gehören Stoffwechsel, Wachstum und Vermehrung. Die Zelle wird von einer Membran eingeschlossen, die sie von der Umgebung abgrenzt und die notwendigen Transportprozesse spezifisch regelt. Im Zytoplasma der Zelle befinden sich neben dem Hauptbestandteil Wasser auch Salze, globuläre Proteine, Energieträger und die höheren Funktionseinheiten, die Organellen. Diese Kompartimente werden von eigenen Membranen eingeschlossen und stellen optimale Bedingungen für die dort stattfindenden Prozesse sicher. Im Kern wird die genetische Information jeder Zelle aufbewahrt und ihre Expression geregelt. Diese erfolgt mit dem Zwischenschritt der Transkription, bei der die mRNA gebildet wird, über Proteine, die stets die Effektoren der zellulären Prozesse sind. Die unterschiedlichen Regionen der tierischen Zelle sind durch die geordnete Struktur des proteinösen Zytoskeletts miteinander verbunden. Es ist hauptverantwortlich für die mechanische Stabilität der Zelle (JANMEY ET AL. 1991). Weiterhin werden Vorgänge wie die Zellteilung, Zellbewegung und die Veränderung der Zellgestalt vom Zytoskelett bewerkstelligt. Der Transport von Vesikeln entlang des Zytoskeletts ermöglicht den intrazellulären Informationsaustausch.

Bis Mitte der 70er Jahre waren nur zwei Komponenten des Zytoskeletts bekannt. Die Aktinfilamente ( $\varnothing$ : 6 nm) und die Mikrotubuli ( $\varnothing$ : 23 nm) sind aus den globulären Einheiten Aktin bzw. Tubulin aufgebaut. Dann wurde ein dritter Bestandteil des Zytoskeletts in tierischen Zellen entdeckt. Die Intermediärfilamente ( $\varnothing$ : 10 nm) konnten erstmals in den Neuronen von Vertebratenzellen identifiziert werden. Bis heute sind über 150 verschiedene IF-Proteine isoliert oder kloniert worden, allein im Menschen beträgt ihre Anzahl 62 (HESSE ET AL. 2001). Die IFs haben aufgrund ihrer Struktur, sie bestehen aus lang gestreckten Proteinen, am deutlichsten den Charakter der stabilisierenden Elemente. Ihre Widerstandsfähigkeit gegenüber mechanischer Belastung ist höher als die der Aktinfilamente und Mikrotubuli (JANMEY ET AL. 1991). Im Gegensatz zu Aktin und Tubulin zeigen die Proteine der zytoplasmatischen Intermediärfilamente (IF<sup>zyt</sup>-Proteine) einen geringen Grad an Sequenzkonservierung von teilweise nur noch 20 %. Sie werden z.B. bei Chordaten gewebespezifisch (FUCHS & WEBER 1994, PARRY & STEINERT 1995, KARABINOS ET AL. 2000, WANG ET AL. 2002) und teilweise zu unterschiedlichen Zeitpunkten der Ontogenese exprimiert woraus sich unter anderem eine Unterteilung in fünf IF-Typen ergibt. Für diese Zuordnung sind darüber hinaus Übereinstimmungen der Aminosäuresequenz, Genorganisation und des Expressionsmusters von Bedeutung (FUCHS & WEBER 1994, QUINLAN ET AL. 1994).

In Epithelien findet man Keratinfilamente. Sie werden aus sauren (Typ I: pI= 4-6) und neutral bis basischen (Typ II: pI= 6-8) Keratinproteinen aufgebaut. Die Keratine stellen die größte Gruppe der menschlichen IF-Proteine. Sie bilden heteropolymere Filamente aus Typ I- und Typ II-Proteinen im molaren Verhältnis 1:1, d.h. der doppelsträngige coiled coil ist ein Heterodimer (HATZFELD & WEBER 1990). Bei den Filamenten aus Typ III-Proteinen, zu denen unter anderem Desmin, Vimentin und Peripherin zählen, werden doppelsträngige homopolymere coiled coils gebildet. Sie werden bei Vertebraten hauptsächlich in mesenchymalen Zellen exprimiert. In den Neuronen der Vertebraten findet man lange Filamente vor allem aus den Typ IV-Proteinen NF-L, NF-M und NF-H. Die Lamine bilden die fünfte Gruppe der IF-Proteine (AEBI ET AL. 1986). Sie werden nur im Zellkern angetroffen und sind essentielle Bestandteile der Kernlamina. Sie können in Vertebraten in die zwei Subtypen A und B unterteilt werden. In undifferenzierten Zellen findet man während der frühen Embryogenese nur B-Typ Lamine. Es konnte gezeigt werden (RÖBER ET AL. 1989), dass die A-Typ Lamine in Mausembryonen nicht simultan in allen Geweben exprimiert werden. So beginnt die Lamin A-Expression in Herzmuskel und Gehirn erst fünf, bzw. sieben Tage nach der Geburt. Trotz der unterschiedlichen Expressionsmuster der IF-Proteine und ihrer sehr geringen Sequenzidentität untereinander lassen sich wichtige strukturelle Gemeinsamkeiten feststellen (Figur 1).

IF-Proteine gliedern sich in eine head-, rod- und tail-Domäne. Die head- und tail-Domänen haben oft unterschiedliche Längen, ihre Sequenzen zeigen hohe Variabilitäten und es gibt nur in der tail-Domäne der Lamine zwei echte Konsensussequenzen. Das Kernlokalisierungssignal (NLS) ist das Erkennungsmotiv eines zytosolischen Carrierproteins das den Kerntransport initiiert. Die carboxyterminale CaaX-Box (Cys-aliphatisch-aliphatisch-beliebig) ist der Ort mehrerer posttranslationaler Modifikationen. Durch Isoprenylierung des Cysteins, anschließende Abspaltung der letzten drei AS-Reste und Methylierung der endständigen Carboxylgruppe wird die Wechselwirkung des Lamins mit der Kernmembran ermöglicht.

Die große Variabilität der tail-Domäne spiegelt sich im Spektrum der IF-Molekulargewichte wieder, welche einen Bereich von 40 kD für manche Keratine, bis zu 110 kD beim humanen NF-H umfassen. Für die Neurofilamente NF-M und NF-H konnte in elektronenmikroskopischen Aufnahmen gezeigt werden, dass die tail-Domänen seitlich vom Filament abstehen (FUCHS UND CLEVELAND 1998) und darüber möglicherweise einen Einfluss auf den Durchmesser der Neuronen und die davon abhängige Leitungsgeschwindigkeit haben (OHARA ET AL. 1993).



Figur 1: Schematisches Strukturmodell der zyttoplasmatischen IF-Proteine und der kernlokalisierten Lamine. Für alle drei Typen gilt ein dreigeteiltes Domänenorganisationsprinzip. Ein zentraler stabartiger Abschnitt (rod) wird in drei Subdomänen unterteilt (coil 1a, 1b, 2). Diese  $\alpha$ -helikalen Segmente werden durch die zwei kurzen nicht helikalen Abschnitte L1 und L2 miteinander verbunden. Die coil 1b-Subdomäne in den IF<sup>zyt</sup>-Proteinen der Chordaten ist gegenüber den protostomischen IF<sup>zyt</sup>-Proteinen und den Laminen um 42 AS verkürzt. Lamine haben in der tail-Domäne einen konservierten ca. 105 AS langen Abschnitt. Diesem homolog ist das Laminhomologiesegment (LHS) der protostomischen IF<sup>zyt</sup>-Proteine. Dem IF-Protein der Chordaten fehlt das LHS in der tail-Domäne.

Die Zeichnung ist nur schematisch, die tail-Domäne neuronaler Typ IV IF-Proteine kann wesentlich länger ausfallen als die der Lamine. CaaX-Box und NLS sind lamintypische Sequenzmotive die mit der Kernlokalisation im direkten Zusammenhang stehen.

Der zentrale Abschnitt der Proteine ist trotz großer Variabilität in der Sequenz weitgehend längenkonstant. Die rod-Domäne besteht aus den drei helikalen Subdomänen coil 1a, 1b und 2, welche durch die nicht helikalen Bereiche L1 und L2 verbunden sind. Die drei Subdomänen haben eine typische Heptadenstruktur: Ein apolarer Aminosäurerest befindet sich an erster und vierter Position solch einer Siebenerwiederholung. Dies bedeutet, dass die hydrophoben AS-Reste auf der Außenseite der Helix in einer Linie liegen. Die  $\alpha$ -Helices können mit den analogen Bereichen eines zweiten IF-Proteins in Wechselwirkung treten und bilden mit einem parallelen coiled coil Dimer den ersten „Baustein“ der Intermediärfilamente.

Der weitere Aufbau der Filamente beginnt mit der schrittweisen lateralen Zusammenlagerung von Dimeren zu einem Protofilament (HERRMANN & AEBI 1999). In Folge der weiteren Prozesse entsteht zum Schluss eine Filament-Untereinheit (ULF) mit bis zu 16 Dimeren im Querschnitt und einer Länge von ungefähr 46 nm (HERRMANN ET AL. 1996). Diese Einheiten lagern sich longitudinal aneinander, durch interne Umstrukturierung und teilweise Überlappung der head- und tail-Domänen kommt es zu einer Verdichtung des Filaments und einer Reduzierung des Durchmessers auf ca. 10 nm (HERRMANN & AEBI 1998 UND 1999).

Aktuelle Röntgenstrukturanalysen der isolierten rod-Enden des humanen Vimentins deuten darauf hin, dass der coil 1b-Subdomäne eine besondere Bedeutung bei der Ausbildung von Wechselwirkungen zwischen zwei Dimeren zukommt (STRELKOV ET AL.

2002). Der isolierte carboxyterminale Bereich der coil 2-Subdomäne hat hingegen eine starke Tendenz coiled coil Strukturen zu bilden. Das Fragment ist in der Lage *in vitro* mit natürlichen Vimentin-Filamenten zu interagieren und im molaren Überschuss sogar geordnete IF-Strukturen zu stören (HATZFELD & WEBER 1992).

Es wurde auch gezeigt, dass IF-Proteine deren rod-Enden mutiert sind, *in vitro* keine höher geordneten Strukturen bilden können (HATZFELD & WEBER 1992, GEISLER ET AL. 1993). In der Praxis bedeutet dies, dass punktmutierte epidermale Keratinproteine zu schwerwiegenden vererbaren Hautkrankheiten wie „Epidermolysis Bullosa simplex“ und „Epidermolysis Hyperkeratosis“ führen können. Bereits geringe Scherkräfte sind dann ausreichend um die Haut abzulösen. Auch das Darmgewebe kann von mutierten Keratingenen (K8) und die Herzmuskulatur von fehlendem Desmin betroffen sein. Mittlerweile können 19 IF-Genen aus Mensch und Maus bestimmte pathologische Phänotypen zugeordnet werden (FUCHS & CLEVELAND 1998). Es zeigt sich deutlich, dass die Funktion der IFs mit der Gewährleistung mechanischer Stabilität im Zusammenhang steht und Störungen der Filamentbildung oft schwerwiegende Folgen haben.

Die IF-Proteine vom Typ IV werden nur in Neuronen exprimiert und stellen vor allem in den langen myelinisierten Axonen die entscheidende stabilisierende Struktur dar. Auch für Mutationen der Neurofilamente wurden verschiedene, zum Teil tödlich verlaufende Krankheiten beschrieben. Das veränderte Protein kann keine funktionierenden Filamente bilden, wodurch der axonale Transport möglicherweise gestört wird und die Neuronen infolge dessen degenerieren (JULIEN 1997).

Im Unterschied zu den zytoplasmatischen IF-Proteinen der Typen I-IV sind die Lamine (Typ V) in der Kernmembran lokalisiert. Sie bilden die innere Schicht der Kernhülle bei allen Eumetazoa und sind auch an der Organisation des Chromatins beteiligt. Sie gewährleisten unter anderem eine gleichmäßige Verteilung der Kernporenkomplexe (NPC). Ein entsprechender Phänotyp wurde in *Drosophila* nachgewiesen, bei dem sich die NPCs in unregelmäßigen Haufen ansammelten (LENZ-BÖHME ET AL. 1997). Auch DNA-Bindungseigenschaften wurden bei Laminen gezeigt, so dass sie möglicherweise eine Funktion bei der Initiierung der DNA-Replikation übernehmen (STUURMAN ET AL. 1998).

Die große Anzahl der IF-Proteine und die hohe Komplexität der humanen IF-Multigenfamilie wurden zum Anlass genommen, um die Evolution der IF-Proteine in Vertebraten und Evertrebraten zu studieren. Dabei ist zu bedenken, dass die Evertrebraten nur ein Paraphylum darstellen, Tierstämme wie Protostomier, Deuterostomier, Chordaten und Vertebraten hingegen phylogenetisch begründet sind. Die gegen Ende der 80er Jahre

zur Verfügung stehenden Sequenzdaten für IF-Proteine stammten vor allem aus Vertebraten und zum Teil auch aus Mollusken und Nematoden. Dabei wurde deutlich, dass die IF<sup>zyt</sup>-Proteine der Protostomier bezogen auf Länge und Organisation der Domänen größere Ähnlichkeit mit den Laminen als den IF<sup>zyt</sup>-Proteinen der Vertebraten hatten. Diese waren in der coil 1b-Subdomäne gegenüber den Laminen und protostomischen IF<sup>zyt</sup>-Proteinen um 42 AS, also genau sechs Heptaden verkürzt. Die protostomischen IF<sup>zyt</sup>-Proteine hatten außerdem häufig ein ungefähr 105 AS langes Segment in der tail-Domäne, welches auch bei Laminen, nicht jedoch bei den IF<sup>zyt</sup>-Proteinen der Vertebraten vorhanden war und als Laminhomologiesegment (LHS) bezeichnet wird. Die strukturelle Ähnlichkeit der protostomischen IF<sup>zyt</sup>-Proteine zu den Laminen führte zu der Hypothese, dass diese aus einem mutierten Lamin entstanden sein könnten (WEBER ET AL. 1989, DÖRING UND STICK 1990, DODEMONT ET AL. 1990 & 1994). Dazu wäre der Verlust von CaaX-Box und NLS nötig gewesen.

Mit Zunahme der verfügbaren IF-Sequenzen aus protostomischen und deuterostomischen Tieren und den ergebnislosen Versuchen diese aus Pflanzen, Pilzen und einzelligen Eukaryonten zu isolieren oder zu klonieren zeigte sich, dass nur in den Metazoa IF-Proteine vorhanden sind. Einzellige Eukaryonten und Pflanzen bauen ihr Zytoskelett demnach nur aus Aktin und Tubulin auf (ERBER ET AL. 1998). Weiterhin war in allen isolierten Laminen die lange coil 1b-Variante vorhanden. Die protostomischen IF<sup>zyt</sup>-Proteine haben wie die Lamine eine lange coil 1b-Subdomäne und meistens ein LHS in der tail-Domäne. Von den 11 IF-Proteinen des Nematoden *C. elegans* zeigen sechs das LHS (KARABINOS ET AL. 2001). Die verkürzte coil 1b-Variante war jedoch nicht nur in Vertebraten, sondern auch in allen bearbeiteten Cephalochordaten (RIEMER ET AL. 1992) und Urochordaten (WANG ET AL. 2000) vorhanden.

Mit einer Datenbasis von mittlerweile über 150 IF<sup>zyt</sup>-Proteinen konnten nun zwei zytoplasmatische IF-Prototypen klar definiert werden. Der erste IF<sup>zyt</sup>-Prototyp ähnelt der aus allen Tieren bekannten Laminstruktur (Figur 1, Mitte) und ist typisch für protostomische IF<sup>zyt</sup>-Sequenzen. Die coil 1b-Subdomäne ist immer lang. Der zweite Prototyp ist charakteristisch für Chordaten und weicht in folgenden Punkten vom protostomischen Typ ab: Die coil 1b-Subdomäne ist immer um 42 AS verkürzt und auch das LHS fehlt immer.

Nahezu alle zytoplasmatischen IF-Proteine der Vertebraten können eindeutig einer der vier IF-Proteintypen I-IV zugeordnet werden. In den IF<sup>zyt</sup>-Proteinen der übrigen Chordaten (Cephalo- und Urochordaten) lassen sich bisher nur die Typen I-III unterscheiden. Eine ausschließlich sequenzbasierte Einteilung der IF-Proteine primitiver Chordaten in die Typen I-III ist jedoch schwierig. Deshalb erfolgte die Zuordnung bei

primitiven Chordaten aufgrund starker Sequenzunterschiede u.a. aufgrund der erfolgreichen Bildung chimärer Filamente mit Keratin K8, einem Typ II-Keratin. Dazu wurde rekombinant exprimiertes Keratin K8 aus *X. laevis* mit *Styela* D-Protein gemischt, was zur Bildung von Filamenten führte und das Protein *Styela* D folglich als Typ I-Keratin identifizierte. Vergleichbare Experimente ermöglichten die Einordnung der *Branchiostoma*-IF-Proteine E1/Y1 und D1 als Typ I-, bzw. Typ II-IF-Proteine (KARABINOS ET AL. 2002, WANG ET AL. 2000).

In dieser Arbeit habe ich mit zwei unterschiedlichen methodischen Ansätzen versucht, cDNA-Banken verschiedener Invertebraten zu durchsuchen. Das Ziel war die Identifizierung und Klonierung neuer IF-Proteine. Zu den untersuchten Organismen gehörte zunächst der Protostomier *P. caudatus* aus dem bislang nur ein Lamin aber keine zytoplasmatischen IF-Proteine bekannt sind.

Viel umfassender jedoch waren die Arbeiten im deuterostomischen Tierreich, wo die wichtige Frage nach dem Ursprung der absolut chordatenspezifischen kurzen coil 1b-Subdomäne noch unbeantwortet ist. Die Überprüfung primitiver Deuterostomier und der entsprechende Nachweis hätte diese Deletion zum Merkmal aller Deuterostomier erhoben. Im anderen Fall, einer langen coil 1b-Subdomäne und evtl. einem LHS in der tail-Domäne, wären die entsprechenden Deletionen ausschließlich chordatenspezifisch geblieben. Die bislang bestehende Literatur bevorzugt die erste Hypothese (ERBER 1999). Für die Experimente standen unter anderem zwei cDNA-Banken aus *Asterias rubens* und *Astropecten brasiliensis* zur Verfügung. Vom entwicklungsbiologisch relevanten Modellorganismus *Strongylocentrotus purpuratus* waren isolierte Saugfüßchen vorhanden aus denen ich eine cDNA-Bank erstellen konnte. Die Hemichordaten bilden mit den Echinodermen die Schwestergruppen der Chordaten und so wurde auch *Saccoglossus kowalevskii* wegen der bereits vorhandenen cDNA-Bank ausgewählt um dort nach IF-Proteinen zu suchen und die korrespondierenden cDNA-Klone zu isolieren.

Außerdem fehlten bislang noch Typ III und Typ IV IF-Proteine in den primitiven Chordaten (Cephalochordaten und Urochordaten) weshalb jeweils aus dem Darmgewebe der Ascidie *C. intestinalis* und dem Lanzettfischchen *B. lanceolatum* Expressionsbanken erstellt und untersucht wurden.

Im Mittelpunkt der vorliegenden Arbeit steht die Klonierung neuer IF<sup>zyt</sup>-Proteine aus verschiedenen Invertebraten. Durch die Vergrößerung des Datenmaterials, insbesondere bei Hemi- und Urochordaten, sollen der Ursprung der IF-Multigenfamilie bei Chordaten bestimmt werden und weitere Rückschlüsse auf die IF-Evolution möglich gemacht werden.







## 2 Material und Methoden

### 2.1 Materialien

#### 2.1.1 Chemikalien

Salze und Lösungsmittel wurden, wenn nicht anders beschrieben, in *pro analysi*-Qualität verwendet. Sie stammten in der Regel von den Firmen Merck (Darmstadt), Sigma (St. Louis, USA) und Baker (Deventer, NL).

#### 2.1.2 Wachstumsmedien und Zusätze

Die Wachstumsmedien für Bakterien wurden mit Bactoagar, Hefeextrakt und Trypton von Difco-Becton Dickinson (Le Pont de Claix, F) hergestellt. Zur Selektion wurden die Antibiotika Ampicillin von USB-Amersham (Braunschweig, D) und Tetrazyklin von Boehringer (Mannheim, D) verwendet.

#### SOB-Medium für *E. coli* TOP10-F':

20g	Trypton
5g	Hefeextrakt
0,59 g	NaCl
0,19 g	KCl
add 1,0 l	H <sub>2</sub> O

Für MRF'-Zellen wurden 0,2 % (w/v) Maltose und 10 mM MgSO<sub>4</sub> zugesetzt.

#### SOC-Medium:

20 g	Trypton
5 g	Hefeextrakt
0,59 g	NaCl
0,19 g	KCl
20 mM	MgSO <sub>4</sub>
20 mM	Glucose
add 1,0 l	H <sub>2</sub> O

#### NZY-Medium für die Arbeit mit phagenkompetenten Zellen *E. coli* MRF':

10 g	Trypton
5 g	Hefeextrakt
5 g	NaCl
2 g	MgSO <sub>4</sub> (7 x H <sub>2</sub> O)
add 1,0 l	H <sub>2</sub> O, pH 7,5

**LB-Medium** (SAMBROOK ET AL. 1989)

ist aufgrund seines höheren Salzgehaltes gut geeignet für die präparative Plasmidisolierung:

10 g	NaCl
10 g	Trypton
5 g	Hefeextrakt
add 500 ml	H <sub>2</sub> O, pH 7,0

Für Kulturplatten wurde den Medien 1,5 % Bacto-Agar zugesetzt. Top-Agar für die Plattierung von  $\lambda$ -cDNA-Banken enthielt 0,7 % Agarose. Selektivmedien enthielten 50  $\mu$ g/ml Ampicillin, bzw. 15  $\mu$ g/ml Tetrazyklin. Für einen Blau/Weiss-Screen wurden auf den Agarplatten ( $\varnothing$  8,5 cm) jeweils 40  $\mu$ l einer 0,1 M wässrigen IPTG-Lösung und einer 2 %igen (w/v) X-Gal-Lösung (in DMF) verteilt.

**2.1.3 Enzyme und Proteine**

Für die PCR wurde ausschließlich TaKaRa Ex Taq Polymerase, TaKaRa, (JP) unter Verwendung der mitgelieferten Puffer und Nukleotide eingesetzt. Die Restriktionsenzyme Bam HI und Nde I, GibcoBRL (Eggenstein, D), und T4-Ligase, Invitrogen (Leek, NL) wurden für die Umklonierung in Expressionsvektoren eingesetzt. Um  $\lambda$ -cDNA-Banken zu untersuchen wurde an Streptavidin gekoppelte Meerrettichperoxidase, Amersham Life-Science Pharmacia (Freiburg, D), benutzt. Peptidsynthese und Antikörperproduktion: BioScience (Göttingen, D)

**2.1.4 DNA-Oligonukleotide**

DNA-Oligonukleotide für PCR-Reaktionen wurden von IBA/NAPS (Göttingen, D) oder der MWG-Biotech AG (Ebersberg, D) hergestellt. Die Lösungen wurden stets auf eine gebrauchsfertige Konzentration von 10 pmol/ $\mu$ l verdünnt.

**2.1.5 Verbrauchsmaterial**

PCR-Reaktionsgefäße 200  $\mu$ l: Eppendorf (Hamburg)

Polypropylenröhrchen Falcon 14 ml: Becton Dickinson Labware (New Jersey, USA)

Pipettenspitzen und Reaktionsgefäße 1,5 ml / 2,0 ml / 50 ml: Greiner (Frickenhausen, D)

Enzympipettenspitzen 10  $\mu$ l: Gilson (Villers Le Bel, F)

Nitrozellulosemembranen BA 85, Porengröße 0,45  $\mu$ m: Schleicher & Schuell (Dassel, D)

Chromatografiepapier: 1MM und 3MM: Whatman (Maidstone, GB)

Agarose für die Elektrophorese: Pharmacia (Uppsala, S)

Petrischalen  $\varnothing$  8,5 cm, 13 cm: Nalge Nunc (Wiesbaden)

Hyperfilm ECL für Chemilumineszenzexperimente (Amersham Pharmacia Biotech)

### 2.1.6 Geräte

Spannungsversorgung:	EPS 600, Pharmacia Biotech (Uppsala, S)
FPLC-System:	Pharmacia (Uppsala, S)
FPLC-Säulen:	mono Q 5/5 (Anionenaustauscher, 1 ml Säulenvol.) mono S 5/5 (Kationenaustauscher, 1 ml Säulenvol.)

#### Zentrifugen:

- Sorvall RC-5 Superspeed, Du Pont de Nemours (New Town, Connecticut, USA)
- Tischzentrifuge 5415 C (Eppendorf, Hamburg, D)
- Speed Vac Vakuum-Konzentrationszentrifuge (Savant-Instruments, Farmingdale, New York, USA).
- Heraeus Varifuge RF (Osterode, D)

Rotoren:	Sorvall SS34 und GSA, Du Pont de Nemours GmbH (Bad Nauheim, D)
Pipetten:	P2, P20, P100, P200, P1000, P5000, Gilson (Bad Camberg, D)
Inkubator:	Typ B12. Heraeus Instruments (Hanau, D)
Schüttelinkubator:	Incubator Shaker Model G25 New Brunswick Scientific Co. Inc. (Edison, N.J., USA)
Photometrie:	Ultrospec 2000, Pharmacia (Freiburg, D)
PCR-Gerät:	GeneAmp 2400, Perkin-Elmer
Peptidsequenzierung:	Procise Protein Sequencer, Perkin-Elmer
Filmentwicklung:	Curix 60, AGFA

### 2.1.7 Wasser

Für Arbeiten mit DNA und die Herstellung der benötigten Lösungen wurde entmineralisiertes autoklaviertes Wasser aus dem Milli-Q-Filtrationssystem von Millipore (Eschborn, D) verwendet. Das Wasser für alle Arbeiten mit RNA wurde zur Deaktivierung sämtlicher RNAsen zusätzlich mit 0,5 ‰ DEPC, Sigma (Deisenhofen, D), behandelt. Das Gemisch wurde zwei Stunden bei RT gerührt, dann 20 Minuten gekocht und anschließend autoklaviert.

### 2.1.8 Organismen

Für die Herstellung von  $\lambda$ -cDNA-Banken und die Extraktion genomischer DNA wurden tiefgefrorene Gewebe folgender Organismen verwendet.

*Ciona intestinalis*: Darm. Die Tiere stammten von der „Meeresbiologischen Station Helgoland“ und wurden von Prof. Dr. Thomas Bartholomaeus zur Verfügung gestellt. Die lebenden Tiere wurden nach Göttingen transportiert. Die einzelnen Gewebe (Darm, Epidermis, Ganglien) wurde unmittelbar nach der Entnahme in flüssigem Stickstoff schockgefroren.

*Saccoglossus kowalevskii*: Die ganzen Tiere wurden vom „Woods Hole Marine Laboratory“, Mass., USA durch Dr. E. Ungewickel geliefert. In der Regel wurden die Tiere nach dem Fang in flüssigem Stickstoff schockgefroren, auf Trockeneis verschickt und dann bei -80 °C bis zur Verwendung gelagert.

### 2.1.9 Standard- und Pufferlösungen

10 x PBS, pH 7,3	57,5 g 10 g 400 g 10 g add 5 l	$\text{Na}_2\text{HPO}_4$ $\text{KH}_2\text{PO}_4$ ; NaCl KCl $\text{H}_2\text{O}$
PBST 0,2%	Der einfach konzentrierten PBS-Lösung wurden 0,2 % Tween 20 zugesetzt.	
1 x SM	0,58 g 0,2 g 5 ml 0,5 ml 2 % (w/v) add 100 ml	NaCl $\text{MgSO}_4 (7x\text{H}_2\text{O})$ 1 M Tris-HCl, pH 7,5 Gelatine $\text{H}_2\text{O}$
10 x TBE, pH 8,3	215,6 g 119,8 g 14,8 g add 1000 ml	Tris Borsäure EDTA $\text{H}_2\text{O}$
20 x SSC, pH 7,0	876,5 g 441 g 10 ml add 5 l	NaCl Na-Citrat (2 x $\text{H}_2\text{O}$ ) 0,5 M EDTA $\text{H}_2\text{O}$
1 x TE	10 mM 1 mM	Tris-HCL, pH 8,0 EDTA, pH 8,0
1 x TC	10 mM 1 mM	Tris-HCL, pH 8,0 CDTA, pH 8,0

Acrylamid-Bis-Lösung	30 % (w/v) 0,8 % (w/v)	Acrylamid N,N'-Methylenbisacrylamid
Trenngelpuffer	1,5 M 0,4 % (w/v)	Tris-HCl, pH 8,8 SDS
Sammelgelpuffer	0,5 M 0,4 % (w/v)	Tris-HCl, pH-6,8 SDS
APS-Lösung	10 % (w/v)	APS in H <sub>2</sub> O
SDS-Probenpuffer	5 g 25 mg 1,5 g 10 g 80 ml	SDS Bromphenolblau Tris-Base Glycerin H <sub>2</sub> O
Den gesamten Ansatz mischen, dann 20' kochen		
	add 100 ml	H <sub>2</sub> O, pH 6,8
Elektrophorese Laufpuffer	3,0 g 14,4 g 1 g add 1 l pH add 8,8	Tris-Base Glycerin SDS H <sub>2</sub> O HCl <sub>konz.</sub>
Coomassie-Blue Färbelösung	5 g 184 ml 900 ml add 2 l	Coomassie blue (Serva, Heidelberg) Essigsäure Methanol H <sub>2</sub> O
Entfärberlösung	190 ml 125 ml add 2 l	Essigsäure Methanol H <sub>2</sub> O
100 x Denhardts	2 % 2 % 2 %	BSA (w/v) Polyvinylpyrrolidon-40 Ficoll-400 (v/v)

Prähybridisierungslösung:

3,75 ml	Formamid
75 µl	0,5 M EDTA, pH 8,5
150 µl	ssss-DNA (6 mg/ml)

15 Minuten kochen

übrige Komponenten mischen:

4,75 ml	H <sub>2</sub> O
3,75 ml	20 x SSC
75 µl	0,5 M EDTA, pH 8,5
750 µl	100 x Denhardts-Lösung
750 µl	1 M Na-Phosphat, pH 6,8
150 µl	0,1 M Na-Pyrophosphat

auf 37 °C erwärmen

Zugabe von

750 µl	10 % SDS (w/v)
--------	----------------

beide Lösungen vereinigen

Hybridisierungslösung:

2 ml	Formamid
80 µl	ssss-DNA (6 mg/ml)
20 µl	0,5 M EDTA, pH 8,5

20 Minuten kochen

dann auf Eis stellen

1,5 µg add 450 µl	DIG-markierte DNA-Probe H <sub>2</sub> O
----------------------	---

5 Minuten kochen

dann auf Eis stellen

2,64 ml	H <sub>2</sub> O
160 µl	1 M Na-Phosphat, pH 6,8
80 µl	0,1 M Na-Pyrophosphat
2 ml	20 x SSC
80 µl	0,5 M EDTA, pH 8,5
80 µl	100 x Denhardts-Lösung

auf 37 °C erwärmen

400 µl	10 % SDS (w/v)
--------	----------------

Lösungen erst am Schluss zusammen mischen

MAN

23,2 g	Maleinsäure
17,5 g	NaCl
add 2,0 l	H <sub>2</sub> O, pH 7,5

TNM	0,1 M	Tris-Cl, pH 9,5
	0,1 M	NaCl
	50 mM	MgCl <sub>2</sub>
	add 500 ml	H <sub>2</sub> O

## 2.1.10 Plasmide, Bakterienstämme und $\lambda$ -Phagen

### 2.1.10.1 Plasmide

pCR 2.1 TOPO	TA-Klonierungen von PCR-Produkten (Invitrogen, Leek, NL)
pET 23a <sup>+</sup>	Rekombinante Proteinexpression (Novagen, Merck Darmstadt, D)

### 2.1.10.2 Bakterienstämme

Für die Klonierung von PCR-Produkten und die Vermehrung von Expressionsvektoren wurden ultrakompetente TOP10F'-Zellen, Invitrogen (Leek, Niederlande) verwendet:

*E. coli* F' {*lacI<sup>q</sup>* Tn10(*Tet<sup>R</sup>*)} *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\Phi$  $\pi$ <sub>180</sub>*lacZ* $\Delta$ M15 *DlacX74*  
*recA1 deoR araD139*  $\Delta$ (*ara-leu*)7697 *galJ galK rpsL* (*Str<sup>R</sup>*) *endA1 nupG*

Zur Expression rekombinanter Proteine wurden kompetente BL21-Zellen, Novagen (Madison, WI) verwendet. Sie zeichnen sich durch hohe Proteinexpression bei geringer Proteaseaktivität aus:

*E. coli* B F- *dcm ompT hsdS*(rB- mB-) *gal*(DE3)pLysS )

Als Wirtstamm für  $\lambda$ -Phagen dienten XL1-Blue MRF'-Zellen, Stratagene (Heidelberg):

*E. coli*  $\Delta$ (*mcrA*)183  $\Delta$ (*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1*  
*recA1 gyrA96 relA1 lac* [F' *proAB lacI<sup>q</sup>Z* $\Delta$ M15 Tn10(*Tet'*)]

### 2.1.10.3 $\lambda$ -Phagen

Zur Herstellung von cDNA-Banken wurden  $\lambda$ -ZAP Express Phagen, Stratagene (Heidelberg) verwendet.

## 2.2 Protein- und Biochemische Methoden

### 2.2.1 Proteinfällung

Bei geringen Proteinkonzentrationen wurde die wässrige Lösung mit neun Volumenteilen Aceton gemischt, eine Stunde bei -20 °C gekühlt und zentrifugiert (SS34, 10.000 U/min, 4 °C, 30 min). Der Überstand wurde verworfen, das Pellet eine Stunde luftgetrocknet, dann in Probenpuffer aufgenommen und 20 Minuten gekocht.

### 2.2.2 Antikörper (AK)

Zur Durchmusterung verschiedener  $\lambda$ -cDNA-Banken wurde der monoklonale IFA-Antikörper (PRUSS ET AL. 1981) benutzt. Als sekundärer Antikörper wurde biotinylierter Schwein- $\alpha$ -Maus-AK verwendet. In Western-Blots erfolgte die Detektion des primären AK mit einem an Meerrettichperoxidase gekoppelten Schwein- $\alpha$ -Kaninchen-AK, bzw. Kaninchen- $\alpha$ -Maus-AK, Dako (Hamburg, D). Nach Isolierung eines zytoplasmatischen IF-Proteins aus *S. kowalevskii* wurde ein polyklonaler Antikörper in Kaninchen erstellt.

#### 2.2.2.1 Erstellung und Reinigung polyklonaler Peptidantikörper

Ein polyklonaler Antikörper dient der Detektion des Zielproteins in histologischen Schnitten fluoreszenzmikroskopischer Untersuchungen und dem Nachweis des Proteins in Gewebeextrakten. Die Auswahl einer Peptidsequenz aus den carboxyterminalen 14 bis 16 Aminosäuren ergab in der Vergangenheit stets gute und spezifische Antikörper. Das Peptid mit der Sequenz NH<sub>2</sub>-CKTTHFSKSKRVLEYA-COOH wurde synthetisiert, gereinigt und über das aminoterminal Cystein an Hemocyanin gekoppelt. Ein Kaninchen wurde immunisiert und das antikörperhaltige Serum nach ca. sechs Wochen entnommen und getestet. Diese Arbeiten wurden von der Firma Bioscience (Göttingen) durchgeführt.

#### 2.2.2.2 Kopplung des Peptids an BSA und Sepharose

Da im Serum immunisierter Kaninchen viele unspezifische Antikörper enthalten sind wurde der gesuchte Antikörper mittels Bioaffinitäts-Chromatografie über peptidbestückte Sepharosesäulen konzentriert. Dazu wurde zunächst (1) BSA an Sulfo-MBS gekoppelt. Anschließend (2) konnten die synthetischen Peptide über das aminoterminal Cystein an die Sulfo-MBS Komponente gekoppelt werden. Dieses Konjugat wurde dann über die Lysinseitenketten des BSA-Moleküls an die BrCn-aktivierte Sepharose gekoppelt. Die synthetischen Peptide hingen also an einem langen Molekülkomplex aus der Sepharosematrix, BSA und Sulfo-MBS wodurch eine gute Erreichbarkeit für den Antikörper sichergestellt wurde



## Benötigte Puffer:

	50 mM	Natriumphosphat-Puffer, pH 6,0
	0,5 M	Natriumphosphat-Puffer, pH 7,5
	0,1 M	Natriumhydrogencarbonat (NaHCO <sub>3</sub> ) + 0,5 M NaCl, pH 8,3
	1M	Tris-HCl, pH 8,0
Lösung 1	0,1 M	HAc + 0,5 M NaCl, pH 4,8
Lösung 2	0,2 M	HAc + 0,5 M NaCl, pH 2,6
		PBS
	1 mM	HCl

**Das Verfahren****1.) Kopplung von BSA an Sulfo-MBS**

25 mg BSA in 1 ml PBS lösen, 9 mg Sulfo-MBS in ½ ml PBS lösen, beide Lösungen mischen und ½ Stunde rotierend mischen.

Eine PD-10 Säule mit 25 ml 50 mM Natriumphosphat-Puffer (pH= 6,0) äquilibrieren. Dem Kopplungsgemisch 1 ml PBS zufügen, alles in die Säule füllen und eindringen lassen. Dann 4 ml Natriumphosphat-Puffer auftragen und den Durchfluss in acht Aliquots zu je ½ ml auffangen.

Mittels Dot-Blot die Fraktionen auf BSA/Sulfo-MBS Konjugat überprüfen.

**2.) Kopplung des Peptids an BSA/Sulfo-MBS**

5 mg Peptid in 4 ml PBS lösen und ein Aliquot mit BSA/Sulfo-MBS zugeben. Dazu 300 µl 0,5 M Natriumphosphat-Puffer (pH= 7,5), in einem 15 ml-Röhrchen verschließend und bei RT drei Stunden rotierend inkubieren. Einen Dialyseschlauch („cut-off“ bei 12-14 kD) in 100 ml einer EDTA-Lösung (250 mM; pH= 8,0) für 20 min kochen und anschließend mit Wasser spülen.

Die Kopplungslösung in den Schlauch füllen und ü.N. bei 4 °C zwei mal gegen 1,0 l Natriumhydrogencarbonat/NaCl-Puffer dialysieren. Das Dialysat aus dem Schlauch in ein 15 ml Röhrchen überführen und bei 4 °C lagern.

**3.) Kopplung an BrCn-aktivierte Sepharosematrix**

1 g BrCn-aktivierte Sepharose in 1 mM HCl suspendieren und in einem Sarstedt-Röhrchen (15 ml) absinken lassen. Mehrmals mit frischem HCl [1 mM] suspendieren, bei geringer Drehzahl abzentrifugieren und erneut 1 mM HCl zugeben. Zum Schluss die 6 ml aus der Dialyse (Abschnitt 2) zugegeben, und gut verschlossen ü.N. gemeinsam inkubieren.

Am nächsten Morgen eine Polystyrolsäule verschließen, mit Wasser füllen und eine Fritte hineinlegen damit sich keine Luftblasen unter dem Filter sammeln. Die Matrix in die vorbereitete Säule geben und sedimentieren lassen. 250 ml 1 mM HCl zugeben und durchlaufen lassen, Matrix in ein 15 ml-Röhrchen überführen, zentrifugieren (600 U/min, 5'), den Überstand abnehmen.

Dazu 6 ml der Kopplungslösung aus 2) geben, mit Parafilm abdichten und ü.N. bei 4 °C rotieren lassen. Die Matrix sedimentieren (600 U/min, 5', 4 °C), Überstand abnehmen, auf einem SDS-Gel überprüfen, dann die Sepharosematrix mit 10 ml 1 M Tris-HCl, pH 8,0 versetzen und 1 h rotieren lassen. Matrix sedimentieren, vier mal mit je 10 ml PBS waschen und anschließend die Matrix in ein 50 ml-Röhrchen überführen und sedimentieren. Die Überstände jew. sammeln und auf einem SDS-Gel analysieren.

### **2.2.2.3 Affinitätsreinigung polyklonaler Antikörper**

Zu der sedimentierten Matrix aus Kap. 2.2.2.2 20 ml Kaninchenserum geben und ü.N. bei 4 °C rotieren. Sephadex sedimentieren, 3x mit PBS waschen und anschließend in ein Plasiksälchen (Pierce) überführen. Mit PBS auffüllen und mit einer Fritte die Matrix abdecken und überstehendes PBS auslaufen lassen. Dann mit 10 ml Lösung 1 die unspezifisch gebundenen Antikörper eluieren, und bei 4 °C lagern. Mit 12 ml Lösung 2 die spezifischen Antikörper von der Matrix lösen, die ersten zwei Milliliter verwerfen. In einem Röhrchen (15 ml) 200 µl 1 M Tris-Base vorlegen und die nächsten 10 ml darin auffangen. Dialyse der erhaltenen Antikörper gegen zwei mal ein Liter PBS ü.N. bei 4 °C. Den Schlauch zuvor 20 min in 250 mM EDTA, pH 8,0 kochen und gründlich mit H<sub>2</sub>O spülen. Die Affinitätssäule mit weiteren 10 ml Lösung 2 waschen, ausgiebig mit PBS waschen und dann mit 3 mM Na-Azid stabilisieren, abdichten und bei 4 °C lagern.

### **2.2.3 Diskontinuierliche SDS-PAGE**

Die diskontinuierliche SDS-PAGE (LAEMMLI 1970) trennt denaturierte Polypeptide aufgrund ihrer zum Molekulargewicht proportionalen Ladung. Die Abmessung der Gele betrug 8,2 x 10,2 x 0,05 cm. Die Trennung erfolgte über eine Distanz von 5 cm, das Sammelgel hatte eine Lauflänge von 1,3 cm. Das Taschenvolumen bei analytischen Gelen betrug 20 µl, präparative Gele konnten mit 50 µl beladen werden.

Das **Trenngel** bestand aus folgenden Komponenten:

1040 µl	H <sub>2</sub> O
622 µl	Trenngelpuffer
830 µl	Acrylamid-Bis-Lösung
20 µl	APS-Lösung
1 µl	TEMED (Sigma, Deisenhofen)

Vortexen und in die Glasapparatur pipettieren

mit 1 ml H<sub>2</sub>O überschichten

nach 30' das Wasser abgießen

den Glasplattenzwischenraum mit Filterpapier trocknen

Dann das **Sammelgel** vorbereiten:

744 µl	H <sub>2</sub> O
200 µl	Sammelgelpuffer
156 µl	Acrylamid-Bis-Lösung
24 µl	APS-Lösung
1,2 µl	TEMED

Einen Kamm zwischen die Glasplatten klemmen und das flüssige Sammelgel blasenfrei einfüllen. Nach einer halben Stunde den Kamm entfernen und in die Gelapparatur einbauen.

Die Proteinproben wurden in 1x SDS-Probenpuffer mit 5 % β-ME 5 Minuten aufgekocht, kurz auf Eis gestellt und sofort in die Taschen des Gels eingefüllt. Zur Abschätzung des Molekulargewichts der zu analysierenden Proben wurde in die äußeren Taschen ein standardisiertes Gemisch verschiedener Proteine gefüllt. Es wurde immer der „Low-Molecular-Weight-Marker“ von Pharmacia Biotech verwendet: Das lyophilisiert vorliegende Proteingemisch wurde in 500 µl 1x Probenpuffer (5 % β-ME) gelöst, 5 min gekocht und bei 4 °C kurzfristig bis zur Verwendung gelagert. Für analytische Gele wurden 5 µl aufgetragen. Die Elektrophorese erfolgte bei konstanter Stromstärke von I=12 mA für 1 ½ Stunden. Die Gelapparatur wurde auseinander gebaut, die Glasplatte getrennt und das Sammelgel abgeschnitten. Nach Färbung in Coomassie-Blue-Lösung (20 ') musste das Gel mindestens 20 bis 60 Minuten entfärbt werden. Die Gele wurden mit einer Digitalkamera fotografiert. Die Bilder wurden gespeichert und ausgedruckt.

#### 2.2.4 Dot-Blot

Für die qualitative Analyse des Proteingehalts von wässrigen Lösung, in der Regel stammten diese aus einer Säulenchromatografie, wurden Dot-Blots durchgeführt. Zwei Mikroliter der proteinhaltigen Lösung trockneten auf Whatman-Papier 3 MM, welches

anschließend 5' in Coomassie-Blue gefärbt und dann eine Stunde entfärbt wurde. Anhand der verbleibenden Blaufärbung an den aufgetropften Stellen, ließen sich die Fraktionen mit dem höchsten Proteingehalt bestimmen.

### 2.2.5 Western-Blot Analyse

Der Transfer der im SDS-Gel getrennten Proteine auf eine Nitrozellulosemembran erfolgte in einem horizontalen Semi-Dry-System. Nacheinander wurden auf die Anode folgende Lagen gelegt:

6 Blatt Whatman 3 MM mit Transferpuffer 1,  
3 Blatt Whatman 3 MM mit Transferpuffer 2,  
Nitrozellulosemembran mit Transferpuffer 2,  
das ungefärbte SDS-PAGE,  
6 Blatt Whatman 3 MM mit Transferpuffer 3.

Benötigte Lösungen:

Transferpuffer 1	0,3 M 20 % (v/v)	Tris-HCl, pH 10,4 Methanol
Transferpuffer 2	25 mM 20 % (v/v)	Tris, pH 10,4 Methanol
Transferpuffer 3	40 mM 25 mM 20 % (v/v)	Aminocaprinsäure Tris-HCl, pH 9,4 Methanol

Die Kathode wurde auf den Stapel gelegt und mit einer Eisschüssel beschwert. Der Transfer erfolgte eine Stunde bei  $I=100$  mA. Dann wurde die Membran 10 Minuten in Ponceau-Lösung gefärbt und kurz in mit 2 Tropfen HCl angesäuertem  $H_2O$  entfärbt. Die nun sichtbaren Banden wurden mit Kugelschreiber markiert und eingescannt. Anschließend wurde die Membran vollständig in 0,2 % PBST entfärbt und blockiert (2% BSA in 0,1 % PBST; 1 h bei RT und ü.N. bei 4 °C). Am nächsten Morgen je 5 Minuten in 0,1 % PBST und PBS gespült und dann mit dem primären Antikörper (IFA) inkubiert (1 : 30 in 4 % Magermilchpulver in 0,2 % TBST; 1 h; RT). Die Membran wurde anschließend vier mal in 0,2 % TBST für jeweils 15 Minuten gewaschen und anschließend mit dem zweiten Antikörper inkubiert (1 : 20.000 in 4 % Magermilchpulver in 0,2 % TBST; 1 h; RT). Die Detektion der gebundenen Antikörper erfolgte mit dem ECL-Chemilumineszenz-Kit (Amersham) gemäß den Angaben des Herstellers.

### 2.2.6 Mikrosequenzierung von Proteinen und Peptiden

Der Edman-Abbau sowie die Analyse der HPLC-gereinigten Peptide wurde mit dem automatischen Protein Sequencer „Procise“ (Perkin-Elmer) durchgeführt.

## 2.2.7 Expression rekombinanter Proteine

Die funktionelle Analyse eines isolierten Proteins schließt die rekombinante Expression ein. Dazu wurden in die isolierten cDNA-Moleküle an den Enden mit Mutagenese-Primern spezielle Restriktionsstellen eingebaut, die eine direktionale Klonierung in einen speziellen Expressionsvektor ermöglichten. Mit diesem Konstrukt transfizierte Bakterienzellen exprimierten das Protein. Es wurde anschließend isoliert, gereinigt und stand für weitere Experimente zur Verfügung. In dieser Arbeit wurde es benutzt um zu testen, ob das Protein *in vitro* homopolymere Filamente bilden kann.

### 2.2.7.1 Induktion

Wie in Kap. 2.3.11 beschrieben, wurden die chemokompetenten Bakterienzellen mit der in den Expressionsvektor klonierten cDNA-Sequenz transformiert, selektiert (Kap. 2.3.4.2) und eine Flüssigkultur ü.N. in 3 ml SOB/Amp-Medium erstellt. Am nächsten Morgen wurde die Kultur im Verhältnis 1:5 mit 2 ml frischem SOB-Medium gemischt. Nach einer Stunde bei 37 °C wurde die Proteinexpression mit 1 mM IPTG induziert. Ein Aliquot (500 µl) dieser Kultur wurde nach vier Stunden pelletiert und die Zellen in 50 µl Laemmli-Puffer (incl. 5 % β-ME) resuspendiert, fünf Minuten lang gekocht, zentrifugiert und 5 µl dieser Lösung auf einem SDS-PAGE (Kap. 2.2.3) analysiert. Nach erfolgreicher Detektion einer Proteinbande in der erwarteten Größenordnung wurde eine neue Kultur in 50 ml SOB-Medium mit den aus der analytischen Expression verbliebenen Zellen inokuliert und die Expression analog dem analytischen Verfahren in 250 ml fortgesetzt. Vier Stunden nach der Induktion wurden die Zellen zentrifugiert und eingefroren.

### 2.2.7.2 Inclusionbody Präparation

#### Geräte:

Dounce-Homogenisator (20 ml) mit Stempel

Sorvall-Zentrifuge, SS34-Rotor

#### Lösungen:

IB-Lysispuffer (25 ml)	50 mM	Tris/HCl, pH=8,0
	25 %	Sucrose
	1 mM	EDTA, pH=8,0
	0,5 % (w/v)	Lysozym
	½ Tablette	Complete Proteaseinhibitor (Boehringer Mannheim)
	add 25 ml	H <sub>2</sub> O

Detergentpuffer (50 ml)	1 %	Nonidet P 40
	1 %	Deoxycholic Acid (DOC)
	20 mM	Tris/HCl, pH 7,5
	2 mM	EDTA, pH 7,5
	0,2 M	NaCl
	1 Tablette add 50 ml	Complete H <sub>2</sub> O
IB-Waschpuffer (25 ml)	10 ml (0,5 %)	Triton (20 %ig)
	0,8 ml (1 mM)	EDTA [0,5 M], pH 8,0
	½ Tablette	Complete
	add 25 ml	H <sub>2</sub> O
Mono-S UREA-Puffer	255,25 g (8,5 M)	Harnstoff
	add 500 ml	H <sub>2</sub> O
Phosphatpuffer 10 mM, pH 6,6	5,05 g	NaH <sub>2</sub> PO <sub>4</sub> (2 x H <sub>2</sub> O)
	3,13 g	Na <sub>2</sub> HPO <sub>4</sub> (2 x H <sub>2</sub> O)
	1 mM	b-Mercaptoethanol (spaltet Disulfidbrücken in cysteinhaltigen Proteinen)
	add 500 ml	H <sub>2</sub> O

**Protokoll:**

Das gefrorene Bakterienpellet (Kap. 2.2.7.1) wurde im Dounce-Homogenisator in Lysis-Puffer vollständig suspendiert. Danach wurde der Detergenzpuffer zugegeben, gründlich gemischt und für 10 min auf Eis inkubiert. Nach Zugabe von Benzonase (3U/ml Endkonzentration), einem hochaktiven DNA/RNA spaltenden Enzym, sollte nach 30 min eine deutliche Abnahme der Viskosität eintreten. Gegebenenfalls wurde die Enzymkonzentration auf 5 U/ml erhöht. Die homogene Lösung wurde aliquotiert, zentrifugiert (SS34-Rotor, 10.000 U/min, 4 °C, 25 min), die Überstände verworfen und die Pellets vereinigt. Das Präzipitat wurde zwei mal in IB-Waschpuffer homogenisiert und wieder zentrifugiert. Abschließend mussten die Pellets in Mono-S-Puffer resuspendiert und eine Stunde rotierend bei Raumtemperatur gelöst werden. Die Suspension wurde mit einem Spritzenfilter (Porengröße 22 µm) sterilisiert und direkt in einer Harnstoff-FPLC gereinigt.

**2.2.7.3 Harnstoff-FPLC****Lösungen** (immer frisch zubereitet)

MONO-S		
Puffer A	8,5 M	Harnstoff, pH 6,6
	10 mM	NaPO <sub>4</sub>
	1 mM	β-ME

Puffer B	8,5 M	Harnstoff, pH 6,6
	10 mM	NaPO <sub>4</sub>
	1 mM	β-ME
	0,4 M	NaCl
MONO-Q		
Puffer A	8,5 M	Harnstoff
	10 mM	Tris-HCl, pH 8,0
	1 mM	β-ME
Puffer B	8,5 M	Harnstoff
	10 mM	Tris-HCl, pH 8,0
	1 mM	β-ME
	0,4 M	NaCl

### Protokoll:

Die MONO-S-Säule wurde benutzt um das in der IB-Präparation eingesetzte Lysozym zu entfernen. Nach dem Äquilibrieren mit MONO-S-Puffer A wurde die Probe auf die Säule geladen. Puffer A wurde fortlaufend durch die Säule gepumpt, bis der Schreiber des Durchflussphotometers ruhig blieb. Dann wurde Puffer B innerhalb von 36 min auf einen Anteil von 50 % zugemischt bei einer Rate von 2 ml/min und einem Druck < 2 MPa. Nachdem die Proteinfractionen gesammelt waren wurden sie in einer SDS-PAGE analysiert. Die Konzentration des Proteins war in drei Fraktionen sehr hoch und sauber, so dass eine weitere Reinigung auf einer MONO-Q-Säule nicht mehr durchgeführt wurde.

## 2.3 Molekularbiologische Methoden

### 2.3.1 Konzentrationsbestimmung von Nukleinsäuren

Die Konzentrationsbestimmung von DNA in wässriger Lösung erfolgte photometrisch bei einer Wellenlänge von 260 nm in einer Quarzglas-Küvette, Hellma (Müllheim, D). Das Volumen betrug stets 250 µl bei einer Schichtdicke von 1 cm. Dabei gilt folgender Zusammenhang:

Die  $OD_{260}=1,0$  entspricht 50 µg doppelsträngiger DNA (SAMBROOK ET AL. 1989). Bei RNA sind es 40 µg. Das Verhältnis der Absorption ( $^{260}/_{280}$ ) sollte im besten Fall bei 1,8 liegen.

## 2.3.2 RNA-Isolierung

### 2.3.2.1 Gewinnung von gesamt-RNA aus Tiergeweben

Um cDNA-Banken verschiedener Tiere und Gewebetypen herzustellen, wurden 0,5 g tiefgefrorenes Gewebe mit einem Hammer zertrümmert und im auf -80 °C vorgekühlten Mörser zu sehr feinem Pulver verrieben. Nach Überführung in ein 50 ml Reaktionsgefäß mit 5 ml TRIzol-Reagenz (Gibco-BRL, Eggenstein) wurde das Gemisch gründlich mit dem UT-Dispergierwerkzeug (Janke & Kunkel, Staufen) homogenisiert. Das Homogenat wurde in fünf Teile aliquotiert. Die weitere Verarbeitung folgte den Angaben des Herstellers.

### 2.3.2.2 Anreicherung von poly A<sup>+</sup>RNA aus Gesamt-RNA

modifiziert nach AVIV & LEDER (1972)

Aus der zuvor isolierten gesamt-RNA wurde in Vorbereitung einer cDNA-Synthese mit oligo(dT)-Zellulose (Stratagene, Heidelberg) die poly A<sup>+</sup>RNA angereichert. Alle Lösungen wurden mit DEPC behandeltem H<sub>2</sub>O hergestellt. Alle Zentrifugationen wurden in einer Tischzentrifuge bei 6.000 U/min für jeweils eine Minute durchgeführt.

<b>Puffer 1</b>	500 µl	5 M NaCl
	250 µl	1 M Tris-HCl (pH 7,5)
	250 µl	0,5 M EDTA (pH 7,5)
	250 µl	10 % (w/v) Sarcosyl
	3,75 ml	H <sub>2</sub> O

<b>Puffer 2</b>	500 µl	5 M NaCl
	250 µl	1 M Tris-HCl (pH 7,5)
	250 µl	0,5 M EDTA (pH 7,5)
	4,0 ml	H <sub>2</sub> O

Die Gesamt-RNA aus 100 mg Gewebe wurde in 450 µl H<sub>2</sub>O<sub>DEPC</sub> gelöst. Nach Zugabe von je 30 µl Tris, EDTA und Sarcosyl sechs Minuten bei +70 °C erhitzt und anschließend eine Minute in Eiswasser gekühlt. Währenddessen wurden 50 mg oligo(dT)-Zellulose mit 1 ml Puffer 1 äquilibriert, zentrifugiert und der Überstand verworfen.

Nach dem Abkühlen der Probe wurden 65 µl NaCl zugesetzt und mit der feuchten oligo(dT)-Zellulose gemischt. Die poly A<sup>+</sup>RNA band im Verlaufe von 15 Minuten bei kontinuierlichem Mischen an die Zellulose. Die Zellulose mit den gebundenen mRNA-Molekülen wurde sedimentiert und drei mal mit je 1 ml Puffer 1 gewaschen. Die ÜS wurden auf Eis gelagert.



Anschließend wurde die Zellulose drei mal mit jeweils 1 ml Puffer 2 gewaschen. Die RNA wurde fünf mal mit 100 µl H<sub>2</sub>O<sub>DEPC</sub> (+70 °C) herausgewaschen und mit 0,25 M NaCl und 2 Volumen Ethanol ü.N. bei -20 °C präzipitiert und abzentrifugiert (SS34, 14.000 UpM, 4 °C, 45'). Die mRNA wurde in 40 µl H<sub>2</sub>O<sub>DEPC</sub> aufgenommen, 2,5 µl wurden auf einem Agarosegel analysiert, die verbliebenen 37,5 µl bei -80 °C gelagert oder direkt für die Konstruktion einer λ-cDNA-Bank eingesetzt.

### **2.3.3 Arbeiten mit λ-cDNA-Banken**

#### **2.3.3.1 Erstellung von λ-cDNA-Banken**

Genbanken auf der Grundlage von cDNA bieten die Möglichkeit einen Großteil der Gene, die in einem Organismus oder speziellen Gewebe exprimiert werden aufzubewahren und zu untersuchen. Das „ZAP Express cDNA Synthesis Kit“ und das „ZAP Express cDNA Gigapack Gold Cloning Kit“ von Stratagene (Heidelberg) wurde für die Erstellung der Genbanken benutzt. Die Prozedur folgte weitgehend den Angaben des Herstellers:

Nach Umschreiben der poly A<sup>+</sup>RNA in doppelsträngige cDNA und Ligation mit EcoRI-Adaptoren wurden die DNA-Moleküle asymmetrisch mit den Restriktionsenzymen BamHI und EcoRI geschnitten. Die abgetrennten Adaptoren und kleinere DNA-Spezies wurden, in Abweichung vom Herstellerprotokoll, auf einem präparativen Agarosegel von den größeren Molekülen getrennt und diese aus dem Gelstück eluiert (Kap. 2.3.7). Anschließend wurde die cDNA an die λ-Arme ligiert und mit dem „Gigapack Gold III“-Kit in Phagenpartikel verpackt. Durch Infektion der λ-kompetenten *E. coli* XL1-blue MRF'-Zellen und Plattierung auf zehn bis 15 großen Petrischalen wurde die Genbank amplifiziert und die nach Lyse der Bakterien freigesetzten Phagen mit jeweils 10 ml SM-Puffer abgewaschen. Abgelöste Bakterien und Agarbestandteile wurden in der Zentrifuge pelletiert, der phagenhaltige Überstand mit 0,7 % CHCl<sub>3</sub> stabilisiert und bei +4 °C gelagert. 10 ml der Genbank wurden mit 7 % DMSO versetzt, aliquotiert, anschließend schockgefroren und bei -80 °C gelagert.

#### **2.3.3.2 λ-kompetente *E. coli* Zellen und Plattierung von λ-cDNA-Banken**

Nach konfluentem Wachstum ü.N. von *E. coli* XL1-blue MRF'-Zellen auf einer SOB/Tetrazyklin-Platte wurden 25 ml SOB/Mg/Maltose-Medium inokuliert. Die Zellen wuchsen drei Stunden bei 37 °C und 250 U/min im Schüttelinkubator. Nach Zentrifugation (4 °C, 2500 U/min, 15', Varifuge RF) wurden sie in 10 mM MgSO<sub>4</sub> aufgenommen und die OD<sub>600</sub> auf 0,5 eingestellt. Für die Plattierung auf kleinen (großen) Petrischalen wurden

240 µl (600 µl) Bakteriensuspension in 14 ml-Röhrchen mit maximal 120 µl (300 µl) Phagenlösung gemischt und für 15' bei 37 °C inkubiert. Anschließend wurden 4 ml (10 ml) warme NZY-Top-Agarose (45 °C) zugegeben, kurz gemischt und dann gleichmäßig auf den Petrischalen verteilt. Nach weiteren 15' wurde kondensiertes Wasser von den Deckeln gewischt und die Platten bei 37 °C ü.N. gelagert. Nach erfolgter Plaquebildung wurden die Platten bei 4 °C bis zur Analyse gelagert. Abhängig von der benötigten Dichte der Plaques wurden die Genbank sowie primäre und sekundäre Isolate titriert und ggf. mit SM-Puffer verdünnt.

### 2.3.3.3 Isolierung von IF-Kandidaten aus $\lambda$ -cDNA-Banken mit monoklonalen Antikörpern

Ein „Expressionsscreen“ ermöglicht das Auffinden von Genen anhand der synthetisierten Proteine. Dabei ist nicht die genaue DNA-Sequenz, sondern ein kurzes Peptid ausschlaggebend für die Reaktivität mit einem (monoklonalen) Antikörper. Die von Phagen infizierten Bakterien exprimieren die fremden Gene, diese werden bei der Lyse freigesetzt und können mit einem AK detektiert werden.

Die Prozedur folgte der Beschreibung in Kap. 2.3.3.2, mit einer Phagendichte von ca. 18.500 pfu/Platte wurden gute Resultate erzielt (10 große Petrischalen). Sechs bis sieben Stunden nach der Plattierung bildeten sich Plaques im konfluenten Bakterienrasen. Zu diesem Zeitpunkt wurden Nitrozellulosefilter 15 Minuten in 10 mM IPTG äquilibriert, auf Filterpapier luftgetrocknet und blasenfrei auf die Agaroseoberfläche der Platten gelegt. Nach erfolgter Induktion wuchsen die Zellen für weitere 12 Stunden. Am nächsten Morgen mussten die Kulturplatten zwei Stunden bei 4 °C gekühlt und die mit Punkten markierten und nummerierten Filter abgezogen werden. Die anhaftenden Proteine wurden zwei Stunden im Vakuumofen bei +80 °C immobilisiert.

Der Antikörperscreen benötigte folgende Lösungen (Abweichungen für den *Saccoglossus*-Screen in Klammern):

0,2 % und 0,1 % Tween 20 in PBS  
2 % BSA (5 % Magermilchpulver) in PBS;  
2 % BSA (5 % Magermilchpulver) + 0,1 % Tween 20 in PBS;

IFA-AK 1:30 in 2 % BSA (5 % Magermilchpulver) in PBS

Schwein- $\alpha$ -Maus-AK 1:500 in 2 % BSA (5 % Magermilchpulver) in PBS

Streptavidingekoppelte Meerrettichperoxidase (HRP) 1:1000 in 2 % BSA  
(5 % Magermilchpulver) in PBS

Farbreaktionslösung	190 ml	PBS
	12 ml	Chlornaphtol (0,3 %) in Methanol
	220 µl H <sub>2</sub> O <sub>2</sub>	

### Prozedur:

2 x 15' in 400 ml PBS/0,2 % Tween 20, alle Filter gemeinsam

2 x 1 h in je 7,5 ml 2 % BSA/PBS mit 0,1 % Tween 20; 2.

Inkubation auch ü.N. bei 4 °C möglich. Für jeden Nitrozellulosefilter wurde eine Petrischale verwendet.

15' in 400 ml 0,1 % Tween 20/PBS

15' in 400 ml PBS

1,5 h in je 7,5 ml IFA-AK-Lösung bei RT (ü.N. bei +37 °C)

15' in 400 ml 0,1 % Tween 20/PBS

15' in 400 ml PBS

1,5 h in je 7,5 ml α-Maus-AK-Lösung

15' in 400 ml 0,1 % Tween 20/PBS

15' in 400 ml PBS

1,5 h in je 7,5 ml Streptavidin-HRP-Lösung

15' in 400 ml 0,1 % Tween 20/PBS

15' in 400 ml PBS

Die Filter auf Chromatografiepapier zwei Minuten antrocknen lassen und die Farbreaktion in je 20 ml Farbreaktionslösung ablaufen lassen, gelegentlich kontrollieren und die Reaktion in Wasser abstoppen. Dann die Filter vollständig trocknen lassen, positive Signale markieren, auf eine Kopierfolie übertragen und die entsprechenden Bereiche in der Petrischale ausstechen und den Agar in 300 µl SM-Puffer lagern. Diese primären Isolate werden nach 12-stündiger Inkubation bei +37 °C titriert, ca. 3000 pfu werden einem weiteren Expressionsscreen unterzogen und die sekundären Isolate mit 300 pfu schließlich so weit vereinzelt, dass frei liegende Plaques ausgestochen werden können. Diese werden dann in 100 µl aufgenommen und ihre Inserte amplifiziert (2.3.4.5), kloniert (2.3.10, 2.3.11) und sequenziert.

#### 2.3.3.4 cDNA-Screen von λ-cDNA-Genbanken

Eine Alternative zum Antikörperscreening stellt das Untersuchen einer cDNA-Bank mit einer cDNA-Sonde dar. Das Verfahren für sechs Petrischalen zur Plattierung der Phagen war wie in Kap. 2.3.3.2 beschrieben. Die Dichte der Phagenpartikel wurde auf 300.000 pfu/Platte eingestellt.

Die unbehandelten Nitrozellulosefilter wurden erst am nächsten Morgen auf die gekühlte Agaroseschicht gelegt. Nach der Aufbringung von Positionsmarken wurden die Filter abgezogen, 15 Minuten luftgetrocknet und die anhaftende DNA mit 120 mJoule UV-Licht kovalent an die Nitrozellulosemembran gebunden (Stratalinker, Stratagene).

Die auf den Filtern gebundene DNA wurde 2x5 Minuten auf nassem Whatman 3 MM-Papier denaturiert (0,5 M NaOH; 1,5 M NaCl) und anschließend 2x5 Minuten renaturiert (1 M Tris; 1,5 M NaCl; pH 6,8). Nach einem Waschschrift (20 min bei 70 °C; 5xSSC; 0,1 % SDS) wurden die Filter luftgetrocknet, paarweise in Beuteln verschweisst und 2 Stunden bei 37 °C in jeweils 5 ml Prähybridisierungslösung inkubiert. Nach dem Austausch gegen jeweils 2,5 ml Hybridisierungslösung wurden die Filter ü.N. bei 37 °C im Wasserbad geschaukelt.

Nach der Hybridisierung wurden die Filter fünf Minuten in 2 x SSC/0,1 % SDS bei 37 °C gewaschen. Mehr oder weniger stringentes Waschen über zwei mal 15 Minuten löste schwach gebundene DNA-Proben bei 50 °C bis 65 °C und 2 x SSC/0,1 % SDS bis 0,1 x SSC/0,1 % SDS. Es folgte zweimaliges Waschen für fünf Minuten in MAN-Puffer, dann eine halbstündige Inkubation in DIG-Blockierungsreagenz (1 % DIG-Block (w/v) in MAN). Anschließend wurden die Filter 45 Minuten bei RT mit dem anti-DIG-AK (1:500 in 1 % DIG-Block in MAN) inkubiert. Die Filter wurden dann sechs mal zehn Minuten in MAN gewaschen, zwei mal fünf Minuten in TNM äquilibriert und dann fünf Minuten in der CSPD-Lösung (Chemilumineszenz-Substrat 1:100 in TNM) im Dunkeln inkubiert. Die Filter wurden drei Minuten auf Filterpapier getrocknet, auf frisches Filterpapier umgesetzt und dann mit Klarsichtfolie eingeschlagen. Nach einer 15-minütigen Inkubation bei 37 °C wurde ein Röntgenfilm aufgelegt und ü.N. bei RT exponiert. Nach der Filmentwicklung wurden die positiven Signale ermittelt, die entsprechenden Bereiche der Kulturplatte ausgestochen, in SM-Puffer aufgenommen und weiter vereinzelt, bis eine spezifische PCR-Amplifikation des klonierten cDNA-Fragments möglich war.

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

Die enzymatische Polymerase-Kettenreaktion (PCR) wurde benutzt um *in vitro* einen definierten DNA-Bereich zu amplifizieren (MULLIS & FALOONA, 1987). In durchschnittlich 35 Zyklen wurde der von zwei begrenzenden Oligonukleotiden (Primern) definierte DNA-Abschnitt exponentiell vermehrt. Die verschiedenen Anwendungen der PCR sind aufgeführt, für gewöhnlich betrug die Annealingtemperatur 60 °C und es wurde 35 Zyklen lang amplifiziert.

### 2.3.4.1 Genomische PCR

Die Analyse der genomischen Organisation zuvor klonierter Gene war ein wichtiger Bestandteil bei phylogenetischen Studien. Mit aus der cDNA-Sequenz abgeleiteten Primern wurden mit Hilfe der PCR Genabschnitte amplifiziert. Um spezifische Produkte zu erhalten wurden bevorzugt Primer mit einer Länge von 28 nt bis 30 nt ausgewählt. Der Primer sollte einen G/C-Anteil von 50 % nicht unterschreiten, die letzten beiden Basen sollen ebenfalls G oder C sein. Im Reaktionsansatz für 50 µl wurden 400 ng genomischer DNA als „Template“ eingesetzt, die Primermenge lag bei 10 nmol, die Zugabe von 5 % DMSO diente dazu, die Bindung der Polymerase an die DNA zu stabilisieren.

### 2.3.4.2 Kolonie-PCR

Das Testen von transformierten Bakterienkolonien auf die richtige Größe des klonierten Inserts wurde ausschließlich mit der Kolonie-PCR-Methode durchgeführt. Zwei bis vier weiße Kolonien wurden in 10 µl H<sub>2</sub>O suspendiert, davon diente jeweils 1 µl als „Template“ in der PCR-Reaktion. Ein gemeinsamer Reaktionsansatz („MasterMix“) wurde vorbereitet und in zehn Aliquots zu je 9 µl verteilt.

PCR-Ansatz für 10 Kolonien:

10 µl	10 x ExTaq-Puffer (Mg <sup>2+</sup> : 20 mM)
8 µl	dNTP (G-,A-,T-,CTP : jeweils 2,5 mM)
1 nmol	Sense Primer
1 nmol	Antisense Primer
2,5 U	TaKaRa ExTaq-DNA-Polymerase
add 90 µl	H <sub>2</sub> O

Temperaturprogramm der Mini-PCR (35 Zyklen):

Temp.:	Dauer:
95 °C	1'45"
95 °C	15"
55 °C	15"
72 °C	1' bis 4' je nach Größe des klonierten Inserts
4 °C	

Die restlichen 9 µl Bakteriensuspension verblieben bis zur Auswertung der PCR-Ergebnisse bei 4 °C. Nach Analyse auf einem Agarosegel (Kap. 2.3.6) wurden jene Klone, deren Reaktionen keine Nebenprodukte bei gleichzeitig hoher Produktmenge zeigten für eine 5 ml ü.N.-Kultur in SOB/Amp-Medium vorbereitet.

### 2.3.4.3 Mutagenese-PCR

Für die Proteinexpression der Gene mussten diese in ein anderes Vektorsystem umklont werden (pCR TOPO 2.1→pET23a<sup>+</sup>). Der Einbau spezifischer Restriktionsstellen in die PCR-Primer ermöglichte die Umklonierung unter Berücksichtigung des durch den Expressionsvektor definierten Leserasters. Die so gewonnenen modifizierten PCR-Produkte wurden zunächst TA-kloniert, und dann in einer Midi-Präparation (2.3.12.2) vermehrt.

### 2.3.4.4 Halbspezifische PCR-Reaktion

Die mit cDNA-Sonden oder dem IFA-Antikörper isolierten Klone waren am 5'-Ende z.T. nicht vollständig (*Saccoglossus*) oder sie lagen als Fusionsprodukt von zwei verschiedenen cDNA-Molekülen vor. Die vollständige cDNA-Sequenz musste gefunden werden, mindestens bis zum Beginn der kodierenden Sequenz. Weil bei der Konstruktion der verwendeten cDNA-Expressionsbanken das 5'-Ende der cDNA-Moleküle immer nahe der Eco-RI Schnittstelle lag (direktionale Klonierung) konnte ein bestimmter Forward-Primer eingesetzt werden. Dieser war lediglich ein verlängerter T3-Primer (GCGCG-CCTGCAGGTCGACACTAGTGG), die Bindung erfolgte unspezifisch an alle im Reaktionsansatz vorhandenen cDNA-Moleküle. Ihm gegenüber stand ein für das gesuchte Molekül sequenzspezifischer Gegenstrangprimer. Die multiplen PCR-Produkte wurden gereinigt, das größte Produkt wurde kloniert und sequenziert. Dieses beinhaltete stets die fehlenden Sequenzbereiche.

### 2.3.4.5 PCR auf $\lambda$ -Phagen

Für die Amplifizierung der cDNA-Fragmente auf isolierten Phagen wurden 8 % Phagenisolat als „Template“ eingesetzt. Die Primer waren immer der T3 (5'-AATTA-ACCCTCACTAAAGGG-3') und T7 (5'-GTAATACGACTCACTATAGGG C-3') Primer. Die Dauer der Strangsynthese bei 72 °C lag bei 4 Minuten. Die Reaktionsprodukte wurden auf einem Agarosegel analysiert und präparativ gereinigt, siehe Kap. 2.3.7.

### 2.3.4.6 Synthese von DNA-Sonden mit DIG-dUTP

Die Synthese von DNA-Sonden für Hybridisierungsexperimente erfolgte über 40 Zyklen ausschließlich auf Plasmid-DNA. Im Reaktionsansatz wurden die Standardnukleotide durch den 10-fach „PCR DIG Labeling Mix“ (Boehringer Mannheim) ersetzt. Die PCR-Produkte wurden mit dem QiaQuick Nucleotide Removal Kit (Qiagen)

gereinigt, auf einem Agarosegel analysiert und UV-photometrisch quantifiziert. 1,5 µg wurden für die Hybridisierung auf sechs großen Nitrozellulosefiltern eingesetzt, siehe Kap. 2.3.3.4.

### 2.3.5 Extraktion von genomischer DNA

modifiziert nach BLIN & STAFFORD (1976)

Homogenisationspuffer, 10 ml:

100 µl	1 M Tris-HCL pH 8,0
2 ml	0,5 M EDTA (pH 8,0)
6,4 ml	H <sub>2</sub> O
1 mg/ml	Proteinase K
1 ml	Sarcosyl 10 %

Die primären Verarbeitungsschritte waren indentisch mit der in Kap. 2.3.2.1 beschriebenen Prozedur. Das fein zerriebene Gewebepulver wurde aus dem Mörser direkt in den auf 58 °C vorgewärmten Homogenisationspuffer überführt und nach Zugabe der Proteinase durch vortexen gemischt. Nach Zugabe von Sarcosyl verlief die Lysisreaktion über drei Stunden bei 58 °C und gelegentlichem vorsichtigen Mischen. Die DNA konnte nun bei 4 °C gelagert werden.

Die reine genomische DNA wurde in Aliquots zu je 2 x 600 µl in 2 ml-Safe-Lock Reaktionsgefäßen extrahiert. Die Lösung wurde auf eine Konzentration von 0,25 mol/l NaCl eingestellt und mit 1 Volumen puffergesättigtem Phenol pH > 7,5 (USB, Cleveland, Ohio) zwei Minuten vorsichtig gemischt. Nach Zentrifugation (1', 12.000 g, RT) wurde die obere wässrige Phase noch zwei mal analog extrahiert. Anschließend wurde der Überstand noch zwei mal mit CHCl<sub>3</sub>/Isoamylalkohol (24:1) gereinigt und zentrifugiert. Die Überstände wurden vereinigt und mit die DNA mit zwei Volumen Ethanol gefällt. Das Knäuel aus genomischer DNA wurde mit einem Glashaken aufgenommen, kurz in Ethanol (70 %) gewaschen, an der Luft getrocknet und dann in 60 µl TE-Puffer bei 37 °C gelöst. Die DNA wurde gebrauchsfertig bei 4 °C gelagert oder bei -20 °C in 0,25 mol/l NaCl und zwei Volumen Ethanol gefällt und eingefroren.

### 2.3.6 Agarosegelelektrophorese

Die Größenfraktionierung von DNA-Molekülen in Agarosegelen (1 x TBE-Puffer; 1 % (w/v) Agarose; 0,5 µg/ml Ethidiumbromid) durch ein elektrisches Gleichspannungsfeld ermöglicht sowohl eine Abschätzung der Größe, als auch der Menge an vorhandener DNA in wässriger Lösung.

Analytische Gele hatten ein Volumen von 25 ml und verfügten über acht Taschen mit maximal 20 µl Ladevolumen. Nach 60 bis 90 Minuten bei einer Spannung von 60 V wurde das Gel unter UV-Licht analysiert und mit einer Digitalkamera dokumentiert.

Präparative Agarosegele hatten ein Volumen von 50 ml und fünf Taschen mit einer Kapazität von 60 µl. Zur Klonierung bestimmte PCR-Produkte wurden 75 Minuten lang bei 125 V getrennt.

5-fach Probenpuffer:	0,5 M	Tris
	50 mM	CDTA
	2,5 % (w/v)	Sarcosyl
	50 % (w/v)	Glycerin
	0,05 %	Bromphenolblau;
		pH 8,3

Längen- und Massenstandard: GeneRuler DNA Ladder Mix, MBI Fermentas  
(St.Leon-Rot, F).

### 2.3.7 DNA-Extraktion aus Agarosegelen

Für TA-Klonierungen wurden PCR-Amplifikate auf einem Agarosegel getrennt und die beabsichtigten DNA-Banden von den Nebenprodukten durch Ausschneiden getrennt. Die DNA wurde mit dem „QiaQuick Gel Extraction Kit“ (Qiagen, Hilden) protokollgemäß isoliert.

Für die Erstellung von  $\lambda$ -cDNA Genbanken wurden größere cDNA Mengen benötigt. Das ausgeschnittene Gelstück wurde mit einem Skalpell zerschnitten und durch eine mittelgroße Kanüle in Aliquots von je ca. 300 mg gequetscht. Nach Zugabe von 2 Vol. TC-Puffer und 0,5 % Sarcosyl wurde das Gel weiter zerkleinert und ü.N. bei 37 °C inkubiert. Danach erfolgte die Abtrennung der DNA mit Phenol und Chloroform wie in Kap. 2.3.5 beschrieben. Das Gesamtvolumen der wässrigen cDNA-Lösung wurde mit Isobutanol in zwei Aliquots auf je 600 µl eingeeengt und die cDNA mit 0,25 M NaCl und 2 Volumen Ethanol ü.N. bei -20 °C gefällt und anschließend pelletiert.

### 2.3.8 Fällung von Nukleinsäuren

Zur Präzipitation von DNA- und RNA-Molekülen wurde die Salzkonzentration mit der 5 M NaCl-Lösung auf 0,25 M eingestellt und die Nukleinsäuren mit 2 Volumenteilen absolutem Ethanol gemischt. DNA-Lösungen wurden 30 Minuten auf Eis, RNA-Lösungen über Nacht bei -20 °C gekühlt und anschließend 45 Minuten bei 4 °C und mindestens 24.000 g zentrifugiert. Das Pellet wurde zwei mal mit 2 Ausgangsvolumen 75 %igem Ethanol (RT) gewaschen. Bei gut sichtbaren Pellets wurde die DNA mit dem Ethanol durch vortexen gemischt und bei RT in einer Tischzentrifuge pelletiert. RNA-Präzipitaten wurde



der mit  $\text{H}_2\text{O}_{\text{DEPC}}$  gemischte 75 %ige Ethanol vorsichtig zugesetzt und unmittelbar ohne weiteres vortexen zentrifugiert. RNA-Pellets wurden im Heizblock bei 37 °C, DNA-Pellets in einer Vakuumzentrifuge getrocknet und im geeigneten Volumen aufgenommen.

### 2.3.9 Restriktionsverdau

Restriktionsendonukleasen spalten DNA hochspezifisch an charakteristischen Erkennungssequenzen. So können Gensequenzen unter Einhaltung des Leserasters in gleichermaßen geschnittene Expressionsvektoren kloniert werden, um eine korrekte Synthese des rekombinanten funktionsfähigen Proteins zu gewährleisten.

Dazu wurden je 1 µg Expressionsvektor (pET 23a<sup>+</sup>) und 1 µg klonierte Plasmide in einem Gesamtvolumen von 10 µl geschnitten. Die Enzyme BamHI, NdeI und der Reaktionspuffer ReAct 3 stammten von Gibco-BRL. In diesem Reaktionspuffer konnten beide Enzyme gleichzeitig eingesetzt werden, da die Effizienz hinreichen groß war und lediglich die Inkubationszeit von vier auf 16 Stunden bei 37 °C verlängert wurde. Die verbliebenen Vektorfragmente und das herausgeschnittene Insert wurden in einem präparativen Agarosegel (Kap. 2.3.7) voneinander getrennt, eluiert und für die Ligation (Kap. 2.3.10) aufbewahrt.

### 2.3.10 Ligation von DNA-Molekülen

DNA-Moleküle mit einfach überstehenden Adenylatresten wurden mit dem „TA-TOPO“ Klonierungssystem (Invitrogen) protokollgetreu ligiert. Produkte mit kohäsiven Enden aus einer Reaktion mit Restriktionsenzymen wurden von einer T4-Ligase (Invitrogen) mit einem gleichermaßen geschnittenen und gelgereinigten Vektor verbunden. Es wurde darauf geachtet, dass das benötigte ATP nicht öfter als zwei mal aufgetaut worden war, da sich dann bereits eine geringere Ligationseffizienz zeigte. Das Reaktionsvolumen betrug 10 µl, das Verhältnis von Insert zu Vektor ungefähr 1:1, die Reaktion verlief über 14 Stunden bei 14 °C. Die Ligationsreaktionen bei der Erstellung von λ-cDNA Banken dauerten 36 Stunden bei 4 °C.

### 2.3.11 Transformation von *E. coli* Zellen

Für Routineklonierungen von PCR-Produkten zu Vermehrung der DNA wurden *E. coli* OneShot TOP10F'-Zellen (Invitrogen) protokollgemäß verwendet. Mit 2 µl der Ligrationsreaktion konnte eine ausreichende Zahl von transgenen Zellen erzeugt werden. Nach dem Hitzeschock (30 Sek.; 42 °C) wurden die Zellen für fünf Minuten auf Eis gekühlt und 250 µl SOC-Medium zugefügt. Nach einstündiger Inkubation bei 37 °C im Schüttelinkubator („Recovery“) wurden 50 µl und 200 µl Zellsuspension auf zwei mit

jeweils 40 µl IPTG und X-Gal behandelten SOB/Amp-Agarplatten verteilt und ü.N. bei 37 °C wachsen gelassen. Die Transformation der speziell für Proteinexpressionen geeigneten *E. coli* BL21-Zellen verlief analog.

### **2.3.12 DNA-Isolierung**

Plasmid-DNA von guter Qualität wird benötigt um Sequenzierungen und erneute Transformationen durchzuführen. Die Plasmide wurden mit Qiagen-Produkten extrahiert. Bis zu einer Ausbeute von 20 ng Plasmid-DNA wurde das „Qiagen Plasmid Mini Kit“ verwendet, höhere Ausbeuten konnten mit dem „Qiafilter Plasmid Midi Kit“ erzielt werden.

#### **2.3.12.1 Analytische Plasmidisolierung (Mini-Präparation)**

Für kurze Sequenzierungen wurden kleine Plasmid-DNA Mengen mit dem „Qiaprep Spin Miniprep Kit“ isoliert. Die Bakterienzellen einer vereinzelt liegenden Kolonie wuchsen in 5 ml LB/Amp-Medium bei 37 °C ü.N. im Schüttelinkubator bei 250 U/min heran. Die Zellen wurden pelletiert und die DNA den Angaben des Herstellers folgend isoliert. Die DNA wurde in 32 µl H<sub>2</sub>O eluiert, ein Mikroliter wurde auf einem analytischen Agarosegel analysiert (Kap. 2.3.6), weitere 2,5 µl wurden im UV-Photospektrometer quantifiziert.

5 ml LB/Amp-Medium (Ampicillin: 100 µg/ml) wurden mit 200 µl der Vorkultur gemischt und wuchsen bei 37 °C bis zu einer OD<sub>600</sub>>1,5 im Schüttelinkubator heran. Anschließend wurden 4,5 ml der Kultur pelletiert (1 min, 20 °C, 13.000 g) und nach den Angaben im Protokoll (QIAGEN, März 1997) weiterverarbeitet. Die eluierte DNA wurde mit 0,7 Volumen absolutem Isopropanol gefällt und abzentrifugiert (30 min, 15 °C, 25.000 g). Nach doppeltem Waschen mit 1 ml Ethanol (70 %) wurde das Pellet bei 37 °C im Heizblock getrocknet und in 50 µl TE-Puffer gelöst. Die Ausbeute und Reinheit der DNA wurde photometrisch bestimmt.

#### **2.3.12.2 Präparative Plasmidisolierung (Midi-Präparation)**

Die Ermittlung vollständiger cDNA und Gensequenzen auf beiden Strängen setzte größere DNA-Mengen voraus. Diese wurden mit dem „QIAfilter Plasmid Midi-Kit“ gewonnen. 25 ml LB/Amp-Medium wurden mit 250 µl der Mini-Kultur inokuliert und wuchsen 18 Stunden bei 30 °C und 250 U/min im Schüttelinkubator. Die Zellen wurden pelletiert und protokollgemäß verarbeitet. Die DNA wurde in 5 ml eluiert, mit 3,5 ml Isopropanol gefällt und im Falconröhrchen (14 ml) pelletiert (30 min, SS34-Rotor, 4 °C, 7000 U/min). Das Präzipitat wurde zwei mal mit 2 ml Ethanol (70 %) gewaschen, zentrifugiert, und in 250 µl TE-Puffer aufgenommen und analysiert (Kap. 2.3.12.1).

## **2.4 Datenverarbeitung**

Sequenzanalyse: GCG Wisconsin-Package V10.1

Bildbearbeitung: The GIMP 1.2 (GIMP community)

Textverarbeitung und Grafiken: OpenOffice 1.0.1 für Linux

Betriebssystem: SuSE-Linux 7.3 für x86 und PPC



## 3 Ergebnisse

### 3.1 Isolierung von IF-cDNA-Klonen bei ausgewählten Invertebraten

Zytoplasmatische Intermediärfilamente (IF<sup>zyt</sup>) sind ein wichtiger Bestandteil des Zytoskeletts, dessen Funktion vor allem in der Strukturierung des Zytoplasmas und der Stabilisierung der Zelle gegenüber von außen einwirkendem mechanischen Stress besteht (FUCHS & CLEVELAND 1998). Im Gegensatz dazu sind die Lamine in der Kernhülle lokalisiert und an deren Aufbau und Gestaltung beteiligt (HUTCHISON ET AL. 2001).

Das Ziel dieser Arbeit war es, Intermediärfilamentproteine in deuterostomischen Tieren zu klonieren und zu charakterisieren um die Evolution der IF-Multigenfamilie besser zu verstehen. Dabei ging es hauptsächlich darum, endgültig zu klären, zu welchem Zeitpunkt der Evolution die Deletion der 42 AS in der coil 1b-Subdomäne von Chordaten-IF<sup>zyt</sup>-Proteinen stattfand und wann, ggf. gleichzeitig, das Laminhomologiesegment aus der tail-Domäne verschwand. Unter der Voraussetzung, dass Lamine der Grundtyp der IF<sup>zyt</sup>-Proteine sind, zeigt Figur 1 (Seite 3) eine schematische Übersicht der Domänenorganisation in den bislang bekannten IF-Proteinen.

Zur Beantwortung der Frage, wann in der coil 1b-Subdomäne und der tail-Domäne die Deletionen stattgefunden haben, ist es notwendig, über die entsprechenden Sequenzinformationen zu verfügen. Die Klonierung neuer IF-Proteine aus primitiven Deuterostomiern (Urochordaten, Hemichordaten und Echinodermen) stellte die erste Aufgabe dar.

Zuvor verwendete Methoden waren PCR-Experimente mit degenerierten Konsensusprimern aus den hochkonservierten Bereichen zu Beginn und am Ende der rod-Domäne (ERBER ET AL. 1998). Außerdem wurde versucht, IF-Proteine mit heterologen cDNA-Sonden zu isolieren. Jene Proben, die letztendlich zum Erfolg führten stammten oft von stammesgeschichtlich weit entfernten Tieren und es zeigte sich eine gewisse Beliebigkeit in der Wahl guter DNA-Sonden (SZARO ET AL. 1991, RIEMER ET AL. 1992, BOVENSCHULTE ET AL. 1995). Dies hatte seinen Grund in der geringen Sequenzidentität bei IF-Proteinen, die teilweise nur bei 20 % liegt. Darüber hinaus ergab sich durch den degenerierten genetischen Code ein weiteres Problem, da oft mehrere Codons für ein und die selbe AS kodieren und sich damit die Sequenzübereinstimmungen zwischen der cDNA-Sonde und der Zielsequenz des IF-Proteins weiter verringerten.

Ein bislang wenig genutzter methodischer Ansatz war die Durchmusterung von Expressionsbanken mit dem monoklonalen Antikörper IFA (PRUSS ET AL. 1981). Mit diesem ist es möglich, viele, aber bei weitem nicht alle, zytoplasmatischen und kernlokalisierten

IF-Proteine zu isolieren (RIEMER ET AL. 1991). Der IFA-Antikörper erkennt die Sequenz YRKLEEGE, welche sich bei IF-Proteinen am Ende der rod-Domäne befindet. Im Unterschied zur PCR basierten Methode mit Konsensusprimern, besteht beim Antikörper-screening von Expressionsbanken aber nicht das Problem des degenerierten DNA-Codes, welches die Anzahl der möglichen Primerpaare bei einer vertretbaren Anzahl von Wobble-Positionen in der Primersequenz massiv erhöht.

Die Untersuchungen umfassten Vertreter folgender Tiergruppen: Urochordaten (*Ciona intestinalis*), Hemichordaten (*Saccoglossus kowalevskii*), Cephalochordaten (*Branchiostoma lanceolatum*), Echinodermen (*Strongylocentrotus purpuratus*, *Asterias rubens*, *Astropecten brasiliensis*), Priapulida (*Priapulid caudatus*) und Cnidarier (*Hydra attenuata*).

Der Invertebrat *P. caudatus* ist Mitglied der protostomischen Tiergruppe. Die Untersuchung der bereits vorhandenen cDNA-Bank diente dazu die Datenbasis für IF<sup>zyt</sup>-Proteine der Ecdysozoa zu vergrößern. Für diese den Lophotrochozoa benachbarte Tiergruppe waren bislang nur Sequenzen aus den Nematoda und Nematomorpha verfügbar. Eine verlässliche Analyse der gesamten IF-Evolution ist aber nur mit einer großen Auswahl von IF-Proteinsequenzen aus möglichst vielen Organismengruppen möglich.

### 3.1.1 Arbeiten mit $\lambda$ -cDNA-Expressionsbanken

#### 3.1.1.1 Erstellung von $\lambda$ -cDNA-Expressionsbanken

Genbanken auf der Basis von cDNA sind Ansammlungen willkürlich klonierter cDNA-Fragmente die mit einem Vektorsystem (hier:  $\lambda$ -Phage des „ $\lambda$ ZAP Express cDNA Synthesis und Gigapack Gold III-Kits“, Stratagene) verbunden werden. Im optimalen Fall ermöglichen cDNA-Banken die Konservierung der räumlich-zeitlichen Genexpression in einem Organismus. Dabei kann die Auswahl der Gene sowohl durch die Isolierung bestimmter Gewebetypen, als auch durch Festlegung auf einen zeitlichen Abschnitt der Ontogenese eingegrenzt werden. Auf der Suche nach IF-Proteinen des Typ III wurde das Darmgewebe von *C. intestinalis* ausgewählt, was durch seinen hohen Anteil an glatter Muskulatur das Vorhandensein homopolymerer Intermediärfilamente erwarten ließ. Dazu wurde der Darm eines Tieres präpariert, schockgefroren und anschließend pulverisiert. Nach Extraktion der Gesamt-RNA wurde die mRNA von der rRNA separiert. Über den poly A-Rest des 3'-UTR wurde die mRNA spezifisch an oligo(dT)-Zellulose gebunden und

eluiert. Anschließend wurde die cDNA synthetisiert und die Erstellung der Expressionsbank begonnen. Die genaue Prozedur einschließlich der Abweichungen vom Herstellerprotokoll (Stratagene) sind im Abschnitt „Material und Methoden“ (→ Kap. 2.3.3.1) dieser Arbeit erläutert.

### 3.1.1.2 Plattierung von $\lambda$ -cDNA-Expressionsbanken

Das effiziente Arbeiten mit den klonierten cDNA-Fragmenten setzte die Amplifikation und Stabilisierung der Phagenpartikel voraus. Dazu wurden Zellen des in dem Kit enthaltenen *E. coli* MRF' Stamms auf einer SOB/Tetrazyklin-Platte angezogen und am nächsten Morgen eine SOB/Mg<sup>2+</sup>-Flüssigkultur (25 ml) mit 0,2 % Maltose inokuliert. Die daraufhin von den Bakterienzellen exprimierten und auf der Zelloberfläche lokalisierten Maltoserezeptoren sind die Voraussetzung für die Phagenkompetenz, also die Möglichkeit für die  $\lambda$ -Phagen, Bakterien erfolgreich infizieren zu können. Nach vierstündigem Wachstum bei 37 °C wurden die Bakterienzellen pelletiert, in 10 mM MgSO<sub>4</sub> suspendiert und auf eine durchschnittliche optische Dichte von OD<sub>600</sub> = 0,5 eingestellt.

Zur Amplifikation wurden die  $\lambda$ -kompetenten *E. coli* Zellen (MRF'-Stamm) von den Phagenpartikeln infiziert, ausplattiert und ü.N. bei 38,5 °C inkubiert. Am nächsten Morgen wurden die Platten auf 4 °C abgekühlt und die freigesetzten Phagen in SM-Puffer aufgenommen. Die amplifizierte Expressionsbank hatte einen Titer von 10<sup>5</sup> pfu/ $\mu$ l (Darmgewebe *C. intestinalis*). Zehn Aliquots zu je 1,5 ml wurden mit 7 % DMSO versetzt, in flüssigem Stickstoff schockgefroren und dauerhaft bei -75 °C gelagert. Diese Form der Lagerung ermöglicht einen Zugriff auf repräsentatives Material über einen Zeitraum von bis zu zehn Jahren. Die verbliebenen 50 ml wurden direkt für die anschließenden Experimente verwendet.

Im Verlauf der Arbeit wurden weitere cDNA-Expressionsbanken von *B. lanceolatum* (Darmgewebe: 7,5 x 10<sup>4</sup> pfu/ $\mu$ l) und *C. intestinalis* (Gesamtgewebe: 2,8 x 10<sup>5</sup> pfu/ $\mu$ l) erstellt. Aufgrund elektronenmikroskopischer Untersuchungsergebnisse von Seeigelfüßchen, in denen Filamente mit einem Durchmesser von ca. 5 nm gefunden wurden, (BARTNIK UND WEBER 1989) sollte eine cDNA-Bank dieses speziellen Gewebes ebenfalls erstellt und mit einem Antikörperscreen untersucht werden. Das sehr wasserhaltige Ausgangsmaterial ermöglichte eine cDNA-Bank, deren primärer Titer von 8 x 10<sup>4</sup> pfu/ $\mu$ l weit davon entfernt war, alle exprimierten Gene zu repräsentieren. Die Untersuchungen des Hemichordaten *S. kowalevskii*, des Priapuliden *P. caudatus* und des Cnidariers *H. attenuata* erfolgten unter Verwendung bereits bestehender cDNA-Banken (ERBER ET AL. 1998).

### 3.1.1.3 Isolierung von IF-cDNAs mit dem IFA-Antikörper

Die exprimierten Gene der untersuchten Tiere waren als cDNA-Konstrukt in die DNA von  $\lambda$ -Phagen ligiert worden. Diese konnten nun kompetente *E. coli*-Zellen infizieren, welche dann ausplattiert wurden und sowohl die Vermehrung der Phagen als auch die spätere Expression der klonierten cDNA-Fragmente übernahmen. Die Untersuchung der  $\lambda$ -cDNA-Banken erfolgte auf jeweils 10 großen Petrischalen. Dazu wurde die Dichte der Plaques auf durchschnittlich 18.500 pfu/Platte eingestellt, nach erfolgreicher Infektion der  $\lambda$ -kompetenten Bakterien wuchsen diese sieben Stunden bei 38,5 °C heran bis kleine Plaques sichtbar wurden. Dann wurde die Proteinexpression mit IPTG-getränkten Nitrozellulosefiltern (10 mM) induziert. Die Filter wurden am nächsten Morgen von den zuvor gekühlten Petrischalen abgezogen, fixiert, gewaschen und blockiert. Die Detektion der Intermediärfilamentproteine erfolgte analog der beschriebenen Prozedur (ERBER ET AL. 1998) in vier Stufen:

1.) Die Nitrozellulosefilter wurden zunächst mit 2 %igem BSA in PBS blockiert und 1½ Stunden mit dem IFA-Antikörper in einer Verdünnung von 1:30 bei Raumtemperatur inkubiert. Im Unterschied zu den meisten verwendeten Expressionsbanken wurde für die Suche nach IF-Proteinen von *S. kowalevskii* die Inkubation mit 5 % Magermilchpulver in PBS bei 37 °C durchgeführt. Im Anschluss an die Blockierung und die Antikörperinkubation folgten zwei Waschschriffe.

2.) Ein biotinylierter Schaf-anti-Maus Antikörper wurde benutzt um den murinen IFA-Antikörper zu detektieren. Die Verwendung des biotinylierten Sekundärantikörpers ermöglichte den Einsatz verschiedener Detektionssysteme. Hier wurde das Peroxidasesystem verwendet welches sich durch eine besonders hohe Empfindlichkeit auszeichnet.

3.) Die verwendete Meerrettichperoxidase (HRP) war an Streptavidin gekoppelt, welches seinerseits eine hochspezifische Bindung mit dem Biotin des Sekundärantikörpers eingehen konnte. Die Inkubationszeiten betragen jeweils 1 bis 1½ Stunden.

4.) Die Filter wurden anschließend gewaschen und einzeln für 20 Minuten in je 25 ml Chlornaphtol/H<sub>2</sub>O<sub>2</sub>-Reagenz gelegt. Auf den Filtern lagerte sich an den Stellen mit Protein-Antikörperkomplexen violette Präzipitate ab. Die korrespondierenden Bereiche wurden aus den Agarplatten ausgestochen und in SM-Puffer überführt. Die Phagen diffundierten bei 38,5 °C (ü.N.) in den Puffer, wurden am nächsten Morgen titriert und in zwei weiteren Selektionen auf eine Dichte von unter 300 pfu/Platte vereinzelt. Diese solitären Plaques wurden ausgestochen, die nun monoklonalen Phagen eluiert und für die Klonierung vorbereitet.



#### 3.1.1.4 Klonierung der isolierten cDNA-Fragmente mittels PCR

Die Phagen eines solitären Plaques wurden in 100 µl SM-Puffer eluiert. Einem Standard-PCR-Ansatz von 25 µl wurden 4 µl der Phagenlösung zugesetzt und mit Standard-T3/T7-Primern (T3: 5'-AATTAACCCTCACTAAAGGG-3', T7: 5'-GTAATACGACTCACTATAGGGC-3') eine Amplifikation über 35 Zyklen durchgeführt (→ Kap. 2.3.4.5). Die Reaktionsprodukte wurden auf einem Agarosegel gereinigt (→ Kap. 2.3.7) und mit einem TA-Klonierungssystem in den bakteriellen Vektor pCR 2.1 ligiert. Chemokompetente Bakterienzellen vom Typ *E. coli* TOP10 F' (Invitrogen) wurden mit 2 µl Ligationsprodukt transfiziert und ausplattiert. Die Vektorinserte der farblosen Bakterienkolonien wurden mit der schnellen Kolonie-PCR-Methode (→ Kap. 2.3.4.2) überprüft und ü.N. in 25 ml-Kulturen zur präparativen Plasmidisolierung aufgezogen. Die Plasmid-DNA wurde extrahiert (→ Kap. 2.3.12.2), quantifiziert und zur Sequenzanalyse eingesetzt. Die Sequenzierungsarbeiten wurden von SeqLab (Göttingen) oder MWG-Biotech (Ebersbach) als Auftragsarbeiten durchgeführt.

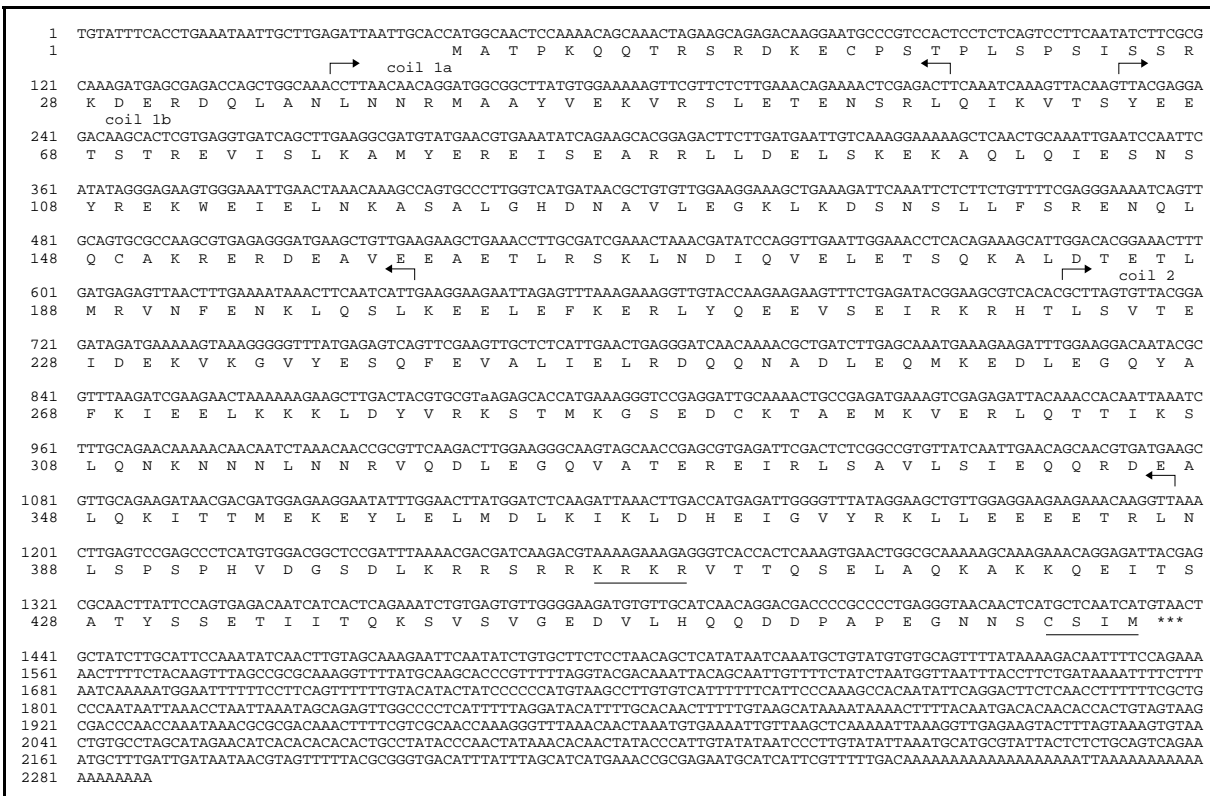
### 3.2 Zwei neue IF-Klone bei *Ciona intestinalis*

Die Untersuchung der cDNA-Bank (Darmgewebe) mit dem IFA-Antikörper ergab fünf primäre Signale, von denen vier vereinzelt und ihre Inserte kloniert werden konnten. Nach der Sequenzanalyse wurden zwei neue IF-Klone identifiziert. Ein IF-Protein vom Lamintyp (*Ciona* Lamin L2: Figur 2) und ein zytoplasmatisches IF-Protein (*Ciona* IF A2: Figur 7).

#### 3.2.1 Charakterisierung von *C. intestinalis* Lamin L2

Das Lamin L2 (Figur 2) lag als Vollängenklon vor. Die Gesamtlänge der klonierten cDNA beträgt 2288 bp, der offene Leserahmen umfasst den Bereich von nt. 41 bis nt. 1438. Die berechnete molekulare Masse beträgt 58042 D, der isoelektrische Punkt liegt bei 4,89.

Der zentrale Bereich (rod) des Proteins Lamin L2 unterteilt sich in die Subdomänen coil 1a, coil 1b und coil 2 und entspricht damit dem bekannten Schema. Die für IF-Proteine charakteristischen Konsensussequenzen befinden sich bei der AS-Position 35-48 und am Ende der coil 2b-Subdomäne (AS 376-387). Zusätzlich sind in der tail-Domäne das Kernlokalisationssignal (KRKR) und das C-terminale CaaX-Motif (CSIM) vorhanden welche das neue Protein eindeutig als Lamin (Typ V IF-Protein) definieren. In deutlicher

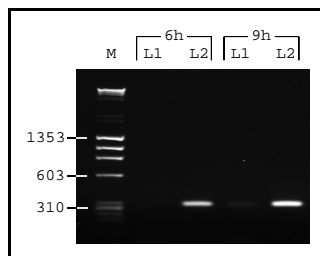


Figur 2: *Ciona intestinalis*, Lamin L2, (AC #: AJ251957). cDNA und korrespondierende AS-Sequenz des neuen Lamin L2. Das Kernlokalisierungssignal und die carboxyterminale CaaX-Box (beide unterstrichen) definieren dieses IF-Protein als Lamin. Die drei Subdomänen coil 1a, 1b und coil 2 sind durch horizontale Pfeilpaare gekennzeichnet.

Übereinstimmung mit den bereits bekannten Laminen aus *C. intestinalis*, *Styela clava* und *Molgula oculata* (RIEMER ET AL. 2000), zeigt die tail-Domäne auch hier die für alle Urochordaten spezifische Deletion von 90 AS. Somit fehlt der größte Teil der 105-Box, die bei allen übrigen deuterostomischen und protostomischen Laminen vorhanden ist.

### 3.2.2 Räumlich-zeitliche Analyse der Expression von Lamin L2

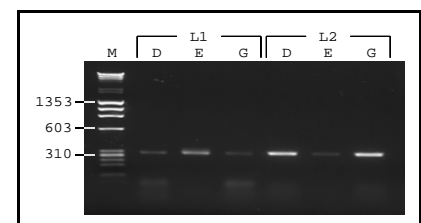
Beim Vorhandensein von zwei Lamintypen in einem Organismus stellte sich die Frage nach Sinn und Funktion dieser Zwillinge. Deshalb wurden mit hochspezifischen Primern für Lamin L1 (forward: 5'-GTCTATCGAAAGCTGTTGGAAGAA-3', reverse: 5'-CAAATATTGAGGCATTTTCGTCATTGG-3') und dem neuen Lamin L2 (forward: 5'-GTTTATAGGAAGCTGTTGGAGGAA-3', reverse: 5'-CAAGTTGATATTTGGAATGC-AAGATAGC) zunächst cDNA-Banken von frühen Entwicklungsstadien überprüft (Figur 3).



Figur 3: *Ciona intestinalis*, Analytisches Agarosegel der halb quantitativen PCR-Reaktionen zum Nachweis von Lamin L1 und L2 in Expressionsbanken von sechs (6 h) und neun Stunden (9 h) alten Embryonen. M: DNA-Standard

Zur Verfügung standen Expressionsbanken von sechs und neun Stunden alten Embryonen. In beiden Fällen war Lamin L2 dominant vertreten. Auch PCR-Produkte vom Lamin L1 konnten in beiden cDNA-Banken, wenn auch nur in sehr geringer Menge, amplifiziert werden. Alle PCR-Produkte wurden kloniert und ihre Identität durch Sequenzierung bestätigt. Außerdem wurden Analysen über die relative Häufigkeit der beiden Lamine in ausgesuchten Geweben von *C. intestinalis* durchgeführt. Die halb quantitative PCR-Untersuchung zur relativen Häufigkeit von Lamin L1 und L2 in der

cDNA-Bank aus Epidermisgewebe (RIEMER ET AL. 2000) zeigte mehrheitlich PCR-Produkte von Lamin L1 gegenüber dem geringer repräsentierten Lamin L2. Die ebenfalls untersuchten gewebespezifischen cDNA-Expressionsbanken aus Darmgewebe und Ganglien hatten ein genau gegensätzliches Ergebnis. Die Produktmenge der Lamin L2-Fragmente überstieg die der L1-Produkte um ein Vielfaches (Figur 4). Eine direkte Schlussfolgerung aus dem alternierenden Auftreten der Lamine L1 und L2 in den adulten Geweben und eine durchgängig schwache Expression des Lamin L1 in embryonalen Entwicklungsstadien ist nicht möglich. Es ist jedoch denkbar, dass es sich beim Lamin L1 um eine Kontamination der Epidermis-Bank handelt.



Figur 4: *Ciona intestinalis*. Halb quantitative PCR der Lamine L1 und L2 in Expressionsbanken von drei verschiedenen Geweben (D: Darmgewebe, E: Epidermis, G: Ganglien) M: DNA-Standard.

### 3.2.3 Charakterisierung von Ciona IF A2

Das zweite IF-Protein aus *C. intestinalis* (Figur 7) gehört zur Gruppe der zytoplasmatischen Intermediärfilamentproteine. Weder das Kernlokalisierungssignal, noch die CaaX-Box sind vorhanden. Die Länge beträgt 1884 bp, der offene Leserahmen umfasst den Bereich von nt. 26 bis nt. 1558. Die berechnete molekulare Masse beträgt 58042 D, der isoelektrische Punkt liegt bei 4,89. Die cDNA von IF A2 konnte auch in der aus Epidermisgewebe erstellten cDNA-Bank nachgewiesen werden.

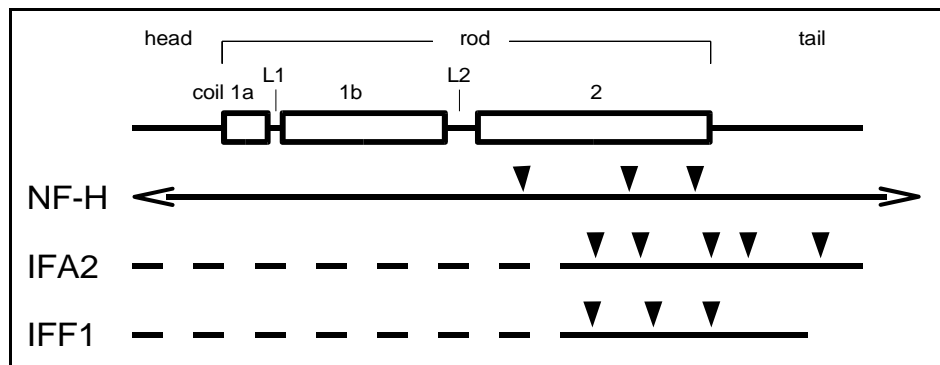
### 3.2.4 Weitere Untersuchungen der Tunikaten-IF-Proteine

Auf der Basis der in dieser Arbeit isolierten IF-Proteinsequenzen von Lamin L2 und IF A2 konnten in der Arbeitsgruppe weitere IF-Proteine kloniert werden: IF B2, IF C2, IF D2, IF F1 und IF F2 (WANG ET AL. 2002). Die Proteine IF F1 und IF F2 konnte auf der Basis der Primärsequenz und prognostizierten Sekundärstruktur, ebenso wie IF A1/A2 keinen bereits bekannten IF-Sequenzen außerhalb der Urochordaten funktionell zugeordnet werden. Deshalb wurde die Verbreitung von IF F2 und IF A2 cDNA-Klonen in den vorhandenen gewebespezifischen Expressionsbanken aus Darmgewebe und Ganglien mit PCR-Methoden nachgewiesen. Vor allem das gehäufte Vorkommen in der aus Ganglien erstellten cDNA-Bank veranlasste eine Untersuchung der genomischen Organisation um mögliche Übereinstimmungen mit den Neurofilamenten (Typ IV IF-Proteinen) der Vertebraten zu finden. Diese haben in der coil 2-Subdomäne zwei, bzw. drei charakteristische Intronpositionen (LEES ET AL. 1988, MYERS ET AL. 1987, LEWIS UND COWAN 1986).

Dazu wurde der genomische rod- und tail-Abschnitt der IF-Proteine IF A2 und IF F1 abschnittsweise kloniert, sequenziert (Figuren 16 & 17, S. 76 ff.) und die gefundenen Intronpositionen mit denen der Neurofilamentproteine verglichen (Figur 6). Eine Verwandtschaft der untersuchten zytoplasmatischen IF-Proteine mit den Neurofilamenten der Vertebraten basierend auf charakteristischen Intronpositionen konnte nicht festgestellt werden.

1	GAAGTTATAACAAGGTAACCTCAAGATGGCGCAAAGGAGCAGCAGAAAAACGACCAGAGCGGTAgCtCAAGCTACAAACGTCACTTCGGTGAAGGTCCGGCCAGATCACTTATTCAAG
1	M A Q R S S R K T T E E R S S S Y K R H F G E G P A R F T Y S S
121	CAAGTCATCTCCAGGCCGATCCAGTCCGCAACCGCATCGTACAGGTACTCAAGCAGCGCGGTGGCGGGCGGTGGAGCTGCCATGGCCGGTGGCCATGGGTGGCGATGGGCAT
33	K S S P G R S Q S A T R S Y R Y S S S G G G G G G G A A M A G G A M G G G M G M
241	GGGTGCTGGCATGGGTATGGGTGCTGGTATGGGGGCGAAAGGATTAGGATGACCCGAGCCGACCACTGAGCAGCGGGATCTCAATGCGAGACATCAACCTAATGTCCGATGCTTTGGG
73	G A G M G M G A G M G G E R I R M T R A A P V S S G I S M R D I N L M S D A L G
361	ACCGGATTTCGAAAGAATGAGGATTAATGAAAAGGAAGAACTTCAAGTCTCAATACAAGATTCGCTGGCTACATCCAAAAGGTTGAGCACTTGAACAAGCCAATAAGATCTTGGAAAT
113	P D F E R M R I N E K E E L Q V L N T R F A G Y I Q K V R A L E Q A N K I L E V
481	TCAACTTGAACAACCTGAGCGCCCAACAACCTGGACGTTGCTGATATGTTGCAAGATGAATGGCAAGACTCCGCCGAGAAGTCCGACCAGATCACCAAAAGAAAAGGCGAACATCCTGAT
153	Q L E Q L S A Q Q P G R V A D M F E D E L A R L R R E V D Q I T K E K A N I L M
601	GCAGTTGGAAAACGACGCTTGAAGCAAACTGGAAGACAAATACGAAGAAGCAATCAGCAAAACGGGAACCGAAGACGACCTCAACTCCATCGAAAAGATTGCGATGACGC
193	Q L E N A A L E A N N W K D K Y E E E A I S K R E L E D D L N S M R K D C D D A
721	AACCTTGGTTCGCTTACCTTGAACGAAGATTGGAACACTTCAAGAAGAAATGAATTCCTTAAGAGAACTCATGACCAGGAGGTTTCTGAGCTCATGGAACAAGTCAGGAAGACGGA
233	T L V R L D L E R R L E T L Q E E I E F L K R T H D Q E V S E L M E Q V R K T E
841	GATAAAGATCGAGATGGCTCCCGACCTGATCTGGAGAACTTGTGAAGACATGAGAAAACAATACGAGAACTCGCATCCAAGAACCAGCTGATGCTGAGGAATGTTACACTGAAAA
273	I K I E M A P G P D L E E L V E D M R K Q Y E K L A S K N R A D A E E W Y T E K
961	GGTGTCCGGTCTCCAAGACCAAGCCGTCGAAGAACGACGATCTCGTACGTAACGTTGGAAGTGAAGTGAAGCAATAACAGAAAGAGTGTCCAAAACCTTGAATATGGAGATTGACTCTTTGAG
313	V S G L Q D Q A V K N D D L V R N V R S E L S E Y R K S V Q T L N M E I D S L R
1081	AGGAACAAACGACTCCCTTCAACGAGCAATGGGTGAACCTGAAGAGAGATACAACAGAGATACTCCGATTACCAAGATACCATCATGAACCTTGCAGGCTGAGTCCGACGACCTGAAGAC
353	G T N D S L Q R A M G E L E E R Y N R D T S D Y Q D T I M N L Q A E C D D L K T
1201	ACAGATGGCTCAACACCTCCGTCATACCAAGAACCTCATGGACGTAAGATGGCTTTGGACATCGAGATCGCTACTTACAGGAaGCTGTTGGAGGGAGAGGAATCAAGGCTTGCTGAACG
393	Q M A Q H L R Q Y Q E L M D V K M A L D I E I A T Y R K L L E G E E S R L A E R

Figur 5: *Ciona intestinalis*, cDNA (oben) und korrespondierende Aminosäuresequenz (unten) des zytoplasmatischen IF-Proteins IF A2 (AC #: AJ298329).



Figur 6: Schematischer Vergleich der genomischen Organisation von drei Chordaten-IF<sup>zyl</sup>-Proteinen: Humanes Neurofilament NF-H, *Ciona intestinalis* IF A2 und IF F1. Die AS-Sequenz der IF-Proteine wurde mit der Domänendarstellung zur Deckung gebracht. Die Länge des NF-H-Proteins übertrifft die anderen Proteine bei weitem. Dies wird durch die Pfeilspitzen verdeutlicht. Weil die Funktion der beiden Urochordaten-Proteine aus Sequenzvergleichen zunächst nicht ersichtlich war, wurden die Intronpositionen der entsprechenden Gene ermittelt, da diese bei Neurofilamenten besonders charakteristisch sind. Zur Identifizierung der diagnostischen Intronpositionen war eine Überprüfung der coil 2-Abschnitte ausreichend. Die weiter aminoterminal liegenden Abschnitte der Proteine wurden deshalb nicht bezüglich ihrer Genorganisation untersucht und sind als gestrichelte Linie dargestellt. Offensichtlich haben die Neurofilamentgene der Vertebraten- und die Urochordaten-IF-Gene keine gemeinsamen Intronpositionen. Die Urochordaten-IF-Gene haben zwei identische Positionen für das erste und dritte (v. l.) hier dargestellte Intron.

### 3.3 Das erste zytoplasmatische IF-Protein des Hemichordaten *S. kowalevskii*

Die Durchmusterung der von A. Erber erstellten cDNA-Expressionsbank (ERBER ET AL. 1998) mit dem IFA-Antikörper ergab 31 primäre Signale die vereinzelt und kloniert wurden. Ein partieller IF-Klon wurde identifiziert. Um die Sequenz zu vervollständigen wurde die cDNA-Bank mit einer DIG-UTP markierten DNA-Probe des neuen IF-Klons untersucht (→ Kap. 2.3.3.4). Die unter halb stringenten Bedingungen erhaltenen Primärisolate wurden unter Umgehung der zeitaufwendigen Folgeschritte direkt mit spezifischen Primern auf ihre Übereinstimmung mit der partiellen Sequenz getestet. Mit der „nested“-PCR (→ Kap. 2.3.4.4) auf den Primärisolaten konnte ein Klon mit vollständiger Sequenz (Figur 5) gefunden werden.

Das Protein Sacco IF1 hat eine Länge von 543 AS. Das berechnete Molekulargewicht beträgt 63090 D, der isoelektrische Punkt liegt bei 4,89. Das Protein hat die für IF-Proteine übliche dreigeteilte Struktur aus head-, rod-, und tail-Domäne (FUCHS UND WEBER 1994, PARRY UND STEINERT 1995). Die Sequenz zeigt deutliche Übereinstimmungen mit zytoplasmatischen IF-Proteinen und darüber hinaus auch mit Laminen.

#### 3.3.1 Sacco IF1 head-Domäne

Die AS-Sequenzen von IF-Proteinen sind in der head- und tail-Domäne sehr variabel, wobei die head-Domäne in der Regel einen Überschuss an basischen AS-Resten zeigt. Beim Sacco IF1-Protein findet sich hier eine Überzahl von 24 negativ gegenüber 15 positiv geladenen Resten. Zusätzlich befindet sich zwischen AS 23 und AS 32 ein ungewöhnliches Cluster von zehn Serinresten. Ein Sequenzmuster dieser Art und Größe ist für IF-Proteine selten, eine Datenbanksuche mit den zehn Serinresten fand ein vergleichbares Motiv nur beim IF B2 von *C. intestinalis* (AC#: AJ298330) dessen Serincluster sich jedoch im carboxyterminalen Bereich der tail-Domäne befindet.

```

1 TGACACAGAGACAGCGCGTAAAAATAGAAAGTCGAAAAACAAAGACGATAGAAAGTTGGAGACAGCAAAGTATCTGTTGAAACTAGCTTGACCATTGAGGGCGCATCTTGACGACATGTC
121 AGCTGAGAAAATGCAGCAGATCTTAGACGGTGCACAGGATATGATAAGTAGTGGcGATGCAGAGGAAGACCCGAAAAGCCCGAGGTAAAAGACGACAGGGCGGTTCAATCTAGCAGTTC
1 M I S S G D A E E E P E K P E V K D D R R G S S S S
241 CTCAAGCAGTTCAAGTGATGAAGATGAATCCGCGTCGCTAAAGGAGAAGATTGACGTCGAAAAATCTGTAGACATTCTTAAGGAGCCAAACAGCGAAGTCTAGAAGATGAAGTGGAAAA
28 S S S S S D E D E S A S L K E K I D V E K S V D I S K E A K Q R S L E D E V E N
361 TAGCGGTACCAGAGATAAAGATGTCAGTTTACCTAGACACTACGAGGGCAATACGTATCTGAAACGACCTCTCGCATGTATGTTGGATGTTGTTGGGGAAGATTGAAACTGGAGACGA
68 R R T R D K D V S L A R H Y E G E L R I L N E R L A C Y V D V V G K I E T G D E
481 GAGAGTAAAAAAGCTCTGGGGTAAACTAAACTAGTGTAGGAAACGGGTAGCTAATGCAATTGGAGACTTGAAGGATGATACGAAATATCATTACGTGACATCCGTGAACCTTATCAAGAC
108 R V K K V W G K L N T S E E R V A N A I G D L K D E Y E I S L R D I R E L I K T
601 GATCAATGAAGAGAGACAACAGATGACGATGAGAGGGATGACCGGAATAAGAAAATGGAGACCTCCGGCCAGAGCTAAAAATCAGTAGAAATTGATTGACGAAATAAACCCGTAGAAGT
148 I N E E R Q Q N D D E R D D R N K K I G D L R A E L K S V E I D S A N K T V E V
721 CAAAATATTAAATTTGAAAAAGACACGCTCAATGAAGAAGTGGCGAGCAAAGCTAAGGCACCTGGTGCAGCTCGTAAAAATAGCGGAACAGGAGCGGAAGCTGCGTGAAGAATCGGAGAG
188 K T I N L E K D T L N E E V A S K A K A L A A A R K I A E Q E R K L R E E L E S
841 TACCTTCAGAAAAGAAATGCTACTACAAGTAAATCGAGAAGCAGTGGGAAGAAATCAAGATGGCAAAGAACAGATCATATCTGGTATTCTTGACCAGGACAAGAAGCGAGCGGAAGA
228 T F R K E M S L Q D E I E K Q W E E I K M A K K Q I I S G I L D Q D K K A A E E
961 AGAAAAGATGAAACTGAAAGCGACATTCAGAAAATGGAGTGAACATTAGCGAAGAGCTTAAGAAATACAAAAAGACTATTGAAAATGAGTACGGCTTGAATGAAAAAATGAGACG
268 E K M K L K A T L Q K W S E H Y S E A E T I E N E Y G F E L K K M R R
1081 ACAAGTTGACGAATTCAGAGCAGATCTAAGCAACGGAATACGAACTGAAACTTGGCCCTTCTCAAGATTGGTGAATGGATGATCAGCACAGATCAGTTTTAGCAAGAAATGGTTTCAT
308 Q V D E F K S R S N A T E Y E L K L A L L K I G E M D D Q H R S V L A K I G F M
1201 GGAAGAACCGGTTCAAGAACTCGAAGAAAGAAAAAGAAATATTAGAAGACCAGAAACGCCAATTCGGTCTTATTAGAGAGAAGGACGAAGAAATCCGCGAATTACTCAATCGACTAGA
348 E E R V Q E L E E R K K K L L E D Q E R Q F G L I R E K D E E I R E L L N R L D
1321 CAGACAAGCGGGAGTTCCCGAAGCCGAGAGACCGAAATTTGCTTGACAAAAGAAATTCGCAAAATTCAGAAAATTTGTTGAGATGGAAGAAAACAGGCTGAATTTGGATGAAATGGA
388 R Q A R E F A E A E R D R I L L D K E I A K F R K L L E M E E N R L N L D E M E
1441 AATGGAAGAGCAGAAATCAAGCTTTGAAATTTCAATCGATGTATCCAAAAAAGTGTAAAGATAGTTAACAAGGGAGATTCCGGATCAAACTGGAAGTTACAACTACGAAGGACAA
428 M E E Q K S S F E I S I D V S K K S V K I V N K G D S D Q T L E G Y K L R R S K
1561 GGGAGGCAGGACTTCAACTTCAGCTTCAAGTCTGGACAGGTATTGAGGGCAGGCTCAAGTATGCAATCAAATGAAAGGACAAACCGTCAATAAAGACGATGACGAAGATGAATGGGA
468 G G R D F N F S F K S G Q V L R A G S S M T I K M K G Q T V N K D D D E D E W D
1681 CGATATGAGTGACATCAGTGATGACGAAAGTGAACATGCTTGGTTGATACAAGTGGAAACGGTCAAGACAACCCACTTTAGCAAGTCCAAAAGAGTATTAGAATACGCTTAGGTATCCAA
508 D M S D I S D D E S E T C L V D T S G T V K T T H F S K S K R V L E Y A
1801 ACTGATGAATCCCGTTGAAGACAACGATTTTATACATGCGATAAAAACACTTAGATGTGTAATTTTTCATAAGTGTGAACATATCAAGCAGTCTGGTGTGCGCATTATATAGACCATGAC
1921 TTTCTCTTTTATATCCAGTAGTTAATTTGATTTGTTGTTTATGACTTACTTAAATTCATTAAATGCTGTAACAACATATATACTGAATGTGCTCCTAGTCTACAGTGTGATGT
2041 GAAGTATTATTATATAAAGCTTTGTCACACTCATAGCGCTGCCAAAATAAATGAGTTCAGAAAAATCCATAAATTTGCGTGGATTATCCCCAGAACCAACAATCTGCTACCTACTA
2161 AGTCTTTCACATTATCCCGTCTGTTTACTAGCCTTGAGTTATAAGCATAACATTTCCGGTTCTTATCTAATTCATTGCTCGAATTTGTTGAATGGGACCCCTCCGCGCATCATGGAT
2281 CTGAGAATATATTTACATGTTTCAAGTCAATGTGTCTCTGGGCTTACATTTGCTTCCATTTGGACAATGGTTCATGATATATTTGACAGCACGTATTAACTGAATATATCC
2401 TAGAAATCTTTGTAATAACAGATACAGTGTGTGCTGAATATAATGAAATAAAGGATAGTTGTTGTTTCAATCCAAAAAATAAAAAAAAAAAAAA

```

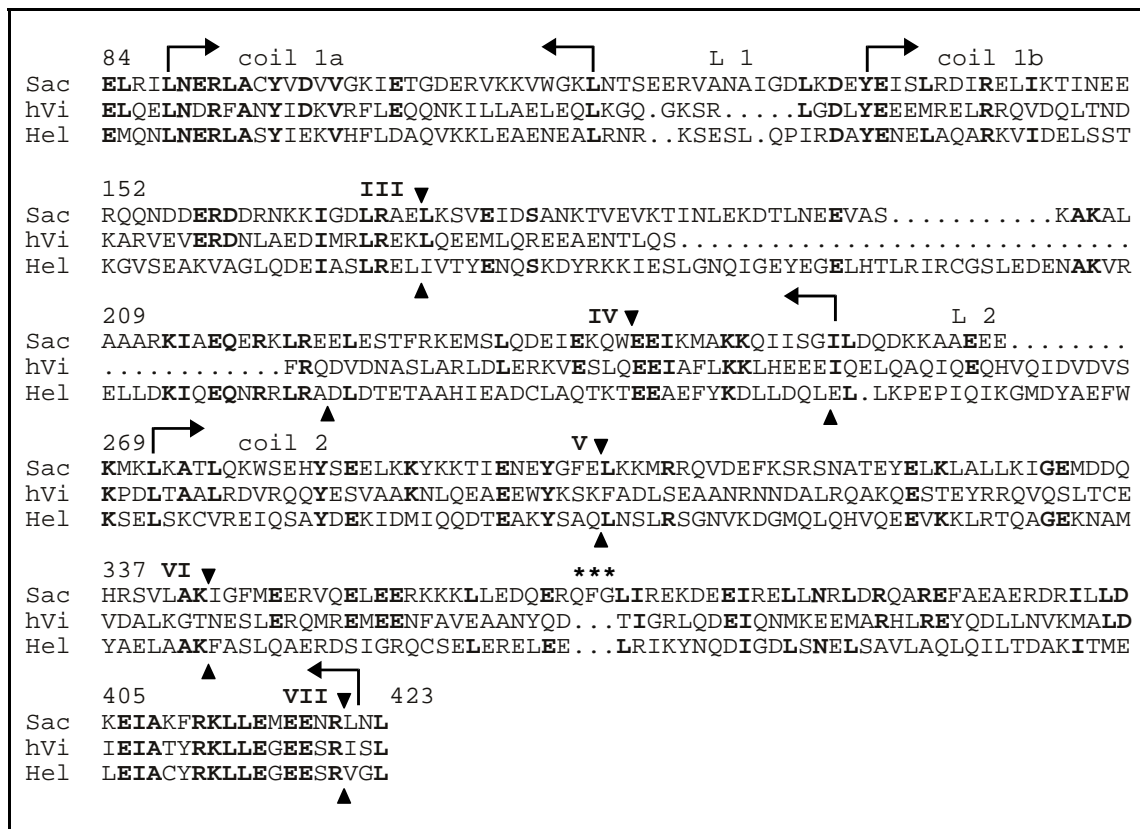
Figur 7: *S. kowalevskii*, cDNA-Sequenz (oben), darunter die korrespondierende AS-Sequenz des zytoplasmatischen IF-Proteins Sacco IF1 (AC #: AJ421616). Dreiecke markieren die Positionen der neun Introne. Die Sequenz konnte ab nt. 213 in sich überlappenden genomischen Fragmenten kloniert werden. In der head-Domäne ist ein durchgehender Bereich von zehn Serinresten markiert. Das IFA-Epitop FRKLEEMEE stimmt gut mit mit der Konsensussequenz (RIEMER ET AL. 1991) überein.

### 3.3.2 Die Organisation der rod-Domäne

Der 333 AS lange zentrale Bereich (rod) des IF-Proteins lässt sich in drei helixbildende Subdomänen unterteilen (coils), die durch nicht helikale „Linker“ variabler Länge miteinander verbunden sind. Die für IF-Proteine charakteristischen Erkennungssequenzen (GEISLER & WEBER 1982) am Beginn von „coil 1a“ (LNERLAXYXDXV) und am Ende von „coil 2“ (EIAXXRKLEEXE) sind auch hier vorhanden (Figuren 5 und 8). Die ungewöhnliche Zusammenstellung von Veränderungen innerhalb der Subdomänen zeigt sich erst beim direkten Vergleich der gesamten rod-Domäne mit einem protostomischen IF<sup>zyt</sup>-Protein und Vimentin, einem charakteristischen IF<sup>zyt</sup>-Protein der Chordaten (Figur 8).

Dem Hemichordatenprotein und dem humanen Vimentin fehlen in der coil 1b-Subdomäne gegenüber der protostomischen Sequenz des Helixproteins 11 AS bzw. 42 AS. Das Muster, welches in „coil 1b“ normalerweise eine sehr regelmäßige Abfolge von hydrophoben AS-Resten an den Positionen 1 und 4 jeder Heptade hat wird im Gegensatz zu Vimentin beim Protein Sacco IF1 durch die Deletion von nur 11 AS durchbrochen.

Darüber hinaus gibt es eine ungewöhnliche aber auf der Grundlage von zwei unabhängigen genomischen Analysen bestätigte Insertion von drei AS (QFG) in „coil 2“. Außerdem sind in den Linkerbereichen Längenvarianzen zu den beiden Vergleichssequenzen zu sehen. Auffällig ist das Fehlen von Prolinen in diesen Bereichen. Prolin unterbricht die Bildung von Helixstrukturen und ist in manchen IF-Linkersequenzen zu finden.



Figur 8: Alignment der rod-Domäne des Hemichordatenproteins Sacco 1 („Sac“) mit humanem Vimentin („hVi“: PERREAU ET AL. 1988) und dem protostomischen IF-A-Protein aus *Helix aspersa* („Hel“: DODEMONT ET AL. 1990). Aufeinander zeigende horizontale Pfeile markieren die drei Subdomänen coil 1a, 1b und coil 2. Ein- oder zweifach konservierte Reste in der Sacco IF1-Sequenz sind fett gedruckt. Die für das optimale Alignment eingefügten Leerstellen sind mit Punkten markiert. Doppelte Dreiecke markieren die vier identischen Intronpositionen aller drei Sequenzen. Einzelne Dreiecke unter dem Alignment zeigen Introne an, die nur in „Hel“ und „hVi“ vorkommen. Die Intronpositionen von Sacco IF1 sind mit römischen Zahlen durchnummeriert (vgl. Figur 5). Die einmalige Insertion von drei AS in coil 2 (QFG) ist mit " \* " markiert. Das zweite Intron der Vergleichssequenzen fehlt beim Sacco IF1-Protein, das Intron IV scheint um 13 AS in N-terminaler Richtung verschoben zu sein.



### 3.3.3 Die Intronverteilung des Sacco IF1 Gens

Die genomische Sequenz für den cDNA-Abschnitt von nt. 213 bis nt. 1823 wurde analysiert (Figur 15, S. 75). Dabei wurden insgesamt neun Introne gefunden (Figur 5), die Länge der Introne liegt zwischen 94 bp (Intron III) und 332 bp (Intron VII). Über möglicherweise vorhandene Introne vor dem nt. 213 der cDNA-Sequenz kann keine Aussage gemacht werden, weil die dafür erforderliche Klonierung des genomischen Bereichs nicht erfolgreich war.

Figur 8 zeigt einen Vergleich der Sacco IF1-Sequenz (rod) mit analogen Sequenzabschnitten aus humanem Vimentin („hVi“: PERREAU ET AL. 1988) und einem protostomischen IF<sup>zyt</sup>-Protein (*H. aspersa*, „Hel“: DODEMONT ET AL. 1990). Insgesamt sind in der rod-Domäne des Sacco IF1-Gens fünf Introne (III-VII) vorhanden. Die zum Vergleich gewählten Gene „hVi“ und „Hel“ beinhalten im gleichen Abschnitt sechs Introne (DODEMONT ET AL. 1990). Die Introne III, V, VI und VII der Sacco IF1-Sequenz stimmen in ihrer Position mit den Intronen der Vergleichssequenzen exakt überein, die coil 2-Subdomäne ist somit identisch organisiert. In „coil 1b“ zeigen sich jedoch einige Unterschiede in der Genorganisation: Das vierte Intron scheint gegenüber den Vergleichssequenzen in N-terminaler Richtung um 13 AS verschoben zu sein. Davor fehlt ein den unteren Sequenzen entsprechendes Intron und wurde vermutlich deletiert. Das sechste Intron der *Saccoglossus*-Sequenz hat eine Abweichung beim Übergang vom Exon zum Intron. Die beiden gut konservierten Nukleotide GT zu Beginn der meisten Introne sind hier durch GC ersetzt worden (Figur 15). Diese Variation wurde durch mehrfache unabhängig voneinander durchgeführte Klonierungen bestätigt.

#### 3.3.3.1 Sacco IF1 ist das erste deuterostomische IF<sup>zyt</sup>-Protein mit einem Laminhomologiesegment in der tail-Domäne

Die tail-Domäne hat eine Länge von 121 AS und ist wesentlich länger als die bislang bekannten Sequenzen der Chordaten-IF-Proteine. Der Vergleich mit vier Lamingenen aus protostomischen und deuterostomischen Organismen zeigt eine Sequenzübereinstimmung von sechs identischen (\*) und 22 hochkonservierten (+) AS-Resten, entsprechend 23 % (Figur 9). Die Organisation des Sacco IF1-Gens zeigt weiterhin, dass deutliche Übereinstimmungen mit der 105-Box der Lamine bestehen: Die meisten Lamingene haben in der 105-Box drei Introne. Das Gen von Sacco IF1 enthält zwei Introne, die phasengenau an der gleichen Position lokalisiert sind wie das erste und dritte Intron der Lamingene. Jedoch sind zwei charakteristische Merkmale der Lamine natürlich

nicht vorhanden. Sowohl das Kernlokalisierungssignal (NLS) als auch die carboxyterminale CaaX-Box fehlen und weisen Sacco IF1 als echtes zytoplasmatisches IF-Protein aus.

Die Analysen der Proteinsequenz und den genomischen Merkmalen verdeutlichen, dass das IF<sup>zyt</sup>-Protein Sacco IF1 größere Ähnlichkeit mit den zytoplasmatischen IF-Proteinen der Protostomier hat als mit denen der Chordaten. Dadurch wird offensichtlich, dass sowohl die Deletion von 42 AS in der coil 1b-Subdomäne als auch das Verschwinden des Laminhomologiesegments in der tail-Domäne als molekulare Merkmale der Chordaten betrachtet werden dürfen.

Sac	431	+++	+++ *	+ *	+ + + + *	+ + +	+++	+				
Sac		<b>QKSS</b>	<b>FEISI</b>	<b>DV</b>	<b>SK</b>	<b>KSVKIVNKGDS</b>	<b>VDQ</b>	<b>TLEGYKLRRSKG</b>	<b>GRDFNFSFKSQVLRAGSSM</b>			
Pri L		<b>TDSSKGE</b>	<b>VRVMECD</b>	<b>PDGK</b>	<b>FVKLSNTSDK</b>	<b>DIHIGTWQLKRITG</b>	<b>DKEITYKFHRNI</b>	<b>ILKAGQHI</b>				
Tea L		<b>TTAE</b>	<b>GAIQIVESN</b>	<b>PEGK</b>	<b>FIKLFNSGDK</b>	<b>DEPLGGWTIQ</b>	<b>RQVGKEDPI</b>	<b>VYKFTPKYVLRAGSSM</b>				
Xen L3		<b>QASSTGPV</b>	<b>SVEDID</b>	<b>PEGNYV</b>	<b>RLLNNTTE</b>	<b>DFSLHGWVVK</b>	<b>RMHMSLPEIA</b>	<b>FKLPCR</b>	<b>FILKSSQRV</b>			
Mus LB2		<b>QTVATGVVN</b>	<b>IDVDPEGR</b>	<b>FVRLKNS</b>	<b>SDK</b>	<b>ADQ</b>	<b>SLGNWRIK</b>	<b>RQVLEGED</b>	<b>IAYKFTPKYVLRAGQTV</b>			
Sac		<b>T</b>	<b>IKMKGQ</b>	<b>TVNKDDDED</b>	<b>EWDDMSDI</b>	<b>SDDESETCL</b>	<b>VDTS</b>	<b>SGT</b>	<b>VVKTT</b>	<b>HFSKSKRVLE</b>	<b>YA</b>	543
Pri L		<b>N</b>	<b>VWSSDVKDV</b>	<b>THNPPS</b>	<b>DLVMKNQ</b>	<b>KWYSGESM</b>	<b>KTI</b>	<b>LIDSEGE</b>	<b>EVGERENFR</b>	<b>QLRASSFQ</b>	<b>...</b>	
Tea L		<b>T</b>	<b>IYSSDGGG</b>	<b>QHK</b>	<b>PPAELLYK</b>	<b>HLSWGS</b>	<b>GNEVRTAL</b>	<b>INASGE</b>	<b>ELAALTEE</b>	<b>KIFEQYCKD</b>	<b>...</b>	
Xen L3		<b>T</b>	<b>IWATGAGAV</b>	<b>HSPPTDL</b>	<b>VWKSQKT</b>	<b>WGTGDN</b>	<b>KITLLD</b>	<b>STGE</b>	<b>ECAERTLY</b>	<b>RVI</b>	<b>...GEEG</b>	<b>...</b>
Mus LB2		<b>T</b>	<b>AVWAAGAG</b>	<b>ATHS</b>	<b>PPSTLV</b>	<b>WKSQTN</b>	<b>WGPGESFR</b>	<b>TALVSAD</b>	<b>GE</b>	<b>EVAVKAAKH</b>	<b>SSVQGRE</b>	<b>...</b>

Figur 9: Vergleich der 112 carboxyterminalen AS-Reste des Sacco IF1-Proteins mit dem mittleren tail-Abschnitt von vier Laminen. *Priapulus caudatus* Lamin (Pri L), *Tealia* Lamin (Tea L), *Xenopus laevis* Lamin L3 (Xen L3) und Maus Lamin B2 (Mus LB2). Die Sequenzen der Lamingene sind der Arbeit von ERBER ET AL.(1998) entnommen. AS-Reste von Sacco IF1 die mit mindestens einer Vergleichssequenz übereinstimmen sind fett gedruckt. Gut konservierte AS-Positionen sind mit (+), die sechs strikt konservierten AS-Reste mit (\*) markiert. Dreiecke markieren die Intronpositionen, doppelte Dreiecke zeigen die beiden Introne an, die auch im Gen von Sacco IF1 nachgewiesen werden konnten. Das erste Intron im Lamingen von *Priapulus* ist deletiert, die bei *Saccoglossus* verbliebenen Intronpositionen stimmen mit den Intronpositionen I und III der Chordatengene phasengenau überein.

### 3.3.4 Rekombinante Expression des Sacco IF1-Proteins

Um das neue IF-Protein zu analysieren wurde der gesamte ORF von Sacco IF1 in den pET-23a<sup>+</sup>-Vektor umklont. Dazu wurden mit modifizierten Primersequenzen (sense:GTGCACAG**CATATG**AATAAGTAGTGG, antisense:GTCTTCAACGG**GATCC**ATCAGTTTGG) neue Schnittstellen (fett gedruckt) für die Restriktionsenzyme BamH I und Nde I in die Sequenz eingebaut. Die modifizierte cDNA-Sequenz wurde zunächst in einen TA-Vektor kloniert, präparativ isoliert und nach dem Restriktionsverdau in den geschnittenen Expressionsvektor (pET 23a<sup>+</sup>) ligiert. Der Vektor wurde in chemokompetenten *E. coli* TOP10-F'-Zellen vermehrt, isoliert und die Sequenz verifiziert. Für die rekombinante Expression wurden *E. coli*-Zellen des BL21-Stamms verwendet, welche sich durch besonders niedrige Proteaseaktivität auszeichnen. Nachdem die *E. coli* BL21-Zellen transfiziert waren, wurde die Expression zunächst im analytischen Maßstab durchgeführt. Die Induktion mit 1 mM IPTG (Endkonzentration) erfolgte eine Stunde nach dem Neustart einer 2,5 ml Kultur und dauerte vier Stunden. Die Zellen wurden abzentrifugiert, in Laemmli-Puffer aufgenommen, 5 min gekocht und durch diskontinuierliche SDS-PAGE analysiert (→ Kap. 2.2.3). Eine deutliche Bande war im erwarteten Molekulargewichtsbereich (60-65 kD) erkennbar. Die präparative Expression wurde in 250 ml Medium durchgeführt und die erhaltenen Zellen eingefroren. Die als Inclusion-Bodies in den Zellen vorliegenden Proteine wurden extrahiert (→ Kap. 2.2.7.2) und über eine FPLC-Anlage gereinigt (→ Kap. 2.2.7.3). Das Protein lag nun in hoher Reinheit in 8 M Harnstoff vor und wurde durch Zugabe von 10 mM Tris-Base stabilisiert. Die Proteinfraction reagierte im Immunoblot wie erwartet positiv mit dem monoklonalen Antikörper IFA. Untersuchungen zur Filamentbildung des rekombinanten Proteins wurden nach Entfernung des Harnstoffs durch Dialyse gegen Tris-HCl-Puffer pH= 7 - 7,8 möglich. Im Elektronenmikroskop konnten nur Aggregate und keine Filamente nachgewiesen werden. Daraus folgt, dass Sacco IF1 unter den gegebenen Pufferbedingungen keine homopolymeren Filamente bilden kann oder dass möglicherweise ein anderes IF-Protein erforderlich ist um heteropolymere Filamente zu erreichen.

### 3.3.5 Erstellung eines polyklonalen anti-Sacco IF1 Peptidantikörpers

Mit immunfluoreszenzmikroskopischen Untersuchungen sollte das Expressionsmuster des Proteins analysiert werden. Die Voraussetzung für immunhistologische Untersuchungen ist ein spezifischer Antikörper. In der Arbeitsgruppe wurden in der Vergangenheit gute Ergebnisse mit Kaninchenantikörpern gegen kurze Peptide aus dem hoch variablen carboxyterminalen Abschnitt von IF-Proteinen gemacht. Das für die Immunisierung benutzte Peptid hatte die Sequenz NH<sub>2</sub>-CKTTHFSKSKRVLEYA-COOH. Dies entspricht den AS-Resten 429-543. Peptide dieser geringen Größe können jedoch keine ausreichende Immunreaktion auslösen. Die Antigenizität wurde durch Kopplung an Hemocyanin über das zusätzlich angefügte aminotermine Cystein des Peptids erhöht. Dieses Konjugat wurde einem Kaninchen injiziert. Nach sechs Wochen lagen ausreichend hohe Antikörperkonzentrationen vor. Die zellulären Bestandteile des Blutes wurden durch Zentrifugation sedimentiert und das Serum bei -80 ° C bis zur Verwendung gelagert.

Ein Teil der vom Kaninchen erzeugten Antikörper war gegen Hemocyanin gerichtet weshalb eine Konzentrierung der anti-Sacco IF1-AK nötig war. Dazu wurde das Peptid über das zusätzliche aminotermine Cystein über Sulfo/MBS an bromcyanaktivierte Sepharose gekoppelt. Das Rohserum des Kaninchens wurde mit der Sepharose inkubiert wodurch die spezifischen Antikörper mit dem Peptid reagieren konnten. Ungebundene Antikörper wurden durch mehrfaches Waschen auf einer Säule entfernt. Die monospezifischen Antikörper konnten bei niedrigem pH eluiert und sofort durch Zugabe von Tris-Base bei neutralen Pufferbedingungen stabilisiert werden.

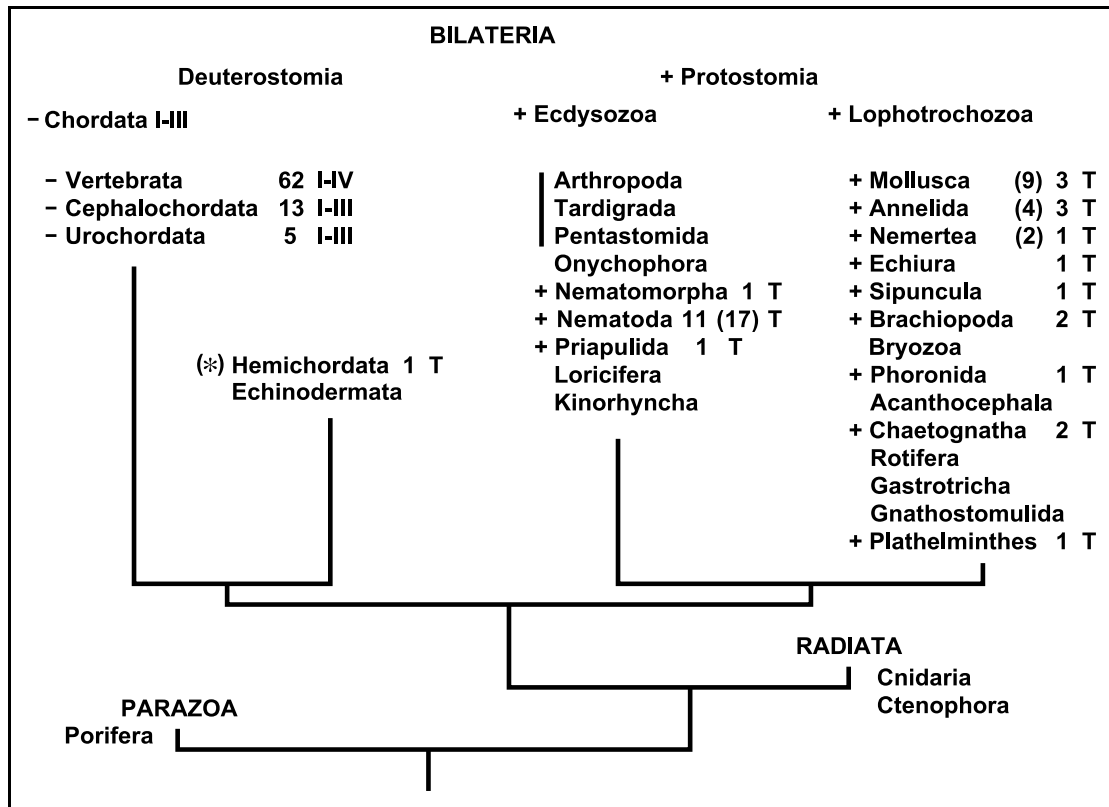
Leider war die Spezifität der Antikörper gering: Es wurde gereinigtes rekombinantes IF-Protein im Western-Blot erkannt, jedoch reagierten die AK im Test mit Gesamtproteinextrakt aus *S. kowalevskii*-Gewebe auch mit einigen unbekanntem Proteinbanden im Molekulargewichtsbereich von 14000 und 96000. Die Antikörper waren nicht monospezifisch und konnten weder für Immunfluoreszenzen noch für Expressionsanalysen eingesetzt werden.



### 3.5 Die Phylogenie der Intermediärfilamentproteine

Die hier betrachteten zytoplasmatischen Intermediärfilamente haben unterschiedliche Eigenschaften. So ist bei protostomischen IF<sup>zyt</sup>-Proteinen meistens ein Laminhomologiesegment in der tail-Domäne zu finden, das in keinem Chordaten-IF<sup>zyt</sup>-Protein vorhanden ist. Alle bekannten Proteine der Protostomier haben immer die lange coil 1b-Subdomäne. Das bislang einheitliche Bild der IF<sup>zyt</sup>-Proteine in Chordaten mit der um 42 AS kürzeren coil 1b-Subdomäne und der kurzen tail-Domäne ohne LHS ist nun um eine Facette reicher geworden. Das Hemichordaten-IF-Protein Sacco IF1 hat nur eine Deletion von 11 AS in der coil 1b-Subdomäne und besitzt einen dem LHS homologen Abschnitt im „tail“.

Die daraus folgende neue Betrachtung der Evolution der zytoplasmatischen IF-Proteine ist in Figur 11 zusammengefasst. Sie basiert auf allen gegenwärtig verfügbaren Sequenzen zytoplasmatischer IF-Proteine. Sicher ist jetzt auch, dass zumindest in *D. melanogaster* nach abgeschlossener Analyse des Genoms (RUBIN ET AL. 2000) keine IF<sup>zyt</sup>-Proteine vorhanden sind und damit die früheren elektronenmikroskopischen Befunde (BARTNIK & WEBER 1989) bestätigt werden. Die Funktion der Intermediärfilamente wird in Arthropoden möglicherweise vollständig durch Mikrotubulibündel ersetzt. Dazu ist bereits bekannt, dass diese in den Nematoda noch in Verbindung mit IF-Fasern auftreten können (KARABINOS ET AL. 2001). Die Grafik zeigt auch, dass wie ZUVOR (ERBER ET AL. 1998) für Echinodermata und Radiata keine Aussage darüber gemacht werden kann, ob IF<sup>zyt</sup>-Proteine vorhanden sind und welche strukturellen Eigenschaften vorliegen. Ebenso konnte bei einem in dieser Arbeit durchgeführten Expressionsscreen der Hydra-cDNA-Bank nur das bekannte Lamin aber kein zytoplasmatisches IF-Protein kloniert werden. Bei den Echinodermata habe ich bereits Expressionsbanken von drei Tieren (*S. purpuratus*, *A. brasiliensis*, *A. rubens*) mit dem IFA-Antikörper und/oder heterologen DNA-Sonden ergebnislos durchsucht. Weiterhin wurden in einer Analyse von 8.293 EST-Sequenzen aus *S. purpuratus* (ZHU ET AL. 2001) keine eindeutigen IF-Proteinsequenzen ermittelt. Ein *S. purpuratus*-EST-Projekt mit über 50.000 isolierten Sequenzen enthielt keine über das bekannte Lamin hinausgehenden Hinweise auf zytoplasmatische IF-Proteine (CAMERON ET AL. 2000A).



Figur 11: Aktualisierte phylogenetische Verteilung der kurzen (-) und langen (+) coil 1b-Subdomäne und die Verbreitung des Laminhomologiesegments in der tail-Domäne (T) zytoplasmatischer IF-Proteine (ERBER ET AL. 1998). Die Aufteilung der protostomischen Tiere in die beiden Gruppen Ecdysozoa und Lophotrochozoa folgt den Ergebnissen der 18S rDNA-Sequenzanalysen von AGUINALDO ET AL. (1997). Die monophyletischen Gruppen Echinodermata und Hemichordata stehen als gemeinsame Schwestergruppe dem Superphylum Chordata gegenüber (TURBEVILLE ET AL. 1994), welches die Vertebrata, Cephalochordata und Urochordata zusammen fasst (CAMERON ET AL. 2000b). Die Gesamtzahl der bekannten IF-Proteinsequenzen protostomischer Phyla ist in Klammern angegeben. Die maximal in einer Spezies bekannten IF<sup>zyt</sup>-Sequenzen sind ohne Klammern geschrieben. Zusätzlich zu den bekannten Referenzen (ERBER ET AL. 1998) wurden der Grafik folgende Informationen hinzugefügt:

Die Anzahl der bekannten IF-Proteinsequenzen für den Anneliden *H. medicinalis* erhöhte sich von eins auf drei (XU ET AL. 1999). Das Genom des Nematoden *C. elegans* enthält 11 zytoplasmatische IF-Gene (KARABINOS ET AL. 2001). Die Anzahl der bekannten IF<sup>zyt</sup>-Proteine von *B. lanceolatum* (Cephalochordata) und *S. clava* (Urochordata) erhöhte sich jeweils um zwei (KARABINOS ET AL. 2000, WANG ET AL. 2000). Die Anzahl von 62 Sequenzen beim Menschen (Vertebraten) folgt aus der Analyse der Rohdaten des Humangenoms (HESSE ET AL. 2001). Die IF-Unterfamilien I-IV sind für die Chordaten angegeben. Der schwarze Balken (|) vor den ersten drei Gruppen der Ecdysozoa veranschaulicht das Fehlen jeglicher zytoplasmatischer IF-Proteine auf der Basis elektronenmikroskopischer Studien (BARTNIK UND WEBER 1989) und der Genomanalyse von *D. melanogaster* (Arthropoda) in RUBIN ET AL. (2000). Neu hinzugekommen ist aus dieser Arbeit das zytoplasmatische IF-Protein des Protostomiers *P. caudatus* und das IF<sup>zyt</sup>-Protein des Hemichordaten *S. kowalevskii*. Das Laminhomologiesegment (T) ist vorhanden, auf die Deletion von 11 AS in „coil 1b“ (siehe Figur 8) wird mit einem Stern (\*) verwiesen. Diese Merkmalskombination ist bislang einmalig. Alle IF-Proteine der Chordaten haben eine Deletion von 42 AS in „coil 1b“ (-) und kein LHS in der tail-Domäne. Sacco IF1 ist das erste deuterostomische IF-Protein, welches im Vergleich mit protostomischen IF<sup>zyt</sup>-Sequenzen nur eine Deletion von 11 AS in „coil 1b“ und darüber hinaus noch das LHS hat.





## 4 Diskussion

Die Intermediärfilamente der tierischen Zellen sind ein wichtiger Bestandteil des Zytoskeletts. Sie ergänzen das bei allen Eukaryonten vorhandene Netzwerk aus Aktinfilamenten und Mikrotubuli, wobei sie im Besonderen die mechanische Widerstandsfähigkeit der Zelle erhöhen. Die Proteine der Intermediärfilamente haben eine langgestreckte Gestalt, die gebildeten Filamente sind dehnbar und aufgrund dieser Eigenschaft gut geeignet um Scherkräften entgegen zu wirken (JANMEY ET AL. 1991). Die IF-Proteine werden in die strukturellen Hauptbereiche head, rod und tail unterteilt. In der rod-Domäne wiederum können die drei  $\alpha$ -helikalen Subdomänen coil 1a, 1b und coil 2 unterschieden werden. Die Wechselwirkung von zwei IF-Proteinen führt zur Ausbildung einer coiled coil-Quartärstruktur im Bereich der rod-Domänen. Durch weitere Zusammenlagerung dieser Dimere entstehen Intermediärfilament-Fasern mit einer Dicke von ca. 10 nm. Die Bildung des Dimers hängt in hohem Maße von der Funktionstüchtigkeit der N- und C-terminalen Abschnitte der rod-Domäne ab. Mutationen in diesen gut konservierten Bereichen führen oft zu schwerwiegenden Krankheiten (siehe Einleitung). Diese medizinische Relevanz und die sich bereits zu Beginn der Untersuchungen abzeichnende hohe Komplexität der humanen IF-Multigenfamilie bildeten einen Ausgangspunkt für Untersuchungen zur molekularen Entstehungsgeschichte der IF-Proteine.

Mit zunehmender Anzahl der klonierten und sequenzierten IF-Proteine stellte sich heraus, dass eine Ableitung hochauflösender molekularer Stammbäume aus den verfügbaren Sequenzen wenig erfolgversprechend war. Dies liegt in mehreren Umständen begründet. Erstens bilden die IF-Proteine Multigenfamilien. Anders als bei den „single-copy-genes“ der ribosomalen RNA gibt es hier schon innerhalb eines Organismus mehrere IF-Proteine die starke Sequenzabweichungen voneinander haben. Selbst bei einer Eingrenzung auf nur einen IF-Proteintyp ist bei allein über 50 humanen Keratinen (HESSE ET AL. 2001) die Erstellung eines molekularen Stammbaumes kaum möglich. Hinzu kommt, dass ein Grund für die Bildung funktionsfähiger Filamente sowohl in der Primärstruktur, als auch in einer einheitlichen Sekundärstruktur der Polypeptide zu suchen ist. Im direkten Vergleich mit „head“ und „tail“ sind die rod-Domänen der IF-Proteine zwar relativ gut konserviert, jedoch nur im Bezug auf die kanonische Abfolge der Heptaden und, selbstverständlich, in den Konsensussequenzen an den Flanken der rod-Domäne. Um die IF-Proteinsequenzen phylogenetisch zu verarbeiten ist eine Konzentration auf molekulare Marker nötig, die jedoch nur eine grobe systematische Einteilung zulassen. Dies bedeutet die Definition diskreter Merkmale, deren Gegenwart oder Abwesenheit die Zuordnung zu großen monophyletischen Abstammungsgemeinschaften begründet.

In diesem Sinn kann man feststellen, dass die coil 1b-Subdomänen der Lamine und protostomischen IF<sup>zyt</sup>-Proteine die gleiche Länge haben, bei den IF<sup>zyt</sup>-Proteinen der Chordaten in diesem Abschnitt jedoch immer eine Deletion von genau 42 AS, also sechs Heptaden vorhanden ist. Auch die tail-Domäne der protostomischen IF-Proteine entspricht in den meisten Fällen jener der Lamine. Dort ist ein Laminhomologiesegment (LHS) vorhanden, welches in den IF<sup>zyt</sup>-Proteinen der Chordaten immer fehlt. Somit können zwei molekulare Marker definiert werden, die protostomische und Chordaten-IF<sup>zyt</sup>-Proteine voneinander trennen.

Die Ähnlichkeit der protostomischen IF<sup>zyt</sup>-Proteine zu den Laminen führte zu der Hypothese, dass Lamine der Ursprung der IF-Proteine seien und aus diesen durch Genduplikation, Deletion von CaaX-Box, LHS und Sequenzdrift die zytoplasmatischen IF-Proteine entstanden sein könnten (WEBER ET AL. 1989, DODEMONT ET AL. 1990, DÖRING UND STICK 1990). Wann jedoch die für Chordaten typischen Deletionen des LHS in der tail-Domäne und vor allem der 42 AS in „coil 1b“ stattgefunden haben, war bis jetzt ungewiss.

#### **4.1 Charakterisierung eines zytoplasmatischen Hemichordaten-IF-Proteins**

In dieser Arbeit habe ich die Klonierung des ersten zytoplasmatischen IF-Proteins aus einem primitiven Deuterostomier beschrieben. Denn Hemichordaten und Echinodermata bilden die gemeinsamen Schwestergruppen des Superphylums Chordata, welches seinerseits aus Uro-, Cephalochordaten und Vertebraten besteht (TURBEVILLE ET AL. 1994). Das IF-Protein Sacco 1 unterscheidet sich in zwei wichtigen Punkten von allen bislang aus dem deuterostomischen Tierstamm charakterisierten IF-Proteinen.

Zunächst ist die coil 1b-Subdomäne nicht um 42 AS sondern lediglich um 11 AS gegenüber den Laminen und protostomischen IF<sup>zyt</sup>-Proteinen verkürzt. Überraschend ist, dass dadurch die kanonische Abfolge hydrophober Aminosäurereste an einer Stelle unterbrochen wird. Die Aufteilung der rod-Sequenz in Heptaden mit unpolaren Aminosäureresten an jeweils erster und vierter Position ist jedoch ein grundlegendes Motiv, welches für die Ausbildung der coiled coil-Konformation notwendig ist (GEISLER UND WEBER 1982).

Zum anderen konnte in der tail-Domäne des Sacco IF1-Proteins das Laminhomologiesegment (LHS) eindeutig nachgewiesen werden. Dieser Abschnitt hat für gewöhnlich eine durchschnittliche Sequenzkonservierung von über 20 % gegenüber den Laminen (WEBER ET AL. 1989). Figur 9 (S. 50) zeigt für Sacco IF1, dass in diesem kurzen Sequenzabschnitt von 112 AS-Resten sechs Aminosäuren invariant und 22 AS-Reste

hoch konserviert sind. Darüber hinaus stimmt auch die Genorganisation weitgehend mit den tail-Domänen der Lamine überein. Die zwei Introne der Sacco IF1-tail-Domäne befinden sich an identischen Positionen wie die Introne der Lamingene. Ein ähnliches Bild zeigt die Genorganisation der rod-Domäne. Auch hier konnten vier hochkonservierte und eine verschobene Intronposition nachgewiesen werden.

Die weiter gehenden Schritte der Charakterisierung umfassten die rekombinante Expression des Sacco 1 Proteins. Mit dem bakteriell produzierten Protein wurde nach chromatografischer Reinigung versucht, *in vitro* Filamente zu erzeugen. Da dies nicht gelang, können für das Sacco 1-Protein keine Aussagen bezüglich des homo- oder heteropolymeren Charakters seiner Filamente gemacht werden.

#### **4.2 Die kurze coil 1b-Variante entstammt den ursprünglichen Chordaten**

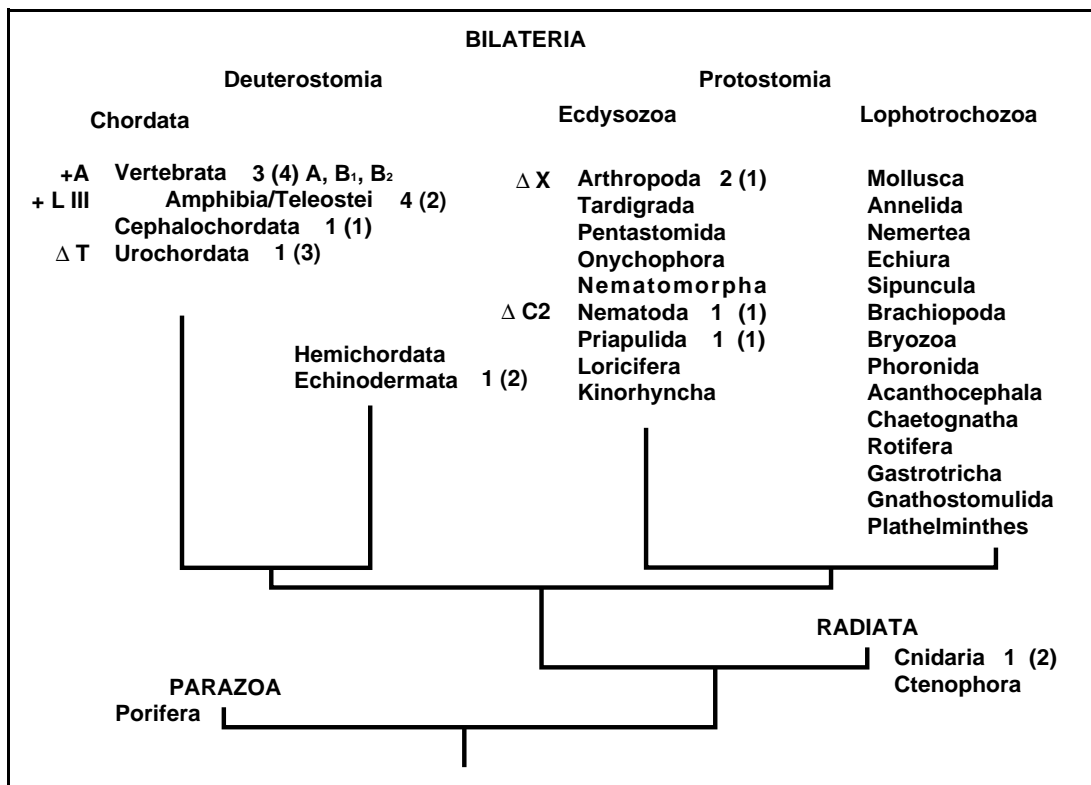
Im Mittelpunkt der meisten Experimente dieser Arbeit stand die Klärung der Frage, ob die Deletion von 42 AS in der coil 1b-Subdomäne zytoplasmatischer IF-Proteine chordatenspezifisch ist, oder gar als ein molekulares Merkmal aller Deuterostomier gelten kann. Die Analyse der Sacco 1-Proteinsequenz zeigt, dass dort eine coil 1b-Subdomäne vorhanden ist, die länger als jene der Chordaten ist. Zusätzlich ist im „tail“ ein LHS vorhanden. Trotz einer kleinen Deletion von 11 AS ist es klar, dass der entscheidende molekulare Marker, die Verkürzung der coil 1b-Subdomäne um sechs Heptaden ein Merkmal der Chordaten bleibt. Damit wird eine monophyletische Betrachtung der Chordaten unterstützt, die bereits durch die diagnostischen Aminosäuren der Muskelaktine nahe gelegt wurde (BOVENSCHULTE & WEBER 1997). Deshalb steht nun außer Frage, dass ein einziges IF-Protein mit kurzer coil 1b-Subdomäne in der Stammart der ursprünglichen Chordaten vorhanden war. Aus diesem entwickelte sich bereits in den primitiven Urochordaten eine Multigenfamilie mit mindestens fünf Mitgliedern und einer Aufteilung in drei IF-Typen (I-III). Sowohl die Typ III-Proteine der homopolymeren Filamente als auch die Keratine (I, II) sind dort schon vorhanden (WANG ET AL. 2000 UND 2002). Der Tunikat *C. intestinalis* wurde bereits in der Vergangenheit auf die zytoplasmatischen IF-Proteine untersucht. Bislang fehlten jedoch die Proteine der homopolymeren IFs und der allgemeine Nachweis von Neurofilament-Proteinen in primitiven Chordaten. Das Typ III IF-Protein IF A2 und das Lamin L2 konnten in dieser Arbeit kloniert werden. Sie bildeten in Zusammenarbeit mit Dr. J. Wang den Ausgangspunkt für die Klonierung von vier weiteren zytoplasmatischen IF-Proteinen IF B, IF C, IF D und IF F, die zunächst als 2. Gruppe

geführt wurden (RIEMER ET AL. 2000). Kürzlich durchgeführte Untersuchungen des Expressionsmusters deuten aber darauf hin, dass jeweils nur ein Gen für diese fünf zytoplasmatischen IF Proteine und das Lamin in *Ciona* vorhanden ist (WANG ET AL. 2002).

### 4.3 Lamine: Der Grundtyp und seine Varianten

Die im Tierreich am weitesten verbreitete Laminvariante ist der B-Typ. Er wurde bereits in Cnidariern nachgewiesen und das Protein kann folgendermaßen charakterisiert werden. Es hat eine vergleichsweise kurze head-Domäne von 23 bis 50 AS Länge und einen rod-Abschnitt mit weitgehend konstanter Länge von 365 AS. In der tail-Domäne befinden sich NLS, 105-Box und die CaaX-Box (ERBER ET AL. 1999). Im Lamin von *C. elegans* ist die coil 2-Subdomäne um 14 AS (zwei Heptaden) verkürzt (RIEMER ET AL. 1993). Ein weiteres ungewöhnliches Lamin wurde beim Arthropoden *D. melanogaster* nachgewiesen. Die Fruchtfliege hat nachweislich zwei Lamingene. Zum einen das B-Typ Lamin Dmo und ein Lamin C, welches als Besonderheit keine CaaX-Box hat (RIEMER UND WEBER 1994). In der Gruppe der Ecdysozoa ist noch ein Lamin von *P. caudatus* bekannt, aus der Gruppe der Lophotrochozoa konnten bislang keine Lamine isoliert werden (Figur 12). Zwar gibt es biochemische Hinweise auf laminverwandte Proteine in Mollusken (MAUL ET AL. 1987, DESSEV ET AL. 1990, DESSEV & GOLDMANN 1990) es wurden aber bislang keine Sequenzen publiziert.

Im Unterschied zu den Protostomiern wurde bei den Chordaten eine größere Anzahl von Laminen gefunden und teilweise ausführlich charakterisiert. Auch hier kann man von den B-Typ-Laminen als Grundtyp ausgehen weil in zwei Echinodermata jeweils ein Lamin nachgewiesen wurde. Bei den Tunikaten als wichtigen Mitgliedern der primitiven Chordaten konnten in der Vergangenheit je ein Lamin aus *S. clava* und *M. oculata* isoliert werden (RIEMER ET AL. 2000). Im Rahmen dieser Arbeit wurde ein Lamin aus *C. intestinalis* kloniert. In allen Tunikaten-Laminen konnte immer das Fehlen eines ca. 90 AS langen Teilstücks in der tail-Domäne gezeigt werden. Aufgrund der fehlenden Sequenzinformationen aus den Nachbargruppen der Tunikaten ist eine allgemeine Gültigkeit dieses Merkmals für alle Urochordaten noch nicht gesichert. Da die Deletion aber zumindest für alle bislang überprüften Ascidien gut konserviert ist, kann sie als molekularer Marker der Tunikaten bezeichnet werden.



Figur 12: Lamine und ihre Varianten. Nach der lateinischen Bezeichnung der Tiergruppen ist zunächst die Gesamtzahl der Lamine angegeben, die in jeweils einer Art dieser Tiergruppe gefunden wurde. In Klammern steht die Anzahl der verschiedenen Arten aus denen Lamine nachgewiesen sind.  $\Delta T$  repräsentiert die Deletion von 90 AS in der tail-Domäne der Tunikaten (RIEMER ET AL. 2000). Jede bislang untersuchte Vertebraten-Spezies hat die drei Laminvarianten B<sub>1</sub>, B<sub>2</sub> und A. Der A-Typ setzt sich von den B-Typ-Laminen durch eine zusätzliche Insertion von ungefähr 90 AS (+ A) kurz vor der CaaX-Box ab (STICK 1992, LIN & WORMAN 1993). Darüber hinaus haben die Vertebraten-Lamine einen Abschnitt von ca. 15 sauren AS-Resten (ERBER ET AL. 1999). Bei Fischen und Amphibien kann man ein weiteres Lamin (L III) vom B-Typ nachweisen, welches vorwiegend in Oozyten exprimiert ist (HOFEMEISTER ET AL. 2000).

Im protostomischen Tierreich wurden beim Arthropoden *D. melanogaster* zwei Lamine kloniert. Im Lamin C fehlt die CaaX-Box ( $\Delta X$ ). Der Nematode *C. elegans* verfügt nur über ein Lamin, bei dem eine Deletion von 14 AS in coil 2 entdeckt wurde ( $\Delta C2$ ). In Echinodermata und Cnidariern wurde in zwei Arten jeweils ein Lamin, in den Lophotrochozoa bisher kein Lamin kloniert.

Bei den Cephalochordaten gibt es keine Abweichungen vom Grundtyp, diese lassen sich erst bei den wesentlich höher entwickelten Vertebraten feststellen. Folgende Ergebnisse sind dokumentiert: Erstens werden bei Vertebraten mehrere Lamine exprimiert (GERACE AND BLOBEL 1980). Die Echinodermata und primitiven Chordaten haben im Gegensatz dazu jeweils nur ein Lamingen (ERBER ET AL. 1999). Zweitens befindet sich in der tail-Domäne der Vertebraten-Lamine ein Cluster von sieben bis 15 sauren AS-Resten zwischen der 105-Box und der CaaX-Box (ERBER ET AL. 1999). Dieses ist sowohl bei den Laminen B<sub>1</sub> und B<sub>2</sub> als auch den zusätzlichen Typ-A-Laminen vorhanden, welche die dritte Abweichung vom Grundtyp der Lamine aufweisen. Die Typ A-Lamine haben eine einmalige Insertion von ca. 90 AS-Resten in der tail-Domäne und werden nur in differenzierten Zellen der Vertebraten exprimiert (ERBER ET AL. 1999). Undifferenzierte



Daraus ergibt sich eine Übersicht (Figur 12), die die strukturelle Vielfalt der Lamine beschreibt. Diese als Ausgangspunkt der IF-Evolution diskutierten Proteine haben ihrerseits unabhängig von der Vielfalt zytoplasmatischer IF-Proteine eigene strukturelle Modifikationen erfahren. Alle diese Änderungen sind jedoch spezifisch für ein Phylum (Urochordata, Vertebrata, etc.) und nicht wie bei den zwei Prototypen der zytoplasmatischen IF-Proteine für einen gesamten Zweig des Tierreichs (Figur 13).

#### 4.4 Die Evolution der zytoplasmatischen IF-Proteine

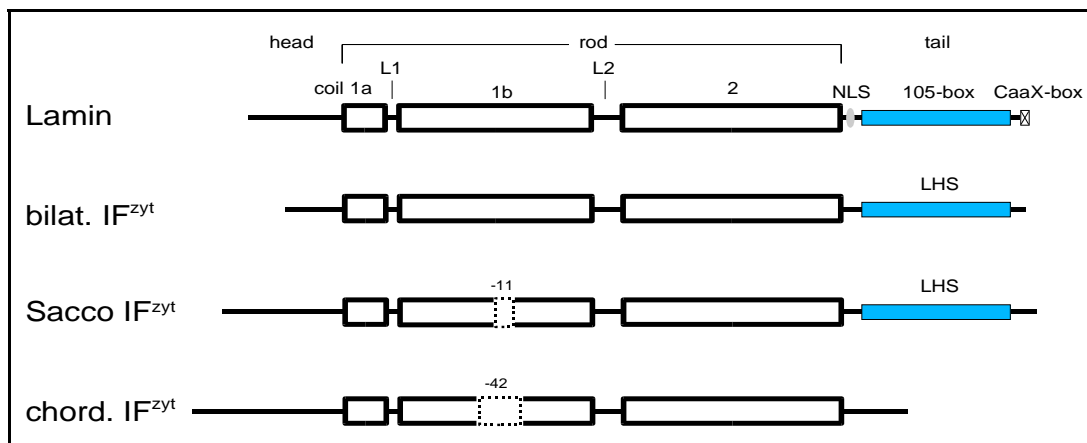
In Anbetracht der Veränderungen, die bei der Laminstruktur dokumentiert sind, erscheint die Theorie von einer Entstehung der zytoplasmatischen IF-Proteine aus einem Lamin durchaus denkbar (WEBER ET AL. 1989). Wie beim Lamin C in *Drosophila* kann es zum Verlust der CaaX-Box gekommen sein. Und durch Punktmutationen oder Deletionen ging das Signal zur Kernlokalisierung verloren. Äquivalente Transfektionsexperimente mit cDNA-Mutanten wurden bereits 1988 durchgeführt (LOEWINGER UND McKEON 1988). Es konnte gezeigt werden, dass Laminproteine ohne CaaX-Box noch in den Kern gelangen und im Nukleoplasma filamentöse Strukturen bilden. Die Entfernung größerer Stücke der tail-Domäne einschließlich des Kernlokalisierungssignals führte zur Ausbildung zytoplasmatischer Filamentstrukturen aus dem modifizierten Lamin.

In der vorliegenden Arbeit konnte das IF-Protein eines primitiven Deuterostomiers kloniert werden. Die coil 1b-Subdomäne ist wesentlich länger als die aller bekannten Chordaten-IF-Proteine. Durch den zusätzlichen Nachweis des LHS ist es schlüssig, von einem gemeinsamen zytoplasmatischen IF-Protein für die Stammarten der Protostomier und Deuterostomier auszugehen. Es ist in Figur 14 mit „bilat. IF<sup>zyt</sup>“ bezeichnet, weil dieser IF-Proteintyp in den Stammarten der Protostomier und Deuterostomier, also beiden großen bilateral symmetrischen Tierstämmen wahrscheinlich vorhanden war. Die Struktur dieses Proteins blieb im protostomischen Zweig weitgehend unverändert. Alle IF-Proteine der Lophotrochozoa sind in ihrer Domänenstruktur gleich gestaltet (BOVENSCHULTE ET AL. 1995, ERBER ET AL. 1998). Deletionen in der tail-Domäne treten nur bei fünf der elf IF<sup>zyt</sup>-Proteine in *C. elegans* auf (KARABINOS ET AL. 2001). Auch das in dieser Arbeit klonierte IF<sup>zyt</sup>-Protein aus *P. caudatus* weist keine strukturellen Veränderungen auf. Die Arthropoda, als insgesamt artenreichste Gruppe exprimieren überraschenderweise keine zytoplasmatischen IF-Proteine (BARTNIK UND WEBER 1989). Auch die IF-Gene sind vollständig aus dem Genom verschwunden. Dies ist zumindest durch das komplett sequenzierte Genom von

*D. melanogaster* dokumentiert (RUBIN ET AL. 2000). Die mechanische Stabilität auf zellulärer Ebene wird bei Arthropoden durch Bündel aus Mikrotubuli im nötigen Umfang gewährleistet (KARABINOS ET AL. 2001).

Ausgehend von der Stammart der Bilateria wurde die IF-Proteinstruktur auch im frühen deuterostomischen Zweig des Tierreichs nur geringfügig verändert. Das LHS und weite Teile der coil 1b-Subdomäne sind bei dem Hemichordaten-IF<sup>zyt</sup>-Protein noch vorhanden. Der erfolglose Versuch mit verschiedenen Techniken weitere IF<sup>zyt</sup>-Proteine zu klonieren lässt vermuten, dass es sich bei Hemichordaten möglicherweise um keine große IF-Multigenfamilie handelt.

Da die Deletion beim Sacco IF1-Protein (Figur 14) kein Vielfaches von sieben ist, kommt es wie in der coil 2-Subdomäne nach der achten Heptade, hier nach der zehnten Heptade der coil 1b-Subdomäne, zu einem „stutter“ der die  $\alpha$ -helikale Sekundärstruktur unterbricht. Die pelagische Lebensweise der Tiere könnte ein Grund für die geringe Präsenz der IF<sup>zyt</sup>-Proteine sein. Denn Hemichordaten leben im Druckgleichgewicht mit ihrer Umgebung in relativ starren Röhren, die Enteropneusta im, und die Pterobranchia über dem Boden. Möglicherweise sind sie kaum auf die Stabilisierung ihrer zellulären Struktur durch die hoch belastbaren Intermediärfilamente angewiesen.



Figur 14: Strukturprinzipien der Intermediärfilamentproteine (Legende: siehe Figur 1, S. 3). Diese neue Fassung wurde gegenüber der Darstellung zu Beginn dieser Arbeit um die neu entdeckte Proteinvariante Sacco IF 1 im Hemichordaten *S. kowalevskii* erweitert. Das LHS im „tail“ und die kleine Deletion von 11 AS in „coil 1b“ setzen die Gegenwart eines langen IF-Proteins, so wie es in rezenten Protostomiern auftritt, auch für die Stammart der Deuterostomier voraus. Aus diesem Grund wird die lange IF<sup>zyt</sup>-Variante nicht länger als protostomisch sondern als „IF<sup>zyt</sup>-Protein der Bilateria“ bezeichnet. IF-Proteine mit einer um 42 AS verkürzten rod-Domäne und ohne LHS wurden erstmals in der Stammart der Chordaten etabliert. Alle IF<sup>zyt</sup>-Proteine rezenter Chordaten stammen deshalb von einem einzigen Gen ab, in dem beide Bereiche deletiert wurden. Innerhalb einer vergleichsweise kurzen Zeit konnten in Cephalochordaten und den primitiveren Urochordaten Multigenfamilien entstehen. Neurofilamente (Typ IV IF-Proteine) werden nach erfolglosem Nachweis entsprechender Proteine in Cephalochordaten und Urochordaten weiterhin als Merkmale der Vertebraten geführt.



#### 4.5 Wie viele IF-Proteine braucht ein Tier ?

Die im Jahre 1989 (BARTNIK UND WEBER) durchgeführten elektronenmikroskopischen Untersuchungen bei 15 verschiedenen Echinodermata lieferten nur schwache Hinweise auf zytoplasmatische IFs. Die Filamente hatten einen Durchmesser von 5 nm, zeigten nur in einem Fall Reaktivität mit dem IFA-Antikörper und wurden damals zumindest als IF-ähnlich bewertet. Mittlerweile wurden in vielen Arbeitsgruppen zehntausende von Sequenzen aus Echinodermen ermittelt (CAMERON ET AL. 2000a, ZHU ET AL. 2001), zytoplasmatische IF-Proteine konnten jedoch nie nachgewiesen werden. Die drei in dieser Arbeit überprüften cDNA-Banken von *A. brasiliensis*, *A. rubens* und *S. purpuratus* enthalten ebenfalls keine cDNAs von IF<sup>zyt</sup>-Proteinen.

Selbstverständlich ist es viel schwieriger das Fehlen eines Proteins als dessen Präsenz zu beweisen. Letztlich ist dazu ein abgeschlossenes Genom-Sequenzprojekt (wie im Fall von *Drosophila*) notwendig. Ich spekuliere aber, dass zytoplasmatische IF-Proteine auch in Echinodermata fehlen. Vor allem im Hinblick auf das abgeschlossene *Drosophila* Genomprojekt sehe ich aufgrund der konvergenten Etablierung eines Exoskeletts die Möglichkeit einer konvergenten Deletion aller zytoplasmatischen IF-Proteine in den Echinodermata. Denn wenn die Hauptfunktion der Filamente in der Stabilisierung gegenüber mechanischer Belastung steht (JANMEY ET AL. 1991), sollte ein nahezu kräftefreier Zellverband mit einer geringeren Anzahl von IF<sup>zyt</sup>-Proteinen auskommen können. Zudem kann die Synthese von Proteinen, die keinen Vorteil in der Fitness bedeuten umgekehrt zum unnötigen Verbrauch biologischer Ressourcen führen und damit einen evolutiven Nachteil darstellen. Die Repression der Gene und ihr sukzessives Verschwinden aus dem Genom wären die Folge.

Im Folgenden erläutere ich eine Hypothese, die eine Korrelation zwischen der Größe einer IF-Multigenfamilie und den mechanischen Belastungen des Organismus durch seinen Lebensraum herstellt: So steht die Körperhülle des Nematoden *C. elegans* unter einem erhöhten Innendruck. Der Nematode zeigt eine hohe Beweglichkeit die möglicherweise zu Scherkräften innerhalb des Zellverbands führt. Dort konnte eine IF-Multigenfamilie von 11 Mitgliedern nachgewiesen werden von denen alleine vier für die Ontogenese unabdingbar sind (KARABINOS ET AL. 2001).

In Arthropoden zeigt sich ein umgekehrtes Bild. Sie besitzen ein sehr steifes Exoskelett und nur die einzelnen Segmente sind beweglich miteinander verbunden. Die von außen und innen aufkommenden Kräfte müssen also möglicherweise nicht allein durch den Zellzusammenhalt aufgefangen werden. Die Anzahl der zytoplasmatischen IF-Proteine ist null. Anstelle der dehnbaren IFs gibt es dort steife Mikrotubuli-Bündel.

Im deuterostomischen Zweig des Tierreichs lässt sich folgendes feststellen. Die mechanische Belastung der Zellstruktur steigt meiner Ansicht nach von den Urochordaten über die Cephalochordaten zu den Vertebraten stetig an. Leben z.B. Tunikaten noch in zähen Hüllen aus Polysacchariden und sind nur der Meeresströmung ausgesetzt, so gilt es bei *B. lanceolatum* bereits die Kräfte schneller Muskelbewegungen über die Chorda auf den gesamten Körper zu verteilen und zur Fortbewegung zu nutzen. Die an Land lebenden Vertebraten müssen darüber hinaus die gesamte Zellmasse des Körpers an das Skelett binden um z.B. der Schwerkraft und mechanischen Belastungen entgegen zu wirken. Im Sinne dieser Reihenfolge findet man auch eine stete Zunahme exprimierter IF<sup>zyt</sup>-Proteine.

Zur Problematik der hohen Anzahl zytoplasmatischer IF-Proteine bei Vertebraten, bzw. der Gegenwart einer komplexen Multigenfamilie existiert eine andere attraktive Hypothese (HESSE ET AL. 2001). Ein wesentlicher Aspekt in der Evolution des Tierreichs ist demnach die Zunahme unterschiedlicher Zelltypen und Gewebe. Sie geht Hand in Hand mit der Etablierung von Multigenfamilien, deren Proteine bei ähnlicher Struktur unterschiedlich ausgeprägte Funktionen wahrnehmen können. So auch die Multigenfamilie der zytoplasmatischen IF-Proteine. Die hohe Zahl der IF<sup>zyt</sup>-Proteine bei den Vertebraten wäre also eine Konsequenz der gestiegenen Anzahl unterschiedlicher Zell- und Gewebetypen.

#### **4.6 Kein Königsweg zur Identifizierung von IF-Proteinen**

Die Frage, ob zytoplasmatische IF-Proteine in Echinodermata exprimiert werden, oder ob sie möglicherweise gar nicht mehr vorhanden sind, wird wahrscheinlich unbeantwortet bleiben, bis entsprechende genomische Sequenzierungsprojekte abgeschlossen sind. Auch für die Aufklärung des eigentlichen Ursprungs der IF<sup>zyt</sup>-Proteine werden noch umfangreichere Untersuchungen der Radiata notwendig sein, weil die bisherigen Klonierungsversuche bei *H. attenuata* erfolglos waren.

Zusammenfassend lässt sich also feststellen, dass es keinen selektiven methodischen Ansatz gibt, der das Potential hat, alle IF-Proteine einschließlich der Lamine eines Tieres zu klonieren. Die geringe Sequenzkonservierung der IF-Proteine steht einer

systematischen Durchmusterung mit cDNA-Sonden entgegen obwohl dieser Ansatz z.B. acht der elf IF<sup>zyt</sup>-Proteine in *C. elegans* identifizierte (DODEMONT ET AL. 1994). Die Konzentration auf die gut konservierten Bereiche an den Flanken der rod-Domäne und die Verwendung entsprechender PCR-Primer stößt mit der durch die Degeneration der AS-Codons bedingten Vielzahl an Nukleotidkombinationen schnell an ihre Grenzen. Vor allem die in den Flanken vorhandenen Aminosäuren Arginin und Leucin, für die jeweils sechs Codons existieren, führen zu einer erschwerenden Vielzahl an Primerkombinationen. Trotzdem war dieser Ansatz in einigen Fällen erfolgreich (ERBER ET AL. 1998 UND 1999).

Der IFA-Expressionsscreen nutzt im Unterschied zu den genannten Methoden die Spezifität des monoklonalen Antikörpers IFA um das charakteristische Epitop „YRKLEEGEE“ am Ende der rod-Domäne zu erkennen. Unabhängig von der immer gegebenen Notwendigkeit die Reaktionsbedingungen während der Antikörperinkubation und Detektion optimal einzustellen, existieren jedoch auch native IF-Proteine mit denen der IFA-Antikörper nicht reagiert. Dazu ist bereits der Austausch des Arginins (R) durch ein Lysin (K) ausreichend (RIEMER ET AL. 1991). Deshalb müssen die Methoden sich gegenseitig ergänzend verwendet werden, um die Unspezifitäten jeder Technik zu überbrücken.

Übrig bleibt die Genomanalyse zur umfangreichen Identifikation neuer IF-Proteine. Dieser methodische Ansatz erfährt eine zunehmende Popularität, weil ausgezeichnete Sequenzier-Apparaturen immer erschwinglicher werden. Der große Vorteil eines vollständig sequenzierten Genoms ist sein Potential, gleichzeitig viele offene Fragen beantworten zu können und noch viel mehr neue Fragen für die Forschung zu stellen. Es bleibt zu hoffen, dass entsprechende Projekte bei Cnidariern und Echinodermata mittelfristig initiiert und abgeschlossen werden können.

## 5 Zusammenfassung

Intermediärfilamente sind ein wichtiger Bestandteil des Zytoskeletts tierischer Zellen. Ihre Hauptfunktion besteht in der Begegnung mechanischer Belastung um den Zellen ein hohes Maß an Stabilität zu verleihen. Die Proteine der zytoplasmatischen Intermediärfilamente der Vertebraten werden auf Grund der gewebe- und zellspezifischen Expressionsmuster, der Aminosäuresequenzen und der Genorganisationen in vier Typen (I-IV) unterteilt. Die Lamine bilden die fünfte Gruppe der IF-Proteine und sind am Aufbau der inneren Kernmembran beteiligt. Die zellspezifischen Expressionsmuster der IF-Proteine und ihre Organisation in einer komplexen Multigenfamilie regten am Ende der 80er Jahre Untersuchungen zur IF-Evolution an.

Es gibt zwei Prototypen zytoplasmatischer IF-Proteine. Die protostomischen IF-Proteine teilen viele strukturelle Eigenschaften mit den Laminen, die IF-Proteine der Chordaten sind jedoch immer um 42 AS in der coil 1b-Subdomäne verkürzt und in der tail-Domäne fehlt das Laminhomologiesegment. Die vorliegende Arbeit liefert weitere Sequenzinformationen über IF-Proteine in deuterostomischen und protostomischen Invertebraten und beantwortet die Frage nach dem Ursprung der IF-Multigenfamilie im Superphylum Chordata.

Die Arbeiten mit dem Priapuliden *P. caudatus* ermöglichten zunächst die Klonierung eines zytoplasmatischen IF-Proteins. Es entspricht dem protostomischen Grundtyp und bestätigt die weite Verbreitung der langen IF-Proteinvariante für das Superphylum Ecdysozoa. Sowohl die coil 1b-Subdomäne als auch die tail-Domäne zeigen keine Deletionen. Im deuterostomischen Tierreich wurde aus dem primitiven Chordaten (Urochordaten) *C. intestinalis* das zytoplasmatische IF-Protein IF A kloniert und damit die Grundlage zur anschließenden Klonierung und Charakterisierung einer mindestens fünfköpfigen IF-Multigenfamilie gelegt. Das ebenfalls klonierte *Ciona*-Lamin hat eine Deletion von 90 AS in der tail-Domäne, die bereits bei den Tunikaten *S. clava* und *M. oculata* bekannt ist und eine Einordnung dieser Deletion als molekularer Marker der Tunikaten rechtfertigt.

Die Analyse des Hemichordaten *S. kowalevskii* führte schließlich zur ersten Klonierung und Charakterisierung eines zytoplasmatischen IF-Proteins in den primitiven Deuterostomiern. Das IF-Protein Sacco 1 hat ein Laminhomologiesegment in der tail-Domäne und die coil 1b-Subdomäne ist nicht um 42 AS verkürzt. Deshalb können die bislang für alle IF<sup>zyt</sup>-Proteine der Chordaten belegten Deletionen in der rod- und tail-Domäne nicht mehr als Merkmal aller Deuterostomier diskutiert werden. Statt dessen ist nun sicher, dass der kurze IF-Prototyp (Deletion von 42 AS) mit der Stammart der

---

Chordaten entstanden ist und sich daraus durch Genduplikationen und Sequenzdrift die Multigenfamilie mit den zytoplasmatischen IF-Typen I-IV entwickelte. Im Gegensatz dazu ist für die primitiven Deuterostomier von einem langen IF-Protein auszugehen, welches mit den protostomischen IF-Proteinen vergleichbar ist.

Die Versuche mit unterschiedlichen Techniken zytoplasmatische IF-Proteine aus verschiedenen Echinodermata und Cnidariern zu klonieren zeigen, dass die abschließende Auflösung der noch unbeantworteten Fragen zur IF-Evolution wahrscheinlich der vollständigen Genomsequenzierung bedürfen. Möglicherweise wird erst dann deutlich, ob Echinodermata, wie die Arthropoda, ihre zytoplasmatischen IF-Proteine nachträglich verloren haben und zu welchem Zeitpunkt der Evolution IF-Proteine als dritte Komponente des Zytoskeletts erstmalig auftauchten. Hierzu stehen die Cnidarier als primitive Eumetazoa noch zur Diskussion. An einem gemeinsamen Ursprung der zytoplasmatischen IF-Proteine für Proto- und Deuterostomier aus einem mutierten Lamingen sollte es hingegen keine Zweifel mehr geben.

## 6 Referenzen

- Aebi, U., Cohn, J., Buhle, L., Gerace, L. (1986)**: The nuclear lamina is a meshwork of intermediate-type filaments. *Nature* **323**, 560-564.
- Aguinaldo, A.M.A., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., Lake, J.A. (1997)**: Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489-492.
- Aviv, H., Leder, P. (1972)**: Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. USA*, **69**, 1408-1412.
- Bartnik, E., Weber, K. (1989)**: Widespread occurrence of intermediate filaments in invertebrates; common principles and aspects of diversification. *Eur. J. Cell. Biol.* **50**, 17-33.
- Blin, N., Stafford, D.W. (1976)**: A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.* **3**, 2303-2308.
- Bonne, G., Di Barletta, M.R., Varnous, S., Becane, H.M., Hammouda, E.H., Merlini, L., Muntoni, F., Greenberg, C.R., Gary, F., Urtizberea, J.A., Duboc, D., Fardeau, M., Toniolo, D., Schwartz, K. (1999)**: Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **21**, 285-288.
- Bovenschulte, M., Riemer, D., Weber, K. (1995)**: The sequence of a cytoplasmic intermediate filament (IF) protein from the annelid *Lumbricus terrestris* emphasizes a distinctive feature of protostomic IF proteins. *FEBS letters* **360**, 223-226.
- Bovenschulte, M., Weber, K. (1997)**: Deuterostomic actin genes and the definition of the chordates: cDNA cloning and gene organization for cephalochordates and hemichordates. *J. Mol. Evol.* **45**, 653-660.
- Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray, G.A., Etensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H., Hood, L. (2000a)**: A sea urchin genome project: sequence scan, virtual map, and additional resources. *Proc. Natl. Acad. Sci. USA* **97**, 9514-9518.
- Cameron, C.B., Garey, J.R. and Swalla, B.J. (2000b)**: Evolution of the chordate body plan: New insights from phylogenetic analyses of deuterostome phyla. *Proc. Natl. Acad. Sci. USA* **97**, 4469-4474.
- Dessev, G., Goldman, R. (1990)**: The oocyte lamin persists as a single major component of the nuclear lamina during embryonic development of the surf clam. *Int. J. Dev. Biol.* **34**, 267-274.
- Dessev, G.N., Iovcheva-Dessev, C., Goldman, R.D. (1990)**: Lamin dimers. Presence in the nuclear lamina of surf clam oocytes and release during nuclear envelope breakdown. *J. Biol. Chem.* **265**, 12636-12641.
- Dodemont, H., Riemer, D. and Weber, K. (1990)**: Structure of an invertebrate gene encoding cytoplasmic intermediate filament proteins: implications for the origin and the diversification of IF proteins. *EMBO J.* **9**, 4083-4094.
- Dodemont, H., Riemer, D., Ledger, N., Weber, K. (1994)**: Eight genes and alternative RNA processing pathways generate an unexpectedly large diversity of cytoplasmic intermediate filament proteins in the nematode *Caenorhabditis elegans*. *EMBO J.* **13**, 2625-2638.
- Döring, V., Stick, R. (1990)**: Gene structure of nuclear lamin LIII of *Xenopus laevis*: a model for the evolution of IF proteins from a lamin-like ancestor. *EMBO J.* **9**, 4073-4081.

- Erber, A., Riemer, D., Bovenschulte, M., Weber, K. (1998)** : Molecular phylogeny of metazoan intermediate filament proteins. *J. Mol. Evol.* **47**, 751-762
- Erber, A. (1999)** : Die Multigenfamilie der Intermediärfilamentproteine der Metazoa; Eine phylogenetische Analyse. *Cuvillier* **3-89712-631-1**, 8.
- Erber, A., Riemer, D., Hofemeister, H., Bovenschulte, M., Stick, R., Panopoulou, G., Lehrach, H., Weber, K. (1999)** : Characterization of the Hydra lamin and its gene: A molecular phylogeny of metazoan lamins. *J. Mol. Evol.* **49**, 260-271.
- Fuchs, E., Weber, K. (1994)** : Intermediate filaments: Structure, dynamics, function and disease. *Ann. Rev. Biochem.* **63**, 345-382
- Fuchs, E., Cleveland, D.W. (1998)** : A structural scaffolding of intermediate filaments in health and disease. *Science* **279**, 514-519.
- Geisler, N., Weber, K. (1982)** : The amino acid sequence of chicken muscle desmin provides a common structural model for intermediate filament proteins. *EMBO J.* **1**, 1649-1656.
- Geisler, N., Heimburg, T., Schünemann, J., Weber, K. (1993)** : Peptides from the conserved ends of the rod domain of desmin disassemble intermediate filaments and reveal unexpected structural features: a circular dichroism, Fourier transform, infrared, and electron microscopic study. *J. Struct. Biol.* **110**, 205-214.
- Gerace, L., Blobel, G., (1980)** : The nuclear envelope lamina is reversibly depolymerized during mitosis. *Cell* **19**, 277-287.
- Harborth, J., Elbashir, S.M., Bechert, K., Tuschl, T., Weber, K. (2001)** : Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J. Cell Sci.* **114**, 4557-4565.
- Hatzfeld, M., Weber, K. (1990)** : The coiled coil of *in vitro* assembled keratin filaments is a heterodimer of type I and II keratins; use of site-specific mutagenesis and recombinant protein expression. *J. Cell Biol.* **110**, 1199-1210.
- Hatzfeld, M., Weber, K. (1992)** : A synthetic peptide representing the consensus sequence motif at the carboxy-terminal end of the rod domain inhibits intermediate filament assembly and disassembles preformed filaments. *J. Cell Biol.* **116**, 157-166.
- Herrmann, H., Häner, M., Brettel, M., Müller, S.A., Goldie, K.N., Fedtke, B., Lustig, A., Franke, W.W., Aebi, U. (1996)** : Structure and assembly properties of the intermediate filament protein vimentin: the role of its head, rod and tail domains. *J. Mol. Biol.* **264**, 933-953.
- Herrmann, H., Aebi, U. (1998)** : Intermediate filament assembly: fibrillogenesis is driven by decisive dimer-dimer interactions. *Curr. Opin. Struct. Biol.* **8**, 177-185.
- Herrmann, H., Aebi, U. (1999)** : Intermediate filament assembly: Temperature sensitivity and polymorphism. *Cell. Mol. Life Sci.* **55**, 1416-1431.
- Hesse, M., Magin, T.M. and Weber, K. (2001)** : Genes for intermediate filament proteins and the draft sequence of the human genome; novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. *J. Cell Sci.* **114**, 2569-2575.
- Hofemeister, H., Weber, K., Stick, R., (2000)** : Association of prenylated proteins with the plasma membrane and the inner nuclear membrane is mediated by the same membrane-targeting motifs. *Mol. Biol. Cell.* **11**, 3233-3246.

- Hofemeister, H., Kuhn, C., Franke, W.W., Weber, K., Stick, R., (2002)** : Conservation of the gene structure and membrane-targeting signals of germ cell-specific lamin LIII in amphibians and fish. *Eur. J. Cell Biol.* **81**, 51-60.
- Hutchison, C.J., Alvarez-Reyes, M., Vaughan, O.A. (2001)** : Lamins in disease: why do ubiquitously expressed nuclear envelope proteins give rise to tissue-specific disease phenotypes? *J. Cell S.* **114**, 9-19.
- Janmey, P., Euteneuer, U., Traub, P., Schliwa, M. (1991)** : Viscoelastic properties of vimentin compared with other filamentous biopolymer networks. *J. Cell Biol.* **113**, 155-160.
- Julien, J.P. (1997)**: Neurofilaments and motor neuron disease. *Trends Cell Biol.* **7**, 243-249.
- Karbinos, A., Riemer, D., Panopoulos, G., Lehrach, H., Weber, K. (2000)** : Characterization and tissue specific expression of the two keratin subfamilies of intermediate filament proteins in the cephalochordate *Branchiostoma*. *Eur. J. Cell Biol.* **79**, 17-26.
- Karabinos, A., Schmidt, H., Harborth, J., Schnabel, R. and Weber, K. (2001)** : Essential roles for four cytoplasmic intermediate filament proteins in *Caenorhabditis elegans* development. *Proc. Natl. Acad. Sci. USA* **98**, 7863-7868.
- Karabinos, A., Schunemann, J., Parry, D.A., Weber, K. (2002)** : Tissue-specific co-expression and in vitro heteropolymer formation of the two small branchiostoma intermediate filament proteins A3 and B2. *J. Mol. Biol.* **316**, 127-137.
- Laemmli, U.K. (1970)** : Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lees, J.F., Shneidman, P.S., Skuntz, S.F., Carden, M.J., Lazzarini, R.A. (1988)** :The structure and organization of the human heavy neurofilament subunit (NF-H) and the gene encoding it. *EMBO J.* **7**, 1947-1955.
- Lenz-Böhme, B., Wismar, J., Fuchs, S., Reifgerste, R., Buchner, E., Betz, H., Schmitt, B. (1997)** : Insertional mutation of the Drosophila nuclear lamin Dmo gene results in defective nuclear envelopes, clustering of nuclear pore complexes, and accumulation of annulate lamellae. *J. Cell Biol.* **137**, 1001-1016.
- Lewis, S.A., Cowan, N.J. (1986)** : Anomalous placement of introns in a member of the intermediate filament multigene family: an evolutionary conundrum. *Mol. Cell Biol.* **6**, 1529-1534.
- Lin, F., Worman, H.J., (1993)** : Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J. Biol. Chem.* **268**, 16321-16326.
- Loewinger, L., McKeon, F., 1988**: Mutations in the nuclear lamin proteins resulting in their aberrant assembly in the cytoplasm. *EMBO J.* **7**, pp. 2301-2309.
- Maul, G.G., Schatten, G., Jimenez, S.A., Carrera, A.E. (1987)** : Detection of nuclear lamin B epitopes in oocyte nuclei from mice, sea urchins, and clams using a human autoimmune serum. *Dev. Biol.* **121**, 368-375.
- Mullis, K., Faloona, F. (1987)** : Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaktion. *Meth. Enzymol.* **55**, 335-350.
- Myers, M.W., Lazzarini, R.A., Lee, V.M.-Y., Schlaepfer, W.W., Nelson, D.L. (1987)** : The human mid-size neurofilament subunit: a repeated protein sequence and the relationship of its gene to the intermediate filament gene family. *EMBO J.* **6**, 1617-1626.



- Ohara, O., Gahara, Y., Miyaka, T., Teraoka, H., Kitumura, T. (1993)** : Neurofilament deficiency in quail caused by nonsense mutation in neurofilament-L gene. *J. Cell Biol.* **121**, 387-395.
- Parry, D.A.D., Steinert, P.M. (1995)** : Intermediate filament structure. *Springer Verlag*, New York.
- Perreau, J., Lilienbaum, A., Vasseur, M., Paulin, D. (1988)** : Nucleotide sequence of the human vimentin gene and regulation of its transcription in tissues and cultured cells. *Gene* **62**, 7-16.
- Pruss, R.M., Mirsky, R., Raff, M.C., Thorpe, R., Dowding, A.J., Anderton, B.H. (1981)** : All classes of intermediate filaments share a common antigenic determinant defined by a monoclonal antibody. *Cell* **27**, 419-428.
- Quinlan, R., Hutchinson, C., Lane, B. (1994)** : Intermediate Filament Proteins. *Protein Profile* **1**, 779-911.
- Riemer, D., Dodemont, H., Weber, K. (1991)** : Cloning of the non-neuronal intermediate filament protein of the gastropod *Aplysia californica*; identification of an amino acid residue essential for the IFA epitope. *Eur. J. Cell Biol.* **56**, 351-357.
- Riemer, D., Dodemont, H., Weber, K. (1992)** : Analysis of the cDNA and gene encoding a cytoplasmic intermediate filament (IF) protein from the cephalochordate *Branchiostoma lanceolatum*; implications for the evolution of the IF protein family. *Eur. J. Cell Biol.* **58**, 128-135.
- Riemer, D., Dodemont, H., Weber, K. (1993)** : A nuclear lamin of the nematode *Caenorhabditis elegans* with unusual structural features; cDNA cloning and gene organization. *Eur. J. Cell Biol.* **62** (2), 214-223.
- Riemer, D., Weber, K. (1994)** : The organization of the gene for *Drosophila* lamin C: limited homology with vertebrate lamin genes and lack of homology versus the *Drosophila* lamin Dmo gene. *Eur. J. Cell Biol.* **63**, 299-306.
- Riemer, D., Wang, J., Zimek, A., Swalla, B.J., Weber, K. (2000)** : Tunicates have unusual nuclear lamins with a large deletion in the carboxyterminal tail domain. *Gene* **255**, 317-325.
- Röber, R.A., Weber, K., Osborn, M. (1989)** : Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. *Development* **105**, 365-378.
- Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., Cherry, J.M., et al. (2000)** : Comparative genomics of the eukaryotes. *Science* **287**, 2204-2215.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989)** : Molecular cloning. A laboratory manual. 2nd Edition. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press.
- Stick, R., (1992)** : The gene structure of *Xenopus* nuclear lamin A: a model for the evolution of A-type from B-type lamins by exon shuffling. *Chromosoma* **101**, 566-574.
- Strelkov, S.V., Herrmann, H., Geisler, N., Wedig, T., Zimbelmann, R., Aebi, U., Burkhard, P. (2002)** : Conserved segments 1A and 2B of the intermediate filament dimer: Their atomic structures and role in filament assembly. *EMBO J.* **21**, 1255-1266.
- Stuurman, N., Heins, S., Aebi, U. (1998)** : Nuclear lamins: Their structure, assembly, and interactions. *J. Struct. Biol.* **122**, 42-66.

- Sullivan, T., Escalante-Alcalde, D., Bhatt, H., Anver, M., Bhat, N., Nagashima, K., Stewart, C.L., Burke, B. (1999)** : Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J. Cell Biol.* **147**, 913-920.
- Szaro, B.G., Pant, H., Way, H., Battey, J. (1991)** : Squid low molecular weight neurofilament proteins are a novel class of neurofilament protein. *J. Biol. Chem.* **266**, 15035-15041.
- Turbeville, J.M., Schulz, J.R., Raff, R.A. (1994)** : Deuterostome phylogeny and the sister group of the chordates: Evidence from molecules and morphology. *J. Mol. Biol. Evol.* **11**, 648-655.
- Wang, J., Karabinos, A., Schünemann, J., Riemer, D., Weber, K. (2000)** :The epidermal intermediate filament proteins of tunicates are distant keratins; a polymerisation-competent hetero coiled coil of the Styela D protein and Xenopus keratin 8. *Eur. J. Cell Biol.* **79**, 478-487.
- Wang, J., Karabinos, A., Zimek, A., Meyer, M., Riemer, D., Hudson, C., Lemaire, P., Weber, K. (2002)** : Cytoplasmic intermediate filament protein expression in tunicate development; a specific marker for the test cells. *Eur. J. Cell Biol.* **81**, 302-311.
- Weber, K., Plessmann, U., Ulrich, W. (1989)** : Cytoplasmic intermediate filament proteins of invertebrates are closer to nuclear lamins than are vertebrate intermediate filament proteins; sequence characterization of two muscle proteins of a nematode. *EMBO Journal* **8**, 3221-3228.
- Xu, Y., Bolton, B., Zipser, B., Jellies, J., Johansen, K.M., Johansen, J. (1999)** : Gliarin and macrolin, two novel intermediate filament proteins specifically expressed in sets and subsets of glial cells in leech central nervous system. *J. Neurobiol.* **40**, 244-253.
- Zhu, X., Mahairas, G., Illies, M., Cameron, R.A., Davidson, E.H., Etensohn, C.A. (2001)** : A large scale analysis of mRNAs expressed by primary mesenchyme cells of the sea urchin embryo. *Development* **128**, 2615-2627.
- Zimek, A., Weber, K. (2002)** : The gene for a cytoplasmic intermediate filament (IF) protein of the hemichordate *Saccoglossus kowalevskii*; definition of the unique features of chordate IF proteins. *Gene* **288**, 187-193.



Figur 16: *Ciona* IF A2, genomische Sequenz des zytoplasmatischen IF-Proteins aus *Ciona intestinalis*. Die Konsensussequenzen zu Beginn und am Ende der Introne sind fett gedruckt. Die Daten der Intronpositionen dienen der Bestimmung des IF-Typs in Figur 6.

```

339 AAGAGTGTCCAAACCTTGAATATGGAGATTGACTCTTTGAGAGGAACAgtaagtttaacataggtgtttaactgtgggctacaaactgtttaagctctgtagaataaatcaattaatgaa
K S V Q T L N M E I D S L R G T

tgtaacttattttatctctgcgtggcggggcagcagcagctogttataatcggggtgtctgtttcatacaccctogtccgcttacgagttaccacgtatgtaactttgtgagtgattatfff
tctgtattgttgacaatttaaacaccataagtaaacactgggtgggaaactgctgcccatttagtctccaccaaggacacatacattcacaatcacaggcagcagcagctttgat
cccgtaacctttgagctaacatattataaccacacagAACGACTCCCTTCAACGAGCAATGGGTGAACCTTGAAGAGAGATACACAGAGATACCTCCGATTACCAAGATACCATCATGA
355 N D S L Q R A M G E L E E R Y N R D T S D Y Q D T I M N

ACTTGCAggtgtataagcatgggagctccatgatacaatgtttaacaataagtgtaaggttaactttataagatatgatcataaggtggtaaatagtttaactgtatttacacctcatg
383 L Q

ccctttaaacttgatgtcttctttatacaccagacaaaatgggtgtattcatacaccctcatgctcactgttgggttataacatgggtgtctcatalacaccagacacaccatgggtgata
catacaoCGTGAAGTGCAGACCTGAAGACACAGATGGCTCAACACTCCGCTCAATACCAAGAATCATGGACGTAAGATGGCTTTGGACATCGAGATCGCTACTTACAGGAAGCTGTT
385 A E C D D L K T Q M A Q H L R Q Y Q E L M D V K M A L D I E I A T Y R K L L

GGAGGGAGAGGAATCAAGgtaatggggacgaataaataaaataataaacatctatgctcatgctacataagtggttcaatgggtcatcgctgctaccattgtggacatgtagatccttgg
423 E G E E S R

gcaagatacttaacggcaattgctccaacaaagtggtcactaatgggtgtccaattgctgccattcagccacaagaatgcacatacgtggtaactcgttaagcaggcagcaggtgtatga
aacaacaccccggtttatataaccgactgttgggtgccccgcaagcgttaggataaaataatgcattcatttattcatttccattcctcaccagcagagtcgaagcgtggaaggaac
aggtaatgacgctcaatcatttcaatttcaattatagtaagtagtacaagttcaagtcactcggccttatgactacaacacacggccttaactcagtgatctcaacaaaatga
acagataaaatcaatcacatcggcctgtattgacaaaatattttgtccaacgctgcaacttttttaaaaggtggcataatgtctctgtataaaaaataaaaaactacacgtgagtagaa
aaactaatacacaatgcaatcccgactaacactatgcacaatcacagtttaagccgggcatacaaatgtaaatgtaacatgcatatcatalatgtaggttacatctgggttaactcgg
taataaatacagGCTTGGCTGAACCCCTCGAACCTCGGGCGCGGTAGCTCCCGTTTCGTCACAAACACATCTAGATCAggtaataattatgtgcgtacgggtatttatctgtctgacta
429 L A E R L R T S G G G S S R F V T T T S R S A

tatatggtcagcgcagcggcgtgtattagacttgacagtagggctggtagcaaaaacatacaacaactgtggggtaaagatgaatttcgttagctcatagctcataataataaaatga
gaaatgagaataaagagcatgtgtacattctgacgggtacgggacaatttatagtggtggggtaaattctaacaacataaaattagCAGGGGGATATTCGTCATCAGCTGGCGGATATTC
452 G G Y S S S A G G Y S

ATCCGGCCGAGGAGGCGCTAGTGACGTCATGGGAAGTGATGACGTAGAAACCGTCAACAACAAGAAGGTTGTTGTCAAGACAATTGAAACAAGGACGGAAAGatcgtcccatattgatag
463 S G R G G A S D V M G S D D V E T V T T K K V V V K T I E T K D G K

taaaccaagattacaaaatataatctgtattataaaactccaacagagctttgttaattgttcaaacacgattaggaataacggaatattatgtgcttacggtatttatcttacact
actataatggtactgcagcgaacccgactgtatagacttgacagtagggctggtagcaaaaacatacaacaactgtggggtaaagatgaatttcgttagctcatagctcataataataaa
atgagaaatgagaataaagacatgtgtacattctgacgggtacgggacaatttatagtggtggggtaaattctaacaacataaaatttagcaggggggatattcgtcatcagctggcggat
attcatccggccgaggaggcctagtgacgtcatgggaagtgatgacgtagaaaacgctcaacaacaagaaggttgtgtcaagaacaatgaaacaaagcggggaaggtggaatgatgtt
tctatgttaactaagaccagagcttacgaaacaaaatgttaattcagataaaaaatccattgaaatattttgaaactaagtaacatacacaaagtaatttcaattttctctcagactacagc
agcgaagatgttatataaaacaagataagataaattggaaaacccggtgcttaaaacactaccatttactatgtggttatatgctgggataacacatttaatttgaactatagca
ataatacgaactctgtgatatgaacactaaccttatgatctctatcacagcaagGTGTGGACCAACCGAGGACGTTCTGTGAGCAATCATCTGTGATATGATGACCAACCGCTCAC
497 V V S Q T E D V R E Q S S G D H D D N A H

```

Figur 17: *Ciona* IF F1, genomische Sequenz des zytoplasmatischen IF-Proteins aus *Ciona intestinalis*. Die Konsensussequenzen zu Beginn und am Ende der Introne sind fett gedruckt. Die Daten der Intronpositionen dienen der Bestimmung des IF-Typs in Figur 6.

```

259 CGAAAGGTTATTCACCAACTTGAATGGAAATGGAGGCTCTTCGTGGAAAGgtaatgatcattttaccaggatatttggaacaaagataatttgccttataatgca
R K V I H Q L E M E L E A L R G K

tgctgcaaaaactaaactgtctttttgattagAATGACGGTTTGGAGCGTAATCAAGATGGTTTGGAGGATCGTCATGCCAGGAAATTCGCCAGTATCAAGAACA
276 N D G L E R N Q D G L E D R H A R E I A E Y Q E Q

AGTTGAGAAAATCGCGTCGGACTTGGATGGAGCGCAACAGgtaacttttagttgcttggattttgtgagcgtatataaagaagcttctactgtcgttcgccc
301 V E K M R S D L D G A K Q

actggcggtaaggattttctctcgtgacgaaatcgcgtgctgttactagtggttagcgcacgtatataatgcccattgatcgtatagtgctcgtttaatcaaatat
atgtgcaacttgtaagtagtagtactgtggggttagatcggatactgttagcactaaatcccatattttctgtcgtgttttaacatttaaccaagatccttaca
attttttaagcaaatatttttagtgctgagcaggtatacaattccgtaaatattctttgtttacttcccaatttaaacaggaaggagaatta
aatgtcccactaaactttagGACATCGGAAAGTGCTTAAAGGACTACCNAGGACTTGGACCACTTGAAGTTAAGTTTGGAGCAAGAAATCGCAATCTACA
314 D I G K C L K G L P R T W T T L K L S L E Q E I A I Y N

ACAAGATCTTGATGGGAGAAGCATAG...ctatcgttaactcgggtgtgtatataactctgttttaataagagatcagtttattccaatttggtaataaga
342 K I L M G E E H R

attgataggtacacatctgtgccataaatgtgagtagtagtaccctgsgggttaagatgggacacatttagcacataatcacaatattcctgatcgtgttttaaa
caattatcgacgctgtatggggctgtgagaatcaggttttaataccctgaaatgttcttggttattacaaaatggagagagaaaatagattgaaaatgtgtc
ccatcttccccccccactatataaaattttatcttaataatgtttagGATCGCGGTGATGATCAAGACACTGGTCTGCAACACGACGCAACCGT
351 I A V I D T K T L V L A N D S K R

TCATCAAGGGCTTCAATCCAAATCATCTTACGATCATCATCCAGCTCTCATCTTCCGAGATCAACAAGCGATATTGTTGAAGAGATGTTTGGAAAAGATACT
368 S S R A S S K S S S R S S S S A R S T S D I V E E M L E K D T

CCAGCCAAGGACCATCTACGGTATGaaacccaatcaaaccttaagctc
403 P A K E P S T V ***

```



Die vorliegende Arbeit wurde durch die Unterstützung ( We-338/7 ) der  
Deutschen Forschungsgemeinschaft  
ermöglicht.