Aus dem Institut für Immunologie des
Universitätsklinikums Hamburg-Eppendorf
Institutsdirektor: Prof. Dr. med. Bernhard Fleischer

Molekulare Klonierung und Charakterisierung neuer ADP-Ribosyltransferasen bei *Homo sapiens* und *Salmonella typhimurium*

Dissertation

zur Erlangung des Grades eines Doktors der Medizin

dem Fachbereich Medizin der Universität Hamburg vorgelegt von
Helge Otto
aus Kassel

Hamburg 2005

Inhaltsverzeichnis

Seite

Abbildungsverzeichnis **5**

Abkürzungsverzeichnis **7**

1 Einführung **9**

1.1 ADP-Ribosylierung **9**

1.2 NAD^+ als Substrat **13**

1.3 Prokaryontische Mono-ADP-Ribosyltransferasen **14**

1.3.1 Der Prototyp Diphterietoxin 15

1.3.2 Identifizierung und Charakterisierung weiterer prokaryontischer ARTs 15

1.4 Topologie der katalytischen Domäne bekannter ARTs **20**

1.5 Die Spezies *Salmonella enterica* **23**

1.5.1 Der Spv-Locus 24

1.6 Eukaryontische ADP-Ribosyltransferasen **25**

1.6.1 Eukaryontische poly-ADP-Ribosyltransferasen 25

1.6.1.1 Die humane PARP 26

1.6.1.2 Die biologische Rolle der PARP 26

1.6.1.3 Die Entdeckung von PARP-Verwandten 27

1.6.2 Eukaryontische mono-ADP-Ribosyltransferasen 28

1.7 Gen-Datenbanken und „*in silico*“-Klonierung **29**

1.7.1 GenBank 30

1.7.2 Die EST-Projekte 31

1.7.3 Die „UniGene“-Datenbank 32

1.7.4 Programme für Datenbankrecherchen 33

1.7.5 Genomprojekte 34

1.8 Das pET-System zur Expression rekombinanter Proteine in *E.Coli* **36**

2 Ziele **38**

3 Materialien und Methoden **39**

3.1 Software und Datenbanken **39**

3.1.1 Identifizierung neuer mono- und poly-ARTs 39

3.1.2 Genstrukturen und chromosomale Lokalisierung 39

3.1.3 Domänenstrukturen 40

3.1.4 Analyse der Sequenzen und Erstellung der Alignments 40

3.1.5 3D-Strukturen 41

3.1.6 Erstellung der Dissertationsarbeit 41

3.2 Kreation von spezifischen Oligonukleotiden (Primern) **41**

3.3 PCR („polymerase chain reaction“, Polymerase-Ketten-Reaktion) **42**

3.3.1 MTC („Multiple Tissue cDNA“)-PCR 42

3.3.2 PCR-Amplifizierung des SpvB 43

3.3.3 PCR-Mutagenese 43

3.3.4 PCR zum Vergleich von cDNA- und genomischen DNA-Fragmenten von humanem pART3 44

3.3.5 Chromosomen-„Mapping“ des humanem *pART3* 45

3.4 Northern Blot ... **45**
3.5 Klonierungsstrategien ... **46**
3.5.1 Klonierungsstrategie für hpART1, hpART3 und hpART4 ... 47
3.5.1.1 Klonierung in den pCR2.1-Vektor und Transformation in INV-Zellen ... 47
3.5.1.2 Klonierung der Expressionsprodukte ... 48
3.5.2 Klonierungsstrategie für SpvB ... 50
3.6 Plasmidaufreinigung ... **50**
3.6.1 Miniprep ... 50
3.6.2 Maxiprep ... 51
3.6.2.1 Maxiprep genomischer PAC-Klone ... 51
3.6.2.2 Maxiprep von Vektoren und cDNA-Konstrukten ... 51
3.7 Expression von rekombinantem Protein in *E.Coli* ... **51**
3.7.1 Transformation der pET26b-/pET32a-Expressionskonstrukte in Bl21(DE3)-/ Bl21(DE3)pLysS-Zellen ... 51
3.7.2 Expression der rekombinanten Proteine ... 52
3.7.2.1 Expression kleiner Proteinmengen und Herstellung von *E.Coli*-Gesamtzell-Lysaten ... 52
3.7.2.2 Expression größerer Proteinmengen und Herstellung von *E.Coli*-Gesamtzell-Lysaten (Midipreps) ... 53
3.8 Proteinanalytik ... **54**
3.8.1 Western-Blot-Analysen ... 54
3.8.2 Silber- und Coomassiefärbung ... 55
3.8.3 Detektion und Chemilumineszenz-Reaktion ... 55
3.8.4 Aufreinigung der Trx-Fusionproteine mit Protein-S-Agarose ... 56
3.9 Enzymassays ... **56**
3.9.1 Enzymassays mit *E.Coli*-Gesamtzell-Lysaten ... 56
3.9.1.1 Nachweis der Enzymaktivität von rekombinanter pART1 und pART4 ... 56
3.9.1.2 Dünnschichtchromatographie zur Analyse der NADase-Aktivität ... 57
3.9.1.3 Aktin-Enzymassay mit den SpvB-Konstrukten ... 57
3.9.2 Enzymassays mit den aufgereinigten Trx-Fusionsproteinen von hpART4 und SpvB ... 58
3.9.2.1 Untersuchung der Automodifikation von Trx-pART4 und Trx-SpvB ... 58
3.9.2.2 Untersuchung der NH_2OH-Sensitivität der durch rekombinantes Trx-SpvB katalysierten ADPR-Bindungen ... 58
3.9.2.3 Nachweis der ADP-Ribosylierung einer 40kD-Bande in Jurkat-Zell-Lysaten ... 59
3.9.3 Autoradiographien ... 59
3.10 Automatische DNA-Sequenzierungen ... **59**
4 Ergebnisse ... **61**
4.1 *In silico* Analyse der poly-ART-Genfamilie ... **61**
4.1.1 Identifikation von RT6- und PARP-verwandten Genen in dbEST ... 61
4.1.2 Die Familie der humanen pARTs und ihre murinen Orthologen ... 65
4.1.3 Chromosomale Lokalisierung der humanen und murinen *pART*s ... 70
4.1.4 *In silico* Analyse und Vergleich der Genstrukturen der humanen *pART*-Paralogen ... 75
4.1.5 Identifikation von pART-Familienmitgliedern in verschiedenen Spezies ... 85
4.1.6 *In silico* Analyse und Vergleich der Genstrukturen von *pART*-Verwandten anderer Spezies ... 93

4.1.7 Analyse der Domänenstruktur verschiedener pART-Familienmitglieder ... 99
4.2 Molekulare Klonierung und rekombinante Expression der humanen ADP-Ribosyltransferasen pART1, pART3 und pART4 ... 109
4.2.1 Herstellung genspezifischer Sonden und Analyse der Expression von pART1, pART3 und pART4 in humanen Geweben ... 109
4.2.2 Molekulare Klonierung und Sequenzierung der cDNA-Sequenz von hpART3, hpART4 und hpART1 ... 112
4.2.3 Herstellung von Expressionskonstrukten und Expression rekombinanter Proteine ... 113
4.2.3.1 Herstellung von Expressionskonstrukten ... 114
4.2.3.2 Expression von rekombinantem pART3, pART4 und pART1 in *E.Coli* ... 117
4.2.4 Enzymassays ... 120
4.2.4.1 Nachweis der ADP-Ribosyltransferase-Aktivität von hpART4 und hpART1 ... 121
4.2.4.2 Nachweis der NAD-Glykohydrolase-Aktivität von hpART4 ... 123
4.2.4.3 Detektion des rekombinanten hpART4 über den „S-tag“ ... 123
4.2.4.4 Nachweis der Enzymaktivität und Auto-ADP-Ribosylierung von aufgereinigtem Trx-pART4 ... 124
4.2.4.5 Ermittlung der Zielregion der ADP-Ribosylierung im rekombinanten Trx-pART4-Fusionsprotein ... 124
4.2.5 Aufschlüsselung der Genstruktur von hpART3 und hpART4 ... 127
4.2.6 Chromosomale Kartierung von *hpART3* ... 130
4.3 Molekulare Klonierung und rekombinante Expression der Salmonellen-ADP-Ribosyltransferase SpvB ... 132
4.3.1 Verknüpfung von Vertebraten-ARTs mit bakteriellen mono-ARTs durch Datenbankanalysen mit PSI-BLAST ... 132
4.3.2 Identifikation von SpvB durch Datenbankanalysen mit PSI-BLAST ... 137
4.3.3 Klonierung, Expression und Aufreinigung von rekombinanten SpvB-Konstrukten in *E.Coli* ... 140
4.3.3.1 Klonierung von drei verschiedenen SpvB-Konstrukten ... 141
4.3.3.2 Rekombinante Expression und Aufreinigung des SpvB ... 141
4.3.4 Enzymassays mit verschiedenen rekombinanten SpvB-Konstrukten ... 143
4.3.4.1 Nachweis der ADP-Ribosyltransferase-Aktivität des rekombinanten SpvB ... 143
4.3.4.2 Nachweis der Arginin-Spezifität der ADP-Ribosyltransferase-Aktivität des rekombinanten SpvB ... 146
4.3.4.3 Ermittlung der Zielregion für die ADP-Ribosylierung im rekombinanten Trx-SpvB-Fusionsprotein ... 147
4.3.4.4 *In vitro*-Nachweis der SpvB-katalysierten ADP-Ribosylierung heterologer Proteine in humanen Jurkat-Zell-Lysaten ... 149
4.3.4.5 Nachweis der ADP-Ribosylierung von Aktin durch rekombinantes SpvB *in vitro* ... 150
4.3.5 Klonierung eines inaktiven Trx-SpvB-Fusionskonstruktes durch Mutagenese ... 152
5 Diskussion ... 154
5.1 Die pART-Familie ... 154
5.1.1 Die humanen pARTs ... 155
5.1.2 pART-Familienmitglieder in verschiedenen Spezies ... 160
5.1.3 pART-Genstrukturen ... 161

5.1.4 Chromosomale Lokalisierung der humanen und murinen *pART*s ... 163
5.1.5 Domänenstrukturen der pARTs ... 164
5.1.6 Weitere Charakterisierung und Funktionen der humanen pARTs ... 166
5.1.6.1 pART1 / PARP 1 ... 167
5.1.6.2 pART2 / PARP 2 ... 168
5.1.6.3 pART3 / PARP 3 ... 169
5.1.6.4 pART4 / vPARP ... 171
5.1.6.5 Die Tankyrasen (pART5 / Tankyrase 1 und pART6 / Tankyrase 2) ... 172
5.1.6.6 pART9 / BAL ... 174
5.1.6.7 pART10 / PARP 10 ... 174
5.1.6.8 pART13 / ZAP ... 175
5.1.6.9 pART14 / Ti PARP ... 176
5.2 Eine neue bakterielle mono-ART: Das Salmonellen-Toxin SpvB ... 178
5.2.1 Ergebnisse der PSI-BLAST-Analysen in der mono-ART-Familie ... 178
5.2.2 Identifizierung des SpvB ... 180
5.2.3 Die Subdomänen von SpvB ... 181
5.2.4 SpvB besitzt eine ARTase-Funktion ... 182
5.2.5 SpvB modifiziert ein wichtiges Wirtszellenprotein ... 182
5.2.6 Das Zielprotein der SpvB-katalysierten ADP-Ribosylierung ist Aktin ... 183
5.2.7 SpvB fördert die Depolymerisierung von Aktin ... 184
5.2.8 SpvB induziert Zytotoxizität und Apoptose in den Wirtszellen ... 185
5.2.9 Der Export-Mechanismus für das SpvB ist bisher ungeklärt ... 186
6 Zusammenfassung ... 190
7 Literaturverzeichnis ... 192
8 Anhang ... 207
9 Danksagung ... 220
10 Lebenslauf ... 221
11 Publikationen ... 222
12 Eidesstattliche Versicherung ... 223

Abbildungsverzeichnis

Seite

Abb. 1: Schematische Darstellung der durch ARTs katalysierten Reaktionen ... 10
Abb. 2: Regulation der Dinitrogenase-Reduktase von *Rhodospirillium rubrum* ... 12
Abb. 3: Schematisches Diagramm von ausgewählten konkurrierenden NAD^+-katabolisierenden Prozessen ... 14
Abb. 4: Schematische Darstellung wichtiger prokaryontischer ARTs und ihrer Wirkungsmechanismen ... 19
Abb. 5: Darstellung der dreidimensionalen Struktur von VIP2, ART2 (A), DT und PARP (B) ... 22
Abb. 6: Graphische Darstellung des Wachstums der GenBank (1982-2002) ... 31
Abb. 7: Expression von rekombinanten Proteinen in *E.Coli* mittels eines regulierten Promoters ... 37
Abb. 8: Schematische Darstellung der Klonierungsstrategien ... 46
Abb. 9: Humane EST-„Cluster“ mit Sequenzidentitäten zu bekannten ADP-Ribosyl-Transferasen in der EST-Datenbank (dbEST) ... 63
Abb. 10: Aminosäuren-Sequenz-Alignment der humanen pART-Familie und ihrer murinen Orthologen ... 67
Abb. 11: Aminosäuresequenzidentität der humanen und murinen pARTs im Bereich der konservierten Sekundärstrukturmotive der katalytische Domäne ... 69
Abb. 12: Schematische Darstellung der chromosomalen Lokalisierung der humanen (A) und murinen (B) *pART*-Gene ... 71
Abb. 13: Vergleich der pART kodierenden Regionen auf dem humanen Chromosom 3q und dem Maus-Chromosom 16B3 ... 73
Abb. 14: Konservierte Aminosäuren, Sekundärstrukturmotive und Lage der Introns in der potentiellen katalytischen Domäne der humanen pART-Paralogen ... 78
Abb.15: Genomische Struktur und konservierte Intronpositionen der humanen *pART*-Paralogen im Bereich der katalytischen Region ... 82
Abb.16: Exonanzahl, Proteinlänge, chromosomale Lage und Zugangsnummern der humanen und murinen pARTs ... 85
Abb.17: PSI-BLAST-Datenbankrecherchen zur Identifizierung neuer pART-Familienmitglieder in verschiedenen Spezies ... 87
Abb. 18: Konservierte Aminosäuren, Sekundärstrukturmotive und Lage der Introns in der potentiellen katalytischen Domäne von pART-Verwandten verschiedener Spezies ... 94
Abb. 19: Schematische Darstellung der Genstrukturen ausgewählter pART-Familienmitglieder aus verschiedenen Spezies (A) sowie Darstellung der konservierten Intronpositionen (B) ... 96
Abb. 20: Schematische Darstellung der Domänenstrukturen ausgewählter Mitglieder der pART-Familie (A) und Auflistung der assoziierten Proteindomänen (B) ... 100
Abb. 21: Expressionsprofil von *hpART3* und *hpART4* in humanen Geweben ... 110
Abb. 22: RT-PCR-Analysen im „MTC“-Panel ... 111
Abb. 23: Klonierung der cDNA-Fragmente in pET-Vektoren ... 114
Abb. 24 Schematische Darstellung der verschiedenen Expressionskonstrukte am Beispiel von hpART3 .. 116
Abb. 25: Expression von hpART3-Konstrukten in *E.Coli* ... 118
Abb. 26: Expression von hpART4-Konstrukten in *E.Coli* ... 119
Abb. 27: Expression von hpART1-Konstrukten in *E.Coli* ... 120
Abb. 28: Nachweis der enzymatischen Aktivität von rekombinantem hpART4 ... 122

Abb. 29: Detektion von rekombinantem Trx-pART4 über den „S-tag“ (A) und Lokalisation der ADP-ribosylierten Region im Trx-pART4-Fusionsprotein (B) ... 126
Abb. 30: Vergleich von cDNA- und genomischen DNA-Fragmenten von hpART3 ... 128
Abb. 31: Schematische Darstellung der Genstrukturen von *hpART1*, *hpART3* und *hpART4* ... 130
Abb. 32: Chromosomale Kartierung von *hpART3* ... 131
Abb. 33: PSI-BLAST-Datenbankrecherchen zur Identifizierung neuer mono-ARTs ... 133
Abb. 34: Schematische Darstellung des katalytischen Zentrums der mono-ARTs (A) und multiples Aminosäuresequenz-Alignment der katalytischen Region verschiedener mono-ARTs (B) ... 136
Abb. 35: Aminosäuresequenz-Alignment von SpvB mit seinen nächsten Verwandten aus der Datenbank ... 138
Abb. 36: Schematische Darstellung der Domänenstrukturen von SpvB und verwandten Proteinen ... 140
Abb. 37: Expression von SpvB-Konstrukten in *E.Coli* ... 142
Abb. 38: Nachweis der ADP-Ribosyl-Transferase-Aktivität in Trx-SpvB enthaltenden *E.Coli*-Lysaten ... 144
Abb. 39: Aufreinigung von Trx-SpvB (A) und NH_2OH-Sensitivität der durch Trx-SpvB katalysierten ADP-Ribosyl-Bindungen (B) ... 147
Abb. 40: ADP-Ribosylierung der assoziierten Trx-Domäne durch rekombinantes Trx-SpvB ... 148
Abb. 41: SpvB-katalysierte ADP-Ribosylierung heterologer Proteine in humanen Jurkat-Zell-Lysaten ... 150
Abb. 42: Aktin als Zielprotein der SpvB-katalysierten ADP-Ribosylierung in *E.Coli* ... 152
Abb. 43: Mutagenese der SpvB-Konstrukte ... 153
Abb. 44: Verwendete Primer ... 207
Abb. 45: Konsensussequenzen der humanen pARTs 1, 3 und 4 und des Salmonellentoxins SpvB ... 208
Abb. 46: Vektorkarten (pCR2.1, pET26b und pET32a) ... 211
Abb. 47: Aminosäuresequenz-Alignments der humanen pART-Subgruppen 1-5 ... 214

Abkürzungsverzeichnis

Abb.	Abbildung
ADPR	ADP-Ribose
ARH	ADP-Ribosylhydrolase
ART	ADP-Ribosyltransferase
ARTase	ADP-Ribosyltransferase
AS	Aminosäuren
ATP	Adenosintriphosphat
BAL	„B aggressive lymphoma“, aggressives B-Zell-Lymphom
BLAST	Basic Local Alignment Search Tool
bp	Basenpaare
cAMP	cyclisches Adenosinmonophosphat
CD	catalytische Domäne
CDART	Conserved Domain Architecture Retrieval Tool
CDD	Conserved Domain Database
cDNA	„complementary DNA“
ci	Chilo iridescent Virus
CT	Choleratoxin
d	Tage („days“)
DDBJ	DNA DataBank of Japan
DNA	„desoxyribonucleic acid“ / Desoxyribonukleinsäure
DRAG	Dinitrogenase-Reduktase-ADP-Ribosylglykohydrolase
DRAT	Dinitrogenase-Reduktase-ADP-Ribosyltransferase
DT	Diphterietoxin
EDTA	„ethylene diamine tetraacetic acid“
EF-2	Elongationsfaktor 2
EK	Enterokinase
EMBL	European Molecular Biology Laboratory
EST	„Expressed Sequence Tag“
ETA	*Pseudomonas aeruginosa* Exotoxin A
EthBr	Ethidiumbromid
GAP	Rho-GTPase-aktivierendes Protein
GPI	Glykosylphosphatidylinositol
h	Stunde („hour“)
hpART	humane poly-ADP-Ribosyltransferase
HTG	„high throughput genomic sequences“
IPTG	Isopropylthiogalactosid
kb =	kilobasen
kbp	kilo Basenpaare
kD	kilo Dalton
k.o.	„knockout“
LT	*E.Coli* „heat labile enterotoxin“
µCi	mikro Curie

MCS	„Multiple cloning site“
MG	Molekulargewicht
min	Minute
mpART	murine poly-ADP-Ribosyltransferase
mRNA	„messenger ribonucleic acid“
MCS	„multiple cloning site“
MGC	Mammalian Gene Collection
MTC	„Multiple tissue cDNA“
NAD^+	Nikotinamid-Adenin-Dinukleotid
NADase	Nikotinamid-Adenin-Dinukleotid-Glykohydrolase
NCBI	National Center of Biotechnology Information
NC	Nitrocellulose
NES	„nuclear export signal“
NLS	„nuclear localization signal“
nr	„non redundant“
PARG	Poly-ADP-Ribose-Glykohydrolase
PARP	Poly-ADP-Ribosyl-Polymerase
PCR	„polymerase chain reaction“, Polymerase-Ketten-Reaktion
[^{32}P]-dCTP	radioaktiv markiertes desoxy-Cytidintriphosphat
pART	poly-ADP-Ribosyltransferase
PPl	Periplasma
PT	Pertussistoxin
PVDF	Polyvinyldendifluorid
OD	optische Dichte
PAC	„P1 artificial chromosome“
rpART	Ratten-poly-ADP-Ribosyltransferase
RT	Raumtemperatur
RZPD	Deutsches Ressourcenzentrum für Genomforschung GmbH
SDS	Sodiumdodecylsulfat
SDS-PAGE	Sodiumdodecylsulfat-Polyacrylamidgel-Elektrophorese
sec	Sekunde
SPI	„salmonella pathogenicity island“
TAE	Tris-Acetat-EDTA-Puffer
TBS	„TRIS Buffered Saline“
TE	Tris-EDTA-Puffer
TiPARP	„TCCD-inducible PARP“
Trx	Thioredoxin
TTSS	Typ-III-Sekretionssystem
ü.N.	über Nacht
vgl.	vergleiche
vPARP	„vault PARP“
X-Gal	5-Bromo-4-Chloro-3-Indolyl-β-D-Galactopyranosid
ZAP	„zinc finger antiretroviral protein“

1 Einleitung

1.1 ADP-Ribosylierung

ADP-Ribosylierung ist neben der Phosphorylierung eine der enzymatisch-katalysierten Reaktionen, die es ermöglicht, spezifische Aminosäurenreste von Zielproteinen posttranslational zu modifizieren und somit die Funktion dieser Zielproteine zu verändern. Die reversible, stereospezifische Reaktion, in der Nikotinamid-Adenin-Dinukleotid (NAD^+) als Substrat dient, wird durch ADP-Ribosyltransferasen (ARTs) katalysiert. Die Bindung des Substrates NAD^+ im aktiven Zentrum der ARTs ermöglicht einen nukleophilen Angriff eines Akzeptormoleküls auf die glykosidische Bindung zwischen Nikotinamid und Ribose. Unter Freisetzung von Nikotinamid wird der ADP-Ribose-Anteil auf spezifische Aminosäurenreste übertragen (**Abbildung 1A**). Häufig wird das Zielprotein durch die ADP-Ribosylierung inaktiviert (Haag and Koch-Nolte 1997, Jacobson and Jacobson 1989, Moss and Vaughan 1990).

Während einige ARTs verschiedene Zielproteine akzeptieren, weisen andere Familienmitglieder eine ausgeprägte Spezifität für ein bestimmtes Zielprotein auf: So können einige Arginin-spezifische Transferasen wie z.B. das Exoenzym S von *Pseudomonas aeruginosa* und das murine ART2 verschiedene Zielproteine bzw. einfache Aminosäurenreste ADP-ribosylieren (Koch-Nolte *et al.* 1996a, Radke *et al.* 1999), während der Elongationsfaktor 2 das einzige Zielprotein des Diphterietoxins und des *Pseudomonas aeruginosa* Exotoxin A darstellt (Wilson and Collier 1992) oder das *Clostridium botulinum* C2-Toxin spezifisch Aktin ADP-ribosyliert (Aktories *et al.* 1986). Einige ARTs sind in der Lage, einzelne Aminosäuren oder Wasser als Zielmolekül zu verwenden. Bei Nutzung von Wasser als Akzeptor kommt es als Resultat der enzymatischen Reaktion zur Glykohydrolyse von NAD^+ zu Nikotinamid und freier ADP-Ribose. Dies wird als NADase-Aktivität des Enzyms bezeichnet (**Abbildung 1B**). Verschiedene ARTs besitzen zudem die Fähigkeit, sich durch die kovalente Übertragung von ADP-Ribose auf Aminosäurenreste der eigenen Polypeptidketten selbst zu modifizieren (sog. Automodifikation). Die bekannten ARTs unterscheiden sich deutlich in ihrer Transferase-, Automodifikations- und NADase-Aktivität.

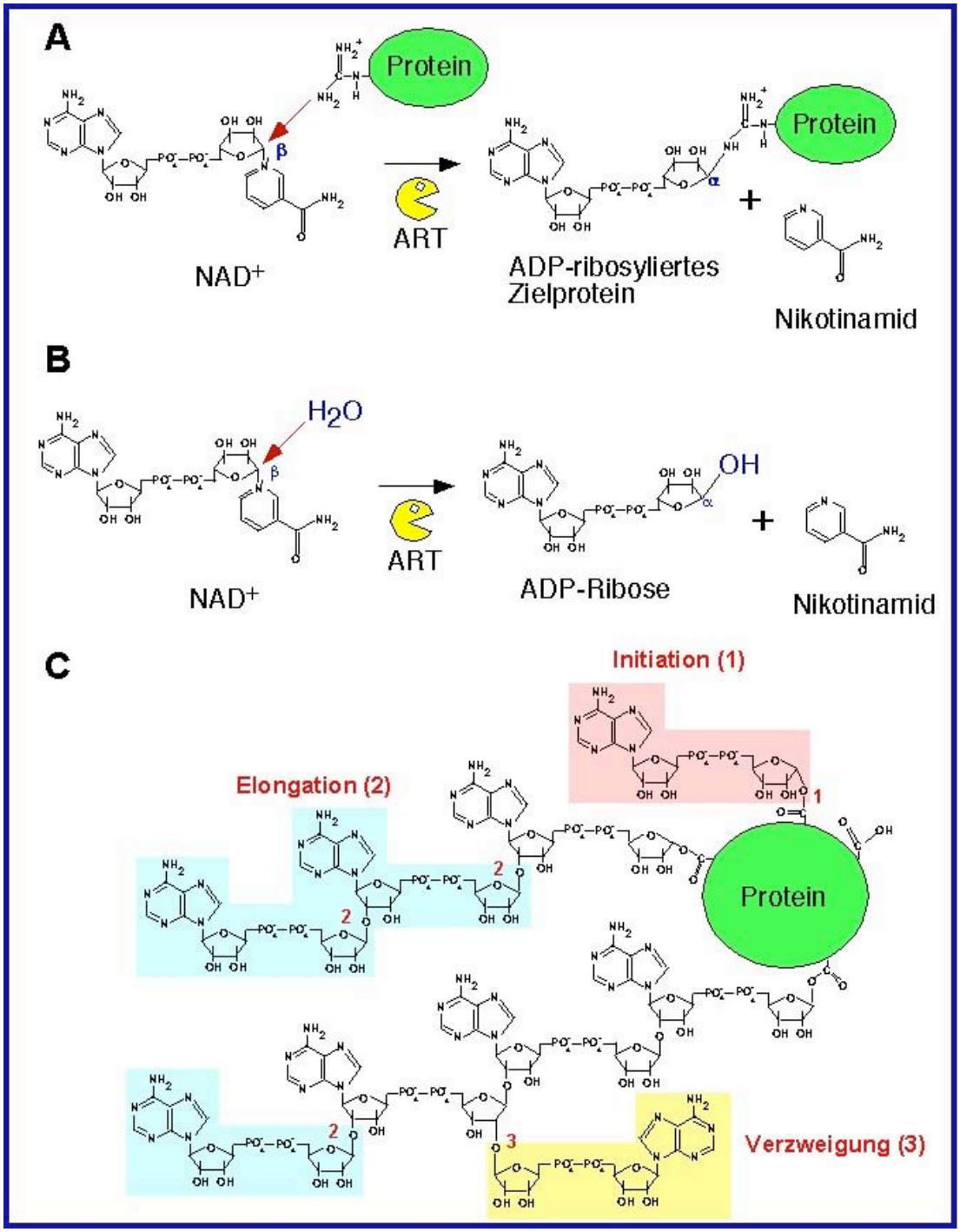


Abbildung 1: Schematische Darstellung der durch ARTs katalysierten Reaktionen

Das Substrat Nikotinamiddinukleotid (NAD^+) wird durch das ART-Enzym (gelber „Pacman") gebunden. Durch den nukleophilen Angriff eines Aminosäurenrestes (z.B. Arginin) auf die β-glykosidische Bindung zwischen Nikotinamid und Ribose wird NAD^+ gespalten und die ADP-Ribose-Gruppe auf die Zielaminosäure übertragen. Dabei kommt es zur Freisetzung von Nikotinamid (**A**). Einige ARTs besitzen eine NADase-Aktivität. Durch Glykohydrolyse von NAD^+ entsteht Nikotinamid und freie ADP-Ribose (**B**). **C** verdeutlicht die Polymerbildung durch die poly-ARTs. Nach der Ankopplung der ersten ADP-Ribose (Initiation) werden die ADP-Ribose-Reste über Ribose-Ribose 1''-2'-Bindungen verknüpft (Elongation). Das Polymer kann *in vitro* mehr als 200 Einheiten lang werden. Ca. alle 20-50 linearen ADP-Ribose-Einheiten tritt eine Verzweigung der Kette („branching") auf.

Chambon und Mitarbeitern gelang in den 60er Jahren der Nachweis eines zellulären Polymers, das initial als Homopolymer riboadenylierter Einheiten (poly A) identifiziert wurde (Chambon 1963). Durch genauere Analysen der Zusammensetzung des Polymers zeigte sich aber, dass es sich um ein Homopolymer aus ADP-Ribose-Einheiten handelte, die untereinander durch glykosidische Bindungen verknüpft waren (Reeder *et al.* 1967). Im Anschluss wurde neben den bekannten ARTs, die lediglich ein ADP-Ribose-Monomer übertragen können, ein Enzym nachgewiesen, das die Polymerisation von ADP-Ribose-Untereinheiten katalysiert. Diese sog. Poly-ADP-Ribosyl-Polymerase (PARP, oder auch Poly-ADP-Ribosyl-Transferase) besitzt die Fähigkeit, lange Zuckerketten an Zielproteine zu koppeln und diese Ketten zu verzweigen („branching“) (**Abbildung 1C**). Die Kettenlänge ist unterschiedlich und kann *in vitro* mehr als 200 Einheiten umfassen. Die Verzweigungen („branching“) treten bei längeren Ketten mit einer Verzweigungsfrequenz von ca. einer Verzweigung pro 20-50 linear verknüpften ADP-Ribose-Einheiten auf (D'Amours *et al.* 1999).

Die ADP-Ribosylierung ist ein reversibler Mechanismus der Proteinmodifikation. Die angekoppelte ADP-Ribose kann durch ADP-Ribosyl-Hydrolasen (ARH) wieder abgespalten werden, wodurch das Zielprotein in vielen Fällen reaktiviert wird. Für den Abbau der durch die PARP-Katalyse synthetisierten Polymere werden zwei verschiedene Enzyme benötigt: Die Poly-ADP-Ribose-Glykohydrolase (PARG) hydrolysiert die glykosidischen Bindungen am Ende bzw. inmitten des Polymers (Exo- und Endoglykosidase-Aktivität), während die ADP-Ribosyl-Protein-Lyase für die Hydrolyse der proximalsten ADP-Ribose-Einheit am Proteinakzeptor verantwortlich ist (D'Amours *et al.* 1999).

Die reversible Modifikation zellulärer Proteine stellt einen essentiellen Mechanismus in der Kommunikation einer Zelle mit ihrer Umgebung dar. Ein klassischer Mechanismus der reversiblen posttranslationalen Proteinmodifikation liegt bei der Phosphorylierung zellulärer Proteine durch Kinasen bzw. der Dephosphorylierung durch Phosphatasen vor (Harnett 1994). Die enzymatischen Komponenten für einen ADP-Ribosylierungs-Kreislauf konnten in Prokaryonten und Eukaryonten nachgewiesen werden. In dem photosynthetischen Bakterium *Rhodospirillium rubrum* wurde von P.W. Ludden der erste physiologische Mono-ADP-Ribosylierungskreislauf nachgewiesen (**Abbildung 2**, Ludden 1994).

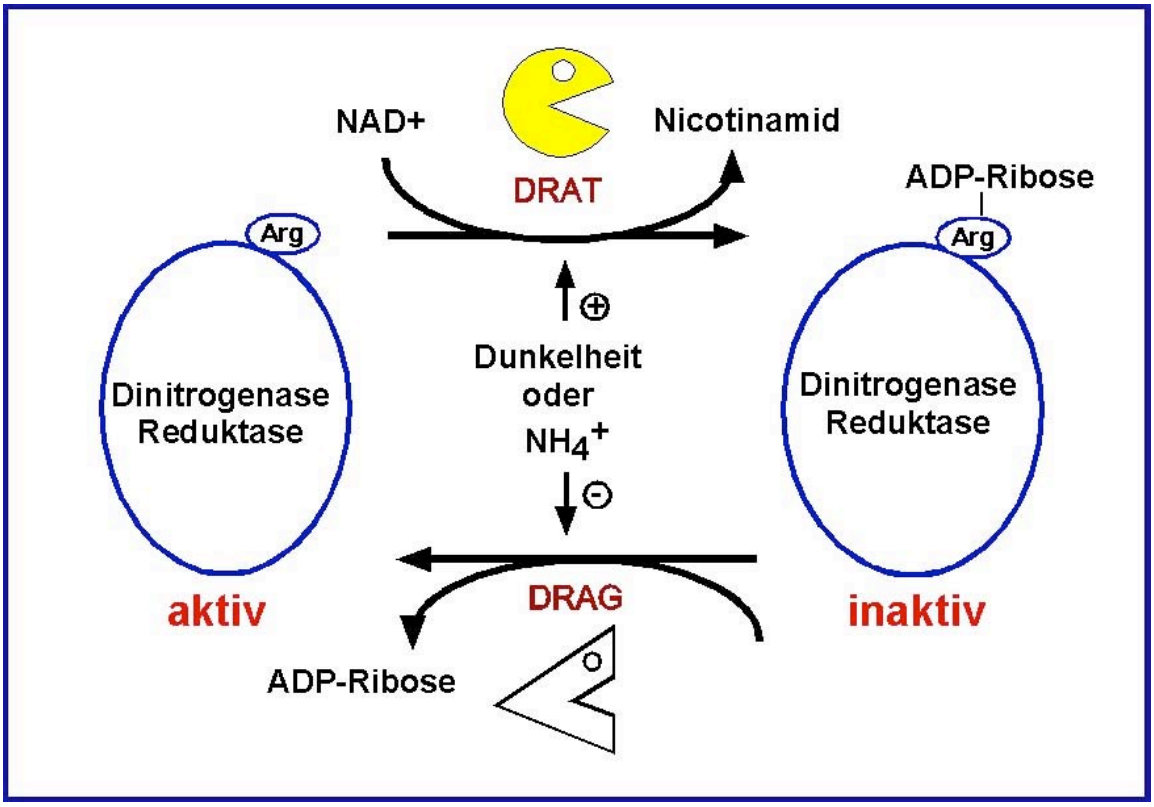


Abbildung 2: Regulation der Dinitrogenase-Reduktase von *Rhodospirillium rubrum*
Die **D**initrogenase-**R**eduktase-**A**DP-Ribosyl**t**ransferase (**DRAT**) ist als Pacman-Symbol dargestellt. DRAG = **D**initrogenase-**R**eduktase-**A**DP-Ribosyl**g**lykohydrolase. Arg = Argininrest des Zielproteins.

Die Stickstofffixierung des Bakteriums wird durch den sog. Dinitrogenase-Reduktase-Enzymkomplex reguliert. Bei ausreichenden Stickstoffreserven (z.B. in Form von Ammonium, Glutamin oder Asparagin) bzw. bei Dunkelheit wird ein spezifischer Argininrest des Enzymkomplexes durch die **D**initrogenase-**R**eduktase-**A**DP-Ribosyl**t**ransferase (**DRAT**) ADP-ribosyliert und somit inaktiviert. Unter Stickstoffmangelbedingungen bzw. bei Licht wird der „Gegenspieler“ der DRAT, die **D**initrogenase-**R**eduktase-**A**DP-Ribosyl-**g**lykohydrolase (**DRAG**), aktiviert und katalysiert die Freisetzung der gebundenen ADP-Ribose-Einheit. Durch die Regeneration der Guanidinogruppe des Arginins ist das Schlüsselenzym Dinitrogenase-Reduktase wieder zur Katalyse der Stickstofffixierung bereit.
In den letzten Jahrzehnten wurden ADP-Ribosyl-Transferasen und -Hydrolasen zunächst in Truthahnerythrozyten und schließlich in höheren Vertebraten sowie beim Menschen nachgewiesen (Williamson and Moss 1990, Yost and Moss 1983). Somit kann in Eukaryonten ein ähnlicher ADP-Ribosylierungs-Kreislauf zur Regulierung biologischer Funktionen postuliert werden, über das funktionelle Zusammenspiel der Enzyme ist im Detail aber noch wenig bekannt.

1.2 NAD^+ als Substrat

Der klassische intrazelluläre Metabolit NAD^+ wurde ursprünglich als „Cozymase“ aus dem Hefepilz isoliert und zunächst als Kofaktor bei Redoxreaktionen und Wasserstoffübertragungen des Intermediärstoffwechsels bekannt. Obwohl die Zellmembran von lebenden Zellen für NAD^+ undurchlässig ist, zeigte sich, dass NAD^+ analog zum ATP als Substrat auch in extrazelluläre Prozesse involviert ist. Erst kürzlich konnte Connexin 43 (Cx43) als transmembranöser Transporter für NAD^+ identifiziert werden: Connexin 43 ermöglicht durch Bildung von Hexameren (sog. Connexone oder Hemikanäle) einen bidirektionalen NAD^+-Transport, der u.a. durch die Ca^{2+}-Konzentration regulierbar ist (Bruzzone *et al.* 2001). Während die Connexine bisher als strukturelle Bestandteile von Zell-Zell-Kanälen zur interzellulären Kommunikation bekannt waren, konnte erstmals gezeigt werden, dass sie auch in isolierten Zellen einen transmembranösen NAD^+-Transport und somit potentielle parakrine Funktionen von NAD^+ ermöglichen. Mit der Freisetzung von NAD^+ im Rahmen von Gewebsschädigung und Zelluntergang (z.B. bei Entzündungen) bestehen zudem Mechanismen, die unabhängig von einem transmembranösen Transportsystem zu einem erhöhten Substratangebot für die extrazellulären ARTs führen.

Um den Metaboliten NAD^+ konkurrieren neben den mono- und poly-ARTs noch andere katabolisierende Enzyme: ADP-Ribosyl-Cyclasen wie das Zelloberflächenmolekül CD38 katalysieren die Bildung einer intramolekularen Bindung im ADP-Ribose-Anteil des NAD^+, so dass zyklische ADP-Ribose (cADPR), ein potenter „second messenger“ zur intrazellulären Ca^{2+}-Mobilisierung, entsteht (Guse 2004, Partida-Sanchez *et al.* 2003). Zudem besitzen sie eine NADase-Funktion. Als weiterer potentiell konkurrierender Mechanismus wurde *in vitro* die nicht-enzymatische Ankopplung von ADP-Riboseeinheiten an nukleophile Aminosäureketten, meist Cystein- oder Lysinreste, nachgewiesen. In **Abbildung 3** sind die vorgestellten NAD^+-katabolisierenden Pfade schematisch abgebildet.

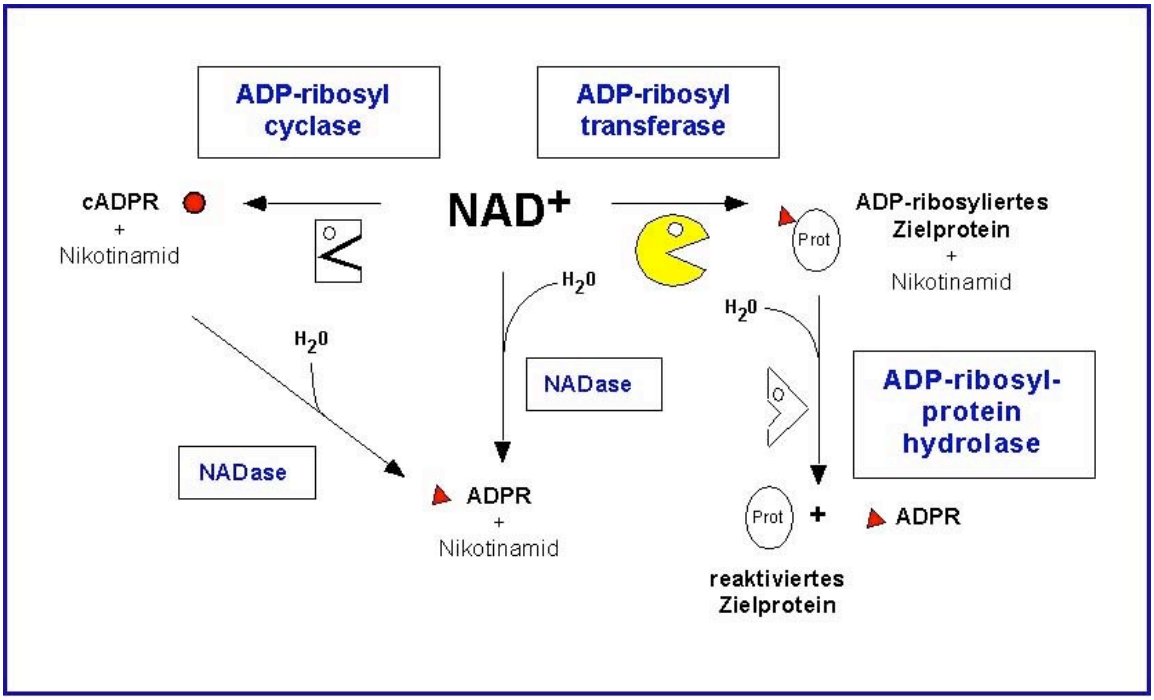


Abbildung 3: Schematisches Diagramm von ausgewählten konkurrierenden NAD^+-katabolisierenden Prozessen

Modifiziert aus: ADP-Ribosylation in animal tissues, Haag, F. und Koch-Nolte, F., Plenum Press, New York (Haag and Koch-Nolte 1997).

1.3 Prokaryontische mono-ADP-Ribosyltransferasen

Die bedeutende biologische Rolle der mono-ADP-Ribosylierung zeigte sich erstmals 1968 durch die Entdeckung, dass das Diphterietoxin des Erregers *Corynebacterium diphteriae* über diesen Mechanismus irreversibel einen Translationsschritt der Proteinbiosynthese in humanen Zellen blockiert (Honjo *et al.* 1968). Seitdem konnten ART-Familienmitglieder in vielen Prokaryonten identifiziert werden, darunter diverse bakterielle Toxine wie das Cholera Toxin, die *Pseudomonas aeruginosa* Exotoxine A und S und das Pertussis Toxin, die aufgrund ihrer human-pathogenen Wirkung von großer Bedeutung sind. Die meisten dieser bakteriellen mono-ARTs sind sezernierte Enzyme, die Proteine anderer Organismen modifizieren. Die bakteriellen Toxine können basierend auf ihrer Zielprotein- und Aminosäurenspezifität oder bezügl. ihrer Struktur und Aminosäuresequenzhomologie in Subgruppen eingeteilt werden.

1.3.1 Der Prototyp Diphterietoxin

Das Diphterietoxin (DT), ein von lysogenen *Corynebacterium diphteriae*-Stämmen sezerniertes 58kD Protein, ist das beherrschende pathogene Prinzip bei der Entstehung der Diphterie und war das erste im Detail studierte ADP-ribosylierende bakterielle Toxin. Das Diphterietoxin wird als inaktives Proenzym synthetisiert und besteht aus drei funktionell verschiedenen Domänen (Bell and Eisenberg 1996, Choe *et al.* 1992): Die C-terminale Domäne bindet an einen Zelloberflächenrezeptor der Zielzelle und ermöglicht so die rezeptorvermittelte Endozytose des Toxins. Durch die pH-Änderung im Endosom kommt es nach Konformationsänderung der Translokationsdomäne zum Transfer der N-terminalen katalytischen Untereinheit durch die endosomale Membran in das Zytoplasma. Dort katalysiert sie die ADP-Ribosylierung eines Diphtamid-Restes (posttranslational modifiziertes Histidin) des Elongationsfaktors 2 (EF-2) und blockiert somit irreversibel die Proteinbiosynthese der Zielzelle (van Ness *et al.* 1980). Dies führt zum Tod der Zelle. Der Mechanismus ist so effektiv, dass eine einzelne katalytische Domäne des Diphterietoxins die Wirtszelle töten kann (Yamaizumi *et al.* 1978).

Hinsichtlich des Krankheitsbildes muss zwischen lokaler Infektion und systemischer Intoxikation unterschieden werden. Die lokale Infektion betrifft zumeist die Rachenschleimhaut, insbesondere die Tonsillen. Das am Ort der Infektion produzierte Diphterietoxin schädigt die Schleimhautzellen. Die folgende Entzündungsreaktion führt zur Ausbildung der charakteristischen diphterischen Pseudomembran. Als systemische toxische Komplikationen kann das DT durch hämatogene Ausbreitung u.a. eine Myokarditis, eine Nephropathie und eine Lähmung der Hirnnerven (Polyneuritis diphterica) verursachen.

1.3.2 Identifizierung und Charakterisierung weiterer prokaryontischer ARTs

In den letzten Jahrzehnten konnten für viele wichtige bakterielle Toxine ähnliche Wirkungsmechanismen nachgewiesen werden. Bezüglich ihrer Proteinsubstrate lassen sich die humanpathogenen Toxine in mindestens 5 Gruppen unterteilen:

a) Toxine, die den Elongationsfaktor 2 (EF-2) ADP-ribosylieren

Das *Pseudomonas aeruginosa* Exotoxin A (ETA) besteht analog dem DT aus drei Domänen, von denen die Domänen 1 und 2 für das Einschleusen der katalytischen Domäne 3 in das

Zytoplasma der Wirtszelle verantwortlich sind (Kounnas *et al.* 1992). Dort wird die katalytische Domäne durch proteolytische Spaltung des Holotoxins aktiviert und durch die ADP-Ribosylierung eines Dipthamid-Restes von EF-2 irreversibel die Protein-biosynthese blockiert. Im Gegensatz zum DT bildet die katalytische Domäne beim ETA das C-terminale Ende des sezernierten Proteins.

b) Toxine, die heterotrimere G-Proteine ADP-ribosylieren

Cholera-Toxin (CT), Pertussis-Toxin (PT) und das *E.Coli* „heat labile enterotoxin" (LT) sind aus mehreren Untereinheiten zu $A(B_5)$-Komplexen zusammengesetzt. Das B-Pentamer vermittelt die Penetration des katalytischen A-Monomers in das Zytoplasma der Wirtszelle. Auch bei diesen Toxinen ist zuvor eine proteolytische Spaltung sowie die Reduktion einer Disulfidbindung erforderlich. CT und LT modifizieren in der Wirtszelle durch ADP-Ribosylierung einen Arginin-Rest der regulatorischen α-Untereinheit heterotrimerer GTP-bindender Proteine, was eine permanente Aktivierung der α-Untereinheit bewirkt. Dies führt über eine massiv gesteigerte Adenylatzyklase-Aktivität zu hohen cAMP-Spiegeln und somit zum Efflux von Wasser und Ionen aus der betroffenen Zelle. Die Folge dieses Mechanismus sind gastrointestinale Symptome, die von einer milden Reisediarrhoe bis zu ausgeprägten reiswasserartigen Durchfällen mit massiver Dehydratation und möglicher Todesfolge reichen können. Das Pertussis-Toxin ADP-ribosyliert die α-Untereinheit bestimmter heterotrimerer G-Proteine an einem Cystein-Rest, wodurch es zu einer Entkopplung der G-Proteine von ihren Rezeptoren und damit zur veränderten Regulation der Signaltransduktion kommt (Passador and Iglewski 1994).

c) Toxine, die die GTP-bindende Rho-Familie ADP-ribosylieren

Zu dieser Gruppe zählen das *Clostridium botulinum* C3-Exotoxin (C3), die *Clostridium limosum* Transferase, die *Bacillus cereus* transferase (Just *et al.* 1995) und der „epidermal differentiation inhibitor" (EDIN) von *Staphylococcus aureus* (Sugai *et al.* 1992). Diese Toxine modifizieren Rho-Proteine an einem Asparagin-Rest. Da die Rho-Proteine eine Rolle in der Regulierung des Aktinskeletts spielen, kommt es durch die Inaktivierung zu einer Depolymerisierung des Aktinskeletts. Zudem werden über die Rho-ADP-Ribosylierung weitere zelluläre Signaltransduktionsprozesse beeinflusst. Beim C3-Exoenzym konnten die für andere Toxine typischen Zellbindungs- und Translokationskomponenten nicht

identifiziert werden, so dass bisher unklar ist, ob es sich bei C3 um ein unabhängiges Toxin oder um die Komponente eines Holotoxins handelt und über welchen Mechanismus es in die Wirtszelle gelangt.

d) Toxine, die Proteine der ras-Familie ADP-ribosylieren

Neben dem oben beschriebenen Exotoxin A sezerniert *Pseudomonas aeruginosa* mit den Exoenzymen S und T zwei weitere ADP-Ribosyltranferasen, die beide Proteine der ras-Familie ADP-ribosylieren. Exoenzym S ist ein bifunktionelles Enzym, das sich aus einem N-terminalen Rho-GTPase-aktivierenden Protein (GAP) und der ART-Domäne zusammensetzt. In die Wirtszelle gelangt das Toxin mit Hilfe des von verschiedenen gramnegativen Bakterien ausgebildeten, komplexen Typ-III-Sekretionsapparates, einer Art „molekularen Spritze“, die aus hochkonservierten Komponenten besteht (Kubori *et al.* 1998). Neben der ADP-Ribosylierung von ras-Proteinen besitzt das Exoenzym S die Fähigkeit, sich an der GAP-Domäne selbst zu ADP-ribosylieren und somit seine GAP-Aktivität zu inhibieren (Riese *et al.* 2002). Exoenzym S induziert über seine Wirkungsmechanismen Apoptose in den Wirtszellen und Untersuchungen mit Exoenzym S-Mutanten zeigten, dass die ART-Aktivität für diese Induktion essentiell ist (Kaufman *et al.* 2000).

e) Toxine, die Aktin ADP-ribosylieren

Das *Clostridium botulinum* C2-Toxin (C2), das *Clostridium perfringens* Iota-Toxin, das *Bacillus cereus* VIP2-Toxin und weitere Clostridien-Toxine ADP-ribosylieren G-Aktin an einem Arginin-Rest und hemmen dadurch die Aktin-Polymerisierung. Das C2-Toxin besteht wie das VIP2-Toxin aus zwei separaten Proteinkomponenten (Aktories and Wegner 1992): C2II bindet an die Zelloberfläche und ermöglicht die endozytotische Aufnahme der katalytischen Komponente C2I (Simpson 1989).

Desweiteren wurde einige bakterielle Toxine als ARTs identifiziert, die sich den genannten Hauptgruppen nicht zuordnen lassen: So modifiziert z.B. das ADP-ribosylierende Enzym eines *E.Coli* Phagen (ALT) eine RNA-Polymerase als Zielprotein. Bei dem endosymbiontischen marinen Bakterium *Vibrio fisheri* konnten zwei immunologisch verschiedene ARTs, sog. Halovibrine (Halovibrin αund β), charakterisiert werden, die

nachweislich eine Rolle in der Entwicklung des Lichtorgans des Wirtsorganismus (z.B. bei bestimmten Tintenfischen) spielen (Reich *et al.* 1997).

Während alle bisher beschriebenen ARTs die Übertragung von ADP-Ribose auf Aminosäurenreste von Proteinen oder einzelne Aminosäuren katalysieren, wurden auch zwei ARTs entdeckt, die andere Zielmoleküle modifizieren: *Mycobacterium smegmatis* sezerniert eine ART, die anstelle eines Proteins das Antibiotikum Rifampin an einer OH-Gruppe ADP-ribosyliert und somit inaktiviert (Quan *et al.* 1999). Das Toxin Pierisin, ein 98kD-Protein des Schmetterlings *Pieris rapae*, transferiert die ADP-Ribose auf 2‘-Desoxy-Guanosin-Reste doppelsträngiger DNA und induziert Mutationen und Apoptose in Säugetierzelllinien (Takamura-Enya *et al.* 2001, Totsuka *et al.* 2003). Pierisin zeigt eine bemerkenswerte Sequenzidentität mit dem moskitoziden Toxin von *Bacillus sphaericus* SSII-1 (MTX), für dessen N-terminale Domäne ebenfalls eine argininspezifische, promiskuitive ART-Aktivität nachgewiesen wurde, die für die Zytotoxizität des Toxins in Säugetierzellen essentiell ist (Schirmer *et al.* 2002a, Schirmer *et al.* 2002b, Thanabalu *et al.* 1993).

Im Gegensatz zu den sezernierten mono-ARTs ist die bereits ausführlicher vorgestellte DRAT aus *Rhodospirillium rubrum* die einzige bekannte prokaryontische ART, die ihre Wirkung intrazellulär entfaltet. Über Datenbanksuchen in den inkompletten prokaryontischen Genomen konnten während der Fertigstellung der Arbeit weitere putative ARTs gefunden werden (Pallen *et al.* 2001), so dass die bisher charakterisierten Toxine möglicherweise lediglich die „Spitze eines Eisbergs“ darstellen. In **Abbildung 4** ist der Wirkungsmechanismus ausgewählter Toxine schematisch dargestellt.

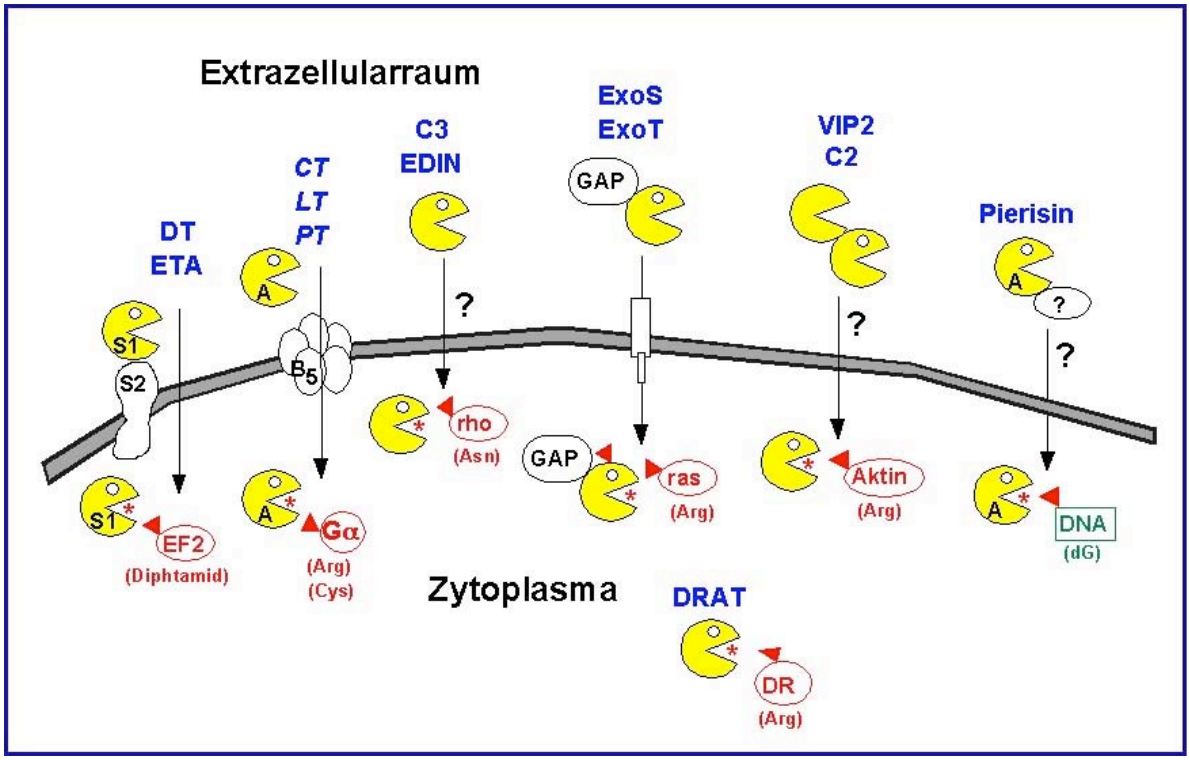


Abbildung 4: Schematische Darstellung wichtiger prokaryontischer ARTs und ihrer Wirkungsmechanismen

Die prominentesten bakteriellen ARTs sind nach ihren Zielproteinen in fünf Gruppen eingeteilt. Die Abbildung zeigt die Zielproteine mit den jeweiligen Zielaminosäuren. Desweiteren ist, soweit bekannt, der Mechanismus, mit dem die katalytischen Domänen ins Zytoplasma der Wirtszelle gelangen, schematisch wiedergegeben. Aufgrund ihrer besonderen Eigenschaften sind zudem die intrazelluläre DRAT von *Rhodospirillium rubrum* und das Pierisin des Schmetterlings *Pieris rapae* abgebildet. Pierisin ADP-ribosyliert DNA anstelle eines Zielproteins. Die katalytischen Domänen der ART-Enzyme sind als „Pacman“, NAD^+ als rotes Sternchen und die gebundene ADP-Ribose als rotes Dreieck dargestellt. Bei den binären Toxinen VIP2 und C2 ist die strukturhomologe, katalytisch inaktive Domäne als „Pacman ohne Auge“ abgebildet. GAP = Rho-GTPase-aktivierende Domäne. Gα = regulatorische α-Untereinheit heterotrimerer GTP-bindender Proteine. Abbildung modifiziert aus Koch-Nolte *et al.*, Immunology Today, 1996 (Koch-Nolte *et al.* 1996a).

Die vorgestellten Toxine besitzen zumeist ein hohe Aminosäurespezifität. Während eine Vielzahl der mono-ARTs Arginin-spezifisch ist (z.B. CT, LT, C2, VIP2, ALT, DRAT), werden von den anderen mono-ARTs Diphtamid-Reste (DT, ETA), Asparagin-Reste (C3) und Cystein-Reste (PT) ADP-ribosyliert. Die sezernierten bakteriellen mono-ARTs werden typischerweise durch mobile genetische Elemente wie Bakteriophagen (DT, ALT, C3 Exotoxin), Plasmide oder sog. pathogenetische Inseln (CT) kodiert. Einige Toxine (z.B. PT, CT) werden heute in der Molekular- und Zellbiologie als Werkzeuge zur Aufklärung von Signaltransduktionswegen in Eukaryonten eingesetzt (Jacobson and Jacobson 1989). In

absolutem Gegensatz zu der Vielzahl an mono-ADP-Ribosyltransferasen konnte durch Sequenzanalysen bisher in prokaryontischen Genomen mit der potentiellen PARP des Insektenvirus Chilo iridescent lediglich ein PARP-Homologes ermittelt werden. Auch wurden bisher keine prokaryontischen ADP-Ribosyl-Cyclasen beschrieben.

1.4 Topologie der katalytischen Domäne bekannter ARTs

Der Vergleich der Aminosäuresequenzen zeigte für einzelne Toxine moderate bis starke Sequenzidentitäten, während die Toxine unterschiedlicher Subgruppen auf den ersten Blick kaum Sequenzhomologien aufzuweisen schienen. Nach genauerer Analyse der vorhergesagten Sekundärstrukturmotive sowie der Tertiärstruktur war jedoch eine starke Konservierung der katalytischen Domäne über die Grenzen biologischer Klassen hinweg zu erkennen. Zu Beginn dieser Arbeit waren fünf Kristallstrukturen von bakteriellen Toxinen (ETA, LT, DT, PT, CT) sowie die Kristallstruktur der Hühner-PARP bekannt (Allured *et al.* 1986, Choe *et al.* 1992, Ruf *et al.* 1996, Sixma *et al.* 1991, Stein *et al.* 1994, Zhang *et al.* 1995). Im Verlauf konnte die Struktur von VIP2 aufgeklärt werden, und erst kürzlich wurden die Strukturen von *Clostridium botulinum* C3 Exotoxin, *Clostridium perfringens* Iota-Toxin, Staphylococcus aureus EDIN B (auch C3stau2 genannt), murinem PARP-2 sowie der ersten eukaryontischen mono-ART, dem Ratten-ART2.2, publiziert (Evans *et al.* 2003, Han *et al.* 1999, Han *et al.* 2001, Mueller-Dieckmann *et al.* 2002, Oliver *et al.* 2004, Tsuge *et al.* 2003). Bei einigen der ARTs gelang die Co-Kristallisierung von NAD^+ im Komplex mit dem Enzym, so dass sich weitere Aufschlüsse über die Bindung des Substrates ergaben.

Sämtliche Kristallstrukturen zeigen eine hochkonservierte „Pacman“-artige Faltung der katalytischen Domäne (**Abbildung 5**). Das Gerüst des katalytischen Zentrums ist durch die charakteristische Anordnung von sechs bzw. sieben β-Strängen gekennzeichnet, die in zwei β-Faltblättern angeordnet den katalytischen Spalt für die NAD^+-Bindung ausbilden (Ober- und Unterkiefer des Pacman). Die Glutaminsäure (E) im β5-Strang ist in allen Strukturen streng konserviert. Die Wichtigkeit dieser Glutaminsäure für die katalytische Aktivität wurde zunächst durch gezielte Mutagenese beim Diphterietoxin demonstriert (Carroll and Collier 1984) und konnte bei der Untersuchung weiterer ARTs bestätigt werden (Bazan and Koch-Nolte 1997, Domenighini and Rappuoli 1996). Selbst ein konservativer Austausch dieser Aminosäure führt in der Regel zum Verlust der Aktivität. Anhand der konservierten

Positionierung weiterer charakteristischer Aminosäurenreste lässt sich die ART-Familie in zwei Untergruppen einteilen: Bei der ersten Gruppe, zu der die meisten mono-ARTs zählen, sind ein Arginin-Rest (R) in β1 und ein Serin-Rest (S) in β2 konserviert, so dass ein sog. R-S-E-Motiv ausgebildet wird (**5A**). Bei den Diphtamid-spezifischen mono-ARTs Diphterietoxin und *Pseudomonas* Exotoxin A und den bekannten poly-ARTs sind diese Positionen in β1 durch Histidin (H) und in β2 durch Tyrosin (Y) ersetzt, die zweite Gruppe wird somit durch ein H-Y-E-Motiv charakterisiert (**5B**). Bei allen bekannten Arginin-spezifischen mono-ARTs ist zwei Positionen N-terminal der katalytischen Glutaminsäure in β5 eine weitere, für die Transferaseaktivität wichtige Glutaminsäure konserviert. Dies wird als R-S-EXE-Motiv bezeichnet (Bazan and Koch-Nolte 1997, Domenighini and Rappuoli 1996). Interessanterweise konnte u.a. für das *Clostridium botulinum* C2-Toxin und für das ART2 der Ratte gezeigt werden, dass eine Mutation dieser zweiten Glutaminsäure zu Glutamin die Transferaseaktivität weitgehend in eine NAD-Glykohydrolase-Aktivität umwandelt (Barth *et al.* 1998b, Bredehorst *et al.* 1997, Hara *et al.* 1996, Koch-Nolte *et al.* 2001, Maehama *et al.* 1996). Als weiteren auffälligen strukturellen Unterschied zeigt eine große Untergruppe der RSE-ARTs (z.B. VIP2 und ART2) einen siebten β-Strang, der sich zwischen β6 und β3 schiebt (**5A und C**, siehe auch Abb.34). Die HYE-Gruppe besitzt eine längere konservierte α-Helixregion im Anschluß an β2 (gelb markiert). Die Verbindung („loop“) zwischen β4 und β5, deren Länge zwischen den verschiedenen ARTs deutlich differiert und die vermutlich für die Zielspezifität wichtig ist (Barbieri and Sun 2004, Han *et al.* 2001, Holbourn *et al.* 2005, Koch-Nolte *et al.* 2001), ist in der Abbildung durch einen Pfeil markiert. In der 3D-Struktur der Hühner-PARP mit dem Liganden carba-NAD (3pax) zeigte sich, dass der ADP-Anteil des Liganden mit dieser peripher gelegenen Region und der α-Helixregion (α2) im Anschluß an β2 interagiert. Daraus wurde geschlossen, dass es sich um die Acceptor-Bindungsstelle für die Poly-ADP-Ribosylierung handelt (Ruf *et al.* 1998). Die Beobachtung, dass bestimmte Mutationen in der α2-Helix zu einem Verlust der Fähigkeit zur Bildung von Polymeren führt, bestätigte diese Annahme. Zudem wurde gezeigt, dass die katalytische Glutaminsäure (E988) der PARP eine Doppelfunktion besitzt: Neben der Polarisierung des Donors NAD+ steigert sie durch Bildung von Wasserstoffbrücken die Nukleophilie des Acceptors (Ruf *et al.* 1998).

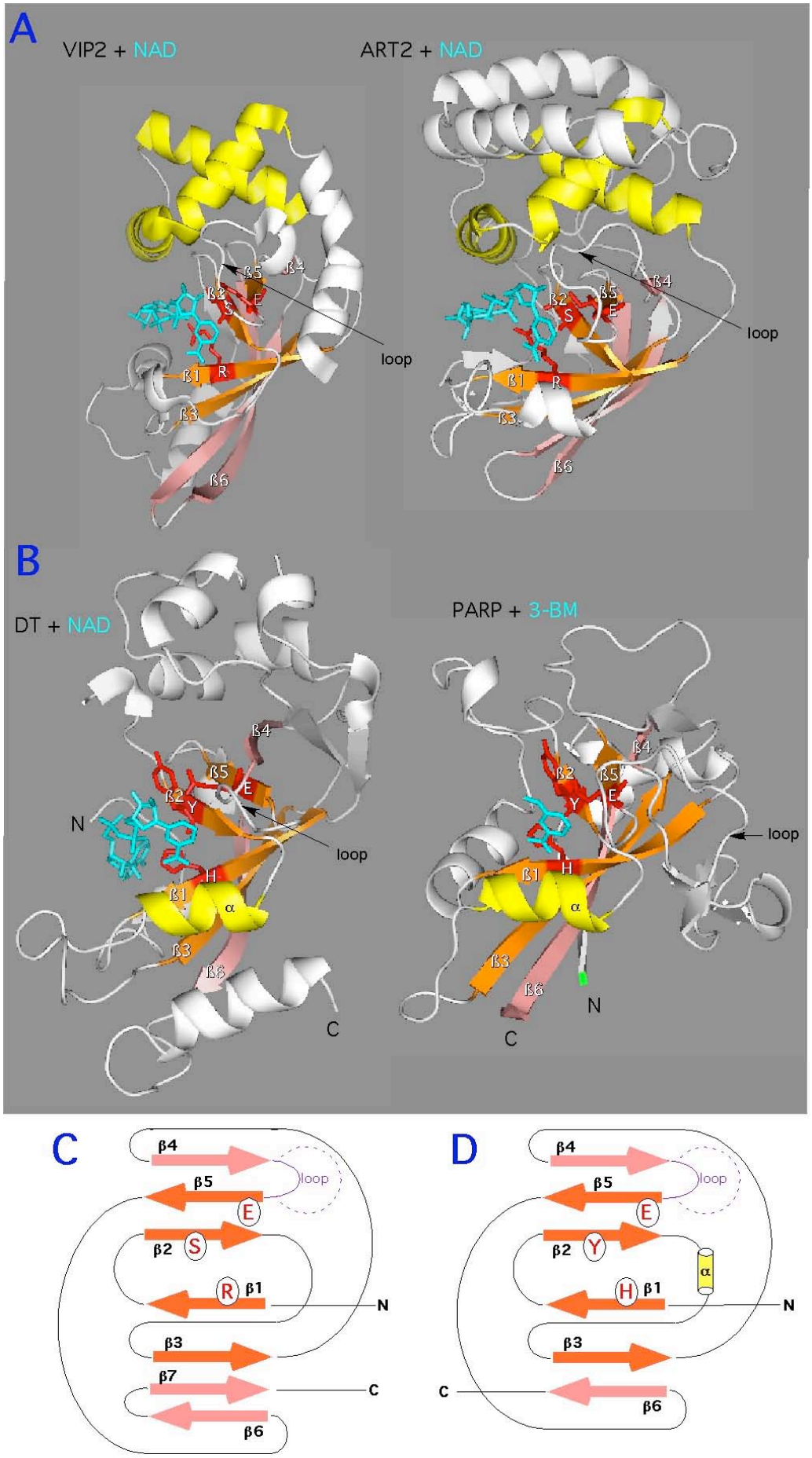


Abbildung 5: Darstellung der dreidimensionalen Struktur von VIP2, ART2 (A), DT und PARP (B)

Die Abbildung zeigt ein Modell der dreidimensionalen Strukturen von VIP2, ART2 (A), DT und PARP (B) in Cokristalisierung mit den Substraten NAD bzw. 3-MB (3-Methoxybenzamid). Die Strukturen sind in der Vorderansicht abgebildet, so dass die gebundenen Liganden (türkis) im aktiven Zentrum gut sichtbar sind. Durch die Darstellung der Liganden wird die durch die „Pacman"-artige Faltung der Moleküle hervorgerufene „Tasche" mit den charakteristischen Aminosäuren (RSE bzw. HYE) verdeutlicht. Zudem ist die Anordnung der entscheidenden Sekundärstrukturmotive für eine große Untergruppe der RSE-ARTs (C) bzw. für die HYE-ARTs (D) zur Verdeutlichung schematisiert dargestellt. Die Kristallstrukturen sind in der Proteindatenbank (PDB) unter folgenden Zugangsnummern deponiert: VIP2 1qs2, Ratten-ART2.2 1og3, DT 1tox, Hühner-PARP 3pax. Die Modelle (A,B) wurden mit dem Programm Pymol erstellt (DeLano 2002). Die konservierten β-Stränge der katalytischen Domäne sind farblich hervorgehoben, wobei die zentralen β-Stränge β1, β2, β3 und β5 orange und die weiter peripher gelegenen β4 und β6 bzw. im Falle von VIP2 und mART2 zusätzlich β7 lachsfarben markiert sind. Die Lage der charakteristischen Aminosäuremotive mit den substratbindenden Endgruppen ist eingezeichnet und rot markiert. Der „loop" zwischen β4 und β5 ist durch einen Pfeil gekennzeichnet (A,B) bzw. violett markiert (C,D). Die konservierte α-Helix zwischen β2 und β3 (DT und PARP) ist gelb markiert. Der N-Terminus der katalytischen Domäne der PARP entspricht dem Phase-0-Intron vor dem β1-Strang (vergleiche Abb.14 und 15). Die Struktur der katalytischen Domäne von Diphtherie-Toxin ist C-terminal durch die proteolytische Spaltstelle begrenzt, die diese Domäne von der Translokationsdomäne trennt.

1.5 Die Spezies *Salmonella enterica*

Die Salmonellen gehören zu einer für die Humanmedizin bedeutenden Bakterienfamilie, den *Enterobacteriacae*. Salmonellen sind gramnegative, fakultativ intrazelluläre Stäbchenbakterien. Die *Salmonella enterica* werden in sieben Subspezies unterteilt: Der Großteil der humanpathogenen Salmonellen gehört zur Subspezies I, während die Subspezies II, IIIa, IIIb, IV, VI und VII zumeist bei kaltblütigen Vertebraten vorkommen. Desweiteren werden die Salmonellen, beruhend auf den unterschiedlichen spezifischen Polysaccharidketten des Lipopolysaccharid (LPS)-Komplexes der äußeren Membran (O-Antigene) sowie den Geißelantigenen (H-Antigene), in sog. Serovare eingeteilt (Kauffmann-White-Schema) (Kayser *et al.* 1993). Die ausgeprägte Antigenvariation spiegelt sich in über 2000 verschiedenen Serotypen wider. Bezüglich ihrer Pathogenese und der hervorgerufenen Krankheitsbilder lassen sich zwei Gruppen von Salmonellen unterscheiden:

Die typhösen Salmonellen *S.typhi* und *S.paratyphi* A,B,C können nach Eindringen über den Gastrointestinaltrakt durch hämatogene Streuung den Typhus bzw. den Paratyphus verursachen, eine systemische Erkrankung mit Fieberkontinua, Benommenheit und schwerem Krankheitsgefühl, die im Falle des Typhus unbehandelt in ca. 15% zum Tode führt.

Zur zweiten Gruppe, den enteritischen Salmonellen, gehören unter anderem die Serovare *S.enteritidis, S.typhimurium* und *S.dublin*. Während für die typhösen Salmonellen der Mensch

das einzige Erregerreservoir darstellt, sind die enteritischen Salmonellen tier- und menschenpathogen und können abhängig von Serovar und Wirtsspezies ein breites Spektrum an Krankheiten verursachen. Die *S.enterica* werden in Mitteleuropa als die wichtigsten Erreger bakteriell bedingter Gastroenteritiden beim Menschen angesehen. Die Infektion erfolgt meistens über Nahrungsmittel, in denen sich die Salmonellen vermehren konnten. Durch die Invasion der Darmschleimhaut und verschiedene Enterotoxine wird eine Enteritis hervorgerufen, die mit Diarrhoen, abdominellen Schmerzen, Nausea, Emesis und Fieber einhergeht. In der Regel bleibt die Infektion beim Menschen lokal auf den Darm beschränkt und heilt unbehandelt rasch aus. Komplikationen sind vor allem durch den Wasser- und Elektrolytverlust bedingt. Insbesondere bei immunsupprimierten Individuen können die enteritischen Salmonellen zudem septische Krankheitsbilder verursachen.

1.5.1 Der Spv-Locus

Viele *S.enterica*-Serovarianten tragen Virulenzplasmide, die bei diesen Stämmen für die Pathogenese einer systemischen Erkrankung eine essentielle Rolle spielen. Obwohl diese Virulenzplasmide 50 bis 96kb umfassen, konnte nachgewiesen werden, dass die Gene einer hochkonservierten, 8kb großen Region dieser Plasmide ausreichen, um plasmidfreie Stämme für Mäuse virulent zu machen (Gulig *et al.* 1993, Williamson *et al.* 1988). Diese Region besteht aus insgesamt fünf Genen, dem regulatorischen *spv* (**s**almonella **p**lasmid **v**irulence) *R* und den vier strukturellen Genen *spv A,B,C* und *D*. *SpvR* kodiert für einen Transkriptionsregulator, der zur LysR/MetR-Familie der prokaryontischen regulatorischen Proteine gehört und der die Expression des *spvABCD*-Operons aktiviert (Fang *et al.* 1991, Taira *et al.* 1991b). Für eine effiziente Expression ist zudem der alternative Sigmafaktor RpoS erforderlich (Fang *et al.* 1992). Die *spv*-Gene wurden inzwischen in vielen verschiedenen *S.enterica* Subspezies nachgewiesen. Interessanterweise ist der *spv*-Locus bei den bei Kaltblütern verbreiteten Subspezies II, IIIa, IV und VII nicht auf Plasmiden, sondern auf dem Chromosom lokalisiert (Boyd and Hartl 1998).

Nachdem es Hinweise gab, dass die Fähigkeit virulenter Salmonellenstämme, in phagozytotischen Zellen des Wirts zu überleben und sich zu vermehren, entscheidend für die Entwicklung einer schweren systemischen Erkrankung ist, zeigten Richter-Dahlfors *et al.* zunächst bei Mäusen, dass die virulenten Salmonellen während einer systemischen Infektion

in Makrophagen der Milz und der Leber lokalisiert sind, dort proliferieren und zytotoxische Effekte verursachen (Richter-Dahlfors *et al.* 1997). Die Beobachtung, dass die Expression von SpvB nach der Phagozytose durch Makrophagen oder nach der Invasion in epitheliale Zellen induziert wird (Fierer *et al.* 1993), ließ eine wichtige Rolle des SpvB für die Proliferation nach intrazellulärer Infektion vermuten. Nach Untermauerung dieser Vermutung im Mausmodell (Gulig and Doyle 1993) wurde in den letzten Jahren gezeigt, dass die *spv*-Gene für die Vermehrung der Bakterien in Makrophagen notwendig sind (Libby *et al.* 1997) und dass deren Expression für die Ausbildung zytotoxischer Effekte in infizierten humanen Makrophagen essentiell ist (Libby *et al.* 2000). Trotz dieser Erkenntnisse blieben die molekularen Wirkungsmechanismen der von den *spv*-Genen kodierten Proteine ungeklärt. Im Rahmen dieser Arbeit wird gezeigt, dass SpvB eine arginin-spezifische ART ist und als Zielprotein Aktin modifiziert. Die Tatsache, dass die bisher beschriebenen, humanpathogenen ADP-ribosylierenden Enzyme alle aus Bakterien stammen, die extrazelluläre Infektionen bewirken, und dass sich bei Sequenzanalysen keine bedeutenden Homologien zu Salmonellenproteinen zeigten, verdeutlichen, warum es vor Beginn dieser Arbeit keinen Hinweis für das Vorkommen von mono-ARTs in der Salmonellenfamilie gab.

1.6 Eukaryontische ADP-Ribosyltransferasen

Der Umstand, dass die bakteriellen Toxine durch die ADP-Ribosylierung entscheidend in die Regulation verschiedener eukaryontischer Proteinfunktionen eingreifen, warf alsbald die Frage auf, ob bei höheren Organismen endogene Verwandte der bakteriellen Toxine vorkommen, die ebenfalls regulatorische Funktionen besitzen. Tatsächlich führte die Inkubation verschiedener Zelltypen (u.a. Lymphozyten, Granulozyten, Makrophagen, Lymphozyten, Myozyten, Neurozyten und Fibroblasten) mit exogenem NAD^+ oder bekannten ART-Inhibitoren zu funktionellen Veränderungen der Zellen. Desweiteren wurde nach Zugabe von radioaktiv markiertem ^{32}P-NAD^+ die Inkorporation von ADP-Ribose in Proteine nachgewiesen (Haag and Koch-Nolte 1998, Okazaki and Moss 1996).

1.6.1 Eukaryontische poly-ADP-Ribosyltransferasen

Vor fast 40 Jahren wurden erstmals ADP-Ribose-Polymere in eukaryontischen Zellen entdeckt (siehe 1.1). Seitdem konnte die Wichtigkeit der poly-ADP-Ribosylierung für verschiedene zelluläre Prozesse erforscht werden (D'Amours *et al.* 1999). Als erste

eukaryontische ADP-Ribosyltransferase wurde eine poly-ART bei der Ratte kloniert (Hayaishi and Ueda 1977). Ende der 80er Jahre gelang schließlich die Aufreinigung und Klonierung der cDNA einer poly-ART aus humanen Geweben (Burtscher *et al.* 1986, Schneider *et al.* 1987, Uchida *et al.* 1987). In den nächsten Jahren wurde dieses Enzym, genannt PARP (**P**oly-**A**DP-**R**ibosyl-**P**olymerase), intensiv charakterisiert.

1.6.1.1 Die humane PARP

Das humane *PARP*-Gen ist 43kb lang, in 23 Exons unterteilt und auf Chromosom 1q41-q42 lokalisiert (Auer *et al.* 1989, Herzog *et al.* 1989). Das *PARP*-Gen kodiert für ein 113kD-Protein, das in den meisten Geweben reichlich vorhanden ist. PARP ist ein multifunktionales Enzym, das sich aus drei Domänen zusammensetzt: Über die N-terminale Domäne, die zwei strukturell und funktionell einzigartige Zink-Finger-Motive enthält, kann die PARP DNA binden (DNA-bindende Domäne). Die zentrale Domäne beinhaltet den Großteil der Regionen, an denen sich die PARP auto-ADP-ribosylieren kann, sowie eine BRCT-Domäne, die Protein-Protein-Interaktionen ermöglicht (Automodifikationsdomäne). Das C-terminale Ende bildet die 55kD umfassende katalytische Domäne. Die Kristallstruktur der katalytischen Domäne der Hühner-PARP bestätigte eindrucksvoll die vorhergesagte Strukturverwandtschaft zu den bakteriellen Toxinen (Ruf *et al.* 1996).

Für die PARP konnten *in vivo* und *in vitro* mehr als 30 Substrate (z.B. Histone, Topoisomerasen, DNA-Ligasen, DNA-Polymerasen, p53) identifiziert werden, von denen die meisten physiologischen Substrate nukleäre Proteine sind. Einer der Hauptakzeptoren ist aber die PARP selbst (Auto-ADP-Ribosylierung). Die Zielaminosäure der PARP ist zumeist Glutaminsäure, daneben wurde die Modifizierung anderer Reste, wie z.B. Aspartat, beschrieben (D'Amours *et al.* 1999, de Murcia and Menissier de Murcia 1994). In den folgenden Jahren wurden Homologe der PARP in Wirbeltieren, Insekten (z.B. *Drosophila melanogaster*), Pilzen (z.B. *Dictyostelium discoideum*) und Pflanzen (z.B. *Arabidopsis thaliana*) kloniert und charakterisiert (Hanai *et al.* 1998, Kofler *et al.* 1993, Lepiniec *et al.* 1995).

1.6.1.2 Die biologische Rolle der PARP

In den letzten Jahren konnte das Wissen über die biologische Funktion der PARP durch die Kombination von zellbiologischen, biochemischen und molekularbiologischen Techniken

sowie die Generierung von „knockout“-Modellen vermehrt werden. Unter normalen Bedingungen ist die PARP in weitgehend katalytisch inaktiver Form im Nukleoplasma lokalisiert. In zahlreichen Experimenten zeigte sich, dass DNA-Strang-Brüche, die z.B. durch Alkylantien oder ionisierende Bestrahlung induziert werden, zu einer Bindung der PARP an die unterbrochene DNA führen und es dadurch zu einer Aktivierung der katalytischen Domäne der PARP kommt. Die resultierende Auto-ADP-Ribosylierung hat, wahrscheinlich über die elektrostatische Abstoßung des negativ geladenen Polymers, das Ablösen der PARP von der DNA zur Folge. Obwohl die molekularen Mechanismen noch weitgehend unklar sind, gibt es zahlreiche Hinweise, dass die transiente Bindung der PARP an beschädigte DNA eine entscheidende Rolle in der Regulation der verschiedenen Mechanismen zur DNA-Reparatur spielt (D'Amours *et al.* 1999, Smith 2001). Diese Rolle wurde durch Versuche mit drei unabhängig voneinander generierten *PARP*-„knockout“-Mäusen (*PARP* -/-) untermauert, die eine Hypersensitivität für Alkylantien und ionisierende Strahlen zeigen (Shall and de Murcia 2000). Desweiteren besteht bei den Mäusen eine gesteigerte genomische Instabilität.

Neben der DNA-Reparatur werden der PARP weitere wichtige biologische Funktionen in der Regulation der Zellproliferation, -differenzierung und Genexpression sowie der Induktion des Zelluntergangs (Apoptose *vs.* Nekrose) zugeschrieben. Aufgrund dieser Multifunktionalität ist die PARP u.a. in die Pathogenese des Schlaganfalls, der Myokardischämie, verschiedener inflammatorischer Prozesse, bösartiger Neoplasien und von Autoimmunerkrankungen wie Diabetes mellitus Typ 1 involviert, so dass das molekulare Verständnis der PARP-Funktionen sowie die Etablierung potentieller therapeutischer Ansätze mit PARP-Inhibitoren Gegenstand intensiver Forschung sind (Virag and Szabo 2002).

1.6.1.3 Die Entdeckung von PARP-Verwandten

Nach der Klonierung der PARP beim Menschen und in anderen Eukaryonten wurde jahrelang angenommen, dass lediglich ein einziges poly-ADP-ribosylierendes Enzym existiert. Erste publizierte Hinweise auf weitere poly-ARTs waren die Klonierung einer zweiten, strukturell verschiedenen poly-ART aus der Pflanze *A.thaliana* (Babiychuk *et al.* 1998) sowie die überraschende Beobachtung, dass in Mausembryo-Fibroblasten von *PARP*-„knockout“-Mäusen noch poly-ART-Aktivität nachzuweisen war (Shieh *et al.* 1998). Schließlich wurden während dieser Arbeit von verschiedenen Arbeitsgruppen fünf neue humane PARP-Familienmitglieder publiziert: PARP-2 (Ame *et al.* 1999) , PARP-3 (Johansson 1999), vault

PARP (Kickhoefer *et al.* 1999), Tankyrase 1 (Smith *et al.* 1998) und zuletzt Tankyrase 2 (Lyons *et al.* 2001).

1.6.2 Eukaryontische mono-ADP-Ribosyltransferasen

In den 90er Jahren gelang aus dem Skelettmuskel von Kaninchen (RABNAART, heute ART1) und aus dem Knochenmark von Hühnern (CHAT 1 und CHAT2, heute ART6a und ART6b) schließlich auch die Aufreinigung, Sequenzierung sowie der Nachweis der Enzymaktivität der ersten endogenen mono-ARTs (Tsuchiya *et al.* 1994, Zolkiewska *et al.* 1992). Bei der Suche nach verwandten Proteinen in den Datenbanken zeigte sich eine signifikante Homologie zu RT6 (heute Ratten-ART2), einem Glykosyl-phosphatidylinositol (GPI)-verankerten, membranständigen T-Zell-Oberflächenmolekül unbekannter Funktion aus der Ratte, für das im Anschluss die ART-Aktivität nachgewiesen wurde (Haag *et al.* 1995, Koch *et al.* 1990, Thiele *et al.* 1986) .

In den letzten Jahren gelang die Identifizierung und Charakterisierung weiterer eukaryontischer mono-ARTs: Das menschliche Genom enthält vier funktionelle *ART*-Gene (*ART 1, ART 3, ART 4* und *ART 5*) und zwei verwandte, durch Introns unterbrochene Pseudogene (*ART2AP* und *ART2BP*) (Braren *et al.* 1997, Glowacki *et al.* 2002, Koch-Nolte *et al.* 1997, Okazaki and Moss 1998). Desweiteren konnten sechs funktionelle Orthologe bei der Maus sowie drei Familienmitglieder beim Huhn (*ART 6a, ART 6b, ART 7*) nachgewiesen werden (Braren *et al.* 1998, Glowacki *et al.* 2001, Glowacki *et al.* 2002, Hollmann *et al.* 1996, Okazaki and Moss 1998). Zwei humane (ART1, ART5) und vier murine (ART1, ART2.1, ART2.2, ART5) ART-Proteine sind durch das R-S-EXE-Motiv argininspezifischer ARTs charakterisiert und besitzen die vorhergesagte Enzymaktivität. Alle humanen und murinen ARTs sind Ektoenzyme, die mit Ausnahme von ART5 eine Signalsequenz für einen GPI-Anker besitzen. Im Gegensatz zu den Untersuchungen bei der PARP zeigen Expressionsanalysen, dass sich die Expression der meisten ARTs, mit starker Konservierung zwischen Mensch und Maus, auf bestimmte Gewebe beschränkt: Während *ART1* am stärksten im Skelettmuskel, *Art2* in T-Zellen und *ART3* bzw. *ART5* im Hoden exprimiert wird, wurde für *ART4* ein breiteres Expressionsprofil (u.a. Erythrozyten, Herz, Leber) nachgewiesen (Glowacki *et al.* 2002, Gubin *et al.* 2000). Mit Hilfe von Sekundärstruktur-Vorhersage-Programmen (PHDsec) konnte für die eukaryontischen mono-ARTs eine, mit der

Kristallstruktur einiger bakterieller Toxine vergleichbare Faltung prognostiziert werden (Koch-Nolte *et al.* 1996b), die sich in der vor kurzem publizierten Kristallstruktur von murinem ART2.2 bestätigte (Mueller-Dieckmann *et al.* 2002).

Es gibt zahlreiche Hinweise, dass die endogenen ARTs eine zentrale Rolle in der Regulierung bestimmter Zellfunktionen spielen. Am intensivsten wurde bisher das T-Zell-Membran-Protein RT6 (ART2) untersucht, dem eine wichtige Funktion in der Regulation des Immunsystems zugeschrieben wird. So sind z.B. die Überexpression dieses T-Zell-Differenzierungsmarkers oder Defekte in dessen Genstruktur mit dem Auftreten verschiedener Autoimmunerkrankungen wie z.B. Diabetes mellitus Typ 1 in bestimmten Ratten- und Maus-Stämmen assoziiert (Haag *et al.* 1993, Haag and Koch-Nolte 1997). Weitere Untersuchungen zeigten, dass murines ART2.2 nach Aktivierung der T-Zellen in enzymatisch aktiver Form von der T-Zell-Oberfläche abgespalten wird und dass die Inkubation von T-Zellen mit NAD^+ die T-Zell-Aktivierung blockiert und T-Zell-Apoptose induziert (Adriouch *et al.* 2001, Kahl *et al.* 2000). Da *Art2*-defiziente T-Zellen aus *Art2.1/Art2.2*-„knockout"-Mäusen für diesen Mechanismus der NAD^+-induzierten Apoptose resistent sind (Ohlrogge *et al.* 2002), ist eine entscheidene Rolle von *Art2* bei der T-Zell-Regulation zu postulieren. Desweiteren wurden ART-vermittelte Reaktionen mit der Regulation der Myogenese und der Langzeitpotenzierung in Hippocampusneuronen in Verbindung gebracht (Schuman *et al.* 1994, Zolkiewska and Moss 1993). Weiterführende Erkenntnisse über die Funktion der eukaryontischen mono-ARTs werden in der Zukunft durch die Entwicklung von neuen, spezifischen ekto-ART-Antikörpern und die Etablierung von *ART*-„knockout"-Modellen erwartet.

1.7 Gen-Datenbanken und „*in silico*"-Klonierung

Nukleotidsequenzen bilden das Fundament zum Verständnis der Struktur, Funktion und Entwicklung unterschiedlicher Organismen und sind von unschätzbarem Wert für die Identifizierung und Charakterisierung wichtiger Gene. Die Entwicklung einer effektiven DNA-Sequenzierungs-Methode durch Sanger leitete 1977 die moderne Ära der DNA-Sequenzierung ein. Nachdem die DNA-Sequenzierung in der Folge zur Routine geworden war, wurden computerisierte Sequenzdatenbanken aufgebaut, um diese Informationen zu katalogisieren und der gesamten Wissenschaft zugänglich zu machen. Die Nutzung dieser

Datenbanken zur Identifizierung und Klonierung neuer Gene bezeichnet man als „*in silico*“-Klonierung. Eine dieser Datenbanken ist die Anfang der 80er Jahre in den USA gegründete „GenBank“.

1.7.1 GenBank

Die GenBank ist eine öffentliche Datenbank aller bekannten Nukleotid- und Proteinsequenzen mit begleitenden bibliographischen und biologischen Anmerkungen, die u.a. vom „National Center of Biotechnology Information“ (NCBI) am „National Institutes of Health“ (NIH) in den USA verwaltet wird (www.ncbi.nlm.nih.gov).

Die GenBank wird vor allem durch direkte Eintragungen ermittelter Sequenzen durch die Autoren und von großen Sequenzierungsprojekten gespeist. Zudem werden im Rahmen der „International Nucleotide Sequence Database Collaboration“ täglich Daten mit zwei anderen großen Datenbanken, der „DNA DataBank of Japan“ (DDBJ) und dem „European Molecular Biology Laboratory“ (EMBL), ausgetauscht (Benson *et al.* 2003). Die GenBank wächst in den letzten Jahren exponentiell (**Abbildung 6**): Während 1982 lediglich 680.338 Basenpaare aus 606 Sequenzen eingetragen waren, enthielt die Datenbank Ende 2002 bereits über 28,5 Milliarden Basenpaare aus 22,3 Millionen Sequenzen. Ende 2004 hatte sich die Zahl der Sequenzen im Vergleich zu 2002 erneut fast verdoppelt (37,3 Millionen (Benson *et al.* 2005)). Inzwischen sind über 165000 Spezies in der GenBank vertreten. Ca. 19% aller Sequenzen sind humanen Ursprungs, 13% aller Sequenzen sind humane ESTs (siehe 1.7.2). Andere stark vertretene Spezies sind z.B. *Mus musculus*, *Rattus norvegicus, Danio rerio* und die Fruchtfliege *Drosophila melanogaster*. Einen großen Anteil der Datenbank nehmen inzwischen vollständige Genome ein. Neben den mehr als 180 kompletten mikrobiellen Genomen (u.a. humanpathogene Erreger wie *Escherichia coli, Clamydia trachomatis, Pseudomonas aeruginosa* und *Haemophilus influenzae*) wurden auch die bekannten Genome von Mensch, Maus und niederen Eukaryonten (z.B. *Drosophila melanogaster, Caenorhabditis elegans, Arabidopsis thaliana* und *Saccharomyces cerevisiae*) deponiert. Die Eintragungen der GenBank wurden traditionell nach taxonomischen Kriterien in sog. „divisions“ unterteilt (z.B. Bakterien, Viren, Nager). Um spezifische Sequenzierungs-Strategien zu unterstützen, wurden im Verlauf weitere „divisions“, u.a. für ESTs oder „high throughput genomic sequences“ (HTG), hinzugefügt (Benson *et al.* 2005).

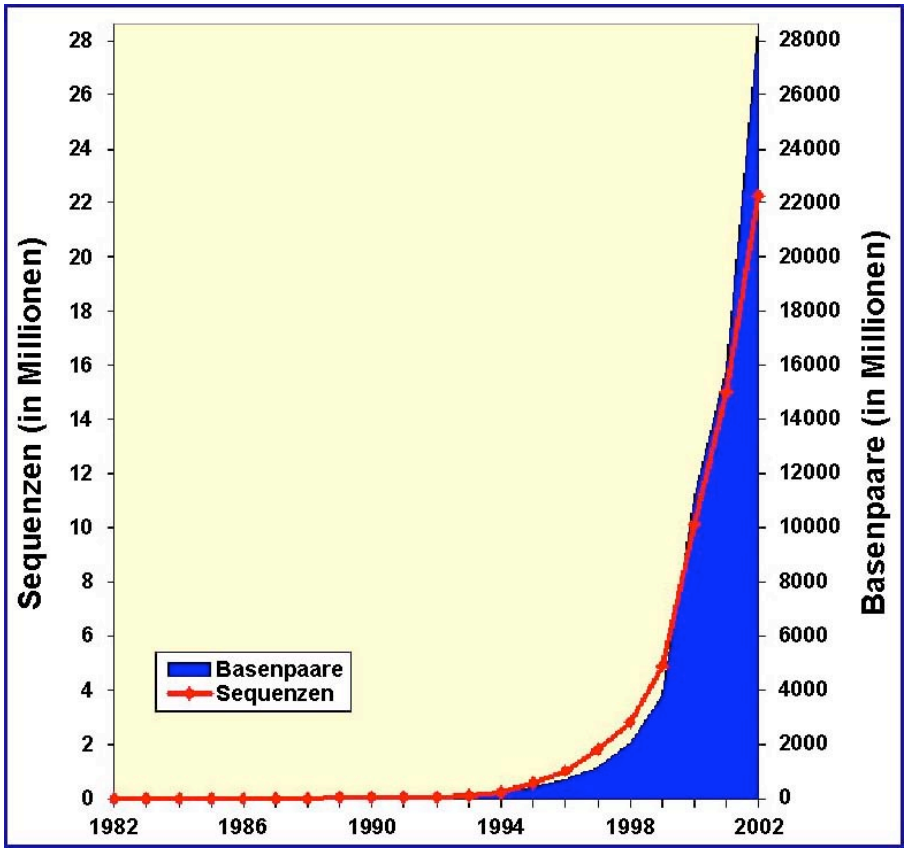


Abbildung 6: Graphische Darstellung des Wachstums der GenBank (1982-2002)
Abbildung modifiziert von der GenBank Homepage/GenBank Overview, NCBI.

1.7.2 Die EST-Projekte

Unter der Annahme, dass nur 2% des humanen Genoms aus proteinkodierenden Nukleotiden besteht, während die restliche DNA weitgehend unbekannte Funktionen besitzt (oft auch als „junk-DNA“ bezeichnet), begann 1991 die Ära des „high throughput sequencing“ von cDNAs, um zunächst die kodierenden Gene zu identifizieren. Nach randomisierter Selektion wurden cDNA-Klone mit vektorspezifischen Oligonukleotiden von der 5‘- und 3‘-Seite des Inserts aus sequenziert. Es wurden also keine genspezifischen Primer benötigt. Die erhaltenen Sequenzen, sog. „Expressed Sequence Tags“ (ESTs), sind kurz (300-500bp) und mit ca. 2% Fehlerquote relativ ungenau (Schuler 1997). Unter der Verwaltung von GenBank wurde 1992 dbEST, eine Datenbank für ESTs, aufgebaut (Boguski *et al.* 1993), in der auch die Herkunftsgewebe bzw. –zellinie der ESTs sowie mögliche Homologien zu bekannten Proteinen angegeben werden. Im August 2004 waren bereits 23,4 Millionen ESTs aus über 740 verschiedenen Organismen deponiert. ESTs sind zur Zeit die größte Quelle neuer Sequenzen in GenBank und eine kostbare Ressource, die in den letzten Jahren die „*in silico*“-Klonierung von vielen Genen ermöglichte (Benson *et al.* 2005, Boguski *et al.* 1994).

Durch die Analyse einer großen Anzahl von ESTs eines bestimmten Genes erhält man Informationen über Transkriptionsstärke, alternatives Spliceverhalten und Gewebespezifität der Genexpression. Um die Wahrscheinlichkeit der Sequenzierung von raren Transkripten zu erhöhen, wurden neue Techniken zur Erstellung von cDNA-Bibliotheken wie z.B. sog. Normalisierungsprozeduren entwickelt, die die Transkripte der hochexprimierten Gene zuvor reduzieren (Soares *et al.* 1994). Ein Nachteil dieser Techniken ist, dass die Transkriptionsstärke eines Genes anhand der Anzahl der ESTs in dbEST lediglich eingeschränkt beurteilbar ist.

1.7.3 Die „UniGene"-Datenbank

Mit dem Ziel der Entwicklung eines Kataloges transkribierter Gene wurde das „UniGene"-System erstellt, in dem die EST-Sequenzen der GenBank in genorientierte Gruppen (sog. Cluster) unterteilt werden (Schuler 1997). Mit Hilfe von Computerprogrammen werden von einem Gen stammende ESTs und längere cDNA-Sequenzen speziesspezifisch zu einem UniGene-Cluster zusammengestellt. Jedes UniGene-Cluster enthält somit die Sequenzen eines einzigen Genes sowie zusätzliche Informationen wie z.B. Gewebeexpression, chromosomale Lokalisierung und putative verwandte Proteine. Die UniGene-Datenbank hat die Zuordnung von Sequenzen zu bestimmten Genen und damit die Identifizierung neuer Gene wesentlich vereinfacht. Inzwischen sind UniGene-Cluster für mehr als 45 Spezies angelegt (Stand: 07/2004), darunter die Säuger Mensch, Maus, Ratte, Kuh und Hausschwein, niedere Eukaryonten wie die Fruchtfliege und der Schleimpilz *Dictyostelium discoideum* sowie Pflanzen wie *Arabidopsis thaliana*. In der humanen UniGene-Datenbank konnten die ESTs in ca. 107.000 Cluster unterteilt werden (Wheeler *et al.* 2005). Da das humane Genom nach Schätzungen ca. 30 - 40.000 Gene enthält, sind wahrscheinlich einzelne Gene durch mehrere Cluster repräsentiert, die in der Zukunft zu einem Cluster zusammengefasst werden können.

Über die Identifizierung konservierter Strukturen durch Sequenzvergleiche zwischen verschiedenen Arten lassen sich besondere Informationen bzgl. einzelner Gene und mögliche funktionelle Verwandtschaften ableiten. Das „HomoloGene"-System ordnet aus den UniGene-Clustern der verschiedenen Spezies orthologe und homologe Gene zu, so dass über dieses System die Möglichkeit besteht, zu einem neu klonierten Gen schnell und effektiv Homologe in den anderen Spezies zu finden (Wheeler *et al.* 2005).

1.7.4 Programme für Datenbankrecherchen

Die effektive Nutzung der Sequenz-Ressourcen in der GenBank ermöglichen verschiedene Computerprogramme. Das integrierte Datenbank-Abfrage-System „Entrez“ von NCBI ermöglicht u.a. den Zugriff auf DNA- und Protein-Sequenzen, Genomkarten, Daten zur Gen-Expression, Protein-Strukturen aus der „Molecular Modeling Database“ (MMDB) und *via* PubMed auch auf Fachliteratur. Für die Recherche nach Sequenzhomologien zu einer eingegebenen Sequenz stehen mit den sog. BLAST („Basic Local Alignment Search Tool“)-Programmen eine ganze Reihe an Suchprogrammen zur Verfügung (Altschul *et al.* 1990). Als Ergebnis der BLAST-Suche erhält man eine Auflistung von Alignments zwischen der Eingabesequenz und ähnlichen Sequenzen (sog. „hits“) in der Datenbank. Vorteile der BLAST-Programme sind die hohe Geschwindigkeit bei minimalem Verlust an Sensivität sowie ein statistisches Fundament, das die Voraussetzungen schafft, um reale Treffer von falschen „hits“ zu unterscheiden. Je nach Zielsetzung können mit den BLAST-Programmen verschiedene Datenbanken durchsucht werden: Neben der nr („non redundant“)-Datenbank, die alle GenBank-Sequenzen enthält, stehen z.B. Nukleotid-Datenbanken wie dbEST, Protein-Datenbanken wie „swissprot“ oder Datenbanken mit genomischen Sequenzen für Analysen zur Verfügung. Die BLAST-Familie wird fortlaufend um Programme mit neuen oder verbesserten Funktionen erweitert. Für die Datenbankanalysen dieser Arbeit wurden folgende BLAST-Programme verwendet:

- **BLASTn** vergleicht eingegebene Nukleotid-Sequenzen mit einer Nukleotid-Sequenz-Datenbank. Durch eine Recherche in dbEST kann so die Anzahl der zugehörigen ESTs ermittelt werden. Die Suche mit einer cDNA gegen Datenbanken mit genomischen Sequenzen erbringt Aufschlüsse über die Exon-Intron-Struktur, da das Alignment im Bereich der Introns abbricht.
- Mit **tBLASTn** werden ebenfalls Nukleotid-Datenbanken durchsucht. Allerdings übersetzt das Programm die offenen Leseraster von Nukleotid-Sequenzen (z.B. aus dbEST) in eine Aminosäuresequenz, so dass mit einer eingegebenen Aminosäuresequenz nach Homologien gesucht werden kann. Da sich z.B. durch die Übersetzung von ESTs die Anzahl von „Proteinsequenzen“ gegenüber Proteindatenbanken deutlich erhöht, ermöglicht diese Methode die Identifizierung zuvor unbekannter, naher Familienmitglieder einer Proteinfamilie, für die bisher keine bekannten Proteinsequenzen in der Datenbank deponiert wurden.

- Durch Nutzung von **BLASTp** wird eine Aminosäuresequenz mit einer Protein-Datenbank verglichen. Eine Weiterentwicklung von BLASTp ist das **p**ositions**s**pezifische, **i**terative Suchprogramm **PSI-BLAST**: Die initialen Verknüpfungen bei der PSI-BLAST-Suche entsprechen den Proteinen, die eine signifikante Sequenzidentität zum jeweils eingegebenen Protein aufweisen und die auch mit dem traditionellen BLASTp-Programm ermittelt werden können. Anschließend erstellt dieses Programm aus den im ersten Suchdurchlauf erhaltenen Aminosäuresequenzen verwandter Proteine ein multiples Alignment und generiert aus diesen Informationen eine neue Matrix, die die positionsspezifischen Sequenzhomologien berücksichtigt. Dies bedeutet, dass bei der Erstellung der Matrix Positionen mit konservierten Aminosäureresten stärker gewichtet werden als Positionen mit nicht-konservierten Aminosäuren. Mit der neu generierten Matrix wird die Suche anschließend fortgesetzt und so lange wiederholt (Iterationen), bis sie keine neuen Ergebnisse erbringt und somit konvergent ist (Altschul *et al.* 1997). PSI-BLAST ermöglicht damit die Aufspürung relativ geringer, jedoch biologisch relevanter Sequenzhomologien in entfernt verwandten Proteinen (Altschul and Koonin 1998).

1.7.5 Genomprojekte

Anfang der 80er Jahre wurde erstmals ein Projekt mit dem Ziel der Entschlüsselung der kompletten Nukleotid-Sequenz des menschlichen Genoms vorgeschlagen. Aufgrund der geringeren Nukleotidanzahl wurden zunächst wenige bakterielle und virale Genome, wie z.B. das 49kbp umfassende Genom des lamda-Phagen, vollständig sequenziert. Mit der Entwicklung neuer Technologien wie der automatischen DNA-Sequenzierung nahmen Forschergruppen schließlich erste eukaryontische Genome in Angriff. 1990 wurde in den USA unter der Führung des „National Institutes of Health“ und des „U.S. Department of Energy“ das Humane Genom-Projekt (HUGO) initiiert. Zudem wurden in Zusammenarbeit verschiedener Forschungseinrichtungen gezielt eukaryontische Organismen zur Sequenzierung ausgewählt, die als Modelle in der Forschung verwendet werden. Im Jahr 1997 wurde als erstes eukaryontisches Genom das Genom der Hefe *Saccharomyces cerevisiae* in „Nature“ veröffentlicht (Mewes *et al.* 1997). Ein Jahr später war das 97 Megabasen umfassende Genom von *C.elegans* bis auf wenige Lücken („gaps“) vollständig sequenziert. Als Test-Spezies für das komplette Zusammensetzen eines komplexen eukaryontischen Genoms wurde schließlich die Fruchtfliege *Drosophila melanogaster* ausgewählt. Die Sequenzierung

des euchromatischen Anteils des Genoms konnte innerhalb eines Jahres abgeschlossen werden. Das ca. 120 Megabasen große Genom enthält ca. 13600 Gene (Adams *et al.* 2000). Ebenfalls im Jahr 2000 wurde das Genom der Pflanze *Arabidopsis thaliana* publiziert (Arabidopsis Genome Initiative 2000).

Als bahnbrechenden Höhepunkt in der Genomforschung stellten im Februar 2001 C.Venter und Kollegen sowie das „International Human Genome Sequencing Consortium“ zeitgleich die erste, 2,91 Milliarden Basenpaare umfassende Konsensussequenz des euchromatischen Anteils des humanen Genoms vor. Nach konservativen Kriterien konnten 26383 Gene und nach Computeranalysen ca. 12000 weitere mögliche Gene vorhergesagt werden, so dass von einer maximalen Anzahl von ca. 40.000 Genen auszugehen ist. Lediglich 1,1% des Genoms sind Exons. Erste vergleichende Genomanalysen weisen daraufhin, dass es bei den Vertebraten gegenüber niederen Eukaryonten zu einer Expansion von Genen gekommen ist, die mit neuronalen Funktionen, der Regulation der Gewebedifferenzierung oder dem Immunsystem assoziiert sind (Lander *et al.* 2001, Venter *et al.* 2001).

Einen weiteren Meilenstein stellte die Publikation des Maus-Genoms im Dezember 2002 dar (Waterston *et al.* 2002), da die Maus einerseits ein intensiv erforschtes Tiermodell ist und andererseits nun erstmals ein Vergleich von zwei verschiedenen Säugetiergenomen möglich war. Vielerseits unerwartet zeigte sich zum Beispiel, dass das Mausgenom ca.14% kleiner ist als das humane Genom.

Inzwischen sind mehr als 90 mikrobielle Genome und weitere eukaryontische Genome komplettiert und über Datenbanken als Ressource verfügbar, so dass erweiterte Möglichkeiten für die „*in silico*“-Klonierung und Charakterisierung von Genen bestehen. Für viele weitere Organismen (u.a. Schimpanse, Hund, Kuh) laufen zur Zeit Genomprojekte, so dass in den nächsten Jahren die Zahl der komplettierten Genome deutlich zunehmen wird.

Die für das „Puzzle des Lebens“ benötigten Puzzleteile einiger Organismen sind aufgeklärt. Die große Aufgabe der Zukunft ist es nun, diese Teile richtig zusammenzufügen, die Funktion der kodierten Proteine zu erforschen (Zeitalter der „Proteomics“) und über den Vergleich der Puzzles verschiedener Organismen neue evolutionsgeschichtliche Erkenntnisse zu erlangen.

1.8 Das pET-System zur Expression rekombinanter Proteine in *E.Coli*

Um rekombinante Proteine in Bakterienstämmen effizient zu exprimieren, benötigt man spezielle Expressionvektoren. Da gezeigt werden konnte, dass viele Fremdproteine toxisch sind und somit das Wachstum des Bakterienwirts sowie die Etablierung des Plasmids hemmen, sind an die Konstruktion dieser Vektoren besondere Anforderungen zu stellen: Sie müssen einerseits zunächst eine ungestörte Vermehrung der transformierten Bakterien und andererseits im Anschluss an die Vermehrung eine effiziente Transkription der Fremd-DNA und Produktion des rekombinanten Proteins ermöglichen. Um dies zu gewährleisten, benötigen die Vektoren einen regulierten Promotor. Das klassische Beispiel eines *E.Coli*-Promoters unter Kontrolle des *lac*-Operators ist in **Abbildung 7** dargestellt. Für *E.Coli*-Promotoren wurde aber gezeigt, dass sie in Abwesenheit des Induktors ein gewisses Level an basaler Transkription der Fremd-DNA ermöglichen und somit bei toxischen Produkten die stabile Klonierung verhindert wird. Deshalb besitzt das in dieser Arbeit verwendete, initial von Studier und Kollegen entwickelte pET-System von Novagen den effektiven Promotor des Bakteriophagen T7, der durch die RNA-Polymerase von *E.Coli* nicht erkannt wird (Studier and Moffatt 1986). Die Transkription der Fremd-DNA setzt erst dann ein, wenn T7-RNA-Polymerase zur Verfügung gestellt wird. Dazu werden in dieser Arbeit λDE3-lysogene *E.Coli*-Stämme verwendet, die ein T7-RNA-Polymerase-Gen besitzen, das unter Kontrolle der *lacUV5*-Promoters steht und dessen Expression somit, analog zum Beispiel in der Abbildung 7, über die Zugabe von IPTG induziert werden kann.

Das pET-System beinhaltet eine Vielzahl von Vektoren und *E.Coli*-Stämmen, um eine für das jeweilige Genprodukt optimale Klonierung und Expression zu gewährleisten sowie die anschließende Detektion und Aufreinigung zu ermöglichen (Mierendorf 1994). Um die Basalexpression der Fremd-DNA im nicht-induzierten Status weiter zu minimieren, wurden u.a. pLysS-Stämme verwendet, die ein T7-Lysozym kodierendes Plasmid enthalten. T7-Lysozym ist ein natürlicher Inhibitor der T7-RNA-Polymerase. Desweiteren sind die für die Expression verwendeten BL21-Stämme defizient für *lon* und *ompT*-Proteasen, um die Proteolyse des rekombinanten Proteins zu verringern. Die Selektionierung der transformierten Bakterien wird wie bei klassischen Vektoren durch ein Antibiotika-Resistenz-Gen gewährleistet. Das pET-System ist eines der effektivsten Systeme zur Klonierung und Expression von rekombinanten Proteinen in *E.Coli*.

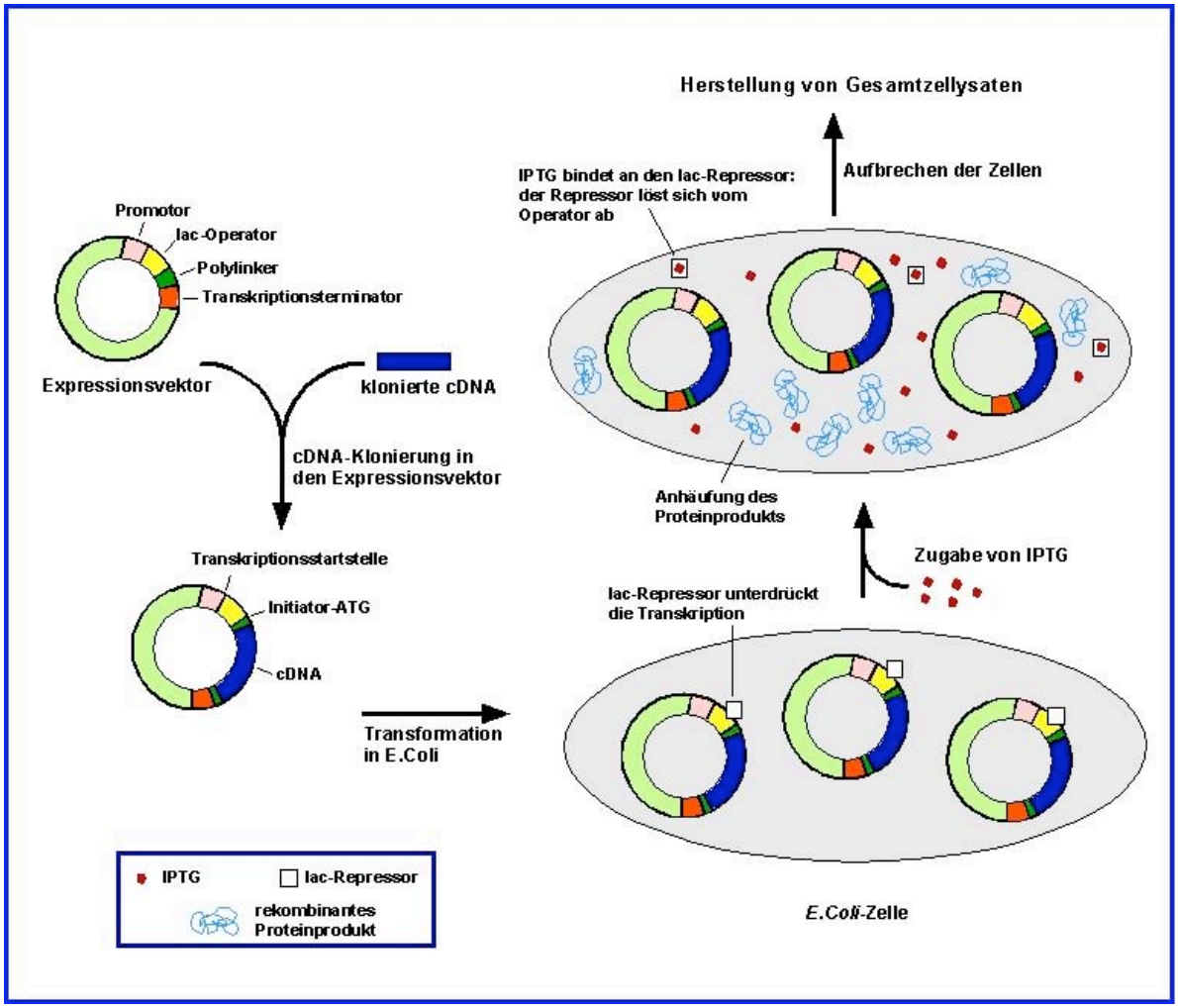


Abbildung 7: Expression von rekombinanten Proteinen in *E.Coli* mittels eines regulierten Promoters
Der gewählte Expressionsvektor besitzt einen *E.Coli*-Promotor, der unter Kontrolle des *lac*-Operators steht. Die klonierte cDNA wird mit Hilfe von Restriktionsenzymen in der Polylinker-Region in gewünschter Orientierung eingebaut. Mit den Plasmiden wird ein *E.Coli*-Stamm transformiert, der den lac-Repressor überexprimiert. Der Repressor heftet sich an eine Operatorsequenz, die dicht am Promotor des zu regulierenden Operons liegt. Durch die Bindung des Repressors wird die RNA-Polymerase daran gehindert, sich an den Promotor zu binden. Während der Wachstumsphase bleibt somit die Transkription der cDNA stark reprimiert, so dass ein möglicherweise toxisches rekombinantes Protein die Vermehrung nicht behindert. Sobald eine ausreichende Menge an Bakterienkultur angezüchtet wurde, wird Isopropylthiogalactosid (IPTG) zugesetzt, um den Repressor zu inaktivieren. Somit kann eine effiziente Transkription der Fremd-DNA sowie die anschließende Bildung und Anreicherung des rekombinanten Proteins stattfinden. Die Kontrolle eines Promotors durch einen Repressor ist ein klassisches Beispiel für eine negative Regulation (Die Abbildung wurde modifiziert aus: Rekombinierte DNA, James D. Watson, 1993 (Watson *et al.* 1993)).

2 Ziele

Die vorliegende Arbeit hatte folgende Zielsetzung:

1) Identifizierung neuer homologer Gene aus den Familien der poly-ADP-Ribosyl-Transferasen und der mono-ADP-Ribosyl-Transferasen in den Datenbanken

2) „In silico“-Charakterisierung der Gen- und Domänenstrukturen der humanen poly-ADP-Ribosyl-Transferasen und ausgewählter poly-ADP-Ribosyl-Transferasen aus anderen Spezies

3) Chromosomale Lokalisierung der humanen und murinen poly-ADP-Ribosyl-Transferasen

4) Klonierung, molekulare Charakterisierung und rekombinante Expression ausgewählter neu identifizierter Familienmitglieder

5) Untersuchung der vorhergesagten ADP-Ribosyltransferaseaktivität

3 Materialien und Methoden

3.1 Software und Datenbanken

3.1.1 Identifizierung neuer mono- und poly-ARTs

Wir benutzten die BLAST-Programme tBLASTn, BLASTn und PSI-BLAST, um die "Expressed Sequence Tags" (EST)- und "non redundant" (nr)-Sektionen der Datenbank GenBank des "National Center of Biotechnology Information" (NCBI, http://www.ncbi.nlm.nih.gov) auf homologe Sequenzen der mono-ARTs und poly-ARTs zu durchsuchen (Altschul *et al.* 1990, Altschul *et al.* 1997, Wheeler *et al.* 2005). Die „Unigene"-Datenbank und das „Homologene"-Projekt auf der NCBI-Homepage wurden verwendet, um zugehörige EST-Cluster zu identifizieren und Homologe in verschiedenen Spezies zuzuordnen.

3.1.2 Genstrukturen und chromosomale Lokalisierung

Um die genomischen Sequenzen für die humanen und murinen pARTs zu ermitteln, wurden die „human genome resources"- und „mouse genome resources"-Datenbanken mit tBLASTn bzw. BLASTn und den Aminosäure- bzw. mRNA- oder cDNA-Nukleotidsequenzen der jeweiligen pART als Eingabe durchsucht. Aus den Alignments der mRNA- bzw. cDNA-Sequenzen und der genomischen Sequenzen konnten die Genstrukturen aufgeschlüsselt werden, da es im Bereich von Introns zu Unterbrechungen des Alignments kommt. Die „Genome view"-Funktion ermöglichte die chromosomale Lokalisierung der Gene. Mit Hilfe des „UCSC genome browser"-Links wurden die Gene der humanen und murinen pARTs sowie die entsprechenden chromosomalen Regionen genauer analysiert (Karolchik *et al.* 2003). Dieser "Browser" ermöglicht die Darstellung der Lage von bekannten Genen (basierend auf Swiss-Prot, TrEmbl, mRNA und Ref Seq) zusammen mit zugehörigen mRNAs und ESTs für einen ausgewählten chromosomalen Bereich im Genom von Maus oder Mensch. Desweiteren ist ein vergrößern bzw. verkleinern („zoomen") des betrachteten Abschnitts sowie ein seitliches „wandern" auf dem Chromosom möglich, um entweder genauere Informationen über einzelne Gene oder über die Nachbarschaft der untersuchten Gene zu erhalten. Zur Ermittlung der Genstrukturen von Familienmitgliedern bei der Fliege *Drosophila melanogaster*, dem Nematoden *Caenorhabditis elegans* und der Pflanze *Arabidopsis thaliana* wurden zunächst genomische Teilsequenzen analysiert, aus denen die kodierenden Regionen mit dem Computerprogramm „Genefinder" oder anderen Vorhersageprogrammen festgelegt

worden waren. Nach Fertigstellung der Genomprojekte dieser Spezies und Verfügbarkeit der entsprechenden Datenbanken im Internet wurden BLASTn-Suchen gegen die genomischen Datenbanken durchgeführt, mit denen die Genstruktur für die meisten pARTs bestätigt werden konnte. Lediglich die Gene der pARTs von *C.elegans* scheinen noch in Lücken des Genomprojektes zu liegen, so dass sie weitgehend auf einer computergestützten Vorhersage beruhen.

3.1.3 Domänenstrukturen

Um mit der pART-Domäne fusionierte Domänen und Motive in den pART-verwandten Proteinen aufzuklären, wurde zunächst die „nr"-Datenbank mit BLASTp und den N-terminal der pART-Domäne gelegenen Aminosäuresequenzen als Eingabe durchsucht. So wurden Homologien dieser Bereiche zu anderen, zum Teil bereits charakterisierten Proteinen ermittelt, aus denen sich die Konservierung einzelner Domänen ableiten ließ. Desweiteren wurde mit Hilfe der „conserved domain search"-Funktion in der „conserved domain database (CDD)" gezielt nach charakterisierten konservierten Bereichen in den pART-Proteinen gesucht (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=cdd) (Marchler-Bauer *et al.* 2005). Die meisten der Domänenmodelle in der CDD sind Importe aus den Datenbanken Pfam, COG und SMART. Mit Hilfe von CDART (Conserved Domain Architecture Retrieval Tool) gelang zudem die Identifikation von Proteinen verschiedener Spezies, die ähnliche Domänen oder Motive besitzen (Geer *et al.* 2002).

3.1.4 Analyse der Sequenzen und Erstellung der Alignments

Zur Auswertung der DNA- und Proteinsequenzen wurden die ermittelten Sequenzen auf Macintosh Computern mit den Analyse-Programmen DNAStar und Sequencher bearbeitet. Die Aminosäure-Alignments wurden mit der Megalign-Funktion des Programms DNAStar erstellt und mit Hilfe von T-Coffee- und PSI-BLAST-Alignments manuell optimiert. T-Coffee ist ein von der EMBnet-Gruppe im Internet zur Verfügung gestelltes Alignment-Programm (www.ch.embnet.org/software/TCoffee), das besondere Qualitäten beim Alignment von Sequenzen mit niedrigerer Sequenzidentität besitzt (Notredame *et al.* 2000). Zur weiteren Optimierung der Alignments wurden für verschiedene pART-Proteine Sekundärstrukturvorhersagen mit PSIPRED erstellt (http://bioinf.cs.ucl.ac.uk/psipred/). PSIPRED ermöglicht die verlässliche Vorhersage von Sekundärstrukturmotiven aus einer

eingegebenen Aminosäuresequenz (McGuffin *et al.* 2000). GeneTHREADER erkennt die Faltung individueller Proteinsequenzen und erstellt Sequenz-Struktur-Alignments zu bekannten Kristallstrukturen in der Datenbank (McGuffin and Jones 2003). Für alle humanen pARTs wurde von GeneTHREADER eine Alignment mit der Kristallstruktur der Hühner-PARP (1a26) erstellt. Diese wurden in den Alignments der pART-Subgruppen (siehe Anhang D) ebenfalls berücksichtigt.

3.1.5 3D-Strukturen

Die 3D-Modelle von VIP2, ART2, DT und Hühner-PARP (siehe Abb.5) wurden mit dem Programm PyMOL erstellt und bearbeitet (DeLano 2002).

3.1.6 Erstellung der Dissertationsarbeit

Zur Erstellung der Abbildungen wurden die Grafikprogramme Canvas (Versionen 3.5.3, 6.0, 8.0) und Photoshop 5.5 verwendet. Die Textverarbeitung sowie die Erstellung von Tabellen erfolgte mit den Microsoft Office Programmen Word und Excel (Versionen 1998 und 2000).

3.2 Kreation von spezifischen Oligonukleotiden (Primern)

Aus den durch die Datenbankrecherchen ermittelten EST-Clustern für hpART3 und hpART4 wurden durch Alignments vorläufige „full length“ cDNA-Konsensussequenzen für diese beiden pARTs erstellt. Für die hpART1 und das SpvB wurden die publizierten Nukleotidsequenzen („accession numbers“: hpART1 J03473, SpvB Z11557) aus der Datenbank verwendet. Zur Amplifikation der cDNA-Bereiche, die für die katalytische Domäne von pART3, pART4, pART1 und SpvB kodieren, bzw. zur Amplifikation der kompletten cDNA von pART3 wurden spezifische Primer konstruiert, in die Restriktionsenzym-Schnittstellen passend zu den späteren Klonierungsstrategien eingebaut wurden: Die Primer F61, F63, P02, K01 und SP2 enthalten Schnittstellen für das Restriktionsenzym Nde I (5´ ...C A^T A T G ... 3´), die Oligonukleotide K02 und SP1 Schnittstellen für Nco I (5´ ... C^C A T G G ... 3´) und die Primer F85, P11, K11 und SP4 Schnittstellen für die Restriktions-Endonuklease Xho I (5´ ... C^T C G A G ... 3´). Zudem wurden weitere pART3-spezifische Primer erstellt (F62, F64, F82, F86, F87), die u.a. für die Amplifikation von pART3-Fragmenten aus cDNA-Templates bzw. genomischen Templates zur Aufklärung der genomischen Struktur und zur Sequenzierung verwendet wurden. Die

Oligonukleotide wurden von uns entworfen und freundlicherweise im Service Labor des Instituts für Zellbiochemie und Klinische Neurobiologie des Universitätskrankenhauses Hamburg-Eppendorf synthetisiert. Die vollständigen Primersequenzen sind im Anhang A, die cDNA-Konsensussequenzen im Anhang B dargestellt. Die Lage der Primer ist in die cDNA-Konsensussequenzen eingezeichnet. Für die DNA-Sequenzierung wurden die genannten Gen-spezifischen Primer oder die „Vektorprimer“ TAF und TAR (pCR2.1) bzw. „T7 promotor“- und „T7 reverse“ (pET-Vektoren) verwendet. Ihre Bindungstellen in der Vektorsequenz ist aus den Vektorkarten ersichtlich (siehe Anhang C).

3.3 PCR („polymerase chain reaction“, Polymerase-Ketten-Reaktionen)

3.3.1 MTC („Multiple tissue cDNA“)-PCR

Die PCR wurde in 20µl Ansätzen mit dem Advantage-Polymerase-Mix (Clontech) durchgeführt. Der Mix enthält die Taq-Polymerase, kleinere Mengen einer kontrolllesenden („proof-reading“) Polymerase und den TaqStart-Antikörper für den sog. „Hotstart“. Der Taq-Start-Antikörper ist ein neutralisierender monoklonaler Antikörper, der die Taq-Polymerase spezifisch bindet und damit die Polymeraseaktivität während des Ansetzens der PCR-Reaktionen blockiert, so dass eine Synthese ausgehend von unspezifisch gebundenen Primern vermieden wird. Durch die Erhitzung zur Initiierung der PCR wird das Enzym reaktiviert. Somit kann eine spezifische und effiziente Amplifikation erfolgen. Mit Hilfe dieser Methode können unspezifische Amplifikationsprodukte sowie Primerdimere deutlich reduziert werden. Von der verwendeten cDNA-Matrize verschiedener menschlicher Gewebe (fetale Niere, fetale Milz, fetale Lunge, fetale Leber, adulte Milz, Leukozyten) aus dem MTC fetal panel #2 bzw. MTC immune panel (Clontech) wurde je 0,2µg cDNA pro Ansatz verwendet. Es wurden je 75ng spezifische Primer eingesetzt (Für hpART3: F61 x F82, F61 x F85 („full length cDNA“), F63xF85 (CD). Für hpART4: K01xK11, K02xK11 (CD). Für hpART1: P02xP11 (CD), wobei einige dieser Primer die unter 3.2 beschriebenen Restriktionsenzym-Schnittstellen enthalten). Zur Amplifikation wurde folgendes Touchdown-Programm verwendet: Zu Beginn wurde die DNA 60sec bei 94°C denaturiert. Die initiale „Annealing“-Temperatur wurde nach 4 Zyklen reduziert (4x 70°C, 4x 65°C), anschließend wurde eine konstante „Annealing“-Temperatur von 60°C gewählt. Die Extensionstemperatur betrug 68°C. Nach 29, 35 und 41 Zyklen wurden Aliquots entnommen und gelelektrophoretisch mit einem 1.0%igen TAE-Agarose-Gel (DNA Typing Grade

Agarose / DNA Typing Grade TAE Buffer, Gibco BRL) aufgetrennt. Das Agarosegel wurde anschließend mit Ethidiumbromid (Sigma) gefärbt und analysiert.

3.3.2 PCR-Amplifizierung des SpvB

Die Arbeiten an dem Salmonellentoxin SpvB erfolgten in Kooperation mit der Arbeitsgruppe von M.Rhen am Karolinska Institut in Stockholm, Schweden. Die Region, die für die ART-verwandte, katalytische Domäne des SpvB kodiert, wurde mit den Primerpaaren SP1 (Nco I-Schnittstelle) und SP4 (Xho I-Schnittstelle) bzw. SP2 (Nde I-Schnittstelle) und SP4 aus einem SpvB-cDNA enthaltenden Plasmid (Taira and Rhen 1989, freundlicherweise zur Verfügung gestellt von M.Rhen) amplifiziert. Für die 50µl Ansätze wurden je 2µl 1:10 verdünnte Plasmid-DNA und je 200ng Primer eingesetzt. Die Amplifikation erfolgte mit der Expand-Polymerase (Boehringer). Als PCR-Programm wurde das bereits unter 3.3.1 beschriebene Touchdown-Programm verwendet. Nach 29, 35 und 41 Zyklen wurden Aliquots entnommen, mittels Gelelektrophorese aufgetrennt und die PCR-Produkte mit Ethidiumbromid markiert. Die Einzelbanden der Aliquots nach 29 Zyklen wurden aus dem Gel ausgeschnitten, mit dem „Qiaquick PCR Purification Kit“ (Qiagen) nach Herstellerangaben aufgereinigt und in 50µl Elutionspuffer („elution buffer“, Qiagen) eluiert.

3.3.3 PCR-Mutagenese

Zur Erzeugung von Mutationen in den verschiedenen SpvB-Expressionskonstrukten benutzten wir eine PCR-Methode. Die *in-vitro*-Mutagenesen wurden mit Hilfe des „QuickChange™ Site-directed Mutagenese Kits“ (Stratagene) nach Anleitung durchgeführt. Diese Methode erlaubt die gezielte Mutagenese bestimmter Nukleotide in doppelsträngigen Plasmiden ohne weitere Subklonierungsschritte. Die Amplifikation der DNA-Stränge erfolgt mit der "proof-reading" Polymerase Pfu ausgehend von zwei "mutierten" Primern, die jeweils zu einem Strang des Plasmids komplementär sind. Das PCR-Programm umfasst 16 PCR-Zyklen. Die Extensionszeit wird mit 2min pro Kilobase DNA kalkuliert. Für die Selektion der mutierten Plasmidstränge macht man sich zunutze, dass die DNA, die aus *E.Coli*-Stämmen isoliert wird, bei nahezu allen *E.Coli*-Stämmen methyliert ist. Die Proben werden deshalb im Anschluss an die PCR mit der Endonuklease Dpn I (20.000U/ml, NEB) verdaut, die spezifisch methylierte DNA und somit dass „Mutter“-Plasmid ohne Mutationen zerschneidet, während das neu synthetisierte Plasmid mit den eingebauten, gezielten

Mutationen unbeschädigt bleibt. Die geringe Menge an eingesetztem DNA-Template, die Genauigkeit der Pfu-Polymerase und die geringe Anzahl an PCR-Zyklen führen einerseits zu einer hohen Effizienz der gezielten Mutationen und andererseits zu einer minimalen Rate unerwünschter Mutationen. Nach dem Verdau wurden die mutierten Plasmide mittels „Heat shock“ in „Solo Pack Gold supercompetent cells“ (Stratagene) transformiert (nach Anleitung des Herstellers) und auf Carbenicillin-Agarplatten ausplattiert. Statt der empfohlenen 50µl Zellsuspension pro Ansatz wurden für die Transformation lediglich 25µl eingesetzt. Nach Picken der Kolonien, Anlage von ü.N.-Kulturen und Miniprep wurden die Plasmide sequenziert.

3.3.4 PCR zum Vergleich von cDNA- und genomischen DNA-Fragmenten von humanem pART3

Um Informationen über die Genstruktur von humanem *pART3* zu erhalten, wurden mit verschiedenen spezifischen Primerpaaren Fragmente aus cDNA- und genomischen DNA-Templates amplifiziert. Als cDNA-Template wurde ein pCR2.1-Plasmid (Miniprep, 1:10 verdünnt), das als Insert die komplette cDNA von pART3 enthält, und als genomisches Template der PAC-Klon M06108 des Deutschen Ressourcenzentrums für Genomforschung GmbH (RZPD) eingesetzt (Maxiprep, 1:50 verdünnt). Als Polymerase wurde die Expand-Polymerase (Boehringer) mit „proof-reading“ Funktion verwendet. Das gewählte PCR-Programm umfasste insgesamt 34 Zyklen: Zu Beginn wurde die DNA 2min bei 94°C in Einzelstränge aufgetrennt. Die initiale „Annealing“-Temperatur wurde im Verlauf reduziert (2 Zyklen 65°C, 2x 60°C, 30x 55°C). Die Extensionstemperatur lag bei 68°C. Anschließend wurden die PCR-Produkte auf einem 0.8%igen TAE-Agarosegel aufgetrennt und mit Ethidiumbromid markiert. Die Größenunterschiede der Fragmente, die mit identischen Primernpaaren aus den beiden Templates amplifiziert worden waren, wiesen bereits auf die ungefähre Lage und Länge der Intronsequenzen hin. Die Banden mit den genomischen Fragmenten wurden aus dem Gel ausgeschnitten und mit dem „Qiaquick PCR Purification Kit“ (Qiagen) nach Herstellerangaben aufgereinigt. Die Fragmente wurden wie unter 3.5.1.1 beschrieben mit dem „TA^{R} Cloning Kit“ (Invitrogen) in den pCR2.1-Vektor kloniert, in INV-Zellen (Invitrogen) transformiert und etabliert. Nach Aufreinigung der Plasmide und Kontroll-Verdau mit dem Restriktionsenzym EcoR I (NEB, 20000U/ml) wurden die Inserts ausgehend von verschiedenen Primern sequenziert (siehe 3.10). Aus dem Vergleich der genomischen

Sequenzen und der cDNA-Konsensussequenz für pART3 konnte die Anzahl und Größe der kodierenden Exons sowie die Lage und Größe der Introneinschübe bestimmt werden.

3.3.5 Chromosomen-„Mapping“ des humanem *pART3*

Somatische Mensch/Nager-Zell-Hybride aus dem „NIGMS mapping panel #2“ (The Corriel Institute, Camden, NJ, USA) wurden mit den *pART3*-spezifischen Primern F65 und F85 (je 100ng/Ansatz) per PCR gescreent. Es wurden 250ng genomische DNA pro Ansatz eingesetzt, als Polymerase wurde AmpliTaq-Gold (Perkin Elmer) verwendet. Das PCR-Programm umfasste 34 Zyklen: Die initiale „Annealing“-Temperatur (65°C) wurde nach zwei Zyklen reduziert (2x60°C) und für die weiteren 30 Zyklen auf 55°C eingestellt. Die Extensionstemperatur betrug 72°C. Die Reaktionsprodukte wurden durch Gelelektrophorese aufgetrennt (siehe 3.3.1) und mit Ethidiumbromid (Sigma) angefärbt.

3.4 Northern Blot

Eine *pART3*-spezifische cDNA-Sonde (F62 x F82, ca. 600bp) sowie ein *pART4*-spezifische cDNA-Sonde (K02 x K11, ca. 1000bp) wurden mit 50µCi [^{32}P]-dCTP unter Verwendung des „rediprime™ labeling kits“ (Amersham Pharmacia) nach Herstellerangaben radioaktiv markiert. Zur Aufreinigung der Proben wurden die nicht-inkorporierten Nukleotide über „MicroSpin™“-Säulen (Amersham Pharmacia) entfernt. Die verwendeten Northern-Blot-Filter (Clontech) enthalten 2µg poly(A)+ aus den angegebenen Geweben pro Spur. Die Hybridisierung und das Waschen fanden unter Bedingungen geringer Stringenz statt (Hybridisierung: 16h in 7% SDS, 0.5M NaPhosphat, 1mM EDTA bei 62°C. Waschen: 3 x 15min in 1% SDS, 40mM NaPhosphat, 100mM NaCl bei 60°C). Die gebundenen Sonden wurden autoradiographisch mit 2 bzw. 7 Tage bei -80°C aufgelegten X-Omat-AR Filmen (Kodak) detektiert. Im Falle eines starken Hintergrunds wurden die Blots anschließend unter hoch stringenten Bedingungen gewaschen (2 x 15min in 1% SDS, 40mM Natriumphosphat bei 70°C). Danach erfolgte erneut eine Exposition zum autoradiographischen Nachweis. Die Blots wurden in 0.5% SDS 6min lang gekocht, um die gebundenen Sonden wieder zu lösen („stripping“). Zur Überprüfung der Qualität und Quantität der RNA wurde nach dem „strippen“ eine Kontrollhybridisierung mit einer radioaktiv markierten G3PDH (Glycerolaldehyd-3-Phosphat-Dehydrogenase)-Sonde durchgeführt (nicht gezeigt).

3.5 Klonierungsstrategien

Für die Klonierung der Mitglieder der pART-Familie, hpART1, hpART3 und hpART4, und die Klonierung des Salmonellentoxins SpvB wurden zwei unterschiedliche Klonierungsstrategien verwendet. Die Klonierungsstrategien sind schematisch in **Abbildung 8** dargestellt.

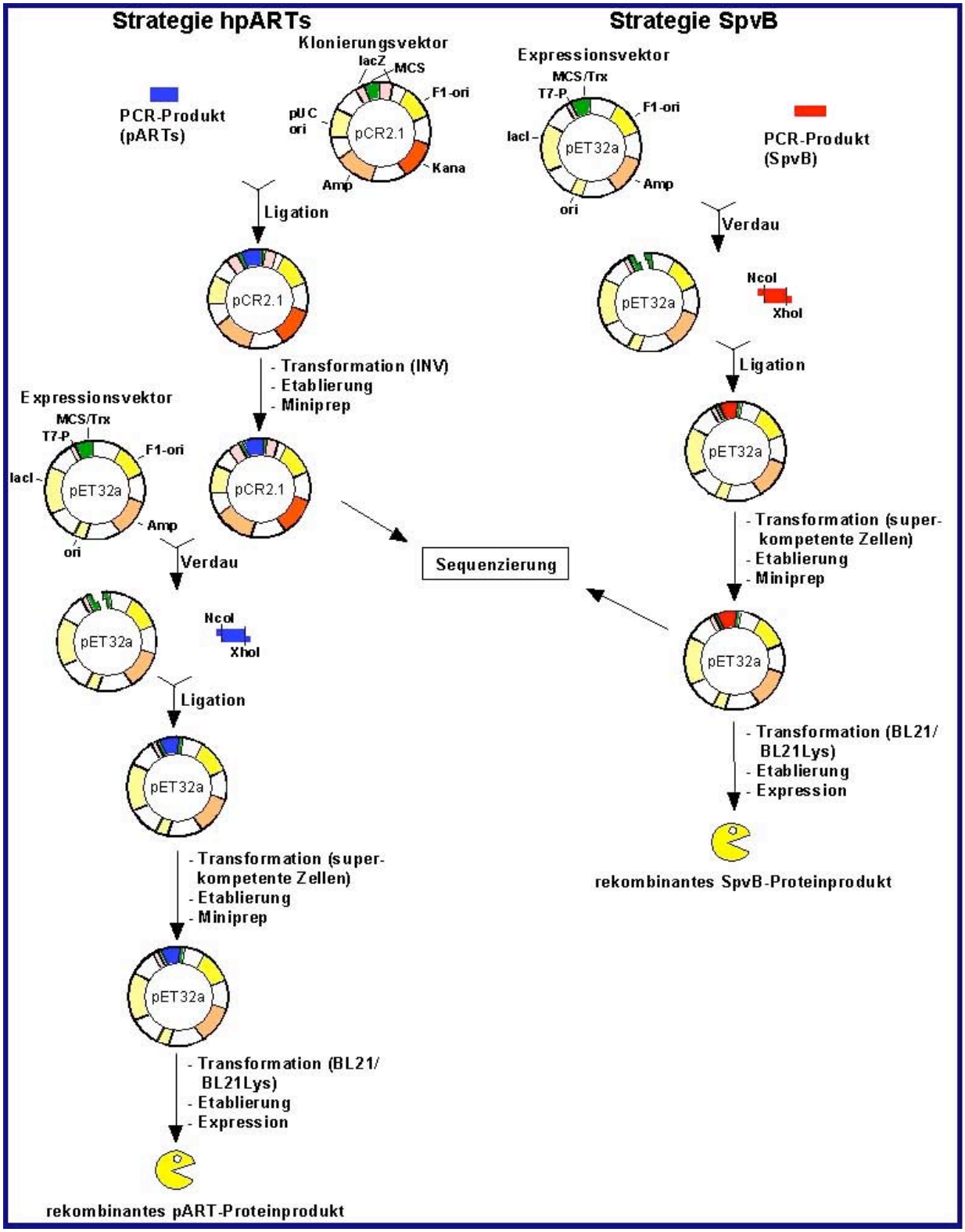


Abbildung 8: Schematische Darstellung der Klonierungsstrategien

Die Abbildung zeigt die unterschiedlichen Klonierungsstrategien für die rekombinante Expression von pART3, pART4 und hpART1 einerseits und des Salmonellentoxins SpvB andererseits. Die Klonierung in die Expressionsvektoren ist am Beispiel des Vektors pET32a dargestellt. Im Gegensatz zu den humanen poly-ARTs wurde das amplifizierte SpvB-PCR-Produkt direkt verdaut und in die Expressionsvektoren pET32a bzw. pET26b ligiert. Die Sequenzierung und damit die Auslese der mutationsfreien Klone erfolgte erst nach der Etablierung der vollendeten Expressionsvektoren mit Hilfe vektorspezifischer Primer, während die Inserts der humanen poly-ARTs bereits vor der Ligation in die Expressionvektoren sequenziert wurden. MCS = „multiple cloning site“ / „Polylinker“. Amp = Ampicillin-Resistenz-Gen. Kana = Kanamycin-Resistenz-Gen. Ori = Replikationsursprung, der die effiziente Vermehrung des Plasmids ermöglicht. lacZ = *lacZ*-Gen, das β-Galaktosidase kodiert. Trx = kodierende Sequenz für den Trx-Tag. T7-P = T7-Promotor.

3.5.1 Klonierungsstrategie für hpART1, hpART3 und hpART4

3.5.1.1 Klonierung in den pCR2.1-Vektor und Transformation in INV-Zellen

Aus den Aliquots des ersten und zweiten Schrittes der unter 3.3.1 beschriebenen Touchdown-PCR wurden die amplifizierten DNA-Fragmente mit dem „TA^R Cloning Kit“ (Invitrogen) in den pCR2.1-Vektor (Invitrogen) ligiert. Die Ligationsprodukte wurden nach Protokoll mittels „Heat shock“ (30sec bei 42°C) in INV-Zellen (Invitrogen) transformiert. Nach Inkubation auf Eis für 2min wurden die Proben mit 250µl SOC-Medium aufgefüllt und 1h bei 37°C geschüttelt. Je 100µl des Ansatzes wurde auf, mit dem Antibiotikum Carbenicillin (Serva) versetzten, Luria-Agarplatten (Miller`s LB Agar, Gibco BRL) ausgestrichen und über Nacht bei 37°C bebrütet. Die Agarplatten wurden zuvor mit 2%igem X-Gal (5-Bromo-4-Chloro-3-Indolyl-β-D-Galactopyranoside, Sigma) behandelt, um lac$^+$-Kolonien blau anzufärben und damit die Unterscheidung von transformierten Kolonien und Kolonien ohne Insert zu ermöglichen. Die weißen bzw. leicht bläulichen Kolonien wurden gepickt und in 2ml, mit 200µg Carbenicillin (Serva) versetztem LB-Medium (Luria Broth Base, Gibco BRL) bei 37°C über Nacht geschüttelt. Nach der Plasmidaufreinigung (siehe 3.6.1) wurde zur Kontrolle der Insertion ein Restriktionsenzymverdau mit EcoR I (NEB, 20000U/ml) durchgeführt. Die Proben wurden 1h bei 37°C inkubiert. Da in der „multiple cloning site“ der pCR2.1-Vektoren beidseits des eingebauten Inserts eine Schnittstelle für EcoR I besteht (siehe Anhang C), wurden die Inserts durch den Verdau aus dem Vektor herausgeschnitten und linearisierter Vektor-„Backbone“ und Insert konnten anschließend mittels Gelelektrophorese analysiert werden. Aus den Minipreps, die Plasmide mit den erwarteten Inserts enthielten, wurden Aliquots zur DNA-Sequenzierung (siehe 3.10) entnommen. Die Plasmid-Minipreps wurden bis zur weiteren Verwendung bei –20°C

gelagert. Die Klone, bei denen sich in der Sequenzierung keine Mutationen zeigten, wurden für die Subklonierung in die Expressionsvektoren verwendet.

3.5.1.2 Klonierung der Expressionsprodukte

Um die katalytischen Domänen von hpART1, hpART3 und hpART4 sowie das komplette („full length“) Protein hpART3 in *Eschericia Coli* zu exprimieren, entschieden wir uns wegen seiner Vielseitigkeit und Effizienz für das bereits unter 1.8. beschriebene pET-Expressionssystem der Firma Novagen. Aus dem großen Sortiment von pET-Plasmidvektoren waren für unsere Klonierungsstrategie die Vektoren pET26b(+) und pET32a(+) am besten geeignet (Vektorkarten: siehe Anhang C). Zur Selektion enthält der pET26b-Vektor ein Kanamycin-, der pET32a-Vektor ein Ampicillin-Resistenz-Gen. Durch Nutzung der Restriktionsenzym-Schnittstellen Nde I, Nco I und Xho I in der „multiple cloning site“ von pET26b bzw. Nco I und Xho I in der von pET32a planten wir drei verschiedene Konstrukte mit den katalytischen Domänen (CD) der pARTs und ein Konstrukt zur Produktion des kompletten („full length“) pART3-Proteins (schematische Darstellung und Beschreibung am Beispiel von pART3: siehe Abb.24):

Konstrukt A: pET26b.CD.NdeI/XhoI. Bei dieser Variante wird die katalytische Domäne von hpART1, hpART3 bzw. hpART4 isoliert im Zytoplasma produziert.

Konstrukt B: pET26b.CD.PPl.NcoI/XhoI. Die katalytische Domäne wird an einen sogenannten "pelB-Leader (pel)" fusioniert, so dass das Fusionsprotein nach der Expression in das Periplasma ausgeschleust wird. Bei diesem Vorgang wird der „pelB-Leader“ durch eine Signalpeptidase abgespalten.

Konstrukt C: pET32a.Trx-CD.NcoI/XhoI. Die katalytische Domäne wird an Thioredoxin fusioniert. Die Fusion mit Thioredoxin erhöht nach der Expression die Löslichkeit des Fusionsproteins im Zytoplasma von *E.Coli*. In das Konstrukt sind ein "S-tag", der die Aufreinigung des Proteins sowie die Markierung über Protein-S ermöglicht, sowie ein interner "His-tag" eingebaut. Der Trx-Anteil mit den beiden „tags“ kann durch Behandlung mit Enterokinase (EK) abgespalten werden.

Konstrukt D: pET26b.hpART3.long.NdeI/XhoI. Das komplette hpART3-Protein (katalytische Domäne mit der N-terminalen putativen regulatorischen Domäne) wird im Zytoplasma exprimiert.

Alle Konstrukte besitzen zudem einen C-terminalen "His-tag". Da hpART3 eine interne Nco I-Schnittstelle im vorderen Bereich des zur Expression der katalytischen Domäne amplifizierten Inserts besitzt, fehlt bei den mit Nco I klonierten Konstrukten im Gegensatz zu den Nde I-Konstrukten ein kleiner N-terminaler Teil (40 AS) der alpha-helicalen Region von pART3 („PARP regulatory“). Dies erklärt die unterschiedlich großen Banden nach der elektrophoretischen Auftrennung.

Zur Klonierung der geplanten Konstrukte wurden die mutationsfreien cDNA-Inserts mit den Restriktionsenzymen Nde I (NEB, 20.000 U/ml) und Xho I (NEB, 20.000 U/ml) bzw. Nco I (NEB, 10.000 U/ml) und Xho I aus den pCR2.1-Vektoren herausgeschnitten. Durch Zugabe des Restriktionsenzyms Sca I (Gibco, 10 U/ml) wurden die pCR.2.1-Vektor-„Backbones“ verdaut. Da das Insert der hpART4 eine interne Sca I-Schnittstelle besitzt, wurde der pCR2.1-Vektor in diesem Fall stattdessen mit dem Enzym Bgl II (NEB, 50.000 U/ml) zerschnitten. Die Vektoren pET26b und pET32a wurden ebenfalls mit Nde I und Xho I bzw. Nco I und Xho I vorbehandelt und somit linearisiert. Nach dem Verdau wurden die Ansätze mit den linearisierten Vektoren 20min bei 37°C mit 1ml Alkalischer Phosphatase (AP) aus dem Kälberdarm (Boehringer, 1000 U/ml) behandelt, um eine Religation der Vektoren zu verhindern. Um zu vermeiden, dass die AP-Aktivität die Ligation der Inserts in die Expressionsvektoren behindert, wurde die AP anschließend durch Hitze (10min bei 65°C) inaktiviert und dann durch Fällung mit StrataClean Resin (Stratagene) entfernt (Proben + 20ml StrataClean Resin, 15sec Vortex, 1min Inkubation bei RT, 1min Zentrifugation bei 3000g, Pellet verwerfen). Die verdauten Vektoren und Inserts wurden gelelektrophoretisch aufgetrennt. Die signifikanten Banden wurden ausgeschnitten und mit dem „Qiaex II Gel Extraction Kit“ bzw. dem „Qiaquick Gel Extraction Kit“ (Qiagen) nach Herstellerangaben aufgereinigt. Die Ausbeute wurde gelelektrophoretisch überprüft.

Nach der Aufreinigung wurden die Inserts mit einer T4-DNA-Ligase (Invitrogen) bei 14°C ü.N. in die Expressionsvektoren ligiert. Dabei wurden Inserts und Vektoren ungefähr in einem Verhältnis von 2:1 eingesetzt, um die Ligationsbedingungen zu optimieren. Anschließend erfolgte die Transformation der ligierten Konstrukte in „Solo Pack Gold Supercompetent Cells“ (Stratagene) nach Herstellerangaben. Statt der empfohlenen 50ml superkompetente Zellen je Transformationsansatz wurden lediglich 25ml eingesetzt. Die Zellsuspensionen mit

den pET26b-Konstrukten wurden auf mit dem Antibiotikum Kanamycin (Boehringer) versetzten LB-Agarplatten, die mit den pET32a-Konstrukten auf LB-Carbenicillin-Agarplatten ausgestrichen und ü.N. bei 37°C bebrütet. Die Ausbeute lag zwischen 7 und >300 Kolonien pro Agarplatte. Es wurden 3-6 Kolonien pro Agarplatte gepickt und in 2ml LB-Medium + Kanamycin (pET26b) bzw. LB-Medium + Carbenicillin (pET32a) ü.N. bei 37°C geschüttelt. Die Miniprep erfolgte wie unter 3.6.1 beschrieben. Zur Insertkontrolle wurden die Konstrukte anschließend mit Nde I und Xho I bzw. Nco I und Xho I verdaut und gelelektrophoretisch aufgetrennt. Die Klone mit den erwarteten Inserts wurden für die Expression der rekombinanten Proteine verwendet.

3.5.2 Klonierungsstrategie für SpvB

Im Gegensatz zu der Klonierungsstrategie für die humanen Konstrukte (3.5.1) wurden die SpvB-PCR-Produkte aus 3.3.2 nicht zunächst in pCR2.1-Vektoren kloniert und sequenziert, sondern direkt für die weitere Klonierung in Expressionsvektoren verwendet. Die SpvB-PCR-Produkte und die Vektoren wurden wie unter 3.5.1.2 beschrieben verdaut, allerdings entfiel der Einsatz von Sca I bzw. Bgl II zur Zerschneidung der pCR2.1-Vektoren. Anschließend erfolgten gelelektrophoretische Auftrennung, Ausschneidung der Banden, Gelaufreinigung, Ligation, Transformation, Miniprep und gelelektrophoretische Überprüfung der Inserts analog zu den humanen Konstrukten. Die Klone mit den erwarteten Inserts wurden mit dem „T7-Terminator“-Primer und den randständigen Primern SP1 und SP4 bzw. SP2 und SP4 sequenziert. Die mutationsfreien Klone wurden für die Expression des rekombinanten Proteins verwendet.

3.6 Plasmidaufreinigung

3.6.1 Miniprep

Die Miniprep wurde mit dem „QIAprep Spin Miniprep Kit“ (Qiagen) nach Protokoll des Herstellers durchgeführt. Die Methode basiert auf der alkalischen Lyse der Bakterien und der anschließenden Absorption der freigesetzten DNA an Kieselsäure-Membranen (in Form von Säulen) unter Bedingungen mit hohem Salzgehalt. Durch optimierte Puffer und mehrere Waschschritte wird sichergestellt, dass RNA, Proteine und Metaboliten nicht an die Membran binden und sich im Durchfluss sammeln. Nach Entfernung der Salze durch einen weiteren Waschschritt wurde die Plasmid-DNA in 50(l Elutionspuffer (EB, Qiagen) eluiert.

3.6.2 Maxiprep

3.6.2.1 Maxiprep genomischer PAC-Klone

Der genomische Sequenzen von pART3 enthaltende PAC-Klon LLNLP709M06108Q2 wurde identifiziert und vom RZPD erworben. Aus dem Agarstock wurden 3-Ösenausstriche auf LB-Kanamycin-Agarplatten angefertigt und ü.N. bei 37°C bebrütet. Eine Kolonie wurde gepickt und in 2ml LB + Kanamycin bei 37°C bis zu einer OD600 von 0,8 geschüttelt. Aus dieser Kultur wurde zum einen ein Glycerolstock angelegt (820ml Kultur + 180ml Glycerol) und bei –80°C gelagert. Zum anderen wurde ein 5ml-Aliquot der Kultur in 500ml frisch angesetztes 2xYT-Medium gegeben. Die Induktion der Kulturen mit IPTG und die Maxiprep erfolgte nach dem adaptierten Plasmidprotokoll „Isolation of Bacteriophage P1-Derived Constructs“ der Firma Qiagen mit dem „Plasmid Midi Kit“ (Qiagen). Nach dem zweiten Zentrifugationsschritt (15min bei 20000g) wurde der Überstand zusätzlich mit Millex-HA 0,45µm-Filtern (Millipore) steril filtriert. Die Plasmid-DNA wurde in 500µl TE eluiert.

3.6.2.2 Maxiprep von Vektoren und cDNA-Konstrukten

Um größere DNA-Mengen für einige Konstrukte (z.B. pCR2.1.hpART3.F63xF85) zu produzieren und die Expressionsvektoren pET26b und pET32a zu vervielfältigen, wurden Maxipreps mit dem „Endo-free Plasmid Maxi Kit“ (Qiagen) nach Herstellerangaben durchgeführt. Die Plasmide wurden in 500µl TE eluiert.

3.7 Expression von rekombinantem Protein in *E.Coli*

3.7.1 Transformation der pET26b-/pET32a-Expressionskonstrukte in Bl21(DE3)-/ Bl21(DE3)pLysS-Zellen

Die pET26b- und pET32a-Konstrukte der katalytischen Domänen von hpART3, hpART4, hpART1 und SpvB sowie von „full length“ hpART3 wurden nach dem Transformationsprotokoll von Novagen in BL21(DE3)- und BL21(DE3)pLysS-Zellen (Novagen) transformiert: Die kompetenten Zellen wurden aufgetaut und in 20µl Aliquots aufgeteilt. Im Anschluss wurde 1µl Plasmid-DNA (ca.100ng) in Elutionspuffer aus der Miniprep des jeweiligen Konstruktes hinzugegeben. Nach 30min Inkubation auf Eis wurden die Zellen für 40sec auf 42°C erhitzt („Heatshock“). Anschließend wurden die Zellen 2min auf Eis gekühlt und mit 100µl frischem SOC-Medium (Gibco BRL, Raumtemperatur) aufgefüllt. Die Kulturen wurden bei 37°C 60min lang geschüttelt und schließlich auf mit

Carbenicillin (pET32a-Konstrukte) bzw. Kanamycin (pET26b-Konstrukte) versetzten LB-Agar-Platten ausgestrichen. Die Platten wurden über Nacht bei 37°C bebrütet.

3.7.2 Expression der rekombinanten Proteine

3.7.2.1 Expression kleiner Proteinmengen und Herstellung von *E.Coli*-Gesamtzell-Lysaten

Von den über Nacht bei 37°C bebrüteten Agarplatten wurde je eine Kolonie gepickt und in 5ml LB-Carbenicillin-Medium (pET32a-Konstrukte) bzw. in 5ml LB-Kanamycin-Medium (pET26b-Konstrukte) gegeben. Die Bakterienkulturen wurden bei 37°C bis zu einer mit dem Photometer gemessenen Ziel-OD_{600} von 0.4-0.6 geschüttelt. Als dieser OD_{600}-Wertebereich erreicht war, wurden Glycerolstocks (1ml = 850µl Kultur + 150µl Glycerol) hergestellt und als Sicherheitskopie bei -80°C eingefroren. Die Kulturen wurden mit LB-Carbenicillin- bzw. LB-Kanamycin-Medium auf 9ml Gesamtvolumen aufgefüllt und mit je 10µl 1M Isopropylthiogalactosid (IPTG, Sigma) versetzt, um die Expression der rekombinanten Proteine zu induzieren. Die Kulturen wurden drei Stunden bei 37°C geschüttelt und anschließend bei 4°C und 1000g (Sorvall RC26 Plus Zentrifuge, Rotor: SA-300) abzentrifugiert. Der Überstand wurde dekantiert und verworfen. Die Pellets wurden in 500µl 1xTBS gelöst. Nach Zugabe von 5µl 100mM AEBSF (Proteasehemmer, Sigma) wurden die Zellen in 50µl und 450µl Aliquots bei -80°C eingefroren.

Zur Herstellung von Gesamtzell-Lysaten wurden die 450µl Aliquots im 37°C warmen Wasserbad aufgetaut, in einem Trockeneis-Ethanol-Bad schockgefroren und erneut aufgetaut. Die 450µl Zell-Suspensionen wurden anschließend in 50µl und in 400µl Aliquots aufgeteilt, um Gesamtproteine und lösliche Proteine getrennt zu untersuchen.

A) Gesamtproteine

Zu den 50µl Aliquots wurden 50µl 2xSDS-PAGE-Probenpuffer und 20µl 1xSDS-PAGE-Probenpuffer (Invitrogen-Novex) gegeben. Die Proben wurden 2 x 20sec mit Ultraschall (maximale Stärke, Ultraschallgerät der Firma MSE) behandelt. Zwischen den Ultraschall-Behandlungen wurden die Proben für ca. 5min auf Eis gekühlt. Anschließend wurde ein 50µl Aliquot entnommen und 5µl des Reduktionsmittels DTT (Invitrogen-Novex) zugesetzt. Die Proben wurden 5min auf 95°C erhitzt und anschließend 3min bei 13.000g abzentrifugiert.

B) Lösliche Proteine

Die 400µl Aliquots wurden 2 x 20sec mit Ultraschall (maximale Stärke, Ultraschallgerät der Firma MSE) behandelt. Zwischen den Ultraschallbehandlungen wurden die Proben für ca. 5min auf Eis gekühlt. Nach Zusatz von 450µl 1xTBS / 1mM AEBSF (Sigma) wurden die Proben 30min bei 13.000g abzentrifugiert. 20µl des Überstandes wurden abgenommen und mit 20µl 2xSDS-PAGE-Probenpuffer und 4µl DTT gemischt. Die Proben wurden 10min auf 70°C erhitzt und anschließend 3min bei 13.000g abzentrifugiert. Die restlichen, geklärten 380µl der Proben wurden bei -80°C zur weiteren Verwendung gelagert.

3.7.2.2 Expression größerer Proteinmengen und Herstellung von *E.Coli*-Gesamtzell-Lysaten (Midipreps)

Nachdem sich durch die Proteinanalytik der Gesamtzell-Lysate gezeigt hatte, dass mit den meisten Konstrukten in beiden gewählten Zellinien ausreichend rekombinantes Protein exprimiert wird, wurden für die zytoplasmatischen Konstrukte von pART3, pART4 und SpvB Gesamtzell-Lysate aus größervolumigen Kulturen hergestellt, um größere Mengen an rekombinantem Protein für Aufreinigung und Enzymassays zu produzieren. Hierzu wurden aus den Glycerolstocks der BL21-Lys-Zellen (siehe 3.7.2.1) 3-Ösen-Ausstriche auf Carbenicillin- bzw. Kanamycin-Agarplatten angefertigt. Nach Bebrütung über Nacht bei 37°C wurde je 1 Kolonie gepickt und in 10ml, mit dem entsprechenden Selektions-Antibiotikum versetztes LB-Medium gegeben. Die Kulturen wurden bei 37°C bis zu einer Ziel-OD_{600} von 0,5 geschüttelt. Nach Erreichen der angestrebten optischen Dichte wurden die Kulturen 10min bei 4°C und 1500g (Beckmann GS-6KR Zentrifuge) abzentrifugiert. Die Pellets wurden in 10ml LB-Medium resuspendiert und in Erlenmeyerkolben mit 500ml frischem LB-Carbenicillin- bzw. LB-Kanamycin-Medium überführt. Nach 3h bei 37°C auf dem Schüttler (OD_{600} bei 0,4-0,6) wurde die Expression der rekombinanten Proteine durch Zugabe von je 700µl 0,1M Isopropylthiogalactosid (IPTG, Sigma) induziert. Die Kulturen wurden nach der Induktion für weiter 2,5h bei 30°C geschüttelt. Anschließend wurden die Kulturen in Sorvall-Zentrifugen-Becher umgefüllt und 10min bei 4°C und 1500g (Sorvall RC26 Plus Zentrifuge, Rotor: SLA-3000) abzentrifugiert. Nach Verwerfen des Überstandes wurden die Pellets in Lyse-Puffer (20ml eiskaltes TBS, 1mM AEBSF (Sigma), 1% Triton-X-100 (Sigma), 50µg Lysozym (150.000U/mg, Serva)) resuspendiert. Die Proben wurden auf je 2 Falcon-Röhrchen (Becton Dickinson) aufgeteilt, um bei kleinerer Menge eine effektivere Lyse zu ermöglichen,

und 2x im Trockeneis/Ethanol-Bad schockgefroren bzw. im 37°C warmen Wasserbad wieder aufgetaut. Anschließend wurden die Proben 2 x 20sec mit Ultraschall behandelt (maximale Stärke, Ultraschallgerät der Firma MSE). Zwischen den Ultraschallbehandlungen wurden die Proben auf Eis abgekühlt. Die lysierten Proben aus den zwei Falcon-Röhrchen wurden zusammen in Sorvall-Zentrifugenröhrchen gegeben und 30min bei 4°C und 42.000g (Sorvall RC26 Plus Zentrifuge, Rotor: SA-300) abzentrifugiert. Der Überstand, der die löslichen Proteine enthält, wurde dekantiert, mit einem 0,22µm-Filter (Millex-GV, Millipore) steril filtriert und für die Aufreinigung der rekombinanten Proteine an Protein-S-Agarose und die meisten Enzymassays verwendet (Lagerung bei 4°C).

3.8 Proteinanalytik

3.8.1 Western-Blot-Analysen

Beim Protein-„Blotting“, dem sog. Western-Blot, werden die einzelnen Komponenten einer Proteinmischung gelelektrophoretisch aufgetrennt und anschließend auf eine Membran überführt, um sie auf dieser mit spezifischen Reaktionen nachzuweisen. Für die SDS-PAGE-Analyse in dieser Arbeit verwendeten wir das "NuPAGE"-System (Invitrogen-Novex). Die Proben mit den rekombinanten Proteinen wurden unter denaturierenden Bedingungen (SDS-Lade-Puffer + 1mM DTT) 10min bei 70°C bzw. 5min bei 95°C erhitzt (siehe Herstellung der Zelllysate) und anschließend 3min bei 13.000g abzentrifugiert. Jeweils 20µl des Überstandes wurde in die Taschen eines 10%igen Tricin-Geles (NuPage, Novex) geladen. Als Marker wurden „Mark12 Wide Range Molecular Weight Standard“ (Novex) und „SeeBlue Pre-Stained Standard“ (Novex) verwendet. Die Gelkammern wurden mit Laufpuffer (NuPage MOPS SDS running buffer) gefüllt. Der MOPS-Puffer eignet sich, um Proteine mit einem Molekulargewicht von >200kD bis 14kD effizient aufzutrennen. Der inneren Gelkammer wurden noch je 500µl „NuPage Running Buffer antioxidant“ pro 200ml Laufpuffer zugesetzt. Die Proteine wurden bei 200 Volt 30min fraktioniert. Zum Blotten wurde das "X Cell II Blotting"-System (Invitrogen-Novex) benutzt, bei dem der Transfer über das sog. "Wet"-Verfahren erfolgt. Es wurden 2 Typen von Membranen verwendet: Einerseits NC-Membranen (Nitrocellulose-Hybond-C, Amersham-Pharmacia), die eine niedrige Bindekapazität (bis zu 180µg/cm^2) besitzen und im Anschluss silbergefärbt wurden, um das rekombinante Proteine als “neu“ aufgetretene Bande nachzuweisen sowie das Verhältnis des exprimierten Proteins zum gesamten Proteingemisch darzustellen. Andererseits PVDF-

Membranen (Polyvinyldendifluorid-ImmobilonP, Boehringer), die eine hohe Bindekapazität (bis zu 600µg/cm^2) auszeichnet und die für die Immundetektion des rekombinanten Proteins verwendet werden können. Die PVDF-Membran wurde intial 15–20sec in Methanol aktiviert. Anschließend wurden Membranen und Filterpapiere in Blotpuffer (NuPage Transfer Buffer) angefeuchtet. Der Blot wurde als „Sandwich“ luftblasenfrei aufgebaut: Die NC-Membran befand sich auf der Anoden-Seite, die PVDF-Membran auf der Kathoden-Seite des Tricin-Geles. Bei 30 Volt wurden die Proteine 2min auf die NC-Membran und 90min auf die PVDF-Membran transferiert. Die PVDF-Membran wurde hinterher mit 10%igem Ziegenserum (Sigma) geblockt, mit TBS gewaschen, für die Detektion mittels Chemilumineszenz-Reaktion verwendet oder getrocknet und zur weiteren Analyse aufbewahrt.

3.8.2 Silber- und Coomassiefärbung

Die Effizienz des Blotvorganges und die Expression der rekombinanten Proteine, wenn diese in ausreichender Menge produziert werden, kann durch Protein-anfärbende Farbreagenzien überprüft werden. Die NC-Membran wurde 1-2min in 10ml frisch angesetzter Silberlösung (9ml H_2O, 0,5ml Natriumcitrat 40%, 0,4ml Eisensulfat 20%, 0,1ml Silbernitrat 20%) geschwenkt, mit destilliertem H_2O gewaschen und luftgetrocknet. Die Tricin-Gele wurden je nach Proteindichte zwischen 6 und 20 Stunden in einer 100ml Coomassie-Färbe-Lösung (Colloidal Coomassie Stain, Novex) geschwenkt. Die Entfärbung der unselektiv angefärbten Gelabschnitte erfolgte mit aktivierten Kohle-Beuteln in H_2O über Nacht. Anschließend wurden die Gele 10min in 10ml Fixierungslösung ("DryEase Gel drying system", Invitrogen-Novex) inkubiert und zwischen zwei Cellophanfolien luftgetrocknet.

3.8.3 Detektion und Chemilumineszenz-Reaktion

Für die spezifische Detektion der rekombinanten Proteine wurde die PVDF-Membran verwendet. Nach dem “Blotten“ wurden die unspezifischen freien Bindungstellen der Membran 10min in 10%igem Ziegenserum (Sigma) geblockt. Die Membran wurde mit TBS gewaschen und anschließend mit 1:5000 verdünnter, Protein-S-gekoppelter Meerrettich-Peroxidase (Novagen) 60min bei 4°C rotierend inkubiert. Protein-S bindet an den S-tag, der mittels des pET32a-Vektors N-terminal in die rekombinanten Fusions-Proteine eingebaut wurde. Um unspezifisch gebundenes Protein-S zu entfernen, wurden nach der Inkubation

fünf Waschschritte (2 x 2min / 3 x 10min) mit TBS / 0,5%Tween (USB) durchgeführt. Die an das gebundene Protein-S gekoppelte Meerrettich-Peroxidase ermöglicht es, das rekombinante Protein mittels einer enzymatischen Chemilumineszenz-Reaktion nachzuweisen. Hierzu wurde die Membran mit "ECL Western Blotting detection reagent" (Amersham-Pharmacia) bedeckt, 60sec inkubiert und in Plastikfolie eingeschlagen. Je nach Intensität wurde für 5 - 60sec ein Hyperfilm ECL (Amersham-Pharmacia) exponiert und sofort entwickelt. Die Vorteile dieses Nachweises sind hohe Sensitivität (<25 altomol Meerrettich-Peroxidase), kurze Entwicklungszeit und die Detektion ohne Einsatz von radioaktiven Stoffen.

3.8.4 Aufreinigung der Trx-Fusionproteine mit Protein-S-Agarose

Der „S-tag“ in den pET32a-Konstrukten ermöglicht eine Aufreinigung der Fusionsproteine mittels Protein-S-Agarose. Die Aufreinigung basiert auf der spezifischen Komplexbildung des „S-tag“ mit dem 104 AS umfassenden S-Protein-Anteil der Ribonuklease A, der kovalent an Agarose-Kügelchen gebunden und somit immobilisiert ist. Nach Entfernung der ungebundenen Proteine ermöglicht die Lage der Enterokinase-Schnittstelle in den Fusionsproteinen eine Abspaltung der aufgereinigten, rekombinanten katalytischen Domäne von der Agarosematrix durch einen Enterokinase-Verdau. Für die Aufreinigung der pET32a-Konstrukte von pART4, SpvB sowie der Negativkontrolle pET32a (ohne Insert) wurden je nach Experiment 20 – 2500µl *E.Coli*-Gesamtzell-Lysat eingesetzt. Die Lysate wurden mit 20-100µl Protein-S-Agarose (Novagen) 2h bei 4°C rotierend inkubiert. Nach Abzentrifugieren der Matrix (10min bei 1000g) wurde der Überstand verworfen und die Pellets 4x (2 x 2min, 2 x 5-10min) in 1x PBS (Gibco BRL) / 0,5% Triton-X-100 (Sigma) gewaschen, um die ungebundenen Proteine zu entfernen. Anschließend wurden die Proben mit den aufgereinigten Konstrukten je nach Experiment einem Enterokinaseverdau zugeführt oder direkt für die Enzymassays mit radioaktiv markiertem NAD^+ verwendet.

3.9 Enzymassays

3.9.1 Enzymassays mit *E.Coli*-Gesamtzell-Lysaten

3.9.1.1 Nachweis der Enzymaktivität von rekombinanter pART1 und pART4

Je 50µl *E.Coli*-Gesamtzell-Lysat (BL21Lys-Zellen) mit den Konstrukten pET26b.hpART1.CD, pET26b.hpART4.CD bzw. der Negativkontrolle pET26b ohne Insert sowie 50µl Periplasma-Lysat aus NM522-Zellen, das als Positivkontrolle Ratten-ART2.2

enthält (pASK.rART2.2-FH/NM522), wurden in „Labeling“-Puffer (50µl PBS, 2mM „kalte“ ADP-Ribose (Sigma), 2,5mM DTT) für 1h bei 37°C mit 1µl [^{32}P]-NAD^{+} (1:10) (Amersham-Pharmacia) inkubiert. Nach Zugabe von je 15µl StrataClean Resin (Stratagene) zur Fällung der Proteine wurden die Proben 60sec bei 5.000g abzentrifugiert. Die Pellets wurden 1x mit PBS gewaschen und anschließend einer SDS-PAGE-Analyse mit Autoradiographie zugeführt. Der gleiche Versuchsansatz wurde zudem mit 50µl Maus-ART2.2-Periplasma-Lysat (pASK.mART2.2-FH/NM522) durchgeführt. Die pASK.ART2.2-Konstrukte wurden in der Arbeitsgruppe kloniert (Bredehorst 2002).

3.9.1.2 Dünnschichtchromatographie zur Analyse der NADase-Aktivität

50µl *E.Coli*-Lysate aus den mit pET26b.hpART4.CD bzw. mit der Negativkontrolle pET26b.ohne Insert transformierten BL21Lys-Stämmen sowie 50µl Ratten-ART2.2-Periplasma-Lysat (pASK.rART2.2-FH/NM522) wurden mit 50µl PBS, 2mM „kalter“ ADP-Ribose (Sigma), 2,5mM DTT und 1µl [^{32}P]-NAD^{+} (1:10) (Amersham-Pharmacia) 60min bei 37°C inkubiert. Die Proteine wurden durch Inkubation mit StrataClean Resin (Stratagene) und anschließende Zentrifugation (60sec bei 5000g) gefällt und einer SDS-PAGE-Analyse zugeführt. Je 1µl der Überstände wurde auf eine PEI-Cellulose-Folie (Merck) aufgetragen und in eine Glaskammer mit 100ml 1M Essigsäure-Puffer gestellt. Nachdem die Laufgrenze die Auftragshöhe der Proben überschritten hatte (ca. 5min), wurde die Folie entnommen, der Puffer gegen einen Puffer mit 3M Lithiumchlorid / 1M Essigsäure ausgetauscht und die Folie erneut ca. 50min in die Kammer gestellt, bis der Laufrand ca. 5cm unter der Oberkante der Membran lag. Nach Lufttrocknung der Folie wurde sie in Plastikfolie eingeschweißt und bei –80°C für 14h ein Röntgenfilm (Biomax MR, Kodak) aufgelegt.

3.9.1.3 Aktin-Enzymassay mit den SpvB-Konstrukten

Zu je 50µl der SpvB-Bakterienlysate wurde 50µl 2xEnzympuffer (100µM Tris pH8, 2mM DTT, 0.16% Tween 20, 4mM ADPR) gegeben. Es wurde ein Versuchsansatz ohne externes Substrat und einer mit je 0,6µg Kaninchen- und 0,6µg Rinder-Skelettmuskel-Aktin (Sigma) durchgeführt. Anschließend wurden die Ansätze mit je 1µl eines NAD-Gemisches (1µl [^{32}P]-NAD^{+} (Amersham-Pharmacia) + 5µl 100µM kaltes NAD^{+} (Sigma) + 4µl H_2O) versetzt und 60min bei 37°C inkubiert. Nach Inkubation mit je 15µl StrataClean Resin (Stratagene) zur

Fällung der Proteine und anschließendem Abzentrifugieren wurde das Pellet 2x mit PBS / 1%Triton-X-100 (Sigma) gewaschen und dann mit 20µl SDS-Puffer (Novex) versetzt. Die Proben wurden auf ein 10%iges Tricin-Gel geladen. Anschließend wurden SDS-PAGE-Analyse, Blotting und Färbungen wie oben beschrieben durchgeführt. Auf die gefärbten Gele wurden Röntgenfilme (BioMax, Kodak) aufgelegt und über 9 Tage bei –80°C exponiert.

3.9.2 Enzymassays mit den aufgereinigten Trx-Fusionsproteinen von hpART4 und SpvB

3.9.2.1 Untersuchung der Automodifikation von Trx-pART4 und Trx-SpvB

Rekombinante Trx-Tag-Fusionsproteine wurden aus *E.Coli*-Lysaten (BL21-Lys-Zellen nach Transformation mit den Konstrukten pET32a.Trx-pART4.CD, pET32a.Trx-SpvB.CD bzw. als Kontrolle pET32a.ohne Insert) mittels Protein-S-Agarose aufgereinigt (siehe 3.8.4). Nach dem 4.Waschschritt in PBS / 0,5% Triton-X-100 (Sigma) wurden die Pellets in 100µl PBS + 2mM ADP-Ribose (Sigma) aufgenommen und für 1h bei 37°C mit 3µl [^{32}P]-NAD^{+} (1:10) (Amersham-Pharmacia) inkubiert. Im Anschluss wurden die Agarose-Pellets 2x in PBS und 1x in 100µl 1x Enterokinase (EK)-Puffer (Cleavage Buffer, Novagen) gewaschen und schließlich in 100µl EK-Puffer aufgenommen. Nach Abnahme eines Aliquots wurden die Proben mit 0,5µl rekombinanter Enterokinase (rEK 2U/µl, Novagen) über Nacht bei RT inkubiert. Nach Zentrifugation wurde der Überstand dekantiert. Die Pellets und die vor dem EK-Verdau abgenommenen Aliquots wurden in SDS-Lade-Puffer (Novex) aufgenommen, die Proteine im Überstand wurden mit 10µl StrataClean Resin (Stratagene) gefällt und ebenfalls einer SDS-PAGE-Analyse mit anschließendem Western Blot und Autoradiographie zugeführt.

3.9.2.2 Untersuchung der NH_2OH-Sensitivität der durch rekombinantes Trx-SpvB katalysierten ADPR-Bindungen

300µl Trx-SpvB-Gesamtzell-Lysat (BL21-Lys-Zellen) wurde an Protein-S-Agarose aufgereinigt (siehe 3.8.4). Nach den Waschschritten wurden die Pellets in 250µl PBS + 2mM ADP-Ribose 1h bei 37°C mit 5µl [^{32}P]-NAD^{+} (1:10) (Amersham-Pharmacia) inkubiert. Nach 3x Waschen in PBS / 0,5% Triton-X-100 (Sigma) wurde der Ansatz in fünf Aliquots aufgeteilt. Nach erneutem Abzentrifugieren wurden die Agarose-Pellets mit den gebundenen Proteinen in fünf verschiedenen 50µl-Lösungen aufgenommen und ü.N. bei RT rotierend

inkubiert: 10mM Tris pH 8 (1), 10mM NAD^+ in 10mM Tris pH 8 (2), 1M NaCl in 10mM Tris pH 8 (3), 1M NH_2OH pH 7,0 (4) bzw. 10mM $HgCl_2$ in 10mM Tris pH 8 (5). Am nächsten Tag wurden die Proben abzentrifugiert, die Pellets in SDS-Lade-Puffer (Novex) aufgenommen und mittels SDS-PAGE und Autoradiographie analysiert.

3.9.2.3 Nachweis der ADP-Ribosylierung einer 40kd-Bande in Jurkat-Zell-Lysaten

Zur Herstellung von Lysaten einer humanen T-Zellinie wurden 1-2x10^7 Jurkat-Zellen in 500µl PBS (Gibco BRL) + 1% Triton-X-100 (Sigma) + 1mM AEBSF (Sigma) lysiert. 300µl pET32a.Trx-SpvB Gesamtzell-Lysat (BL21-Lys-Zellen) wurden an Protein-S-Agarose aufgereinigt, nach den Waschschritten in 100µl EK-Puffer 1:10 (Cleavage Buffer, Novagen) aufgenommen und ü.N. bei RT mit 1µl rekombinanter Enterokinase (rEK 2U/µl, Novagen) rotierend inkubiert. Am nächsten Tag wurde die Agarose-Matrix abzentrifugiert (10min bei 1000g) und der Überstand dekantiert. Zu je 50µl Überstand wurden 50µl PBS bzw. 50µl Jurkat-Zell-Lysat gegeben. Als Kontrolle wurden zudem 50µl Jurkat-Zelllysat mit 50µl PBS ohne Zugabe von Trx-SpvB-*E.Coli*-Lysat vermischt. Anschließend wurden die Proben mit 2mM ADP-Ribose und 2µl [^{32}P]-NAD^+ (1:10) (Amersham-Pharmacia) 1h bei 37°C inkubiert. Die Proteine wurden mit je 15µl StrataClean Resin (Stratagene) gefällt und abzentrifugiert (60sec bei 5000g). Nach 2x Waschen in PBS wurden die Pellets in SDS Lade Puffer (Novex) aufgenommen und mittels SDS-PAGE, Western-Blot und Autoradiographie analysiert.

3.9.3 Autoradiographien

Zum Nachweis der inkorporierten radioaktiven Aktivität wurden Coomassie-gefärbte Gele bzw. Silber-gefärbte NC-Membranen in eine Röntgenfilmkassette geklebt. Anschließend wurden in der Dunkelkammer Röntgenfilme (Biomax MR, Xomat) der Firma Kodak aufgelegt. Nach lichtdichtem Verschluss der Kassetten wurden die Filme bei –80°C exponiert und anschließend nach Übertragung des Markers von den Gelen bzw. NC-Membranen analysiert.

3.10 Automatische DNA-Sequenzierungen

Die DNA-Sequenzierungen erfolgten nach der Didesoxy-Methode mit Hilfe des „Dye-Terminator“-Sequenzierungskits (Applied Biosystems) mit entsprechenden Oligonukleotiden

und 500ng Template-DNA nach Herstellerangaben. Statt der vom Hersteller empfohlenen 8µl des Reaktionsmixes (Big Dye Ready Mix) wurden 4µl eingesetzt. PCR-Programm: 40sec 96°C, 15sec 50°C, 4min 60°C, 27 Zyklen. Die im „Big Dye“-Mix enthaltene Polymerase synthetisiert zum Template komplementäre Nukleotidstränge, in die auch vier verschiedene fluoreszierende Didesoxynukleotide eingebaut werden. Dies führt zu Kettenabbrüchen. Alle Sequenzierungen wurden an beiden DNA-Strängen ausgeführt. Nach der PCR-Reaktion wurden die 20µl Sequenzreaktionen mit 50µl Ethanol 100% und 2µl 3M Natriumacetat versetzt, 10min auf Eis gelagert und anschließend 30min bei 13000g und 4°C abzentrifugiert, um die Reaktionsprodukte zu fällen. Die Überstande wurden verworfen. Nach Zugabe von 70%igem Ethanol wurden die Probe erneut abzentrifugiert (10min, 13000g, 4°C). Nach Vakuumtrocknung des Präzipitates erfolgte die Sequenzanalyse mit dem „Modell 370A DNA-Sequencing System“ (ABI) im Service Labor des Instituts für Zellbiochemie und Klinische Neurobiologie des Universitätskrankenhauses Hamburg-Eppendorf.

4 Ergebnisse

Der Ergebnisteil dieser Arbeit gliedert sich in drei Abschnitte.

Im ersten, bioinformatischen Teil werden Datenbankanalysen vorgestellt, mit deren Hilfe neue putative ADP-Ribosyl-Transferasen identifiziert wurden. Durch *in silico* Experimente ließen sich diese neuen ADP-Ribosyltransferasen bestimmten Proteinfamilien zuordnen. Es wurden Orthologe bei Maus und Ratte sowie homologe Proteine bei weiteren eukaryontischen Spezies von *Drosophila melanogaster* bis zum Wurm *Caenorhabditis elegans* und bei Pflanzen (z.B. *Arabidopsis thaliana*) nachgewiesen. Mit Hilfe von Alignments wurden die Aminosäuresequenzen der Familienmitglieder verglichen und konservierte Regionen identifiziert. Desweiteren wird die chromosomale Lokalisation der humanen poly-ART-Gene sowie ihrer murinen Orthologen dargestellt. Bei ausgewählten Familienmitgliedern werden die ermittelten Exon-Intron-Strukturen der Gene sowie die Domänenstrukturen der kodierten Proteine vorgestellt.

Im zweiten Abschnitt wird das experimentelle Vorgehen zur molekularen Klonierung und rekombinanten Expression von zwei neuen putativen poly-ADP-Ribosyl-Transferasen des Menschen, pART3 und pART4, abgehandelt. Die ART-Enzymaktivität des rekombinanten hpART4 wird *in vitro* nachgewiesen. Weiterhin werden Ergebnisse zur Aufklärung der Genstruktur von *pART3* und dessen chromosomaler Lokalisation gezeigt.

Im dritten Teil wird die molekulare Klonierung und rekombinante Expression einer neuen bakteriellen mono-ADP-Ribosyl-Transferase, des Salmonellen-Toxins SpvB, dargestellt. Neben dem *in vitro* Nachweis der Enzymaktivität wird gezeigt, dass dieses Toxin in humanen Zelllinien ein Zielprotein ADP-ribosyliert und dass Aktin *in vitro* ein Zielprotein von SpvB ist.

4.1 *In silico* Analyse der poly-ART-Genfamilie

4.1.1 Identifikation von RT6- und PARP-verwandten Genen in dbEST

Um neue RT6- und PARP-verwandte Gene im menschlichen Genom zu identifizieren, durchsuchten wir mit tBLASTn und den Sequenzen der katalytischen Domäne bekannter ADP-Ribosyltranserasen als “Input“ die EST-Datenbank der GenBank nach Sequenzen, die Sequenzidentitäten mit bekannten ADP-Ribosyl-Transferasen zeigen. Die ESTs, die eine signifikante Sequenzidentität aufwiesen, wurden untereinander mit Sequenz-Alignment-Programmen verglichen. ESTs, die offensichtlich vom selben Gen abstammten, wurden in

Gruppen (sog. cluster) eingeordnet. Aus diesen überlappenden ESTs wurde für jede neue putative ART eine vorläufige konsensuelle cDNA-Sequenz entworfen. Mit dieser cDNA-Sequenz wurde dbEST mittels BLASTn erneut durchsucht, um die genaue Anzahl der zugehörigen ESTs auf Nukleotidbasis zu ermitteln. Über den „Unigene"-Link wurden zudem die zugehörigen „Unigene-Cluster" der einzelnen Gene ausfindig gemacht. Die Untersuchungen wurden 1998 begonnen und seit 1999 jährlich wiederholt, um mögliche neue Mitglieder zu identifizieren und die Anzahl der zugehörigen ESTs im Verlauf zu dokumentieren. Auf diese Weise wurden 4 verschiedene humane Gene, die mit der RT6-ekto-ADP-Ribosyl-Transferase verwandt sind (*ART1*, *ART3*, *ART4*, *ART5*), und 17 verschiedene PARP-verwandte humane EST-„Cluster" identifiziert. Die PARP-verwandten putativen poly-ARTs wurden pART (für p**oly**-**A**DP-**r**ibosyl**t**ransferase) genannt. In **Abbildung 9** sind die Ergebnisse der Datenbankrecherchen tabellarisch zusammengefasst (letzte Aktualisierung: Oktober 2004). Die humane PARP wurde mit pART1 bezeichnet, die Nummerierung der weiteren pARTs (pART 2-17) spiegelt die Sequenzähnlichkeit zur pART1 und der einzelnen pARTs untereinander im Bereich der konservierten Sekundärstrukturmotive der katalytischen Domäne wider. Zudem wurden für die Nummerierung die in den folgenden Kapiteln vorgestellten Ergebnisse zu den Untergruppen der pART-Familie berücksichtigt.

Name des Genes	Anzahl der ESTs in dbEST						Unigene "cluster"
	0kt 99	Jun 00	Jun 01	Feb 02	0kt 03	0kt 04	
A) mono-ART-Familie							
ART1	**0**	**0**	**0**	**0**	**1**	**1**	**Hs.382188**
ART2P	**0**	**0**	**0**	**0**	**0**	**0**	**--**
ART3	**21**	**35**	**66**	**74**	**87**	**95**	**Hs.24976**
ART4	**9**	**9**	**11**	**11**	**16**	**23**	**Hs.13776**
ART5	**3**	**4**	**3**	**10**	**13**	**17**	**Hs.125680**
B) poly-ART-Familie							
1	**157**	**339**	**492**	**527**	**756**	**862**	**Hs.177766**
2	**54**	**58**	**78**	**84**	**120**	**122**	**Hs.409412**
3	**23**	**24**	**40**	**64**	**109**	**109**	**Hs.271742**
4	**106**	**132**	**187**	**202**	**282**	**337**	**Hs.117825**
5	**81**	**108**	**181**	**188**	**260**	**298**	**Hs.370267**
6			**129**	**134**	**205**	**430**	**Hs.329327**
7		**7**	**10**	**12**	**8**	**12**	**Hs.120250**
8		**10**	**241**	**252**	**298**	**331**	**Hs.518203**
9			**122**	**132**	**226**	**243**	**Hs.518200**
10	**47**	**25**	**67**	**152**	**279**	**290**	**Hs.348609**
11		**10**	**12**	**16**	**33**	**99**	**Hs.504538**
12	**51**	**67**	**115**	**134**	**183**	**187**	**Hs.12646**
13		**62**	**93**	**96**	**288**	**312**	**Hs.133512**
14	**39**	**87**	**140**	**145**	**196**	**214**	**Hs.12813**
15					**107**	**116**	**Hs.30634**
16			**61**	**64**	**121**	**166**	**Hs.369581**
17			**161**	**175**	**286**	**304**	**Hs.270244**
C) andere Gene							
ARH 1	**0**	**0**	**1**	**1**	**71**	**54**	**Hs.99884**
PARG	**28**	**34**	**73**	**88**	**82**	**77**	**Hs.91390**
HPRT	**46**	**67**	**148**	**178**	**273**	**299**	**Hs.412707**
G3PDH	**>2000**	**>3700**	**>10000**	**>10000**	**>15000**	**>15000**	**Hs.479728**

Abbildung 9: Humane EST-„Cluster“ mit Sequenzidentitäten zu bekannten ADP-Ribosyl-Transferasen in der EST-Datenbank (dbEST)

Die Abbildung fasst die Ergebnisse der Datenbankrecherchen tabellarisch zusammen. Neben der Anzahl der ESTs in dbEST zu den jeweiligen Zeitpunkten der Suchen ist für die identifizierten Gene die Zugangsnummer („accession number“) des aktuellen „Unigene clusters“ aufgeführt (Stand Februar 2005). Die pART-Subgruppen, die im weiteren Verlauf des Kapitels vorgestellt werden, sind farblich hervorgehoben: Subgruppe 1 rot, Subgruppe 2 rosa, Subgruppe 3 orange, Subgruppe 4 grün, Subgruppe 5 grau. Die pARTs, die zur weiteren Charakterisierung ausgewählt wurden, sind gelb markiert. G3PDH = Glycerinaldehyd-3-Phosphat-Dehydrogenase, HPRT = Hypoxanthin-Guanin-Phosphoribosyltransferase, ARH 1 = Arginin-ADP-Ribosyl-Hydrolase, PARG = Poly-(ADP-Ribose)-Glykohydrolase.

Im Einklang mit dem exponentiellen Wachstum der GenBank nahm die Anzahl der deponierten Sequenzen der meisten ARTs in den letzten Jahren deutlich zu. Nachdem im Oktober 1999 lediglich acht verschiedene PARP-verwandte EST-„Cluster“ in dbEST zu ermitteln waren, konnten bei den folgenden Suchen Sequenzen weiterer Gene (z.B. *pART6*, *pART9*) der pART-Familie zugeordnet werden, die in den vorherigen Jahren entweder noch nicht in der Datenbank deponiert oder mit den verwendeten Suchprogrammen noch nicht aufzufinden waren. Es fällt auf, dass einige *pART*-Gene bereits durch eine hohe Anzahl an ESTs repräsentiert sind (z.B. *pART1*, *pART4*, *pART6*, *pART8*, *pART13* und *pART17* mit >300 ESTs), während für *ART5* und *pART7* lediglich siebzehn bzw. zwölf ESTs gefunden wurden und für *ART1* im Februar 2002 sogar noch keine einzige EST-Sequenz vorhanden war. Erst bei der Suche im Herbst 2003 wurde ein signifikantes EST für den vorderen Teil der Sequenz von *ART1* ermittelt. Desweiteren wurden im Oktober 2003 erstmals ESTs einer siebzehnten humanen pART nachgewiesen (*pART15*). Bei der jüngsten Suche im Oktober 2004 waren für alle pARTs, mit Ausnahme von pART7, mehr als 100 ESTs nachweisbar. Eine Analyse der ermittelten ESTs zeigte, daß die ESTs für die pARTs aus vielen verschiedenen Geweben stammen. Dies weist auf eine ubiquitäre Expression der *pART*-Gene hin. Die humanen Mitglieder der mono-ART-Familie, deren Expression auf spezielle Gewebe beschränkt ist (Glowacki *et al.* 2002), sind durch eine deutlich geringere Anzahl an ESTs repräsentiert. Da es sich bei dem humanen *ART2P* um ein Pseudogen handelt, existieren keine transkribierten Sequenzen.

Um die Anzahl der gefundenen Sequenzen in Relation setzen zu können, führten wir zum Vergleich BLASTn Suchen für zwei humane "Haushälter"-Enzym-Gene in der EST-Datenbank durch: Die Glycerinaldehyd-3-Phosphat-Dehydrogenase (G3PDH) ist ein Schlüsselenzym der Glykolyse, die Hypoxanthin-Guanin-Phosphoribosyltransferase (HPRT) ein Schlüsselenzym des Purinmetabolismus. Während die pARTs im Vergleich zum ubiquitär exprimierten *G3PDH* in deutlich geringerer Anzahl repräsentiert sind, ist die Expression der *HPRT* mit dem Expressionsniveau einiger pARTs vergleichbar. Zur Gegenüberstellung wurden die BLASTn Suchen auch für die zwei katabolischen Enzyme Arginin-ADP-Ribosyl-Hydrolase (ARH 1) und Poly-(ADP-Ribose)-Glykohydrolase (PARG) durchgeführt. Diese Enzyme fungieren als Gegenspieler der ARTs, indem sie die glykosidischen Bindungen spezifisch spalten, die ADP-Ribose freisetzen und somit evtl. eine

Inaktivierung des Zielproteins aufheben. Während für *PARG* schon seit 1999 mehrere ESTs in der Datenbank deponiert sind, wurde für *ARH 1* bis zum Februar 2002 lediglich ein einziges EST ermittelt. Bei der jüngsten Suche wurden 54 zugehörige ESTs identifiziert.

Bei verschiedenen Genen (z.B. *PARG*, *pART7, ARH 1)* fällt zudem auf, dass bei aktuelleren Suchen weniger ESTs als zuvor identifiziert werden konnten. Dies ist am ehesten durch die Entfernung redundanter bzw. kurzer, nicht signifikanter ESTs oder chimärer Klonierungsartefakte aus den „Clustern“ und die dadurch bedingte Neuordnung von „Unigene Clustern“ anhand aktueller Erkenntnisse zu erklären. In der Abbildung sind die aktuellen „Unigene-Cluster“ der einzelnen Gene mit der jeweiligen Zugangsnummer aufgeführt (Stand: Februar 2005).

4.1.2 Die Familie der humanen pARTs und ihre murinen Orthologen

In der EST-Datenbank wurden mit tBLASTn und den Konsensussequenzen der 17 humanen pARTs als „input“ hochkonservierte murine Orthologe für alle pARTs mit der Ausnahme von pART7 aufgespürt. Durch wiederholte Suchen mit den Konsensussequenzen in der „nr-Datenbank“ konnten schließlich allen humanen pARTs und mit der Ausnahme von mpART7 und mpART13 auch den murinen pARTs mRNAs und zugehörige Proteinsequenzen zugeordnet werden. Für mPART13 wurden lediglich partielle Sequenzen gefunden, bei denen die C-terminale Region mit der pART-katalytischen Domäne fehlt. Deshalb wurde weiterhin die aus den ESTs erstellte Konsensussequenz verwendet. Die Aminosäuresequenzen der potentiellen katalytischen Domäne der humanen und murinen pARTs wurden mit Hilfe von Sequenz-Alignment-Programmen (T-Coffee, Megalign) unter Einbeziehung der Alignments der BLAST-Suchen verglichen (**Abbildung 10**). Im Bereich der Sekundärstrukturmotive, die in den bekannten Kristallen der Hühner-pART1, der Maus-pART2, des *Pseudomonas aeruginosa* Exotoxin A (ETA) und des Diphterietoxins (DT) das katalytische Zentrum bilden, ist die Sequenzidentität der verschiedenen pARTs am stärksten ausgeprägt. Analog zur humanen pART1 besitzen fünf weitere humane pARTs das charakteristische HYE-Motiv (pART 2-6). Während das Tyrosin (Y) in β2 bei allen Mitgliedern hochkonserviert ist, finden sich an Stelle der Glutaminsäure (E) in β5 verschiedenste Varianten. So weisen zum Beispiel sechs pARTs (pART 10, 11, 12, 14, 16 und 17) an dieser Stelle ein Isoleucin (I) und pART13 ein Valin (V) auf. Interessanterweise ist auch das RT6-verwandte ART3 durch ein

Isoleucin (I) in dieser Position charakterisiert (Glowacki *et al.* 2002). Mit der Ausnahme von pART9 (Q) und humanem pART13 (Y) besitzen alle pARTs das Histidin (H) in β1, auch das Maus-Orthologe pART13. Die Motive der murinen pARTs entsprechen bis auf eine weitere Ausnahme (pART8: konservativer Austausch von Leucin (L) zu Isoleucin (I) in β5) denen ihrer humanen Orthologen.

Die "loops" zwischen den Sekundärstrukturmotiven sind weniger stark konserviert und enthalten eine sehr unterschiedliche Anzahl von Aminosäuren: So liegen bei pART1, pART2 und pART3 mehr als 30 Aminosäuren zwischen β4 und β5, während diese beiden Motive bei vielen pARTs nur durch 6 Aminosäuren getrennt sind. Bei pART5 und pART6 bzw. pART16 und pART17 ist die Größe der „loops“ identisch, so dass sich allein aufgrund der „loops“ eine enge Verwandtschaft zwischen den jeweiligen Proteinen vermuten lässt. Bei den bakteriellen Toxinen DT und ETA sind zwischen β5 und β6 zwei zusätzliche Aminosäuren eingeschoben. Die Anzahl der Aminosäuren C- und N-terminal der katalytischen Region ist in der Abbildung für jedes Protein angegeben und zeigt, dass die pARTs in ihrer Gesamtgröße stark variieren. Bei den kleineren Proteinen wie z.B. humanem pART7 bleibt allerdings unklar, ob es sich bei den bisher erfassten Sequenzen um die vollständige Proteinsequenz oder lediglich um einen Teil der Proteinsequenz handelt. Die Länge der Proteine und die Zugangsnummern sind in Abbildung 16 tabellarisch aufgeführt.

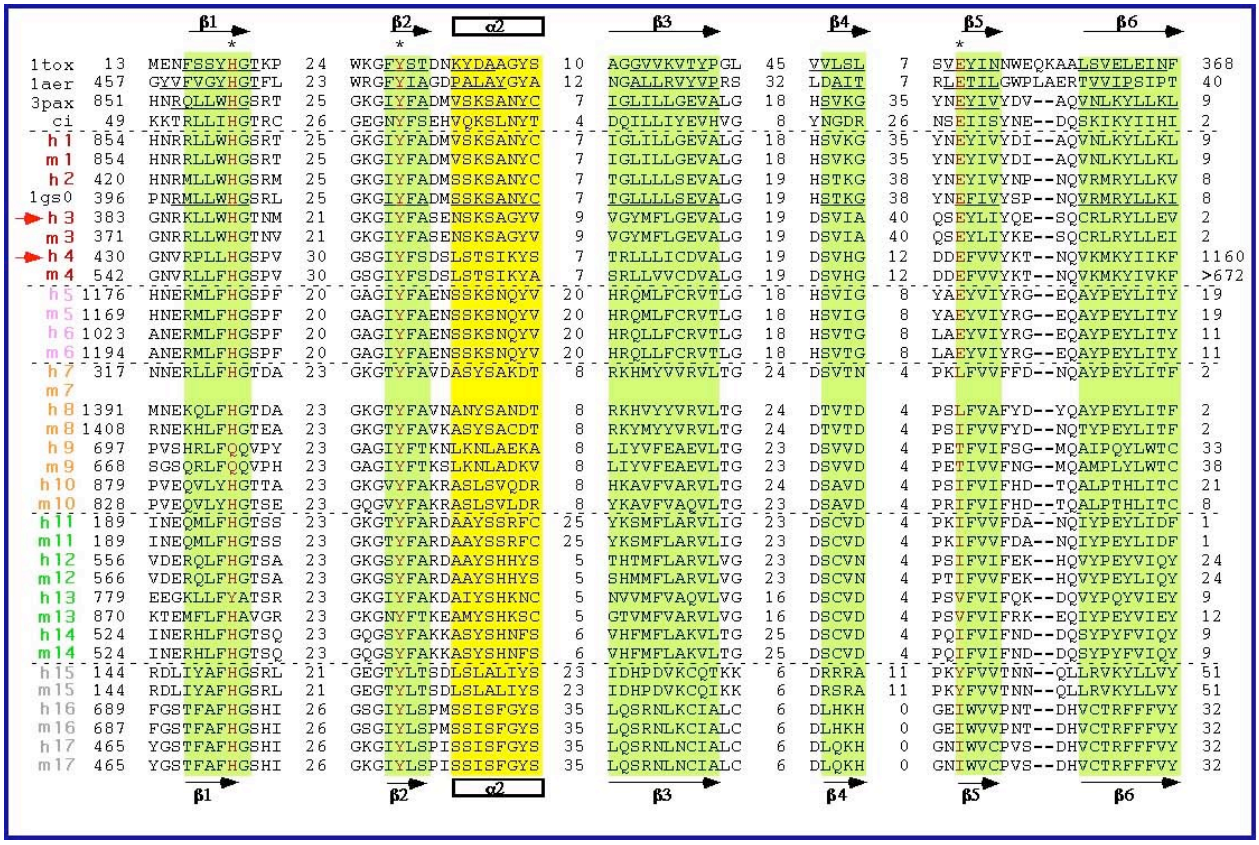


Abbildung 10: Aminosäuren-Sequenz-Alignment der humanen pART-Familie und ihrer murinen Orthologen

Die Abbildung zeigt ein Aminosäuresequenz-Alignment der katalytischen Region der humanen und murinen pART-Familienmitglieder, wobei das Alignment auf die Sekundärstrukturmotive beschränkt ist, die in den Kristallstrukturen des humanen pART1 und des Diphterietoxins das katalytische Zentrum umgeben. Diese Sekundärstrukturmotive sind unterhalb des Alignments gekennzeichnet und farbig markiert (Pfeile / grün: β-Stränge. Kasten / gelb: α-helix). Die humanen pARTs sind oben, die jeweiligen Maus-Orthologen darunter aufgeführt. In das Alignment sind zudem die Sequenzen von Diphterietoxin (DT, 1tox), *Pseudomonas aeruginosa* Exotoxin A (ETA, 1aer), Hühner-pART1 (3pax) und der putativen pART des Virus Chilo iridescent (Ci, siehe 4.1.5) aufgenommen. Für die Proteine mit bekannter 3D-Struktur sind die pdb („protein data bank“)-Zugangsnummern angegeben (s.o., mpART2 1gs0). Die Aminosäuren, die in den 3D-Strukturen die charakteristischen Sekundärstrukturmotive bilden, wurden unterstrichen. In diesen Bereichen ist die Sequenzidentität der verschiedenen pARTs am stärksten ausgeprägt. Sechs pARTs enthalten das charakteristische HYE-Motiv (die entsprechenden Positionen sind rot und mit Sternchen markiert). Die Anzahl der Aminosäuren, die die "loops" zwischen den Sekundärstrukturmotiven bilden, sowie die Anzahl der Aminosäuren C- und N-terminal der katalytischen Region sind angegeben. Die pART-Subgruppen (vgl. Abb.9) sind farblich hervorgehoben und durch gestrichelte horizontale Linien voneinander abgetrennt. Die beiden pARTs, die zur weiteren Charakterisierung ausgewählt wurden, sind durch rote Pfeile gekennzeichnet.

Wie bereits erwähnt spiegelt die Nummerierung der pARTs den Grad der Sequenzidentität zur hpART1 und der pARTs untereinander in den Bereichen mit den charakteristischen Sekundärstrukturmotiven der katalytischen Domäne wider. Für die Errechnung der Sequenzidentität mit der DNA Star Software (Programm Megalign) wurden nur die in Abbildung 10 direkt zugeordneten Regionen der einzelnen Proteine verwendet (insgesamt 66 Aminosäuren). **Abbildung 11** zeigt eine Übersicht der errechneten Aminosäuresequenzidentitäten. Die Sequenzidentität der hpART1 zu den bakteriellen Toxinen beträgt weniger als 20%. Die nächsten Verwandten hpART2 und hpART3 besitzen eine Sequenzidentität von über 50% zur hpART1. Dagegen weisen neun Familienmitglieder weniger als 30% Sequenzidentität zur hpART1 auf. Am stärksten weichen pART9 und pART15 von der hpART1 ab (jeweils 23%). Außerdem wird deutlich, dass sich einzelne pARTs in diesem Bereich sehr ähnlich sind, während andere sich ausgeprägt voneinander unterscheiden: Einerseits zeigen hpART5 und hpART6 eine Sequenzidentität von 94% bzw. hpART16 und hpART17 eine Identität von 86%, andererseits beträgt z.B. die Identität zwischen hpART8 und hpART16 lediglich 14%. Zudem lassen sich einige pARTs, wie z.B. die pARTs 11, 12, 13 und 14, einander aufgrund einer größeren Ähnlichkeit zuordnen. Dies war der erste Hinweis auf mögliche Untergruppen näher verwandter Mitglieder in der pART-Familie. Aus den Erkenntnissen dieser Analyse wurde die pART-Familie zunächst in vier Untergruppen unterteilt: Der ersten Gruppe wurden pART 1-4 (rot), der zweiten Gruppe pART5 und 6 (rosa), der dritten Gruppe pART 7-14 (orange und grün) und der vierten Gruppe pART 15-17 (grau) zugeordnet. Zur Verdeutlichung sind die potentiellen Untergruppen in der Abbildung durch verschiedene Farben hervorgehoben. Die farbliche Trennung der großen Gruppe in orange (pART7-10) und grün (pART11-14) greift den Ergebnissen weiterer Analysen zu den Subgruppen vor, die eine Auftrennung der großen Gruppe in zwei separate Untergruppen ermöglichten (siehe 4.1.4).

Der Großteil der murinen pARTs (dreizehn) weist eine Sequenzidentität von 90-100% zu dem jeweiligen humanen Orthologen auf. pART8 und 9 (82%) sowie pART 13 (70%) weichen stärker von den nächsten humanen Verwandten ab. Desweiteren wurde untersucht, welche Aminosäuren in den zugeordneten konservierten Bereichen gegeneinander ausgetauscht wurden. Häufig handelt es sich um einen konservativen Austausch, z.B. den Ersatz der hydrophoben Aminosäure Valin (V) durch das hydrophobe Isoleucin (I) oder der

basischen Aminosäure Lysin (K) durch das basische Arginin (R). Dies lässt vermuten, dass die Proteine trotz der abweichenden Aminosäuresequenz eine ähnliche 3D-Struktur ausbilden und somit auch funktionelle Gemeinsamkeiten aufweisen.

	tox	aer	ci	g01	h01	h02	h03	h04	h05	h06	h07	h08	h09	h10	h11	h12	h13	h14	h15	h16	h17	
tox	***	30	15	18	18	18	20	15	15	15	17	21	9	14	14	15	9	11	17	11	12	tox
aer	30	***	18	17	17	20	20	18	17	17	15	14	6	20	15	17	12	12	15	9	9	aer
ci	15	18	***	36	38	36	32	36	30	32	26	26	15	21	21	27	24	27	18	17	15	ci
g01	18	17	36	***	97	79	56	47	47	44	33	29	23	26	33	27	26	26	23	26	26	g01
h01	18	17	38	97	***	79	56	49	49	46	35	29	23	24	32	29	26	27	23	26	26	h01
m01	18	17	38	97	100	79	56	49	49	46	35	29	23	24	32	29	26	27	23	26	26	m01
h02	18	20	36	79	79	***	58	50	47	46	33	27	21	24	32	29	26	27	23	30	29	h02
m02	17	21	36	76	76	92	53	52	44	44	35	29	26	27	33	30	27	27	24	29	29	m02
h03	20	20	32	56	56	58	***	36	44	41	36	33	29	32	33	35	36	33	21	23	24	h03
m03	20	20	35	56	58	55	95	41	46	42	38	32	29	32	33	36	33	35	21	23	24	m03
h04	15	18	36	47	49	50	36	***	46	47	38	29	26	26	32	33	27	30	24	30	26	h04
m04	14	17	32	46	47	47	36	91	44	46	39	29	29	26	32	30	29	29	26	30	26	m04
h05	15	17	30	47	49	47	44	46	***	94	46	41	35	38	39	41	32	38	21	24	23	h05
m05	15	17	30	47	49	48	44	46	100	94	46	41	35	38	39	41	32	38	21	24	23	m05
h06	15	17	32	44	46	46	41	47	94	***	46	42	35	38	38	39	30	36	21	24	23	h06
m06	15	17	32	44	46	46	41	47	94	100	46	42	35	38	38	39	30	36	21	23	23	m06
h07	17	15	26	33	35	33	36	38	46	46	***	79	36	55	62	55	47	50	29	17	17	h07
h08	21	14	26	29	29	27	33	29	41	42	79	***	39	55	55	50	42	47	21	14	15	h08
m08	17	15	24	30	30	30	38	33	41	41	79	82	36	55	61	52	44	53	24	18	18	m08
h09	09	06	15	23	23	21	29	26	35	35	36	39	***	46	35	33	39	36	18	15	14	h09
m09	07	05	18	26	24	24	29	27	33	33	36	36	82	41	36	30	35	35	20	20	15	m09
h10	14	20	21	26	24	24	32	26	38	38	55	55	46	***	52	50	46	52	20	15	17	h10
m10	12	20	21	24	23	23	30	26	33	33	50	50	46	91	52	47	47	55	20	15	15	m10
h11	14	15	21	33	32	32	33	32	39	38	62	55	35	52	***	64	53	59	24	20	20	h11
m11	14	15	21	33	32	32	33	32	39	38	62	55	35	52	100	64	53	59	24	20	20	m11
h12	15	17	27	27	29	29	35	33	41	39	55	50	33	50	64	***	62	67	24	23	24	h12
m12	15	17	26	29	30	29	32	33	39	38	56	49	32	47	65	94	59	65	26	24	24	m12
h13	9	12	24	26	26	26	36	27	32	30	47	42	39	46	53	62	***	56	18	17	18	h13
m13	11	7	24	21	21	23	30	26	38	36	47	44	39	47	55	61	70	53	20	15	17	m13
h14	11	12	27	26	27	27	33	30	38	36	50	47	36	52	59	67	56	***	21	26	24	h14
m14	11	12	27	26	27	27	33	30	38	36	50	47	36	52	59	67	56	100	21	26	24	m14
h15	17	15	18	23	23	23	21	24	21	21	29	21	18	20	24	24	18	21	***	30	26	h15
m15	17	15	18	23	23	23	21	24	20	20	29	21	18	20	24	24	18	21	97	30	26	m15
h16	11	9	17	26	26	30	23	30	24	24	17	14	15	15	20	23	17	26	30	***	86	h16
m16	11	9	17	26	26	30	23	30	23	23	17	14	15	15	20	23	17	26	30	100	86	m16
h17	12	9	15	26	26	29	24	26	23	23	17	15	14	17	20	24	18	24	26	86	***	h17
m17	12	9	15	26	26	29	24	26	23	23	17	15	14	17	20	24	18	24	26	86	100	m17
	tox	aer	ci	g01	h01	h02	h03	h04	h05	h06	h07	h08	h09	h10	h11	h12	h13	h14	h15	h16	h17	

Abbildung 11: Aminosäuresequenzidentität der humanen und murinen pARTs im Bereich der konservierten Sekundärstrukturmotive der katalytischen Domäne

Die Tabelle zeigt die prozentuale Aminosäuresequenzidentität der einzelnen Mitglieder der humanen und murinen pART-Familie im Bereich der Sekundärstrukturmotive, die einander in Abb.10 mit Alignment-Programmen zugeordnet wurden. Zur Erstellung der Abbildung wurden ausschließlich die Aminosäuren dieser zugeordneten Bereiche aneinandergefügt (insgesamt 66 Aminosäuren für jedes Protein) und die Sequenzidentität mit dem Programm Megalign (DNA-Star-Software) errechnet. Die Subgruppen sind farblich hervorgehoben. Die prozentuale Sequenzidentität zu den bakteriellen Toxinen DT (tox) und ETA (aer), zur Hühner-pART1 (g01) und zur putativen pART des Chilo iridescent Virus (Ci) ist ebenfalls aufgeführt.

4.1.3 Chromosomale Lokalisierung der humanen und murinen *pARTs*

Um die chromosomale Lokalisierung der humanen *pART*-Gene und ihrer murinen Orthologen zu ermitteln, wurden mit den vorhandenen mRNAs bzw. den aus den ESTs erstellten, konsensuellen cDNA-Sequenzen der verschiedenen pARTs als Eingabe BLASTn-Suchen gegen das humane bzw. murine Genom durchgeführt. Nach Auffinden der entsprechenden genomischen Sequenz wurde über den „Genome View"-Link die chromosomale Lage des jeweiligen *pART*-Genes dargestellt. Um herauszufinden, ob *pArt7* und *pArt13* trotz fehlender Sequenzen in der „nr"-Sektion der Datenbank im Mausgenom vorkommen, wurde zudem eine tBLASTn-Suche mit den humanen Orthologen dieser pARTs gegen das Mausgenom durchgeführt. Mit der tBLASTn-Suche gelang der Nachweis des fehlenden *pART13*-Orthologen auf dem Maus-Chromosom 6. Die aus der genomischen Sequenz übersetzte Aminosäuresequenz entsprach unserer aus den ESTs erstellten Konsensussequenz für die katalytische Domäne. Für murines *pArt13* fehlt somit lediglich eine Sequenz in der „nr"-Datenbank, die die katalytische Domäne einschließt. Dagegen konnte auch mit dieser Methode kein *pART7*-Orthologes bei der Maus aufgespürt werden. **Abbildung 12** zeigt eine schematische Darstellung der Lage der *pART*-Gene auf den humanen (**A**) und murinen (**B**) Chromosomen. Die genaue Bezeichnung der entsprechenden Banden ist zudem in der Abbildung 16 angegeben.

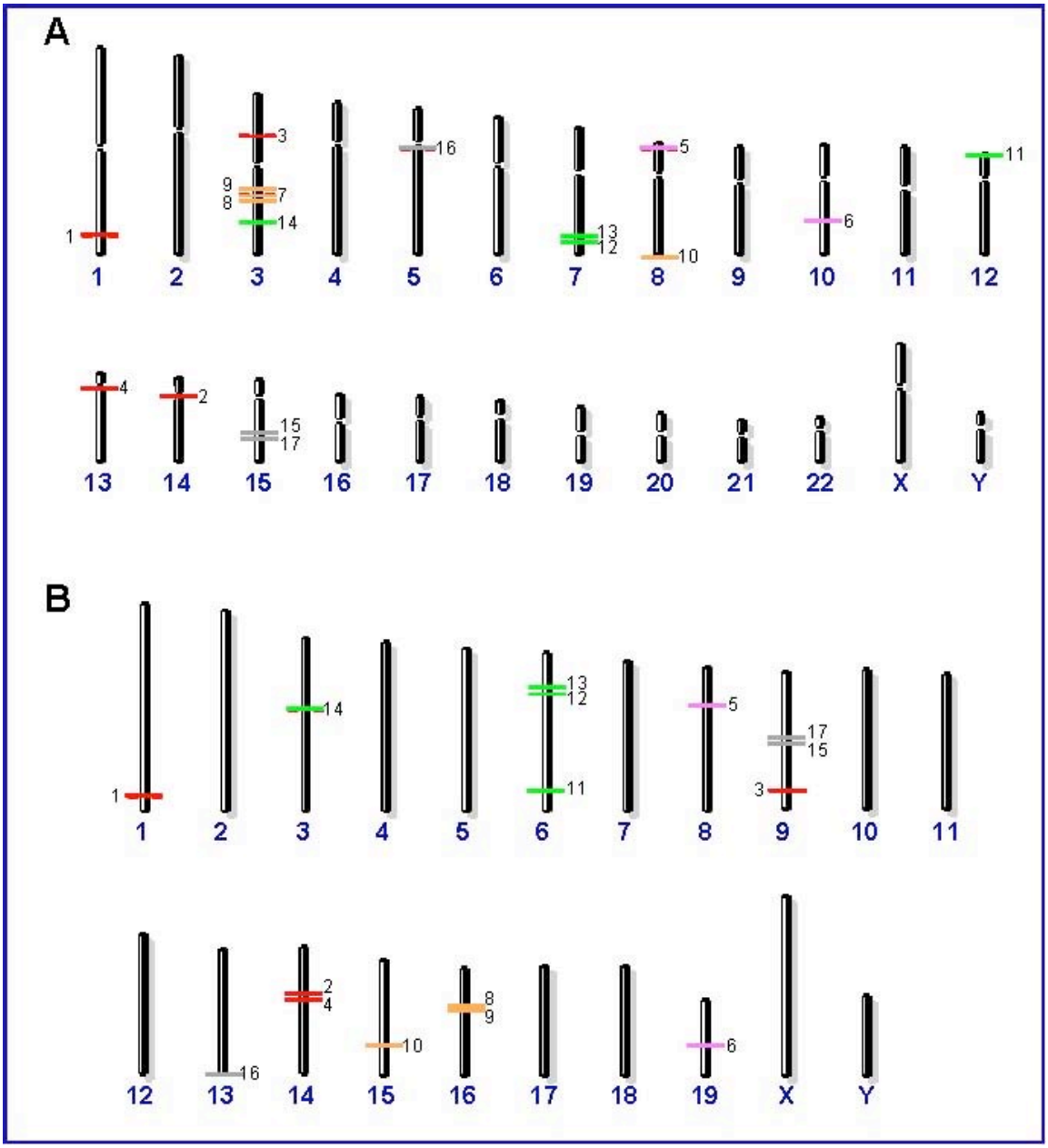


Abbildung 12: Schematische Darstellung der chromosomalen Lokalisierung der humanen (A) und murinen (B) *pART*-Gene

Die Abbildung zeigt den Chromosomensatz des humanen (**A**, 22 homologe Autosomenpaare und 2 Gonosomen) bzw. murinen (**B**, 19 homologe Autosomenpaare und 2 Gonosomen) Genoms, wobei für jedes Autosomenpaar nur jeweils ein Chromosom dargestellt ist. Die Lage der *pART*-Gene auf den Chromosomen ist mit Balken markiert, die den pART-Subgruppen entsprechend farblich markiert wurden: Subgruppe 1 rot, 2 rosa, 3 orange, 4 grün und 5 grau. Das Orthologe von *pART7* fehlt im Maus-Genom. Die Bezeichnung der Chromosomenbanden, in denen die Gene liegen, ist in Abbildung 16 tabellarisch aufgeführt.

Die *pART*-Gene kommen auf einzelnen Autosomen gehäuft vor, während auf anderen Autosomen und den Gonosomen keine *pART*-Gene lokalisiert sind. Auf dem humanen Chromosom 3 liegen fünf *pART*-Gene, vier auf dem langen Arm (q) und eines auf dem kurzen Arm (p), wobei sich drei Gene (*pART7*, *8* und *9*) eng beieinander im Abschnitt 3q21 befinden. In enger Nachbarschaft lokalisierte *pART*s existieren ebenfalls auf den humanen Chromosomen 7 (*pART12* und *13*) und 15 (*pART15* und *17*). Da diese Gene, abgesehen vom fehlenden *pArt7*, auch bei der Maus eng beieinander liegen, ist davon auszugehen, dass diese Gene innerhalb einer sog. Kopplungsgruppe liegen. Der Begriff der Kopplung reicht zurück bis zu frühen Untersuchungen an *Drosophila melanogaster*, in denen gezeigt wurde, dass auf einem Chromosom nahe zusammen liegende Gene häufig in Gruppen (gekoppelt) vererbt werden. Bei der Maus sind zudem noch *pArt2* und *pArt4* (Chr.14) benachbart, die menschlichen Orthologen liegen aber auf verschiedenen Chromosomen. Die eng benachbarten Gene gehören immer derselben, von uns anhand der Sequenzidentitäten eingeteilten Subgruppe an. Es ist somit wahrscheinlich, dass diese Gene durch regionale Genduplikation entstanden sind. Interessanterweise liegen jedoch die einander am nächsten Verwandten pARTs (pART 5 und 6 bzw. pART 16 und 17) auf verschiedenen Chromosomen.

Um die genaue Lage der *pART*-Gene *7*, *8* und *9* auf dem humanen Chromosom 3 zu untersuchen und weitere Hinweise für die Kopplung dieser Region zu finden, wurden die Region q21 des humanen Chromosoms 3 und die Region B3 des murinen Chromosoms 16 mit Hilfe des „UCSC Genome browsers“ genauer analysiert und miteinander verglichen. Dieser "Browser" ermöglicht eine Darstellung der Lage der bekannten Gene (basierend auf Swiss-Prot, TrEmbl, mRNA und Ref Seq) zusammen mit zugehörigen mRNAs und ESTs in bestimmten chromosomalen Bereichen. **Abbildung 13** zeigt eine schematische Darstellung der untersuchten chromosomalen Regionen.

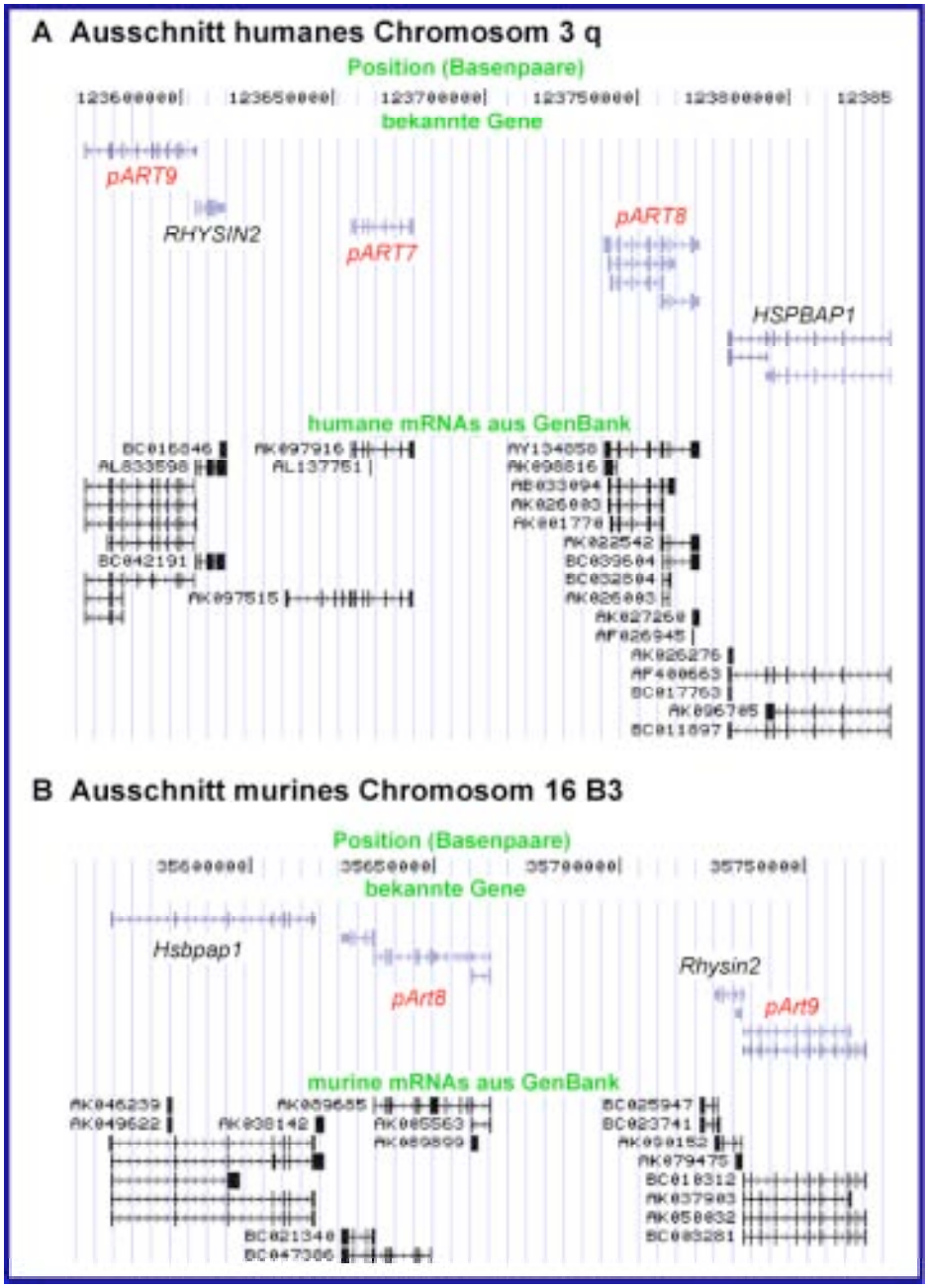


Abbildung 13: Vergleich der pART kodierenden Regionen auf dem humanen Chromosom 3q und dem Maus-Chromosom 16B3

Mit Hilfe des "UCSC genome browsers" wurden das humane Chromosom im Bereich der Region 3q (**A**) und das murine Chromosom 16 im Abschnitt B3 (**B**) genauer analysiert. In der Abbildung sind die analysierten Regionen schematisiert dargestellt. Oben ist die Lokalisation auf dem Chromosom in Basenpaaren angegeben. Die bekannten Gene sind als waagrechte Linien, die Exons als senkrechte Balken dargestellt, wobei die Breite der Balken bei höheren Auflösungen („zoom in“) der Größe der Exons entspricht und die Höhe der Balken eine Unterscheidung zwischen kodierenden (hoch) und 3`-bzw. 5`-gelegenen, nicht-kodierenden Abschnitten (halbhoch) ermöglicht. Die Pfeile zeigen die Transkriptionsrichtung des jeweiligen Genes an. Unterhalb der Gene sind entsprechende mRNAs aus der GenBank mit ihren Zugangsnummern aufgeführt. Bei mehreren der schematisiert angegebenen Sequenzen handelt es sich um partielle mRNAs, die in einigen Fällen lediglich Transkripte eines Exons beinhalten (breite Balken ohne waagrechte Linie). HSPBAP1 = Hitze-Schock 27kD assoziiertes Protein 1.

Die Abbildung zeigt, dass die bekannten Gene in den untersuchten Bereichen durch mehrere mRNAs in ihrer Exon-Intron-Struktur abgesichert sind. Es wird aber z.B. am *pART8*-Gen deutlich, dass die Größe des Genes abhängig von der Einbeziehung bestimmter mRNAs und der angewandten Methode unterschiedlich vorhergesagt wird (**A**). So werden für den *hpART8*-Locus vier verschiedene Varianten des Gens angegeben. Während beim *hpART8* keine mRNA über den vorhergesagten 5´- oder 3´-Bereich hinausragt, zeigt die mRNA AK097515 beim *hpART7*, dass sich das Gen möglicherweise weiter 5`-wärts fortsetzt als bisher angenommen. Dies verdeutlicht, dass sich die genaue Vorhersage der Struktur und Größe eines Genes mit den 5`- und 3`-Enden häufig schwierig gestaltet. Das 3`-Ende des „offenen Leserasters“ ist in der Regel durch das Stop-Codon gut definiert. Der polyA-Schwanz ist zumeist im letzten kodierenden Exon lokalisiert. Die Definition des 5`-Endes gestaltet sich schwieriger. Das offene Leseraster wird „rückwärts“ bis zum Auftreten eines Abbruches des Leserasters verfolgt („in-frame STOP“) und das erste in Leserichtung anschließende Methionin als Startcodon interpretiert. Unterschiedliche Vorhersagen für die Genstruktur können zudem durch alternative Splice-Varianten oder Sequenzierfehler bedingt sein, die möglicherweise zu einem vorzeitigen Abbruch des Leserasters führen. Deshalb ist darauf zu achten, dass die Vorhersagen möglichst durch verschiedene ESTs und mRNAs abgesichert werden.

Beim Menschen liegen die Gene *pART7* und *pART8* in gleicher Orientierung hintereinander. 5`-wärts des *pART7* liegt das Gen *RHYSIN 2*, dessen erst kürzlich identifiziertes Proteinprodukt in Komplexen mit Myosin VI immunpräzipitiert werden kann. *RHYSIN 2* grenzt in einer „Kopf an Kopf (head to head)“-Orientierung an das *pART9*-Gen. 3‘-wärts des *pART8* liegt auf dem komplementären DNA-Strang das Gen für das Protein *HSPBAP1* (Hitze-Schock 27kD assoziiertes Protein 1), das an das kleine Hitze-Schock-Protein hsp27 binden kann. Interessanterweise stellen sich die 5´- und 3´-Randbereiche dieser Region, wenn auch umgekehrt orientiert, auf dem Maus-Chromosom 16 sehr ähnlich dar (**B**). Das 5`-wärts des *mpArt9* auf dem komplementären Strang gelegene, bisher unbenannte Gen konnte über Sequenzvergleiche der mRNAs von Maus und Mensch als das murine *Rhysin 2*-Orthologe identifiziert werden. Weiter 5`-wärts folgt das *mpArt8*-Gen, wobei die Teile des Gens noch nicht als zusammengehörig erkannt sind. Die überlappende mRNA BC047386 zeigt aber, dass vorderer Teil und hinterer Teil, der die potentielle katalytische Domäne kodiert, zu einem Gen gehören. Im Anschluss folgt analog zum Menschen in einer „Schwanz an

Schwanz“-Orientierung das murine *Hsbpap1*. Das beim Menschen zwischen dem *RHYSIN 2*- und dem *pART8*-Gen gelegene *pART7* fehlt. Aus diesem Abschnitt zwischen den Genen sind weder mRNAs noch ESTs in der GenBank vorhanden.

In der „nr“-Sektion der Datenbank konnte den jeweiligen pARTs mit Ausnahme von pART7, 11 und 13 ein orthologes Ratten-Protein zugeordnet werden. Mit zusätzlichen tBLASTn-Suchen gegen das Rattengenom wurden zwar die in der „nr“-Datenbank fehlenden Ratten-Orthologen von *pART11* und *13* aufgespürt, analog zur Maus fehlt bei der Ratte aber das *pART7*-Orthologe. Zusammenfassend wurden somit bei beiden Nagern sechszehn paraloge, zu den humanen Orthologen hoch konservierte Mitglieder der pART-Familie nachgewiesen .

4.1.4 *In silico* Analyse und Vergleich der Genstrukturen der humanen *pART*-Paralogen

Die meisten Gene höherer Eukaryonten sind in kodierende und nicht kodierende Abschnitte gespalten, während niedere Eukaryonten wie Hefen einen weit größeren Anteil an kontinuierlichen Genen enthalten und bei Eubakterien wie z.B. *E.Coli* überhaupt keine unterbrochenen Gene zu finden sind. Durch Vergleiche der DNA-Sequenzen von Genen, die evolutionsgeschichtlich stark konservierte Proteine kodieren, wurde zunächst die These postuliert, dass in den Urgenen Introns vorhanden waren, die im Laufe der Evolution verloren gingen und zwar besonders ausgeprägt bei Organismen, die für rasches Wachstum optimiert wurden (wie z.B. Eubakterien und Hefen). Anfang der 90er Jahre ergaben sich zunehmend Hinweise, dass einige Introns erst später in der Evolution erworben wurden. Dies führte zu einer kontroversen Debatte, in der einige Forschergruppen die Ansicht vertraten, dass alle Introns erst im Laufe der Evolution in die Gene „eingesprungen“ sind (z.B. durch Transposition). Neue Erkenntnisse weisen auf ein gemischtes Modell aus sehr alten Introns und erst kürzlich erworbenen Introns hin (Roy 2003), durch das viele Beobachtungen der letzten Jahre erklärt werden können. Desweiteren drängte sich in den letzten Jahren die Frage auf, ob durch bestimmte Introns gezielt die Regionen getrennt werden, die verschiedene Tertiärstrukturen bzw. Module in den jeweiligen Proteinen kodieren. Es werden verschiedene Arten von Introns unterschieden: Introns, die sich genau zwischen zwei kodierenden Tripletts befinden (Phase-0-Introns), und Introns die innerhalb eines kodierenden Tripletts liegen (Phase-1-Introns: hinter dem ersten Nukleotid des Tripletts, Phase-2-Introns: hinter

dem zweiten Nukleotid des Tripletts). Fedorov und Mitarbeiter zeigten, dass in vielen Genen, die Verwandte in evolutionär weit entfernten Spezies besitzen, eine Korrelation der Phase-0-Introns mit den Regionen besteht, die Module miteinander verbinden (Fedorov *et al.* 2001). Desweiteren zeigt sich für Phase-1-Introns eine Antikorrelation zu diesen Regionen und stattdessen eine Präferenz für Glycin-kodierende Tripletts innerhalb der Module. Trotz vieler ungeklärter Fragen ermöglicht der Vergleich der Exon-Intron-Struktur Aussagen über den Grad der Verwandtschaft von Proteinen und bietet Hinweise auf mögliche gemeinsame Vorläuferproteine.

Um Aufschluss über die genomische Struktur der verschiedenen humanen *pART*s zu bekommen, wurden weitere Datenbankrecherchen durchgeführt. Zur Abgrenzung der Region des Genes, die die katalytische Domäne kodiert, wurden die Aminosäuresequenzen der humanen pARTs mittels Sequenz-Alignment-Programmen mit der minimalen katalytischen Domäne der putativen pART des Chilo iridescent Virus verglichen. Anschließend setzten wir die Nukleotidsequenzen dieser Region ein, um mit BLASTn die humane genomische Datenbank ("human genome resources", NCBI) zu durchsuchen. Die so ermittelten Alignments zwischen cDNA-Sequenz und genomischer Sequenz brechen an Stellen ab, an denen in der genomischen Sequenz Introns eingefügt sind. Die Kenntnis der charakteristischen Nukleotide zu Beginn und am Ende eines Introns bzw. Exons eröffnet die Möglichkeit, die genaue Lage und Größe des Introns zu bestimmen.

Mit dieser Methode gelang es, die Exon-Intron-Struktur im Bereich der katalytischen Region für alle humanen *pART*s *in silico* zu entschlüsseln. Für das zur weiteren Analyse im zweiten Abschnitt ausgewählte *hpART3* korrelieren die *in silico* Ergebnisse mit den durch molekularbiologische Techniken erzielten Ergebnissen zur Exon-Intron-Struktur (siehe 4.2.5). Abschließend nutzten wir die erst seit kurzem verfügbaren, modernen Datenbank-Verknüpfungen auf der NCBI-Homepage wie „Locus link“, „Ensemble“ und „Gene“, um unsere experimentell erzielten Ergebnisse mit computergestützt vorhergesagten Genstrukturen zu vergleichen. Die dort vorhergesagten Exon-/ Intron-Strukturen bestätigten unsere Ergebnisse.

Abbildung 14 zeigt das Aminosäuren-Alignment der potentiellen katalytischen Domänen der humanen pART-Paralogen, in das die Positionen der Introns eingezeichnet sind. Konservierte Intronpositionen zwischen verschiedenen pART-Paralogen sind durch Pfeile markiert und im Alignment farblich hervorgehoben. Zur Optimierung des Alignments der siebzehn humanen pARTs wurde zunächst für jede einzelne Subgruppe ein Alignment erstellt, bei denen auch die mittels "PSIPRED" und „GeneTHREADER" vorhergesagten Sekundärstrukturmotive berücksichtigt wurden. Die Alignments der einzelnen Subgruppen mit den Sekundärstrukturvorhersagen sind in Anhang D abgebildet. β1, β2, β5 und β6 konnten bei allen Subgruppen problemlos zugeordnet werden. Für β3 und insbesondere den kurzen β4-Strang war dies bei den Subgruppen 3-5 deutlich schwieriger, so dass das Alignment erst unter Berücksichtigung der vorhergesagten Sekundärstrukturen gelang. Für die Subgruppe 5 konnte β4 nicht mit ausreichender Sicherheit vorhergesagt werden. Die zwischen den charakteristischen Sekundärstrukturmotiven gelegenen Regionen („loops") zeigen nur innerhalb der Subgruppen eine signifikante Sequenzidentität, so dass sich diese Regionen wahrscheinlich nur bei den nah verwandten pARTs ähnlich falten. Besonders der „loop" zwischen β4 und β5 unterscheidet sich in Länge und Aminosäuresequenz erheblich zwischen den Subgruppen. Bei den pARTs der Subgruppen 3 und 4 umfaßt er nur 6 AS, während er bei pART3 42 AS lang ist. Eine mögliche Erklärung liefert die Betrachtung der Kristallstruktur der hpART1 (siehe Abb.5): Die hochkonservierten β-Stränge 1, 2 und 5 liegen zentral und bilden das charakteristische Aminosäurenmotiv aus. β3, β4 und β6 und insbesondere der die β-Stränge 4 und 5 verbindende „loop" liegen weiter außerhalb. Im Gegensatz zu einer Abwandlung der zentralen β-Stränge hat eine schwächere Konservierung dieser β-Stränge sowie eine Verkürzung des „loops" für die katalytische Funktion möglicherweise nur eine geringe Bedeutung, so dass größere Variationen entstehen konnten. Die Abweichungen in dieser Region könnten vielmehr eine unterschiedliche Substratspezifität ermöglichen.

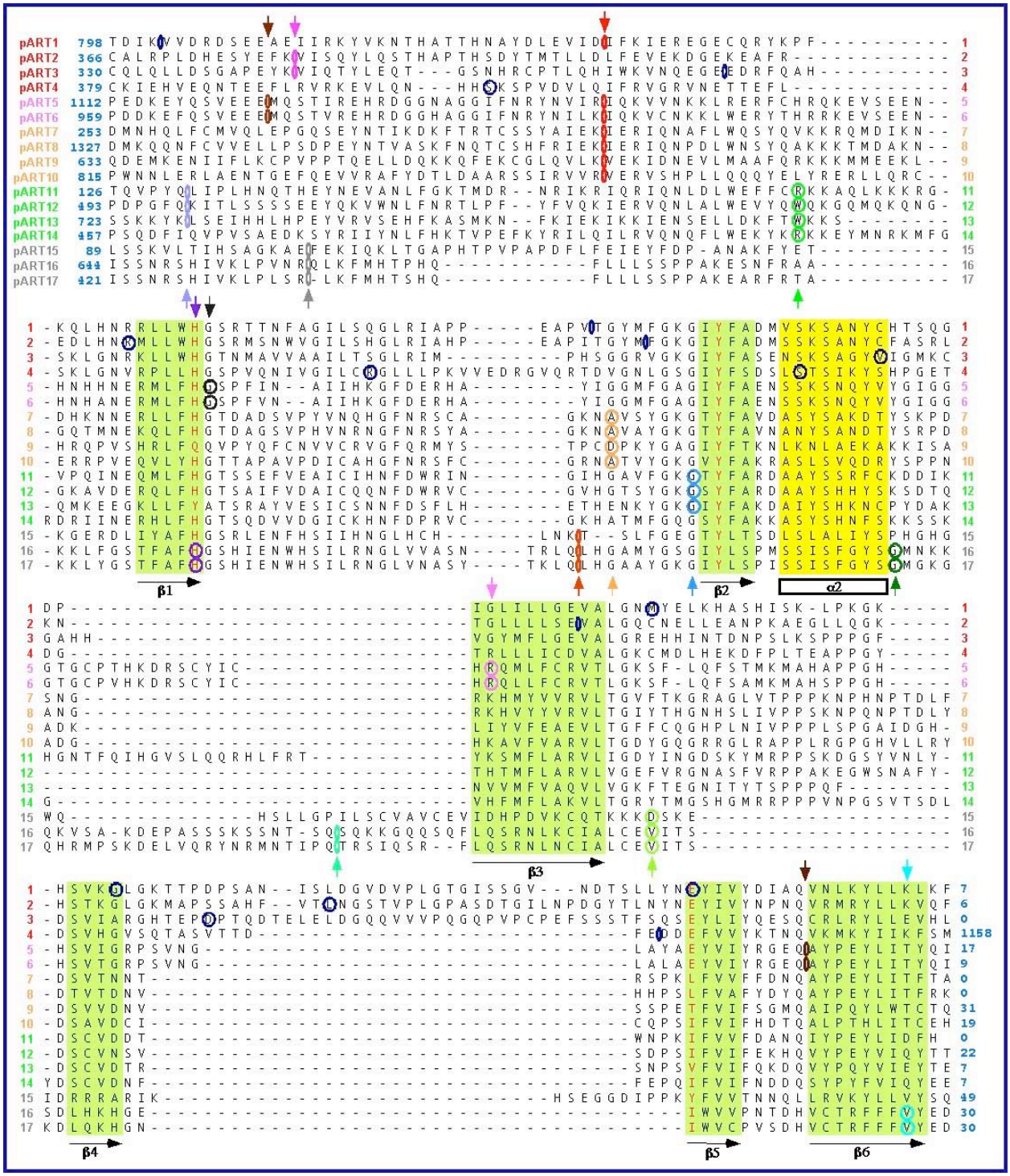


Abbildung 14: Konservierte Aminosäuren, Sekundärstrukturmotive und Lage der Introns in der potentiellen katalytischen Domäne der humanen pART-Paralogen

Die Abbildung zeigt ein mit Hilfe von Sequenz-Alignment-Programmen (T-Coffee, Megalign) erstelltes, sowie unter der Berücksichtigung von PSI-BLAST-Ergebnissen und Sekundärstrukturvorhersagen (PSIPRED, GeneTHREADER) optimiertes Alignment der potentiellen katalytischen Domäne der humanen pART-Paralogen. Die Lage der Introns ist in die Abbildung eingezeichnet, wobei Phase-0-Introns als Ovale zwischen den Aminosäuren und Phase-1- bzw. Phase-2-Introns als Kreise um die jeweilige Aminosäure dargestellt sind. Zwischen verschiedenen pART-Paralogen konservierte Intronpositionen sind farblich hervorgehoben und mit

Pfeilen gekennzeichnet. Die nichtkonservierten Introns sind dunkelblau dargestellt. Die Anzahl der Aminosäuren der einzelnen Proteine N-terminal und C-terminal des Alignments ist angegeben. Die Positionen der bekannten Sekundärstrukturmotive sind analog zur Abbildung 10 eingefärbt und unterhalb des Alignments markiert. Das HYE-Motiv bzw. die bei anderen Paralogen an diesen Positionen liegenden Aminosäuren sind rot hervorgehoben. Die pARTs sind entsprechend ihrer Subgruppe farblich markiert. Die Sequenzen sind in der Datenbank unter den in Abbildung 16 angegebenen Zugangsnummern deponiert.

Die genomische Struktur der für die potentielle katalytische Domäne kodierenden Regionen (**A**) sowie eine Übersicht der konservierten Introns (**B**) ist in **Abbildung 15** schematisch dargestellt. Bei den meisten *hpART*s liegt die für die katalytische Domäne kodierende Region am 3`-Ende des Genes. Lediglich bei *pART4* setzt sich der kodierende Bereich des Genes stromabwärts der katalytischen Region fort. Bei den *pARTs 1-6* und *pART15* wird die katalytische Region durch vier bis fünf Exons kodiert, wobei die Basentripletts für die charakteristischen drei Aminosäuren HYE, bzw. im Falle von *pART15* HYY, in unterschiedlichen Exons liegen. Nur bei *pART14* werden die Regionen der charakteristischen Sekundärstruktur-Motive von einem einzigen Exon kodiert. Im Aminosäuren-Alignment fällt auf, dass die Aminosäuresequenzidentität zwischen den verschiedenen Subgruppen stromaufwärts von β 1 abbricht. Bemerkenswerterweise korreliert der Verlust der Sequenzidentität in diesem Bereich mit dem Vorkommen von Phase-0-Introns in fast allen pARTs. Das stromaufwärts des für β1 kodierenden Exons gelegene Phase-0-Intron von pART1 ist bei den *pARTs 5-10* konserviert (roter Pfeil), so dass hier der Beginn der Region festgelegt wurde, die die pART-katalytische Domäne kodiert.

Die Lage der weiteren konservierten Introns bestätigt im wesentlichen die von uns anhand der Sequenzidentität postulierte Subgruppeneinteilung. Die große Gruppe mit *pART 7-14* konnte anhand der konservierten Intronlagen in zwei Subgruppen aufgetrennt werden. Den fünf Subgruppen wurden folgende pARTs zugeordnet:

Gruppe 1 (rot): *hpART 1-4*

Die Region, die die katalytische Domäne kodiert, wird in dieser Gruppe durch drei (*pART3* und *4*) bzw. vier (*pART1* und *2*) Introns unterbrochen. Überraschenderweise ist trotz der hohen Sequenzidentität der Gruppenmitglieder (einschließlich des HYE-Motives) keine dieser 14 Intronpositionen in der katalytischen Region konserviert. Lediglich *pART2* und *pART3* verfügen über eine konservierte Intronstellung (Phase-0-Intron) kurz vor dem als katalytische Region definierten Bereich (pink).

Gruppe 2 (rosa): *pART 5* und *6*

Bei den beiden *pART*s sind neben dem zur *pART1* konservierten Intron auch die drei weiteren Intronpositionen im Bereich der katalytischen Region konserviert. Beide besitzen das HYE-Motiv.

Gruppe 3 (orange): *pART 7-10*

Das Phase-0-Intron vor der katalytischen Region ist zur hpART1 konserviert. Die Untergruppe ist durch eine zusätzliche konservierte Phase-1-Intronposition vor dem β2-kodierenden Bereich gekennzeichnet (orange). Die Gruppe zeichnet sich durch die größte Variation des HYE-Motives aus: HYL bei *pART7* und *8*, QYT bei *pART9* und HYI bei *pART10*.

Gruppe 4 (grün): *pART 11-14*

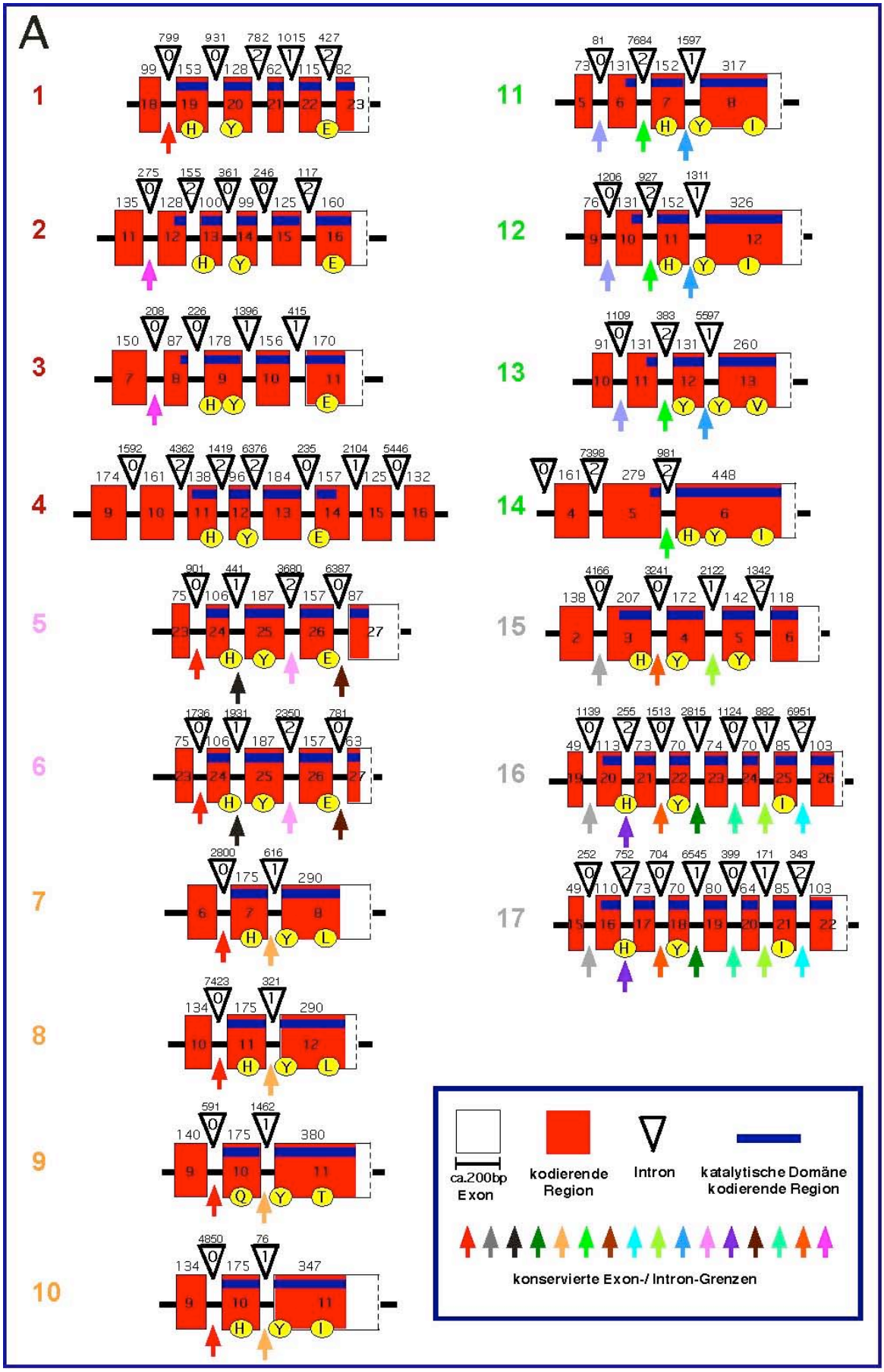
Die Mitglieder dieser Gruppe besitzen ebenfalls ein konserviertes Intron (Phase-2) vor dem Exon, das für β1 kodiert (grüner Pfeil). Zudem besitzen drei Mitglieder (*pART11-13*) stromaufwärts ein konserviertes Phase-0-Intron. pART14 ist die einzige humane pART, bei der das nächste, 5`-wärts gelegene Phase-0-Intron deutlich von der β1-kodierenden Region entfernt ist. Während bei *pART14* der gesamte, für die charakteristischen Sekundärstrukturen kodierende Bereich in einem Exon liegt, besitzen die anderen *pART*s der Gruppe 4 vor dem Triplett für das Tyrosin (Y) einen weiteren konservierten Introneinschub. Auffällig ist außerdem, dass bei *pART11*, *pART12* und *pART14* statt des HYE-Motivs ein HYI-Motiv besteht.

Gruppe 5 (grau): *pART 15-17*

Diese Gruppe zeichnet sich ebenfalls durch ein konserviertes Phase-0-Intron vor dem β1-kodierenden Bereich aus (grau). Die stark sequenzhomologen *pART16* und *17* sind in viele kleine Exons aufgeteilt und weisen multiple konservierte Exon-Intron-Grenzen auf, davon sind die Intronpositionen vor den β2- bzw. β5-kodierenden Bereichen auch bei *pART15* konserviert. *pART16* und *17* besitzen ein HYI-Motiv. Bei *pART15* ist das Motiv zu HYY abgewandelt.

Die konservierten Intronpositionen machen es wahrscheinlich, dass sich die jeweiligen *pART*s im Laufe der Evolution aus einem gemeinsamen Vorläufergen entwickelt haben. Auffällig ist, dass die humanen *pARTs 1* und *3* sowie alle *pARTs* der Gruppen 2, 3 und 5 ein Phase-0-Intron direkt vor dem β1-kodierenden Bereich besitzen. Analysen weiter 5´-wärts gelegener

Intronpositionen zeigten, dass auch die *pART*s *2-4* und *11-13* ein Phase-0-Intron vor dem Bereich aufweisen, der die katalytische Domäne kodiert, wobei das Intron bei *pART4* deutlich weiter entfernt liegt. Mit Ausnahme von *pART4* und *pART14* besitzen folglich alle *pART*s ein Phase-0-Intron 26-64 Codons stromaufwärts des kodierten Histidins in β1. Dies weist darauf hin, dass das „pART-katalytische“ Modul durch diese, zum Teil konservierten Phase-0-Introns von anderen funktionellen Modulen der kodierten Proteine abgetrennt wird. Drei der fünf konservierten Phase-1-Introns liegen in Glycin-kodierenden Tripletts. Interessanterweise bestehen zwischen der *pART1* und den nächsten Verwandten *pART2*, *pART3* und *pART4* keine konservierten Exon-Intron-Grenzen im Bereich der für die pART-Domäne kodierenden Region. Da *pART1* aber eine konserviertes Intron zu den Mitgliedern anderer Gruppen besitzt, ist die Gruppenzuordnung der *pART1* nicht eindeutig. Aufgrund der hohen Sequenzidentität wurde *pART1* aber weiterhin der Gruppe 1 zugeordnet.


A
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
ca.200bp Exon
kodierende Region
Intron
katalytische Domäne kodierende Region
konservierte Exon-/ Intron-Grenzen

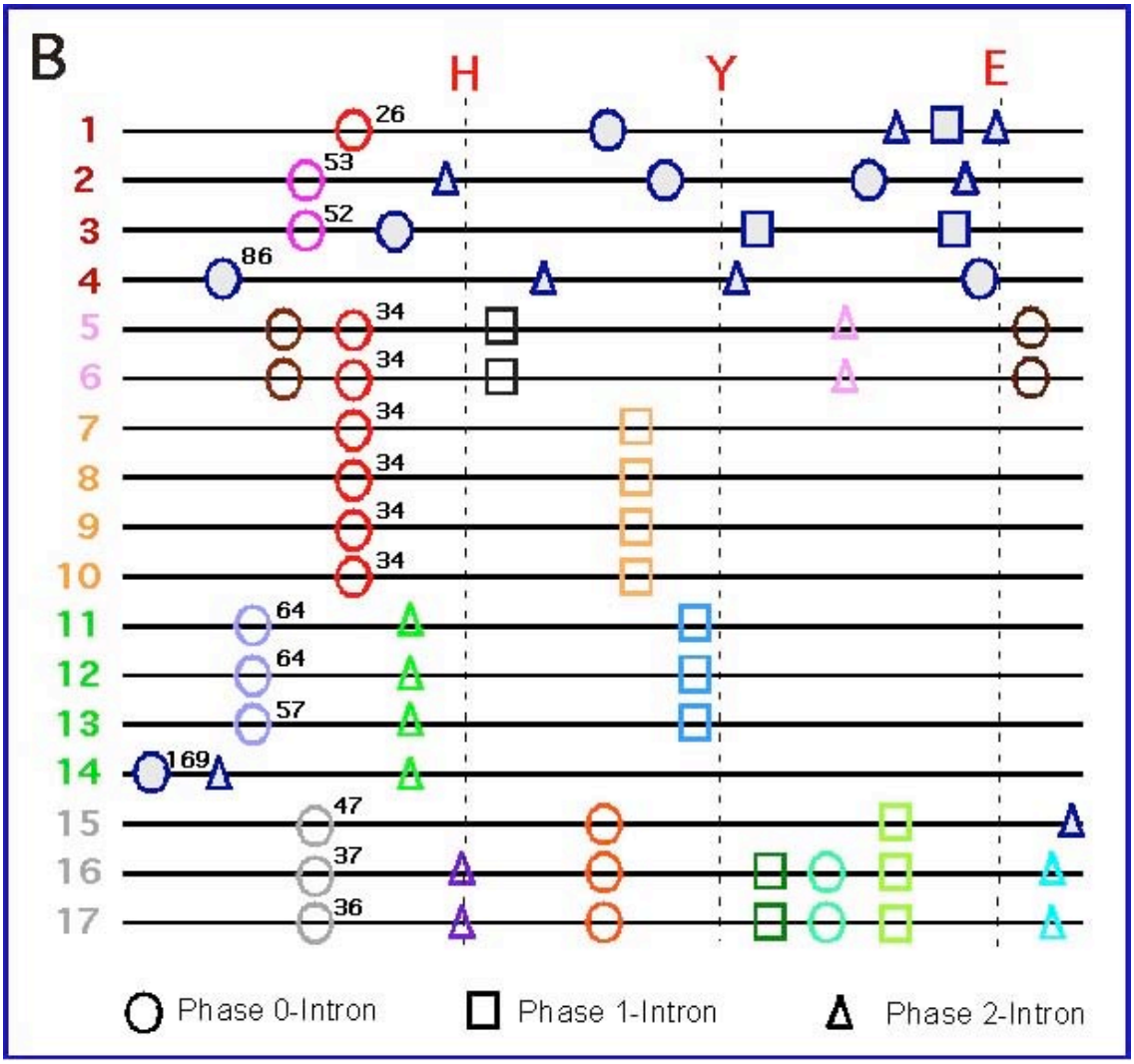


Abbildung 15: Genomische Struktur und konservierte Intronpositionen der humanen *pART*-Paralogen im Bereich der katalytischen Region

Die Exon-/Intron-Struktur konnte für alle humanen *pART*s *in silico* entschlüsselt werden. In dieser Abbildung ist die Exon-/Intron-Struktur der humanen *pART*-Paralogen schematisiert für die Region dargestellt, die die pART-verwandte katalytische Domäne kodiert (**A**). Die Exons sind als Kästen, die Introns als Dreiecke abgebildet. Die Phase des Introneinschubs (0, 1 oder 2) ist in den Dreiecken angegeben. Die Länge der Exons und Introns ist in Basenpaaren (bp) angegeben, die Breite der Kästen spiegelt die Größe der Exons wider. Die kodierenden Regionen der Gene sind rot, die nicht-kodierende 3`-untranslatierte Region ist weiß markiert. Bei den meisten *pART*s endet die kodierende Region im letzten Exon. In diesen Fällen ist nur die Länge der kodierenden Region im letzten Exon (bis zum STOP-Codon), nicht die Gesamtgröße des Exons in Basenpaaren angegeben und die Exons wurden zu den Enden hin durch eine gestrichelte Linie abgeschlossen. Im pART-Alignment (Abbildung 14) wurde der Beginn der katalytischen pART-Domäne durch die Lage des konservierten Phase-0-Introns bei *pART1* (rot) definiert, die entsprechende „katalytische“ Region ist in den Genstrukturen mit einem dunkelblauen Balken gekennzeichnet. Die Positionen der für das "HYE"-Motiv bzw. Varianten dieses Motivs kodierenden Tripletts sind in beiden Abbildungsteilen eingezeichnet. Konservierte Exon-/Intron-Grenzen zwischen den einzelnen humanen *pART*-Genen sind übereinstimmend mit dem Alignment durch farbige Pfeile gekennzeichnet. Die Subgruppen sind durch folgende Farben hervorgehoben: Gruppe 1 rot, 2 rosa, 3 orange, 4 grün und 5 grau. Zur Verdeutlichung wurden die konservierten

Intronpositionen im Abbildungsteil **B** schematisiert dargestellt. Die Intronphasen sind durch unterschiedliche Symbole hervorgehoben (Kreis: Phase-0-Intron, Viereck: Phase-1-Intron, Dreieck: Phase-2-Intron). Die nichtkonservierten Introns sind dunkelblau-grau dargestellt. Die Anzahl der kodierenden Tripletts vom Codon für das Histidin (H) in β1 bis zum nächsten stromaufwärts gelegenen Phase-0-Intron ist angegeben (Ausnahme: *pART2* – Angabe bis zum konservierten Phase-0-Intron).

Analog zu den humanen *pARTs* wurde auch das Maus-Genom mit BLASTn durchsucht und somit die genomische Struktur der Maus-Orthologen ermittelt. Die Intronpositionen im Bereich, der die katalytische Domäne kodiert, sind zwischen den Spezies *Homo sapiens* und *Mus musculus* vollständig konserviert. Die Zugangsnummern der Proteine in der Datenbank, die chromosomale Lokalisation der Gene, die Anzahl der Exons und die Anzahl der Aminosäuren der humanen und murinen pART-Proteine sind in **Abbildung 16** tabellarisch zusammengefasst. Desweiteren sind die zum Teil auf einer kürzlich von Ame *et al.* veröffentlichten Nomenklatur beruhenden, bisherigen Gennamen sowie einige in der Literatur verwendete *alias*-Namen angegeben (Ame *et al.* 2004). Die Unterschiede in der Exonanzahl bei Maus und Mensch sind durch vorhergesagte („Gene“, „Ensemble“), 5`-gelegene , nicht kodierende Exons (z.B. bei *pART11*) oder durch fehlende Sequenzen für den 5`-Bereich bei der Maus bedingt (z.B. *pART4*). Fünf pART-Proteine enthalten bei Maus und Mensch genau die gleiche Anzahl an Aminosäuren (pART 1, 11, 14, 15 und 17). Bei einigen fehlen nur einzelne Aminosäuren (z.B. pART16). Für andere Maus-Proteine sind in der „nr“-Datenbank keine Proteinsequenzen vorhanden, die das mittels EST-Analysen vorausgesagte N-terminale Ende enthalten. Bei diesen Proteinen wurden die Aminosäuresequenzen im Alignment mit den humanen Orthologen durch translatierte EST- und genomische Sequenzen der Maus ergänzt und daraus die Gesamtproteinlänge abgeleitet. Wie bereits erwähnt fehlt in „nr“ für murines pART13 eine Proteinsequenz, die die C-terminale katalytische Domäne umfasst. Dieser Bereich wurde durch unsere, aus überlappenden ESTs translatierte Konsensussequenz ergänzt. Bei einzelnen pARTs wird die genaue Proteinlänge erst in der Zukunft zu ermitteln sein. Für andere pARTs (z.B. pART11, pART16) sind bereits komplette („full length“) cDNAs in der „Mammalian Gene Collection (MGC, http://mgc.nci.nih.gov/)“ vorhanden. Die Zugangsnummern sind ebenfalls in der Abbildung angegeben . MGC-cDNAs, die nach unseren Ergebnissen nicht das vollständige offene Leseraster enthalten, sind in Klammern gesetzt.

pART	Gen	Alias	Zugangsnummer Protein		Chromosom		Exons		Aminosäuren		MGC-cDNA	
	Hs		Hs	Mm	Hs	Mm	Hs	Mm	Hs	Mm	Hs	Mm
1	PARP1	PARP	P09874	NP_031441	1q41-q42	1 H5	23	23	1014	1014	BC037545	BC012041
2	PARP2		CAB41505	NP_033762	14q11.2-q12	14 C1	16	16	583	559	nv	BC062150
3	PARP3		AAM95460	NP_663594	3p21.1-22.2	9 F1	11	11	540	528	(BC014260)	BC014870
4	PARP4	vaultPARP	AAD47250	XP_283217	13q11	14 C1	34	> 28	1724	> 1446	nv	nv
5	TNKS	Tankyrase	AAC79841	AAH57370	8p23.1	8 A4	27	27	1327	1320	nv	BC057370
6	TNKS2	Tankyrase2	NP_079511	XP_129246	10q23.3	19 C2	27	28	1166	1337	nv	nv
7	PARP15		NP_689828	---	3q21.1	---	8	---	444	---	nv	---
8	PARP14		AAN08627	XP_488522	3q21.1	16 B3	12	12	1518	1535	nv	(BC021340)
9	PARP9	BAL	NP_113646	NP_084529	3q13.3-q21	16 B3	11	11	854	830	(BC039580)	BC003281
10	PARP10		BAB55067	AAH24074	8q24.3	15 D3	11	11	1025	960	nv	nv
11	PARP11		AAF91391	NP_852067	12p13.3	6 F3	8	9	331	331	BC017569	BC040269
12	ZC3HDC1		NP_073587	NP_766481	7q34	6 B1	12	12	701	711	BC081541	nv
13	ZC3HAV1	ZAP	NP_064504	BAB32047	7q34	6 B1	13	> 11	902	996	(BC025308)	(BC029090)
14	TIPARP	TiPARP	NP_056323	NP_849223	3q25.31	3 E1	6	6	657	657	BC050350	BC068173
15	PARP16		AAH31074	NP_803411	15q22.2	9 C	6	7	322	322	BC006389	BC055447
16	PARP8		NP_078891	AAH21881	5q11.2	13 D2.3	26	26	854	852	(BC075801)	(BC021315)
17	PARP6		CAB59261	XP_134863	15q22.23	9 C	22	22	630	630	(BC026955)	BC062096

Abbildung 16: Exonanzahl, Proteinlänge, chromosomale Lage und Zugangsnummern der humanen und murinen pARTs

Die pART-Familienmitglieder sind nach den Subgruppen sortiert, die auf der Basis von Aminosäuresequenzidentität, konservierten Intronpositionen und assoziierten Proteindomänen festgelegt wurden. Die Subgruppen sind farblich hervorgehoben: 1 = rot, 2 = rosa, 3 = orange, 4 = grün, 5 = grau. Die offiziellen Gennamen, verwendete *alias*-Namen, die Zugangsnummern der Proteinsequenzen und die chromosomale Lage der pARTs sind aufgeführt. Die Anzahl der Exons und die Länge der offenen Leseraster (in Aminosäuren) der längsten bekannten bzw. vorhergesagten Proteine ist angegeben. Hs = *Homo sapiens*, Mm = *Mus musculus*. Für die MGC-cDNAs sind die zugehörigen Zugangsnummern in der GenBank angegeben. MGC-cDNAs, die nach unseren Ergebnissen nicht das vollständige offene Leseraster enthalten, sind in Klammern gesetzt. nv = nicht vorhanden.

4.1.5 Identifikation von pART-Familienmitgliedern in verschiedenen Spezies

Um die Nutzbarkeit des Programmes PSI-BLAST zur Analyse der strukturellen Verwandtschaft von bekannten ARTs zu erforschen sowie neue Verwandte der humanen pARTs in anderen Spezies aufzuspüren, führten wir PSI-BLAST-Suchen in der "non redundant"-Sektion (nr) der GenBank durch. PSI-BLAST ist ein positionsspezifisches, iteratives Suchprogramm, das durch Generierung einer gewichteten Matrix aus den erhaltenen Ergebnissen die Identifizierung relativ geringer, jedoch biologisch relevanter Sequenzhomologien in entfernt verwandten Proteinen ermöglicht (siehe 1.7.4). Als "Input“ für die PSI-BLAST-Suche wurden die AS-Sequenzen der katalytischen Domäne von

hpART1, hpART9, hpART12, hpART15 verwendet, so dass die zuvor beschriebenen Untergruppen 1, 3, 4 und 5 durch je ein Mitglied vertreten waren. Auf eine Suche mit pART5 oder 6 (Subgruppe 2) wurde verzichtet, da diese beiden pARTs mit >45% noch eine relativ hohe Sequenzidentität zur hpART1 besitzen (vgl. Abb.11).

Abbildung 17 zeigt die Reihenfolge der Detektion der pART-Familienmitglieder in den PSI-BLAST-Analysen, wobei zur besseren Übersichtlichkeit lediglich ein Säuger (*Homo sapiens*) sowie einige ausgewählte Spezies angegeben sind. Aufgrund der Vielzahl der Sequenzen beim Schleimpilz *Dictyostelium discoideum* und der Pflanze *Arabidopsis thaliana* sind diese Organismen in einer eigenen Tabelle aufgeführt (**17B**). Viele pARTs der entfernteren Spezies konnten anhand ihrer Aminosäuresequenz, der Genstruktur (siehe 4.1.6) und der Domänenstruktur (siehe 4.1.7) einer der fünf pART-Subgruppen zugeordnet werden, so dass sie in der Abbildung entsprechend farblich markiert sind. **Abbildung 17C** zeigt eine Auswahl weiterer, mit PSI-BLAST identifizierter pART-Verwandter, die einzelnen pARTs oder den Subgruppen tabellarisch zugeordnet werden konnten. Die Spezies wurden zudem in sechs Gruppen sortiert und das Speziessymbol (z.B. Hs für *Homo sapiens*) wurde entsprechend farblich hervorgehoben: Säugetiere (rot), niedere Wirbeltiere (orange), Insekten (violett), niedere Eukaryonten (blau), Pflanzen (grün), mikrobielle Organismen (braun).

A

Eingabe	Hs.pART1	Hs.pART9	Hs.pART12	Hs.pART15	Ci.pART
Iteration 1 = traditionelle Blast-Programme	Ag.pARTa	Hs.pART8	Hs.pART11	Ag.pARTc	Ag.pARTa
	Dm.pARTa	Hs.pART10	Hs.pART14	Hs.pART17	Hs.pART1
	Hs.pART2	Hs.pART7	Hs.pART13		Hs.pART2
	Ce.pARTa	Hs.pART13	Hs.pART7		Dm.pARTa
	Hs.pART3	Hs.pART14	Hs.pART8		
	Ce.pARTb	Hs.pART12	Hs.pART10		
	Hs.pART4	Hs.pART11	Ag.pARTb		
	Ce.pARTc	Hs.pART6	Hs.pART16		
	Hs.pART5	Hs.pART5	Hs.pART6		
	Ci.pART		Dm.pARTb		
	Ag.pARTb		Hs.pART9		
	Dm.pARTb		Hs.pART5		
Iteration 2	Hs.pART6	Ag.pARTb	Hs.pART3	Hs.pART16	Hs.pART3
	Hs.pART14	Dm.pARTb	Dm.pARTa	Hs.pART2	Ce.pARTa
	Hs.pART13	Hs.pART3	Ce.pARTc	Ag.pARTa	Hs.pART4
	Ba.pART	Hs.pART1	Hs.pART1	Ce.pARTa	Ce.pARTb
	Hs.pART16	Dm.pARTa	Hs.pART4		Ce.pARTc
	Hs.pART7	Ce.pARTb	Ce.pARTb		Dm.pARTb
	Hs.pART11	Hs.pART4	Ag.pARTa		Hs.pART5
	Ag.pARTc	Ce.pARTc	Hs.pART2		Ag.pARTb
	Hs.pART17	Ag.pARTa	Ce.pARTa		Hs.pART6
	Hs.pART12		Ci.pART		Hs.pART14
					Hs.pART16
					Hs.pART7
					Ba.pART
Iteration 3	Hs.pART8	Hs.pART2	Hs.pART17	Dm.pARTa	Hs.pART8
	Hs.pART10	Ce.pARTa	Hs.pART15	Hs.pART1	Hs.pART10
	Hs.pART15	Hs.pART15	Ag.pARTc	Hs.pART3	Hs.pART12
	Hs.pART9	Ci.pART		Ce.pARTc	Hs.pART11
		Ag.pARTc		Ce.pARTb	Hs.pART13
		Hs.pART16		Hs.pART4	Ag.pARTc
		Hs.pART17		Hs.pART5	Hs.pART17
		Ba.pART		Ag.pARTb	Hs.pART15
				Dm.pARTb	
Iteration 4				Hs.pART6	Hs.pART9
				Hs.pART14	
				Hs.pART7	
				Hs.pART13	
				Hs.pART11	
				Ci.pART	
				Hs.pART12	
				Ba.pART	
Iteration 5				Hs.pART8	
				Hs.pART10	
				Hs.pART9	
Iteration 6	konvergent		Ba.pART		konvergent
Iteration 7		konvergent	konvergent		
Iteration 8				konvergent	

B

Eingabe	Hs.pART1	Hs.pART9	Hs.pART12	Hs.pART15	Ci.pART
Iteration 1 = traditionelle Blast-Programme	At.pARTa		Dd.pARTg		At.pARTb
	At.pARTb		Dd.pARTf		Dd.pARTb
	Dd.pARTa				Dd.pARTa
	Dd.pARTb				
	Dd.pARTc				
	Dd.pARTd				
	At.pARTc				
	Dd.pARTe				
Iteration 2	Dd.pARTf	Dd.pARTg	Dd.pARTc	Dd.pARTd	At.pARTa
	Dd.pARTg	Dd.pARTf	At.pARTb		Dd.pARTc
	At.pARTd	Dd.pARTc	At.pARTa		Dd.pARTd
		At.pARTh	Dd.pARTh		At.pARTc
		At.pARTa	Dd.pARTa		Dd.pARTe
		Dd.pARTh	At.pARTg		Dd.pARTf
		At.pARTb	Dd.pARTb		
		Dd.pARTb	At.pARTd		
		At.pARTg	At.pARTf		
		At.pARTf	At.pARTh		
			At.pARTe		
Iteration 3	At.pARTe	Dd.pARTa	Dd.pARTd	At.pARTb	Dd.pARTg
	At.pARTf	At.pARTd	Dd.pARTe	At.pARTa	At.pARTd
		At.pARTe	At.pARTi	Dd.pARTa	At.pARTe
		Dd.pARTd	At.pARTc	Dd.pARTb	
		Dd.pARTe	Dd.pARTi	Dd.pARTc	
		At.pARTi		Dd.pARTe	
				Dd.pARTf	
Iteration 4	At.pARTg	At.pARTc	At.pARTj	At.pARTc	At.pARTf
	At.pARTh	At.pARTj		Dd.pARTg	At.pARTg
				At.pARTd	At.pARTh
				At.pARTe	
Iteration 5	At.pARTi			At.pARTf	At.pARTi
				At.pARTh	
Iteration 6	konvergent			At.pARTg	konvergent
Iteration 7		konvergent	konvergent	At.pARTi	
Iteration 8				konvergent	

C

pART	Gen	Mensch *Homo sapiens*	Maus *Mus musculus*	Huhn *Gallus gallus*	Fisch *Tetraodon nigroviridis*	Fliege *Drosophila melanogaster*	Moskito *Anopheles gambiae*
1	PARP1	P09874	NP_031441	NP_990594	CAG09179	P35875_a	XP_312938_a
2	PARP2	Q9UGN5	NP_033762		CAF92030		
3	PARP3	AAM95460	NP_663594		CAG06805		
4	PARP4	AAD47250	XP_283217	XP_417150	CAG08214		
5	TNKS	AAC79841	AAH57370	NP_989671	CAG12585	AAF56487_b	XP_321116_b
6	TNKS2	NP_079511	XP_129246	NP_989672	CAG04910		
7	PARP15	NP_689828	—				
8	PARP14	AAN08627	XP_488522	XP_422113	CAF98988		
9	PARP9	NP_113646	NP_084529	XP_422116	CAG12587		
10	PARP10	BAB55067	AAH24074		CAG05989		
11	PARP11	AAF91391	NP_852067	XP_416489	CAG01913		
12	ZC3HDC1	NP_073587	NP_766481	XP_416333	CAF98285		
13	ZC3HAV1	NP_064504	BAB32047	XP_423977	CAF96305		
14	TIPARP	NP_056323	NP_849223	XP_422828	CAF96664		
15	PARP16	AAH31074	NP_803411	XP_413903	CAG05566		XP_308419_c
16	PARP8	NP_078891	AAH21881	XP_424786	CAG05573		
17	PARP6	CAB59261	XP_134863		CAF95416		

pART	Wurm *Caenorhabditis elegans*	Schleimpilz *Dictyostelium discoideum*	Pilz *Gibberella zeae*	Amoebe *Entamoeba histolytica*	Pflanze *Arabidopsis thaliana*	Virus	Bakterium
1	AAM27195_a	CAD59237_a	EAA75569_a	EAL47198_a	NP_850165_a	AAB94432	1AERA 760286A
2		CAD58666_c		EAL50270_b	CAA88288_b	AAQ17796	AAW80252
3	Q09525_b	CAD59238_d			BAB09119_c		
4		CAD59240_e					
5	AAC04454_c	CAD59239_b					
6		AA051129_f					
7							
8		AAS38928_g					
9							
10							
11				EAL43406_c			
12				EAL50270_d	NP_849739_d		
13				EAL49071_e	AAC36170_e		
14				EAL45174_f			
15							
16			EAA73885_c				
17							

Abbildung 17: PSI-BLAST-Datenbankrecherchen zur Identifizierung neuer pART-Familienmitglieder in verschiedenen Spezies

In den Tabellen ist eine Auswahl der Ergebnisse der PSI-BLAST Datenbankanalysen zur Entdeckung neuer Mitglieder der poly-ART-Familie zusammengefasst. Die PSI-BLAST-Suchen wurden mit den AS-Sequenzen der katalytischen Domäne von hpART1, hpART9, hpART12, hpART15 sowie der Chilo iridescent Virus pART als "Input" in der "non redundant"-Sektion (nr) der GenBank durchgeführt (zuletzt aktualisiert im Februar 2005). In Tabelle **A** und **B** sind die pARTs in der Reihenfolge ihrer Detektion aufgeführt. Die initialen Verknüpfungen entsprechen den Proteinen, die eine signifikante Sequenzidentität zum jeweils eingegebenen Protein aufweisen und die auch mit traditionellen BLAST-Programmen ermittelt werden können. Der "threshold" zur Abgrenzung von signifikanten Ergebnissen lag bei 0.005. Die PSI-BLAST-Suche wurde so lange wiederholt (Iterationen), bis sie keine neuen Ergebnisse erbrachte und somit konvergent war. Die Mitglieder der verschiedenen Spezies wurden sechs Speziesgruppen zugeordnet, die Speziessymbole wurden entsprechend farblich markiert: Säugetiere (rot), niedere Wirbeltiere (orange), Insekten (violett), niedere Eukaryonten (blau), Pflanzen (grün), mikrobielle Organismen (braun). Zur besseren Übersichtlichkeit sind in den Tabellen **A** und **B** für jede Gruppe nur ausgewählte Spezies dargestellt, z.B. sind die Orthologen von Maus, Ratte und Huhn nicht aufgeführt. Aufgrund der diversen Verwandten bei *D.discoideum* und *A.thaliana* wurde für diese Spezies eine eigene Tabelle erstellt (**B**). In Tabelle **C** sind die Familienmitglieder aus den

verwandten Spezies aufgeführt, die einer bestimmten pART oder zumindest einer pART-Subgruppe zugeordnet werden konnten. Die pART-Subgruppen sind farblich hervorgehoben: 1 rot, 2 rosa, 3 orange, 4 grün, 5 grau. Da die pART-Verwandten in den entfernteren Spezies nur schwierig einer bestimmten humanen pART zuzuordnen waren, wurden sie nach der Reihenfolge ihres Auffindens mit der pART1 als Matrize als pARTa, pARTb, etc. bezeichnet. Die meisten Zugangsnummern für die Proteinsequenzen sind in Tabelle **C** angegeben. Weitere Zugangsnummern sind: *Dictyostelium discoideum* (Dd, Schleimpilz) Dd.pARTh AAL92335, Dd.pARTi AAL92334, *Arabidopsis thaliana* (At) At.pARTf AAF79590, At.pARTg AAC18815, At.pARTh NP_201058, At.pARTi T07711, At.pARTj AAF16645, Chilo iridescent (Ci, Insektenvirus) Ci.pART AAB94432, Bacteriophage Aeh1 (Ba) Ba.pART AAQ17796.

Bei Verwendung der Aminosäuresequenz der katalytischen Domäne von hpART1, hpART9, hpART12 und hpART15 wurden alle humanen pARTs gefunden (**A**). Es wird deutlich, dass in Abhängigkeit von der pART-Matrize zunächst die Mitglieder der jeweiligen Untergruppen detektiert wurden. Entferntere humane Verwandte ließen sich erst in den späteren Iterationen ermitteln. Während PSI-BLAST z.B. hpART16 und hpART17 mit der Matrize hpART15 bereits initial bzw. in der zweiten Iteration aufspürt, wurde die signifikante Homologie zu den humanen pARTs 8, 9 und 10 erst im fünften Durchgang erkannt. Mit den Vertretern hpART9 und hpART12 als Eingabe ermittelte PSI-BLAST alle Paralogen der Subgruppen 3 und 4 bereits in der ersten Iteration, wobei mit hpART9 zunächst pART 7, 8 und 10 (Subgruppe 3) und mit hpART12 die pARTs 11, 13 und 14 (Subgruppe 4) gefunden wurden. Dies bestätigt erneut unsere Subgruppeneinteilung.

Weiterhin wurden Homologe in vielen verschiedenen Spezies wie Maus, Ratte, Huhn, Insekten, Würmern und Pflanzen identifiziert. In vielen Fällen wurden Verwandte aus entfernten Spezies in früheren Durchläufen gefunden als humane Paraloge aus anderen Subgruppen. Wie bereits erwähnt konnten bei der Maus (*mus musculus*), mit Ausnahme von pART7 und pART13 (Stand 04/2005), für alle humanen pARTs hochkonservierte Orthologe (Sequenzidentität meist zwischen 80 und 90%) in der „nr“-Sektion der GenBank nachgewiesen werden. Bei der Ratte (*rattus norvegicus*) fehlte neben diesen beiden pARTs zusätzlich das Orthologe von pART11 (Stand 04/2005). Unsere Genomanalysen (siehe 4.1.3) zeigten, daß das pART7-Orthologe in den beiden Nagergenomen tatsächlich fehlt, während die anderen *pArt*-Gene in den Genomen enthalten sind und bisher lediglich Proteinsequenzen in der „nr-„Datenbank fehlen, die die katalytische Domäne umfassen.

Auch bei *Gallus gallus* und dem Fisch *Tetraodon nigroviridis* gelang es, den meisten humanen pARTs ein Orthologes zuzuordnen (**17C**). Schwierigkeiten ergaben sich bei den Paralogen mit hoher Sequenzidentität (z.B. pART5 und 6 bzw. pART16 und 17), die beim Fisch nur noch der Subgruppe, nicht aber dem einzelnen Gen zugeteilt werden konnten. Interessanterweise wurde auch bei Huhn und Fisch kein pART7-Orthologes nachgewiesen. Möglicherweise kommt diese pART nur bei Primaten vor. Niedere Eukaryonten besitzen meist eine geringere Anzahl an pARTs. Im komplettierten Genom des Nematoden *Caenorhabditis elegans* konnten drei, bei der Fliege *Drosophila melanogaster* zwei, bei der Mücke *Anopheles gambiae* drei und im Genom der Pflanze *Arabidopsis thaliana* zehn potentielle Familienmitglieder nachgewiesen werden. Beim Schleimpilz *Dictyostelium discoideum* wurden bisher neun, bei dem Ascomyzeten *Gibberella zeae* drei und bei dem Einzeller *Entamoeba histolyticum* sechs pARTs ermittelt. Zudem wurden zwei virale Homologe, die pARTs des Chilo iridescent Virus und des Bakteriophagen Aeh1, entdeckt. Bei beiden handelt es sich um Doppelstrang-DNA-Viren mit relativ großen Genomen. Der Chilo iridescent Virus, dessen komplettes Genom (Gesamtgröße ca. 212kb) seit 2001 bekannt ist, ist ein großer, ikosaedrischer Virus, der verschiedene Insektenarten wie z.B. Grillen und Heuschrecken befällt und aufgrund seiner hohen Pathogenität als mögliches biologisches Insektizid erforscht wird (Jakob and Darai 2002). Der Bakteriophage Aeh1, der aus dem Bakterium *Aeromonas hydrophila* isoliert wurde, gehört zu den T4-verwandten, Bakterien-infizierenden Viren und besitzt ein ca. 230kb umfassendes Genom.

Als weiteres PSI-BLAST-Experiment führten wir den Versuch einer „gegensinnigen“ Verknüpfung, also von den niederen Spezies zu den menschlichen Sequenzen, durch. Mit der Iridovirus pART als Eingabe konnten alle humanen pARTs aufgespürt werden (**Abb.17 A**, rechte Spalte). Dies unterstreicht die Signifikanz der Verknüpfungen. Trotz der bestehenden Ähnlichkeit gelang es bei den PSI-BLAST-Analysen mit pARTs als „input“ nicht, die pART-Familie mit Diphterietoxin oder *Pseudomonas aeruginosa* Exotoxin A zu verknüpfen. Mit einer chimären Aminosäuresequenz, in der die Sequenzen des Diphterietoxins bei drei der konservierten Sekundärstrukturmotive (β1, β5 und β6) durch die entsprechende Sequenz der Ci.pART ersetzt wurden, konnten sowohl alle humanen pARTs als auch die beiden bakteriellen Toxine gefunden werden (nicht gezeigt). Durch die Suche mit der Chimäre wurde

zudem eine DT/ETA-ähnliche ADP-Ribosyltransferase in *V.cholerae* entdeckt (Zugangsnummer: AAW80252). Weitere neue mikrobielle Proteine wurden nicht gefunden.

Interessanterweise wurden keine Verwandten in den vollständig entschlüsselten Genomen der Hefe (*Saccharomyces cerevisiae*) oder der Archae ermittelt. Im auffälligen Gegensatz dazu wurden mit der humanen ART1 als "Input" in vielen Mikroben Verwandte der RT6-mono-ARTasen aufgespürt, während in keinem der komplettierten Genome niederer Eukaryonten ein Familienmitglied zu lokalisieren war (siehe Abb.33). Die DT/ETA/pARTs-Gruppe konnte durch PSI-BLAST nie mit einem bekannten Mitglied aus der RSE-Motiv-Familie der bakteriellen Toxine und Vertebraten-mono-ARTs verknüpft werden.

Der Grad der Sequenzidentität mehrerer Verwandter zur hpART1 beträgt deutlich weniger als 20%, so dass sie folglich mit traditionellen Datenbank-Suchprogrammen nicht zu detektieren gewesen wären. Da insbesondere die pART-Verwandten in den entfernteren Spezies, wie z.B. *A.thaliana,* alleine von der Aminosäuresequenz nur schwierig einer bestimmten humanen pART zuzuordnen waren, wurden sie nach der Reihenfolge ihres Auffindens mit der pART1 als Matrize als pARTa, pARTb, etc. bezeichnet. Aus den PSI-BLAST Analysen wird deutlich, dass mit der pART1 als Eingabesequenz zunächst die pART1-Orthologen in verschiedenen Spezies (z.B. Hühner-pART1, *Drosophila* pARTa, *Anopheles* pARTa, *C.elegans* pARTa) gefunden werden. Somit ist anzunehmen, dass die erstgefundenen Homologen bei *A.thaliana* und *D.discoideum* ebenfalls pART1-Orthologen entsprechen. Erwähnenswert ist zudem, dass mit pART15 als Eingabe als erster Treffer die *Anopheles* pARTc aufgespürt wurde. Aufgrund der hohen Sequenzidentität im Bereich der katalytischen Domäne ist davon auszugehen, dass es sich um das Orthologe der humanen pART15 handelt. Da die beiden anderen *Anopheles*- und *Drosophila*-Orthologen jeweils eine hohe Ähnlichkeit zueinander aufweisen, wurde mit der Aminosäuresequenz der *Anopheles* pARTc und tBLASTn das Fliegen-Genom durchsucht. Es konnte aber keine homologe Sequenz nachgewiesen werden, so dass der Fliege diese dritte pART möglicherweise fehlt oder das Gen in einer bisher nicht sequenzierten Lücke des Fliegengenoms liegt. Einige der Familienmitglieder aus den verschiedenen Spezies wurden zur weiteren Charakterisierung ausgewählt (siehe 4.1.6 und 4.1.7). Die Erkenntnisse aus den Gen- und Domänenstrukturen wurden bereits für die Zuordnung in der Abbildung 17 berücksichtigt.

4.1.6 *In silico* Analyse und Vergleich der Genstrukturen von *pART*-Verwandten anderer Spezies

Einige der mit PSI-BLAST neu identifizierten Familienmitglieder aus den Spezies *D.melanogaster, C.elegans* und *A.thaliana* sowie die mikrobielle, nur aus der katalytischen Domäne bestehende Chilo iridescent pART wurden zur weiteren Analyse ausgewählt. Die Aminosäuresequenzen der katalytischen Domänen wurden mit Hilfe von Sequenz-Alignment-Programmen mit ausgewählten humanen pARTs verglichen, wobei das Alignment die Ergebnisse aus den PSI-BLAST-Suchen sowie die PSIPRED- und GeneTHREADER-Analysen berücksichtigend manuell nachgebessert wurde (**Abbildung 18**). Übereinstimmend mit dem Vergleich der verschiedenen humanen Paralogen zeigt sich auch bei den Verwandten der anderen Spezies die ausgeprägteste Sequenzidentität im Bereich der Sekundärstrukturmotive, die in der Kristallstruktur der Hühner-pART1 das katalytische Zentrum bilden. Dies unterstreicht die Signifikanz der gefundenen Verknüpfungen. In den „loops“ zwischen den Motiven ist die Anzahl der konservierten Aminosäuren deutlich geringer. Alle für die weitere Analyse ausgewählten Verwandten weisen das charakteristische HYE-Motiv auf. Dies sollte aber nicht darüber hinwegtäuschen, dass viele der mit PSI-BLAST gefundenen Verwandten nicht das klassische Motiv besitzen: Während bei fast allen Verwandten das katalytische Histidin (H) in β1 und das Tyrosin (Y) in β2 konserviert ist, kommen anstelle der katalytischen Glutaminsäure (E) häufiger andere Aminosäuren wie Isoleucin (I), Valin (V), Leucin (L) oder Threonin (T) vor. Die PSI-BLAST-Analyse ließ wegen der hohen Sequenzidentität vermuten, dass es sich bei der pARTc von *Anopheles gambiae* um das Orthologe des humanen pART15 handelt. Interessanterweise besitzen beide ein HYY-Motiv.

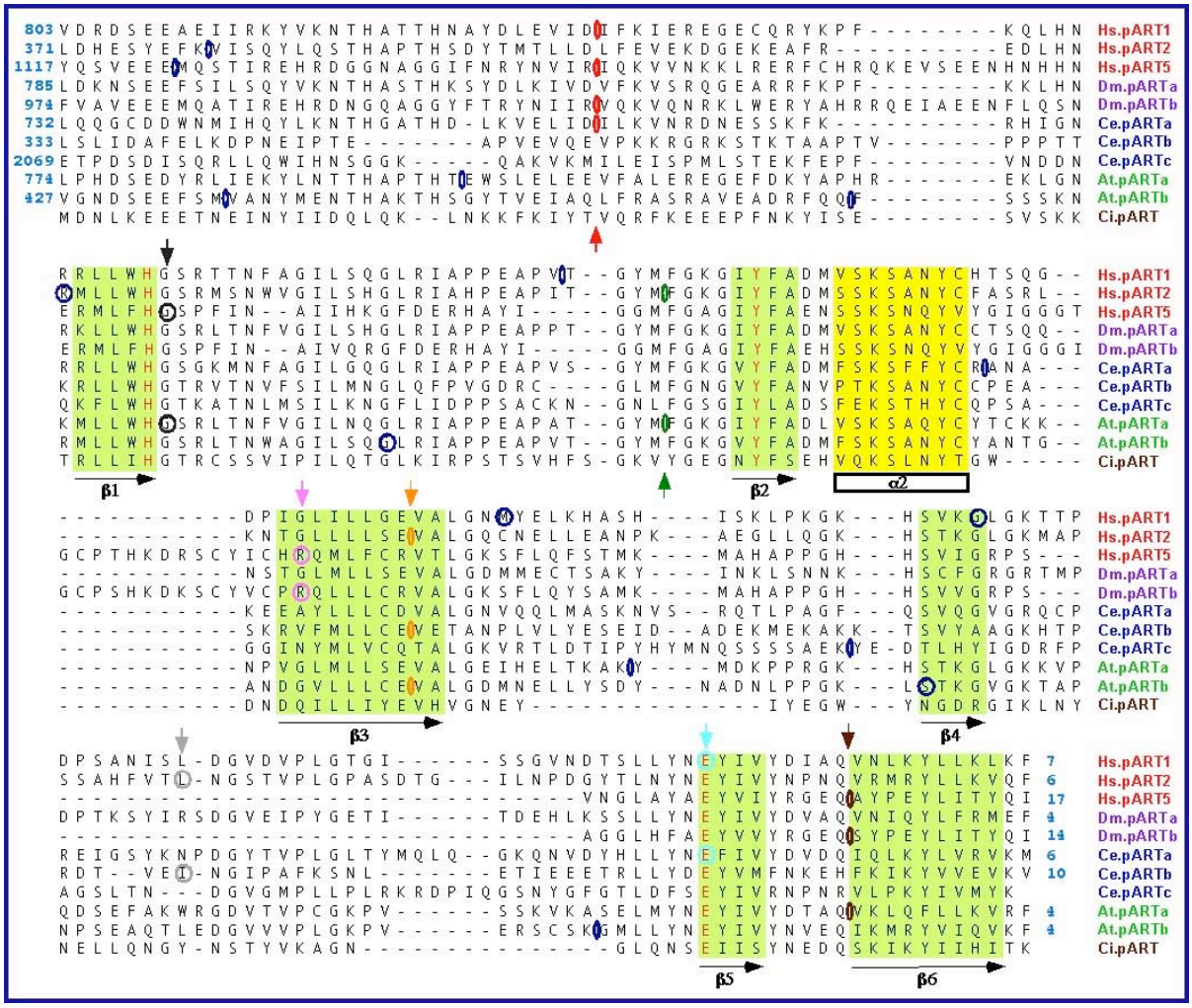


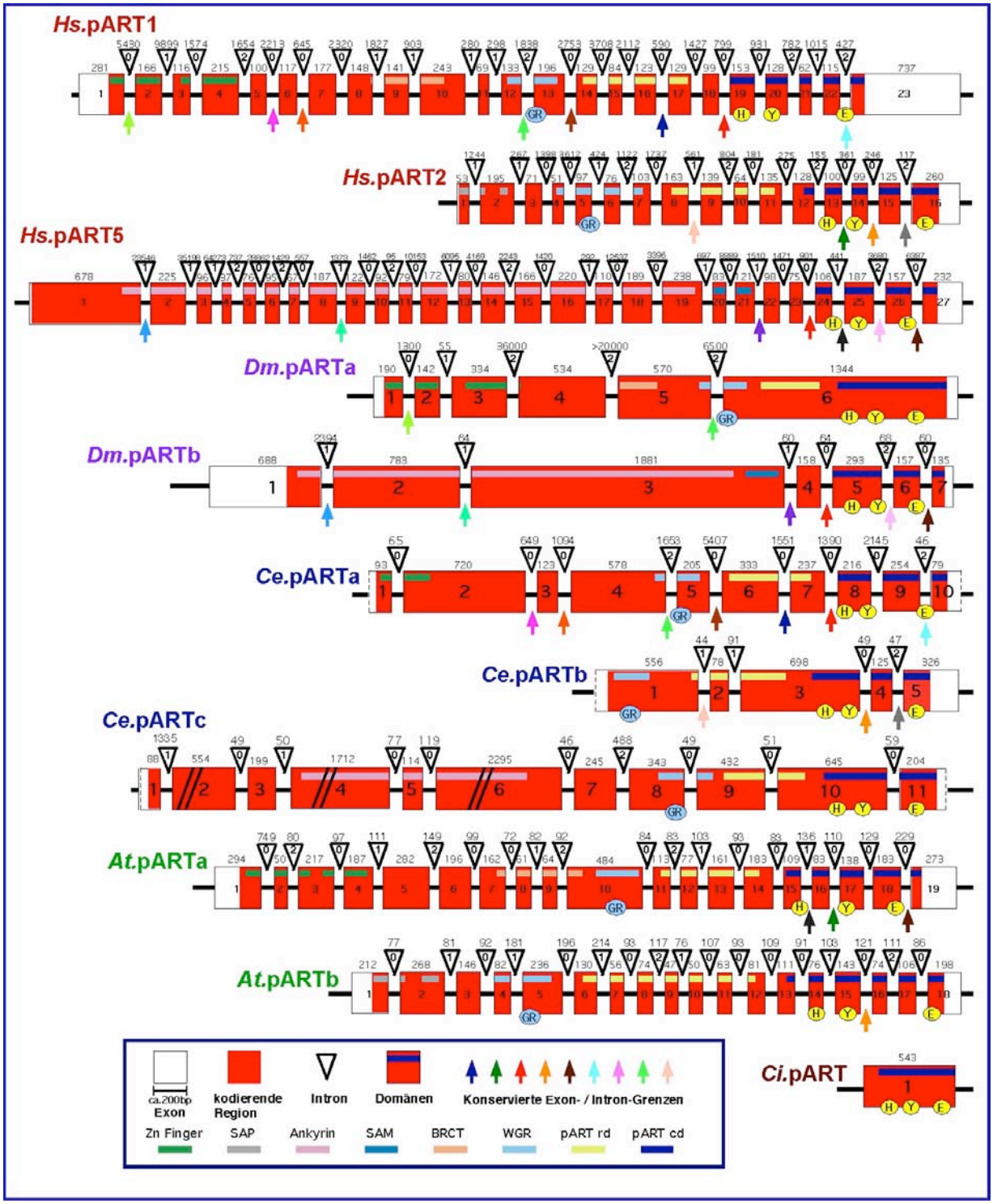
Abbildung 18: Konservierte Aminosäuren, Sekundärstrukturmotive und Lage der Introns in der potentiellen katalytischen Domäne von pART-Verwandten verschiedener Spezies

Die Abbildung zeigt ein mit Hilfe von Sequenz-Alignment-Programmen (T-Coffee, Megalign) erstelltes sowie nach PSI-BLAST-Ergebnissen manuell nachgebessertes Alignment der katalytischen Domäne von ausgewählten pART-Familienmitgliedern verschiedener Spezies. Die Speziesgruppen sind übereinstimmend mit Abbildung 17 farbig markiert: Säugetiere rot, Insekten violett, niedere Eukaryonten blau, Pflanzen grün, mikrobielle Organismen braun. Die Lage der Introns ist in die Abbildung eingezeichnet, wobei Phase-0-Introns als Ovale zwischen den Aminosäuren und Phase-1- bzw. Phase-2-Introns als Kreise um die jeweilige Aminosäure dargestellt sind. Konservierte Intronpositionen zwischen verschiedenen Familienmitgliedern wurden farblich hervorgehoben und mit Pfeilen gekennzeichnet. Die nichtkonservierten Introns sind dunkelblau dargestellt. Die Anzahl der Aminosäuren der einzelnen Proteine N-terminal und C-terminal des Alignments ist angegeben. Für die pART des Virus Chilo iredescent (Ci) ist die gesamte Proteinsequenz angegeben. Die aus der Kristallstruktur der Hühner-pART1 bekannten Sekundärstrukturmotive sind, analog zur Abbildung 10, eingefärbt und unterhalb des Alignments markiert (β-Stränge grün/Pfeil, α-Helix gelb/Kasten). Das HYE-Motiv

ist rot markiert. Die Sequenzen sind in der Datenbank unter den in Abbildung 17C angegebenen Zugangsnummern deponiert.

Nach den aufschlussreichen Ergebnissen für die humanen pARTs stellte sich die Frage, ob auch die Familienmitglieder anderer Spezies konservierte Introns aufweisen, aus denen sich Hinweise auf die nächsten humanen Verwandten und gemeinsame Vorläufergene ergeben. Zu den bei der PSI-BLAST-Suche erhaltenen Proteinsequenzen wurden deshalb die entsprechenden Nukleotid-Sequenzen ermittelt. Mit den Nukleotid-Sequenzen als Matrize führten wir eine BLASTn-Suche nach bereits in der Datenbank deponierten genomischen Sequenzen dieser Spezies durch. Unter Nutzung der zunehmenden Ressourcen durch die fortschreitenden Genomprojekte konnten die zugehörigen genomischen Sequenzen in der Datenbank ermittelt werden. Somit war es möglich, die Exon-Intron-Struktur im kodierenden Bereich aufzuklären und dadurch Aussagen über die Exongrößen und die Konservierung einzelner Intronpositionen zu erhalten. Die Genstrukturen sind in **Abbildung 19A** schematisch dargestellt. Die konservierten Intronpositionen sind durch farbige Pfeile markiert und zusätzlich in dem Alignment von Abbildung 18 eingezeichnet. Zur Verdeutlichung sind die konservierten Intronpositionen in **Abbildung 19B** (analog zur Abb.15B) schematisiert dargestellt. Zudem sind in der Abbildung die Regionen gekennzeichnet, die konservierte Proteindomänen kodieren, die einzelnen Domänen werden unter 4.1.7 genauer erläutert.

Die pART-Gene unterscheiden sich stark in der Anzahl der Exons. Vergleichbar mit vielen anderen Genen enthalten die pART-Gene der „niederen“ Spezies zumeist weniger und kürzere Introneinschübe als ihre humanen Verwandten. So wird die pART1 beim Menschen durch 23 Exons, hpART2 durch 16 Exons und hpART5 durch 27 Exons kodiert, während für die *Drosophila pARTa* lediglich 6 Exons ermittelt wurden. Das für die Iridovirus pART kodierende Gen ist kontinuierlich. Die meisten Exons der humanen *pART*-Gene sind 100 – 300 Basenpaare (bp) groß und damit deutlich kleiner als einige Exons der Homologen in den anderen Spezies (z.B. bei *C.elegans* bis zu 3000bp). Bei den gezeigten Verwandten wird die katalytische Domäne mit dem charakteristischen HYE-“Fingerabdruck“ neben dem Iridovirus lediglich bei der *Drosophila pARTa* durch ein einziges Exon zusammenhängend kodiert. Im Gegensatz dazu liegen die Tripletts für diese drei Aminosäuren bei der *hpART1* und weiteren humanen *pART*s (siehe Abbildung 15) in drei unterschiedlichen Exons. Dies ist ein auffälliger Unterschied zu den humanen RT6-verwandten mono-ARTs, bei denen die gesamte katalytische Domäne von einem großen Exon kodiert wird.


Hs.pART1
Hs.pART2
Hs.pART5
Dm.pARTa
Dm.pARTb
Ce.pARTa
Ce.pARTb
Ce.pARTc
At.pARTa
At.pARTb
Ci.pART
ca.200bp Exon
kodierende Region
Intron
Domänen
Konservierte Exon- / Intron-Grenzen
Zn Finger
SAP
Ankyrin
SAM
BRCT
WGR
pART rd
pART cd

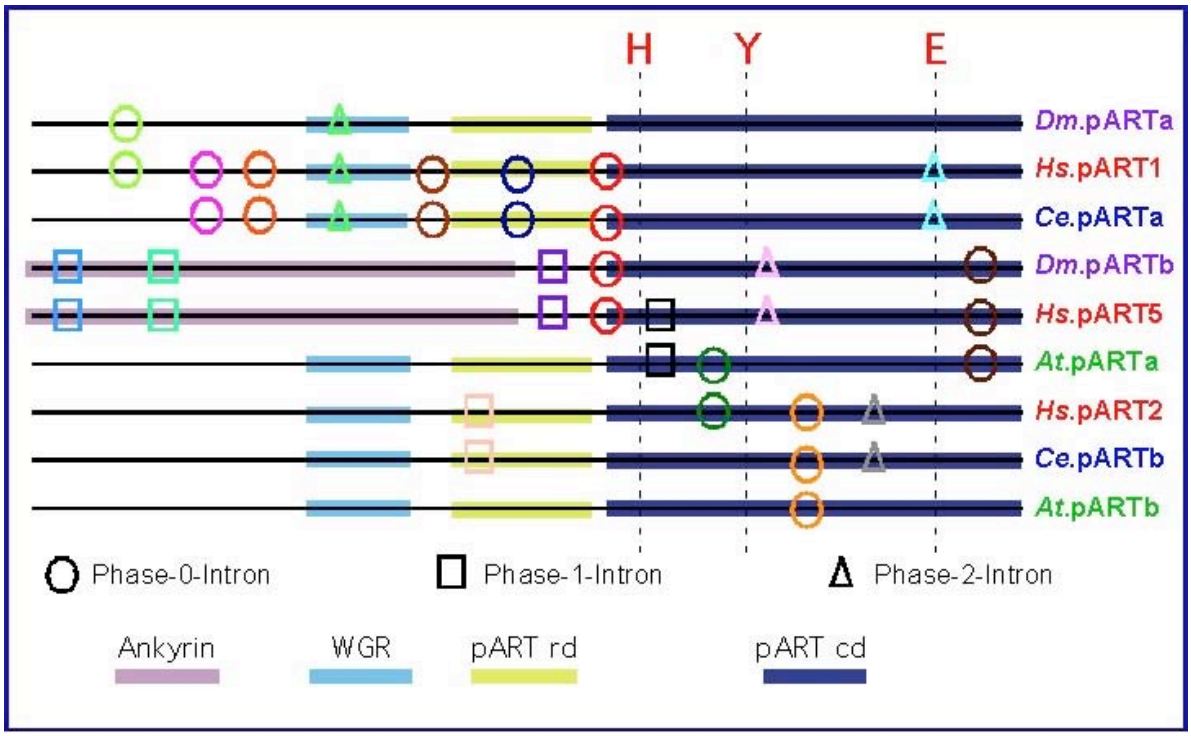


Abbildung 19: Schematische Darstellung der Genstrukturen ausgewählter pART-Familienmitglieder aus verschiedenen Spezies (A) sowie Darstellung der konservierten Intronpositionen (B)

Für ausgewählte pARTs aus verschiedenen Spezies wurde die Exon-/Intron-Struktur des kodierenden Bereiches *in silico* aufgeklärt. In der schematischen Darstellung sind die Exons als Kästen und die Introns als Dreiecke abgebildet. Die Intronphase (0,1 oder 2) ist in den Dreiecken angegeben. Die Länge der Exons und Introns ist in Basenpaaren (bp) angegeben. Die kodierenden Regionen der Gene sind rot markiert. Wenn die Größe der Exons am 5´- bzw. 3‘-Ende nicht zu ermitteln war, wurden die Exons zu den Enden hin durch eine gestrichelte Linie abgeschlossen und lediglich die Länge der kodierenden Region in Basenpaaren angegeben. Die Bereiche, die die katalytische Domäne kodieren, sind mit einem blauen Balken gekennzeichnet. Regionen, die wichtige assoziierte Domänen kodieren (siehe 4.1.7), sind ebenfalls durch farbliche Balken hervorgehoben. Die Lage der für das "HYE"-Motiv bzw. für das „GR“-Motiv der WGR-Domäne kodierenden Tripletts ist eingezeichnet. Konservierte Exon-/Intron-Grenzen zwischen den einzelnen Familienmitgliedern sind übereinstimmend mit dem Alignment durch farbige Pfeile markiert. Zusätzlich wurden auch die konservierten Intronlagen 5`-wärts der katalytischen Region ermittelt. Die großen Exons bei der *C.elegans pARTc* wurden nicht im Maßstab wiedergegeben und deshalb mit schwarzen Schrägbalken kenntlich gemacht. Die Speziesgruppen sind durch folgende Farben markiert: Säugetiere rot, Insekten violett, niedere Eukaryonten blau, Pflanzen grün, mikrobielle Organismen braun. Zur Verdeutlichung wurden die konservierten Intronlagen im Abbildungsteil **B** schematisiert dargestellt und die Familienmitglieder nach den konservierten Intronlagen sortiert. Die Intronphasen sind durch unterschiedliche Symbole hervorgehoben. (Kreis: Phase-0-Intron, Viereck: Phase-1-Intron, Dreieck: Phase-2-Intron). Die Lage des „HYE“-Motivs ist durch senkrechte Linien gekennzeichnet.

Über Speziesgrenzen hinweg bestehen sowohl im Bereich, der die katalytische Domäne kodiert, als auch in der 5‘-Region konservierte Exon-Intron-Grenzen, die durch farbige Pfeile gekennzeichnet sind. Eine deutliche Variation zeigt sich aber bezüglich der Größe der Introneinschübe an den konservierten Intronpositionen. Der bei sieben humanen *pART*s konservierte Phase-0-Introneinschub vor dem β1-kodierenden Bereich ist auch bei der *Drosophila pARTb* und der *C.elegans pARTa* konserviert (roter Pfeil). Die *C.elegans pARTa* besitzt sechs weitere konservierte Intronpositionen mit der h*pART1*, wobei das 3´-gelegene Phase-2-Intron direkt im Triplett für die katalytische Glutaminsäure liegt (hellblau). Bei der *Drosophila pARTa* sind zwei ihrer fünf Introns zur *hpART1* konserviert. Dies ist neben den Sequenzidentitäten ein weiterer Hinweis, daß es sich bei Ce.pARTa und Dm.pARTa um Orthologe der hpART1 handelt. *C.elegans pARTb*, *A.thaliana pARTb* und die humane *pART2* zeichnen sich durch ein, in allen drei Spezies konserviertes Intron aus (orange), so dass diese *pART*s, die bei der PSI-BLAST-Suche mit der humanen *pART1* jeweils als zweites ihrer Spezies gefunden wurden, wahrscheinlich die Orthologen der *pART2* darstellen. Dies wird im Falle der *ce.pARTb* dadurch untermauert, dass zwei weitere konservierte Intronpositionen zur *hpART2* bestehen. Die *Drosophila pARTb* ist durch sechs konservierte Introns zur Subgruppe der humanen *pART5* und *pART6* gekennzeichnet, von denen mehrere auch in der vorderen Region liegen. Dies ist ein weiterer Hinweis darauf, dass es sich bei der *Drosophila pARTb* um ein Orthologes der Subgruppe 2 handelt. Obwohl ce.pARTc nach ihrer Domänenenstruktur (siehe Abb.20) zu der Subgruppe 2 gehört, sind keine Intronpositionen zur humanen pART5 konserviert. Die pARTs der Modellpflanze *Arabidopsis thaliana* enthalten eine relativ große Zahl an Introns, von denen aber nur wenige zu humanen pARTs konserviert sind. Bei der *A.thaliana pARTa,* die von der Sequenz am ehesten der pART1 zuzuordnen ist, ist kein einziges der 18 Introns zur humanen pART1 konserviert. Sie besitzt jedoch konservierte Intronlagen zu zwei verschiedenen humanen Subgruppen: Einen gemeinsamen Introneinschub mit der humanen *pART2* (dunkelgrün) und zwei weitere mit der Subgruppe 2 (braun und schwarz).

Bereits in frühen Untersuchungen war uns aufgefallen, dass einige pARTs in N-terminaler Richtung über die katalytische Domäne hinaus bis zu einem Glycin-Arginin-Motiv (GR) eine Homologie mit der hpART1 besitzen. Die Lage der Tripletts für das GR-Motiv ist in der Abbildung markiert. Heute ist diese Region als sog. WGR-Domäne definiert.

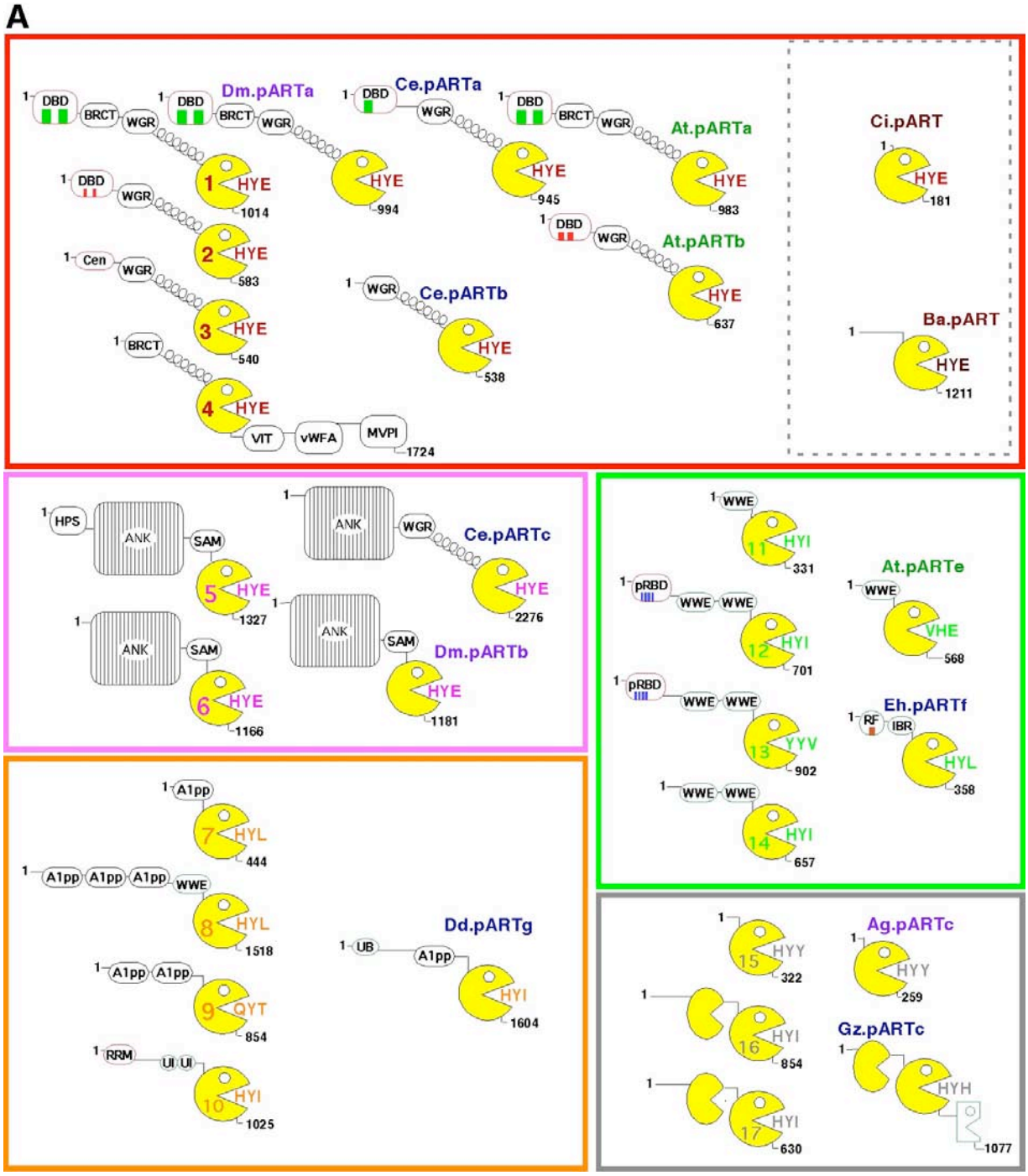
Interessanterweise wird an genau dieser Stelle das Hauptexon 6 der *Drosophila pARTa* durch ein Intron abgetrennt. Diese Grenze ist auch bei der *hpART1* zwischen den Exons 12 und 13 und bei der *ce.pARTa* zwischen den Exons 4 und 5 konserviert. Dies weist auf ein gemeinsames Vorläufergen hin, wobei die weiteren Introns möglicherweise nach der evolutionsgeschichtlichen Trennung auf dem Pfad zur Spezies *Drosophila melanogaster* verloren gingen. Durch die Hervorhebung der Regionen, die konservierte Domänen kodieren, wird deutlich, daß diese Regionen häufig durch Phase-0-Introns voneinander abgetrennt sind.

4.1.7 Analyse der Domänenstruktur verschiedener pART-Familienmitglieder

Eine unter 4.1.4 bereits angeschnittene Hypothese besagt, dass während der Evolution neue Proteine durch Neuordnung von Exons bzw. Exongruppen entstanden sind, die jeweils festgelegte katalytische Regionen, Bindungstellen oder strukturelle Elemente kodieren (sog. „Exon shuffling“). Die Entstehung neuartiger Gene durch die Mischung von Exons ist effizient, da sowohl funktionelle Einheiten erhalten bleiben als auch neue Möglichkeiten der Wechselwirkung eröffnet werden. Einzelne Domänen eines Proteins können vollständig von einem Exon kodiert werden.

Beim Vergleichen der Aminosäuresequenzen innerhalb der pART-Familie fiel auf, dass sich die signifikante Sequenzhomologie zumeist auf den Bereich der katalytischen Domäne beschränkt und die N-terminal gelegenen Bereiche oft keine Ähnlichkeiten aufweisen. Um die Funktion dieser assoziierten Regionen aufzuklären, durchsuchten wir mit den N-terminal der katalytischen Domäne gelegenen Aminosäuresequenzen die „nr“-Datenbank und verwendeten die „Conserved domain search“-Funktion der NCBI-Homepage (Wheeler *et al.* 2005). So konnten für einige Regionen signifikante Sequenzhomologien zu charakteristischen Domänen anderer Proteine nachgewiesen und somit eine Domänenstruktur vorhergesagt werden. Der Vergleich der vorhergesagten Domänenstrukturen zeigt, dass es bei der pART-Familie, abhängig von den Anforderungen im Laufe der Evolution, zu einer Assoziation der ADP-ribosylierenden katalytischen Domäne mit unterschiedlichsten anderen Domänen gekommen ist. Einige der assoziierten Domänen sind in zwei oder mehreren pARTs vorhanden. In **Abbildung 20A** ist die Domänenstruktur ausgewählter pARTs schematisch dargestellt. Die gefundenen assoziierten Domänen sind mit Angabe der ungefähren Größe (in Aminosäuren) und, soweit vorhanden, der Zugangsnummer in den Datenbanken konservierter

Proteindomänen (pfam, cd, smart) aufgelistet (**20B**). Die pART-katalytische Domäne, die die Familie charakterisiert, ist durch ein Pacman-Symbol abgebildet. Die assoziierten Domänen werden nachfolgend näher beschrieben. Die Ergebnisse aus den *in silico*-Experimenten wurden mit den kürzlich von anderen Arbeitsgruppen veröffentlichten Ergebnissen zu einigen pARTs abgeglichen, sich daraus ergebende neue Erkenntnisse wurden in die Abbildung eingearbeitet.



B

Symbol	Domänenname	Größe	Zugangsnummer
katalytische Domänen			
	pART katalytische Domäne	180	pfam00644
	trunkierte pART katalytische Domäne	120	-
	"Ubiquitin-conjugating enzyme", katalytische Domäne (UBCc)	140	cd00195
Protein-Protein-Interaktions-Domänen			
WGR	WGR	85	pfam05406
	regulierende Domäne (pART1)	135	pfam02877
SAM	SAM	65	cd00166
BRCT	BRCT	75	pfam00533
A1pp	A1pp	135	smart00506
HPS	His-Pro-Ser Region	180	-
ANK	Ankyrin	20 x 30	cd00204
VIT	VIT	130	smart00609
vWFA	vWFA	160	cd00198
MVPI	"major vault protein interacting domain"	160	-
WWE	WWE	75	pfam02825
UI	"ubiquitin interaction motif"	18	-
RF	zf-C3HC4 (RING-finger)	45	pfam00097
UB	U-Box	75	pfam04564
IBR	"In Between Ring fingers domain"	65	pfam01485
Nukleinsäuren-bindende Domänen			
DBD	zf-PARP	2 x 90	pfam00645
DBD	SAP	2 x 35	pfam02037
pRBD	zf-CCCH	4 x 30	pfam00642
Cen	Zentriolen-Lokalisierung	55	-
RRM	RNA-recognition motfv	70	pfam00076

Abbildung 20: Schematische Darstellung der Domänenstrukturen ausgewählter Mitglieder der pART-Familie (A) und Auflistung der assozierten Proteindomänen (B)

Die pART-katalytischen Domänen sind durch ein Pacman-Symbol dargestellt. Bei den humanen pARTs ist die jeweilige Nummer (pART1, 2, 3, etc.) im Pacman-Symbol eingezeichnet. Bei den pARTs der anderen Spezies ist der Name entsprechend der bei den PSI-BLAST-Analysen vorgestellten Nomenklatur angegeben, die Namen sind nach ihrer Speziesgruppe eingefärbt (Säugetiere rot, Insekten violett, niedere Eukaryonten blau, Pflanzen grün, mikrobielle Organismen braun). Das HYE-Motiv bzw. die jeweilige Variante dieses Motivs ist im „Maul" des Pacmans eingezeichnet. Die Gesamtanzahl der Aminosäuren des jeweiligen Proteins ist C-terminal angegeben. Die pART-Subgruppen sind durch farblich angepaßte Rahmen hervorgehoben: Subgruppe 1 rot, 2 rosa, 3 orange, 4 grün, 5 grau. Die assoziierten Domänen sind mit den zugehörigen Symbolen, Domänennamen, ungefähren Größen (in Aminosäuren) und Zugangscodes für die verschiedenenen Datenbanken konservierter Domänen (pfam, cd, smart) aufgelistet (**B**). Potentielle Nukleinsäure-bindende Domänen sind rot umrandet und Domänen, die vermutlich mit Ubiquitin interagieren, sind grün markiert.

a) Subgruppe 1

In N-terminaler Richtung an die katalytische Domäne anschließend besitzen alle humanen pARTs der Subgruppe 1 eine pART1-homologe Region, die aufgrund der potentiellen Ausbildung von alpha-Helices in der Abbildung als Spirale dargestellt ist. Diese Region wurde bei der hpART1 als regulatorische Domäne definiert. Mit Ausnahme von pART4, bei der die Sequenzähnlichkeit dann abbricht, schließt sich eine Region an, die sich durch ein konserviertes Tryptophan (W) – Glycin (G) – Arginin (R) - Motiv auszeichnet und inzwischen nach diesem Motiv als WGR-Domäne definiert wurde. Diese Domäne kommt u.a. in verschiedenen polyA-Polymerasen vor, die Funktion ist unklar. Auch die ausgewählten Familienmitglieder anderer Spezies, die aufgrund der Sequenzidentität dieser Subgruppe zugeordnet wurden, besitzen diese beiden assoziierten Domänen. Bei den meisten Proteinen dieser Subgruppe sind noch weitere Domänen assoziiert:

Die humane pART1 besitzt im Anschluß an die WGR-Domäne eine BRCT-Domäne, die Protein-Protein-Interaktionen ermöglicht und von der angenommen wird, dass sie mit anderen BRCT-Domänen Homo-oder Heterodimere ausbilden kann. Die Domäne wurde erstmals im Brustkrebs-Supressor-Protein BRCA1 identifiziert und wird hauptsächlich in Proteinen gefunden, die als Reaktion auf eine Schädigung der DNA in die DNA-Reparatur und in Zellzyklus-„checkpoint"-Prozesse involviert sind (z.B. XRCC1, (Zhang *et al.* 1998)). N-terminal folgt eine DNA-bindende Domäne, die zwei außergewöhnliche Zink-Finger („zf PARP-type") einschließt. Bei mehreren pARTs ist die katalytische Domäne ebenfalls mit einer DNA-bindenden Domäne fusioniert. Dabei weisen die *Drosophila* pARTa und die *Arabidopsis* pARTa eine pART1-analoge Domänenstruktur auf und enthalten zudem die zwei charakteristischen Zink-Finger-Motive. In Abweichung davon besitzt die Ce.pARTa

lediglich ein einzelnes Zink-Finger-Motiv vom PARP-Typ, die BRCT-Domäne ist bei dieser pART nicht zu identifizieren. Die Domänenstrukturen bestätigen erneut, dass es sich bei diesen pARTs um Orthologe der humanen pART1 handelt.

Das humane Paraloge pART2 (als PARP-2 publiziert (Ame *et al.* 1999)) verfügt über eine N-terminale DNA-bindende Domäne mit nukleärem Lokalisationssignal, die interessanterweise keine signifikante Homologie mit der DNA-bindenden Domäne der pART1 aufweist und der somit auch die charakteristischen Zink-Finger-Motive fehlen. Dies könnte auf eine unterschiedliche Substratspezifizät der Proteine hinweisen. Bei der *A.thaliana* pARTb ist die katalytische Domäne mit einer Domäne assoziiert, die zwei sog. SAP-Motive enthält, die in verschiedenen nukleären Proteinen vorkommen und vermutlich ebenfalls eine DNA-bindende Funktion besitzen (Aravind and Koonin 2000a). In der DNA-bindenden Domäne des hpART2 kommen trunkierte Varianten dieses Motivs vor. Vermutlich ist die at.PLARTb das Orthologe der hpART2. *C.elegans* pARTb verfügt N-terminal über keine weitere assoziierte Domäne.

Die kleinste humane pART der Subgruppe 1 ist pART3, deren cDNA unter dem Namen PARP-3 publiziert wurde (Johansson 1999). pART3 ist durch eine kleine N-terminale Domäne von 54 Aminosäuren gekennzeichnet. Die Funktion dieser Domäne war lange unklar, aber unsere Datenbankrecherchen ergaben, dass auch sie Motive enthält, die eine DNA-Bindung ermöglichen könnten. Erst kürzlich zeigten Augustin *et al.*, dass sich die N-terminale Domäne einerseits durch eine DNA-Bindungs-Kapazität auszeichnet und andererseits ein Motiv besitzt, das für die zentrosomale Lokalisation (Cen) von pART3 verantwortlich ist (Augustin *et al.* 2003).

Die von Kickhoefer *et al.* unter dem Namen vault PARP (vPARP) publizierte pART4 ist mit 1724 AS die größte humane pART. Die katalytische Domäne wird N-terminal von einer BRCT-Domäne flankiert. pART4 ist das einzige humane Familienmitglied, bei dem sich das Protein C-terminal der katalytischen Domäne mit weiteren funktionellen Domänen fortsetzt. Neben einer Domäne, die Ähnlichkeit mit einem humanen Plasma-Glykoprotein aufweist („vault protein inter-alpha trypsin domain“, VIT), ist eine Domäne assoziiert, die initial in dem an der Blutgerinnung beteiligten von Willebrand-Faktor gefunden wurde („von

Willebrand factor type A domain“, vWA). Diese Domäne, die mit verschiedenen Domänen anderer Proteine Komplexe bilden kann (z.B. mit der N-terminalen Domäne des Plättchenrezeptor-Glykokoproteins Ib alpha), konnte seitdem in vielen verschiedenen Proteinen mit unterschiedlichen Funktionen z.B. bei der Zelldifferenzierung, Adhäsion oder Hämostase nachgewiesen werden (Huizinga *et al.* 2002). Am C-terminalen Ende von pART4 liegt die MVPI-Domäne, die für die Interaktion mit dem „major vault protein (MVP)“ verantwortlich ist (Kickhoefer *et al.* 1999). Auf diese Interaktion mit dem an der Bildung von Ribonukleoprotein-Komplexen, sog. “vaults“, beteiligten MVP bezieht sich der Name vault PARP.

Die beiden viralen pARTs wurden aufgrund der Sequenzidentität und des konservierten HYE-Motivs der Subgruppe 1 zugeordnet. Während die Ci.pART aus einer isolierten katalytischen Domäne besteht, ist die katalytische Domäne bei der Ba.pART mit einer langen, N-terminal gelegenen Sequenz assoziiert, die entfernte Ähnlichkeiten zu viralen Hüllproteinen aufweist.

b) Subgruppe 2

Weitere humane pARTs mit komplexer Domänen-Struktur stellen pART5 und pART6 dar, die unter den Namen “Tankyrase 1“ und “Tankyrase 2“ veröffentlicht wurden (Kaminker *et al.* 2001, Smith *et al.* 1998). Die Gesamtsequenzidentität der beiden humanen Tankyrasen beträgt über 80%. Charakteristisches Merkmal ist eine große Domäne, die aus sich wiederholenden Ankyrin-Motiven besteht. Aus der Erforschung anderer Proteine ist bekannt, dass Ankyrin-Motive Protein-Protein-Interaktionen vermitteln. Bei pART6 fehlt im Gegensatz zu pART5 die N-terminale Region, die sich durch einen hohen Anteil an Histidin, Prolin und Serin auszeichnet (HPS). Die Protein-Interaktions-Domäne „sterile alpha motif (SAM)“ ist in verschiedenen Proteinen zu finden, die in die Signaltransduktion und die Regulierung der Transkription involviert sind. Sie besitzt die Fähigkeit, mit anderen SAM-Domänen Homo- oder Hetero-Oligomere auszubilden (De Rycker *et al.* 2003).

Interessanterweise kommen auch im Genom von *D.melanogaster* und *C.elegans* Verwandte vor, bei denen die katalytische Domäne an eine Ankyrin-Domäne fusioniert ist. Bei beiden Verwandten fehlt analog zu hpART6 die HPS-Region. Im Gegensatz zu den anderen Tankyrasen besitzt die pARTc von *C.elegans* anstelle des SAM-Motivs die pART1-

homologe, regulatorische Domäne und die WGR-Domäne. Auch die pARTb von *Dictyostelium discoideum* (nicht dargestellt) zeichnet sich durch eine große Ankyrin-Domäne aus. Sie ist mit der gleichen Domänenkomposition ausgestattet wie die Ce.PARTc, auch ihr fehlt das SAM-Motiv.

c) Subgruppe 3

Bei den humanen Paralogen pART7, pART8 und pART9 wurde N-terminal der katalytischen Domäne eine einzelne bzw. in mehreren Kopien vorkommende Domäne bisher unbekannter Funktion identifiziert, die A1pp-Domäne. Sie weist Ähnlichkeiten zu einem Abschnitt der „nicht-Histon„-Region im Makro-Histon H2A (sog. Makro-Domäne) sowie zu bestimmten Polyproteinen verschiedener RNA-Viren auf. Die Makro-Domäne existiert zudem als vollständiges Protein in einigen Bakterien-Spezies (Pehrson and Fuji 1998). Die Konservierung in der Evolution spricht für eine funktionelle Signifikanz der Region. Die Kristalisierung des Proteins AF1521 von *Archaeoglobus fulgidus*, das aus einer isolierten „Makro-Domäne“ besteht, ergab erste Hinweise auf die Funktion dieser Domäne (Allen *et al.* 2003). Seine Struktur gleicht der N-terminalen DNA-bindenden Domäne der Leucin-Aminopeptidase PepA von *E.Coli* und weist zudem eine Ähnlichkeit zu Mitgliedern der „P-loop“-Familie der Nukleotidhydrolasen auf, zu denen auch Phosphoesterasen zählen, die an ADP-Ribose-Derivaten agieren. Möglicherweise besitzt die Makro-Domäne ebenfalls die Fähigkeit, Nukleinsäuren zu binden. Während pART7 lediglich eine einzelne A1pp-Domäne enthält, verfügt pART9 über zwei und pART8 sogar über drei dieser Domänen in Reihe. Eventuell handelt es sich bei der AS-Sequenz von pART7 mit 444 AS noch nicht um die vollständige Proteinsequenz und in der Zukunft sind N-terminal weitere A1pp-Domänen zu ermitteln.

Als weiterer Unterschied ist bei pART8 zwischen die pART-katalytische Domäne und die dritte A1pp-Domäne eine WWE-Domäne eingeschoben. Diese Domäne wurde nach ihrem charakteristischen Aminosäuremotiv (Tryptophan-Tryptophan-Glutaminsäure = WWE) benannt und kommt in Ubiquitin-assoziierten Proteinen wie Deltex, einem positiven Regulator des „Notch“ Signalweges, vor (Aravind 2001). Die WWE-Domäne ermöglicht bestimmte Protein-Interaktionen (u.a. mit Ankyrinen). pART9 entspricht dem sog. BAL-

Protein (B-Aggressive Lymphoma) (Aguiar *et al.* 2000), das in der Originalpublikation nicht als potentielle pART identifiziert wurde.

Die Domänenstruktur der vierten humanen pART der Subguppe 3 unterscheidet sich deutlicher von den anderen: Statt den A1pp-Domänen besitzt pART10 zwei kurze Motive, die eine Ubiquitin-Interaktion ermöglichen könnten (UI). Dies ist nach der WWE-Domäne der zweite Hinweis auf eine Assoziation dieser pARTs mit Ubiquitin. N-terminal zeichnet sich pART9 durch ein RNA-Erkennungs-Motiv aus (RRM), das in verschiedenen RNA-bindenden Proteinen kommt (Birney *et al.* 1993). Als Beispiele sind heterogene nukleäre Ribonukleoproteine (hnRNPs), Proteinkomponenten der kleinen Kern-Ribonukleoproteine (snRNPs) und Proteine, die das alternative Splicen regulieren, zu nennen. Das Motiv kommt aber auch in einigen Proteinen vor, die Einzelstrang-DNA binden können.
Auch in den entfernteren Spezies konnte mit Dd.pARTg eine A1pp-assoziierte pART-katalytische Domäne identifiziert werden. Die N-terminal gelegene sog. U-Box-Domäne ist eine modifizierte Ring-Finger-Domäne, der die charakteristischen Zink-bindenden Aminosäurenreste fehlen (Aravind and Koonin 2000b). Für mehrere U-Box-Proteine wurde gezeigt, dass sie eine wichtige Rolle bei der Ubiquitinierung spielen, wobei die Deletion der U-Box oder die Mutation konservierter Aminosäuren innerhalb der U-Box zum Verlust der Ubiquitinierungs-Aktivität führt (Hatakeyama *et al.* 2001).

d) Untergruppe 4

Alle humanen pARTs der Untergruppe 4 (pART 11-14) sind durch die bereits bei pART8 beschriebene WWE-Domäne gekennzeichnet, wobei pART11 über eine und pART 12-14 über zwei aufeinanderfolgende WWE-Domänen verfügen. Während bei den anderen pARTs, die die WWE-Domäne enthalten, weitere N-terminale Domänen assoziiert sind, bildet die WWE-Domäne bei pART11 das N-terminale Ende des Proteins. Bei pART12 und pART13 liegt N-terminal eine Domäne, in der sich vier potentielle Zink-Finger vom CCCH-Typ ausbilden. Diese wurden bisher nur in wenigen RNA-bindenden Proteinen ausfindig gemacht, so dass für diese Domäne eine potentielle RNA-bindende Funktion postuliert wird. Die N-terminale Region von pART14 konnte bisher keiner bekannten Domänenstruktur zugeordnet werden. Das Ratten-Orthologe von pART13 wurde kürzlich unter dem Namen „Zinc-finger

Antiviral Protein“ (ZAP) näher charakterisiert (Gao *et al.* 2002), wobei die C-terminale Domäne nicht als PARP-verwandte, potentiell katalytische Domäne erkannt wurde.

Auch in der Pflanzenwelt kommt mit At.pARTe eine pART-Variante vor, die eine WWE-Domäne enthält. Die pARTf des Einzellers *Entamoeba histolytica* wurde aufgrund der Sequenzidentität im Bereich der katalytischen Domäne der Gruppe 4 zugeordnet. Sie enthält eine Domäne, die meist zwischen Paaren von Ring-Fingern vorkommt („in between ring fingers“, IBR). Die Funktion dieser Domäne ist unklar. Als Ring-Finger werden Zink-Finger vom C3HC4-Typ bezeichnet. Die Domänen sind reich an Cysteinresten und können zwei Zinkionen koordinieren. Viele Proteine, die einen Ring-Finger enthalten spielen Schlüsselrollen im Ubiquitin-Stoffwechsel. Z.B. konnte gezeigt werden, dass verschiedene, anderweitig nicht verwandte Proteine (z.B. BRCA-1, NF-X1) die E2-abhängige Konjugation des Ubiquitins modulieren können (E2 = Ubiquitin-konjugierendes Enzym 2, (Lorick *et al.* 1999)). Tatsächlich konnte bei Eh.pARTf in aminoterminaler Richtung ein die IBR-Domäne flankierender Ring-Finger entdeckt werden, ein zweiter Ring-Finger zwischen der IBR-Domäne und der pART-katalytischen Domäne fehlt. Für Eh.pARTf kann somit ebenfalls eine Beeinflussung der Ubiquitinierung vermutet werden. Interessanterweise wurden bei der Amöbe drei weitere pARTs gefunden, die eine hohe Sequenzidentität zu Eh.pARTf besitzen und ebenfalls ein HYL-Motiv aufweisen (Eh.pARTc-e). Diese Sequenzidentität bricht aber N-terminal der katalytischen Domäne ab. Trotzdem konnte auch bei Eh.pARTd ein einzelner Ring-Finger nachgewiesen werden (nicht gezeigt), die IBR-Domäne ist nicht vorhanden.

e) Subgruppe 5

Bei den einander nah verwandten pARTs 16 und 17 ist die katalytische Domäne mit einer abgewandelten und trunkierten Version der pART-katalytischen Domäne assoziiert. Wegen der ausgeprägten Sequenzvariation auch in den Bereichen, die in der katalytischen Domäne die zentralen Sekundärstrukturmotive ausbilden, ist eine katalytische Aktivität dieser duplizierten Domäne äußerst unwahrscheinlich. In der Einleitung wurde bereits beschrieben, dass es auch bakterielle ART-Toxine gibt, die eine duplizierte inaktive Domäne besitzen (z.B. *Bacillus cereus* VIP2, (Han *et al.* 1999)). Sowohl bei pART16 und 17 als auch bei pART15 und seinem Insektenorthologen Ag.pARTc sind bisher im N-terminalen Bereich der Proteine

keine konservierten Domänen zu erkennen. pART15 ist mit 322 AS die kleinste humane pART.

Die pARTc des Pilzes *Gibberella zeae* enthält ebenfalls eine duplizierte, abgewandelte pART-katalytische Domäne. Als bis zum jetzigen Zeitpunkt neben der hpART4 einziges pART-Protein setzt sich diese pART C-terminal der katalytischen Domäne mit einer weiteren assoziierten Domäne fort: Die Domäne besitzt eine hohe Sequenzidentität zur katalytischen Domäne des Ubiquitin-konjugierenden Enzymes E2, das eine zentrale Komponente bei der Ubiquitin-abhängigen Degradierung von Proteinen darstellt. Das zentrale, aktive Cystein, das die COOH-Gruppe des Ubiquitins bindet, ist bei GZ.pARTc vorhanden, so dass eine katalytische Aktivität möglich erscheint. Auch beim Schimmelpilz *Aspergillus nidulans* wurde eine pART mit dieser Domänenkomposition gefunden (GenBank: EAA66581).

Zusammenfassend ist aus den unterschiedlichen Domänenstrukturen zu schlussfolgern, dass die pART-katalytische Domäne im Laufe der Evolution im "Baukastenprinzip" mit verschiedensten Domänen, die sich sowohl strukturell als auch funktionell unterscheiden, fusioniert wurde. Das Vorkommen von einigen dieser hochkonservierten Fusionsproteine in niederen Eukaryonten und Pflanzen lässt eine wichtige funktionelle Rolle dieser Proteine vermuten. Auffällig ist, dass innerhalb der Subgruppen, die anhand der Sequenzidentitäten der pART-Domänen sowie der konservierten Intronpositionen gebildet wurden, häufig ähnliche Domänen fusioniert sind und viele pART-Proteine Domänen enthalten, die Nukleinsäuren binden können, mit dem Ubiquitin-Stoffwechsel assoziiert sind oder wichtige Protein-Protein-Interaktionen vermitteln. Dies weist, je nach Domänenkomposition, auf unterschiedliche Zielproteine, Funktionen und eine zentrale Rolle der pARTs in verschiedenen Stoffwechselwegen hin. Im Gegensatz zu den poly-ARTs bestehen alle bekannten mono-ARTs bei Vertebraten lediglich aus der isolierten katalytischen Domäne, während viele bakterielle mono-ARTs, analog zu den poly-ARTs, mit verschiedensten regulatorischen Domänen fusioniert sind (siehe 1.3.2).

4.2 Molekulare Klonierung und rekombinante Expression der humanen ADP-Ribosyltransferasen pART1, pART3 und pART4

Wie unter 4.1.1 beschrieben wurden durch Datenbankanalysen beim Menschen 17 Mitglieder der poly-ART-Familie identifiziert. pART1 und zwei der neuen Familienmitglieder, humane pART3 und humane pART4, wurden zur weiteren Analyse ausgewählt. Mit Hilfe der ESTs erstellten wir eine Konsensusnukleotidsequenz für pART3 und pART4. Diese Konsensussequenzen sowie die Lage der zur Amplifikation verschiedener Abschnitte benutzten Primer sind im Anhang B abgebildet.

4.2.1 Herstellung genspezifischer Sonden und Analyse der Expression von pART1, pART3 und pART4 in humanen Geweben

Um zu untersuchen, in welchen Geweben des Menschen und in welcher relativen Stärke die beiden *hpART*-Gene transkribiert werden, führten wir Northern-Blot- und RT-PCR-Analysen durch. Mit Primern aus verschiedenen Exonen von *hpART3* und *hpART4* wurden Sonden amplifiziert und radioaktiv markiert (siehe 3.4). Mit diesen Sonden wurden anschließend Northern-Blot-Filter auf die Expression von *hpART3* und *hpART4* gescreent. Die spezifisch gebundenen Sonden wurden autoradiographisch detektiert (**Abbildung 21**). Die Autoradiographien der Filter, die mit dem radioaktiv markiertem *hpART3*-Fragment (Primer: F65xF87) hybridisiert wurden, zeigen Banden bei 2.2kb und 2.4kb in vielen verschiedenen Geweben wie z.B. Prostata, Hoden, Ovar, Darm, Skelettmuskel, Niere und Leber. Im Gehirn zeigt sich lediglich das 2.4kb-Transskript; in der Milz wird das 2.2kb-Transskript wesentlich stärker exprimiert. Besonders ausgeprägt ist die Expression in Niere, Leber, Skelettmuskel und Herz.

Die mit einer *hpART4*-Sonde (Primer: K02xK11) hybridisierten Northern-Blots zeigen eine Einzelbande bei ca. 5.5kb, die ebenfalls in verschiedenen Geweben nachgewiesen werden konnte. Die intensivste Expression ist in Niere, Leber, Milz, Hoden und peripheren Lymphozyten zu beobachten. Im Gehirn war keine Expression zu detektieren.

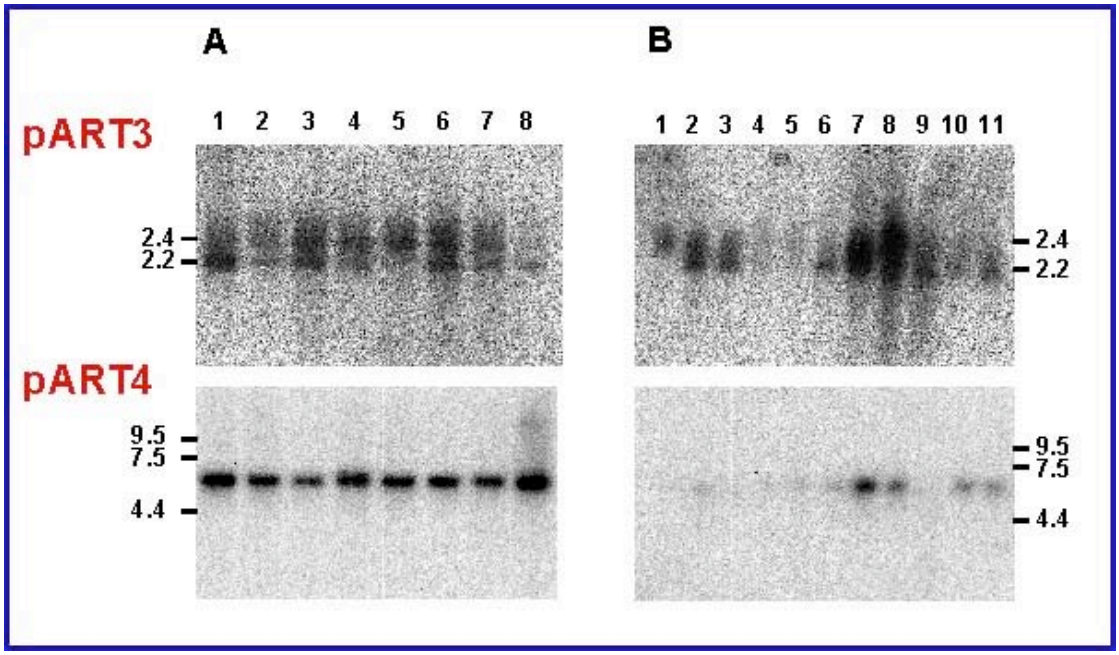


Abbildung 21: Expressionsprofil von *hpART3* und *hpART4* in humanen Geweben

Die Expression von *hpART3* und *hpART4* in verschiedenen humanen Geweben wurde mit Northern-Blot-Analysen untersucht. Es wurden Northern-Blot-Filter (Clontech) verwendet, die pro Spur 2µg poly(A)$^+$ aus den angegebenen Geweben enthalten. Die Filter wurden mit einer radioaktiv markierten *pART3*-spezifischen "F62 x F82" cDNA-Sonde bzw. einer radioaktiv markierten *pART4*-spezifischen "K02 x K11" cDNA-Sonde hybridisiert. Die Hybridisierung und das Waschen fanden unter Bedingungen geringer Stringenz statt. Die gebundenen Sonden wurden autoradiographisch mit 2 bis max. 7 Tage bei -80°C aufgelegten X-Omat-AR Filmen (Kodak) detektiert. Nach dem "Stripping" wurde zur Überprüfung der Qualität und Quantität der RNA eine Kontrollhybridisierung mit einer radioaktiv markierten G3PDH (Glycerolaldehyd-3-Phosphat-Dehydrogenase)-Sonde durchgeführt (nicht abgebildet). Die Spuren enthielten poly(A)$^+$ folgender Gewebe: **A 1** Milz, **2** Thymus, **3** Prostata, **4** Hoden, **5** Ovar, **6** Dünndarm, **7** Dickdarm, **8** periphere Blutlymphozyten, **B 1** Gehirn, **2** Herz, **3** Skelettmuskulatur, **4** Dickdarm, **5** Thymus, **6** Milz, **7** Niere, **8** Leber, **9** Dünndarm, **10** Plazenta, **11** Lunge. Marker in Kilobasen.

Die Vorteile von RT-PCR Untersuchungen bestehen darin, dass sie sensitiver sein können als Northern-Blot-Analysen und dass die amplifizierten Produkte der RT-PCR anschließend kloniert und sequenziert werden können. Es ist wichtig, die Bedingungen für die RT-PCR zu optimieren, da anderenfalls die Spezifität stark abnimmt. Um weitere Informationen über das Expressionsmuster der untersuchten Gene zu erhalten, wurde ein cDNA-Panel (MTC, „Multiple Tissue cDNA“) untersucht, das zusätzlich auch fetale Gewebe enthält. Wir setzten für diese Untersuchungen Primer-Kombinationen ein, die entweder die gesamte kodierende Region (*pART3*) oder selektiver die Region einschließen, die die katalytische

Domäne kodiert (*pART4, pART1*). Nach 29, 35 und 41 Zyklen wurden Aliquots der PCR-Reaktionen auf einem Ethidiumbromid-gefärbten Agarosegel analysiert (**Abbildung 22**).

Mit *hpART3*-spezifischen Primern wurde in allen vier getesteten Geweben (fetale Niere, fetale Milz, adulte Milz und Leukozyten) eine ca. 1600bp große Bande amplifiziert. Besonders starke Signale waren in fetaler Niere und adulter Milz nachzuweisen. Sowohl bei *hpART4* als auch bei *hpART1* konnte in allen vier getesteten Geweben (fetale Lunge, fetale Leber, adulte Milz, Leukozyten) ein ca. 1000bp großes Fragment amplifiziert werden. *pART4* zeigt die stärkste Expression in der adulten Milz. Bei der *hpART1* ist die PCR nach 35 Zyklen bereits gesättigt, so dass ein quantitativer Unterschied in der Expression nicht mehr zu erkennen ist. RT-PCR und Northern Blot Analysen zeigen, dass das *hpART3*-, das *hpART4*- und das *hpART1*-Gen ubiquitär exprimiert werden.

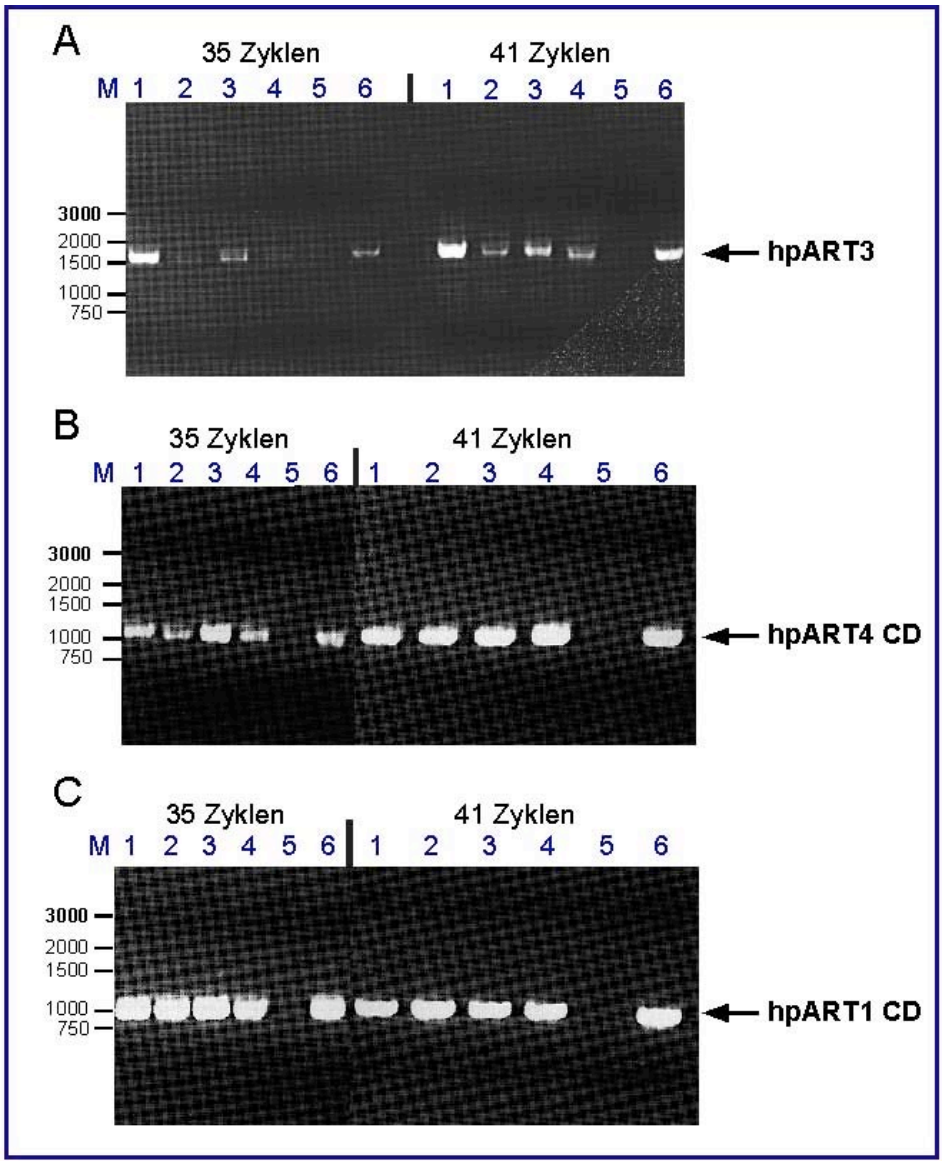


Abbildung 22: RT-PCR-Analysen im „MTC"-Panel

Die RT-PCR-Analysen wurden mit fallenden Annealing-Temperaturen in den ersten Zyklen durchgeführt (Touchdown-PCR), anschließend wurde eine hohe konstante Annealing-Temperatur gewählt. Zur Amplifikation wurde die Advantage Polymerase verwendet. Als Templates wurden cDNAs aus verschiedenen Geweben des „MTC fetal panel“ und des „MTC immune panel“ (Clontech) eingesetzt: **1** fetale Niere (A) / fetale Lunge (B,C), **2** fetale Milz (A) / fetale Leber (B,C), **3** adulte Milz, **4** Leukozyten, **5** H_2O, **6** „mixed cDNA“ Clontech. Für *hpART3* (**A)** wurde eine Primerkombination eingesetzt, die die Amplifikation der gesamten kodierenden cDNA ermöglicht (F61xF85), bei *hpART4* (**B**, K01xK11) und *hpART1* (**C**, P02xP11) eine Kombination, die die für die katalytische Domäne kodierende Region einschließt. Nach 31, 35 und 41 Zyklen wurden Aliquots entnommen und auf EthBr-gefärbten Agarosegelen analysiert. Als Marker (M) wurde der „GeneRuler 1kb ladder (MBI)“ verwendet. Die Abbildung zeigt die Agarosegele mit den Amplifikaten nach 35 bzw. 41 Zyklen.

4.2.2 Molekulare Klonierung und Sequenzierung der cDNA-Sequenz von hpART3, hpART4 und hpART1

Die Klonierung eukaryontischer Gene zur Ermittlung der Proteinsequenz oder zur heterologen Expression erfolgt meist über ihre cDNA. Da über die ESTs verschiedene Teilsequenzen von hpART3 und hpART4 bekannt waren, konnten wir daraus eine vorläufige cDNA-Konsensussequenz erstellen. Durch Vergleich mit der bereits veröffentlichten cDNA-Sequenz der humanen pART1 identifizierten wir die Bereiche in den Konsensussequenzen von hpART3 und hpART4, die für die mutmaßliche katalytische Domäne kodieren. Schließlich benutzten wir die Konsensussequenzen, um neue Primer zu entwerfen, in die Schnittstellen für Restriktionsenzyme zur Klonierung der Produkte in Expressionsvektoren eingebaut wurden. Mit diesen Primern wurde sowohl die für die katalytische Domäne des jeweiligen Proteins kodierende cDNA (pART3: F63xF85, pART4: K01xK11 bzw. K02xK11, pART1: P02xP11) als auch die vollständige cDNA-Sequenz von pART3 (F61xF85) kloniert. Die Primer F63, K02 und P02 besitzen eine NdeI-, der Primer K01 eine NcoI- und die Primer F85, K11 und P11 eine XhoI-Schnittstelle (siehe Anhang A). Als Matrize wurden die cDNAs von adulter Milz und fetaler Niere aus den „MTC-panels“ (Clontech) verwendet. Die PCR wurde als Touchdown-PCR (wie in 4.2.1) durchgeführt. Die Amplifikate wurden nach gelelektrophoretischer Auftrennung in den pCR2.1 Vektor kloniert und anschließend mit Hilfe der Didesoxy-Methode sequenziert. Die Plasmide, die Amplifikate ohne Mutationen, Deletionen oder Insertionen enthielten, wurden für die Klonierung von Expressionskonstrukten weiterverwendet.

An dieser Stelle ist anzumerken, dass im Gegensatz zu der in 4.1. vorgestellten neuen Definition der pART-katalytischen Domäne (Start mit dem konservierten Phase-0-Intron vor

β1 bei pART1) die amplifizierten cDNAs für die „katalytische Domäne“ 5´-wärts noch einen unterschiedlich großen Anteil der anschließenden Region enthalten. Exakt betrachtet handelt es sich in diesem Kapitel folglich bei der Bezeichnung „katalytische Domäne“ (gelbes Pacman-Symbol) um die katalytische Domäne mit verschieden langen Anteilen der N-terminal anschließenden pART1-homologen, regulatorischen Domäne. Die genaue Länge der assoziierten Anteile im jeweiligen Konstrukt geht aus den im Anhang B abgebildeten Konsensussequenzen hervor. Die Funktion und genaue Größe der N-terminalen Domäne des pART3 war zum Zeitpunkt der Durchführung der Experimente noch unbekannt, sie wird hier deshalb als „putative regulatorische Domäne“ bezeichnet.

4.2.3 Herstellung von Expressionskonstrukten und Expression rekombinanter Proteine

Die Charakterisierung des von einer bestimmten cDNA kodierten Genproduktes ist möglich, indem man es in einem gut bekannten Organismus exprimiert. Für die Produktion eines rekombinanten Proteines, d.h. für die gezielte Synthese eines bestimmten fremden Genproduktes, stehen zahlreiche Expressionssysteme zur Verfügung. Abhängig von den Anforderungen an das Expressionskonstrukt muss ein für die Expression geeigneter Organismus bzw. eine passende Zelllinie (Bakterien, Hefen, Insektenzellen, Säugetierzellen) sowie ein geeigneter Expressionsvektor ausgewählt werden. In den letzten Jahrzehnten wurde eine große Anzahl von verschiedenen Plasmidvektoren konstruiert, die eine Expression fremder Proteine mit hoher Ausbeute erlauben und die sich durch unterschiedliche Promotoren, Selektionsmarker und Fusionspartner auszeichnen. Die rekombinanten pARTs, die in ihrer natürlichen Form im Organismus nur in sehr geringen Konzentrationen vorkommen, können durch Wahl des geeigneten Expressionssystems auf das Tausendfache oder noch stärker angereichert werden. Die Fusion des fremden Genproduktes an bestimmte bakterielle Polypeptide kann einerseits die Löslichkeit des rekombinanten Proteins erhöhen und andererseits eine Detektion bzw. Aufreinigung des Proteins ermöglichen. Für die rekombinante Expression von pART3, pART4 und pART1 wählten wir mit dem pET-Expressionssytem (Novagen) eines der effektivsten Systeme für die Klonierung und Expression von rekombinanten Proteinen in *E.Coli* aus (siehe 1.8).

4.2.3.1 Herstellung von Expressionskonstrukten

Aus dem Sortiment von pET-Plasmidvektoren waren für unsere Klonierungsstrategie die Vektoren pET26b und pET32a am besten geeignet. Die Klonierungsstrategie ist im Kapitel 3.5.1 ausführlich beschrieben (siehe Abb.8). Für die Klonierung der Expressionskonstrukte wurden die sequenzierten, auf Mutationsfreiheit geprüften cDNA-Inserts aus den pCR2.1-Vektoren verwendet (siehe 4.2.2). Der Klonierungsvorgang ist am Beispiel der Konstrukte für die katalytische Domäne von pART3 in **Abbildung 23** auf einem Ethidiumbromid-gefärbten Gel zusammengefasst. Die mutationsfreien cDNA-Inserts wurden mit den Restriktionsenzymen NdeI/XhoI bzw. NcoI/XhoI aus den pCR2.1-Vektoren herausgeschnitten (Spur 1/2) und in die vorverdauten, linearisierten pET-Vektoren (Spur 3-5) kloniert. Der Verdau und die Klonierung der pART3- und pART1-Inserts mit NcoI ist über eine interne NcoI-Schnittstelle möglich (siehe Konsensussequenzen im Anhang B). Die Spuren 6-8 zeigen die vollendeten Expressionsvektoren, durch Verdau aufgetrennt in Vektor-"Backbone" und Insert, und damit den Erfolg der Klonierung.

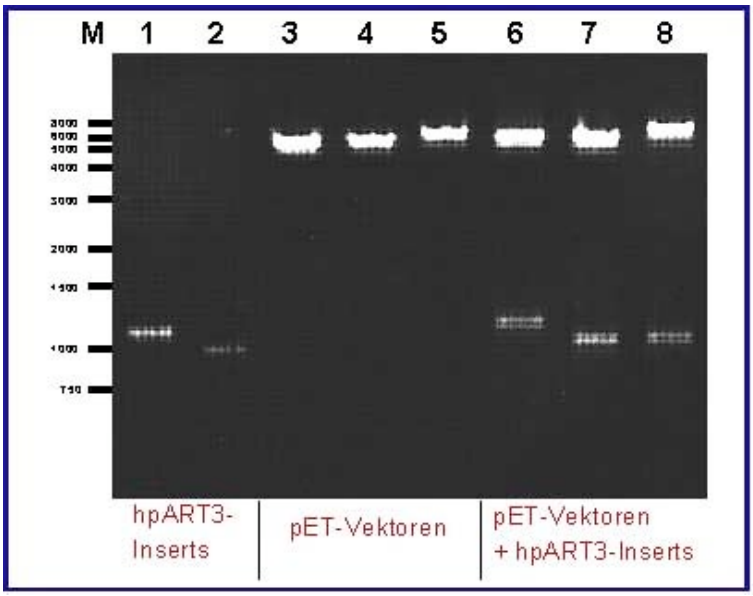


Abbildung 23: Klonierung der cDNA-Fragmente in pET-Vektoren

In der Abbildung ist ein Ethidiumbromid-gefärbtes Agarosegel gezeigt. Zur Herstellung von Expressionskonstrukten für die katalytische Domäne von hpART3 wurden aus den pCR2.1.hpART3.F63xF85-

Plasmiden mit den Restriktionsenzymen NdeI/XhoI bzw. NcoI/XhoI die Inserts herausgeschnitten, gelelektrophoretisch aufgetrennt und aufgereinigt. Aliquots der aufgereinigten Inserts sind in Spur 1 (hpART3.F63xF85.NdeI/XhoI) und Spur 2 (hpART3.F63xF85.NdeI/XhoI) aufgetragen. Die pET-Vektoren wurden ebenfalls mit NdeI/XhoI bzw. NcoI/XhoI verdaut und anschließend mit alkalischer Phosphatase (AP) aus Kälberdarm versetzt, um eine Religation zu verhindern. Nach Hitzeinaktivierung und Fällung der AP mit StrataClean Resin (Stratagene) erfolgte eine Gelelektrophorese. Aliquots der linearisierten pET-Vektoren sind in Spur 3-5 dargestellt: pET26b.NdeI/XhoI (3), pET26b.NcoI/XhoI (4) und pET32a.NcoI/XhoI (5). Vektoren und Inserts wurden ligiert und in superkompetente XL10-Zellen (Stratagene) transformiert. Aliquots der Miniprep wurden verdaut und gelelektrophoretisch analysiert (Spur 6-8): pET26b.hpART3.NdeI/XhoI (6), pET26b.hpART3.NcoI/XhoI (7) und pET32a.hpART3.NcoI/XhoI (8).

Mit den Vektoren pET26b und pET32a wurden auf diese Weise für hpART3 vier und für hpART4 und hpART1 je drei verschiedene Varianten konstruiert. Die Expressionskonstrukte sind am Beispiel von pART3 in der **Abbildung 24** schematisch dargestellt. Folgende Konstrukte wurden entworfen:

Konstrukt A: Expression der katalytischen Domäne (Pacman-Symbol) von hpART3 (pET26b.hpART3.CD.NdeI/XhoI). Bei dieser Variante wird die katalytische Domäne von pART3 isoliert im Zytoplasma produziert.

Konstrukt B: Fusion der katalytischen Domäne von hpART3 an einen sogenannten "pel B-Leader (pel)" (pET26b.hpART3.CD.PPl.NcoI/XhoI). Durch den "pelB-Leader" wird das Fusionsprotein nach der Expression in das Periplasma ausgeschleust. Bei diesem Vorgang wird der „pelB-Leader“ durch eine Signalpeptidase (SP, die Position des Spaltungssignals ist durch einen Pfeil markiert) abgespalten. Dieses Konstrukt ermöglicht es in bestimmten Fällen, durch die Herstellung von Periplasma-Lysaten den prozentualen Anteil von rekombinantem Protein am Gesamtprotein zu vergrößern. Im Falle einer Toxizität des rekombinanten Produktes kann die Auswirkung auf den Zellstoffwechsel durch die Ausschleusung des Proteines in das Periplasma verringert werden. Die Produktion von löslichem Protein im periplasmatischen Spalt dient ferner als grober Test für die Faltbarkeit des Proteins im nichtreduzierten Milieu, das dem Milieu des Extrazellulärraumes entspricht.

Konstrukt C: Fusion der katalytischen Domäne von pART3 mit Thioredoxin (Trx) (pET32a.Trx-hpART3.CD.NcoI/XhoI). Die Fusion mit Thioredoxin erhöht nach der Expression die Löslichkeit des Fusionsproteins im Zytoplasma von *E.Coli.* In das Konstrukt sind ein "S-tag", der die Aufreinigung des Proteins sowie die Markierung über Protein S ermöglicht, sowie ein interner "His-tag" eingebaut. Der Trx-Anteil mit den beiden „tags“ kann durch Behandlung mit Enterokinase (EK) abgespalten werden.

Konstrukt D: Expression des gesamten hpART3-Proteins mit der N-terminalen putativen regulatorischen Domäne (pRD) im Zytoplasma (pET26b.hpART3.long.NdeI/XhoI).

Alle Konstrukte besitzen einen C-terminalen "His-tag". Zur Selektion enthält der pET26b-Vektor eine Kanamycin-, der pET32a-Vektor eine Ampicillin-Resistenz. Die Konstrukte A,B und C wurden analog für hpART4, hpART1 sowie das "Salmonella plasmid virulence B (SpvB)" (siehe 4.3) kloniert.

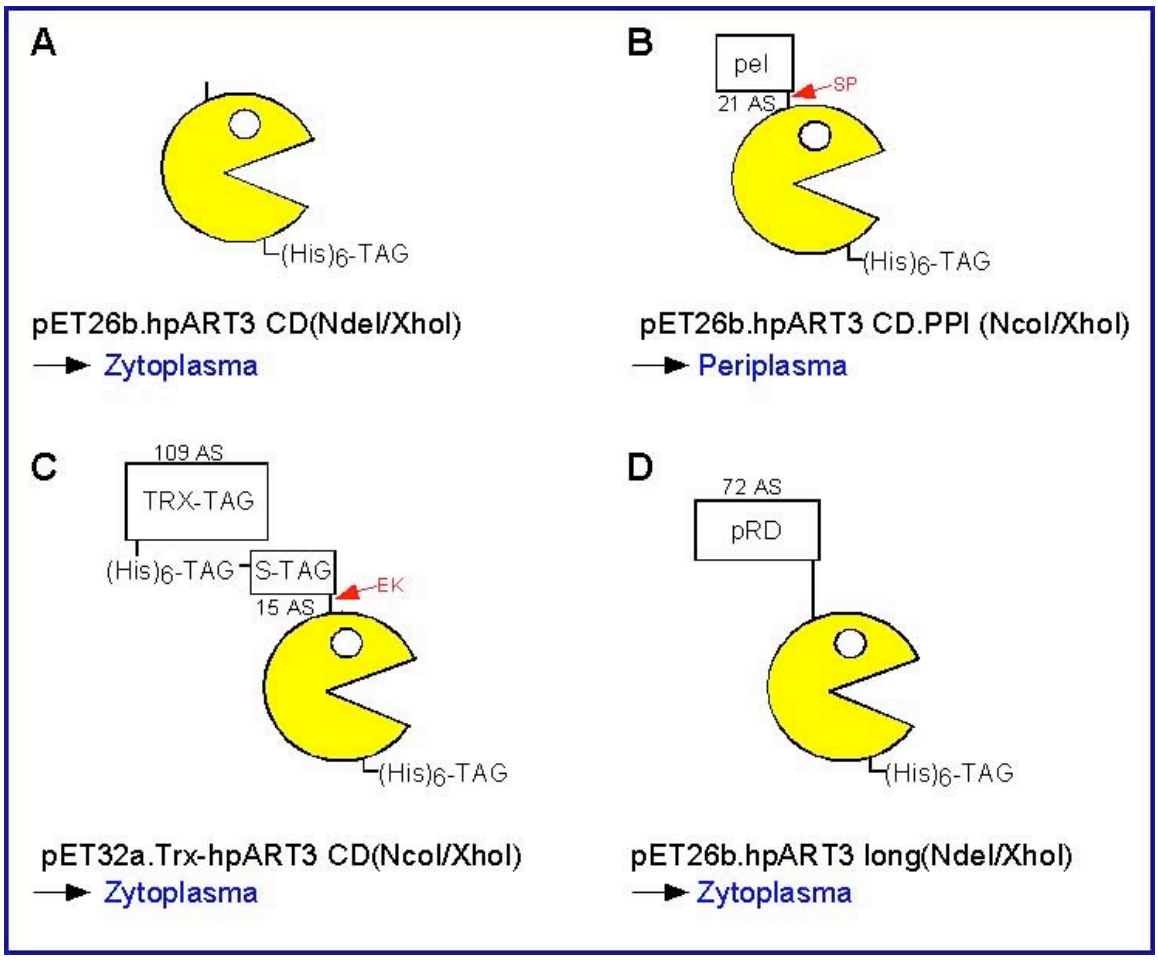


Abbildung 24: Schematische Darstellung der verschiedenen Expressionskonstrukte am Beispiel von hpART3

Für die Expression der rekombinanten Proteine wurde das pET-Expressionssystem (Novagen) ausgewählt. Mit den Vektoren pET26b und pET32a wurden für hpART3 vier und für hpART4 und hpART1 je drei verschiedene Varianten der rekombinanten Proteine konstruiert. In der Abbildung sind die vier verschiedene Konstrukte für hpART3 schematisiert dargestellt. Die Besonderheiten der Konstrukte sind im Text beschrieben. SP = Abspaltungsstelle des „pel B-leaders“ nach Ausschleusung in das Periplasma. EK = Enterokinase-Signalsequenz.

4.2.3.2 Expression von rekombinantem pART3, pART4 und pART1 in *E.Coli*

Als geeignete "Wirtszellen" für die Expression wurden BL21(DE3)- und BL21(DE3)pLysS-*E.Coli*-Stämme ausgewählt. Beide Zelllinien sind Lysogene des *lambda-Phagen DE3* und besitzen somit das T7-RNA-Polymerase-Gen unter Kontrolle des *lac*UV5-Promoters (siehe 1.8.). Durch die regulierbare Expression kann vermieden werden, dass eine für den Bakterienstoffwechsel möglicherweise toxische rekombinante pART während des Wachstums der Kultur die Etablierung des Plasmids negativ beeinflusst. Die Konstrukte wurden mittels Hitzeschock in BL21(DE3)- und BL21(DE3)pLysS-*E.Coli*-Stämme transformiert. Anschließend züchteten wir mit diesen transformierten Stämmen Kulturen an. Kurz vor der Zellernte wurde die Expression des Genproduktes durch Isopropylthiogalactosid (IPTG) induziert. Nach der Ernte stellten wir Gesamtzell-Lysate her, aus denen Gesamtproteine bzw. die Fraktion der löslichen Proteine gewonnen wurden (siehe 3.7.2). Anschließend wurden die Proteine mittels SDS-PAGE größenfraktioniert und durch Coomassie-Färbung sichtbar gemacht.

In den **Abbildungen 25**, **26** und **27** sind die Coomassie-gefärbten Tricin-Gele mit den aufgetrennten Gesamtproteinen (**A**) und löslichen Proteinen (**B**) der einzelnen Lysate dargestellt. Das Auftreten von prominenten zusätzlichen Banden, deren Größe den *in silico* errechneten Molekulargewichten der rekombinanten Proteine entspricht, zeigt, dass sämtliche pART3-Konstrukte in beiden Zellinien im Überschuss produziert wurden (**Abb.25A**). Durch den Vergleich der Banden in der Gesamtproteinfraktion und in der Fraktion der löslichen Proteine wird jedoch deutlich, dass ein Großteil der rekombinanten Proteine unlöslich, am ehesten in Form sog. "inclusion bodies", vorliegt. Der höchste lösliche Anteil wurde beim Trx-pART3-Konstrukt erreicht (**Abb.25B**, Spur 5/6). Insgesamt scheint die Expression von löslichem Protein in den BL21(DE3)-Zellen effizienter zu sein als in den Stämmen, die zusätzlich Lysozym enthalten. Dies fällt besonders beim Konstrukt mit dem höchsten Molekulargewicht auf (**Abb.25B**, Spur7/8). In den Lysaten der pLysS-Stämme ist das Lysozym als zusätzliche kleine Bande sichtbar.

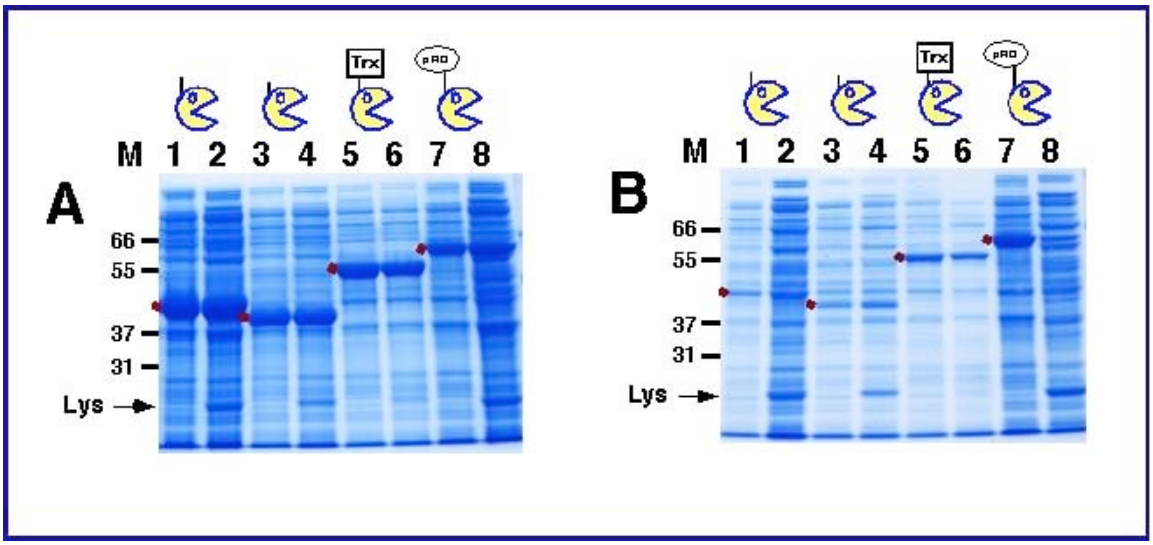


Abbildung 25: Expression von hpART3-Konstrukten in *E.Coli*

Die hpART3-Konstrukte wurden in BL21(DE3)- und BL21(DE3)pLysS-Zellen transformiert. Nach Anlage von Kulturen wurde die Expression der rekombinanten Proteine mit IPTG induziert. Durch mehrmaliges Auftauen und Einfrieren sowie Ultraschallbehandlung wurden die Proteine aus den Bakterienzellen freigesetzt. Um die Gesamtproteine und die Fraktion der löslichen Proteine getrennt untersuchen zu können, wurden die Proben geteilt und die unlöslichen Proteine bei dem einem Teil durch hochtourige Zentrifugation pelletiert. Anschließend wurden die Proteine mittels SDS-PAGE größenfraktioniert und mit Coomassie-Lösung angefärbt. In der Abbildung sind die Coomassie-gefärbten Tricin-Gele mit den aufgetrennten Gesamtproteinen (**A**) bzw. löslichen Proteinen (**B**) abgebildet. Die Banden, die durch die rekombinanten Proteine hervorgerufen werden, sind durch rote Punkte markiert. Die rekombinanten Varianten von pART3 sind oberhalb der Gele schematisch abgebildet. Die Spuren enthalten die Lysate der mit folgenden Konstrukten transformierten *E.Coli*-Stämme (in Klammern sind die *in silico* berechneten Molekulargewichte der rekombinanten Proteine angegeben): **1/2** pET26b.pART3-CD (42.0kD), **3/4** pET26b.pART3-CD.PPl (37.3kD), **5/6** pET32a.Trx-pART3-CD (54.5kD), **7/8** pET26b.pART3.long (61.8kD). In den Spuren 1, 3, 5 und 7 wurden die Lysate der BL21(DE3)-Zellen, in 2, 4, 6 und 8 die der BL21(DE3)pLysS-Zellen aufgetragen. Die durch das Lysozym hervorgerufene Bande ist mit einem Pfeil gekennzeichnet. Als Marker wurde MARK 12 verwendet (Novex, Angabe in kD). pRD = putative regulatorische Domäne.

Bei den pART4-Konstrukten (**Abbildung 26**) zeigt sich ein ähnliches Bild: Es sind zusätzliche, dem errechneten MG entsprechende Banden nachweisbar, die insgesamt weniger prominent sind als bei pART3. Die starke Bande bei den in das Periplasma ausgeschleusten Konstrukten (**A**, Spur 3/4) wird durch einen großen Anteil an unlöslichem Protein

hervorgerufen. Die höchste Konzentration an löslichem rekombinanten Protein wurde mit dem Trx-pART4-Konstrukt in den Lys-Zellen erreicht (**Abb.26B**, Spur 6).

Die Konstrukte der pART1 (**Abbildung 27**) werden ebenfalls stark exprimiert. Analog zu pART3 und pART4 wird bei der Trx-gekoppelten Variante am meisten lösliches Protein produziert (Spur 5/6). Zum Vergleich sind in Spur 7 und 8 die Lysate von *E.Coli* aufgetragen, die mit dem reinen pET32a-Vektor (ohne Insert) transformiert wurden. Als Produkt zeigt sich das durch den Vektor kodierte Trx-Protein mit einer prominenten Bande bei ca. 28kD. Aus den Bakterienkulturen, die mit den „pelB-leader"-Konstrukten (Konstrukt B) transformiert worden waren, wurden zudem Periplasma-Lysate hergestellt. Diese erbrachten aber keine ausreichende Ausbeute an löslichem, rekombinanten Protein (hier nicht gezeigt).

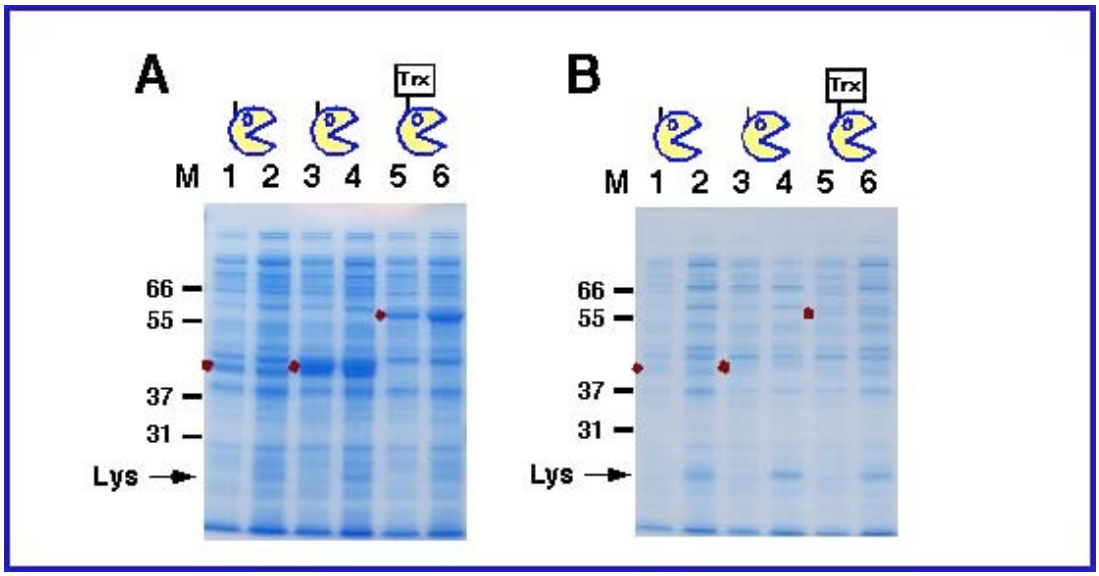


Abbildung 26: Expression von hpART4-Konstrukten in *E.Coli*

Die rekombinante Expression der hpART4-Konstrukte wurde analog zu dem in Abb.24 für hpART3 beschriebenen Verfahren durchgeführt. Die Abbildung zeigt die Coomassie-gefärbten Gele mit den aufgetrennten Gesamtproteinen (**A**) bzw. löslichen Proteinen (**B**). Die Spuren enthalten die Lysate der mit folgenden Konstrukten transformierten *E.Coli*-Stämme (in Klammern sind die *in silico* vorhergesagten Molekulargewichte der rekombinanten Proteine angegeben): **1/2** pET26b.pART4-CD (41.0kD), **3/4** pET26b.pART4-CD.PPl (41.0kD), **5/6** pET32a.Trx-pART4-CD (58.1kD). In den Spuren 1, 3 und 5 wurden die Lysate der BL21(DE3)-Zellen, in 2, 4 und 6 die der BL21(DE3)pLysS-Zellen aufgetragen. Die Banden, die durch die rekombinanten Proteine hervorgerufen werden, sind durch rote Punkte markiert. Die rekombinanten Varianten von pART4 sind oberhalb der Gele schematisch abgebildet. Die durch das Lysozym hervorgerufene Bande ist mit einem Pfeil gekennzeichnet. Als Marker wurde MARK 12 verwendet (Novex, Angabe in kD).

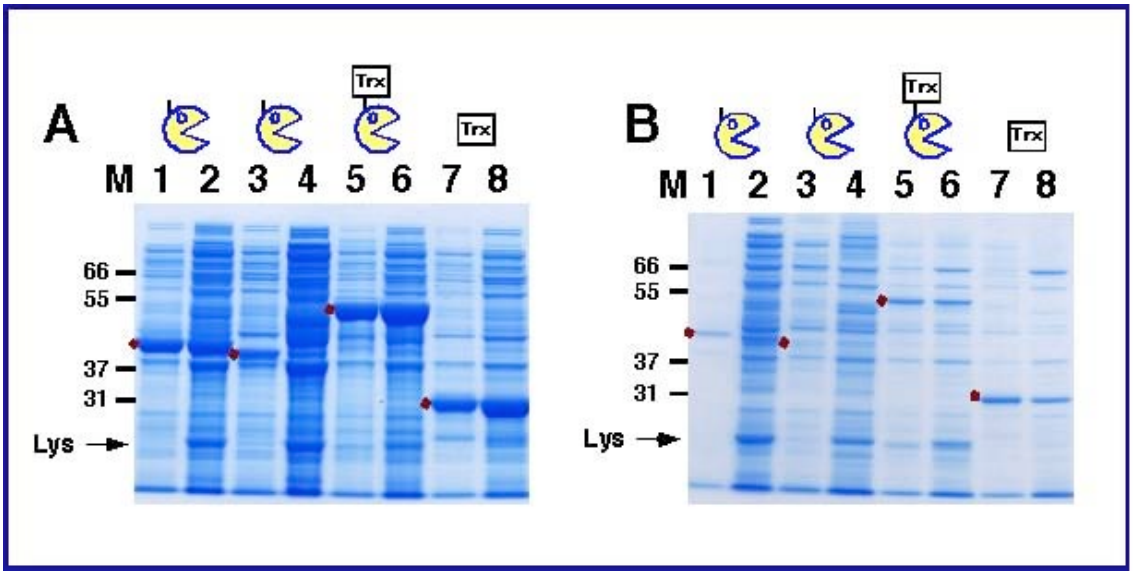


Abbildung 27: Expression von hpART1-Konstrukten in *E.Coli*

Die Expression der hpART1-Konstrukte wurde wie in Abb.24 für hpART3 geschildert durchgeführt. In der Abbildung sind Coomassie-gefärbte Gele mit den aufgetrennten Gesamtproteinen (**A**) bzw. löslichen Proteinen (**B**) abgebildet. Die Banden, die durch die rekombinanten Proteine hervorgerufen werden, sind durch rote Punkte markiert. Die rekombinanten Varianten von hpART1 sind oberhalb der Gele schematisch abgebildet. Die Spuren enthalten die Lysate der mit folgenden Konstrukten transformierten *E.Coli*-Stämme (das *in silico* ermittelte Molekulargewicht der rekombinanten Proteine ist in Klammern angegeben): **1/2** pET26b.pART1-CD (41.8kD), **3/4** pET26b.pART1-CD.PPl (37.8kD), **5/6** pET32a.Trx-pART1-CD(55.0kD),**7/8** pET32a.Trx ohne Insert (20.4kD). In den Spuren 1, 3, 5 und 7 wurden die Lysate der BL21(DE3)-Zellen, in 2, 4, 6 und 8 die der BL21(DE3)pLysS-Zellen aufgetragen. Die durch das Lysozym hervorgerufene Bande ist mit einem Pfeil gekennzeichnet. Als Marker wurde MARK 12 verwendet (Novex, Angabe in kD).

4.2.4 Enzymassays

In früheren Arbeiten hatten Mitarbeiter unserer Arbeitsgruppe beobachtet, dass die Aktivität rekombinanter mono-ARTs in *E.Coli* durch Inkubation von bakteriellen Lysaten mit [^{32}P]-NAD^{+} nachzuweisen ist, indem durch anschließende SDS-PAGE-Autoradiographie die kovalente und damit SDS-resistente Inkorporation von [^{32}P]-ADP-Ribose in Zielproteine sichtbar gemacht wird (Haag *et al.* 1995, Karsten *et al.* 1997, Koch-Nolte *et al.* 1996b). Dabei wird die Markierung der Zielproteine durch radioaktive ADP-Ribose in Anwesenheit eines starken Überschusses von unmarkierter, sogenannter „kalter“ ADP-Ribose durchgeführt. Der Überschuss an „kalter“ ADP-Ribose ist wichtig, um sicherzustellen, dass die zu beobachtende Protein-Markierung enzymatisch und nicht durch nicht-enzymatische

Anheftung von reaktiver ADP-Ribose, die durch Hydrolyse von NAD^+ entstehen kann, bedingt ist.

4.2.4.1 Nachweis der ADP-Ribosyltransferase-Aktivität von hpART4 und hpART1

Um die ADP-Ribosyl-Transferase-Aktivität in *E.Coli*-Lysaten zu untersuchen, die rekombinantes pART4 bzw. pART1 enthalten, wurden verschiedene *E.Coli*-Lysate mit radioaktiv markiertem [^{32}P]-NAD^+ inkubiert. Die Proteine wurden präzipitiert und mittels SDS-PAGE-Autoradiographie analysiert (**Abbildung 28A**). In den mit Gesamtzell-Lysaten beladenen Spuren 1-3 tritt eine markierte endogene Bande bei ca. 60kD auf. Während in der Negativkontrolle (pET26b ohne Insert, Spur1) keine weitere Bande markiert wird, ist in Spur 2 eine Bande mit inkorporierter Aktivität von ca. 41kD nachweisbar, die der Größe des rekombinanten pART4 entspricht. Dies lässt vermuten, dass rekombinantes pART4 die Funktion besitzt, sich selbst zu ADP-ribosylieren.

In den Lysaten mit rekombinanter hpART1 (Spur 3) wird neben der endogenen Bande eine Bande von ca. 44kD markiert. Das *in silico* errechnete Molekulargewicht für das rekombinante hpART1 beträgt allerdings lediglich 41.8kD und auch die beobachtete Bande in der Coomassie-Färbung liegt in diesem Bereich (Abb.27A, Spur 1 / Abb.28A Spur 3). Hierfür gibt es zwei mögliche Erklärungen: 1. Die katalytische Domäne der humanen pART1 katalysiert die ADP-Ribosylierung eines 44kD-*E.Coli*-Proteins. 2. Da bei der hpART1 die Bildung von verzweigten Polymeren („branching“) durch Auto-Poly-ADP-Ribosylierung bekannt ist, könnte sich diese Zunahme des Molekulargewichts durch die Anheftung von mehreren ADP-Ribosegruppen erklären. Die 44kD-Bande könnte somit dem poly-ADP-ribosylierten rekombinanten hpART1 entsprechen. Allerdings würde dies bedeuten, dass nur ein kleiner Teil des exprimierten rekombinanten Proteins ADP-ribosyliert wurde, da die Hauptfraktion bei ca. 41kD (Coomassie-Färbung) zu finden ist.

Auffällig ist außerdem, dass die Markierung der endogenen Bande in den Spuren 2 und 3 deutlich stärker ist als bei der Negativkontrolle in Spur 1, so dass postuliert werden kann, dass rekombinantes hpART4 und hpART1 Zielprotein(e) im Bereich von 60kD ADP-ribosylieren oder durch ihre Anwesenheit die endoge ADP-Ribosylierung aktivieren. Als Positivkontrolle wurden Periplasma-Lysate von Ratten-ART2.2 verwendet (Spur 4), dessen Fähigkeit zur Auto-ADP-Ribosylierung bereits nachgewiesen wurde (Haag *et al.* 1995).

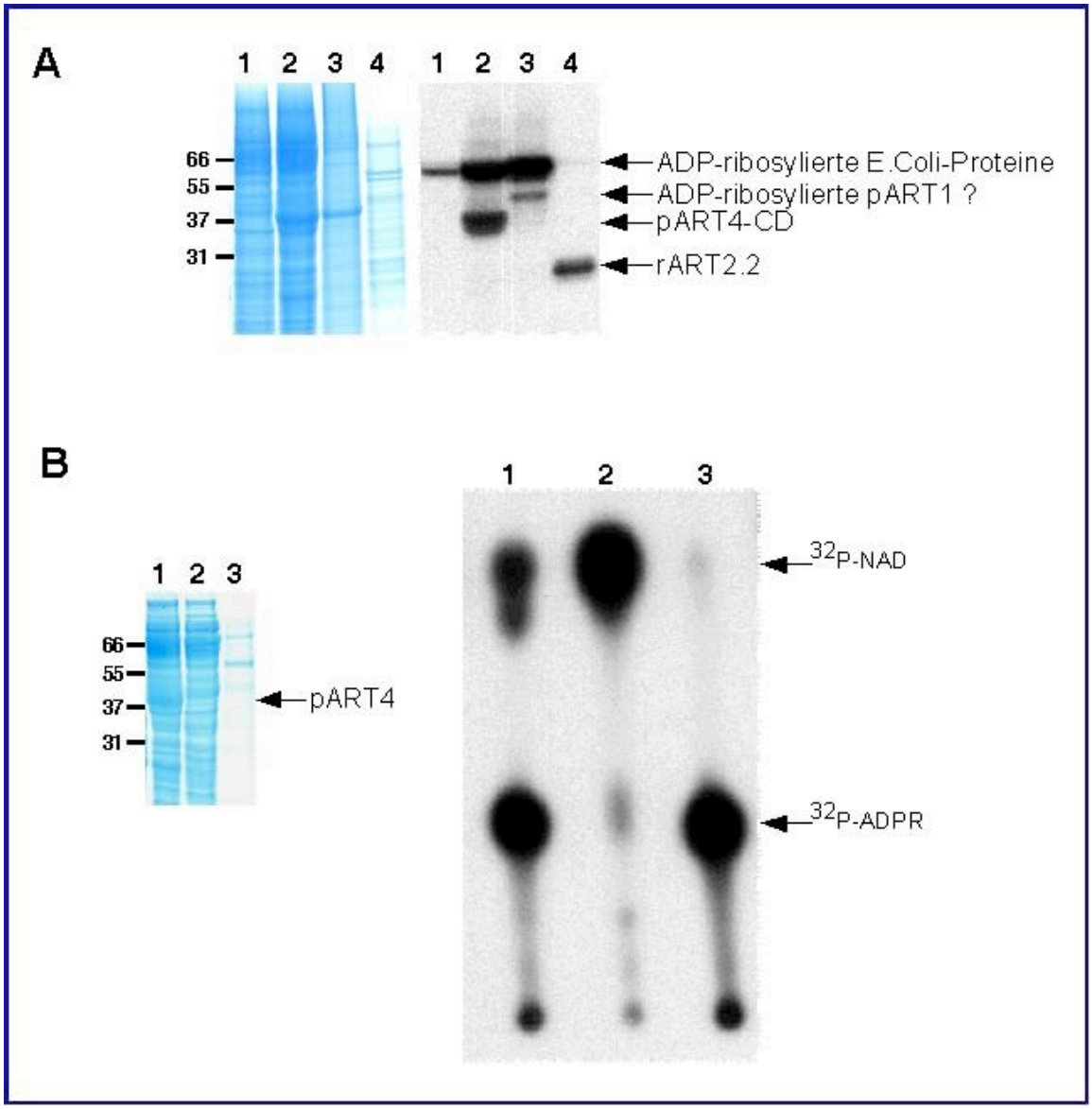


Abbildung 28: Nachweis der enzymatischen Aktivität von rekombinantem hpART4

Die die verschiedenen rekombinanten Proteine enthaltenden *E.Coli*-Lysate wurden 60 Minuten mit radioaktiv markiertem (^{32}P)-NAD^{+} inkubiert. Nach Präzipitation der Proteine mit StrataClean Resin (Stratagene) wurden die Proteine mittels SDS-PAGE größenfraktioniert. **A** zeigt einen Ausschnitt aus dem Coomassie-gefärbten Gel sowie die Autoradiographie dieses Geles nach Exposition eines Röntgenfilmes (BioMax MR, Kodak) für 5 Tage bei -80°C. Auf die Spuren wurden die Lysate der mit folgenden Konstrukten transformierten *E.Coli*-Stämme aufgetragen (*in silico* errechnete MG in Klammern): **1** Negativkontrolle pET26b ohne Insert, **2** pET26b.pART4 CD (41.0kD), **3** pET26b.pART1 CD (41.8kD), **4** pASK.Ratten-ART2.2-FH (28.1kD).

Analog zu A wurden *E.Coli*-Lysate 60 Minuten mit (^{32}P)-NAD inkubiert, die Proteine gefällt und anschließend einer SDS-PAGE-Analyse zugeführt. Die Überstände wurden mittels Dünnschichtchromatographie analysiert. **B** zeigt das Coomassie-gefärbte Gel nach SDS-PAGE sowie eine Autoradiographie der Dünnschichtchromatographie-Platte (Film: BioMax MR, Kodak. Exposition: 14h bei –80°C). Die Spuren enthalten folgende rekombinante Proteine: **1** pET26b.pART4 CD, **2** Negativkontolle pET26b ohne Insert, **3** Ratten-ART2.2. Angabe des Markers (M = Mark 12, Novex) in kiloDalton (kD). ADPR=ADP-Ribose.

4.2.4.2 Nachweis der NAD-Glykohydrolase-Aktivität von hpART4

Um zu prüfen, ob rekombinantes pART4 eine NAD-Glykohydrolase-Aktivität besitzt, wurden die *E.Coli*-Lysate analog zum vorherigen Versuch mit radioaktiv markiertem [^{32}P]-NAD^{+} inkubiert. Anschließend wurden die Proteine präzipitiert und durch SDS-PAGE aufgetrennt. Die Überstände wurden mittels Dünnschichtchromatographie untersucht (**Abbildung 28B**). Die Autoradiographie der Dünnschichtchromatographie-Platte zeigt, dass durch die Katalyse von Ratten-ART2.2 (Spur 3) nahezu das gesamte radioaktiv markierte NAD^{+} in (^{32}P)-ADP-Ribose und Nikotinamid gespalten wurde (bekannte NAD-Glykohydrolase-Aktivität). Rekombinantes pART4 (Spur 1) hat ebenfalls einen großen Anteil des (^{32}P)-NAD^{+} hydrolysiert, während die Überstände bei der pET26b-Negativkontrolle (Spur 2) keine Aktivität zeigen. Rekombinantes pART4 besitzt folglich in *vitro* eine NADase-Aktivität. Diese ist jedoch schwächer als beim Ratten-ART2.2.

Trotz diverser Versuche gelang es uns nicht, für die rekombinanten Varianten des hpART3 eine reproduzierbare ADP-Ribosyltransferase-Aktivität oder NADase-Aktivität nachzuweisen.

4.2.4.3 Detektion des rekombinanten hpART4 über den „S-tag“

Die Trx-gekoppelte Variante der rekombinanten Proteine bietet zwei verschiedene Möglichkeiten zur Detektion und Aufreinigung. Zum einen können das intramolekulare und das C-terminale Polypeptid aus sechs aufeinanderfolgenden Histidin-Resten (sog. „His-tag“) durch anti-His-Antikörper spezifisch erkannt oder über bivalente Kationen (z.B. Ni^{2+}) an eine immobilisierte Matrix gebunden werden. Zum anderen beinhaltet das Trx-Konstrukt den sog. „S-tag“, ein 15 AS umfassendes Peptid, das spezifisch und mit hoher Affinität Protein-S binden kann. Dieser „S-tag„ wurde im nächsten Experiment zur Detektion des rekombinanten Trx-Fusionsproteins verwendet. Nach Größenfraktionierung mittels SDS-PAGE wurden die Proteine der *E.Coli*-Lysate auf eine NC- sowie eine PVDF-Membran geblottet. Die PVDF-Membran wurde zunächst mit Protein-S-gekoppelter Meerrettich-Peroxidase und nach mehrfachem Waschen mit einer Chemilumineszenz-Reagenz inkubiert. Zum Nachweis der enzymatischen Chemilumineszenz-Reaktion wurde eine spezieller Film aufgelegt. **Abbildung 29A** zeigt, dass die bekannte prominente Bande in dem Trx-hpART4 enthaltenden Lysat intensiv Protein-S bindet und deshalb auf dem Film kräftig nachzuweisen ist (Spur 3). Somit bestätigt sich, dass diese Bande dem rekombinanten Trx-hpART4-Fusionsprotein entspricht.

Auch das im dritten Abschnitt des Ergebnisteils vorgestellte Trx-SpvB-Fusionsprotein wird durch die Chemilumineszenz-Reaktion intensiv detektiert (Spur 2). In Spur 1 ist eine schwächere breitere Bande bei ca. 20kD sichtbar, die dem Protein-S-gebundenen isolierten Trx-Protein entsprechen könnte, das ebenfalls den „S-tag“ enthält.

4.2.4.4 Nachweis der Enzymaktivität und Auto-ADP-Ribosylierung von aufgereinigtem Trx-pART4

Durch Inkubation der *E.Coli*-Lysate mit [^{32}P]-NAD^{+} konnte gezeigt werden, dass rekombinantes hpART4 eine ART-Aktivität besitzt oder zumindest endogene ART-Aktivität induziert. Da das ADP-ribosylierte Zielprotein mit ca. 41kD der vorhergesagten Größe von pET26b.hpART4-CD entspricht, wurde postuliert, dass sich die katalytische Domäne von pART4 auto-ADP-ribosyliert. Um diese Vermutung näher zu untersuchen, wurden Trx-pART4 und die Trx-Protein-Kontrolle durch Absorption an immobilisierte Protein-S-Agarose aus den *E.Coli*-Lysaten aufgereinigt. Die aufgereinigten Proben wurden mit [^{32}P]-NAD^{+} inkubiert, anschließend mittels SDS-PAGE aufgetrennt und durch Autoradiographie analysiert (**Abbildung 29B**). In Spur 1 des Coomassie-gefärbten Geles ist die ca. 20kD große Bande der aufgereinigten Trx-Kontrolle zu sehen. Im Gegensatz zu dem in Abb.28 gezeigten Experiment ist nach der Aufreinigung keine stark markierte endogene 60kD-Bande mehr nachweisbar (Spur 1,2). Während in der Negativ-Kontrolle keine Markierung nachweisbar ist, wird beim aufgereinigten Trx-pART4-Lysat eine ca. 55kD große Bande markiert (Spur 2), die ungefähr der *in silico* errechneten Größe von Trx-pART4 (58,1kD) und der zuvor in den Gesamtzell-Lysaten beobachteten Trx-pART4 Fusionsproteinbande entspricht (vgl. Abb.26A, Spur 5,6). Sowohl die isolierte katalytische Domäne von pART4 (Abb.28) als auch aufgereinigtes Trx-pART4 besitzen somit die Fähigkeit, ein Produkt in der jeweiligen Größe des rekombinanten Proteins zu ADP-ribosylieren. Dies lässt auf eine Auto-ADP-Ribosylierungsfunktion des pART4 schließen.

4.2.4.5 Ermittlung der Zielregion der ADP-Ribosylierung im rekombinanten Trx-pART4-Fusionsprotein

Die radioaktive Markierung von Trx-pART4 könnte durch zwei verschiedene Mechanismen erklärt werden:

1) Die katalytische Domäne von pART4 modifiziert sich selbst; analog zur Automodifikation, die bei anderen ARTs, z.B. Ratten-ART2.2, beschrieben wurde.
2) Die assoziierte N-terminale Trx-Domäne wird durch eine Heteromodifikation markiert, da sie zufällig strukturell ein Substrat für die katalytische ADP-Ribosylierung darstellt.

Um zwischen diesen beiden Möglichkeiten zu unterscheiden, versuchten wir die Zielregion für die ADP-Ribosylierung des Trx-pART4-Fusionsproteins zu ermitteln. Die Trx-Domäne und der "S-tag" im Trx-pART4-Fusionsprotein werden durch ein Enterokinase-Spaltungssignal (DDDDK) von der katalytischen Domäne des pART4 abgegrenzt. Dieses Spaltungssignal nutzten wir, um die beiden Anteile des Fusionsproteins zu trennen und anschließend einzeln darzustellen. Aufgereinigtes, radioaktiv markiertes Trx-pART4 wurde durch Inkubation mit Enterokinase am Spaltungssignal proteolytisch gespalten. Anschließend wurden die Spaltungsprodukte mittels SDS-PAGE aufgetrennt und die inkorporierte Aktivität autoradiographisch dargestellt (**Abbildung 29B,** Spur 3 und 4). Zur besseren Übersichtlichkeit sind die einzelnen Proteinanteile vor und nach der Spaltung mit dem jeweils errechneten Molekulargewicht schematisch dargestellt.

Durch den Enterokinase-Verdau wurde ein kleines Fragment des Trx-Kontroll-Proteins abgespalten, so dass die Bande im elektrischen Feld nun etwas weiter gewandert ist (Spur 3). An der Abnahme der 55kD-Bande sowie dem neuen Auftreten der abgespaltenen Trx-Domäne als 17kD-Bande bzw. der isolierten katalytischen Domäne von pART4 als 41kD-Bande wird deutlich, dass ein Teil des Trx-pART4-Fusionsproteins verdaut wurde (Spur 4). Die Autoradiographie von Spur 4 zeigt, dass die Aktivität nicht in der 17kD-Bande (Trx-Domäne), sondern in der 41kD-Bande und in einer kleinen, neu aufgetretenen, ca. 10kD großen Bande inkorporiert ist. Die ADP-ribosylierten Regionen liegen folglich im Bereich der katalytischen Domäne (41kD-Bande). Dieses Experiment zeigt, dass die katalytische Domäne von pART4 die Fähigkeit besitzt, sich selbst zu modifizieren. Dies konnte bereits durch die Automodifikation des Konstruktes mít der katalytischen Domäne ohne Fusionspartner vermutet werden (Abb.28). Das Auftreten der zusätzlichen markierten Banden könnte möglicherweise durch ein internes Enterokinase-Spaltungssignal in der katalytischen Domäne zu erklären sein, so dass Spaltprodukte der katalytischen Domäne entstanden sind. Ein klassisches DDDDK-Spaltungssignal konnte in der Aminosäuresequenz der katalytischen Domäne von pART4 allerdings nicht nachgewiesen werden (siehe Anhang B).

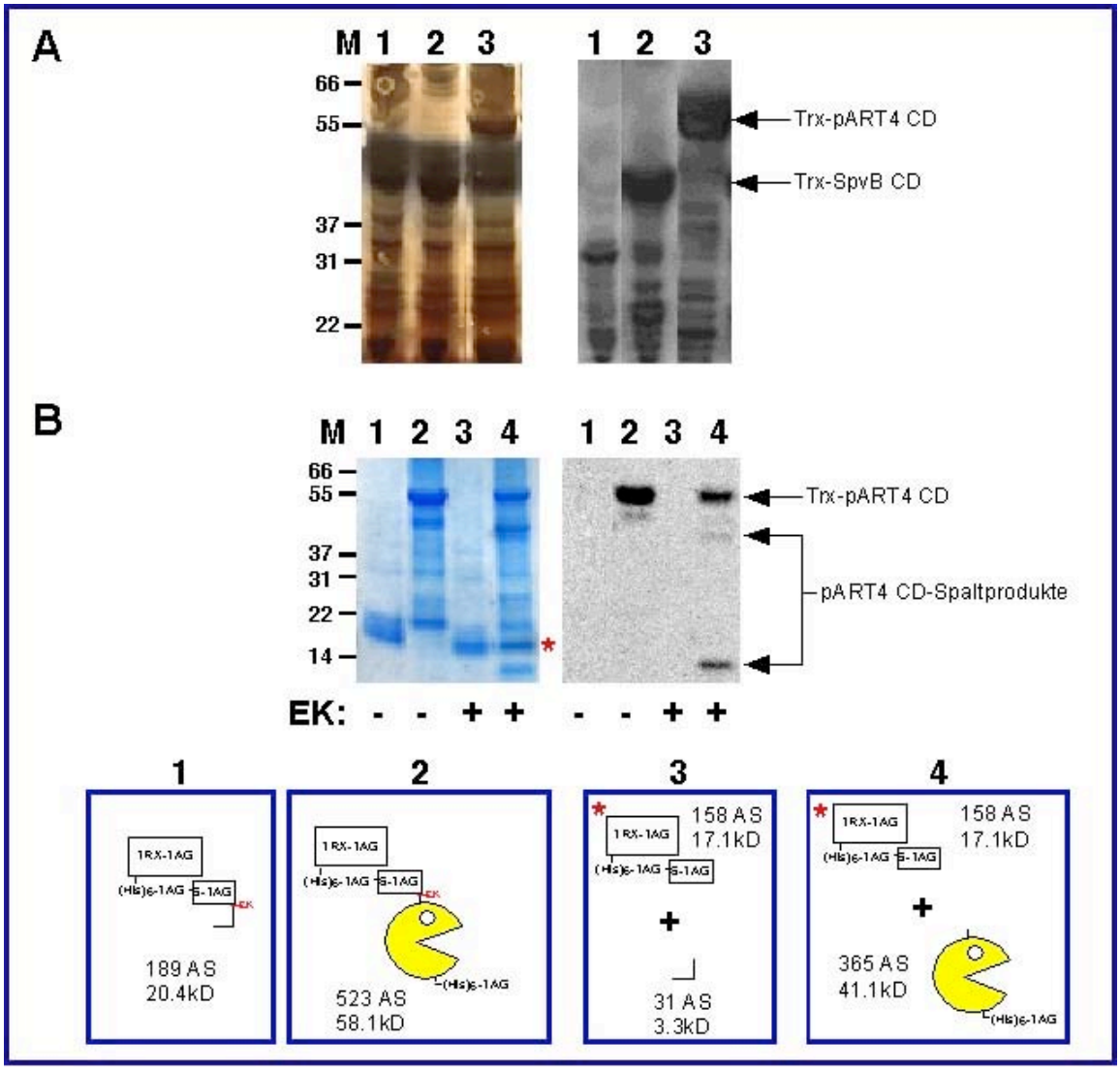


Abbildung 29: Detektion von rekombinantem Trx-pART4 über den „S-tag“ (A) und Lokalisation der ADP-ribosylierten Region im Trx-pART4-Fusionsprotein (B)

E.Coli-Lysate wurden wie unter 3.7.2 beschrieben aufbereitet, die Proteine mittels SDS-PAGE-Analyse größenfraktioniert und auf eine NC- sowie eine PVDF-Membran geblottet. Die NC-Membran wurde silbergefärbt. Die PVDF-Membran wurde für 30 Minuten bei RT mit 10%igem Ziegenserum geblockt und anschließend 60 Minuten rotierend bei 4°C mit 1:5000 verdünnter Protein S-Meerrettich-Peroxidase inkubiert. Nach mehrfachem Waschen mit TBS-Tween 0.5% wurde die Membran mit "ECL Western Blotting detection"-Reagenz (Amershan-Pharmacia) inkubiert, in Plastikfolie eingewickelt, ein Hyperflim ECL (Amershan-Pharmacia) für 5 Sekunden exponiert und sofort entwickelt. Die Abbildung zeigt die silbergefärbte NC-Membran sowie den entwickelten Hyperfilm ECL (**A**). Die Spuren enthalten Lysate, der mit folgenden Konstrukten transformierten *E.Coli*-Kulturen (*in silico* errechnetes MG in Klammern): 1 pET32a.Trx ohne Insert (20.4kD), 2 pET32a.Trx-SpvB-CD (42.6kD), 3 pET32a.Trx-hpART4-CD (58.1kD). Marker (M = Mark12, Novex) in kD.

Zum Nachweis der Lokalisation der ADP-Ribosylierung im Trx-pART4-Fusionsprotein wurden die rekombinanten Proteine aus *E.Coli*-Lysaten durch Absorption an immobilisierte Protein-S-Agarose aufgereinigt. Die aufgereinigten Proteine wurden 60 Minuten bei 37°C mit [^{32}P]-NAD inkubiert und anschließend gewaschen. Die Ansätze in den Spuren 1 und 2 wurden für die SDS-PAGE-Analyse vorbereitet, während die Ansätze in den Spuren 3 und 4 über Nacht mit Enterokinase (EK) inkubiert und erst anschließend der SDS-PAGE-Analyse zugeführt wurden. In der Abbildung sind das Coomassie-gefärbte Gel sowie die zugehörige Autoradiographie nach 48 Stunden Exposition eines Röntgenfilmes (BioMax MR, Kodak) bei -80°C gezeigt (**B**). Die verwendeten Lysate stammen aus mit folgenden Konstrukten transformierten *E.Coli*: **1/3** pET32a.Trx ohne Insert, **2/4** pET32a.Trx-pART4-CD. Marker (M, Mark12) in kD. Die verschiedenen rekombinanten Proteine vor und nach Enterokinase-Spaltung sind schematisiert dargestellt. Die nach EK-Spaltung generierte Trx/His-Tag/S-Tag-Bande ist mit einem roten Stern markiert. Die Anzahl der Aminosäuren sowie das *in silico* errechnete Molekulargewicht sind angegeben.

4.2.5 Aufschlüsselung der Genstruktur von hpART3 und hpART4

Um Informationen über die Exon-Intron-Struktur der neu entdeckten Familienmitglieder zu erhalten, wurden zwei verschiedene Methoden angewendet. Für die Aufschlüsselung der Genstruktur von *hpART3* wurde eine PCR mit verschiedenen, *hpART3*-spezifischen Primer-Paaren vergleichend mit einem genomischen DNA-Template und einem cDNA-Template als Matrize durchgeführt. Nach 34 Zyklen wurden Aliquots der PCR auf einem Ethidiumbromid-gefärbten Agarosegel analysiert und die Größe der Fragmente miteinander verglichen (**Abbildung 30**). Die cDNA von pART3 mit der Primer-Lage ist zur Veranschaulichung schematisiert dargestellt (**B**). Bei allen eingesetzten Primerkombinationen bestehen zwischen den aus der genomischen DNA und den aus der cDNA amplifizierten Fragmenten deutliche Größenunterschiede. So ist das bei Verwendung der Primer F65xF85 aus cDNA vervielfältigte Fragment ca. 600bp groß, während das mit genomischer DNA als Matrize amplifizierte Fragment ca. 2300bp umfasst. Diese Größenunterschiede zwischen den Fragmenten weisen auf den Einschub nicht kodierender Abschnitte in der genomischen DNA hin.

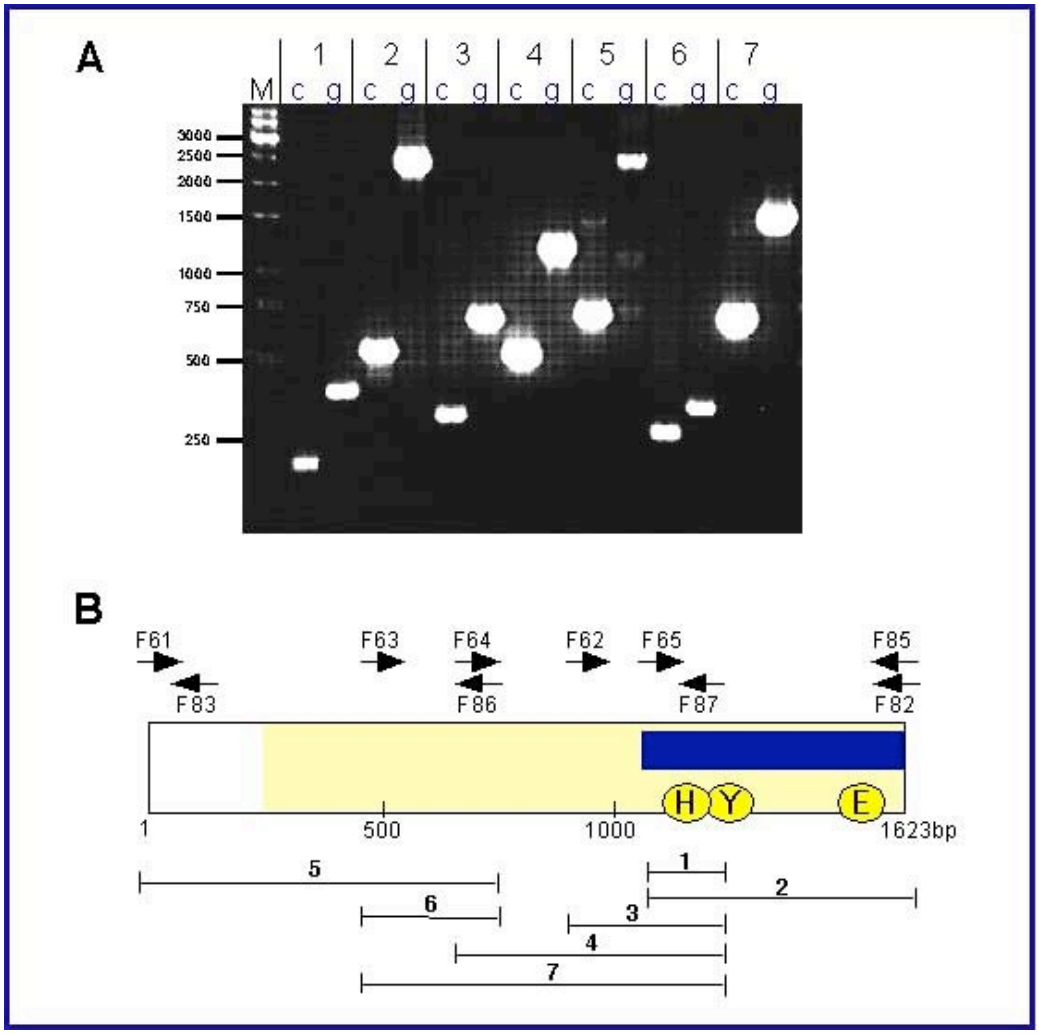


Abbildung 30: Vergleich von cDNA- und genomischen DNA-Fragmenten von hpART3

Um Informationen über die Exon-/Intron-Struktur von *hpART3* zu erhalten, wurden bestimmte Abschnitte eines cDNA-Klons sowie eines genomische DNA enthaltenen Klons amplifiziert und miteinander verglichen. Mit verschiedenen *hpART3*-spezifischen Primer-Kombinationen und der Expand-Polymerase (Boehringer-Ingelheim) mit „proof-reading" Funktion wurde eine PCR durchgeführt. Für die ersten Zyklen wurden fallende Annealing-Temperaturen, für die letzten 30 Zyklen eine Annealing-Temperatur von 55°C gewählt. Nach 34 Zyklen wurden Aliquots auf einem Ethidiumbromid-gefärbten Agarosegel analysiert und die Größe der Fragmente miteinander verglichen (**A**). Als cDNA-Template wurde ein pCR2.1.pART3-Plasmid (**c**), als genomisches DNA-Template (**g**) der PAC-Klon M06108 des RZPD (Deutsches Ressourcenzentrum für Genomforschung GmbH, www.rzpd.de) eingesetzt. Folgende Primer-Kombinationen wurden verwendet: **1** F65xF87, **2** F65xF85, **3** F62xF87, **4** F64xF87, **5** F61xF86, **6** F63xF86, **7** F63xF87. Marker (M): „GeneRuler 1kb-ladder" (MBI) in Basenpaaren. Die cDNA von hpART3, die Lage und Orientierung der verwendeten Primer (Pfeile) sowie die mit den Primer-Kombinationen aus der cDNA-Matrize amplifizierten Fragmente sind schematisiert abgebildet (**B**). Die pART1-verwandte Region von hpART3 ist gelb unterlegt. Die Region, die die katalytische Domäne kodiert, ist zusätzlich mit einem blauen Balken gekennzeichnet. Die Position der kodierenden Tripletts für die drei Aminosäuren, die das charakteristische HYE-Motiv bilden, ist eingezeichnet. Die genomischen Fragmente wurden anschließend aufgereinigt, in pCR2.1 kloniert und sequenziert.

Die mit genomischer DNA als Matrize gewonnenen Fragmente wurden in pCR2.1-Vektoren kloniert und anschließend sequenziert. Dadurch gelang es, die Exon-Intron-Grenzen und die Länge der Introns zu ermitteln. Um die 3`- und 5`-Enden des Genes zu vervollständigen sowie die laborexperimentell ermittelten Intronlokalisationen und -längen zu bestätigen, wurde als ergänzende Methode eine BLASTn-Datenbanksuche mit der vollständigen pART3-Konsensussequenz als Eingabe in der humanen genomischen Datenbank durchgeführt. Die Exon-Intron-Strukturen von *hpART1* und *hpART4* wurden analog durch *in silico* Analysen ermittelt und mit bereits publizierten Ergebnissen abgeglichen. **Abbildung 31** zeigt die genomische Struktur von *hpART1*, *hpART3* und *hpART4* im Vergleich. Das *hpART3*-Gen besteht aus 11, das *hpART4*-Gen aus 34 und das *hpART1*-Gen aus 23 Exons. Bei *hpART1* und *hpART3* liegt die Information für die katalytische Domäne am 3`-Ende des Genes, während sie bei dem deutlich größeren *hpART4*-Gen in der Mitte des Genes lokalisiert ist. Sowohl bei *hpART3* als auch bei *hpART4* besteht die Homologie mit *hpART1* 5`-wärts über die Region der katalytischen Domäne hinaus (vgl. 4.1.4 und 4.1.7). Während bei *hpART4* die Sequenzähnlichkeit 5`-wärts der Region, die die regulatorische Domäne kodiert abbricht, enthält das *hpART3*-Gen wie *hpART1* die Information für die WGR-Domäne. In die Genstrukturen ist die Lage der Regionen eingezeichnet, die für wichtige Motive wie z.B. Zink-Finger oder funktionelle Domänen in den Proteinen kodieren (vgl. Abb.20). Bei *hpART3* werden alle Bereiche, die funktionelle Domänen kodieren, durch Phase-0-Introns voneinander abgetrennt, während die Phase-1-Introns innerhalb dieser Regionen liegen. Auch bei den anderen Genen liegt zwischen diesen konservierten Bereichen zumeist mindestens ein Phase-0-Intron.

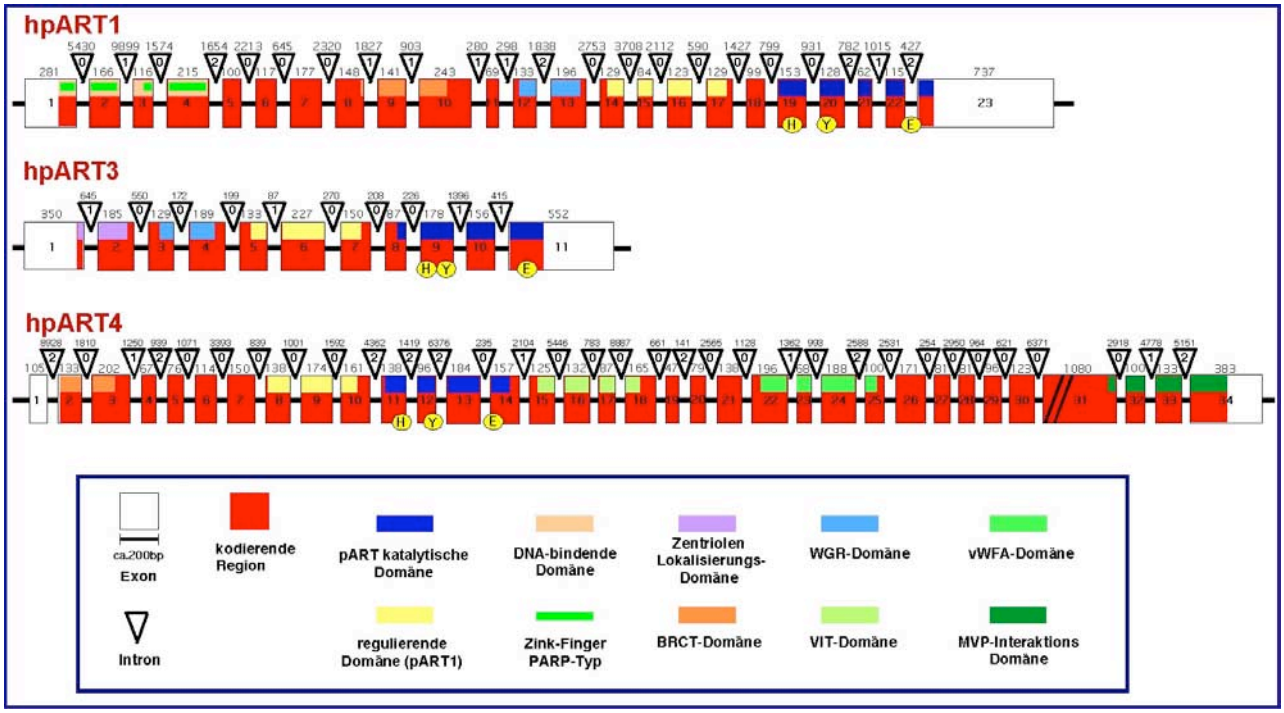


Abbildung 31: Schematische Darstellung der Genstrukturen von *hpART1*, *hpART3* und *hpART4*

Die Abbildung zeigt die schematische Darstellung der Genstrukturen von *hpART3* und *hpART4*. Zum Vergleich ist die Genstruktur der humanen *pART1* abgebildet. Die Exon-Intron-Struktur des *hpART3*-Genes wurde durch Sequenzierung von genomischen DNA-Fragmenten entschlüsselt und *in silico* mit der Nukleotid-Sequenz als "Input" durch BLASTn-Suche in der humanen genomischen Datenbank („human genome resources“) bestätigt. Für *hpART4* und *hpART1* wurde diese *in silico* Analyse mit der *hpART4*-Nukleotid-Sequenz bzw. der *hpART1*-Nukleotid-Sequenz als "Input" durchgeführt. In der schematischen Darstellung sind die Exons als Kästen und die Introns als Dreiecke abgebildet. Die Intronphase (0,1 oder 2) ist in den Dreiecken angegeben. Die Länge der Exons und Introns ist in Basenpaaren (bp) angegeben. Die kodierenden Regionen der Gene sind rot markiert. Die jeweiligen Passagen des Genes, die vorhergesagte Domänen kodieren, sind verschiedenfarbig gekennzeichnet. Die einzelnen Domänen sind unter 4.1.7 genauer beschrieben. Das große Exon 31 der *hpART4* wurde nicht im Maßstab wiedergegeben und deshalb mit schwarzen Schrägbalken kenntlich gemacht. Die Positionen der Tripletts, die das charakteristische HYE-Motiv kodieren, sind eingezeichnet.

4.2.6 Chromosomale Kartierung von *hpART3*

Bevor die Möglichkeit der *in silico*-Kartierung von Genen mit Hilfe von genomischen Sequenzen in der Datenbank bestand, konnte die chromosomale Lokalisation unter Verwendung von Mensch-/Nagetier-Hybridzellen laborexperimentell nachgewiesen werden. Hierbei macht man sich zunutze, dass bei der Verschmelzung von menschlichen Zellen mit Nagetierzellen bevorzugt die menschlichen Chromosomen aus den Hybridzellen verloren gehen. Deshalb kann man Hybridzellen herstellen, die nur ein einziges humanes Chromosom

oder nur den Teil eines humanen Chromosoms enthalten, und diese mit spezifischen Primern mittels PCR screenen.

Um das *hpART3*-Gen zu lokalisieren, screenten wir somatische Mensch-/Nagetier-Hybridzellen unter stringenten Bedingungen mit den *hpART3*-spezifischen Primern F65 und F87, die ein amplifiziertes Fragment von ca. 380bp erwarten ließen (vgl. Abb.30A, Spur 1g). Als Kontrolle wurden Zelllinien verwendet, die den gesamten Chromosomensatz von Mensch, Hamster und Maus enthalten. Die PCR-Produkte wurden gelelektrophoretisch aufgetrennt und mit Ethidiumbromid gefärbt (**Abbildung 32**). Die erwartete Bande ist lediglich in den Hybridzellen, die das humane Chromosom 3 enthalten, und in der humanen Kontrolle mit dem kompletten Chromosomensatz nachweisbar. Das *hpART3*-Gen ist somit auf Chromosom 3 lokalisiert. Homologe bei Maus und Hamster können mit der Primerkombination unter diesen Bedingungen nicht amplifiziert werden. Die chromosomale Lokalisierung von *hpART3* konnte später über *in silico* Analysen bestätigt werden (3p21.1-22.2). Das *pART4*-Gen wurde mittels *in silico* Analysen auf dem langen Arm von Chromosom 13 (13q11) lokalisiert.

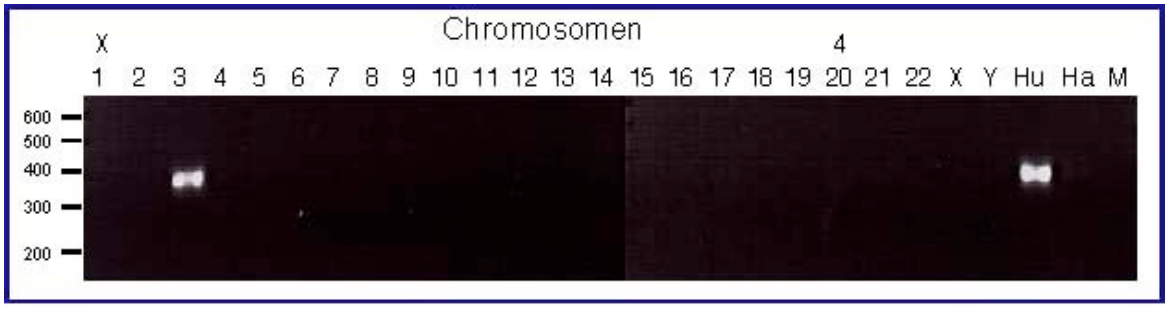


Abbildung 32: Chromosomale Kartierung von *hpART3*

Somatische Mensch/Nager-Hybrid-Zellen aus dem "NIGMS mapping panel #2" (the Corriel institute, Camden, NJ, USA) wurden mittels PCR mit den *hpART3*-spezifischen Primern F65 und F87 gescreent. Es wurden pro Ansatz 250ng genomische DNA eingesetzt. Zur Amplifikation wurden 30 PCR-Zyklen durchgeführt. Die Produkte wurden anschließend auf EthBr-gefärbten Agarosegelen aufgetrennt und analysiert. Die einzelnen bzw. zwei humanen Chromosomen, die die jeweilige Zelllinie enthält, sind oberhalb der Spuren angegeben. Zur Kontrolle sind Zelllinien gezeigt, die den gesamten Chromosomensatz von Mensch (Hu), Hamster (Ha) bzw. Maus (M) enthalten. Marker: „100bp ladder“ (MBI).

4.3 Molekulare Klonierung und rekombinante Expression der Salmonellen-ADP-Ribosyltransferase SpvB

4.3.1 Verknüpfung von Vertebraten-ARTs mit bakteriellen mono-ARTs durch Datenbankanalysen mit PSI-BLAST

Die in der Einleitung beschriebene, wichtige Rolle von mono-ARTs in der Zellphysiologie veranlasste uns nach Strategien zu suchen, um das Vorkommen dieser Enzyme in pro- und eukaryontischen Zellen zu erforschen. Zur Identifizierung von Verwandten der RT6-ekto-ADP-Ribosyl-Transferase in anderen Spezies benutzten wir, wie bei der Erforschung der pART-Familie, das positionsspezifische, iterative Suchprogramm PSI-BLAST (siehe 1.7.4). Wir verwendeten die Aminosäuresequenzen der katalytischen Domänen bekannter pro- und eukaryontischer mono-ARTs als „Input“ und durchsuchten mit PSI-BLAST die öffentlichen Datenbanken. Die Suche wurde in der „nr“-Sektion der GenBank im Dezember 1997 begonnen und in den folgenden Jahren regelmäßig wiederholt. Bei der Suche im Februar 2000 enthielt die „nr“-Datenbank die Sequenz eines neuen mono-ART-Familienmitgliedes, dem *Bacillus cereus*-Toxin VIP2, dessen 3D-Struktur erst kurz zuvor aufgeklärt worden war (Han *et al.* 1999). Die Ergebnisse unserer PSI-BLAST-Suche sind in **Abbildung 33** tabellarisch zusammengefasst.

Eingabe	*Hs*.ART1	*Bc*.VIP2	*St*.SpvB	*Cb*.C3	*Pa*.exoS	*Ec*.ALT	*Vc*.CT
initiale Verknüpfungen → traditionelle BLAST Suchen	*Mm*.ART2	*Cb*.C2	*Cp*.iota	*Sa*.EDIN	Pa.exoT	*Cb*.C2	*Ec*.LT
	Hs.ART3	*Cp*.iota	*Bc*.VIP2	*Cp*.iota		*Cp*.iota	
	Hs.ART4	*Cb*.C3		*Bc*.VIP2		*Cs*.Toxin A	
	Mm.ART5	*Sa*.EDIN					
	Gg.ART6	*Cs*.Toxin A	*Pi*.TcaC				
	Gg.ART7						
Iteration 2	*Pa*.exoS	*Pa*.exoS	*Cb*.C2	*Pa*.exoS	Hs.ART1	*Bc*.VIP2	*Bp*.PT
	Pa.exoT	*Pa*.exoT	*Cb*.C3	*St*.SpvB	Hs.ART2		
		St.SpvB			Hs.ART4		
					Mm.ART5		
					Gg.ART6		
					Gg.ART7		
					Vc.CT		
Iteration 3	*Cb*.C2	*Ec*.ALT	*Ec*.ALT	*Pa*.exoT	Hs.ART3	*Cb*.C3	*Bs*.MTX
			Sa.EDIN	*Vc*.CT	Cp.iota	*Sa*.EDIN	*Pr*.Pierisin
					Bc.VIP2	*St*.SpvB	
					Cb.C2		
					Ec.LT		
Iteration 4	*St*.SpvB	*Gg*.ART7	*Pa*.exoS	*Ec*.LT	Cb.C3	*Pa*.exoS	*Pa*.exoS
	Bc.VIP2		*Pa*.exoT	*Ec*.ALT	Ec.ALT		*Pa*.exoT
Iteration 5	*Cb*.C3	*Hs*.ART1	**konvergent**	*Hs*.ART4	Sa.EDIN	*Pa*.exoT	**konvergent**
	Cp.iota	*Mm*.ART2			Cs.Toxin A		
		Mm.ART5					
		Gg.ART6					
Iteration 6	*Sa*.EDIN	*Hs*.ART3		*Hs*.ART1	St.SpvB	**konvergent**	
				Hs.ART2			
				Mm.ART5			
				Gg.ART6			
				Gg.ART7			
Iteration 7	*Ec*.ALT	*Vc*.CT		**konvergent**	**konvergent**		
Iteration 8	*Cs*.ToxinA	*Ec*.LT					
Iteration 9	**konvergent**	**konvergent**					

Abbildung 33: PSI-BLAST-Datenbankrecherchen zur Identifizierung neuer mono-ARTs

Um bisher unbekannte Familienmitglieder der mono-ART-Familie aufzuspüren, wurden PSI-BLAST-Analysen in der "non redundant"-Datenbank (nr, Stand Februar 2000) mit sieben verschiedenen mono-ART-Sequenzen als Eingabe durchgeführt. Die initialen Verknüpfungen (gelb unterlegt) entsprechen den Proteinen, die eine hohe Sequenzidentität zum jeweils eingegebenen Protein aufweisen und auch mit dem traditionellen BLASTp-Programm ermittelt werden können. Für Gene, die bei mehr als einer Spezies gefunden werden konnten, ist lediglich ein Orthologes aufgeführt. Die Speziesgruppen sind in Anlehnung an Abb.17 farblich markiert: Säugetiere rot, andere Vertebraten orange, mikrobielle Organismen braun. Das identifizierte offene Leseraster des SpvB von *Salmonella typhimurium* ist blau fettgedruckt hervorgehoben. Die Sequenzen sind in der Datenbank unter folgender Zugangsnummer deponiert (Gen/Spezies/„accession number" des Proteins bzw. der Kristallstruktur): ART1/*Homo sapiens*/NP_004305, ART2.2/*Mus musculus*/CAA60948, ART2.2/*Rattus*

norvegicus/1OG3, ART3/*Homo sapiens*/NP_001170, ART4/*Homo sapiens*/NP_066549, ART5/*Mus musculus*/CAA69285, ART6/*Gallus gallus*/NP_990186, ART7/*Gallus gallus*/Q92080, Exoenzym S/*Pseudomonas aeruginosa*/CAA67834, Exoenzym T/*Pseudomonas aeruginosa*/AAB07232, ALT/*Escherichia coli* T6 Phage/CAA49518, Pertussis toxin/*Bordetella pertussis*/1BCP, SpvB/*Salmonella typhimurium*/CAA77649, Hitzelabiles Enterotoxin (LT)/*Escherichia coli*/1LTS, Cholera Toxin (CT)/*Vibrio cholerae*/1XTC, iota Toxin/*Clostridium perfringens*/1GIQ, Toxin A/*Clostridium spiriforme*/CAA66611, C2/*Clostridium botulinum*/BAA09942, Exoenzym C3/*Clostridium botulinum*/1G24, VIP2/*Bacillus cereus*/1QS2, "Epidermal differentiation inhibitor"(EDIN)/*Staphylococus aureus*/AAA26616, moskitozides Toxin (MTX)/*Bacillus sphaericus*/S27514, Pierisin/*Pieris rapae*/BAA84491, TcaC/*Photorhabdus luminescence*/AAC38625.

Durch PSI-BLAST gelang die Verknüpfung von Vertebraten-ARTs mit bakteriellen mono-ARTs. Die initial gefundenen Verwandten zeigen jeweils signifikante Sequenzidentitäten (>20% Sequenzhomologie), so dass sie auch von traditionellen BLAST-Programmen detektiert werden können (Altschul *et al.* 1990). Mit den Wiederholungen, bei denen das Programm jeweils die neu generierte Matrix zur Suche nutzt, wurden viele bekannte ARTs verschiedener Vertebraten und Bakterien ermittelt. Die Sequenzidentität dieser neuen potentiellen ART-Verwandten zu den Input-Sequenzen beträgt oft weniger als 10%; sie wären von traditionellen BLAST-Programmen folglich kaum noch aufzuspüren („deep in the twilight zone of traditional search programs“, (Rost 1999)). Während durch die PSI-BLAST-Analyse die bekannten Säugetier-mono-ARTs verlässlich mit den meisten der bekannten eubakteriellen mono-ARTs verknüpft wurden, fällt auf, dass keine signifikanten offenen Leseraster in den vollendeten Genomen der Fliege *Drosophila melanogaster* und niederer Eukaryonten wie der Hefe *Saccharomyces cerevisiae* und des Nematoden *Caenorhabditis elegans* gefunden wurden. Auch bei den Archaebacteriae konnte keine verwandte mono-ART ermittelt werden. Die entsprechenden mono-ART-Gene müssen in diesen Spezies entweder verloren gegangen sein oder haben sich in ihrer Sequenz so weit von den bekannten mono-ARTs entfernt, dass eine Verknüpfung durch PSI-BLAST nicht möglich war. Desweiteren wurden in der PSI-BLAST-Analyse mit den verschiedenen mono-ARTs als Input weder die Diphtamid-spezifischen Toxine Diphterietoxin und *Pseudomonas aeruginosa* Exotoxin A noch Mitglieder der pART-Familie ausfindig gemacht. Dies ist besonders bemerkenswert, da die Ähnlichkeiten in den bekannten Tertiärstrukturen einen gemeinsamen Vorfahren vermuten lassen.

Auch bei den gefundenen mono-ARTs ist die Sequenzidentität der verwandten Proteine vor allem in Regionen ausgeprägt, die den Sekundärstruktur-Einheiten entsprechen, die an der Bildung des aktiven Zentrums in bakteriellen Toxinen mit bekannter 3D-Struktur beteiligt sind. Das minimale Gerüst dieses katalytischen Zentrums ist bei den bekannten Kristallstrukturen ADP-ribosylierender bakterieller Toxine durch die charakteristische Anordnung von sechs bis sieben β-Strängen gekennzeichnet (siehe 1.4). Die Aminosäuresequenzen der durch PSI-BLAST ermittelten mono-ARTs wurden im Bereich dieser Sekundärstruktureinheiten mittels Sequenz-Alignment-Programmen verglichen (**Abbildung 34**). Zur Verdeutlichung ist die Topologie der Stränge in der katalytischen Region schematisiert abgebildet (**A,** vgl. auch Abb.5).

Die höchste Sequenzidentität zwischen den verschiedenen mono-ARTs besteht im Bereich der drei β-Stränge (β1, β2 und β5), die den Spalt des katalytischen Zentrums umgeben. In diesen am ausgeprägtesten konservierten Sekundärstruktureinheiten liegen drei Aminosäuren, für die in früheren Arbeiten gezeigt wurde, dass sie für die katalytische Aktivität der bakteriellen Toxine essentiell sind: Arginin (R) in β1, Serin (S) in β2 und Glutaminsäure (E) in β5 (Bazan and Koch-Nolte 1997, Domenighini and Rappuoli 1996). Mit Ausnahme von humanem ART3 (RLI) ist das charakteristische RSE-Motiv bei allen aufgeführten mono-ARTs konserviert. Während sich β3 und β6 bei CT, PT und LT analog zu den HYE-ARTs direkt nebeneinander befinden (**Abb.34 A1**), ist bei den anderen RSE-ARTs ein weiterer β-Strang (β7) zwischen β3 und β6 eingeschoben (**Abb.34 A2**, vergl. Abb.5). Diese aus den Kristallstrukturen gewonnene Erkenntnis erklärt, dass die Aminosäuresequenzen der beiden Untergruppen im Bereich des β6-Stranges deutlich voneinander abweichen (**Abb.34B**).

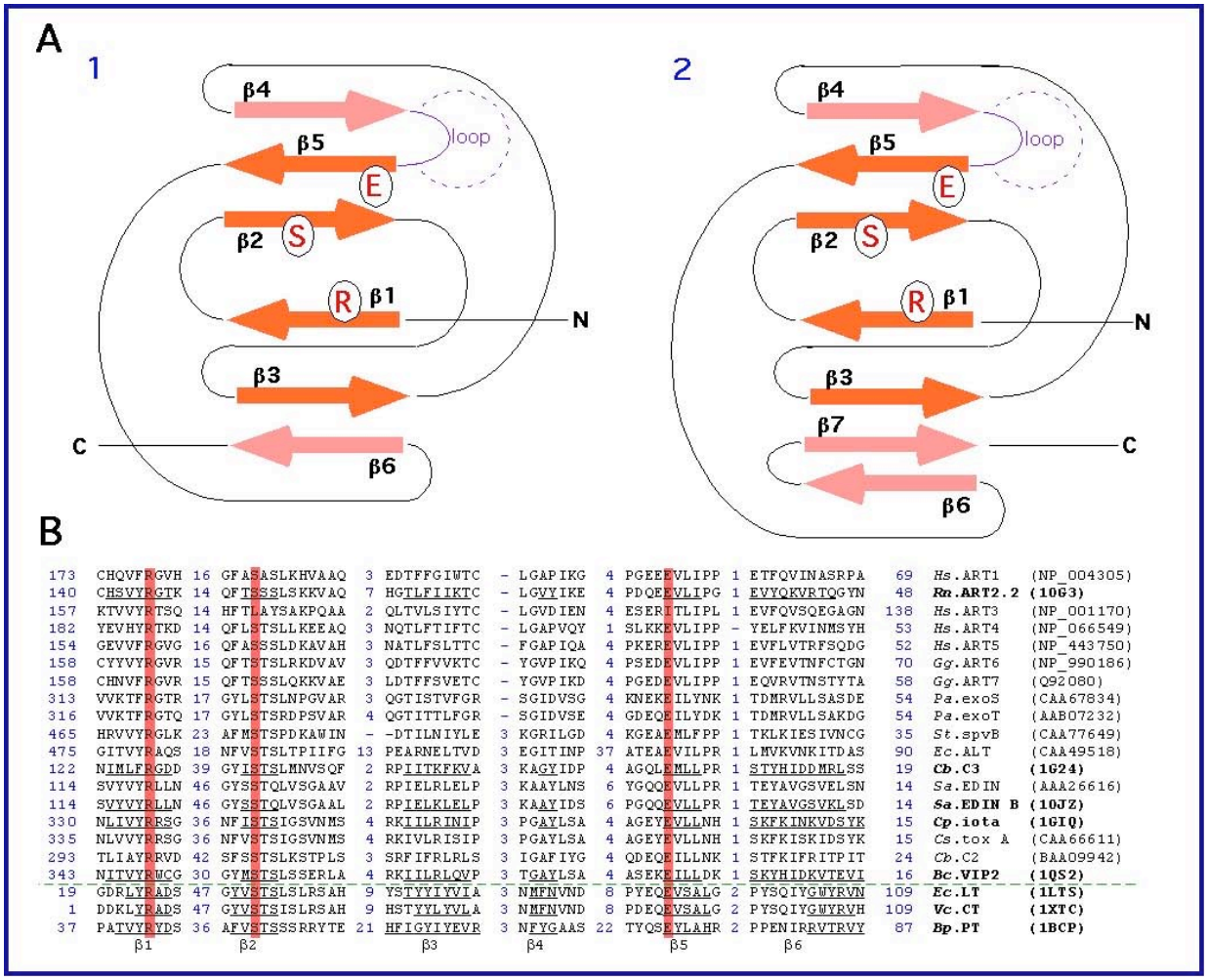


Abbildung 34: Schematische Darstellung des katalytischen Zentrums der mono-ARTs (A) und multiples Aminosäuresequenz-Alignment der katalytischen Region verschiedener mono-ARTs (B)

Das katalytische Zentrum bakterieller ARTs mit bekannter 3D-Struktur ist durch ein Gerüst von sechs (**A1** - CT, LT, PT) bis sieben (**A2**) charakteristisch angeordneten β-Strängen gekennzeichnet . Das auf der PSI-BLAST-Analyse basierende Aminosäuren-Alignment der verschiedenen mono-ARTs (**B**) zeigt die höchste Sequenzidentität im Bereich der drei β-Stränge (β1, β2 und β5), die den Spalt des aktiven Zentrums umgeben. Die drei Aminosäuren, die das charakteristische RSE-Motiv der mono-ARTs bilden (R in β1, S in β2 und E in β5) sind rot markiert. Die Zahlen (blau) geben die Anzahl der Aminosäuren-Reste am N- bzw. C-terminalen Ende der Proteine sowie die Länge der zwischen den Sekundärstruktureinheiten gelegenen Schleifen ("loops") an. Die Namen der mono-ARTs mit bekannten 3D-Srukturen sind fettgedruckt, die an der Bildung der β-Stränge beteiligten Aminosäuren-Reste sind unterstrichen. Am rechten Rand sind die Zugangsnummern angegeben, unter denen die Proteinsequenzen bzw. die Kristallstrukturen in der Datenbank deponiert sind (vgl. Abb.33). Die beiden Untergruppen der RSE-ARTs sind im Alignment durch eine gestrichelte, grüne Linie getrennt. Die Schleife („loop“) zwischen β4 und β5 ist violett markiert.

4.3.2 Identifikation von SpvB durch Datenbankanalysen mit PSI-BLAST

Die PSI-BLAST-Suche förderte ein offenes Leseraster bisher unbekannter Funktion aus der Spezies *Salmonella typhimurium* zu Tage, das eine signifikante Sequenzidentität zu den bekannten mono-ARTs aufweist. Dieses offene Leseraster wird durch das *spvB*-Gen („**s**almonella **p**lasmid **v**irulence“) kodiert. Das *spvB*-Gen ist auf einem Virulenzplasmid lokalisiert, über das viele klinisch relevante Serovarianten von *Salmonella enterica* verfügen. Das *spvB*-Gen bildet mit vier weiteren Genen den *spv*-Locus, dessen Nukleotidsequenz seit mehr als zehn Jahren bekannt ist und der eine essentielle Rolle bei der Virulenz von Salmonellen spielt (siehe 1.5.1).

Nach der Entdeckung dieses offenen Leserasters von *Salmonella typhimurium* führten wir eine PSI-BLAST-Suche mit der SpvB-Sequenz selbst als Input durch, um genauere Informationen über die funktionellen Anteile des SpvB-Proteins zu gewinnen. Neben dem Auffinden von bekannten bakteriellen ARTs kam es über den N-terminalen Anteil des SpvB bereits initial zur Verknüpfung mit dem TcaC-Protein von *Photorhabdus luminescence*. Die Analyse des Sequenz-Alignments der gefundenen Verwandten zeigte, dass SpvB aus zwei verschiedenen Subdomänen zusammengesetzt ist. In **Abbildung 35** ist ein Aminosäuresequenz-Alignment zwischen der SpvB-Sequenz und den beiden nächsten Verwandten der Subdomänen in der Datenbank dargestellt.

Die mono-ART-verwandte Region ist im C-terminalen Anteil von SpvB enthalten (Aminosäuren-Reste 374-591). Die nächste Verwandte dieser Domäne in der Datenbank ist die katalytische Domäne von VIP2, einem von *Bacillus cereus* produzierten insektiziden Toxin (Han *et al.* 1999). Die Sequenzidentität beträgt insgesamt 19%. Die katalytische Domäne wird durch einen auffälligen, starren Polyprolin-Abschnitt vom N-terminalen Anteil des SpvB getrennt. Der aus sieben Prolin-Resten bestehende Abschnitt wird von zwei Methioninen flankiert und wäre somit ein typischer α-Helix-Brecher. Die N-terminale Domäne des SpvB (Aminosäuren-Reste 1-366) besitzt lediglich zu einem anderen bekannten Protein in der öffentlichen Datenbank Sequenzähnlichkeit: Zu TcaC, das Bestandteil eines großen insektiziden Toxin-Komplexes ist, der von *Photorhabdus luminescence* produziert wird, besteht eine Sequenzidentität von 47%. Das Bakterium *P.luminescence* lebt im Darm von Würmern, die Insekten befallen, und wird nach Invasion des Wirts aus dem Darm des

Wurmes freigesetzt, so dass es sich vermehren kann. Über die Sekretion von insektiziden Proteinkomplexen tötet das Bakterium das Insekt (Bowen *et al.* 1998).

Die Sequenzidentität der katalytischen Domänen von SpvB und VIP2 ist besonders ausgeprägt im Bereich der Sekundärstruktur-Einheiten, die in der aufgeklärten Kristallstruktur von VIP2 den katalytischen Spalt umgeben. Das R-S-E-Motiv der mono-ARTs ist auch beim SpvB konserviert (**R**471, **S**501 und **E**538). Überdies weist SpvB das für Arginin-spezifische mono-ARTs typische, aus zwei kurz aufeinanderfolgenden Glutaminsäure-Resten bestehende EXE-Motiv (**E**536-A537-**E**538) auf.

```
SpvB:   1  MLILNGFSSATLALITPPFLPKGG-------KALSQSGPDGLASITLPLPISAERGFAPALALHYSSGGGNGPF
                 ++++    IT   LPKGG       +AL+ +GPDG+AS++LPLP+S  RG AP L+L YS+  GNGPF
TcaC:   1        MQDSPEVSITTLSLPKGGGAINGMGEALNAAGPDGMASLSLPLPLSTGRGTAPGLSLIYSNSAGNGPF

SpvB:  68 GVGWSCATMSIARSTSHGVPQYNDSDEFLGPDGEVLVQTLSTGDAPN-PVTCFAYGDVSFPQSYTVTRYQPRTES
          G+GW C  MSI+R T HG+PQY + D FL P GEV+   L+    P+          V+ P SYTVTRYQ R
TcaC:  69 GIGWQCGVMSISRRTQHGIPQYGNDDTFLSPQGEVMNIALNDQGQPDIRQDVKTLQGVTLPISYTVTRYQARQIL

SpvB: 142 SFYRLEYW--VGNSNGDDFWLLHDSNGILHLLGKTAAARLSDPQAASHTAQWLVEESVTPAGEHIYYSYLAENGD
           F ++EYW       G  FWL+   +G LH+LGKTA A L++PQ     AQWL+EE+VTPAGEH+ Y Y AE+
TcaC: 144 DFSKIEYWQPASGQEGRAFWLISSPDGQLHILGKTAQACLANPQNDQQIAQWLLEETVTPAGEHVSYQYRAEDEA

SpvB: 215 NVDLNGNEAGRDRSAMRYLSKVQYGNATPAADLYLWTSATPA-VQWLFTLVFDYGERGVDPQVPPAFTAQNS-WL
          + D N   A  + +A RYL +V YGN  P A L++  +A PA  +WLF LVFD+GER       P + A  + W
TcaC: 219 HCDDNEKTAHPNVTAQRYLVQVNYGNIKPQASLFVLDNAPPAPEEWLFHLVFDHGERDTSLHTVPTWDAGTAQWS

SpvB: 288 ARQDPFSLYNYGFEIRLHRLCRQVLMFHHFPDELGEADT-----LVSRLLLEYDENPILTQLCAARTLAYEGDGY
           R D FS Y YGFE+R  RLC+QVLMFH      GEA T     LV RL+LEYD+N  +T L   R L++E DG
TcaC: 294 VRPDIFSRYEYGFEVRTRRLCQQVLMFHRTALMAGEASTNDAPELVGRLILEYDKNASVTTLITIRQLSHESDGS

SpvB: 358 RRAPVNNMMPPPPPPP
                  +
TcaC: 369 PVTQPPLEL

SpvB: 374 MMGGNSSRPKSKWAIVEESKQIQALRYYSAQGYSVINKYLRGDDYPETQAKETLLSRDYLSTNEPSDEEFKNAMS
                 +  ++WA      Q +AL  Y+ Q Y  IN YLR                  S NE  D + KN
VIP2: 275 EAHSWGMKNYEEWAKDLTDSQREALDGYARQDYKEINNYLRNQGG---------------SGNEKLDAQIKN---

SpvB: 449 VYINDIAEGLSSLPETDHRVVYRGLKLDK---------PALSDVLKEYTTIGNIIIDKAFMSTSPDKAWI-----
               I++ L   P  ++  VYR   + +         P+L D  +++  +  I  DK +MSTS     +
VIP2: 335 -----ISDALGKKPIPENITVYRWCGMPEFGYQISDPLPSLKDFEEQF--LNTIKEDKGYMSTSLSSERLAAFGS
                                β1                                    β2

SpvB: 510 NDTILNIYLEKGHKGRILGDVAHFKGEAEMLFPPNTKLKIESIVNCGSQDFASQLSKLRLSDDATADTNRIKRII
            IL + + KG  G  L  +  F  E E+L   ++K  I+ + +   + + + + + +L  +
VIP2: 400 RKIILRLQVPKGSTGAYLSAIGGFASEKEILLDKDSKYHIDKVTEVIIKGVKRYVVDATLLTN
                β3       β4          β5        β6          β7
```

Abbildung 35: Aminosäuresequenz-Alignment von SpvB mit seinen nächsten Verwandten aus der Datenbank

Durch PSI-BLAST-Suche in der "non redundant“-Datenbank (nr) mit der AS-Sequenz von SpvB als Eingabe wurde ermittelt, dass SpvB aus zwei verschiedenen Subdomänen besteht: Die N-terminale Domäne des SpvB (Aminosäurereste 1-366) zeigt eine signifikante Sequenzhomologie (47% Sequenzidentität) zu TcaC, das eine der Komponenten des insektiziden Toxin-Komplexes des Bakteriums *Photorhabdus luminescence* bildet. Diese TcaC-verwandte Domäne wird durch einen starren Polyprolin-Abschnitt (Aminosäurereste 367-373, blau markiert) von der C-Domäne (Aminosäuren-Reste 374-592) des SpvB getrennt, die eine Sequenzidentität von 19% mit der katalytischen Untereinheit des *B.cereus*-Toxins VIP2 (Aminosäurenreste 275-461 von VIP2) aufweist. Die sieben charakteristischen β-Stränge des VIP2 (vgl. Abb.34, Abb.5) sind gelb unterlegt und mit Pfeilen gekennzeichnet. Weitere Sekundärstruktureinheiten des VIP2 (Helices und 1 weiterer β-Strang) sind unterstrichen. Das charakteristische RSE-Aminosäuren-Motiv der mono-ARTs ist rot markiert. Die aufgeführten Sequenzen sind unter den folgenden Zugangsnummern ("accession numbers") in der öffentlichen Datenbank deponiert: SpvB/*Salmonella typhimurium*/Z11557, VIP2/*Bacillus cereus*/1QS2, TcaC/*Photorhabdus luminescence*/AF046867.

Die Subdomänenstruktur von SpvB, TcaC und VIP2 ist in **Abbildung 36** schematisch dargestellt. Die Sequenzähnlichkeit von SpvB zu TcaC und VIP2 bricht vor bzw. nach dem Polyprolin-Abschnitt ab. Die TcaC-Sequenz setzt sich zum C-terminalen Ende hin noch in einen Anteil fort, der 1100 Aminosäurereste umfasst. Für diesen C-terminalen Abschnitt von TcaC konnte keine Sequenzidentität mit einem in der Datenbank eingetragenen Protein nachgewiesen werden.

VIP2 gehört zur Klasse der sog. binären Toxine. An das vordere Ende der katalytischen Domäne ist eine weitere, aus 230 Aminosäuren bestehende Domäne gekoppelt. Diese N-terminale Domäne von VIP2 ist zwar strukturell mit der ADP-Ribosyl-Transferase-Domäne verwandt, zeigt aber selbst keine katalytische Aktivität. Stattdessen ist diese Domäne für die Bindung von VIP2 an VIP1 verantwortlich, damit das VIP1-VIP2-Holotoxin entstehen kann, wobei beide Polypeptid-Komponenten eine separate Funktion besitzen. Das Membran-bindende VIP1 bereitet dabei den Weg für die Einschleusung von VIP2 in das Zytoplasma (Han *et al.* 1999). Die N-terminale Domäne von VIP2 zeigt 22% Sequenzidentität mit der C-terminalen, katalytischen Domäne von VIP2, aber nur weniger als 5% Sequenzidentität mit der katalytischen Domäne von SpvB.

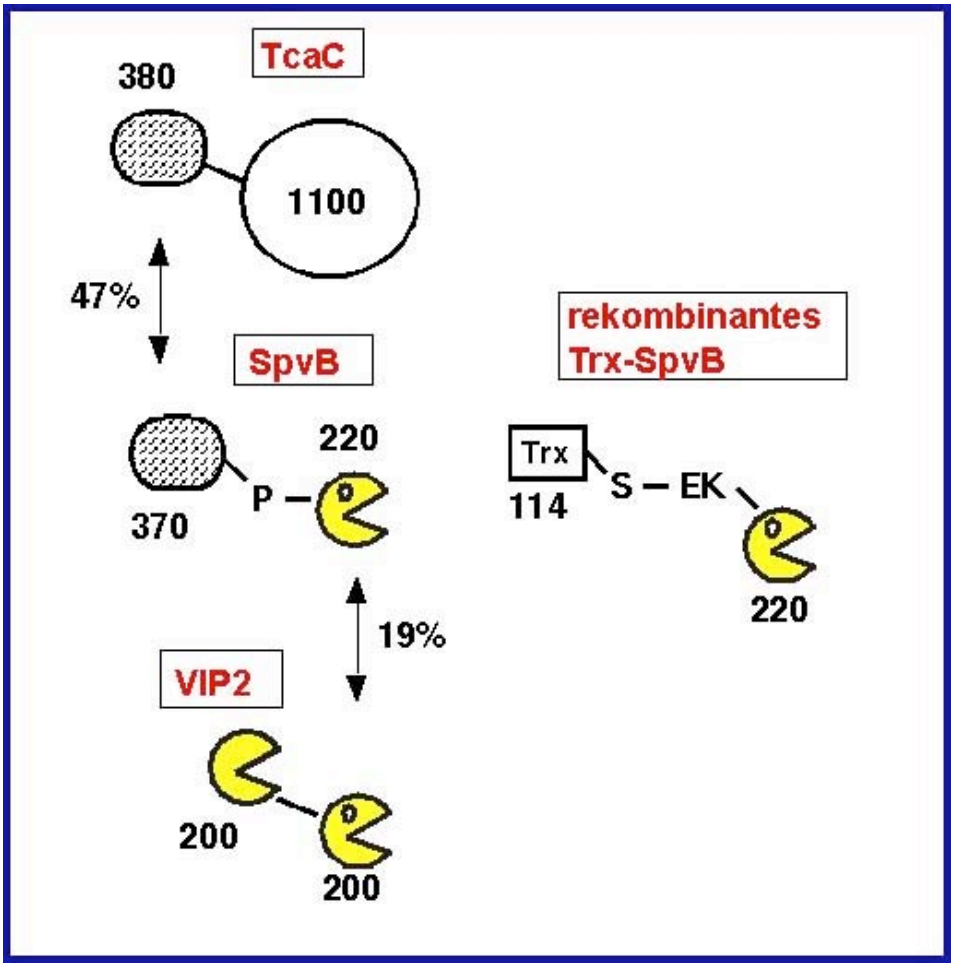


Abbildung 36: Schematische Darstellung der Domänenstrukturen von SpvB und verwandten Proteinen
In der Abbildung sind die Subdomänen von SpvB, TcaC und VIP2 dargestellt. Die Zahlen geben die Anzahl der Aminosäurereste an, die die einzelnen Subdomänen bilden, und verdeutlichen so die Größe der Subdomänen. Die den katalytischen Domänen der mono-ARTs verwandten Subdomänen sind durch ein Pacman-Symbol dargestellt. Die Aminosäuresequenzidentitäten der verwandten Subdomänen sind in Prozent angegeben. SpvB setzt sich aus zwei Subdomänen zusammen: Die C-terminale, putative katalytische Domäne zeigt eine Sequenzidentität von 19% zur katalytischen Domäne von VIP2, die N-terminale Region eine Sequenzidentität von 47% zur N-terminalen Domäne von TcaC. An der rechten Seite ist zum Vergleich das rekombinante Trx-SpvB-Fusionsprotein schematisch dargestellt. Die TcaC-verwandte Domäne ist in diesem Konstrukt durch Thioredoxin, einen S-tag (S) zur Affinitätsaufreinigung sowie ein fünf Aminosäuren umfassendes Erkennungssignal (DDDDK) für Enterokinase (EK) ersetzt, wobei als Fusionspunkt der erste Methionin-Rest nach dem Polyprolin-Abschnitt (P) des SpvB ausgewählt wurde.

4.3.3 Klonierung, Expression und Aufreinigung von rekombinanten SpvB-Konstrukten in *E.Coli*

Um die nach der Sequenz vorhergesagte ADP-Ribosyl-Transferase-Aktivität von SpvB experimentell zu untersuchen, entwickelten wir Strategien um rekombinantes SpvB in löslicher Form zu produzieren. Da vorherige Versuche anderer Arbeitsgruppen, das vollständige SpvB-Protein in *E.Coli* zu exprimieren, nur zur Produktion von unlöslichem

SpvB geführt hatten (Taira *et al.* 1991a), entschlossen wir uns, Expressionskonstrukte mit der isolierten katalytischen Domäne von SpvB herzustellen.

4.3.3.1 Klonierung von drei verschiedenen SpvB-Konstrukten

Aus einer SpvB-Plasmide enthaltenden Präparation (freundlicherweise zur Verfügung gestellt von M.Rhen, Karolinska Institut, Stockholm, Schweden) wurde durch PCR mit zwei verschiedenen Primerpaaren (Sp1 x Sp4 / Sp2 x Sp4) und der Expand-Polymerase (Boehringer) die katalytische Domäne von SpvB amplifiziert. Die genauen Primerpositionen sind zusammen mit der SpvB-Sequenz im Anhang B dargestellt. Im Gegensatz zu der bei der Klonierung der pARTs verfolgten Klonierungsstrategie wurden die amplifizierten Produkte nicht zunächst in den Vektor pCR2.1 kloniert, sondern mit Restriktionsenzymen verdaut und direkt in die Expressionsvektoren ligiert (siehe Abb.8). Auch für die Expression des SpvB benutzten wir das bereits ausführlich unter 1.8 und im zweiten Abschnitt des Ergebnisteils vorgestellte pET-Expressionssystem von Novagen. Wir entschieden uns analog zu den pART-Konstrukten für drei verschiedene Expressionsvarianten, die schematisch am Beispiel des humanen pART3 in **Abbildung 24** dargestellt wurden:

1) Expression der isolierten katalytischen Domäne von SpvB im Zytoplasma
2) Fusion der katalytischen Domäne von SpvB an einen „pel B-Leader“, so dass das Fusionsprotein nach der Expression in das Periplasma ausgeschleust wird
3) Fusion der katalytischen Domäne von SpvB mit Thioredoxin (Trx), das die Löslichkeit des Fusionsproteins im Zytoplasma von *E.Coli* erhöht

Als Fusionspunkt der katalytischen SpvB-Domäne wählten wir den ersten Methionin-Rest nach dem Polyprolin-Abschnitt. Da uns noch keine Antikörper gegen natives SpvB zur Verfügung standen, wurde für das Trx-SpvB-Expressionskonstrukt analog zu den Trx-pART-Varianten der pET32a-Vektor ausgewählt, der es ermöglicht, in das Fusionsprotein verschiedene Epitop-„tags“ einzubauen. Dies eröffnet die Gelegenheit, das rekombinante SpvB spezifisch zu detektieren sowie durch eine Affinitätsaufreinigung anzureichern.

4.3.3.2 Rekombinante Expression und Aufreinigung des SpvB

Die SpvB-Konstrukte wurden mittels Hitzeschock in BL21(DE3)- und BL21(DE3)pLysS-*E.Coli*-Stämme transformiert. Analog zu den rekombinanten pARTs wurden die Kulturen

angezüchtet, die Expression des Genproduktes durch IPTG induziert, die Gesamtzell-Lysate hergestellt, die Proteine mittels SDS-PAGE aufgetrennt und auf eine NC-Membran geblottet. **Abbildung 37** zeigt die Silberfärbungen der NC-Membranen. Die rekombinanten SpvB-Varianten wurden zwar exprimiert, jedoch ist durch den Vergleich der Gesamtproteine (**A**) und der löslichen Proteine (**B**) deutlich sichtbar, dass ein umfangreicher Anteil des rekombinanten Proteins in unlöslicher Form vorliegt. Dies zeigt sich auch bei dem am effizientesten produzierten Trx-SpvB-Fusionsprotein.

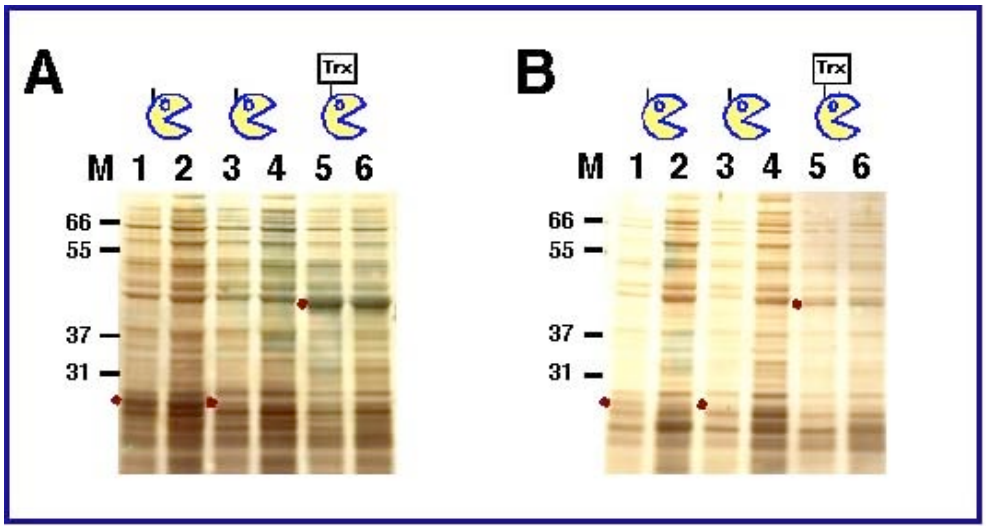


Abbildung 37: Expression von SpvB-Konstrukten in *E.Coli*

Die Expression und die SDS-PAGE-Analyse der rekombinanten Proteine erfolgte wie in Abb.25 für die Konstrukte der hpART3 beschrieben. Die Proteine wurden anschließend auf eine Nitrocellulose-Membran geblottet und silbergefärbt. Die Abbildung zeigt die silbergefärbten Nitrocellulose-Membranen mit den aufgetrennten Gesamtproteinen (**A**) bzw. löslichen Proteinen (**B**). Oberhalb der Membranen sind die rekombinanten SpvB-Varianten, die in den jeweiligen Lysaten enthalten sind, schematisch abgebildet. Die roten Punkte kennzeichnen die Banden, die durch die rekombinanten Proteine hervorgerufen werden. Die Spuren enthalten die Lysate der mit folgenden Konstrukten transformierten *E.Coli*-Stämme (in Klammern sind die *in silico* vorhergesagten Molekulargewichte der rekombinanten Proteine angegeben): **1/2** pET26b.SpvB (25.5kD), **3/4** pET26b.SpvB.PPl (25.5kD), **5/6** pET32a.Trx-SpvB (42.6kD). In den Spuren 1, 3 und 5 wurden die Lysate der BL21(DE3)-Zellen, in 2, 4 und 6 die der BL21(DE3)pLysS-Zellen aufgetragen. Als Marker (M) wurde MARK 12 verwendet (Novex, Angabe in kD).

Wie das Trx-pART4 konnte auch das Trx-SpvB-Fusionsprotein über die Bindung des „S-tag“ an Protein-S-gekoppelte Meerrettich-Peroxidase in einer Chemilumineszenz-Reaktion detektiert werden. Dies wurde bereits in **Abbildung 29A** (Spur 2) gezeigt. Zudem wurde der „S-tag“ genutzt, um das rekombinante Protein durch eine “single step“-Affinitätschromatographie an Protein-S-Agarose zu binden und anschließend aufgereinigt und angereichert zu eluieren (**Abbildung 39A**). Während das SpvB im Gesamtzell-Lysat lediglich als eine prominente Bande unter vielen zu erkennen ist (Spur 1), zeigt sich in Spur 2 deutlich eine Einzelbande, die dem aufgereinigten Trx-SpvB-Fusionsprotein mit einer Größe von 42kD entspricht.

4.3.4 Enzymassays mit verschiedenen rekombinanten SpvB-Konstrukten

4.3.4.1 Nachweis der ADP-Ribosyltransferase-Aktivität des rekombinanten SpvB

Um zu ermitteln, ob die verschiedenen rekombinanten SpvB-Konstrukte ADP-Ribosyl-Transferase-Aktivität besitzen, wurden die löslichen Proteine aus *E.Coli*-Lysaten in Gegenwart von 2mM ADP-Ribose mit radioaktiv markiertem [^{32}P]-NAD^{+} 60 Minuten bei 37°C inkubiert. Nach Präzipitation wurden die Proteine mittels SDS-PAGE und Autoradiographie analysiert (**Abbildung 38A**). Während in den Lysaten, die die isolierte Trx-Domäne enthalten (Spur 7/8), keine Markierung mit [^{32}P]-ADP-Ribose sichtbar wird, ist in den rekombinantes Trx-SpvB-Fusionsprotein enthaltenden Lysaten autoradiographisch eine einzelne, ca. 42kD große markierte Bande nachzuweisen (Spur 5/6). Da das *in silico* vorhergesagte Molekulargewicht des Trx-SpvB 42.6kD beträgt, könnte es sich bei der markierten Proteinbande um ADP-ribosyliertes Trx-SpvB-Fusionsprotein handeln. In den Spuren mit der isolierten katalytischen Domäne des SpvB ohne Fusionspartner (pET26b-Konstrukte, Spur 1-4) zeigt sich, wie bei der Trx-Kontrolle, keine Aktivität, so dass diese Konstrukte entweder keine katalytische Aktivität besitzen oder die Lysate kein potentielles Zielprotein enthalten. Bzgl. dieses Experimentes ist kritisch anzumerken, dass aufgrund der sehr unterschiedlichen Mengen an geladenen Proteinen in den einzelnen Spuren die Vergleichbarkeit eingeschränkt ist. Allerdings gehören die Spuren, in denen die Aktivität nachweisbar ist, zu den am schlechtesten beladenen Spuren, so dass die Aussagekraft des Experimentes nicht beeinträchtigt ist.

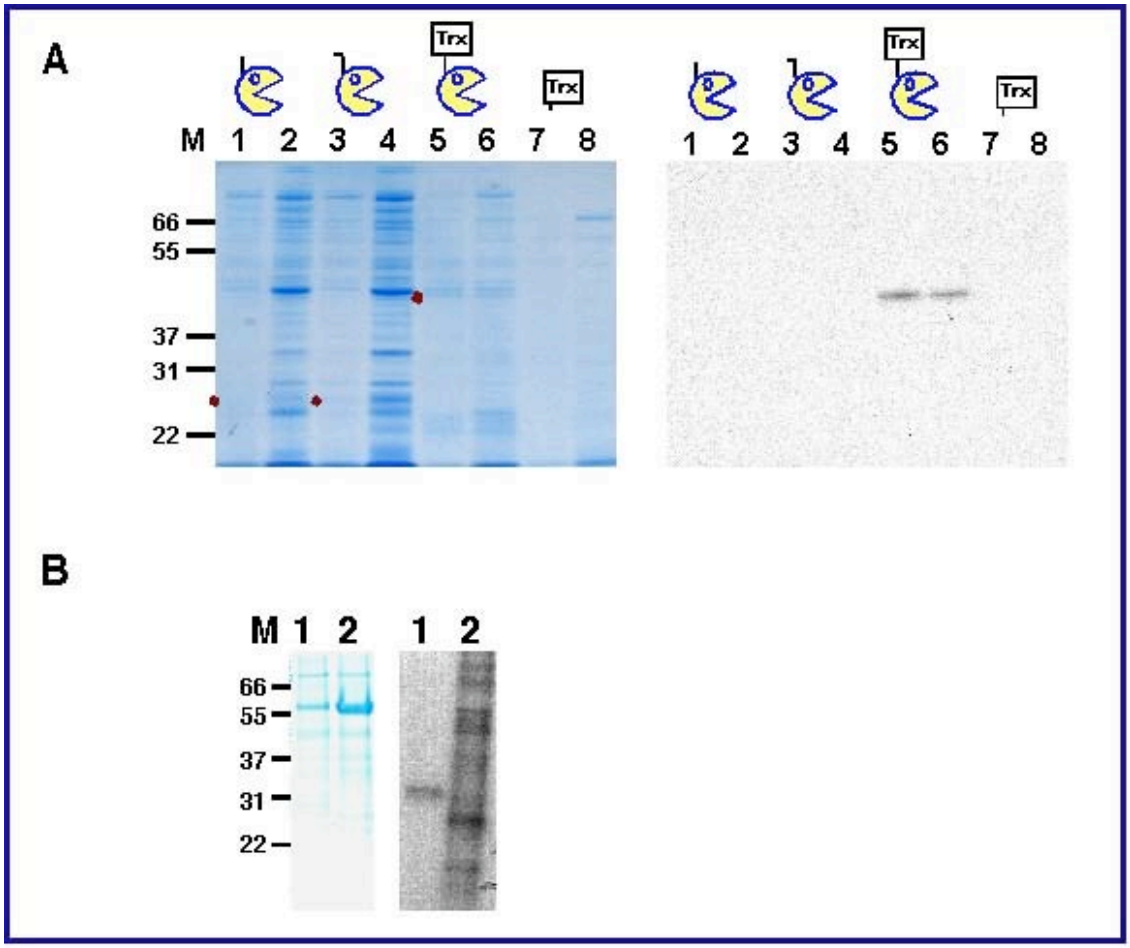


Abbildung 38: Nachweis der ADP-Ribosyl-Transferase-Aktivität in Trx-SpvB enthaltenden *E.Coli*-Lysaten

Die löslichen Proteine aus *E.Coli*-Lysaten wurden in Enzympuffer 60 Minuten bei 37°C mit [^{32}P]-NAD^{+} inkubiert. Nach Präzipitation mit StrataClean Resin (Stratagene) wurden die Proteine mittels SDS-PAGE aufgetrennt und autoradiographisch analysiert. Das Coomassie-gefärbte Gel sowie die dazugehörige Autoradiographie (Film: Biomax MR, Kodak / Exposition: 9d bei -80°C) sind abgebildet (**A**). Die Spuren enthalten die mit folgenden Konstrukten transformierten Lysate (*in silico* errechnetes MG in Klammern): **1/2** pET26b.SpvB-CD (25.5kD), **3/4** pET26b.SpvB-CD.PPl (25.5kD), **5/6** pET32a.Trx-SpvB-CD (42.6kD), **7/8** pET32a.Trx ohne Insert (20.4kD). In den Spuren 1, 3, 5 und 7 wurden die Lysate der BL21(DE3)-Zellen, in 2, 4, 6 und 8 die der BL21(DE3)pLysS-Zellen aufgetragen. Die verschiedenen Konstrukte sind oberhalb der Spuren schematisch dargestellt. Die roten Punkte markieren die Höhe, in der sich die rekombinanten Proteine nach der Größenfraktionierung befinden (vgl. Abb.37). Zum Vergleich wurden Periplasma-Lysate von NM522/pASK.rART2.2-FH (**B**, Spur 1) sowie NM522/pASK.mART2.2-FH (**B**, Spur 2) 60 Minuten bei 37°C mit [^{32}P]-NAD^{+} inkubiert und mittels SDS-PAGE und Autoradiographie (Film: Biomax MR, Kodak / Exposition: 5d bei -80°C) analysiert. Die pASK.ART2.2-Konstrukte wurden in der Arbeitsgruppe kloniert (Bredehorst 2002). Das *in silico* errechnete MG beträgt 28.1kD für Ratten-ART2.2 und 28.4kD für Maus-ART2.2. Marker (M, Mark12) in kD.

Zum Vergleich wurde das Experiment mit Periplasma-Lysaten von *E.Coli* durchgeführt, die mit Expressionskonstrukten der mono-ARTs Maus-ART2.2 bzw. Ratten-ART2.2 transformiert worden waren (**Abbildung 38B**). In den Lysaten mit Maus-ART2.2 sind durch den Einbau radioaktiver ADP-Ribose in Proteine verschiedener Größe multiple Banden nachweisbar (Spur 2), während in den Lysaten mit Ratten-ART2.2 lediglich eine einzelne, von der Größe dem rekombinanten Ratten-ART2.2 (*in silico* errechnetes Molekulargewicht: 28.1kD) entsprechende Proteinbande markiert wird.
Die Ergebnisse weisen darauf hin, dass Trx-SpvB eine Auto-ADP-Ribosylierungs-Aktivität besitzt, wie sie für das Ratten-ART2.2 bereits beschrieben wurde (Haag *et al.* 1995). Weiterhin wird deutlich, dass rekombinantes Trx-SpvB im Gegensatz zu Maus-ART2.2 (Karsten *et al.* 1997, Koch-Nolte *et al.* 1996b) nicht die Fähigkeit besitzt, multiple *E.Coli*-Proteine zu modifizieren.

Die Schlussfolgerung, dass die radioaktive Markierung der 42kD Bande in den Lysaten der mit pET32a.Trx-SpvB transformierten *E.Coli* enzymatisch hervorgerufen wird, lässt sich durch folgende, in dieser Arbeit nicht gezeigte Beobachtungen untermauern: 1) Wenn die Lysate vor der Zugabe der Markierungsreagentien 5 Minuten bei 80°C inkubiert werden, findet keinerlei Markierung statt. 2) Durch Zugabe von 5mM Novobiocin, einem bekannten Inhibitor der mono-ADP-Ribosylierung (Banasik *et al.* 1992), vor der Inkubation wird die radioaktive Markierung der 42kD-Proteinbande ebenfalls verhindert.

Mit den Überständen der Reaktionsansätze führten wir zudem eine Dünnschichtchromatographie durch, um die Lysate auf NAD-Glykohydrolase-Aktivität zu testen. Im Gegensatz zu Ratten-ART2.2, aber vergleichbar den Beobachtungen bei Maus-ART2.2 zeigen die SpvB-Konstrukte keine detektierbare NAD-Glykohydrolase-Aktivität (Ergebnisse hier nicht abgebildet).

Mit den bisher vorgestellten Ergebnissen konnte gezeigt werden, dass in Anwesenheit von rekombinantem Trx-SpvB-Fusionsprotein eine ADP-Ribosylierungs-Reaktion in den transformierten Bakterien stattfindet. Ob diese Reaktion *de facto* von Trx-SpvB selbst katalysiert wird oder ob Trx-SpvB möglicherweise endogene ARTs in den *E.Coli*-Zellen

induziert, die in der Folge ein 42kD-Zielprotein modifizieren, kann durch diese Experimente nicht unterschieden werden.

Um zu untersuchen, ob die enzymatische Markierung der 42kD-Proteinbande tatsächlich ein direkt durch rekombinantes Trx-SpvB bedingter Effekt ist, wurde das mittels Protein-S-Agarose aufgereinigte Trx-SpvB (**Abbildung 39A**, Spur 2) mit [^{32}P]-NAD^{+} und 2mM „kalter“ ADP-Ribose 60 Minuten bei 37°C inkubiert. Nach der Größenfraktionierung mittels SDS-PAGE wurde durch Auflegen eines Filmes über Nacht auch bei diesem Experiment eine 42kD-Proteinbande sichtbar, die dem rekombinanten Trx-SpvB-Fusionsprotein entspricht (**Abbildung 39B**). Das Trx-SpvB-Fusionsprotein ist somit enzymatisch aktiv und besitzt die Fähigkeit, sich selbst zu ADP-ribosylieren. Es ist folglich davon auszugehen, dass auch die enzymatische Markierung der 42kD-Proteinbande in den Lysaten einen direkten Effekt des rekombinanten Trx-SpvB darstellt.

4.3.4.2 Nachweis der Arginin-Spezifität der ADP-Ribosyltransferase-Aktivität des rekombinanten SpvB

In der Einleitung wurde bereits darauf hingewiesen, dass die meisten mono-ARTs spezifische Zielaminosäuren modifizieren. Eine große Gruppe stellen die Arginin-spezifischen ARTs dar, zu denen z.B. die bakteriellen Toxine CT, LT und VIP2 sowie humanes ART1 und ART5 gehören (Aktories and Just 2000, Domenighini and Rappuoli 1996, Glowacki *et al.* 2002). Da SpvB das charakteristische R-S-EXE-Motiv der Arginin-spezifischen ARTs aufweist, postulierten wir für das SpvB ebenfalls eine Arginin-Spezifität.

Um diese Spezifität für das rekombinante Trx-SpvB nachzuweisen, wurden die mit Protein-S-Agarose aufgereinigten Trx-SpvB-Lysate im Anschluss an die [^{32}P]-NAD^{+}-Inkubation mit verschiedenen Chemikalien versetzt und über Nacht inkubiert. Nach Auftrennung der Proteine wurde die inkorporierte Radioaktivität autoradiographisch detektiert (**Abbildung 39B**). Die inkorporierte radioaktive Markierung war resistent gegenüber der Inkubation mit 10mM Tris-HCl (pH 8.0), 10mM NAD^{+}, 1M NaCl und 10mM $HgCl_2$ (Spuren 1, 2, 3 und 5), aber sensitiv gegenüber der Inkubation mit 1M NH_2OH, bei der die gesamte inkorporierte Aktivität wieder aus dem Protein ausgebaut wurde (Spur 4). Diese Empfindlichkeit gegenüber NH_2OH ist typisch für Arginin-ADP-Ribosyl-Bindungen (Payne *et al.* 1985). Somit kann das SpvB der Gruppe der Arginin-spezifischen ARTs zugeordnet werden.

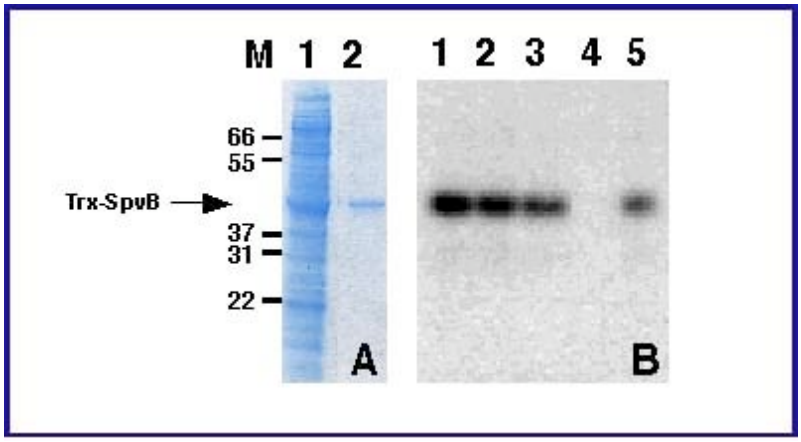


Abbildung 39: Aufreinigung von Trx-SpvB (A) und NH_2OH-Sensitivität der durch Trx-SpvB katalysierten ADP-Ribosyl-Bindungen (B)

Trx-SpvB-Fusionsprotein aus *E.Coli*-Lysaten wurde über den integrierten "S-tag" durch "single step" Affinitätschromatographie an Protein-S-Agarose immobilisiert, aufgereinigt und mittels SDS-PAGE analysiert. Das gezeigte Coomassie-gefärbte Gel (**A**) verdeutlicht die Aufreinigung von Trx-SpvB: **1** *E.Coli*-Gesamtzell-Lysat aus mit pET32a.Trx-SpvB-CD transformierten BL21(DE3)pLysS-Zellen, **2** Aufgereinigtes Trx-SpvB. Das *in silico* errechnete Molekulargewicht von Trx-SpvB beträgt 42.6kD. Um zu ermitteln, dass die in Abb.38 gezeigte ADP-Ribosyl-Transferase-Aktivität in den Trx-SpvB enthaltenden Bakterienlysaten *de facto* ein Effekt des rekombinanten SpvB ist, wurde für diesen Versuch nicht das Gesamtzell-Lysat, sondern das mittels Protein-S-Agarose aufgereinigte Trx-SpvB-Fusionsprotein (**A**, Spur 2) 60 Minuten bei 37°C mit radioaktiv markiertem NAD^+ inkubiert. Anschließend wurde den Proben 10mM Tris-HCl/pH8 (**B**, Spur 1), 10mM NAD^+ (Spur 2), 1M NaCl (Spur 3), 1M NH_2OH (Spur 4) bzw. 10mM $HgCl_2$ (Spur 5) zugesetzt. Nach Inkubation über Nacht wurden die Proben mittels SDS-PAGE-Autoradiographie analysiert. Die Abbildung zeigt die Autoradiographie eines Coomassie-gefärbten Gels (**B,** Exposition 5 Tage bei -80°C). Marker (M = Mark 12, Novex) in kiloDalton (kD).

4.3.4.3 Ermittlung der Zielregion für die ADP-Ribosylierung im rekombinanten Trx-SpvB-Fusionsprotein

Die radioaktive Markierung von Trx-SpvB könnte analog zum Trx-pART4 (siehe 4.2.4.5) durch zwei verschiedene Mechanismen erklärt werden: Entweder modifiziert sich die katalytische Domäne von SpvB selbst oder die assoziierte N-terminale Trx-Domäne stellt zufällig ein potentielles Zielpeptid dar, das durch das SpvB ADP-ribosyliert wird. Um zwischen diesen beiden Möglichkeiten zu unterscheiden, wurde ein Aliquot des aufgereinigten, radioaktiv markierten Fusionsproteins analog zum Trx-pART4 (siehe Abb. 29) mit Enterokinase verdaut, um die katalytische Domäne des SpvB abzuspalten. Nach Größenfraktionierung der matrixgebundenen bzw. der löslichen Proteine aus dem Überstand mittels SDS-PAGE wurden die Proteine auf eine NC-Membran geblottet. **Abbildung 40** zeigt die silbergefärbte NC-Membran (**A**) sowie die Autoradiographie des Blots (**B**). Verglichen mit

der unverdauten Probe (Spur 1) zeigt sich in Spur 2 eine deutliche Abnahme des Trx-SpvB-Fusionsproteins nach Verdau mit Enterokinase (43kD-Bande). Durch die Spaltung treten zwei neue Banden auf, die einerseits den matrixgebundenen Trx-Anteil (17kD-Bande) und andererseits die isolierte katalytische Domäne des SpvB (26kD-Bande) repräsentieren. Autoradiographisch zeigt sich, dass die Aktivität nicht im Bereich der 26kD-Bande, sondern im 17kD-Fragment, das dem Trx-Anteil des Fusionsproteins entspricht, inkorporiert ist (**B**, Spur 2). Während sich bei den löslichen Proteinen (**A**, Spur 3) aus dem Überstand ebenfalls die prominente, dem abgespaltenen und damit nicht mehr matrixgebundenen SpvB entsprechende 26kD-Bande zeigt, bleibt der Trx-Anteil des Fusionsproteins über den "S-tag" matrixgebunden und ist somit im Überstand nicht bzw. nur minimal nachweisbar. Dementsprechend ist in Spur 3 der Autoradiographie auch keine markierte Bande sichtbar. Trx-SpvB modifiziert also im Gegensatz zum Trx-pART4 (vgl. 4.2.4.5) nicht die eigene katalytische Domäne, sondern ADP-ribosyliert die assoziierte Trx-Domäne.

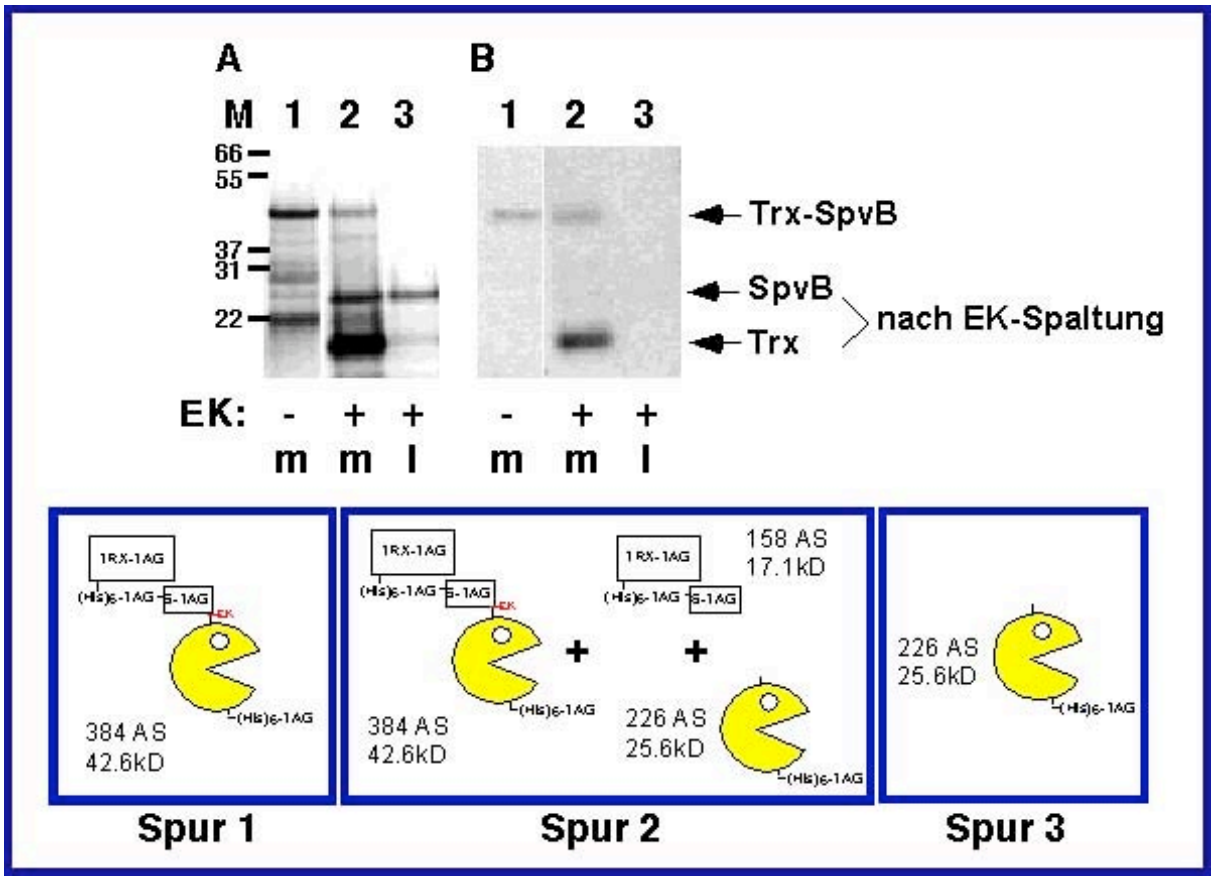


Abbildung 40: ADP-Ribosylierung der assoziierten Trx-Domäne durch rekombinantes Trx-SpvB

Trx-SpvB-Fusionsprotein wurde durch Absorption an immobilisierte Protein-S-Agarose aus *E.Coli*-Lysaten aufgereinigt. Das aufgereinigte Protein wurde in Gegenwart von 2mM ADP-Ribose 60 Minuten bei 37°C mit [^{32}P]-NAD inkubiert. Ein Aliquot der immobilisierten Proteine wurde direkt für die SDS-PAGE-Analyse vorbereitet (Spur 1), die Proben wurden über Nacht bei Raumtemperatur mit Enterokinase (EK) verdaut (Spuren 2,3). Matrix-gebundene (m) Proteine und die löslichen (l) Proteine aus dem Überstand wurden mittels SDS-PAGE aufgetrennt und auf eine NC-Membran geblottet. Die NC-Membran wurde mit Silberlösung gefärbt (**A**). **B** zeigt die Autoradiographie des Blots nach 72h Exposition bei -80°C. Molekulargewichts-Marker (M = Mark 12, Novex) in kD. Unten sind zur Übersicht die rekombinanten Proteine vor und nach der Spaltung schematisiert dargestellt.

Die bisherigen Enzymanalysen zusammenfassend ist folglich festzustellen, dass die C-terminale Domäne von SpvB eine Domäne mit katalytischer Funktion ist, die ADP-Ribosyl-Transferase-Aktivität, aber keine messbare NADase-Aktivität besitzt. Wenn dieses Enzym als rekombinantes Fusionsprotein gekoppelt an eine Trx-Domäne exprimiert wird, katalysiert es die ADP-Ribosylierung der assoziierten N-terminalen Trx-Domäne.

4.3.4.4 *In vitro*-Nachweis der SpvB-katalysierten ADP-Ribosylierung heterologer Proteine in humanen Jurkat-Zell-Lysaten

Mit dem Wissen um die ADP-Ribosyl-Transferase-Aktivität des SpvB stellte sich als nächstes die Frage, ob die isolierte C-terminale Domäne von SpvB fähig ist, die ADP-Ribosylierung heterologer Proteine in den Lysaten von Säugetierzellen zu katalysieren. Zur Klärung dieser Frage benutzten wir erneut Enterokinase, um das aufgereinigte Trx-SpvB-Fusionsprotein in die beiden Anteile zu spalten und den freigesetzten SpvB-Anteil (*in silico* errechnetes MG: 25.6kD) aus dem Überstand zu isolieren (**Abbildung 41A**, Spur 1). Die aufgereinigte C-terminale Domäne des SpvB gaben wir zu den Lysaten von humanen Jurkat-Zellen hinzu. Die anschließende Inkubation mit [^{32}P]-NAD^{+} führte zur radioaktiven Markierung einer 40kD-Proteinbande (**Abbildung 41B**, Spur 3). Im Gegensatz dazu kam es bei der Kontroll-Inkubation mit der SpvB-Domäne ohne Jurkat-Zell-Lysate zu keiner radioaktiven Markierung (**B**, Spur 1). Bei der Kontroll-Inkubation mit Jurkat-Zell-Lysaten ohne SpvB-Domäne zeigte sich lediglich die schwache Markierung einer Proteinbande, die im Bereich größer 100kD liegt (**B**, Spur 2). Die schwache Markierung dieser Bande ist am wahrscheinlichsten auf eine endogene PARP-Aktivität zurückzuführen (D'Amours *et al.* 1999). Diese Vermutung wird durch die Beobachtung unterstützt, dass durch vorsichtiges Entfernen der Zellkerne die Intensität dieser Bande deutlich reduziert werden konnte.

Diese Ergebnisse ermöglichen die Schlussfolgerung, dass die C-terminale Domäne von SpvB die ADP-Ribosylierung eines 40kD Zielproteins in humanen Jurkat-Zell-Lysaten katalysiert. Der Vergleich der Markierung dieses 40kD Proteins mit der Markierung der Trx-Komponente des rekombinanten Trx-SpvB-Fusionsproteins weist darauf hin, dass die ADP-Ribosylierung des Zielproteins in Jurkat-Zell-Lysaten wesentlich effizienter ist als die der Trx-Domäne.

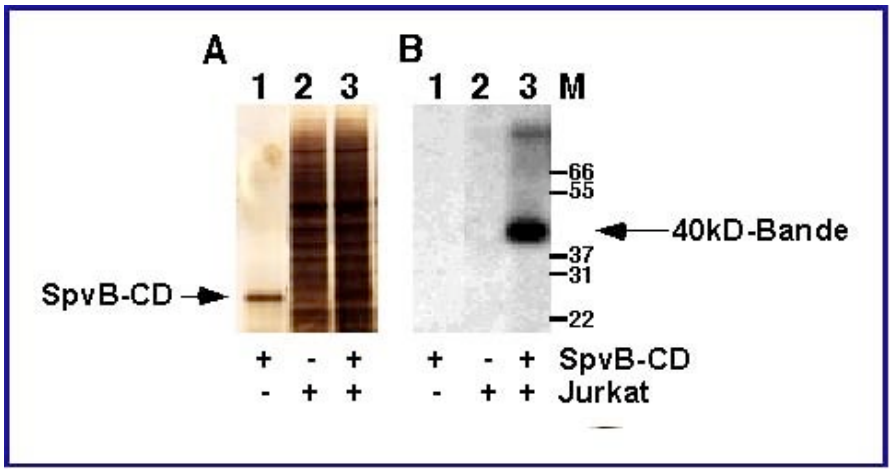


Abbildung 41: SpvB-katalysierte ADP-Ribosylierung heterologer Proteine in humanen Jurkat-Zell-Lysaten

Rekombinantes Trx-SpvB-Fusionsprotein wurde nach Affinitätsaufreinigung mittels Protein-S-Agarose mit Enterokinase inkubiert. Nach der Spaltung des Fusionsproteins wurde die lösliche, isolierte katalytische Domäne des SpvB (SpvB-CD) aus dem Überstand gewonnen. Jurkat-Zellen, eine humane T-Zell-Linie, wurden in 0.5% Triton-X lysiert. Anschließend wurden die Kerne und andere unlösliche Bestandteile der Lysate mit hoher Geschwindigkeit abzentrifugiert. Die lösliche katalytische SpvB-Domäne (Spur 1), ein Aliquot der aufbereiteten Jurkat-Zell-Lysate (Spur 2) sowie ein Aliquot der Jurkat-Zell-Lysate nach Zugabe der löslichen katalytischen SpvB-Domäne (Spur 3) wurden 60 Minuten bei 37°C mit [^{32}P]-NAD inkubiert und einer SDS-PAGE-Analyse zugeführt. Die aufgetrennten Proteine wurden auf eine NC-Membran geblottet und silbergefärbt (**A**). Die Autoradiographie (Film: BioMax MR, Kodak. Exposition: 4 Tage bei -80°C) des Blots ist in **B** gezeigt. Marker (M, Mark 12) in kD.

4.3.4.5 Nachweis der ADP-Ribosylierung von Aktin durch rekombinantes SpvB *in vitro*

Nach der Entdeckung, dass SpvB die ADP-Ribosylierung eines 40kD-Proteins in humanen Jurkat-Zell-Lysaten zu katalysieren vermag, versuchten wir Proteine zu ermitteln, die von SpvB ADP-ribosyliert werden und von der Größe her dieser 40kD-Proteinbande entsprechen könnten. Es zeigte sich, dass die ADP-Ribosylierung einer 40kD-Proteinbande kein für die Jurkat-Zelllinie spezifisches Phänomen ist. Vielmehr wird diese Bande auch in anderen

humanen Zelllinien durch SpvB ADP-ribosyliert (nicht gezeigt). Dies deutet darauf hin, dass es sich um ein ubiquitär verbreitetes Protein handeln muss. Für Bakterientoxine wie VIP2 von *Bacillus cereus* und C2 von *Clostridium botulinum* war bereits bekannt, dass sie in Wirtszellen das Protein Aktin ADP-ribosylieren (Aktories *et al.* 1986, Han *et al.* 1999). Aktin ist ein hochkonserviertes Eukaryonten-Protein. So unterscheidet sich z.B das Aktin des plasmodienbildenden Schleimpilzes *Physarum polycephalum* vom Kaninchenmuskel-Aktin nur in 17 von 375 Aminosäuren (Stryer 1994). Die meisten Eukaryontenzellen enthalten eine große Menge Aktin, das im allgemeinen 10% des gesamten Zellproteins ausmacht. In vielen Zelltypen ist Aktin das Hauptprotein. Aktin ist in den Eukaryontenzellen an der Bildung des Zytoskeletts beteiligt, das den Zellen ihre definierte Gestalt gibt sowie Gestaltänderungen, den Vesikeltransport und die Wanderung von Zellen ermöglicht. Als Folge der ADP-Ribosylierung des Wirtszellen-Aktins durch die o.g. Bakterientoxine wird die Polymerisierung des Aktins inhibiert. Dies führt z.B. im Falle des C2-Toxins zu einer kompletten Depolymerisierung des Aktin-Zytoskeletts und als Folge zum Zelltod (Reuner *et al.* 1987).

Wir untersuchten *in vitro*, ob Aktin ein mögliches Ziel-Protein des SpvB ist. Für diesen Ansatz wurden die Lysate der mit verschiedenen SpvB-Konstrukten transformierten *E.Coli*-Kulturen zusätzlich zu [^{32}P]-NAD^{+} und Enzympuffer mit einem Gemisch aus Kaninchen- und Rinder-Aktin inkubiert. Im Gegensatz zur Inkubation ohne Aktin zeigt die Coomassie-Färbung des Tricin-Geles in allen Spuren eine neue 40kD-Bande, die dem zugegebenen Aktin entspricht (**Abbildung 42**). Diese 40kD-Bande wurde in allen Lysaten, die die katalytische Domäne von SpvB enthielten, ADP-ribosyliert (Autoradiographie**, B** Spur 1-6), während sie in den Kontroll-Lysaten mit isoliertem Trx-Protein nicht markiert wurde (Spur 7,8). Die Markierung der Bande in Spur 3 ist auf dem Originalfilm nur schwach sichtbar und deshalb in der Abbildung kaum zu erkennen. Im Versuchsansatz ohne Aktin zeigt sich in Spur 5 und 6 die bekannte markierte 42kD-Bande, die auto-ADP-ribosyliertem Trx-SpvB entspricht.
Dieses Experiment zeigt, dass Aktin *in vitro* ein Zielprotein der SpvB-katalysierten ADP-Ribosylierung darstellt und dass auch die pET26b-Konstrukte des rekombinanten SpvB eine ART-Aktivität besitzen.

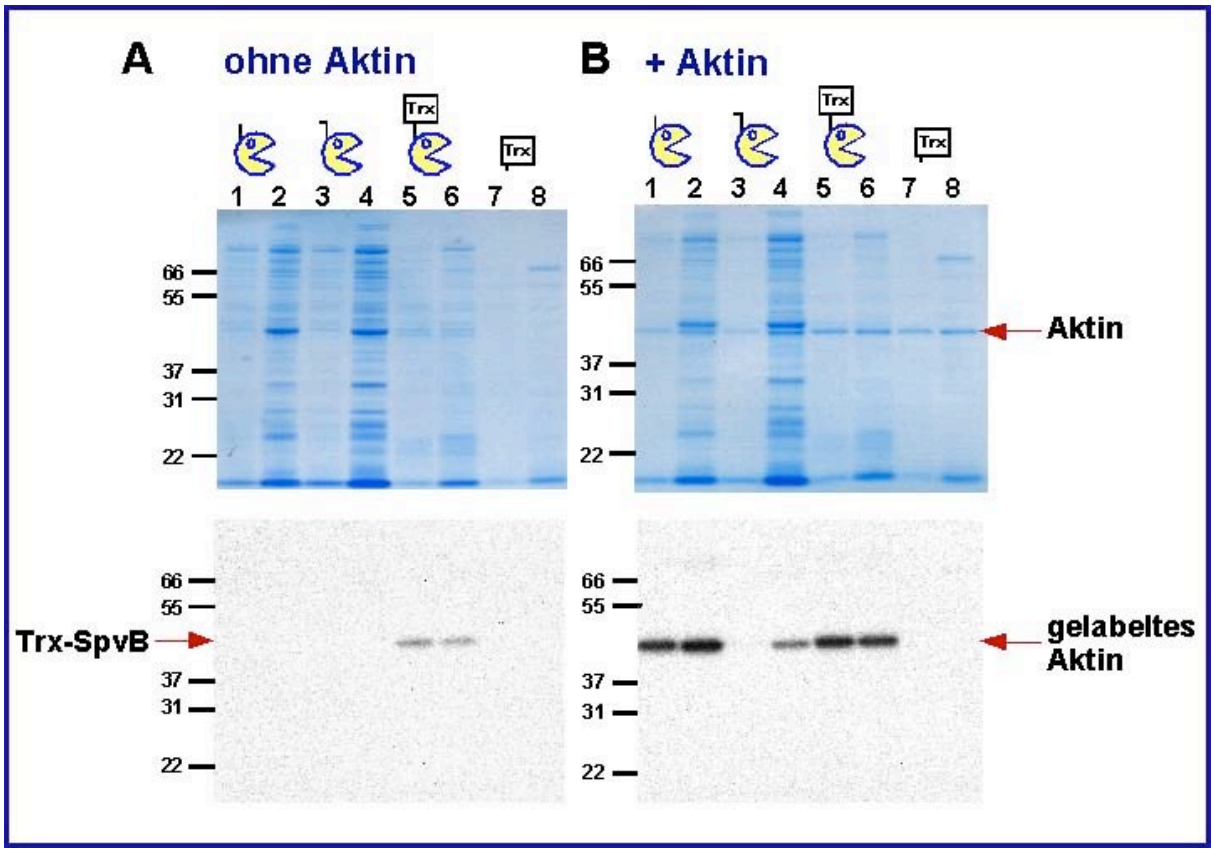


Abbildung 42: Aktin als Zielprotein der SpvB-katalysierten ADP-Ribosylierung in *E.Coli*

In der Abbildung sind Coomassie-gefärbte Tricin-Gele nach SDS-PAGE sowie dazugehörige Autoradiographien dargestellt. Zu den SpvB-Bakterienlysaten wurden Enzympuffer und [^{32}P]-NAD gegeben. Bei Versuchsansatz **A** wurde kein externes Substrat, bei **B** je 0,6µg Kaninchen- und 0,6µg Rinder-Aktin (Sigma) eingesetzt. Anschließend wurden die Ansätze 60min bei 37°C inkubiert. Nach Fällung der Proteine mit StrataClean Resin (Stratagene) wurden diese mittels SDS-PAGE größenfraktioniert. Die Gele wurden Coomassie-gefärbt und ein Röntgenfilm (BioMax, Kodak) über 9d bei -80°C exponiert. Die Spuren enthalten *E.Coli*-Lysate mit folgenden rekombinanten Proteinen (*in silico* errechnetes MG in Klammern): **1/2** pET26b.SpvB-CD (25.5kD), **3/4** pET26b.SpvB-CD.PPl (25.5kD), **5/6** pET32a.Trx-SpvB-CD (42.6kD), **7/8** pET32a.Trx ohne Insert (20.4kD). Spuren 1, 3, 5, 7: BL21(DE3)-Zellen / Spuren 2, 4, 6, 8: BL21(DE3)pLysS-Zellen. Marker (M, Mark 12) in kD. Die produzierten rekombinanten Proteine sind oberhalb der Spuren schematisch abgebildet.

4.3.5 Klonierung eines inaktiven Trx-SpvB-Fusionskonstruktes durch Mutagenese

Wie bereits in der Einleitung beschrieben wurde bei verschiedenen bakteriellen ART-Toxinen nachgewiesen, dass drei Aminosäuren für die katalytische Aktivität essentiell sind: Arginin in β1 (**R**471), Serin in β2 (**S**501) und Glutaminsäure in β5 (**E**538). Um dies auch für das SpvB zu untersuchen, wurden die das Arginin bzw. die Glutaminsäure kodierenden Nukleotid-Tripletts an diesen Positionen durch PCR-Mutagenese gezielt mutiert. In den rekombinanten Proteinen ergibt sich daraus folgender Austausch: Das Arginin (R) in Position 471 wird durch

Asparagin (N) bzw. Lysin (K) ersetzt, die Glutaminsäure (E) in Position 538 durch Alanin (A) bzw. Asparaginsäure (D). **Abbildung 43** zeigt Ausschnitte der entsprechenden Regionen aus den „bunten“ Sequenzen des „Wildtyps“ und der mutierten Konstrukte. Die SpvB-Mutanten stehen nun für weitere Experimente in unserer Arbeitsgruppe zur Verfügung. Sie ermöglichen Untersuchungen zur Klärung, ob die charakteristischen Aminosäuren auch beim SpvB essentiell für die ART-Aktivität sind. Mit Hilfe der Mutanten ohne messbare ART-Funktion kann dann untersucht werden, welche der SpvB-vermittelten Effekte in Wirtszellen *de facto* auf der enzymatischen Aktivität des SpvB beruhen.

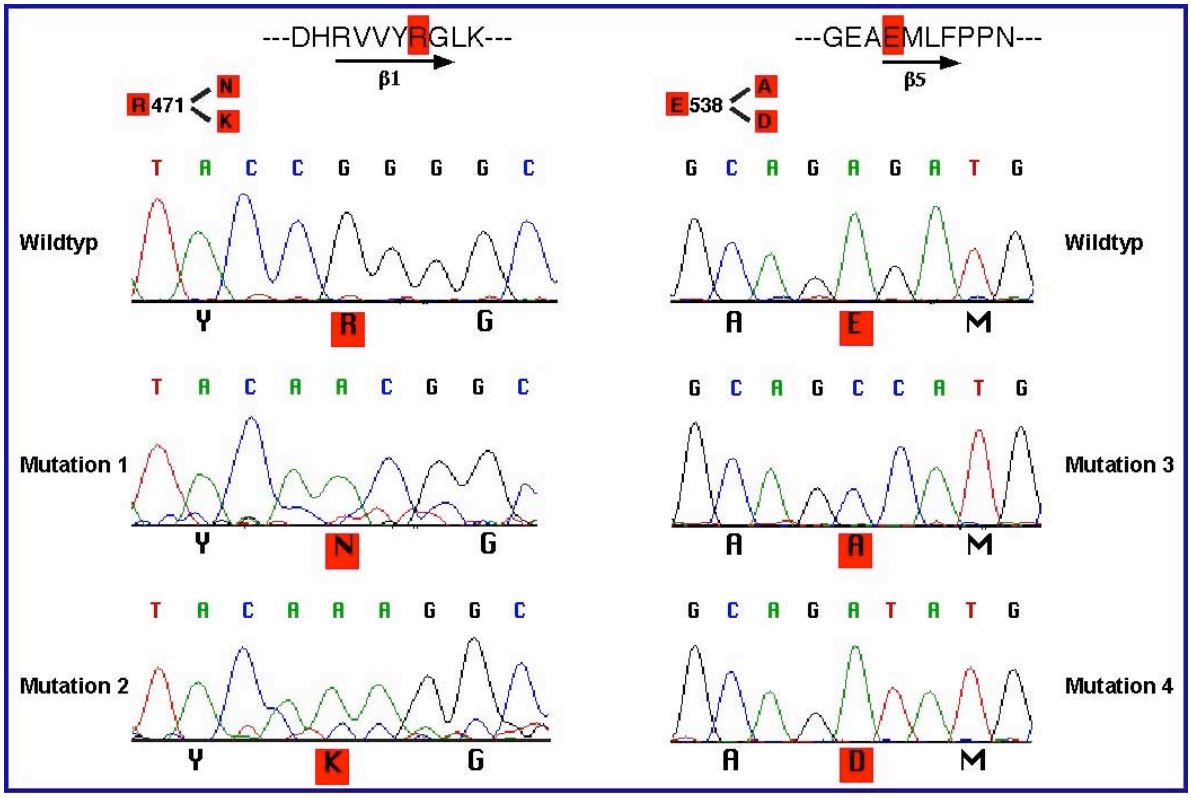


Abbildung 43: Mutagenese der SpvB-Konstrukte

Die *in-vitro*-Mutagenesen wurden mit Hilfe des "QuickChange™ Site-directed Mutagenese Kits" (Stratagene) nach Anleitung durchgeführt. Die Amplifikation des gesamten Plasmids mit den "mutierten" Primern SP11xSP21 bzw. SP12xSP22 erfolgte mit der "proof-reading" Polymerase Pfu und mit 16 PCR-Zyklen. Die Extensionszeit wurde mit 2 Minuten pro Kilobase DNA kalkuliert und betrug somit 14 Minuten. Die Inserts der mutierten „Tochterplasmide“ wurden sequenziert und mit der DNAStar-Software analysiert. In der Abbildung sind Auschnitte der entsprechenden Regionen aus den "bunten" Sequenzen gezeigt. Arginin (R) in Position 471 wurde durch Asparagin (N, Mutation 1) bzw. Lysin (K, Mutation 2) ersetzt, die katalytische Glutaminsäure (E) in Position 536 durch Alanin (A, Mutation 3) bzw. Asparaginsäure (D, Mutation 4). Nukleotide: Grün = Adenin, Blau = Cytosin, Schwarz = Guanin, Rot = Thymidin. Roter Kasten: Aminosäure, die durch die Mutation im Triplett verändert wurde.

5 Diskussion

Die bekannten mono- und poly-ADP-Ribosyltransferasen bilden wichtige Gruppen von Enzymen, die entscheidende Rollen in verschiedensten biologischen Prozessen spielen. Die biologische Bedeutsamkeit der ARTs veranlasste uns, mit den innovativen Möglichkeiten aktueller Suchprogramme in den exponentiell wachsenden Datenbanken nach neuen ART-Familienmitgliedern zu suchen. Die *in silico* Klonierung ist zum einen abhängig von der Qualität der Primärdaten, d.h. z.B. von hochwertigen ESTs aus verschiedenen Geweben, die in Datenbanken wie GenBank, DDBJ und EMBL deponiert sind. Zum anderen werden leistungsstarke Analyseprogramme benötigt, die auch signifikante, niedrige Homologien erkennen und mit den Ausgangsdaten verknüpfen können. Sind diese Voraussetzungen erfüllt, ist die *in silico* Klonierung eine effektive, schnelle und kostensparende Methode um genetische Verwandtschaftsverhältnisse über Speziesgrenzen hinweg aufzuklären und neue Gene zu charakterisieren. Diese Arbeit zeigt den Nutzen des iterativen Datenbanksuchprogrammes PSI-BLAST bei der Identifikation von entfernten Verwandten der mono- und poly-ARTs mit biologisch signifikanten Sequenzhomologien. Ausgewählte neue Familienmitglieder wurden rekombinant exprimiert. Für eine bakterielle mono-ART und eine humane poly-ART konnte die aufgrund der Sequenzhomologie vorhergesagte ADP-Ribosyltransferase-Aktivität mit laborexperimentellen Methoden nachgewiesen werden.

5.1 Die pART-Familie

Nachdem vor fast 40 Jahren erstmals ADP-Ribose-Polymere in eukaryontischen Zellen entdeckt wurden (Reeder *et al.* 1967), konnte die Wichtigkeit der poly-ADP-Ribosylierung für verschiedene zelluläre Prozesse in den letzten Jahrzehnten erforscht werden (D'Amours *et al.* 1999). Das erste klonierte Enyzm beim Menschen, dass diese Poly-ADP-Ribosylierung katalysiert, wurde PARP (**P**oly-**A**DP-**R**ibosyl-**P**olymerase) genannt und intensiv charakterisiert. Da der PARP eine entscheidende Rolle in der DNA-Reparatur, der Regulation der Zellproliferation und –differenzierung, der Genexpression sowie der Induktion des Zellunterganges (Apoptose vs. Nekrose) zugeschrieben wird, ist das molekulare Verständnis der PARP-Funktionen Gegenstand intensiver Forschung.

5.1.1 Die humanen pARTs

Jahrelang wurde angenommen, dass die PARP beim Menschen und anderen Eukaryonten das einzige Enzym mit poly-ART-Funktion sei. Bei der Suche in EST-Datenbanken nach Sequenzähnlichkeiten zur katalytischen Domäne der PARP fielen uns 1997 erstmals mehrere verwandte ESTs auf, so dass wir postulierten, dass beim Menschen mehrere Enzyme mit potentieller poly-ADP-ribosylierender Funktion exprimiert werden. Diese Vermutung wurde durch die Beobachtung, dass in Mausembryo-Fibroblasten von *Parp*-„knockout"-Mäusen weiterhin poly-ART-Aktivität nachzuweisen war (Shieh *et al.* 1998), und durch die Klonierung einer zweiten, strukturell verschiedenen poly-ART in der Pflanze *Arabidopsis thaliana* (Babiychuk *et al.* 1998) unterstützt.

Durch kontinuierliche Datenbanksuchen in den letzten Jahren konnten wir 16 verschiedene, PARP-verwandte humane EST-Cluster in dbEST identifizieren (Abb.9). Die bereits ausführlich charakterisierte *PARP* eingeschlossen werden beim Menschen somit mindestens 17, potentiell zur PARP-Familie gehörende Gene exprimiert. Die PARP-Verwandten wurden pART (für **p**oly-**A**DP-**r**ibosyl**t**ransferase) genannt. Die PARP wurde in unserer Nomenklatur als pART1 definiert. Die Nummerierung der weiteren pARTs (pART2-17) spiegelt die Sequenzidentität zur pART1 und der einzelnen pARTs untereinander im Bereich der Sekundärstrukturmotive des katalytischen Zentrums (66 Aminosäuren) wider (Abb.11). Die 17 humanen pARTs wurden in 5 Subgruppen unterteilt. Die Zuordnung in die Subgruppen wird durch verschiedene, unabhängige Ergebnisse unterstützt:

1. Die Mitglieder einer Subgruppe zeigen eine höhere Sequenzidentität zueinander als zu den pARTs anderer Subgruppen (Abb.10, 11). Die höhere Sequenzidentität innerhalb einer Subgruppe spiegelt sich in den PSI-BLAST-Analysen wider, bei denen die pARTs der zugehörigen Subgruppe zuerst gefunden wurden (Abb.17).

2. Humane pARTs einer bestimmten Subgruppe weisen eine oder mehr spezifische, assoziierte Domänen auf, die in den anderen Subruppen nicht vorkommen (Abb.20). Ausnahmen stellen pART 8, 10 und 15 dar.

3. Mitglieder einer Subgruppe besitzen mindestens eine konservierte Intronposition in dem Genabschnitt, der für die katalytische Domäne kodiert (Abb.15). Ausnahme sind hier die pARTs 1-4 der Subgruppe 1.

4. Wenn mehrere *pART*-Gene in einem Cluster auf einem Chromosom lokalisiert sind, gehören sie derselben Subgruppe an (Abb.12).

Die meisten *pART*-Gene sind bereits durch eine hohe Anzahl an ESTs in dbEST repräsentiert, die in der Regel korrelierend mit dem Wachstum der Datenbank in den letzten Jahren zugenommen haben. Als bisher „jüngste“ *pART* konnte im Oktober 2003 erstmals *pART15* identifiziert werden. Die große Anzahl der pART-ESTs und die Herkunft der ESTs aus vielen verschiedenen Gewebetypen lassen auf eine ubiquitäre Expression der meisten pARTs, vergleichbar mit bekannten Haushältergenen wie z.B. *Glycerinaldehyd-3-Phosphat-Dehydrogenase* (*G3PDH*) und *Hypoxanthin-Guanin-Phosphoribosyl-Transferase* (*HPRT*), schließen. Allerdings ist die *G3PDH* durch wesentlich mehr ESTs repräsentiert. Im Vergleich sind die Mitglieder der eukaryontischen mono-ART-Familie, die eine auf wenige Gewebe begrenzte Expression aufweisen, durch deutlich weniger ESTs repräsentiert (Abb.9).
Interessanterweise wurden für alle pARTs, mit der Ausnahme von pART7, mehr ESTs gefunden als für den Gegenspieler PARG, das einzige bekannte Enzym, das die poly-ADP-Ribose von den Zielproteinen wieder entfernen kann. Es kann vermutet werden, dass die Transkriptionsstärke der hochrepräsentierten *pART*s (wie z.B. *pART1*, *pART4* und die Tankyrasen *pART5* und *pART6*) auf besonders wichtige Rollen der kodierten Enzyme im Stoffwechsel schließen lässt, während die anderen einer besonders strengen Transkriptionskontrolle unterliegen und evtl. eher eine untergeordnete oder eine auf wenige Gewebe beschränkte Funktion besitzen. Bei der Beurteilung der Ergebnisse muss aber berücksichtigt werden, dass aufgrund der sog. Normalisierungsprozeduren zur Erstellung von cDNA-Bibliotheken, die die Transskripte hochexprimierter Gene reduzieren, um die Sequenzierung von seltenen Transskripten wahrscheinlicher zu machen (Soares *et al.* 1994), Aussagen über die Transkriptionsstärke von Genen anhand der Anzahl der ESTs nur eingeschränkt möglich sind. Zudem können selbstverständlich auch Mutationen gering transkribierter Gene zu erheblichen Pathologien führen, wenn die Gene eine Schlüsselfunktion für den Organismus besitzen.

Die Sequenzähnlichkeit der verschiedenen humanen pARTs ist im Bereich der Sekundärstrukturmotive, die bei den bekannten Kristallen der Hühner-pART1 und des Diphterietoxins das katalytische Zentrum umgeben, am stärksten ausgeprägt. Bei der

inzwischen aufgeklärten Kristallstruktur der katalytischen Domäne der murinen PARP-2 (nach unserer Nomenklatur pART2) sind im Vergleich mit der bekannten Kristallstruktur der pART1 alle wichtigen Sekundärstrukturmotive im Zentrum vollständig erhalten (Oliver *et al.* 2004). Überdies zeigte die Analyse der Interaktionen der pART2-Akzeptor-Region mit ADP-Ribose eine Konservierung aller Aminosäurenreste, die in die Koordination mit dem ADP-Anteil involviert sind, und zwar sowohl in ihrer Identität als auch in ihrer räumlichen Anordnung. Auch die „donor"-Region der beiden pARTs im Modell mit gebundenem NAD^+ ist fast identisch gefaltet (Oliver *et al.* 2004). In den Regionen zwischen den zentralen Sekundärstrukturmotiven weichen PARP-1 und PARP-2 stärker voneinander ab. So besitzt PARP-1 z.B. direkt N-terminal vor β5 einen weiteren kurzen β-Strang, der bei PARP-2 fehlt. Interessanterweise handelt es sich hierbei um eine Region, die bei den C3-ähnlichen ART-Toxinen als sog. ARTT-Motiv („ADP-ribosylating turn-turn-motif") beschrieben wurde, und für die zunehmende Hinweise bestehen, dass sie eine gewichtige Rolle bei der Substratspezifität und der Erkennung der Substrate spielt (Barbieri and Sun 2004, Han *et al.* 2001, Han and Tainer 2002, Holbourn *et al.* 2005, Sun *et al.* 2004). Einige ART-Toxine besitzen jedoch trotz ähnlichem ARTT-Motiv unterschiedliche Substratspezifitäten, so dass diese Region nur eine von mehreren Regionen zu sein scheint, die durch Interaktion die Spezifität des Zielproteins determinieren. Bei *Pseudomonas aeruginosa* Exo S und Exo T, die trotz 76%-iger Sequenzidentität unterschiedliche Subtrate ADP-ribosylieren, wurden durch Mutationsexperimente mehrere dieser Regionen identifiziert und durch den Austausch dieser kurzen Regionen (u.a. des ARTT-Motivs und des N-terminal von β1-gelegenen sog. „active site loop") konnte die Substratspezifität entscheidend verändert werden (Sun *et al.* 2004). Die abweichenden Aminosäurenreste in diesen Bereichen führen zur Ausbildung unterschiedlicher elektrostatischer Potentiale in der Enzym-Substrat-Kontaktfläche, die die Substratspezifität beeinflussen (Sun *et al.* 2004).

Aufgrund der Sequenzhomologien im Bereich der entscheidenden Sekundärstrukturmotive ist auch für die anderen pARTs eine vergleichbare Faltung der katalytischen Domäne zu vermuten. Die deutlichen, insbesondere zwischen den einzelnen pART-Subgruppen bestehenden Unterschiede in den „loops" (Abweichung in Länge und Aminosäuresequenz) weisen auf eine abweichende Faltung in diesen Bereichen und damit auf mögliche Unterschiede in der Subtratspezifität hin. Bzgl. der pART-Alignments (Abb.10, 14 und

Anhang D) ist kritisch anzumerken, dass insbesondere bei den pARTs mit niedriger Sequenzhomologie und einigen Regionen, die längere, individuelle Schleifen flankieren, mit den verschiedenen, verwendeten Alignment-Programmen unterschiedliche korrespondierende Regionen zugeordnet wurden. Den gezeigten Alignments liegen in diesen Regionen als Basis zumeist die Ergebnisse des Programmes T-Coffee (www.ch.embnet.org/software/TCoffee) zugrunde, das sich durch besondere Qualitäten beim Alignment niedriger Sequenzhomologien auszeichnet (Notredame *et al.* 2000). Durch Kombination mit den Ergebnissen der PSIPRED- und Gene THREADER-Analysen mit Vorhersage der Sekundärstrukturen gelang es, bei allen pART-Subgruppen (mit der Ausnahme von β4 bei der Subruppe 5) die charakteristischen Sekundärstrukturmotive mit hoher Sicherheit zuzuordnen. Die Aufklärung und vergleichende Analyse von Kristallstrukturen der katalytischen Domänen weiterer pARTs wird genauere Erkenntnisse zu den korrespondierenden Sekundärstrukturmotiven und zur dreidimensionalen Faltung bringen.

Fünf humane pARTs sind analog zur pART1 mit dem charakteristischen HYE-Aminosäuren-Motiv ausgestattet. Im Gegensatz zur Konservierung des Tyrosins (Y) in β2 bei allen Mitgliedern treten an Stelle der Glutaminsäure (E) in β5 verschiedene abweichende Varianten auf; so ist diese Position bei pART 10, 11, 12, 14, 16 und 17 durch Isoleucin (I) und bei pART13 durch Valin (V) besetzt. Interessanterweise konnte bei verschiedenen mono-ARTs, wie z.B. dem Diphterietoxin, und der PARP durch gezielte Mutagenese gerade für diese Glutaminsäure (E) in β5 gezeigt werden, dass sie für die katalytische ART-Aktivität entscheidend ist (Bazan and Koch-Nolte 1997, Carroll and Collier 1984, Domenighini and Rappuoli 1996, Marsischky *et al.* 1995) und dass selbst ein konservativer Austausch dieser Aminosäure zum Verlust von Aktivität führt. Folglich lag zunächst die Vermutung nahe, dass diese Glutaminsäure auch bei den anderen pARTs essentiell ist. Es schien unwahrscheinlich, dass die Funktion des sauren Glutamats durch hydrophobe Aminosäuren wie Isoleucin oder Valin ersetzt werden kann, und ließ vermuten, dass es sich bei den in dieser Position abweichenden pARTs um inaktive Varianten handelt. Diese Vermutung wurde in den letzten Jahren zunächst durch die Tatsache unterstrichen, dass es sich bei allen pARTs, für die bis vor kurzem eine ADP-Ribosyltransferase-Aktivität nachgewiesen wurde, um pARTs mit dem charakteristischen HYE-Motiv handelte (z.B. pART2/PARP-2, pART4/vPARP oder die Tankyrasen pART5 und pART6).

Auch ohne eine ART-Aktivität könnten die „inaktiven“ pARTs wichtige Funktionen erfüllen, z.B. durch allosterische Effekte oder bei ähnlichem Zielprotein als kompetitive Antagonisten von aktiven pARTs. Beispiele für die Ausbildung neuer Funktionen durch inaktivierte Mitglieder von Enzymfamilien finden sich z.B. bei den bakteriellen ART-Toxinen, wo die inaktive N-terminale ART-Domäne der binären bakteriellen Toxine VIP2 von *Bacillus cereus* und C2 von *Clostridium botulinum* eine wichtige proteinbindende Funktion besitzt (Barth *et al.* 1998a, Han *et al.* 1999). Auch in der Kristallstruktur des sog. letalen Faktor („anthrax lethal factor“) des Milzbranderregers *Bacillus anthracis* wurde eine ART-verwandte Domäne entdeckt, die mehrere zentrale Mutationen aufweist, so dass eine katalytische ART-Aktivität unwahrscheinlich ist (Pannifer *et al.* 2001). Die mutierte ART-Domäne scheint aber wichtig für die Substraterkennung zu sein. Auch in anderen Enzymfamilien kommen solche strukturell verwandten, aber inaktiven Mitglieder vor: Z.B. gleichen sog. zelluläre FLIPs strukturell der Caspase 8, besitzen aber keine proteolytische Aktivität und inhibieren deshalb apoptotische Signale des „Todesrezeptors“ Fas, einem Mitglied der Tumor-Nekrose-Faktor (TNF)-Rezeptor-Familie (Tschopp *et al.* 1998).
Das Fehlen der katalytischen Glutaminsäure könnte auch einen Einfluß auf die Elongationsaktivität der pARTs haben, denn die meisten poly-ADP-ribose-Polymere, die durch immunologische Färbungen (polyklonale Antikörper gegen poly-ADP-Ribose) *in vivo* nachgewiesen werden können, scheinen kürzere Polymere zu sein (Gagne *et al.* 2001).

Überraschenderweise wurde 2001 von Ma *et al.* die Klonierung einer neuen PARP-Verwandten, genannt TiPARP, bei der Maus publiziert, deren Expression in Retikulozytenlysaten die PARP-Aktivität um das 4-fache gegenüber den Kontroll-Lysaten erhöhte (Ma *et al.* 2001). Von den Autoren wurde die Zunahme der PARP-Aktivität als Fähigkeit der TiPARP gewertet, Histone zu ADP-ribosylieren. Die TiPARP entspricht nach unserer Nomenklatur dem Maus-Homologen von humanem pART14 und besitzt analog zum humanen Verwandten ein Isoleucin (I) anstelle der Glutaminsäure (E) in β5. Dies wäre somit der bahnbrechende Nachweis, dass auch diese pART-Varianten eine PARP-Aktivität ausbilden können. Allerdings ist kritisch zu betrachten, dass die in der Publikation vorgestellten Ergebnisse keine eindeutigen Rückschlüsse zulassen. Ursache der beobachteten Steigerung der PARP-Aktivität gegen Histone könnte z.B. auch eine, durch die TiPARP vermittelte Induktion der endogenen PARP-Aktivität sein.

In diesem Jahr wurde nun bei humanem pART10 die PARP-Aktivität überzeugend nachgewiesen (Yu *et al.* 2005). pART 10 besitzt ebenfalls ein HYI-Motiv. Dies lässt vermuten, daß auch die anderen pARTs mit diesem Motiv (u.a. pART11, 12, 16 und 17) eine ART-Aktivität aufweisen. Die ebenfalls hydrophoben Aminosäuren Leucin und Valin könnten an dieser Position einen dem Isoleucin vergleichbaren Effekt haben, so dass folglich auch bei diesen pARTs eine ART-Aktivität wahrscheinlich ist (pART7, 8 und 13). Interessanterweise ist auch das RT6-verwandte humane ART3, für das bisher keine Enzymaktivität nachgewiesen werden konnte, durch ein Isoleucin (I) anstelle des Glutamat in β5 charakterisiert (Glowacki *et al.* 2002). Ob weitere, in dieser entscheidenden Position in β5 abweichende mARTs und pARTs eine ADP-ribosylierende Enzymaktivität besitzen oder ob es sich bei diesen pARTs um „loss of function“ (Funktionsverlust) bzw. um „gain of other function“ (Funktionsgewinn)-Varianten handelt, ist bisher ungeklärt und wird in den nächsten Jahren Gegenstand der Forschung sein. Es wird von Interesse sein, ob und wie die positionsspezifische Mutation des HYE-Motivs zu den Motiven von pART7-17 (und umgekehrt) die Enzymaktivität beeinflusst. Besonders reizvoll werden hierbei die Untersuchungen des humanen und murinen pART9 sein, bei denen das Motiv am deutlichsten von dem klassischen HYE-Motiv abweicht (QYT).

5.1.2 pART-Familienmitglieder in verschiedenen Spezies

Mit PSI-BLAST-Analysen wurden homologe pARTs in vielen verschiedenen Spezies aus den Reichen der Tiere, der Pilze, der Pflanzen und der Protisten identifiziert. Dabei beträgt der Grad der Sequenzidentität einiger Verwandter deutlich weniger als 20%, so dass mit traditionellen Datenbank-Suchprogrammen eine Verknüpfung nicht möglich gewesen wäre. Bei Ratte und Maus konnten mit der Ausnahme von pART7 für alle humanen pARTs hochkonservierte Orthologe aufgespürt werden. Auch beim Huhn und dem Fisch *Tetraodon nigroviridis* scheint das pART7-Homologe zu fehlen. Bei *C.elegans* und der Mücke *Anopheles gambiae* wurden je drei, bei *Drosophila melanogaster* zwei, beim Schleimpilz *Dictyostelium discoideum* neun und im Genom von *Arabidopsis thaliana* 10 potentielle Familienmitglieder nachgewiesen. Viele pARTs konnten anhand der Kriterien Sequenzidentität, Domänenstruktur und konservierte Intronpositionen einer bestimmten humanen pART als Homologes zugeordnet werden. Bei den pARTs entfernterer Spezies war häufig nur die Einteilung in die entsprechende Subgruppe möglich (Abb.17).

Interessanterweise wurde bei den Doppelstrang-DNA-Viren Chilo iridescent und Bacteriophagus Aeh1 je eine mikrobielle pART entdeckt. Möglicherweise haben diese großen DNA-Viren ihre pARTs über horizontalen Gentransfer erhalten. Mit einer chimären pART (Iridovirus/DT) gelang uns zudem erstmals die Verknüpfung der bakteriellen Toxine DT, ETA und eines ähnlichen Toxins von *Vibrio cholerae* mit den eukaryontischen pARTs. In den vollständig entschlüsselten Genomen der Hefe und der Archaebakterien konnten dagegen keine Verwandten ermittelt werden. Dies steht im auffälligen Gegensatz zu den Analysen bei den mono-ARTs (siehe 5.2.1). Desweiteren konnte die DT/ETA/pARTs-Gruppe durch PSI-BLAST nie mit einem bekannten Mitglied aus der RSE-Motiv-Familie der bakteriellen Toxine und Vertebraten-mono-ARTs verknüpft werden. Jedoch wurden in einigen Fällen Familienmitglieder der RNA:NAD-2`-Phosphotransferasen gefunden, deren Sequenzidentität in der Region um den β1-Strang nur knapp unterhalb des „threshold“ lag (nicht gezeigt). Diese Enzyme katalysieren eine Reaktion beim tRNA-Splicing, die der durch die ARTs katalysierten Reaktion ähnelt, aber in der die ADP-Ribose statt auf die Aminosäure eines Proteins auf das 2`-Phosphat der tRNA übertragen wird (Spinelli *et al.* 1999). Ob diese niedrig signifikanten Verknüpfungen auf eine konservierte NAD-bindende Struktur hinweisen, müssen weitere Analysen z.B. durch Aufklärung einer Kristallstruktur dieser Enzyme zeigen.

5.1.3 pART-Genstrukturen

Die Aufschlüsselung der Genstruktur im Bereich der katalytischen Domäne zeigte, dass die Subgruppen zumeist durch gemeinsame, konservierte Intronpositionen charakterisiert sind, die es wahrscheinlich machen, dass diese Gene sich im Laufe der Evolution aus einem gemeinsamen Vorläufergen durch Genduplikation entwickelt haben. Interessanterweise fehlen gerade in der Subgruppe 1, die sich durch die höchste Sequenzidentität zur hpART1 auszeichnet, für die gesamte Gruppe charakteristische konservierte Intronpositionen. Bei einzelnen, bzgl. der Sequenzidentität der kodierten Proteine nah verwandten Mitgliedern (z.B. *pART5* und *pART6* sowie *pART16* und *pART17*) sind dagegen alle Intronpositionen hochkonserviert. Dies weist darauf hin, dass diese Gene erst kürzlich in der Evolution durch Genduplikation entstanden sind. Besonders auffällig ist, dass alle humanen *pART*s bis auf *pART4* und *14* ein Phase-0-Intron kurz vor dem β1-kodierenden Bereich besitzen (Abb.14, 15). Das Phase-0-Intron der pART1 ist bei allen pARTs der Subgruppen 2 und 3 konserviert. Dieses Phase-0-Intron wurde verwendet, um das N-terminale Ende der

katalytischen Domäne festzulegen. N-terminal dieser Position schließt sich bei den pARTs der Subgruppe 1 eine alpha-helicale Domäne an (Abb.20), die in mehreren Publikationen der PARP-katalytischen Domäne zugerechnet wurde. Da aber den pART-Familienmitgliedern der anderen Subgruppen diese Region fehlt, sollte sie als unabhängige Domäne von der katalytischen Domäne getrennt werden. Zudem korreliert das durch die Phase-0-Introns festgelegte N-terminale Ende der katalytischen Domäne sowohl mit dem N-terminalen Ende der isolierten pART-Domäne des Chilo iridescent Virus als auch mit dem N-terminalen Ende des Diphterietoxins und des Pseudomonas Exotoxin A nach proteolytischer Entfernung der Signalsequenzen bzw. der Translokationsdomäne. Das Phase-0-Intron ermöglicht die Kombination des Moduls, das die katalytische Domäne kodiert, mit anderen Exons (sog. „Exon-Shuffling“) und damit die Entstehung von ART-Proteinen mit neuen Funktionen in der Evolution.

Mit Ausnahme von *pART4* liegt das für die katalytische Region kodierende Modul der humanen *pART*-Gene am 3`-Ende (Abb.15). Der Vergleich mit den Genstrukturen von pART-Homologen anderer Spezies zeigt, dass sich die *pART*-Gene sowohl in der Anzahl als auch in der Größe der Exons deutlich unterscheiden. Iridovirus *pART* und *Drosophila pARTa* sind die bisher einzigen charakterisierten Familienmitglieder, bei denen die katalytische Domäne durch ein einziges Exon zusammenhängend kodiert wird, während bei vielen anderen die katalytische Region in verschiedene Exons aufgesplittert ist. Dies ist ein weiterer bemerkenswerter Unterschied zu den humanen mono-ARTs, bei denen die katalytische Domäne durch ein einzelnes Exon kodiert wird (Glowacki *et al.* 2002). Einzelne Module der pARTs werden zumeist durch Introns voneinander abgetrennt. An diesen Positionen kommen gehäuft Phase-0-Introns vor (Abb.19).

Die Konservierung von Introns zwischen humanen *pARTs* und *pARTs* anderer Spezies unterstreicht erneut die aufgrund der Sequenzidentität der Proteine postulierte Verwandtschaft und verdeutlicht, dass einige *pART*s Orthologe in anderen Spezies besitzen (z.B. hs*pART2* und ce*pARTb*, Abb.19). Mittels einer Analyse der konservierten Intronpositionen zwischen den humanen *pART*s untereinander und den *pARTs* anderer Spezies versuchten wir die potentiellen Genstrukturen der katalytischen Region von gemeinsamen Vorläufergenen zu ermitteln, aus denen sich die Vielzahl der *pART*-Gene im

Verlauf der Evolution abgeleitet hat. Insbesondere konservierte Intronpositionen zwischen verschiedenen Subgruppen (z.B. besitzt at*pARTa* einerseits zwei konservierte Intronpositionen mit hs*pART5* und andererseits eine gemeinsame Intronposition mit hs*pART2*, während hs*pART5* und hs*pART2* keine gemeinsamen Introns aufweisen, siehe Abb.19) zeigen, dass sich die evolutionsgeschichtliche Entwicklung der Intronpositionen weder alleine durch eine Theorie des „Einspringens“ neuer Introns (sog.“introns late“-Hypothese) noch durch den Verlust von im Urgen vorhandenen Introns (sog. „introns early“-Hypothese oder auch „Exon theory of genes“, (Long *et al.* 1995)) erklären lässt. Vielmehr muss es sich analog zu den neueren Theorien um ein gemischtes Modell handeln (Roy 2003).

5.1.4 Chromosomale Lokalisierung der humanen und murinen *pART*s

Mit Hilfe von Datenbankanalysen konnte die chromosomale Lage der humanen *pART*s und ihrer murinen Orthologen ermittelt werden. Auf den Gonosomen liegen bei diesen Spezies keine *pART*-Gene, während sich auf einzelnen Autosomen gleich mehrere befinden, wobei einige sowohl bei der Maus als auch beim Menschen eng benachbart sind, so dass von Kopplungsgruppen auszugehen ist. Die eng benachbarten Gene gehören immer derselben Subgruppe an. Die Anordnung von Genfamilien in solchen „Genclustern“ lässt sich durch das Modell der regionalen Genduplikation erklären (Ohno 1999). Kommt es bei der Zellteilung zu einer versehentlichen Verdopplung eines Genes, enthält das Tochterzellgenom zwei identische Kopien. Die Folgen der Genduplikation hängen von den bestehenden Selektionsbedingungen ab: Häufig verhindern zunehmende Punktmutationen in einer der beiden Kopien des redundanten Genes z.B. durch Entstehen vorzeitiger Stopkodons, Verlust von Splicestellen oder Verschiebungen im Leseraster eine korrekte Translation und die Genkopie wird somit inaktiviert. Solche Pseudogene finden sich z.B. mit dem humanen *ART2AP* und *ART2BP* in der Familie der mono-*ART*s (Glowacki *et al.* 2002, Haag *et al.* 1994). Andererseits können Mutationen in den Genkopien durch Veränderung der Regulation oder des kodierten Proteinprodukts auch zu einem Funktionszuwachs („gain of function“) wie z.B. gesteigerter Expressionsrate, neuer Gewebespezifität, erhöhter Enzymaktivität oder einem neuen Zielprotein führen. Der Nachweis von ESTs für alle humanen *pART*s in dbEST zeigt, dass auch die eng verwandten humanen *pART*-Gene exprimiert werden (z.B. *pART16* und *pART17*) und es sich somit um „aktive“ Gene handelt, wobei die Funktionen vieler Varianten noch ungeklärt sind.

Bis auf *pART7* konnten die Orthologen aller humanen *pART*s im Genom von Maus und Ratte nachgewiesen werden. Das Vorhandensein der entsprechenden Orthologen in verschiedenen Säugetierspezies weist darauf hin, dass diese Gene z.B. durch Genduplikation während oder vor der Abzweigung der Säugetiere in der Evolution entstanden sind. Da *pART7*, *pART8* und *pART9* beim Menschen eng benachbart auf dem Chromosom 3 liegen, wurde diese Region mit der analogen Region auf dem murinen Chromosom 16 verglichen (Abb.13). Während die Informationen für alle weiteren in dieser Region liegenden Gene (u.a. *RHYSIN 2*, *HSPBAP 1*) bei Mensch und Maus gleichermaßen vorhanden sind, fehlt der Maus das *pART7*-Gen. Da auch bei der Ratte kein *pART7*-Orthologes nachgewiesen werden konnte, muss den Nagern dieses Gen entweder verloren gegangen sein oder nach der evolutionären Abzweigung der Nager erfolgte bei den höheren Säugetieren eine regionale Genduplikation z.B. von *pART8*, das beim Menschen in direkter Nachbarschaft liegt und dessen Proteinprodukt eine hohe Sequenzhomologie zum pART7 besitzt (ca. 80% im Bereich der charakteristischen Sekundärstrukturmotive, Abb.11). Interessanterweise konnte bisher auch beim Fisch *Tetraodon nigroviridis* und beim Huhn kein pART7-Orthologes nachgewiesen werden. Nach Fertigstellung weiterer Genomprojekte, insbesondere von Primaten und anderen höheren Säugern, wird sich zeigen, ob das *pART7*-Gen auch bei anderen Säugetieren vorkommt oder ob es sich um ein auf Primaten (oder sogar den Menschen) beschränktes Gen handelt. Das Beispiel des *pART7* macht deutlich, welche Möglichkeiten „Genom Browser“ und andere *in silico* Programme in der Zukunft bieten werden, um zur Aufklärung evolutionärer Zusammenhänge beizutragen.

5.1.5 Domänenstrukturen der pARTs

In dieser Arbeit wird gezeigt, dass die funktionelle Einheit „katalytische ART-Domäne“ bei den verschiedenen pARTs beeindruckenderweise mit unterschiedlichsten funktionellen Einheiten fusioniert ist. Die schematische Darstellung in Abbildung 20 verdeutlicht, dass die Proteine im Sinne eines „Baukastenprinzips“ („lego-like“) aus unterschiedlichen Bestandteilen zusammengesetzt sind. Mit Ausnahme von pART4 bildet die katalytische ART-Domäne bei allen humanen pARTs das C-terminale Ende des Proteins. In der Regel sind die Subgruppen durch mindestens eine charakteristische Domäne gekennzeichnet, die alle humanen pARTs der jeweiligen Subgruppe aufweisen und die bei pARTs anderer

Subgruppen nicht vorkommen (Subgruppe 1: alpha-helicale, PARP-regulatorische Domäne / Subgruppe 2: Ankyrin-Domäne / Subgruppe 3: A1pp-Domäne / Subgruppe 4: WWE-Domäne / Subgruppe 5: trunkierte pART-Domäne). Ausnahmen von dieser Regel stellen humane pART8 (zusätzliche WWE-Domäne, die eigentlich für die Subgruppe 4 charakteristisch ist), pART10 (die A1PP-Domäne fehlt) und pART15 (N-terminal ist bisher keine konservierte assoziierte Domäne nachweisbar) dar.
Die assoziierten Domänen können vier funktionellen Gruppen zugeordnet werden: Neben Nukleinsäure-bindenden Domänen (z.B. DNA-bindende Domäne bei pART1 und 2) und Domänen, die mit der Ubiquitinierung in Zusammenhang stehen (z.B. WWE-Domäne), kommen Protein-Interaktions-Domänen (z.B. A1pp-, WGR-, BRCT-Domäne) und potentielle katalytische Domänen vor (z.B. UBCc-Domäne). Die einzigartige Struktur einiger Familienmitglieder lässt ein breites Spektrum an unterschiedlichen Funktionen der pARTs vermuten.

Für jede Subgruppe wurde in der Datenbank mindestens eine pART in entfernten Spezies (Insekten, Pflanzen, Pilzen oder Amöben) identifiziert, die eine charakteristische, assoziierte Domäne der jeweiligen Subgruppe enthält. Dies weist darauf hin, dass es sich bei diesen Domänenkombinationen um sehr alte Schöpfungen handelt. Bisher wurde mit der pART des Virus Chilo iridescent lediglich eine pART entdeckt, die vergleichbar den eukaryontischen mono-ARTs lediglich aus einer isolierten katalytischen Domäne besteht (Glowacki *et al.* 2002). Besondere Beispiele für den Austausch von Domänen zwischen den Subgruppen sind die humane pART8 (WWE- und A1pp-Domänen) und die pARTc von *C.elegans*, die die charakteristische Ankyrindomäne der Subgruppe 2 und gleichzeitig, statt der SAM-Domäne, die für die Subgruppe 1 typische Kombination aus alpha-helicaler Domäne und WGR-Domäne aufweist. Neben den pART4-Orthologen sind die pARTc des Mikropilzes *Gibberella zeae* und ihr Orthologes beim Schimmelpilz *Aspergillus nidulans* die einzigen bisher bekannten pARTs, bei denen C-terminal der pART-katalytischen Domäne eine weitere Domäne assoziiert ist (Abb.20). Diese Domäne zeichnet sich durch eine hohe Sequenzidentität zur katalytischen Domäne des Ubiquitin-konjugierenden Enzymes E2 aus, das als eine zentrale Komponente bei der Ubiquitin-abhängigen Degradierung von Proteinen fungiert. Da das zentrale, aktive Cystein, das die COOH-Gruppe des Ubiquitins bindet, bei der GZ.pARTc vorhanden ist, scheint eine katalytische Aktivität möglich. Die katalytische

Aktivität beider Domänen vorausgesetzt wäre dies die bemerkenswerte Variante einer pART mit zwei, in einem Polypeptid zusammengefassten katalytischen Domänen, die auf unterschiedliche Weise Proteine modifizieren (ADP-Ribosylierung und Ubiquitinierung). Neben dieser pART-Variante weisen die vielen pARTs, bei denen die pART-katalytische Domäne mit Domänen fusioniert ist, die mit der Ubiquitinierung in Zusammenhang stehen, auf eine Verbindung dieser Mechanismen der posttranslationalen Proteinmodifikation hin. Tatsächlich wurde kürzlich gezeigt, dass die ADP-Ribosylierung von TRF1 durch Tankyrase 1 (pART5) zu dessen Ablösung von den Telomeren und seiner anschließenden Ubiquitinierung führt (Chang *et al.* 2003).

Nach *atpARTa* und *atpARTb* (Babiychuk *et al.* 1998) sowie *ce.pARTa* und *ce.pARTb* (Gagnon *et al.* 2002) wurde kürzlich auch *ce.pARTc* molekular kloniert und das Proteinprodukt genauer charakterisiert (Gravel *et al.* 2004). Neben der Bestätigung unserer *in silico* Ergebnisse (Exonanzahl, kodierende Sequenz, im Gegensatz zu den Tankyrasen der höheren Eukaryonten fehlende SAM-Domäne) konnte u.a. gezeigt werden, dass durch alternatives Splicen zwei Isoformen produziert werden und dass DNA-Schädigung zu einer gesteigerten Expression des *ce.pARTc*-Genes führt. Interessanterweise fehlt beim Tankyrase-Orthologen in *Dictyostelium discoideum* (nach unserer Nomenklatur dd.*pARTb*) ebenfalls die SAM-Domäne. Während bei höheren Eukaryonten wie Mensch, Maus, Ratte und Huhn zwei *Tankyrasen* im Genom vorhanden sind, wurde in den niederen Eukaryonten *D.melanogaster*, *C.elegans* und *D.discoideum* jeweils erst eine *Tankyrase* detektiert. Dies ist ein Hinweis auf die Entwicklung der *Tankyrase*-Gene höherer Eukaryonten aus einem gemeinsamen Ahnen, der bereits in Organismen vorkommt, die in der Evolution nicht weit entfernt sind. Möglicherweise haben sich die Funktionen, die sonst durch eine einzelnes Gen sichergestellt wurden, bei den höheren Eukaryonten auf zwei Gene aufgeteilt.

5.1.6 Weitere Charakterisierung und Funktionen der humanen pARTs

Die pART1 wurde in den letzten Jahren weiter intensiv erforscht. Wegen der Vielzahl an Publikationen wird hier nur kurz auf die wesentlichen Erkenntnisse eingegangen. Die rekombinante Expression und genauere Charakterisierung von pART3 und pART4 waren Teil dieser Arbeit. Über einige pARTs (z.B. die Tankyrasen und vPARP) wurden in den letzten Jahren von anderen Arbeitsgruppen herausragende Ergebnisse publiziert, die in der

Folge zusammenfassend vorgestellt werden. Dagegen ist über die Struktur und Funktion anderer Mitglieder der pART-Familie bisher wenig bekannt.

5.1.6.1 pART1 / PARP 1

Zur Struktur und den Funktionen des „Familienoberhauptes“ pART1, die bereits in der Einleitung kurz beschrieben wurde (siehe 1.6.1), sind bisher die meisten Erkenntnisse vorhanden. pART1 ist normalerweise in weitgehend inaktiver Form im Nukleoplasma lokalisiert. Das Auftreten von DNA-Strang-Brüchen aktiviert die pART1, die daraufhin sich selbst und verschiedene nukleäre Akzeptorproteine (z.B. XRCC1, DNA-Polymerase β und DNA-Ligase III), die in die DNA-Reparatur involviert sind, ADP-ribosyliert. Diese Modifikationen tragen entscheidend zum Überleben der geschädigten Zellen bei (D'Amours *et al.* 1999). Die Bestätigung der entscheidenden Rolle der pART1 in der DNA-Reparatur zeigte sich bei den pART1-„knockout“-Mäusen, die eine Hypersensitivität auf ionisierende Bestrahlung und alkylierende Substanzen und eine erhöhte genomische Instabilität aufweisen (Shall and de Murcia 2000). Weitere Untersuchungen wiesen auf zunächst unvereinbar erscheinende Funktionen der pART1 bei verschiedenen Prozessen hin, die zum Zelltod führen: Einerseits wird pART1 bei der Apoptose durch Caspasen frühzeitig gespalten und inaktiviert, wobei die Inaktivierung der pART1 für einige Formen der Apoptose zwingend erforderlich ist (Virag and Szabo 2002). Auch bei Caspase-unabhängigen Formen der Apoptose mittels des sog. AIF (**a**poptosis **i**nducing **f**actor) wurden für die pART1 entscheidende Funktionen nachgewiesen (Yu *et al.* 2002). Andererseits führt eine Überaktivierung der pART1 zur folgenschwereren Art des Zelltodes, der Nekrose. Dies wurde durch die Beobachtung unterstützt, dass die Inhibition der pART1-Aktivität bzw. das Fehlen der pART1 (k.o.-Mäuse) insbesondere bei Krankheitsmodellen wie der cerebralen und kardialen Ischämie, die vornehmlich durch nekrotischen Zelltod charakterisiert sind, einen protekiven Effekt hatte (Shall and de Murcia 2000, Virag and Szabo 2002). Die zunächst widersprüchlich erscheinenden Erkenntnisse wurden inzwischen in einem Modell zusammengeführt, nach dem die drei Prozesse Überleben, Apoptose bzw. Nekrose abhängig von der Art des DNA-schädigenden Agens und der Intensität der DNA-Schädigung in den Zellen ausgelöst werden (Virag and Szabo 2002). Es ist anzunehmen, dass pART1 bei dieser Determination eine entscheidende Rolle spielt. Die zentrale Position der pART1 könnte auch

die widersprüchlichen Ergebnisse bzgl. der pART1-Funktion bei der Tumorgenese (Protektion *vs.* Promotion) erklären.

Desweiteren spielt die pART1 eine wesentliche Rolle bei der Regulation der Gentranskription, indem sie direkt mit verschiedenen Transkriptionsfaktoren als Cofaktor interagiert (z.B. NF-κB) oder deren Funktion durch ADP-Ribosylierung modifiziert (z.B. p53) (Kraus and Lis 2003). Die PARP-Aktivität ist dabei nicht in allen Fällen zwingend erforderlich. Über die Regulation der Transkription von verschiedenen Entzündungsfaktoren wie Zytokinen, Adhäsionsmolekülen und Mediatoren (z.B. iNOS) greift pART1 entscheidend in entzündliche Prozesse ein. So weisen PARP-„knockout"-Mäuse z.B. durch die fehlende Aktivierung des Transkriptionsaktivators NF-κB eine deutlich geringe Empfindlichkeit gegenüber dem LPS (Lipopolysaccharid)-induzierten Schock auf (Oliver *et al.* 1999).
Die zentrale Position der pART1 bei so entscheidenden Prozessen wie der DNA-Reparatur, der Induktion des Zelluntergangs (Apoptose *vs.* Nekrose) und der Genexpression erklärt die Beteiligung der PARP in der Pathogenese verschiedenster Erkrankungen (u.a. bei Diabetes mellitus Typ 1, Myokardischämie, Kolitis, Arthritis und Tumoren) und damit das hohe Forschungsinteresse (Virag and Szabo 2002).

5.1.6.2 pART2 / PARP 2

Als zweites humanes poly-ADP-ribosylierendes Protein wurde im Jahr 1999 die pART2 unter dem Namen PARP 2 publiziert (Ame *et al.* 1999). pART2 ist analog zur humanen pART1 im Nucleus lokalisiert, bindet direkt an Einzelstrang-DNA und besitzt eine PARP-Aktivität, die durch beschädigte DNA aktiviert wird. Auch pART2 ADP-ribosyliert verschiedene DNA-Reparatur-Proteine, ist direkt mit dem DNA-Reparatur-Komplex assoziiert und kann dort Homodimere oder Heterodimere mit der pART1 bilden (Schreiber *et al.* 2002), so dass eine wichtige Rolle in der DNA-Reparatur vermutet wurde. Dies konnte im Verlauf mit „knockout"-Experimenten bestätigt werden: *Parp 2* (-/-)-Mäuse weisen eine Hypersensitivität mit erhöhter genomischer Instabilität und verzögerter DNA-Reparatur gegenüber Alkylantien und Bestrahlung auf (Menissier de Murcia *et al.* 2003, Schreiber *et al.* 2002). Interessanterweise sterben *Parp 1* / *Parp 2*- Doppel-„knockout"-Mäuse bereits in der

frühen Embryonalphase, so dass die Expression zumindest eines der beiden Enzyme für die embryonale Entwicklung essentiell ist (Menissier de Murcia *et al.* 2003).

Die pART2 ist auch in andere zelluläre Prozesse involviert: So konnte kürzlich nachgewiesen werden, dass pART2 das Telomer-assoziierte Protein TRF2 bindet und dessen Funktion über ADP-Ribosylierung negativ reguliert (Dantzer *et al.* 2004). Desweiteren zeigten Experimente, dass Inhibitoren der pART2-Aktivität in einem Mausmodell mit chronischer Kolitis zu einer deutlichen Abschwächung der Entzündung mit Normalisierung der Kolonfunktion führen, so dass eine entscheidende Rolle der pART2 bei inflammatorischen Prozessen zu vermuten ist (Popoff *et al.* 2002). Zusammengefasst weisen die Ergebnisse sowohl auf überlappende Funktionen von pART1 und pART2 als auch auf einzelne, selektiv einem der beiden Proteine vorbehaltene Funktionen hin, wie dies bereits durch die Unterschiede in der Struktur der DNA-bindenden Domäne zu vermuten war.

5.1.6.3 pART3 / PARP 3

In dieser Arbeit konnte gezeigt werden, dass das humane pART3 ubiquitär exprimiert wird (Abb.21, 22). Im Northern Blot wurden in den meisten Geweben zwei Banden (2,2kb und 2,4kb) markiert, die vermuten ließen, dass es sich um alternative Transskripte oder Splicevarianten handelt. Interessanterweise kodiert die von Johansson *et al.* unter dem Namen PARP 3 publizierte cDNA-Sequenz für ein Protein, dem der N-terminale Anfang (7 Aminosäurenreste) unserer pART3 fehlt (Johansson 1999). Die Analyse der Genstruktur zeigt, dass direkt 3`-wärts des für diese Aminosäuren kodierenden Bereiches (Exon 1) ein Intron (645bp) liegt (Abb.31). Bei der von Johansson *et al.* publizierten Sequenz handelt es sich somit wahrscheinlich um eine verkürzte Splicevariante, die unserer 2,2kb-Bande entsprechen könnte. In einigen Geweben wird jeweils eine Variante deutlich schwächer exprimiert als die andere (z.B. Milz und Gehirn). Dies könnte auf eine veränderte, dem jeweiligen Gewebe angepasste Funktion der Splicevarianten hindeuten.

Trotz mehrfacher Versuche und unterschiedlichen Versuchsbedingungen konnte für die rekombinant exprimierte katalytische Domäne von hpART3 *in vitro* weder eine ART-Aktivität noch eine NADase-Aktivität nachgewiesen werden. Für die Inaktivität von pART3

in unseren Experimenten wurden verschiedene Erklärungen u.a. anhand von Erkenntnissen bei den bakteriellen ART-Toxinen diskutiert:

1) pART3 besitzt wie z.B. das Diptherietoxin (Wilson and Collier 1992) eine schmale Target-Spezifität, die mit unseren Assays nicht messbar ist.

2) pART3 entwickelt seine ART-Aktivität erst nach Bindung an einen Aktivator (Beispiele: Cholera Toxin und Exo S (Barbieri 2000, Rappuoli and Montecucco 1997)) oder benötigt die Interaktion mit seiner N-terminalen Domäne. Die Bedeutung der N-terminalen Domäne für die ART-Aktivität zeigt sich z.B. bei der humanen pART1, die unter normalen Bedingungen katalytisch weitgehend inaktiv ist und erst durch die Bindung an DNA-Strangbrüche aktiviert wird (Simonin *et al.* 1993, Smith 2001).

3) pART3 ist eine enzymatisch inaktive pART-Variante. Diese Möglichkeit erschien bei hoher Sequenzhomologie zur katalytischen Domäne der hpART1 und erhaltenem HYE-Motiv am unwahrscheinlichsten.

Bis auf eine strukturell abzuleitende potentielle DNA-Bindungskapazität blieb auch die Funktion der kleinen N-terminalen Domäne lange unklar. Augustin *et al.* gelang es schließlich, das vollständige pART3-Protein rekombinant zu exprimieren und *in vitro* eine Auto-ADP-Ribosylierungs-Aktivität nachzuweisen. pART3 konnte zudem in allen Phasen des Zellzyklus als Komponente des Zentrosoms identifiziert werden, wobei die N-terminale Domäne ein Motiv enthält, das für die bevorzugte Lokalisation des pART3 im Bereich der Tochter-Zentriole verantwortlich ist (Augustin *et al.* 2003). Desweiteren gibt es Hinweise, dass pART3 den Übergang aus der G1-Phase in die S-Phase des Zell-Zyklus negativ beeinflusst und am Zentrosom mit der pART1 interagiert (Augustin *et al.* 2003). Trotz dieser weiterführenden Ergebnisse bleibt die biologische Rolle der pART3 am Zentrosom bisher unklar.

Das *pART3*-Gen ist auf dem kurzen Arm des Chromosoms 3 (3p) lokalisiert. Der Verlust von Allelen auf dem 3p-Arm ist interessanterweise eine der am häufigsten und am frühesten auftretenden genetischen Veränderungen in der Pathogenese des Bronchialcarcinomes (Wistuba *et al.* 2000). Dabei ist in einem definierten kurzen Abschnitt der Region 3p21.3, in der auch *pART3* liegt, besonders häufig ein Allelverlust in den Bronchialepithelien von

Rauchern nachzuweisen. Bei weiteren Untersuchungen sollte somit eine mögliche Rolle des *pART3* in der Pathogenese des Bronchialcarcinoms berücksichtigt werden.

5.1.6.4 pART4 / vPARP

Auch das humane *pART4* wird ubiquitär exprimiert (Abb.21, 22). Die rekombinant produzierte katalytische Domäne des pART4 besitzt *in vitro* eine APD-Ribosylierungs-Aktivität. Sowohl durch die isolierte katalytische pART4-Domäne als auch durch die Trx-gekoppelte Variante wird bei den Enzymassays die Bande markiert, die der jeweiligen errechneten Größe des rekombinanten Proteins entspricht (Abb.28A, 29B). Aus diesen Beobachtungen wurde geschlossen, dass pART4 die Fähigkeit zur Auto-ADP-Ribosylierung besitzt. Die weitere Analyse des ADP-ribosylierten Trx-pART4 zeigte, dass die Radioaktivität im Gegensatz zum Trx-SpvB nicht im Bereich der assozierten Trx-Domäne, sondern vor allem in einem kleinen Spaltprodukt (deutlich < 14kD) des rekombinanten pART4 inkorporiert wurde (Abb.29B). Das Auftreten der Spaltprodukte könnte durch das Vorliegen einer internen Enterokinase-Schnittstelle erklärt werden, ein klassisches Enterokinasesignal ist allerdings in der Aminosäuresequenz des rekombinant exprimierten Anteils der pART4 (katalytische Domäne mit kurzem Abschnitt der alpha-helicalen Domäne, vgl. Anhang B) nicht nachzuweisen. Interessanterweise führt die Expression des pART4 in *E.coli* analog zu der rekombinant exprimierten katalytischen Domäne der humanen pART1 zu einer deutlichen gesteigerten ADP-Ribosylierung einer endogenen Bande bei ca. 60kD (Abb.28A). Dies weist darauf hin, dass rekombinantes pART4 die Fähigkeit besitzt, *E.Coli*-Zielprotein(e) mit einer Größe von ca.60kD zu ADP-ribosylieren oder zumindest eine verstärkte endogene ADP-Ribosylierung zu induzieren. Desweiteren besitzt pART4 eine NAD-Glykohydrolaseaktivität (Abb.28B).

Über seine Interaktion mit dem „major vault protein“ (MVP) wurde pART4 als Komponente großer Ribonukleoprotein-Komplexe, sog. vaults, identifiziert und deshalb als vault PARP (vPARP) publiziert (Kickhoefer *et al.* 1999). Neben der korrelierend mit unseren Ergebnissen bestehenden Fähigkeit zur Auto-ADP-Ribosylierung kann die pART4 MVP als Zielprotein ADP-ribosylieren. Die vaults, deren Funktion weitgehend unbekannt ist, bestehen aus einer Vielzahl von Kopien des MVP, dem hochkonservierten Telomerase-assoziierten Protein (TEP1), der pART4 und einer untranslatierten vault RNA (Kong *et al.* 2000). Es konnte

gezeigt werden, dass eine Fraktion der hauptsächlich im Zytoplasma lokalisierten vaults an der zytoplasmatischen Seite der Kernmembran lokalisiert ist und dass die vaults hohl sind (Chugani *et al.* 1993, Kong *et al.* 1999), so dass sie möglicherweise als Transporter fungieren. Die Beobachtungen, dass vaults hormonabhängig mit nukleären Estradiol-Rezeptoren interagieren (Abbondanza *et al.* 1998) und in sog. „multidrug"-resistenten Krebszellen hochreguliert werden (Schroeijers *et al.* 2000), unterstützen diese These. Möglicherweise könnte die pART4-vermittelte ADP-Ribosylierung des vorherrschenden strukturellen Proteins MVP zu einer Konformationsänderung der vaults führen und damit den Transport von Molekülen durch die vaults regulieren (Kickhoefer *et al.* 1999). Die Tatsachen, dass pART4 auch in anderen zellulären Bereichen wie z.B. im Kern und an der Mitosespindel detektiert werden konnte (Kickhoefer *et al.* 1999) und dass pART4 weitere Protein-Protein-Interaktionen vermittelnde Domänen besitzt (z.B. BRCT), weisen auf mögliche weitere Funktionen hin. Der pART4 wurde wie einigen anderen pARTs (z.B. den Tankyrasen, pART2 und pART1) auch eine Assoziation mit der Telomerase-Aktivität zugeschrieben. Neueste Untersuchungen an der pART4-„knockout"-Maus zeigen aber, dass die pART4 für die normale Entwicklung der Mäuse, die Telomerasefunktion, die Stabilität der Telomer-Länge und die Struktur der vaults *in vivo* entbehrlich zu sein scheint (Liu *et al.* 2004b).

5.1.6.5 Die Tankyrasen (pART5 / Tankyrase 1 und pART6 / Tankyrase 2)

pART5 wurde von einer anderen Arbeitsgruppe über seine Interaktion mit TRF1, einem negativen Regulator der Telomerlänge, identifiziert (van Steensel and de Lange 1997). Wegen seiner charakteristischen Ankyrin-Domäne, die u.a. das TRF1 bindet, wurde pART5 Tankyrase genannt (Cook *et al.* 2002, Smith *et al.* 1998). Die für die Stabilität der Chromosomen wichtigen Telomere werden durch die Telomerase, eine spezielle reverse Transkriptase, instand gehalten. Wichtige Regulatoren der Telomerstruktur und -funktion sind die telomer-spezifischen, DNA-bindenden Proteine TRF1 und TRF2, die über Konformationsänderungen den Zugang der Telomerase zu den Telomeren steuern (Smith 2001). Es konnte gezeigt werden, dass Tankyrase *in vitro* sich selbst und das assoziierte TRF1 ADP-ribosyliert. Dies hat eine Ablösung des TRF1 von den Telomeren und damit einen verstärkten Zugang für die Telomerase zur Folge (Smith *et al.* 1998). Die Überexpression von Tankyrase in humanen Zellen führte tatsächlich zur vermuteten Verlängerung der Telomere (Smith and de Lange 2000), so dass die Tankyrase *in vivo* einen

positiven Regulator der Telomer-Länge darstellt. Kürzlich konnte zudem gezeigt werden, dass die Tankyrase-induzierte Inhibition des TRF1 durch das Protein TIN-2 moduliert werden kann, indem es das TRF1 vor der Modifizierung durch die Tankyrase schützt (Ye and de Lange 2004). Die Beobachtung, dass während der Mitose Tankyrase 1 zur Teilung der Telomere der Schwesterchromatiden erforderlich ist, weist auf eine weitere Funktion der Tankyrase an den Telomeren hin (Dynek and Smith 2004).

Neben der Regulation der Telomerlänge ist die Tankyrase auch in andere zelluläre Prozesse involviert: So bindet und modifiziert Tankyrase z.B. die „insulin responsive aminopeptidase" (IRAP), ein Enzym, das in Golgi-assoziierten, sog. GLUT4-Vesikeln vorkommt, die die insulinabhängige Aufnahme von Glucose in das Zytoplasma beeinflussen (Chi and Lodish 2000). Ein weiterer Hinweis auf die Beteiligung der Tankyrase an Signaltransduktionsprozessen ist die Interaktion mit Mitgliedern der Mcl-1-Proteinfamilie (Bae *et al.* 2003). „Myeloid cell leukaemia-1" (Mcl-1)-Proteine weisen eine regulatorische Funktion bei der Apoptose auf, die durch die Überexpression der Tankyrase 1 antagonisiert wird. Desweiteren ist die Tankyrase ein Substrat der MAPK („mitogen-activated protein kinase"), wobei die Phosphorylierung eine gesteigerte PARP-Aktivität der Tankyrase zur Folge hat. Dies weist auf eine potentielle Rolle der Tankyrase als Signalmolekül beim vesikulären Transport hin (Chi and Lodish 2000). Für Tankyrase 1 wurde zudem gezeigt, dass sie über die SAM-Domäne *in vitro* und *in vivo* große Polymer-Komplexe (*in vitro* > 15 Monomere) ausbilden kann und dass die Auto-ADP-Ribosylierung zur Auflösung dieser Komplexe führt (De Rycker and Price 2004). Diese Fähigkeit, Proteine in einem großen Proteinkomplex zu organsieren und wieder aufzulösen, lässt vermuten, dass Tankyrase seine Wirkung in verschiedenen subzellulären Lokalisationen als sog „scaffolding" Protein entfaltet.

Tankyrase 2 (pART6), die bei fehlender HPS-Region über 80% Sequenzidentität zur Tankyrase 1 besitzt, wurde zunächst als ein neues, mit der Immunantwort assoziiertes Antigen bei Patienten mit Brustkrebs bzw. Meningiomen (Kuimov *et al.* 2001, Monz *et al.* 2001) entdeckt und zudem als Bindungspartner des Grb14 detektiert, einem kleinen Adaptorprotein, dem eine Funktion in Insulin-vermittelten Stoffwechselwegen zugeschrieben wird (Lyons *et al.* 2001). Tankyrase 2 zeichnet sich ebenfalls durch die Fähigkeit aus, TRF1 und IRAP zu binden und zu ADP-ribosylieren (Sbodio *et al.* 2002). Beide Tankyrasen sind in

ähnlichen Bereichen der Zelle lokalisiert und können über ihre SAM-Domänen Homo- und Heterodimere bilden (De Rycker *et al.* 2003, Sbodio *et al.* 2002). Dies lässt überlappende Funktionen der beiden Tankyrasen vermuten. Die Beobachtung, dass die Überexpression der Tankyrase 2 im Unterschied zur Tankyrase 1 nekrotischen Zelltod induziert (Kaminker *et al.* 2001), weist aber darauf hin, dass die jeweilige Tankyrase auch einzigartige Funktionen besitzt.

5.1.6.6 pART9 / BAL

Aguiar *et al.* haben im Jahr 2000 auf der Suche nach einer molekularen Basis zur Einschätzung des unterschiedlichen klinischen Risikos bei Patienten mit diffusen großzelligen B-Zell-Lymphomen das *BAL*-Gen (*BAL* = „B aggressive lymphoma“) identifiziert, das in Hochrisiko-Lymphomen deutlich stärker exprimiert wird als in Lymphomen mit niedrigem Risiko (Aguiar *et al.* 2000). Es gelang der Nachweis, dass die Überexpression von *BAL* zu einer gesteigerten B-Zell-Migrationsrate führt. Durch unsere Ergebnisse zeigt sich nun, dass BAL unserem pART9 entspricht und die C-terminale Region von BAL somit eine potentielle ADP-ribosylierende Domäne enthält. Die ADP-Ribosylierungsfunktion könnte folglich eine entscheidende Rolle in der Entwicklung der Aggressivität dieser B-Zell-Lymphome spielen. In einem „yeast two hybrid“-Screening wurde ein BAL-bindendes Protein, genannt BBAP („**B**-lymphoma- and **B**AL-**a**ssociated **p**rotein“) identifiziert, das C-terminal eine hohe Ähnlichkeit mit Mitgliedern der DTX (**De**l**tex**)-Familie aufweist und eine E3-Aktivität (Transfer von Ubiquitin auf ein Substrat) besitzt (Takeyama *et al.* 2003). Auch bei diesem Mitglied besteht somit die interessante Verbindung der potentiellen ART-Aktivität mit der Ubiquitinierung. Zudem besitzt pART9 möglicherweise über die beiden Makro-Domänen (A1pp) die Fähigkeit, Nukleinsäuren zu binden. Interessanterweise kommt auch die WWE-Domäne, die bei pART8 mit den A1pp-Domänen kombiniert ist, in Mitgliedern der DTX-Familie vor (Aravind 2001).

5.1.6.7 pART10 / PARP 10

Das Produkt des Protoonkogens c-Myc ist ein Regulator der Transkription, der u.a. bei der zellulären Proliferation, Differenzierung und Apoptose eine wichtige Rolle spielt. Die Myc-Proteine besitzen die Fähigkeit, Zellen zu transformieren, und sind bei verschiedenen humanen Neoplasien dereguliert. Auf der Suche nach assoziierten Partnern des c-Myc

identifizierten Yu *et al.* ein 150kD-Protein (PARP 10), das unserem pART10 entspricht (Yu *et al.* 2005). Die Myc-bindende Region liegt zwischen dem zweiten Ubiquitin-interagierenden Motiv und dem β1-Strang der katalytischen pART-Domäne (Abb.20). pART10 wird ubiquitär exprimiert mit der höchsten Expression in hämatopoetischen Geweben (v.a. Milz und Thymus) und ist hauptsächlich im Zytoplasma lokalisiert. In weiteren Versuchen konnte pART10 auch im Zellkern angefärbt werden, so dass vermutet wird, dass ein Transportmechanismus durch die Kernmembran besteht (Yu *et al.* 2005). Die PARP-Aktivität wurde überzeugend nachgewiesen: Neben der Fähigkeit zur Automodifikation ADP-riboysliert pART10 Core-Histone (v.a. H2A). Somit kann vermutet werden, dass das Protein an der Regulation des Chromatins beteiligt ist. Der Interaktionspartner c-Myc wird dagegen nicht ADP-ribosyliert. Desweiteren wurde gezeigt, dass pART10, unabhängig von der pART-Aktivität, in ausgewählten Zellinien (u.a. embryonalen Rattenfibroblasten) die Proliferation hemmt. Interessanterweise war bei Mutanten, die nicht aus dem Zellkern exportiert werden (fehlendes NES), die Proliferation ungestört (Yu *et al.* 2005). Möglicherweise entfaltet pART10 seine Wirkung über die Bindung von RNA und ss-DNA über das N-terminale RRM-Motiv.

5.1.6.8 pART13 / ZAP

Gao *et al.* zeigten in einer in *Science* publizierten Arbeit, dass die Expression eines Ratten-Proteins, das wie einige unserer pARTs einen Zink-Finger vom CCCH-Typ aufweist, zu einem Verlust retroviraler mRNA im Zytoplasma führt und somit die Virusreplikation inhibiert (Gao *et al.* 2002). Das Protein wurde „Zinc-finger Antiviral Protein“ (ZAP) genannt. ZAP bindet spezifische virale mRNAs über seine N-terminalen CCCH-Zink-Finger. Die speziellen Zink-Finger sind entscheidend für die antivirale Aktivität: Die Destruktion des zweiten bzw. des vierten Zink-Fingers durch positionsspezifische Mutagenese führt zu einem kompletten Verlust der antiviralen Aktivität (Guo *et al.* 2004). In einer weiteren Arbeit konnte gezeigt werden, dass ZAP auch die Replikation verschiedener Alphaviren inhibiert, indem es die Translation der eingeschleusten viralen RNA blockiert (Bick *et al.* 2003). Andere Viren, wie z.B. Herpes simplex Typ 1, die Polioviren oder das Gelbfiebervirus, können sich in ZAP-exprimierenden Zellen jedoch unverändert replizieren. ZAP ist vornehmlich im Zytoplasma lokalisiert, kann sich aber über einen CRM1-abhängigen Mechanismus zwischen Zytoplasma und Nukleus hin- und herbewegen (Liu *et al.* 2004a). Diese „Shuttle“-Fähigkeit

lässt vermuten, dass ZAP neben der Modulation der viralen mRNA-Level im Zytoplasma auch eine Funktion im Zellkern besitzt. Unsere Datenbankrecherchen ergaben, dass das ZAP-Protein C-terminal eine pART-Domäne besitzt und dass es sich um das Ratten-Orthologe von pART13 handelt. Somit ist zu postulieren, dass humanes pART13 und das in der Domänenstruktur analoge hpART12 ebenfalls eine antivirale Funktion besitzen. Die Aufklärung der molekularen Mechanismen, mit denen ZAP die Replikation der Viren hemmt, könnte zur Entwicklung von Substanzen mit einer breiten Wirkung gegen Retro- und Alphaviren führen.

5.1.6.9 pART 14 / TiPARP

Maus-TiPARP (pART14) wurde als mRNA identifiziert, deren Expression durch den toxischen Chlorkohlenwasserstoff 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) induziert wird (Ma *et al.* 2001). TiPARP wird in einem breiten Spektrum von Geweben (u.a. Herz, Gehirn, Lunge, Leber und Nieren) exprimiert und besitzt eine hohe Sequenzidentität zu Proteinen, die mit der Gedächtnisbildung bzw. der Progression bestimmter Tumoren in Verbindung stehen, so dass möglicherweise auch die TiPARP bei diesen Prozessen eine Rolle spielt (Ma *et al.* 2001). Zudem wurde gezeigt, dass das humane *TiPARP*-Gen zusammen mit dem benachbarten *CCNL* (*Cyclin L*)-Gen bei Plattenepithelkarzinomen im Kopf-Hals-Bereich hochreguliert wird (Katoh 2003).

Zusammenfassend lassen diese neueren Forschungsergebnisse vermuten, dass die humanen pARTs in vielen, auch für die Medizin relevanten Prozessen wie z.B. der Erhaltung der chromosomalen Stabilität, der DNA-Reparatur, dem Zelltod, dem vesikulären Transport, der Inflammation, der Entwicklung von Tumoren und den molekularen Mechanismen antiviraler Effekte eine entscheidende Rolle spielen. Neben der gemeinsamen subzellulären Lokalisation einiger pARTs gibt es weitere Hinweise, dass verschiedene pARTs miteinander interagieren können und somit zelluläre Prozesse durch Poly-ADP-Ribosylierung kooperativ regulieren. Die medizinische Relevanz führte in den letzten Jahren zu einer intensiven Erforschung von suffizienten pART-Inhibitoren zur Entwicklung therapeutischer Strategien, z.B. beim Schlaganfall. Auch das gezielte Design von spezifischen Inhibitoren für die pARTs wird in der Zukunft von Interesse sein. Dabei könnten die Ähnlichkeiten untereinander und auch die Multifunktionalität einzelner pARTs ein Problem darstellen, da durch unspezifische

Inhibitoren eine Störung von Abläufen, die man nicht beeinflussen will, wahrscheinlich erscheint. Z.B. könnten beim Schlaganfall zur Reduzierung des nekrotischen Gebietes oder bei entzündlichen Erkrankungen zur Abschwächung der Inflammation eingesetzte pART-Inhibitoren zu einer erhöhten genomischen Instabilität und damit zur Entstehung von Tumoren führen. Von der Aufklärung weiterer pART-Kristallstrukturen ist die Identifizierung von potentiellen spezifischen Ansatzpunkten für Inhibitoren zu erhoffen.

Die ausführliche Analyse der ESTs kann weitere Informationen über die gewebespezifische Expression und alternative Splicevarianten der bisher weniger bekannten *pART*-Gene ergeben. Das Vorkommen unterschiedlicher *pART*-Kombinationen in verschiedenen Spezies weist darauf hin, dass die Funktionen der *pART*s auf die Bedürfnisse der jeweiligen Spezies angepasst wurden. Die laufenden Genomprojekte bei verschiedenen Organismen werden in der Zukunft genauere Untersuchungen zu der *pART*-Ausstattung der einzelnen Spezies gestatten und zeigen, welche *pART*-Orthologen auch in entfernteren Spezies eine Rolle spielen und welche erst spät in der Evolution entstanden sind.

Eine interessante Frage zum Verständnis der molekularen Mechanismen wird zudem sein, wie die unterschiedlichen pARTs aktiviert werden. Die Beobachtung, dass in ruhenden Zellen kaum PARP-Aktivität nachzuweisen ist, weist darauf hin, dass die pARTs zumeist in einem inaktiven Zustand vorliegen und *in vivo* einer strikten Regulation unterstehen. Die Analyse der pART-katalysierten Modifikationen wird zudem durch den Mangel an effizienten Reagentien zum spezifischen Nachweis von ADP-ribosylierten Proteinen in Zellen kompliziert. Hinzu kommt vermutlich, dass der rasche Abbau der Polymere durch die PARG nur ein kurzes Zeitfenster für die Aktivierung und folgende Poly-ADP-Ribosylierung zulässt. Mono- und Oligo-ADP-Ribose ist wesentlich stabiler und besitzt eine Halbwertszeit von mehreren Stunden (D'Amours *et al.* 1999). Dies lässt vermuten, dass die pARTs insbesondere in der Kurzzeitregulation von Substraten eine entscheidende Rolle spielen. Ob die PARG selbst fähig ist, die Polymere der unterschiedlichen pARTs abzubauen oder ob Isoformen der PARG dafür verantwortlich sind, werden weitere Untersuchungen zeigen. Da die *PARG* bisher in allen entschlüsselten Genomen als Einzelkopiegen vorliegt, könnten mögliche Isoformen durch proteolytische Prozessierung des vollständigen Enzyms entstehen. Im Gegensatz zu der Vielzahl an Arbeiten zu den pARTs und zu den Effekten der Inhibition

der pART1, ist über die biologische Funktion der PARG bisher wenig bekannt. Dies liegt zum Teil an der Schwierigkeit, reines PARG-Enzym zu erhalten, denn die PARG kommt nur in geringen Mengen in den Zellen vor und ist während der Aufreinigung sehr empfindlich gegenüber proteolytischem Abbau (Davidovic *et al.* 2001). Die geringen Mengen des Enzyms werden durch die hohe Spezifität für poly-ADP-Ribose kompensiert. Es wurde gezeigt, dass mehr als 90% der PARG-Aktivität inaktiviert sein muß, damit es zur Akkumulation von poly-ADP-Ribose in den Zellen kommt (Jonsson *et al.* 1988). PARG weist strukturell sowohl ein nukleäres Lokalisationssignal (NLS) als auch ein nukleäres Exportsignal auf (NES), so dass eine „shuttle"-Fähigkeit zwischen Zellkern und Zytoplasma vermutet wird (Shimokawa *et al.* 1999). Dies würde den Abbau sowohl der zellulären als auch der nukleären poly-ADP-Ribose-Derivate durch dasselbe Enzym ermöglichen.

Es gibt paradoxe Hinweise, dass die Inhibition der PARG in oxydativ gestressten Zellen zu einem ähnlichen protektiven Effekt führt wie die Inhibition der pART1 (Ying *et al.* 2001). Eine mögliche Erklärung für diese Beobachtung ist, dass die Entfernung der Poly-ADP-Ribose von der automodifizierten pART1 durch die PARG erforderlich ist, um die pART1 wieder in den aktivierbaren Zustand zu versetzen (Virag and Szabo 2002). Um die Rolle der pARTs zu verstehen, ist auch eine weitere Aufklärung der Funktionen der PARG und damit des poly-ADP-Ribose-Zyklus in den Zellen erforderlich.

5.2 Eine neue bakterielle mono-ART: Das Salmonellen-Toxin SpvB

Die bakteriellen ADP-ribosylierenden Toxine wurden in den letzten Jahrzehnten intensiv erforscht und einige gehören zu den am besten verstandenen mikrobiellen Krankheitsmediatoren. Ein Charakteristikum dieser Toxin-Familie ist ihre Sequenz-Vielfältigkeit, die lange sowohl die Identifikation neuer Mitglieder als auch die Definition essentieller, charakterisierender Regionen erschwerte.

5.2.1 Ergebnisse der PSI-BLAST-Analysen in der mono-ART-Familie

Trotz teilweise nur sehr gering ausgeprägter Gesamtsequenzidentität (<10%) zwischen pro- und eukaryontischen ARTs konnten mit PSI-BLAST bekannte eukaryontische mono-ARTs mit den meisten bakteriellen mono-ARTs verknüpft werden. Die mit PSI-BLAST identifizierten, sequenzhomologen Regionen entsprechen den Bereichen, die bei den ARTs

mit bekannter 3D-Struktur entscheidende Sekundärstrukturmotive im Bereich des katalytischen Zentrums ausbilden. Dies untermauert die biologische Signifikanz der gefundenen Homologien.

Desweiteren zeigen die Ergebnisse unserer PSI-BLAST-Analyse der mono-ART-Familie weitere bemerkenswerte Details auf: Obwohl bei der Suche Familienmitglieder aus so verschiedenen Spezies wie Säugetieren und Bakteriophagen identifiziert wurden, konnten in den komplettierten Genomen der *Archaeabacteria*, der Modellpflanze *Arabidopsis thaliana,* der Hefe *Saccharomyces cerevisiae*, des Nematoden *Caenorhabditis elegans* oder der Fliege *Drosophila melanogaster* kein verwandtes Gen detektiert werden. Somit enthalten diese Genome entweder keine entsprechenden *mono-Art*-Gene, z.B. weil den Organismen diese genetische Information verloren gegangen ist, oder die Sequenz dieser Gene hat sich bei diesen Spezies so weit verändert, dass Homologien mit den uns zur Verfügung stehenden Computerprogrammen nicht aufzuspüren waren. Dass erstaunlicherweise mit verschiedenen mARTs als Input weder einige bakterielle Mitglieder der R-S-E-Familie (z.B. die Halovibrine von *Vibrio fisheri*, mod A und mod B des *E.coli-Phagen*) noch die Diphtamid-spezifischen Toxine Diphterietoxin und *Pseudomonas aeruginosa* Exotoxin A oder Mitglieder der pART-Familie detektiert werden konnten, obwohl die Tertiärstrukturen einen gemeinsamen Vorfahren vermuten lassen, ist ein möglicher Hinweis auf die letztgenannte Erklärung. Folglich könnten weitere „offene Leseraster“ in der Genom-Datenbank bisher unidentifizierte ARTs kodieren, und tatsächlich konnten mit PSI-BLAST-Analysen in bisher unvollständig entschlüsselten prokaryontischen Genomen weitere potentielle bakterielle ARTs identifiziert werden (Pallen *et al.* 2001).

Sollten die Archeaebakterien und niederen Eukaryonten über weiter entfernte mono-ARTs verfügen, stellt sich die Frage, warum die bakteriellen Toxine zu den Vertebraten-ARTs eine höhere Sequenzidentität aufweisen als zu den möglicherweise von diesen Spezies kodierten ARTs. Interessanterweise werden alle gefundenen prokaryontischen mono-ARTs durch mobile genetische Elemente wie z.B. Plasmide, Phagen oder sog. pathogenetische Inseln kodiert. Dies lässt vermuten, dass die mono-ARTs kodierenden Gene durch horizontalen Gentransfer weitergegeben wurden und zwar nicht nur im Reich der Eubakterien, sondern wahrscheinlich auch zwischen Pro- und Eukaryonten (z.B. von Bakterien zu den Vertebraten

bzw. Prävertebraten). Tatsächlich zeigte sich, dass die Genstruktur verschiedener Vertebraten-ARTs partiell für einen horizontalen Gentransfer geeignet ist, da die gesamte katalytische Domäne von einem einzelnen Exon und die N- und C-terminalen Signalpeptide von separaten Exons kodiert werden (Koch-Nolte *et al.* 1996a, Koch-Nolte *et al.* 1997).

Auch bei den ADP-Ribosyl-Hydrolasen (ARH), den Enzymen, die die übertragene ADP-Ribose von Zielproteinen abspalten, gelang mit PSI-BLAST die Verknüpfung von bakteriellen und humanen ARHs, während in den komplettierten Genomen der niederen Eukaryonten analog zu den mono-ARTs keine ARHs detektiert werden konnten (Glowacki *et al.* 2002). Bei der ARH-Familie ist es viel unwahrscheinlicher als bei den ARTs, dass verwandte „offene Leseraster“ lediglich durch PSI-BLAST nicht identifiziert wurden, da die Sequenzidentität zwischen den pro- und eukaryontischen Familienmitgliedern wesentlich höher ist (Glowacki *et al.* 2002). Bei verschiedenen Organismen weist das gleichzeitige Vorhandensein der genetischen Information für ARTs und ARHs auf endogene ADP-Ribosylierungszyklen hin, wie sie bei photosynthetischen Bakterien beschrieben wurden (Ludden 1994). Ein vergleichbares phylogenetisches Verteilungsmuster wie bei den ARHs und ARTs wird für 41 humane Gene beschrieben, die zumeist Enzyme kodieren (Andersson *et al.* 2001, Lander *et al.* 2001).

5.2.2 Identifizierung des SpvB

Neben der Verknüpfung der ARTs verschiedener, entfernter Spezies wurde bei der PSI-BLAST-Suche ein einzelnes offenes Leseraster bisher unbekannter Funktion entdeckt, das signifikante Ähnlichkeit zu bekannten mARTs aufweist. Dabei handelt es sich um das Proteinprodukt des Virulenz-assoziierten *spvB*-Gens von *Salmonella enterica* (Gulig and Doyle 1993). Zwischen dem Proteinprodukt SpvB und bekannten mono-ARTs besteht, bezogen auf die Gesamtsequenz, nur eine geringe Sequenzidentität. Zum nächsten Verwandten in der Datenbank, der katalytischen Untereinheit des VIP2-Toxins von *Bacillus cereus* (Han *et al.* 1999), besteht lediglich eine Sequenzidentität von < 20%. Die Signifikanz unterstreichend ist die Sequenz-Homologie im Bereich der Regionen am höchsten ausgeprägt, die in der Kristallstruktur von VIP2 die wichtigen Sekundärstrukturmotive im katalytischen Zentrum ausbilden (Abb.35). Das charakteristische R-S-EXE-Aminosäurenmotiv der Arginin-spezifischen mono-ARTs ist auch beim SpvB konserviert.

Inzwischen konnte auch die erste Kristallstrukur einer Säugetier-mono-ART, dem ART2.2 der Ratte, aufgeklärt werden (Mueller-Dieckmann *et al.* 2002). Dabei zeigte sich, dass die katalytischen Domänen von zwei bakteriellen ARTs mit bekannter Kristallstruktur, VIP2 und C3, strukturell mit dem rART2.2 besonders nah verwandt sind, während die Struktur der anderen eher der eukaryontischen pART1 ähnelt (vgl. Abb.5). Dies ist ein Hinweis auf das Bestehen von zwei strukturellen Subfamilien der bakteriellen ARTs. Da das VIP2 bzgl. der Sequenz den nächsten Verwandten des SpvB in der Datenbank darstellt, lässt sich vermuten, dass das SpvB strukturell der rART2.2-verwandten Gruppe zuzuordnen ist.

Bei der *S.enterica* Subspezies I sind die *spv*-Gene auf einem selbst-übertragbaren Plasmid lokalisiert (Ahmer *et al.* 1999, Gulig *et al.* 1993). Dies weist darauf hin, dass das SpvB-Protein, wie andere bakterielle ADP-ribosylierende Toxine, durch ein mobiles genetisches Element kodiert wird. Tatsächlich lassen eine Analyse der phylogenetischen Verbreitung des *spv*-Operons in *Salmonella*-Isolaten sowie vergleichende Analysen der Nukleotidsequenz vermuten, dass die *spv*-Region erst sehr spät in der Evolution, nach der Spezifizierung der Salmonellen erworben wurde (Boyd and Hartl 1998). In Experimenten zeigte sich für das Virulenzplamid eine vergleichbare Konjugationsfrequenz wie für andere natürliche selbst-übertragbare Plasmide. Die Konjugationsfrequenz ist damit deutlich geringer als die Übertragungsfrequenz des F-Plasmids von *E.Coli* (Ahmer *et al.* 1999).

5.2.3 Die Subdomänen von SpvB

Zunehmende Aufmerksamkeit in der Analyse von Genomen erhält die Erkenntnis, dass Informationen für auf den ersten Blick unabhängig erscheinende, aber funktionell interagierende Proteine gelegentlich genetisch fusioniert sind und dass Homologien in der Aminosäuresequenz Schlüssel zur Aufklärung der Funktion von Subdomänen darstellen (Marcotte *et al.* 1999). Dies trifft im Besonderen auch auf die bakterielle *mART*-Genfamilie zu, in der die katalytische, ADP-ribosylierende Domäne mit einem sehr variablen Spektrum von regulatorischen Domänen assoziiert ist. Bei der in dieser Arbeit charakterisierten mART SpvB wird die ART-katalytische Domäne durch einen Strang hintereinander angeordneter Prolin-Aminosäurenreste von der TcaC-verwandten N-terminalen Domäne abgegrenzt, wobei das Polyprolin-Stück als starre Verbindung zwischen den beiden Domänen dient. Der Fusionspartner TcaC wurde als Teil des Toxin-Komplexes beschrieben, der von

Photorhabdus luminescence sezerniert wird, einem Bakterium, dass im Darm Insekten infizierender Nematoden lebt. Zwar ist die präzise Funktion von TcaC bisher ungeklärt und die Entdeckung der Sequenzhomologie konnte somit zunächst nur wenig zur Aufklärung der Funktion von SpvB beitragen, aber der Nachweis dieses Toxin-Anteils im Überstand von Kulturen weist auf eine wichtige Rolle in der Interaktion zwischen dem Wirt und seiner Umwelt hin (ffrench-Constant and Bowen 1999).

5.2.4 SpvB besitzt eine ARTase-Funktion

Der Versuch anderer Arbeitsgruppen das SpvB-Protein mit seinen beiden fusionierten Anteilen in *E.Coli* zu exprimieren, führte zunächst nur zu einem unlöslichen Proteinprodukt (Taira *et al.* 1991a). Deshalb entwickelten wir Strategien, um die vorhergesagte C-terminale ART-Domäne isoliert in löslicher Form zu exprimieren. Wir produzierten verschiedene rekombinante SpvB-Varianten und wiesen in Enzymassays die vorhergesagte ADP-Ribosyl-Transferase-Aktivität für die isolierte katalytische Domäne sowie für ein Trx-SpvB-Fusionsprotein *in vitro* nach (Abb.38-40). Die Beobachtung, dass die katalytische Domäne von SpvB in dem von uns konstruierten Trx-SpvB-Fusionsprotein die ADP-Ribosylierung der Trx-Subdomäne katalysiert, weist auf die Möglichkeit hin, dass die katalytische Domäne im nativen SpvB die N-terminale TcaC-Domäne ADP-ribosyliert und somit deren Funktion verändert bzw. sogar reguliert. Unter Berücksichtigung dieser Erkenntnis wurden von der kooperierenden Arbeitsgruppe von M.Rhen am Karolinska Institut in Stockholm weitere Versuche mit nativem SpvB durchgeführt. Die Inkubation von *Salmonella*-Extrakten bzw. des Überstandes aus *Salmonella*-Kulturen mit [^{32}P]-NAD^{+} führte jedoch nicht zur radioaktiven Markierung eines Proteins mit der erwarteten Größe (65kD) von nativem SpvB (Tezcan-Merdol *et al.* 2001). Deshalb hielten wir die Annahme für wahrscheinlicher, dass es sich bei dem physiologischen Zielmolekül von SpvB nicht um die assoziierte TcaC-Subdomäne, sondern vielmehr um ein Protein der Wirtszelle handelt.

5.2.5 SpvB modifiziert ein wichtiges Wirtszellenprotein

Auf der Suche nach diesem Wirtszellenprotein wurden weitere *in vitro*-Experimente mit dem rekombinanten SpvB durchgeführt. Die Inkubation von Jurkat-Zell-Lysaten mit der aufgereinigten, rekombinanten SpvB-C-Domäne hatte als Bestätigung unserer Annahme tatsächlich die effiziente ADP-Ribosylierung eines zellulären 40kD-Zielproteins zur Folge

(Abb.41). Obwohl nicht sicher auszuschließen ist, dass die katalytische SpvB-Domäne lediglich eine endogene ADP-Ribosyl-Transferase-Aktivität in Gang setzt, ist es wahrscheinlicher, dass die Markierung des endogenen Zielproteins durch SpvB selbst katalysiert wird. Die Tatsache, dass die ADP-Ribosylierung der 40kD-Bande durch SpvB kein für die Jurkat-Zellen spezifisches Phänomen ist, sondern auch in Lysaten anderer humaner Zelllinien erfolgte, wies darauf hin, dass es sich um ein ubiquitär verbreitetes Protein handeln muss. Fraglich war zudem, ob es sich bei dieser Markierung lediglich um ein *in vitro*-Phänomen des rekombinanten SpvB ohne biologische Signifikanz handelt oder ob auch das native SpvB in Infektionsmodellen ein entsprechendes Zielprotein modifiziert. Tatsächlich wurde sowohl in Makrophagen-Lysaten, die mit rekombinantem SpvB behandelt wurden, als auch in Lysaten von Makrophagen, die zuvor mit *Salmonella enterica* infiziert wurden, eine stark markierte Bande mit korrelierender Größe nachgewiesen. Im Gegensatz dazu blieb die Markierung in Lysaten von Makrophagen, die mit einem *spvB*-mutierten *S.enterica*-Stamm infiziert worden waren, vollständig aus (Tezcan-Merdol *et al.* 2001).

Salmonellen werden von Makrophagen phagozytiert und können in den endosomalen Vesikeln proliferieren. Aufgrund der beschriebenen Erkenntnisse konnte postuliert werden, dass SpvB durch die ADP-Ribosylierung eines wichtigen Makrophagenproteins eine entscheidende Rolle in der Etablierung der Infektion spielt und dass dieser Vorgang zu einer Unterstützung der intrazellulären Vermehrung der Salmonellen führt. Diese Annahme liegt im Einklang mit früheren detaillierten Untersuchungen, die sowohl genetische Techniken als auch murine Infektionsmodelle einschließen und die zeigten, dass die *spv*-Gene innerhalb von Wirtszellen verstärkt transkribiert werden (Fierer *et al.* 1993, Rhen *et al.* 1993) und dass diese Gene die Vermehrung der Salmonellen in den Zellen des retikuloendothelialen Systems der Mäuse erleichtern (Gulig and Doyle 1993). Innerhalb des murinen retikuloendothelialen Systems dienen dabei insbesondere die Makrophagen als wichtigster Ort der *spv*-getriggerten intrazellulären Vermehrung der Salmonellen (Gulig *et al.* 1998).

5.2.6 Das Zielprotein der SpvB-katalysierten ADP-Ribosylierung ist Aktin

Mit dem Ziel der Identifizierung des ca. 40kD umfassenden Zielproteins wurden weitere Experimente durchgeführt. Verschiedene Hinweise implizierten, dass es sich bei dem gesuchten Zielmolekül um Aktin handelt: Zum einen zeigte sich beim Versuch die bakteriellen

ARTs mit Hilfe der Ergebnisse unserer PSI-BLAST-Analyse in Subfamilien zu gliedern, dass SpvB am ehesten der Gruppe der Aktin-ADP-ribosylierenden Toxine (wie z.B. *Bacillus cereus* VIP2-Toxin, *Clostridium botulinum* C2-Toxin und *Clostridium perfringens* Iota-Toxin) zugeordnet werden konnte. Da die SpvB-vermittelte Markierung einer 40kD-Bande auch in den Lysaten anderer humaner Zellinien in ähnlicher Intensität erfolgte, musste es sich um ein ubiquitär verbreitetes Protein handeln. Desweiteren entsprach das molekulare Gewicht des SpvB-Zielproteins in Jurkat-Zellen ungefähr dem Molekulargewicht von Aktin. In dieser Arbeit konnte schließlich experimentell nachgewiesen werden, dass Aktin *in vitro* ein Zielprotein der SpvB-katalysierten ADP-Ribosylierung darstellt.

Diese Ergebnisse konnten im Verlauf von der kooperierenden Arbeitsgruppe von M.Rehn bestätigt und in ein Infektionsmodell übertragen werden. Mittels Immunopräzipitation und Immunoblotting-Techniken wurde das in den Lysaten von *S.enterica*-infizierten Makrophagen markierte Protein als Aktin identifiziert. Desweiteren zeigte sich, dass die isolierte katalytische Domäne von SpvB *in vitro* den Transfer der ADP-Ribose auf verschiedene Aktine (z.B. Muskelaktin von Rindern und Kaninchen sowie Herzmuskelaktin von Schweinen) katalysiert, während andere eukaryontische Kontrollproteine bzw. bakterielle Proteine nicht modifiziert wurden. Es handelt sich somit um eine spezifische SpvB-vermittelte ADP-Ribosylierung des Aktins (Tezcan-Merdol *et al.* 2001).

5.2.7 SpvB fördert die Depolymerisierung von Aktin

Weiter stellte sich nun die Frage, inwiefern die SpvB-vermittelte ADP-Ribosylierung des Aktins sich auf die Struktur des Aktins auswirkt. *In vitro*-Polymerisations-Assays zeigten, dass die Anwesenheit der katalytischen SpvB-Domäne die Konversion von globulärem G-Aktin in filamentöses F-Aktin verhindert (Tezcan-Merdol *et al.* 2001). Für andere bakterielle Toxine, wie z.B. das C2-Toxin von *Clostridium difficile*, war bereits bekannt, dass die toxin-vermittelte ADP-Ribosylierung des Aktins die Aktin-Polymerisation beeinflusst (Aktories and Wegner 1992) und dass die C2-Toxin-Exposition in Zellkulturen zur Disintegration des Aktin-Zytoskeletts und damit zu definierten morphologischen Veränderungen führt (Barth *et al.* 1998b). Andere Toxine wie z.B. das *Clostridium botulinum* Exoenzym C3 ADP-ribosylieren GTP-bindende Proteine der Rho-Familie und beeinflussen damit indirekt die Aktin-Polymerisierung (Aktories 1994, Wiegers *et al.* 1991). Kürzlich konnte auch für das

SpvB gezeigt werden, dass es einen Verlust des polymerisierten Aktins in epithelialen MDCK-Zellen induziert und zytopathische Effekte, wie z.B. das Abrunden und Ablösen der Zellen in der Zellkultur, auslöst, während Zellen, die mit einem *spvB*-deletierten Salmonellenstamm infiziert worden waren, keine morphologischen Veränderungen aufwiesen (Tezcan-Merdol *et al.* 2001). Der beschriebene Einfluss des SpvB auf die Aktin-Polymerisierung wurde durch Ergebnisse von Lesnick und Mitarbeitern untermauert, die im Jahre 2001 publizierten, dass die Expression von SpvB in CHO-Zellen zu einer Depolymerisierung des Aktins führt (Lesnick *et al.* 2001).

5.2.8 SpvB induziert Zytotoxizität und Apoptose in den Wirtszellen

Die Infektion von humanen Makrophagen mit *S.typhimurium* oder *S.dublin in vitro* führte, analog zu den Beobachtungen bei den epithelialen MDCK-Zellen, zum Ablösen der Zellen als Ausdruck der Zytopathologie. Dabei enthielten die abgelösten Zellen eine hohe Anzahl an proliferierenden Bakterien. Es konnte nachgewiesen werden, dass die zytopathischen Effekte von den *spv*-Genen abhängen und dass inaktivierende Mutationen im transkriptionellen Aktivator *spvR* oder im *spvB* selbst zu einer deutlichen Reduktion des zytopathischen Effekts führen. Die beobachteten zytopathischen Veränderungen wiesen dabei mit dem Austreten der Nukleosomen in das Zytoplasma, Kondensation des Zellkerns und DNA-Fragmentierung mehrere Charakteristika der Apoptose auf (Libby *et al.* 2000). Ungeklärt blieb aber, ob die *spv*-Gene direkt den zytopathischen Effekt auslösen oder ob die Zytopathologie eine Konsequenz der massiven intrazellulären Bakterienproliferation darstellt. In einer kürzlich veröffentlichten Arbeit zeigte sich, dass die intrazelluläre Expression des isolierten SpvB-Proteins in eukaryontischen Zellen den apoptotischen Zelltod induziert (Kurita *et al.* 2003). Die Phagozytose von infizierten, apoptotischen Zellen durch nicht befallene Makrophagen könnte möglicherweise *in vivo* zur Ausbreitung von Zelle zu Zelle führen und damit zur Etablierung der systemischen Infektion beitragen. Obwohl die Signalwege der SpvB-vermittelten Apoptose noch nicht hinreichend aufgeklärt sind, kann zumindest postuliert werden, dass SpvB eine entscheidende Bedeutung bei der Induktion der Zytotoxizität besitzt.

Um nachzuweisen, dass es sich bei den beschriebenen SpvB-Effekten in eukaryontischen Zellen nicht nur um einen strukturellen Effekt des SpvB handelt, sondern dass die ADP-

Ribosyltransferase-Aktivität des SpvB für diese Effekte essentiell ist, wurden auch von anderen Arbeitsgruppen durch gezielte Mutationen im aktiven Zentrum der ART-Domäne inaktive SpvB-Mutanten konstruiert und produziert. Mit Hilfe dieser ART-inaktiven Mutanten konnte im Vergleich mit dem nativen SpvB gezeigt werden, dass die ADP-Ribosylierungs-Aktivität von SpvB für die verzögerte Zytotoxizität und den Verlust an polymerisiertem Aktin in mit Salmonellen infizierten humanen Makrophagen verantwortlich (Browne *et al.* 2002) und für die Virulenz der Salmonellen in Mäusen essentiell ist (Lesnick *et al.* 2001).

Interessanterweise unterscheidet sich die von Libby *et al.* bei humanen Makrophagen beobachtete Zytopathologie signifikant von der zuvor beschriebenen „**s**almonella **p**athogenicity **i**sland“ (SPI)-1-abhängigen Zytopathologie in murinen Makrophagen nach Infektion mit *S.typhimurium* (Chen *et al.* 1996, Monack *et al.* 1996): Während sich die SPI-1-vermittelte Zytotoxizität bereits früh (3-4h) nach der Infektion der murinen Makrophagen manifestiert und zur schnellen Apoptose führt, konnte bei den Experimenten mit den humanen Makrophagen und opsonierten Bakterien keine signifikante „frühe“ Zytotoxizität beobachtet werden; vielmehr zeigte sich der spv-vermittelte Effekt erst verspätet nach ca. 10h und ist charakterisiert durch das Ablösen der Zellen bei zumeist intakter Membranpermeabilität (Libby *et al.* 2000). Dieser „verspätete“ zytopathische Effekt ähnelt der „verspäteten“, in murinen Makrophagen beobachteten Apoptose (van der Velden *et al.* 2000), für die eine Abhängigkeit vom SPI-2-Locus nachgewiesen werden konnte (Lesnick and Guiney 2001). Diese Ergebnisse korrelieren mit verschiedenen Hinweisen, dass das SpvB-Protein erst später im Infektionsprozess, nachdem Invasion und Phagozytose stattgefunden haben, exprimiert wird. Die Tatsache, dass die *spv*-Gene mit einer zentralen Rolle des alternativen Sigma-Faktors RpoS durch einen ganz anderen Mechanismus reguliert werden als die Gene des SPI-1 (Fang *et al.* 1992), unterstützt die These, dass es sich um unterschiedliche Prozesse handelt, die zu verschiedenen Zeitpunkten des Infektionszyklus ablaufen.

5.2.9 Der Export-Mechanismus für das SpvB ist bisher ungeklärt

Da zu beobachten war, dass bei Zusatz von SpvB in das Medium keine Veränderungen des Aktinskeletts der Kulturzellen auftraten (Lesnick *et al.* 2001), ist vermutlich ein aktiver Transport des SpvB in die Wirtszelle durch ein spezialisiertes System erforderlich, damit

SpvB seine Effekte entfalten kann. Ein geeignetes Transportsystem wäre das bereits ausführlich charakterisierte Typ-III-Sekretionssystem (TTSS) (Kubori *et al.* 1998), da dieses System wie eine „molekulare Spritze“ einen Transport durch beide Membranen ermöglicht. Das Chromosom von *Salmonella typhimurium* enthält mindestens 5 „pathogenetische Inseln“ (SPI), die dem Stamm spezifische Virulenzen verleihen und die möglicherweise durch horizontalen Gentransfer erworben wurden (Marcus *et al.* 2000). Zwei dieser pathogenetischen Inseln, SPI-1 und SPI-2, kodieren u.a. für ein TTSS. Das von SPI-1 kodierte TTSS exportiert Proteine, die bei der Invasion des Wirtes eine wichtige Rolle spielen (z.B. SipA, SopE, SipC). Von SPI-1 wird desweiteren auch das SipB-Protein kodiert, das Caspase 1 aktiviert und damit die schnelle Apoptose einleitet (Hersh *et al.* 1999). Die SPI-1 vermittelten Prozesse sind für die systemische Virulenz in Mäusen nicht erforderlich (Galan and Curtiss 1989). Das von SPI-2 kodierte TTSS ist notwendig für die systemische Ausbreitung der Infektion in Mäusen sowie für das intrazelluläre Überleben und die Proliferation der Salmonellen in murinen Makrophagen (Hensel *et al.* 1997, Ochman *et al.* 1996). Ein bekanntes über dieses TTSS (auch als Spi/Ssa-System bezeichnet) exportiertes, SPI-2 kodiertes Protein ist SpiC, das in den intrazellulären Verkehr der Lyso- und Endosomen eingreift, indem es z.B. Fusionen von Endosomen hemmt (Uchiya *et al.* 1999).

Zur Frage der Beteiligung der Transportsysteme bzw. des SipB an der „verspäteten“ SpvB-vermittelten Zytotoxizität in humanen Makrophagen wurden von Browne und Mitarbeitern Versuche mit mutanten Stämmen durchgeführt, bei denen jeweils für die Struktur der TTSS erforderliche Gene des SPI-1-, des SPI-2-Locus oder das Gen *sipB* mutiert und damit inaktiviert wurden: Während die *sipB*- und SPI-1-Mutanten eine vergleichbare Zytotoxizität wie der Wildtyp erzeugten, zeigte sich, dass das SPI-2 TTSS für die SpvB-abhängige Zytotoxizität absolut erforderlich ist (Browne *et al.* 2002). Es bestehen folglich zwei Möglichkeiten: Entweder das SpvB wird selbst über das SPI-2 TTSS transportiert oder die intrazellulären Effekte des SpvB sind von anderen, über das SPI-2 TTSS transportierten Effektorproteinen abhängig.

Da die spontane Sekretion oder Freisetzung von SpvB unter *in vitro* Wachstumsbedigungen bisher nicht beobachtet werden konnte, führte eine japanische Arbeitsgruppe weitere *in vitro*- und *in vivo*-Analysen zur Sekretion des SpvB durch (Gotoh *et al.* 2003). Zum einen konnte

gezeigt werden, dass „Wildtyp“-Salmonellenstämme das SpvB bei Wachstum in einem, die intrazellulären Salzverhältnisse von Eukaryonten imitierenden Medium („intracellular salts medium, ISM“) in den Überstand sezernieren, wobei ein Teil der N-terminalen Region des SpvB (AS 1-229) erforderlich ist, um eine effektive Sekretion zu gewährleisten. Die N-terminale Domäne enthält somit ein wichtiges Transportsignal für die Auschleusung des SpvB in das Zytoplasma. Die Sekretion von SpvB beim Wachstum unter „intrazellulären“ Umweltbedingungen lässt vermuten, dass SpvB erst sezerniert wird, wenn die Bakterien sich in der Wirtszelle befinden.

Interessanterweise konnte im strikten Gegensatz zu den Ergebnissen von Browne *et al.* auch ein *S.typhimurium*-Stamm mit SIP-2-Mutation die gleiche Menge an SpvB sezernieren und zytotoxische Effekte in infizierten MDCK-Zellen erzeugen, so dass SpvB nach diesen Ergebnissen unabhängig von den Typ-III-Sekretionssytemen in das Zytoplasma transportiert zu werden scheint (Gotoh *et al.* 2003). Als alternatives Transportmodell wird stattdessen das Typ-II-Sekretionssystem diskutiert. Die Sekretion setzt sich dabei aus zwei Schritten zusammen: Nachdem die Proteine mittels Signalpeptiden, die mit dem Sec-Transport-System interagieren, die Zytoplasmamembran passiert haben (Pugsley 1993), falten sich die Proteine in ihre ausgereifte Konfiguration. Das Typ II-Sekretionssystem erkennt spezifisch diese gefalteten Proteine und erleichtert ihre Ausschleusung, vermutlich durch „große Poren“ in der äußeren Membran. Über diesen Mechanismus werden z.B. auch andere ADP-ribosylierende Toxine wie das Choleratoxin von *Vibrio cholerae* (Sandkvist *et al.* 1997) und das *Pseudomonas aeruginosa* Exotoxin A exportiert (Bally *et al.* 1992). Der Sekretionsmechanismus des SpvB ist Gegenstand unterschiedlicher Spekulationen, eindeutige überzeugende Ergebnisse stehen aber noch aus.

Zusammenfassend demonstrieren die hier präsentierten Ergebnisse die Nützlichkeit des PSI-BLAST-Programmes zur Identifizierung neuer, entfernter Verwandter mit biologisch signifikanten Sequenzhomologien in den Datenbanken. Der Nachweis, dass das „offene Leseraster“ von SpvB eine ADP-Ribosyltransferase-Domäne kodiert, die an einen möglicherweise sezernierten Toxin-Komplex fusioniert ist, und in eukaryontischen Zellen Aktindepolymerisierung, „verspätete“ Zytotoxizität und Apoptose induziert, eröffnet spannende neue Forschungsansätze zum Verständnis der Pathogenese von Salmonellen und

somit gegebenenfalls mittelfristig neue Aussichten in der Therapie bzw. Prophylaxe humaner Salmonellosen. Bis kürzlich wurden ADP-ribosylierende Toxine lediglich in humanpathogenen Bakterien gefunden, die extrazelluläre Infektionen verursachen. SpvB wäre das erste Beispiel einer bakteriellen ADP-Ribosylierung von Aktin, die mit dem intrazellulären Wachstum des Pathogens assoziiert ist. Die Eigenschaft, den Darm zu verlassen und eine extraintestinale Infektion zu etablieren, wird durch den *spv*-Locus vermittelt. Die Abwesenheit der *spv*-Gene bei S*almonella typhi* ist ein Hinweis darauf, dass die Pathogenese des Typhus sich fundamental von der Bakteriämie durch nichttyphöse Salmonellen unterscheidet (Fierer and Guiney 2001). Die Aquisition von SPI-2 ermöglicht den virulenten Salmonellenstämmen das Eindringen in die Endo- bzw. Phagosomen. Es kann postuliert werden, dass dort das SpvB die intrazelluläre Vermehrung der Salmonellen durch Interaktion mit der zytoskelettalen Funktion des Aktins fördert. Interessanterweise entwickeln auch einige der Proteine, die über das SPI-1 TTSS in die Wirtszelle injiziert werden (z.B. SipA), ihren Effekt über die Interaktion mit Aktin. Die unterschiedliche Regulation und Expression der verschiedenen Effektorproteine scheint eine sequentiell angepasste Modifikation des Aktinzytoskeletts während des Infektionszyklus zu ermöglichen: Zunächst schaffen die SPI-1 assoziierten Faktoren die Voraussetzungen für die Aufnahme der Bakterien in die Wirtszelle, anschliessend fördert SpvB deren intrazelluläre Vermehrung und induziert die Zytotoxizität. Die Interaktion des SpvB mit Aktin könnte Veränderungen im Bereich der Antigenpräsentation, des vesikulären Verkehrs oder der Reifung der Phagolysosomen zur Folge haben und damit einen wichtigen Baustein bei der Etablierung der Infektion darstellen. Die Sekretion eines ART-Toxins durch die Salmonellen aus ihrer intrazellulären Nische im Phagososom heraus, um den Makrophagen kampfunfähig zu machen bzw. sich vor dem wirtseigenen Immunsystem zu verstecken, wäre eine einzigartige Virulenz-Strategie. Diese Zusammenhänge, die molekulare Charakterisierung des SpvB-Transport-Prozesses sowie die molekularen Mechanismen der SpvB-vermittelten Apoptose werden in Zukunft Gegenstand intensiver Forschung sein. Das in dieser Arbeit beschriebene rekombinante SpvB-Protein und seine Mutanten könnten bei diesen Experimenten nützliche Werkzeuge darstellen. Desweiteren erleichtern rekombinante ARTs die Entwicklung von spezifischen Antikörpern und Inhibitoren, die möglicherweise für neue therapeutische Strategien genutzt werden können.

6 Zusammenfassung

ADP-Ribosylierung ist eine posttranslationale Proteinmodifikation, bei der der ADP-Ribose-Anteil von NAD^+ auf spezifische Aminosäurereste von Zielproteinen übertragen wird. ADP-Ribosyltransferasen (ARTs) werden in mono-ARTs und poly-ARTs unterteilt. Ziel der vorliegenden Arbeit war die Identifizierung und molekulare Charakterisierung neuer ART-Genfamilienmitglieder. Mit der Aminosäuresequenz der katalytischen Domäne bekannter ARTs als Input wurden mittels tBLASTn in dbEST 4 humane mono-ART- und 17 poly-ART-kodierende EST-Cluster identifiziert (ART1, 3, 4, 5; pART 1-17). Die höchste Sequenzidentität besteht im Bereich der Sekundärstrukturmotive, die in den aufgeklärten Kristallstrukturen bakterieller ART-Toxine und der Hühner-PARP das katalytische Zentrum bilden. Nur sechs humane poly-ARTs besitzen die für ARTs charakteristische katalytische Glutaminsäure. Mittels PSI-BLAST wurden Homologe der pART-Familie in Wirbeltieren, Insekten, Nematoden, Pflanzen sowie in zwei DNA-Viren nachgewiesen, während in den durchsequenzierten Genomen der Hefe sowie der Archaebakterien keine pART-Homologen detektiert werden konnten. Die durchsequenzierten Genome von Maus und Ratte enthalten Orthologe aller humanen pARTs, mit Ausnahme von pART7. Bei den mono-ARTs wurden dagegen Mitglieder in verschiedenen bakteriellen Genomen, nicht aber in den komplettierten Genomen von Insekten, Nematoden, Pflanzen und Hefe identifiziert.

Die eukaryontischen poly- und mono-ARTs zeigen deutliche Unterschiede in ihrer Gen- und Domänenstruktur: Während die katalytischen Domänen der mono-ARTs durch ein einzelnes großes Exon kodiert werden, wird die katalytische Region der meisten *pART*-Gene durch ein odere mehrere Introns in kleinere Exons aufgesplittert. Die mARTs bestehen lediglich aus einer isolierten katalytischen Domäne. Im Gegensatz dazu sind die katalytischen Domänen der pARTs mit unterschiedlichsten funktionellen, teilweise regulatorischen Domänen verknüpft. Nach dem Grad der Sequenzidenität, konservierten Introneinschüben und assoziierten Proteindomänen wurden die humanen pARTs in fünf Subgruppen unterteilt.

Zwei humane pARTs, pART3 und pART4, wurden molekular genauer charakterisiert. Beide pARTs werden ubiquitär exprimiert. Das *pART3*-Gen besteht aus 11 Exons und kodiert ein Protein von 540 Aminosäuren; das *pART4*-Gen besteht aus 34 Exons und kodiert ein Protein von 1724 AS. Die katalytischen Domänen der beiden pARTs wurden in *E.Coli* rekombinant exprimiert. Für pART4 gelang der Nachweis der vorhergesagten ADP-Ribosyl-Transferase-Aktivität.

Desweiteren wurde eine neu entdeckte, bakterielle Toxin-ART - das auf dem Virulenzplasmid von Salmonellen kodierte SpvB - molekular charakterisiert. Nach rekombinanter Expression und Aufreinigung der katalytischen Domäne von SpvB gelang der Nachweis der vorhergesagten ART-Aktivität. In humanen Zellen ADP-ribosyliert SpvB ein 40kD Zielprotein, das als Aktin identifiziert wurde.

Die hier vorgestellten Ergebnisse bilden die Basis für die künftige Klärung der Funktionen von mono-ARTs und poly-ARTs bei der Wirt-Pathogen-Interaktion (SpvB) sowie bei zellulären Prozessen wie der DNA-Reparatur, dem Zelltod, der Regulation der Telomerlänge, dem zellulären Transport, der Inflammation, der Entwicklung von Tumoren und der antiviralen Abwehr (pARTs). Die Entwicklung von Inhibitoren von mono-ARTs und poly-ARTs könnte neue therapeutische Ansätze eröffnen.

7 Literaturverzeichnis

Abbondanza C., Rossi V., Roscigno A., Gallo L., Belsito A., Piluso G., Medici N., Nigro V., Molinari A.M., Moncharmont B. and Puca G.A. (1998) "Interaction of vault particles with estrogen receptor in the MCF-7 breast cancer cell" *J Cell Biol* **141:** 1301-1310

Adams M.D., Celniker S.E., Holt R.A. *et al.* (2000) "The genome sequence of Drosophila melanogaster" *Science* **287:** 2185-2195

Adriouch S., Ohlrogge W., Haag F., Koch-Nolte F. and Seman M. (2001) "Rapid induction of naive T cell apoptosis by ecto-nicotinamide adenine dinucleotide: requirement for mono(ADP-ribosyl)transferase 2 and a downstream effector" *J Immunol* **167:** 196-203

Aguiar R.C., Yakushijin Y., Kharbanda S., Salgia R., Fletcher J.A. and Shipp M.A. (2000) "BAL is a novel risk-related gene in diffuse large B-cell lymphomas that enhances cellular migration" *Blood* **96:** 4328-4334

Ahmer B.M., Tran M. and Heffron F. (1999) "The virulence plasmid of Salmonella typhimurium is self-transmissible" *J Bacteriol* **181:** 1364-1368

Aktories K., Barmann M., Ohishi I., Tsuyama S., Jakobs K.H. and Habermann E. (1986) "Botulinum C2 toxin ADP-ribosylates actin" *Nature* **322:** 390-392

Aktories K. and Wegner A. (1992) "Mechanisms of the cytopathic action of actin-ADP-ribosylating toxins" *Mol Microbiol* **6:** 2905-2908

Aktories K. (1994) "Clostridial ADP-ribosylating toxins: effects on ATP and GTP-binding proteins" *Mol Cell Biochem* **138:** 167-176

Aktories K. and Just I. (2000) "Bacterial protein toxins", Springer Verlag, Berlin

Allen M.D., Buckle A.M., Cordell S.C., Lowe J. and Bycroft M. (2003) "The crystal structure of AF1521 a protein from Archaeoglobus fulgidus with homology to the non-histone domain of macroH2A" *J Mol Biol* **330:** 503-511

Allured V.S., Collier R.J., Carroll S.F. and McKay D.B. (1986) "Structure of exotoxin A of Pseudomonas aeruginosa at 3.0-Angstrom resolution" *Proc Natl Acad Sci U S A* **83:** 1320-1324

Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. (1990) "Basic local alignment search tool" *J Mol Biol* **215:** 403-410

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs" *Nucleic Acids Res* **25:** 3389-3402

Altschul S.F. and Koonin E.V. (1998) "Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases" *Trends Biochem Sci* **23:** 444-447

Ame J.C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P., Muller S., Hoger T., Menissier-de Murcia J. and de Murcia G. (1999) "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase" *J Biol Chem* **274:** 17860-17868

Ame J.C., Spenlehauer C. and de Murcia G. (2004) "The PARP superfamily" *Bioessays* **26:** 882-893

Andersson J.O., Doolittle W.F. and Nesbo C.L. (2001) "Genomics. Are there bugs in our genome ?" *Science* **292:** 1848-1850

Arabidopsis Genome Initiative (2000) "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana" *Nature* **408:** 796-815

Aravind L. and Koonin E.V. (2000a) "SAP - a putative DNA-binding motif involved in chromosomal organization" *Trends Biochem Sci* **25:** 112-114

Aravind L. and Koonin E.V. (2000b) "The U box is a modified RING finger - a common domain in ubiquitination" *Curr Biol* **10:** 132-134

Aravind L. (2001) "The WWE domain: a common interaction module in protein ubiquitination and ADP ribosylation" *Trends Biochem Sci* **26:** 273-275

Auer B., Nagl U., Herzog H., Schneider R. and Schweiger M. (1989) "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing): organization of the gene" *DNA* **8:** 575-580

Augustin A., Spenlehauer C., Dumond H., Menissier-De Murcia J., Piel M., Schmit A.C., Apiou F., Vonesch J.L., Kock M., Bornens M. and De Murcia G. (2003) "PARP-3 localizes preferentially to the daughter centriole and interferes with the G1/S cell cycle progression" *J Cell Sci* **116:** 1551-1562

Babiychuk E., Cottrill P.B., Storozhenko S., Fuangthong M., Chen Y., O'Farrell M.K., Van Montagu M., Inze D. and Kushnir S. (1998) "Higher plants possess two structurally different poly(ADP-ribose) polymerases" *Plant J* **15:** 635-645

Bae J., Donigian J.R. and Hsueh A.J. (2003) "Tankyrase 1 interacts with Mcl-1 proteins and inhibits their regulation of apoptosis" *J Biol Chem* **278:** 5195-5204

Bally M., Filloux A., Akrim M., Ball G., Lazdunski A. and Tommassen J. (1992) "Protein secretion in Pseudomonas aeruginosa: characterization of seven xcp genes and processing of secretory apparatus components by prepilin peptidase" *Mol Microbiol* **6:** 1121-1131

Banasik M., Komura H., Shimoyama M. and Ueda K. (1992) "Specific inhibitors of poly(ADP-ribose) synthetase and mono(ADP-ribosyl)transferase" *J Biol Chem* **267:** 1569-1575

Barbieri J.T. (2000) "Pseudomonas aeruginosa exoenzyme S, a bifunctional type-III secreted cytotoxin" *Int J Med Microbiol* **290:** 381-387

Barbieri J.T. and Sun J. (2004) "Pseudomonas aeruginosa ExoS and ExoT" *Rev Physiol Biochem Pharmacol* **152:** 79-92

Barth H., Hofmann F., Olenik C., Just I. and Aktories K. (1998a) "The N-terminal part of the enzyme component (C2I) of the binary Clostridium botulinum C2 toxin interacts with the binding component C2II and functions as a carrier system for a Rho ADP-ribosylating C3-like fusion toxin" *Infect Immun* **66:** 1364-1369

Barth H., Preiss J.C., Hofmann F. and Aktories K. (1998b) "Characterization of the catalytic site of the ADP-ribosyltransferase Clostridium botulinum C2 toxin by site-directed mutagenesis" *J Biol Chem* **273:** 29506-29511

Bazan J.F. and Koch-Nolte F. (1997) "Sequence and structural links between distant ADP-ribosyltransferase families" *Adv Exp Med Biol* **419:** 99-107

Bell C.E. and Eisenberg D. (1996) "Crystal structure of diphtheria toxin bound to nicotinamide adenine dinucleotide" *Biochemistry* **35:** 1137-1149

Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J. and Wheeler D.L. (2003) "GenBank" *Nucleic Acids Res* **31:** 23-27

Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J. and Wheeler D.L. (2005) "GenBank" *Nucleic Acids Res* **33:** 34-38

Bick M.J., Carroll J.W., Gao G., Goff S.P., Rice C.M. and MacDonald M.R. (2003) "Expression of the zinc-finger antiviral protein inhibits alphavirus replication" *J Virol* **77:** 11555-11562

Birney E., Kumar S. and Krainer A.R. (1993) "Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors" *Nucleic Acids Res* **21:** 5803-5816

Boguski M.S., Lowe T.M. and Tolstoshev C.M. (1993) "dbEST--database for expressed sequence tags" *Nat Genet* **4:** 332-333

Boguski M.S., Tolstoshev C.M. and Bassett D.E., Jr. (1994) "Gene discovery in dbEST" *Science* **265:** 1993-1994

Bowen D., Rocheleau T.A., Blackburn M., Andreev O., Golubeva E., Bhartia R. and ffrench-Constant R.H. (1998) "Insecticidal toxins from the bacterium Photorhabdus luminescens" *Science* **280:** 2129-2132

Boyd E.F. and Hartl D.L. (1998) "Salmonella virulence plasmid. Modular acquisition of the spv virulence region by an F-plasmid in Salmonella enterica subspecies I and insertion into the chromosome of subspecies II, IIIa, IV and VII isolates" *Genetics* **149:** 1183-1190

Braren R., Firner K., Balasubramanian S., Bazan F., Thiele H.G., Haag F. and Koch-Nolte F. (1997) "Use of the EST database resource to identify and clone novel mono(ADP-ribosyl)transferase gene family members" *Adv Exp Med Biol* **419:** 163-168

Braren R., Glowacki G., Nissen M., Haag F. and Koch-Nolte F. (1998) "Molecular characterization and expression of the gene for mouse NAD+:arginine ecto-mono(ADP-ribosyl)transferase, Art1" *Biochem J* **336 (Pt 3):** 561-568

Bredehorst K., Wursthorn K., Thiele H.G., Haag F. and Koch-Nolte F. (1997) "Using secondary structure predictions and site-directed mutagenesis to identify and probe the role of potential active site motifs in the RT6 mono(ADP-ribosyl)transferases" *Adv Exp Med Biol* **419:** 185-189

Bredehorst K. (2002) "Zielgerichtete Mutagenese der ADP-Ribosyltransferasen RT6.1 / RT6.2: Identifizierung der für die Katalyse essentiellen Aminosäuren" Dissertationsschrift. Institut für Immunologie, Universitätskrankenhaus Eppendorf, Hamburg

Browne S.H., Lesnick M.L. and Guiney D.G. (2002) "Genetic requirements for salmonella-induced cytopathology in human monocyte-derived macrophages" *Infect Immun* **70:** 7126-7135

Bruzzone S., Guida L., Zocchi E., Franco L. and De Flora A. (2001) "Connexin 43 hemi channels mediate Ca2+-regulated transmembrane NAD+ fluxes in intact cells" *Faseb J* **15:** 10-12

Burtscher H.J., Auer B., Klocker H., Schweiger M. and Hirsch-Kauffmann M. (1986) "Isolation of ADP-ribosyltransferase by affinity chromatography" *Anal Biochem* **152:** 285-290

Carroll S.F. and Collier R.J. (1984) "NAD binding site of diphtheria toxin: identification of a residue within the nicotinamide subsite by photochemical modification with NAD" *Proc Natl Acad Sci U S A* **81:** 3307-3311

Chambon P., Weil, J.D. and Mandel,P. (1963) "Nicotinamide mononucleotide activation of a new DNA-dependent polyadenylic acid synthesizing nuclear enzyme." *Biochem. Biophys. Res. Commun.*: 39-43

Chang W., Dynek J.N. and Smith S. (2003) "TRF1 is degraded by ubiquitin-mediated proteolysis after release from telomeres" *Genes Dev* **17:** 1328-1333

Chen L.M., Kaniga K. and Galan J.E. (1996) "Salmonella spp. are cytotoxic for cultured macrophages" *Mol Microbiol* **21:** 1101-1115

Chi N.W. and Lodish H.F. (2000) "Tankyrase is a golgi-associated mitogen-activated protein kinase substrate that interacts with IRAP in GLUT4 vesicles" *J Biol Chem* **275:** 38437-38444

Choe S., Bennett M.J., Fujii G., Curmi P.M., Kantardjieff K.A., Collier R.J. and Eisenberg D. (1992) "The crystal structure of diphtheria toxin" *Nature* **357:** 216-222

Chugani D.C., Rome L.H. and Kedersha N.L. (1993) "Evidence that vault ribonucleoprotein particles localize to the nuclear pore complex" *J Cell Sci* **106 (Pt 1):** 23-29

Cook B.D., Dynek J.N., Chang W., Shostak G. and Smith S. (2002) "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2 at human telomeres" *Mol Cell Biol* **22:** 332-342

D'Amours D., Desnoyers S., D'Silva I. and Poirier G.G. (1999) "Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions" *Biochem J* **342 (Pt 2):** 249-268

Dantzer F., Giraud-Panis M.J., Jaco I., Ame J.C., Schultz I., Blasco M., Koering C.E., Gilson E., Menissier-de Murcia J., de Murcia G. and Schreiber V. (2004) "Functional interaction between poly(ADP-Ribose) polymerase 2 (PARP-2) and TRF2: PARP activity negatively regulates TRF2" *Mol Cell Biol* **24:** 1595-1607

Davidovic L., Vodenicharov M., Affar E.B. and Poirier G.G. (2001) "Importance of poly(ADP-ribose) glycohydrolase in the control of poly(ADP-ribose) metabolism" *Exp Cell Res* **268:** 7-13

de Murcia G. and Menissier de Murcia J. (1994) "Poly(ADP-ribose) polymerase: a molecular nick-sensor" *Trends Biochem Sci* **19:** 172-176

De Rycker M., Venkatesan R.N., Wei C. and Price C.M. (2003) "Vertebrate tankyrase domain structure and sterile alpha motif (SAM)-mediated multimerization" *Biochem J* **372:** 87-96

De Rycker M. and Price C.M. (2004) "Tankyrase polymerization is controlled by its sterile alpha motif and poly(ADP-ribose) polymerase domains" *Mol Cell Biol* **24:** 9802-9812

DeLano W. (2002) "The PyMOL User's Manual" *DeLano Scientific,* San Carlos, California, U S A

Domenighini M. and Rappuoli R. (1996) "Three conserved consensus sequences identify the NAD-binding site of ADP-ribosylating enzymes, expressed by eukaryotes, bacteria and T-even bacteriophages" *Mol Microbiol* **21:** 667-674

Dynek J.N. and Smith S. (2004) "Resolution of sister telomere association is required for progression through mitosis" *Science* **304:** 97-100

Evans H.R., Sutton J.M., Holloway D.E., Ayriss J., Shone C.C. and Acharya K.R. (2003) "The crystal structure of C3stau2 from Staphylococcus aureus and its complex with NAD" *J Biol Chem* **278:** 45924-45930

Fang F.C., Krause M., Roudier C., Fierer J. and Guiney D.G. (1991) "Growth regulation of a Salmonella plasmid gene essential for virulence" *J Bacteriol* **173:** 6783-6789

Fang F.C., Libby S.J., Buchmeier N.A., Loewen P.C., Switala J., Harwood J. and Guiney D.G. (1992) "The alternative sigma factor katF (rpoS) regulates Salmonella virulence" *Proc Natl Acad Sci U S A* **89:** 11978-11982

Fedorov A., Cao X., Saxonov S., de Souza S.J., Roy S.W. and Gilbert W. (2001) "Intron distribution difference for 276 ancient and 131 modern genes suggests the existence of ancient introns" *Proc Natl Acad Sci U S A* **98:** 13177-13182

Ffrench-Constant R. and Bowen D. (1999) "Photorhabdus toxins: novel biological insecticides" *Curr Opin Microbiol* **2:** 284-288

Fierer J., Eckmann L., Fang F., Pfeifer C., Finlay B.B. and Guiney D. (1993) "Expression of the Salmonella virulence plasmid gene spvB in cultured macrophages and nonphagocytic cells" *Infect Immun* **61:** 5231-5236

Fierer J. and Guiney D.G. (2001) "Diverse virulence traits underlying different clinical outcomes of Salmonella infection" *J Clin Invest* **107:** 775-780

Gagne J.P., Shah R.G. and Poirier G.G. (2001) "Analysis of ADP-ribose polymer sizes in intact cells" *Mol Cell Biochem* **224:** 183-185

Gagnon S.N., Hengartner M.O. and Desnoyers S. (2002) "The genes pme-1 and pme-2 encode two poly(ADP-ribose) polymerases in Caenorhabditis elegans" *Biochem J* **368:** 263-271

Galan J.E. and Curtiss R., 3rd (1989) "Cloning and molecular characterization of genes whose products allow Salmonella typhimurium to penetrate tissue culture cells" *Proc Natl Acad Sci U S A* **86:** 6383-6387

Gao G., Guo X. and Goff S.P. (2002) "Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein" *Science* **297:** 1703-1706

Geer L.Y., Domrachev M., Lipman D.J. and Bryant S.H. (2002) "CDART: protein homology by domain architecture" *Genome Res* **12:** 1619-1623

Glowacki G., Braren R., Cetkovic-Cvrlje M., Leiter E.H., Haag F. and Koch-Nolte F. (2001) "Structure, chromosomal localization, and expression of the gene for mouse ecto-mono(ADP-ribosyl)transferase ART5" *Gene* **275:** 267-277

Glowacki G., Braren R., Firner K., Nissen M., Kuhl M., Reche P., Bazan F., Cetkovic-Cvrlje M., Leiter E., Haag F. and Koch-Nolte F. (2002) "The family of toxin-related ecto-ADP-ribosyltransferases in humans and the mouse" *Protein Sci* **11:** 1657-1670

Gotoh H., Okada N., Kim Y.G., Shiraishi K., Hirami N., Haneda T., Kurita A., Kikuchi Y. and Danbara H. (2003) "Extracellular secretion of the virulence plasmid-encoded ADP-ribosyltransferase SpvB in Salmonella" *Microb Pathog* **34:** 227-238

Gravel C., Stergiou L., Gagnon S.N. and Desnoyers S. (2004) "The C. elegans gene pme-5: molecular cloning and role in the DNA-damage response of a tankyrase orthologue" *DNA Repair (Amst)* **3:** 171-182

Gubin A.N., Njoroge J.M., Wojda U., Pack S.D., Rios M., Reid M.E. and Miller J.L. (2000) "Identification of the dombrock blood group glycoprotein as a polymorphic member of the ADP-ribosyltransferase gene family" *Blood* **96:** 2621-2627

Gulig P.A., Danbara H., Guiney D.G., Lax A.J., Norel F. and Rhen M. (1993) "Molecular analysis of spv virulence genes of the Salmonella virulence plasmids" *Mol Microbiol* **7:** 825-830

Gulig P.A. and Doyle T.J. (1993) "The Salmonella typhimurium virulence plasmid increases the growth rate of salmonellae in mice" *Infect Immun* **61:** 504-511

Gulig P.A., Doyle T.J., Hughes J.A. and Matsui H. (1998) "Analysis of host cells associated with the Spv-mediated increased intracellular growth rate of Salmonella typhimurium in mice" *Infect Immun* **66:** 2471-2485

Guo X., Carroll J.W., Macdonald M.R., Goff S.P. and Gao G. (2004) "The zinc finger antiviral protein directly binds to specific viral mRNAs through the CCCH zinc finger motifs" *J Virol* **78:** 12781-12787

Guse A.H. (2004) "Biochemistry, biology, and pharmacology of cyclic adenosine diphosphoribose (cADPR)" *Curr Med Chem* **11:** 847-855

Haag F., Nolte F., Lernmark A., Simrell C. and Thiele H.G. (1993) "Analysis of T-cell surface marker profiles during the postnatal ontogeny of normal and diabetes-prone rats" *Transplant Proc* **25:** 2831-2832

Haag F., Koch-Nolte F., Kuhl M., Lorenzen S. and Thiele H.G. (1994) "Premature stop codons inactivate the RT6 genes of the human and chimpanzee species" *J Mol Biol* **243:** 537-546

Haag F., Andresen V., Karsten S., Koch-Nolte F. and Thiele H. (1995) "Both allelic forms of the rat T cell differentiation marker RT6 display nicotinamide adenine dinucleotide (NAD)-glycohydrolase activity,

yet only RT6.2 is capable of automodification upon incubation with NAD" *Eur J Immunol* **25:** 2355-2361

Haag F. and Koch-Nolte F. (1997) "ADP-Ribosylation in Animal Tissues: Structure, Function and Biology of Mono(ADP-Ribosyl)transferases and Related Enzymes", Plenum Press, New York

Haag F. and Koch-Nolte F. (1998) "Endogenous relatives of ADP-ribosylating bacterial toxins in mice and men: potential regulators of immune cell function" *J Biol Regul Homeost Agents* **12:** 53-62

Han S., Craig J.A., Putnam C.D., Carozzi N.B. and Tainer J.A. (1999) "Evolution and mechanism from structures of an ADP-ribosylating toxin and NAD complex" *Nat Struct Biol* **6:** 932-936

Han S., Arvai A.S., Clancy S.B. and Tainer J.A. (2001) "Crystal structure and novel recognition motif of rho ADP-ribosylating C3 exoenzyme from Clostridium botulinum: structural insights for recognition specificity and catalysis" *J Mol Biol* **305:** 95-107

Han S. and Tainer J.A. (2002) "The ARTT motif and a unified structural understanding of substrate recognition in ADP-ribosylating bacterial toxins and eukaryotic ADP-ribosyltransferases" *Int J Med Microbiol* **291:** 523-529

Hanai S., Uchida M., Kobayashi S., Miwa M. and Uchida K. (1998) "Genomic organization of Drosophila poly(ADP-ribose) polymerase and distribution of its mRNA during development" *J Biol Chem* **273:** 11881-11886

Hara N., Tsuchiya M. and Shimoyama M. (1996) "Glutamic acid 207 in rodent T-cell RT6 antigens is essential for arginine-specific ADP-ribosylation" *J Biol Chem* **271:** 29552-29555

Harnett M.M. (1994) "Assay of protein kinases and protein phosphorylation" *Methods Mol. Biol.* **27:** 213-226

Hatakeyama S., Yada M., Matsumoto M., Ishida N. and Nakayama K.I. (2001) "U box proteins as a new family of ubiquitin-protein ligases" *J Biol Chem* **276:** 33111-33120

Hayaishi O. and Ueda K. (1977) "Poly(ADP-ribose) and ADP-ribosylation of proteins" *Annu Rev Biochem* **46:** 95-116

Hensel M., Shea J.E., Raupach B., Monack D., Falkow S., Gleeson C., Kubo T. and Holden D.W. (1997) "Functional analysis of ssaJ and the ssaK/U operon, 13 genes encoding components of the type III secretion apparatus of Salmonella Pathogenicity Island 2" *Mol Microbiol* **24:** 155-167

Hersh D., Monack D.M., Smith M.R., Ghori N., Falkow S. and Zychlinsky A. (1999) "The Salmonella invasin SipB induces macrophage apoptosis by binding to caspase-1" *Proc Natl Acad Sci U S A* **96:** 2396-2401

Herzog H., Zabel B.U., Schneider R., Auer B., Hirsch-Kauffmann M. and Schweiger M. (1989) "Human nuclear NAD+ ADP-ribosyltransferase: localization of the gene on chromosome 1q41-q42 and expression of an active human enzyme in Escherichia coli" *Proc Natl Acad Sci U S A* **86:** 3514-3518

Holbourn K.P., Sutton J.M., Evans H.R., Shone C.C. and Acharya K.R. (2005) "Molecular recognition of an ADP-ribosylating Clostridium botulinum C3 exoenzyme by RalA GTPase" *Proc Natl Acad Sci U S A* **102:** 5357-5362

Hollmann C., Haag F., Schlott M., Damaske A., Bertuleit H., Matthes M., Kuhl M., Thiele H.G. and Koch-Nolte F. (1996) "Molecular characterization of mouse T-cell ecto-ADP-ribosyltransferase Rt6: cloning of a second functional gene and identification of the Rt6 gene products" *Mol Immunol* **33:** 807-817

Honjo T., Nishizuka Y. and Hayaishi O. (1968) "Diphtheria toxin-dependent adenosine diphosphate ribosylation of aminoacyl transferase II and inhibition of protein synthesis" *J Biol Chem* **243:** 3553-3555

Huizinga E.G., Tsuji S., Romijn R.A., Schiphorst M.E., de Groot P.G., Sixma J.J. and Gros P. (2002) "Structures of glycoprotein Ibalpha and its complex with von Willebrand factor A1 domain" *Science* **297:** 1176-1179

Jacobson M.K. and Jacobson E.L. (1989) "ADP-ribose Transfer Reactions: Mechanisms and Biological Significance", Springer Verlag, New York

Jakob N.J. and Darai G. (2002) "Molecular anatomy of chilo iridescent virus genome and the evolution of viral genes" *Virus Genes* **25:** 299-316

Johansson M. (1999) "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA cloning of two novel poly(ADP-ribose) polymerase homologues" *Genomics* **57:** 442-445

Jonsson G.G., Menard L., Jacobson E.L., Poirier G.G. and Jacobson M.K. (1988) "Effect of hyperthermia on poly(adenosine diphosphate-ribose) glycohydrolase" *Cancer Res* **48:** 4240-4243

Just I., Selzer J., Jung M., van Damme J., Vandekerckhove J. and Aktories K. (1995) "Rho-ADP-ribosylating exoenzyme from Bacillus cereus. Purification, characterization, and identification of the NAD-binding site" *Biochemistry* **34:** 334-340

Kahl S., Nissen M., Girisch R., Duffy T., Leiter E.H., Haag F. and Koch-Nolte F. (2000) "Metalloprotease-mediated shedding of enzymatically active mouse ecto-ADP-ribosyltransferase ART2.2 upon T cell activation" *J Immunol* **165:** 4463-4469

Kaminker P.G., Kim S.H., Taylor R.D., Zebarjadian Y., Funk W.D., Morin G.B., Yaswen P. and Campisi J. (2001) "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes rapid induction of cell death upon overexpression" *J Biol Chem* **276:** 35891-35899

Karolchik D., Baertsch R., Diekhans M., Furey T.S., Hinrichs A., Lu Y.T., Roskin K.M., Schwartz M., Sugnet C.W., Thomas D.J., Weber R.J., Haussler D. and Kent W.J. (2003) "The UCSC Genome Browser Database" *Nucleic Acids Res* **31:** 51-54

Karsten S., Schroder J., da Silva C., Kahlke D., Thiele H.G., Koch-Noite F. and Haag F. (1997) "Expression and comparative analysis of recombinant rat and mouse RT6 T cell mono(ADP-ribosyl)transferases in E. coli" *Adv Exp Med Biol* **419:** 175-180

Katoh M. (2003) "Identification and characterization of human TIPARP gene within the CCNL amplicon at human chromosome 3q25.31" *Int J Oncol* **23:** 541-547

Kaufman M.R., Jia J., Zeng L., Ha U., Chow M. and Jin S. (2000) "Pseudomonas aeruginosa mediated apoptosis requires the ADP-ribosylating activity of exoS" *Microbiology* **146 (Pt 10):** 2531-2541

Kayser F.H., Bienz K.A., Eckert J. and Lindenmann J. (1993) "Medizinische Mikrobiologie", 8.Auflage. Georg Thieme Verlag Stuttgart, New York

Kickhoefer V.A., Siva A.C., Kedersha N.L., Inman E.M., Ruland C., Streuli M. and Rome L.H. (1999) "The 193-kD vault protein, VPARP, is a novel poly(ADP-ribose) polymerase" *J Cell Biol* **146:** 917-928

Koch F., Haag F., Kashan A. and Thiele H.G. (1990) "Primary structure of rat RT6.2, a nonglycosylated phosphatidylinositol-linked surface marker of postthymic T cells" *Proc Natl Acad Sci U S A* **87:** 964-967

Koch-Nolte F., Haag F., Kastelein R. and Bazan F. (1996a) "Uncovered: the family relationship of a T-cell-membrane protein and bacterial toxins" *Immunol Today* **17:** 402-405

Koch-Nolte F., Petersen D., Balasubramanian S., Haag F., Kahlke D., Willer T., Kastelein R., Bazan F. and Thiele H.G. (1996b) "Mouse T cell membrane proteins Rt6-1 and Rt6-2 are arginine/protein mono(ADPribosyl)transferases and share secondary structure motifs with ADP-ribosylating bacterial toxins" *J Biol Chem* **271:** 7686-7693

Koch-Nolte F., Haag F., Braren R., Kuhl M., Hoovers J., Balasubramanian S., Bazan F. and Thiele H.G. (1997) "Two novel human members of an emerging mammalian gene family related to mono-ADP-ribosylating bacterial toxins" *Genomics* **39:** 370-376

Koch-Nolte F., Reche P., Haag F. and Bazan F. (2001) "ADP-ribosyltransferases: plastic tools for inactivating protein and small molecular weight targets" *J Biotechnol* **92:** 81-87

Kofler B., Wallraff E., Herzog H., Schneider R., Auer B. and Schweiger M. (1993) "Purification and characterization of NAD+:ADP-ribosyltransferase (polymerizing) from Dictyostelium discoideum" *Biochem J* **293 (Pt 1):** 275-281

Kong L.B., Siva A.C., Rome L.H. and Stewart P.L. (1999) "Structure of the vault, a ubiquitous celular component" *Structure Fold Des* **7:** 371-379

Kong L.B., Siva A.C., Kickhoefer V.A., Rome L.H. and Stewart P.L. (2000) "RNA location and modeling of a WD40 repeat domain within the vault" *Rna* **6:** 890-900

Kounnas M.Z., Morris R.E., Thompson M.R., FitzGerald D.J., Strickland D.K. and Saelinger C.B. (1992) "The alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein binds and internalizes Pseudomonas exotoxin A" *J Biol Chem* **267:** 12420-12423

Kraus W.L. and Lis J.T. (2003) "PARP goes transcription" *Cell* **113:** 677-683

Kubori T., Matsushima Y., Nakamura D., Uralil J., Lara-Tejero M., Sukhan A., Galan J.E. and Aizawa S.I. (1998) "Supramolecular structure of the Salmonella typhimurium type III protein secretion system" *Science* **280:** 602-605

Kuimov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K., Scanlan M.J., Jongeneel C.V., Lagarkova M.A. and Nedospasov S.A. (2001) "Cloning and characterization of TNKL, a member of tankyrase gene family" *Genes Immun* **2:** 52-55

Kurita A., Gotoh H., Eguchi M., Okada N., Matsuura S., Matsui H., Danbara H. and Kikuchi Y. (2003) "Intracellular expression of the Salmonella plasmid virulence protein, SpvB, causes apoptotic cell death in eukaryotic cells" *Microb Pathog* **35:** 43-48

Lander E.S., Linton L.M., Birren B. *et al.* (2001) "Initial sequencing and analysis of the human genome" *Nature* **409:** 860-921

Lepiniec L., Babiychuk E., Kushnir S., Van Montagu M. and Inze D. (1995) "Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(ADP-ribose) polymerase" *FEBS Lett* **364:** 103-108

Lesnick M.L. and Guiney D.G. (2001) "The best defense is a good offense--Salmonella deploys an ADP-ribosylating toxin" *Trends Microbiol* **9:** 2-4; discussion 4-5

Lesnick M.L., Reiner N.E., Fierer J. and Guiney D.G. (2001) "The Salmonella spvB virulence gene encodes an enzyme that ADP-ribosylates actin and destabilizes the cytoskeleton of eukaryotic cells" *Mol Microbiol* **39:** 1464-1470

Libby S.J., Adams L.G., Ficht T.A., Allen C., Whitford H.A., Buchmeier N.A., Bossie S. and Guiney D.G. (1997) "The spv genes on the Salmonella dublin virulence plasmid are required for severe enteritis and systemic infection in the natural host" *Infect Immun* **65:** 1786-1792

Libby S.J., Lesnick M., Hasegawa P., Weidenhammer E. and Guiney D.G. (2000) "The Salmonella virulence plasmid spv genes are required for cytopathology in human monocyte-derived macrophages" *Cell Microbiol* **2:** 49-58

Liu L., Chen G., Ji X. and Gao G. (2004a) "ZAP is a CRM1-dependent nucleocytoplasmic shuttling protein" *Biochem Biophys Res Commun* **321:** 517-523

Liu Y., Snow B.E., Kickhoefer V.A., Erdmann N., Zhou W., Wakeham A., Gomez M., Rome L.H. and Harrington L. (2004b) "Vault poly(ADP-ribose) polymerase is associated with mammalian telomerase and is dispensable for telomerase function and vault structure in vivo" *Mol Cell Biol* **24:** 5314-5323

Long M., de Souza S.J. and Gilbert W. (1995) "Evolution of the intron-exon structure of eukaryotic genes" *Curr Opin Genet Dev* **5:** 774-778

Lorick K.L., Jensen J.P., Fang S., Ong A.M., Hatakeyama S. and Weissman A.M. (1999) "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination" *Proc Natl Acad Sci U S A* **96:** 11364-11369

Ludden P.W. (1994) "Reversible ADP-ribosylation as a mechanism of enzyme regulation in procaryotes" *Mol Cell Biochem* **138:** 123-129

Lyons R.J., Deane R., Lynch D.K., Ye Z.S., Sanderson G.M., Eyre H.J., Sutherland G.R. and Daly R.J. (2001) "Identification of a novel human tankyrase through its interaction with the adaptor protein Grb14" *J Biol Chem* **276:** 17172-17180

Ma Q., Baldwin K.T., Renzelli A.J., McDaniel A. and Dong L. (2001) "TCDD-inducible poly(ADP-ribose) polymerase: a novel response to 2,3,7,8-tetrachlorodibenzo-p-dioxin" *Biochem Biophys Res Commun* **289:** 499-506

Maehama T., Hoshino S. and Katada T. (1996) "Increase in ADP-ribosyltransferase activity of rat T lymphocyte alloantigen RT6.1 by a single amino acid mutation" *FEBS Lett* **388:** 189-191

Marchler-Bauer A., Anderson J.B., Cherukuri P.F. *et al.* (2005) "CDD: a Conserved Domain Database for protein classification" *Nucleic Acids Res* **33:** D192-196

Marcotte E.M., Pellegrini M., Ng H.L., Rice D.W., Yeates T.O. and Eisenberg D. (1999) "Detecting protein function and protein-protein interactions from genome sequences" *Science* **285:** 751-753

Marcus S.L., Brumell J.H., Pfeifer C.G. and Finlay B.B. (2000) "Salmonella pathogenicity islands: big virulence in small packages" *Microbes Infect* **2:** 145-156

Marsischky G.T., Wilson B.A. and Collier R.J. (1995) "Role of glutamic acid 988 of human poly-ADP-ribose polymerase in polymer formation. Evidence for active site similarities to the ADP-ribosylating toxins" *J Biol Chem* **270:** 3247-3254

McGuffin L.J., Bryson K. and Jones D.T. (2000) "The PSIPRED protein structure prediction server" *Bioinformatics* **16:** 404-405

McGuffin L.J. and Jones D.T. (2003) "Improvement of the GenTHREADER method for genomic fold recognition" *Bioinformatics* **19:** 874-881

Menissier de Murcia J., Ricoul M., Tartier L., Niedergang C., Huber A., Dantzer F., Schreiber V., Ame J.C., Dierich A., LeMeur M., Sabatier L., Chambon P. and de Murcia G. (2003) "Functional interaction between PARP-1 and PARP-2 in chromosome stability and embryonic development in mouse" *Embo J* **22:** 2255-2263

Mewes H.W., Albermann K., Bahr M., Frishman D., Gleissner A., Hani J., Heumann K., Kleine K., Maierl A., Oliver S.G., Pfeiffer F. and Zollner A. (1997) "Overview of the yeast genome" *Nature* **387:** 7-65

Mierendorf R., Yeager K. and Novy R. (1994) "The pET-System: Your Choice for Expression" *inNovations - Newsletter of NOVAGEN inc.* **1:** 1-3

Monack D.M., Raupach B., Hromockyj A.E. and Falkow S. (1996) "Salmonella typhimurium invasion induces apoptosis in infected macrophages" *Proc Natl Acad Sci U S A* **93:** 9833-9838

Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.I., Feiden W., Glass B. and Meese E.U. (2001) "Novel tankyrase-related gene detected with meningioma-specific sera" *Clin Cancer Res* **7:** 113-119

Moss J. and Vaughan M. (1990) "ADP-ribosylating toxins and G proteins: Insights into signal transduction", American Society for Microbiology, Washington DC

Mueller-Dieckmann C., Ritter H., Haag F., Koch-Nolte F. and Schulz G.E. (2002) "Structure of the ecto-ADP-ribosyl transferase ART2.2 from rat" *J Mol Biol* **322:** 687-696

Notredame C., Higgins D.G. and Heringa J. (2000) "T-Coffee: A novel method for fast and accurate multiple sequence alignment" *J Mol Biol* **302:** 205-217

Ochman H., Soncini F.C., Solomon F. and Groisman E.A. (1996) "Identification of a pathogenicity island required for Salmonella survival in host cells" *Proc Natl Acad Sci U S A* **93:** 7800-7804

Ohlrogge W., Haag F., Lohler J., Seman M., Littman D.R., Killeen N. and Koch-Nolte F. (2002) "Generation and characterization of ecto-ADP-ribosyltransferase ART2.1/ART2.2-deficient mice" *Mol Cell Biol* **22:** 7535-7542

Ohno S. (1999) "Gene duplication and the uniqueness of vertebrate genomes circa 1970-1999" *Semin Cell Dev Biol* **10:** 517-522

Okazaki I.J. and Moss J. (1996) "Structure and function of eukaryotic mono-ADP-ribosyltransferases" *Rev Physiol Biochem Pharmacol* **129:** 51-104

Okazaki I.J. and Moss J. (1998) "Glycosylphosphatidylinositol-anchored and secretory isoforms of mono-ADP-ribosyltransferases" *J Biol Chem* **273:** 23617-23620

Oliver A.W., Ame J.C., Roe S.M., Good V., de Murcia G. and Pearl L.H. (2004) "Crystal structure of the catalytic fragment of murine poly(ADP-ribose) polymerase-2" *Nucleic Acids Res* **32:** 456-464

Oliver F.J., Menissier-de Murcia J., Nacci C., Decker P., Andriantsitohaina R., Muller S., de la Rubia G., Stoclet J.C. and de Murcia G. (1999) "Resistance to endotoxic shock as a consequence of defective NF-kappaB activation in poly (ADP-ribose) polymerase-1 deficient mice" *Embo J* **18:** 4446-4454

Pallen M.J., Lam A.C., Loman N.J. and McBride A. (2001) "An abundance of bacterial ADP-ribosyltransferases--implications for the origin of exotoxins and their human homologues" *Trends Microbiol* **9:** 302-307; discussion 308

Pannifer A.D., Wong T.Y., Schwarzenbacher R., Renatus M., Petosa C., Bienkowska J., Lacy D.B., Collier R.J., Park S., Leppla S.H., Hanna P. and Liddington R.C. (2001) "Crystal structure of the anthrax lethal factor" *Nature* **414:** 229-233

Partida-Sanchez S., Randall T.D. and Lund F.E. (2003) "Innate immunity is regulated by CD38, an ecto-enzyme with ADP-ribosyl cyclase activity" *Microbes Infect* **5:** 49-58

Passador L. and Iglewski W. (1994) "ADP-ribosylating toxins" *Methods Enzymol* **235:** 617-631

Payne D.M., Jacobson E.L., Moss J. and Jacobson M.K. (1985) "Modification of proteins by mono(ADP-ribosylation) in vivo" *Biochemistry* **24:** 7540-7549

Pehrson J.R. and Fuji R.N. (1998) "Evolutionary conservation of histone macroH2A subtypes and domains" *Nucleic Acids Res* **26:** 2837-2842

Popoff I., Jijon H., Monia B., Tavernini M., Ma M., McKay R. and Madsen K. (2002) "Antisense oligonucleotides to poly(ADP-ribose) polymerase-2 ameliorate colitis in interleukin-10-deficient mice" *J Pharmacol Exp Ther* **303:** 1145-1154

Pugsley A.P. (1993) "The complete general secretory pathway in gram-negative bacteria" *Microbiol Rev* **57:** 50-108

Quan S., Imai T., Mikami Y., Yazawa K., Dabbs E.R., Morisaki N., Iwasaki S., Hashimoto Y. and Furihata K. (1999) "ADP-ribosylation as an intermediate step in inactivation of rifampin by a mycobacterial gene" *Antimicrob Agents Chemother* **43:** 181-184

Radke J., Pederson K.J. and Barbieri J.T. (1999) "Pseudomonas aeruginosa exoenzyme S is a biglutamic acid ADP- ribosyltransferase" *Infect Immun* **67:** 1508-1510

Rappuoli R. and Montecucco C. (1997) "Guidebook to Protein Toxins and Their Use in Cell Biology", Oxford University Press, Oxford

Reeder R.H., Ueda K., Honjo T., Nishizuka Y. and Hayaishi O. (1967) "Studies on the polymer of adenosine diphosphate ribose. II. Characterization of the polymer" *J Biol Chem* **242:** 3172-3179

Reich K.A., Biegel T. and Schoolnik G.K. (1997) "The light organ symbiont Vibrio fischeri possesses two distinct secreted ADP-ribosyltransferases" *J Bacteriol* **179:** 1591-1597

Reuner K.H., Presek P., Boschek C.B. and Aktories K. (1987) "Botulinum C2 toxin ADP-ribosylates actin and disorganizes the microfilament network in intact cells" *Eur J Cell Biol* **43:** 134-140

Rhen M., Riikonen P. and Taira S. (1993) "Transcriptional regulation of Salmonella enterica virulence plasmid genes in cultured macrophages" *Mol Microbiol* **10:** 45-56

Richter-Dahlfors A., Buchan A.M. and Finlay B.B. (1997) "Murine salmonellosis studied by confocal microscopy: Salmonella typhimurium resides intracellularly inside macrophages and exerts a cytotoxic effect on phagocytes in vivo" *J Exp Med* **186:** 569-580

Riese M.J., Goehring U.M., Ehrmantraut M.E., Moss J., Barbieri J.T., Aktories K. and Schmidt G. (2002) "Auto-ADP-ribosylation of Pseudomonas aeruginosa ExoS" *J Biol Chem* **277:** 12082-12088

Rost B. (1999) "Twilight zone of protein sequence alignments" *Protein Eng* **12:** 85-94

Roy S.W. (2003) "Recent evidence for the exon theory of genes" *Genetica* **118:** 251-266

Ruf A., Mennissier de Murcia J., de Murcia G. and Schulz G.E. (1996) "Structure of the catalytic fragment of poly(AD-ribose) polymerase from chicken" *Proc Natl Acad Sci U S A* **93:** 7481-7485

Ruf A., Rolli V., de Murcia G. and Schulz G.E. (1998) "The mechanism of the elongation and branching reaction of poly(ADP-ribose) polymerase as derived from crystal structures and mutagenesis" *J Mol Biol* **278:** 57-65

Sandkvist M., Michel L.O., Hough L.P., Morales V.M., Bagdasarian M., Koomey M. and DiRita V.J. (1997) "General secretion pathway (eps) genes required for toxin secretion and outer membrane biogenesis in Vibrio cholerae" *J Bacteriol* **179:** 6994-7003

Sbodio J.I., Lodish H.F. and Chi N.W. (2002) "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1 (telomere-repeat-binding factor 1) and IRAP (insulin-responsive aminopeptidase)" *Biochem J* **361:** 451-459

Schirmer J., Just I. and Aktories K. (2002a) "The ADP-ribosylating mosquitocidal toxin from Bacillus sphaericus: proteolytic activation, enzyme activity, and cytotoxic effects" *J Biol Chem* **277:** 11941-11948

Schirmer J., Wieden H.J., Rodnina M.V. and Aktories K. (2002b) "Inactivation of the elongation factor Tu by mosquitocidal toxin-catalyzed mono-ADP-ribosylation" *Appl Environ Microbiol* **68:** 4894-4899

Schneider R., Auer B., Kuhne C., Herzog H., Klocker H., Burtscher H.J., Hirsch-Kauffmann M., Wintersberger U. and Schweiger M. (1987) "Isolation of a cDNA clone for human NAD+: protein ADP-ribosyltransferase" *Eur J Cell Biol* **44:** 302-307

Schreiber V., Ame J.C., Dolle P., Schultz I., Rinaldi B., Fraulob V., Menissier-de Murcia J. and de Murcia G. (2002) "Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1" *J Biol Chem* **277:** 23028-23036

Schroeijers A.B., Siva A.C., Scheffer G.L., de Jong M.C., Bolick S.C., Dukers D.F., Slootstra J.W., Meloen R.H., Wiemer E., Kickhoefer V.A., Rome L.H. and Scheper R.J. (2000) "The Mr 193,000 vault protein is up-regulated in multidrug-resistant cancer cell lines" *Cancer Res* **60:** 1104-1110

Schuler G.D. (1997) "Pieces of the puzzle: expressed sequence tags and the catalog of human genes" *J Mol Med* **75:** 694-698

Schuman E.M., Meffert M.K., Schulman H. and Madison D.V. (1994) "An ADP-ribosyltransferase as a potential target for nitric oxide action in hippocampal long-term potentiation" *Proc Natl Acad Sci U S A* **91:** 11958-11962

Shall S. and de Murcia G. (2000) "Poly(ADP-ribose) polymerase-1: what have we learned from the deficient mouse model ?" *Mutat Res* **460:** 1-15

Shieh W.M., Ame J.C., Wilson M.V., Wang Z.Q., Koh D.W., Jacobson M.K. and Jacobson E.L. (1998) "Poly(ADP-ribose) polymerase null mouse cells synthesize ADP-ribose polymers" *J Biol Chem* **273:** 30069-30072

Shimokawa T., Masutani M., Nagasawa S., Nozaki T., Ikota N., Aoki Y., Nakagama H. and Sugimura T. (1999) "Isolation and cloning of rat poly(ADP-ribose) glycohydrolase: presence of a potential nuclear export signal conserved in mammalian orthologs" *J Biochem (Tokyo)* **126:** 748-755

Simonin F., Hofferer L., Panzeter P.L., Muller S., de Murcia G. and Althaus F.R. (1993) "The carboxyl-terminal domain of human poly(ADP-ribose) polymerase. Overproduction in Escherichia coli, large scale purification, and characterization" *J Biol Chem* **268:** 13454-13461

Simpson L.I. (1989) "The binary toxin produced by Clostridium botulinum enters cells by receptor-mediated endocytosis to exert its pharmacologic effects" *J Pharmacol Exp Ther* **251:** 1223-1228

Sixma T.K., Pronk S.E., Kalk K.H., Wartna E.S., van Zanten B.A., Witholt B. and Hol W.G. (1991) "Crystal structure of a cholera toxin-related heat-labile enterotoxin from E. coli" *Nature* **351:** 371-377

Smith S., Giriat I., Schmitt A. and de Lange T. (1998) "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres" *Science* **282:** 1484-1487

Smith S. and de Lange T. (2000) "Tankyrase promotes telomere elongation in human cells" *Curr Biol* **10:** 1299-1302

Smith S. (2001) "The world according to PARP" *Trends Biochem Sci* **26:** 174-179

Soares M.B., Bonaldo M.F., Jelene P., Su L., Lawton L. and Efstratiadis A. (1994) "Construction and characterization of a normalized cDNA library" *Proc Natl Acad Sci U S A* **91:** 9228-9232

Spinelli S.L., Kierzek R., Turner D.H. and Phizicky E.M. (1999) "Transient ADP-ribosylation of a 2'-phosphate implicated in its removal from ligated tRNA during splicing in yeast" *J Biol Chem* **274:** 2637-2644

Stein P.E., Boodhoo A., Armstrong G.D., Cockle S.A., Klein M.H. and Read R.J. (1994) "The crystal structure of pertussis toxin" *Structure* **2:** 45-57

Stryer L. (1994) "Biochemie", Aufl. 1990, 2. korrigierter Nachdruck. Spektrum Akademischer Verlag GmbH, Heidelberg, Berlin, Oxford

Studier F.W. and Moffatt B.A. (1986) "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes" *J Mol Biol* **189:** 113-130

Sugai M., Hashimoto K., Kikuchi A., Inoue S., Okumura H., Matsumoto K., Goto Y., Ohgai H., Moriishi K., Syuto B. and *et al.* (1992) "Epidermal cell differentiation inhibitor ADP-ribosylates small GTP-binding proteins and induces hyperplasia of epidermis" *J Biol Chem* **267:** 2600-2604

Sun J., Maresso A.W., Kim J.J. and Barbieri J.T. (2004) "How bacterial ADP-ribosylating toxins recognize substrates" *Nat Struct Mol Biol* **11:** 868-876

Taira S. and Rhen M. (1989) "Identification and genetic analysis of mkaA--a gene of the Salmonella typhimurium virulence plasmid necessary for intracellular growth" *Microb Pathog* **7:** 165-173

Taira S., Baumann M., Riikonen P., Sukupolvi S. and Rhen M. (1991a) "Amino-terminal sequence analysis of four plasmid-encoded virulence-associated proteins of Salmonella typhimurium" *FEMS Microbiol Lett* **61:** 319-323

Taira S., Riikonen P., Saarilahti H., Sukupolvi S. and Rhen M. (1991b) "The mkaC virulence gene of the Salmonella serovar typhimurium 96 kb plasmid encodes a transcriptional activator" *Mol Gen Genet* **228:** 381-384

Takamura-Enya T., Watanabe M., Totsuka Y., Kanazawa T., Matsushima-Hibiya Y., Koyama K., Sugimura T. and Wakabayashi K. (2001) "Mono(ADP-ribosyl)ation of 2'-deoxyguanosine residue in DNA by an apoptosis-inducing protein, pierisin-1, from cabbage butterfly" *Proc Natl Acad Sci U S A* **98:** 12414-12419

Takeyama K., Aguiar R.C., Gu L., He C., Freeman G.J., Kutok J.L., Aster J.C. and Shipp M.A. (2003) "The BAL-binding protein BBAP and related Deltex family members exhibit ubiquitin-protein isopeptide ligase activity" *J Biol Chem* **278:** 21930-21937

Tezcan-Merdol D., Nyman T., Lindberg U., Haag F., Koch-Nolte F. and Rhen M. (2001) "Actin is ADP-ribosylated by the Salmonella enterica virulence-associated protein SpvB" *Mol Microbiol* **39:** 606-619

Thanabalu T., Berry C. and Hindley J. (1993) "Cytotoxicity and ADP-ribosylating activity of the mosquitocidal toxin from Bacillus sphaericus SSII-1: possible roles of the 27- and 70-kilodalton peptides" *J Bacteriol* **175:** 2314-2320

Thiele H.G., Koch F., Hamann A. and Arndt R. (1986) "Biochemical characterization of the T-cell alloantigen RT-6.2" *Immunology* **59:** 195-201

Totsuka Y., Kawanishi M., Nishigaki R., Matsukawa K., Yagi T., Takamura-Enya T., Watanabe M., Sugimura T. and Wakabayashi K. (2003) "Analysis of HPRT and supF mutations caused by pierisin-1, a guanine specific ADP-ribosylating toxin derived from the cabbage butterfly" *Chem Res Toxicol* **16:** 945-952

Tschopp J., Irmler M. and Thome M. (1998) "Inhibition of fas death signals by FLIPs" *Curr Opin Immunol* **10:** 552-558

Tsuchiya M., Hara N., Yamada K., Osago H. and Shimoyama M. (1994) "Cloning and expression of cDNA for arginine-specific ADP-ribosyltransferase from chicken bone marrow cells" *J Biol Chem* **269:** 27451-27457

Tsuge H., Nagahama M., Nishimura H., Hisatsune J., Sakaguchi Y., Itogawa Y., Katunuma N. and Sakurai J. (2003) "Crystal structure and site-directed mutagenesis of enzymatic components from Clostridium perfringens iota-toxin" *J Mol Biol* **325:** 471-483

Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S., Suzuki H., Nyunoya H., Miwa M. and Sugimura T. (1987) "Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribose) polymerase" *Biochem Biophys Res Commun* **148:** 617-622

Uchiya K., Barbieri M.A., Funato K., Shah A.H., Stahl P.D. and Groisman E.A. (1999) "A Salmonella virulence protein that inhibits cellular trafficking" *Embo J* **18:** 3924-3933

van der Velden A.W., Lindgren S.W., Worley M.J. and Heffron F. (2000) "Salmonella pathogenicity island 1-independent induction of apoptosis in infected macrophages by Salmonella enterica serotype typhimurium" *Infect Immun* **68:** 5702-5709

van Ness B.G., Howard J.B. and Bodley J.W. (1980) "ADP-ribosylation of elongation factor 2 by diphtheria toxin. NMR spectra and proposed structures of ribosyl-diphthamide and its hydrolysis products" *J Biol Chem* **255:** 10710-10716

van Steensel B. and de Lange T. (1997) "Control of telomere length by the human telomeric protein TRF1" *Nature* **385:** 740-743

Venter J.C., Adams M.D., Myers E.W. *et al.* (2001) "The sequence of the human genome" *Science* **291:** 1304-1351

Virag L. and Szabo C. (2002) "The therapeutic potential of poly(ADP-ribose) polymerase inhibitors" *Pharmacol Rev* **54:** 375-429

Waterston R.H., Lindblad-Toh K., Birney E. *et al.* (2002) "Initial sequencing and comparative analysis of the mouse genome" *Nature* **420:** 520-562

Watson J.D., Gilman M., Witkowski J. and Zoller M. (1993) "Rekombinierte DNA", 2.Auflage. Spektrum Akademischer Verlag GmbH, Heidelberg, Berlin, Oxford. S.111

Wheeler D.L., Barrett T., Benson D.A. *et al.* (2005) "Database resources of the National Center for Biotechnology Information" *Nucleic Acids Res* **33 Database Issue:** D39-45

Wiegers W., Just I., Muller H., Hellwig A., Traub P. and Aktories K. (1991) "Alteration of the cytoskeleton of mammalian cells cultured in vitro by Clostridium botulinum C2 toxin and C3 ADP-ribosyltransferase" *Eur J Cell Biol* **54:** 237-245

Williamson C.M., Pullinger G.D. and Lax A.J. (1988) "Identification of an essential virulence region on Salmonella plasmids" *Microb Pathog* **5:** 469-473

Williamson K.C. and Moss J. (1990) "Mono-ADP-Ribosyl-Transferases and ADP-ribosylarginine hydrolases: A mono-ADP-ribosylation cycle in animal cells". **In:** Moss J. and Vaughan, M. "ADP-ribosylating Toxins and G-Proteins: Insights into signal transduction", American Society for Microbiology, Washington D.C., S.493-510

Wilson B.A. and Collier R.J. (1992) "Diphtheria toxin and Pseudomonas aeruginosa exotoxin A: active-site structure and enzymic mechanism" *Curr Top Microbiol Immunol* **175:** 27-41

Wistuba, II, Behrens C., Virmani A.K., Mele G., Milchgrub S., Girard L., Fondon J.W., 3rd, Garner H.R., McKay B., Latif F., Lerman M.I., Lam S., Gazdar A.F. and Minna J.D. (2000) "High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints" *Cancer Res* **60:** 1949-1960

Yamaizumi M., Mekada E., Uchida T. and Okada Y. (1978) "One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell" *Cell* **15:** 245-250

Ye J.Z. and de Lange T. (2004) "TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex" *Nat Genet* **36:** 618-623

Ying W., Sevigny M.B., Chen Y. and Swanson R.A. (2001) "Poly(ADP-ribose) glycohydrolase mediates oxidative and excitotoxic neuronal death" *Proc Natl Acad Sci U S A* **98:** 12227-12232

Yost D.A. and Moss J. (1983) "Amino acid-specific ADP-ribosylation. Evidence for two distinct NAD:arginine ADP-ribosyltransferases in turkey erythrocytes" *J Biol Chem* **258:** 4926-4929

Yu M., Schreek S., Cerni C. *et al.* (2005) "PARP-10, a novel Myc-interacting protein with poly(ADP-ribose) polymerase activity, inhibits transformation" *Oncogene* **24:** 1982-1993

Yu S.W., Wang H., Poitras M.F., Coombs C., Bowers W.J., Federoff H.J., Poirier G.G., Dawson T.M. and Dawson V.L. (2002) "Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor" *Science* **297:** 259-263

Zhang R.G., Scott D.L., Westbrook M.L., Nance S., Spangler B.D., Shipley G.G. and Westbrook E.M. (1995) "The three-dimensional crystal structure of cholera toxin" *J Mol Biol* **251:** 563-573

Zhang X., Morera S., Bates P.A., Whitehead P.C., Coffer A.I., Hainbucher K., Nash R.A., Sternberg M.J., Lindahl T. and Freemont P.S. (1998) "Structure of an XRCC1 BRCT domain: a new protein-protein interaction module" *Embo J* **17:** 6404-6411

Zolkiewska A., Nightingale M.S. and Moss J. (1992) "Molecular characterization of NAD:arginine ADP-ribosyltransferase from rabbit skeletal muscle" *Proc Natl Acad Sci U S A* **89:** 11352-11356

Zolkiewska A. and Moss J. (1993) "Integrin alpha 7 as substrate for a glycosylphosphatidylinositol-anchored ADP-ribosyltransferase on the surface of skeletal muscle cells" *J Biol Chem* **268:** 25273-25276

8 Anhang

Im Anhang sind die Primersequenzen (A), die Konsensussequenzen von hpART3, hpART4, hpART1 und SpvB (B), die Karten der verwendeten Vektoren (C) sowie die Alignments der pART-Subgruppen 1-5 (D) abgebildet.

A Primersequenzen

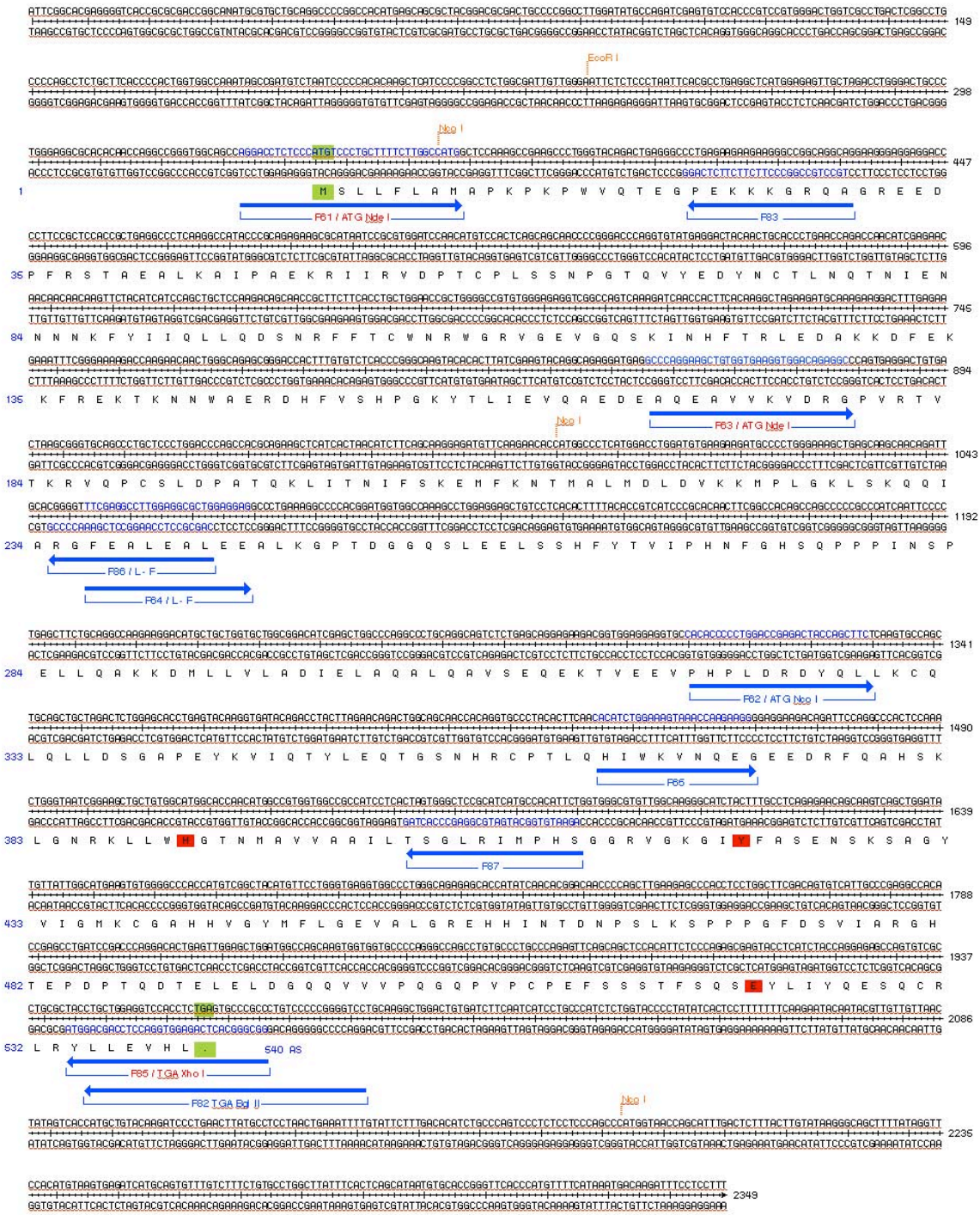
Gen	Primer-name	Orientierung	Länge (bp)	Nucleotidsequenz	RES	Besonderheiten
hpART1	P02	vorwärts	38	CAGGATGAAcAtatgGTGAAGAAGCTGACAGTAAATCC	Nde I	Einbau START
	P11	rückwärts	39	TACCTCTCCCtcgagCCACAGGGAGGTCTTAAAATTGAA	Xho I	Ausbau STOP
hpART3	F61	vorwärts	36	AGGACCTCTcatATGTCCCTGCTTTTCTTGGCCATG	Nde I	natürlicher START
	F63	vorwärts	33	GCCCAGGAAcaTaTGGTGAAGGTGGACAGAGGC	Nde I	Einbau START
	F64	vorwärts	27	TTCGAGGCCTTcGAGGCGCTGGAGGAG	keine	L / F - Mutation
	F62	vorwärts	30	CACACCCCaTGGACCGAGACTACCAGCTTC	Nco I	Einbau START
	F65	vorwärts	26	CACATCTGGAAAGTAAACCAAGAAGG	keine	keine
	F83	rückwärts	27	TGCCTGCCGGCCCTTCTTCTTCTCAGG	keine	keine
	F86	rückwärts	27	CAGCGCCTCgAAGGCCTCGAAACCCCG	keine	L / F - Mutation
	F87	rückwärts	29	AGAATGTGGCATGATGCGGAGCCCACTAG	keine	keine
	F85	rückwärts	33	GGCGGGCACctcGAGGTGGACCTCCAGCAGGTA	Xho I	Ausbau STOP
	F82	rückwärts	36	GGACAGatCttaatCTCAGAGGTGGACCTCCAGCAG	Bgl II	natürliches STOP
hpART4	K01	vorwärts	40	TTTGAAAATcAtATgGAAGAACTGAAGAAACAAGGATTTC	Nde I	Einbau START
	K02	vorwärts	40	TTTGAAAATTccATgGAAGAACTGAAGAAACAAGGATTTC	Nco I	Einbau START
	K11	rückwärts	39	ATGATCACTctcgaGAAAGTCCTTTATCTGATCTCCAGG	Xho I	kein STOP
SpvB	SP1	vorwärts	34	CCTCCTCCGAccATGGGAGGTAATTCATCTCGAC	Nco I	START
	SP2	vorwärts	34	CCTCCTCCGcatATGGGAGGTAATTCATCTCGAC	Nde I	START
	SP11	vorwärts	26	CAGAGTCGTATACaamGGCCTGAAGC	keine	R / N, K - Mutation
	SP12	vorwärts	28	AGGGAGAGGCAGmyATGCTTTTCCCTCC	keine	E / A, D - Mutation
	SP21	rückwärts	26	GCTTCAGGCCkttGTATACGACTCTG	keine	R / N, K - Mutation
	SP22	rückwärts	28	GGAGGGAAAAGCATrkCTGCCTCTCCCT	keine	E / A, D - Mutation
	SP4	rückwärts	36	GATTCTTAGctcgagTGAGTTGAGTACCCTCATGTT	Xho I	Ausbau STOP
Vektor						
pCR2.1	TAF	vorwärts	21	TCACACAGGAAACAGCTATGAC	keine	Sequenzierung
	TAR	rückwärts	22	GACGTTGTAAAACGACGGCCAG	keine	Sequenzierung
pET26b / pET32a	T7 P	vorwärts	20	TAATACGACTCACTATAGGG	keine	Sequenzierung
	T7 T	rückwärts	19	GCTAGTTATTGCTCAGCGG	keine	Sequenzierung

Abbildung 44: Verwendete Primer

Die in dieser Arbeit verwendeten Primer sind tabellarisch aufgeführt. Orientierung, Länge, Nukleotidsequenzen, Restriktionsenzymschnittstellen (RES) und Besonderheiten der einzelnen Primer sind angegeben. Kleingeschriebene Buchstaben in der Nukleotidsequenz zeigen die Abweichung der Primer von der Konsensussequenz an. Rot markiert sind die Restriktionsenzymschnittstellen. Die mutierten Tripletts in den Mutationsprimern sind grün hervorgehoben. A = Adenin, C = Cytosin, G = Guanin, T = Thymin. m = amino (A oder C), y = pyrimidin (C oder T), k = keto (G oder T), r = purin (A oder G). T7 P = T7 Promotor Primer, T7 T = T7 Terminator Primer.

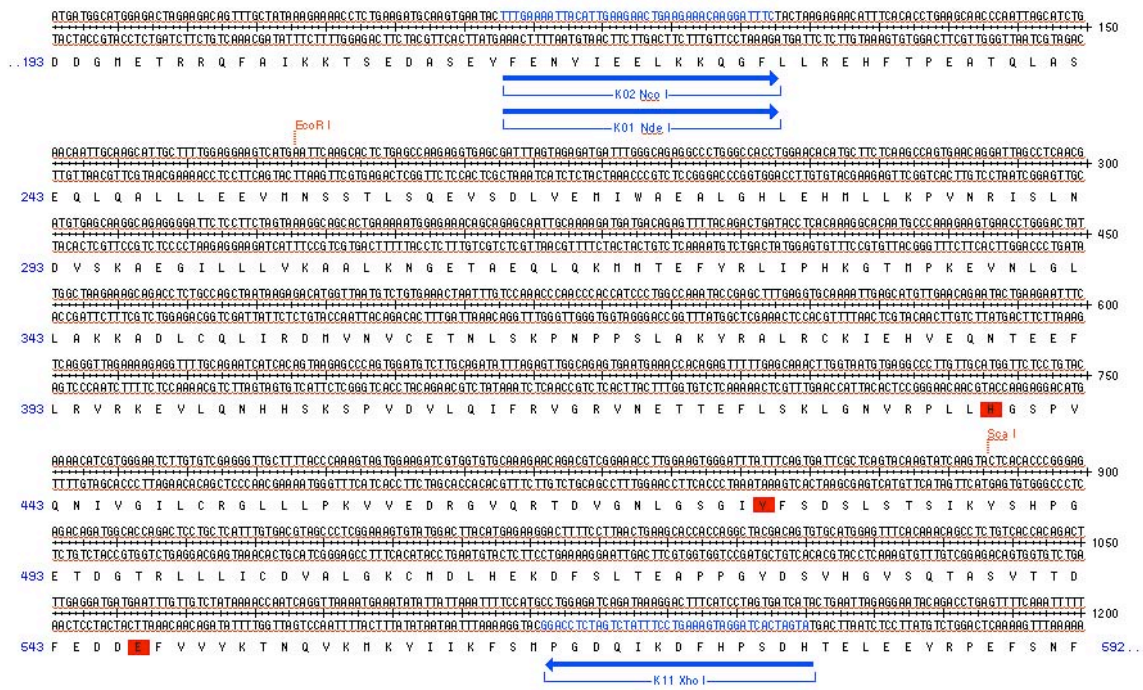
B Konsensussequenzen

- **hpART3**



- **hpART1**

- **hpART4**



- **SpvB**

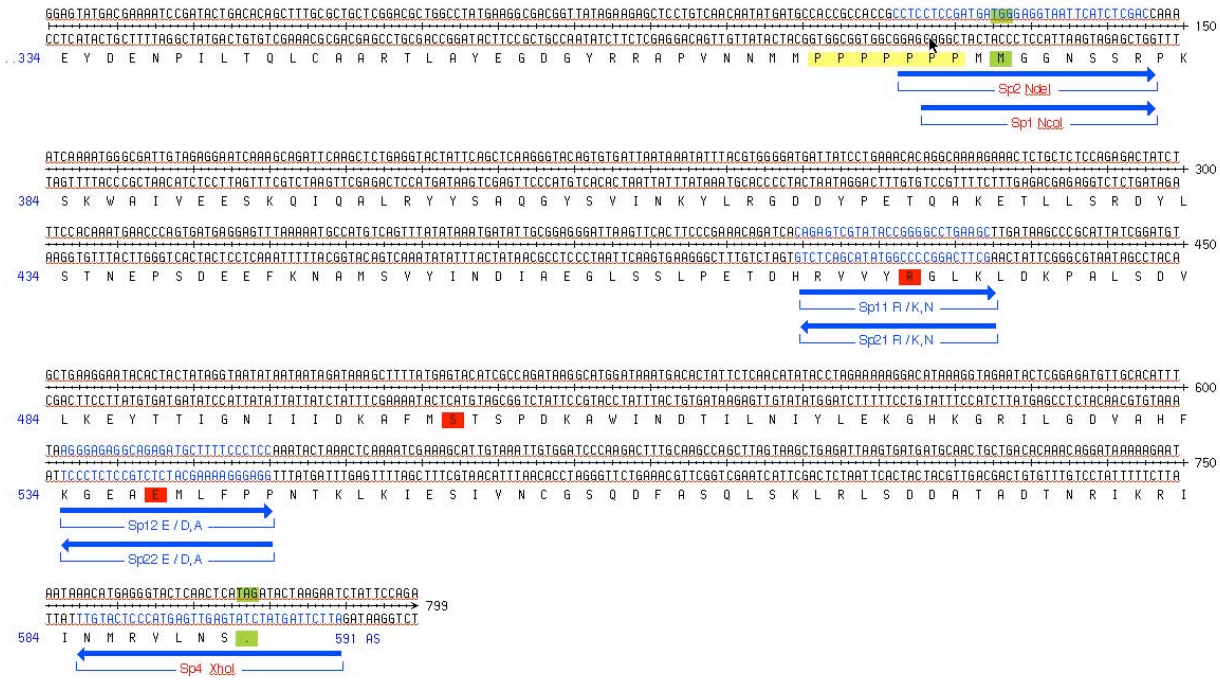
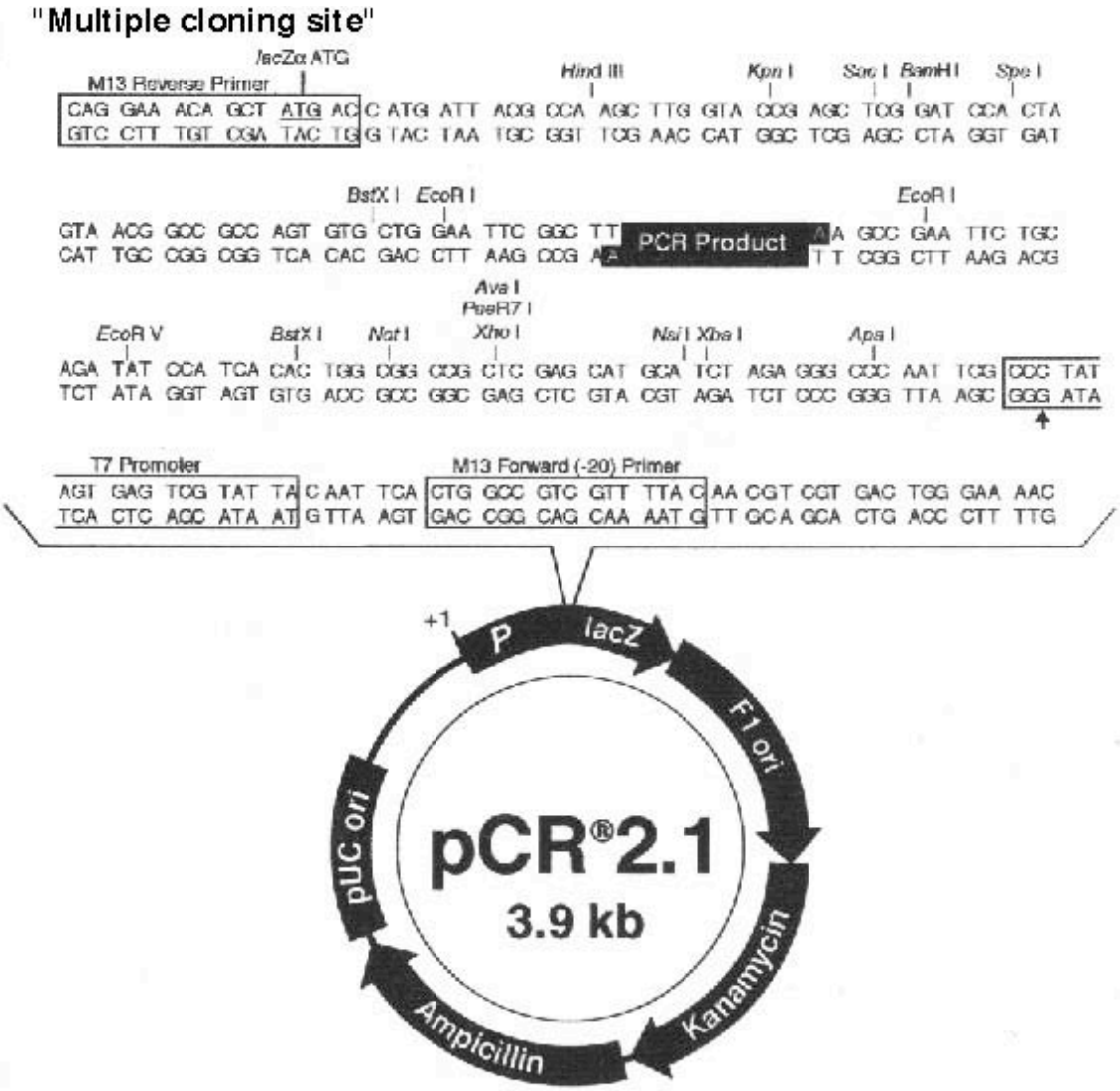


Abbildung 45: Konsensussequenzen der humanen pARTs 1, 3 und 4 und des Salmonellentoxins SpvB
Die Abbildung zeigt die Nukleotid- und translatierten Aminosäuresequenzen der humanen pARTs 1, 3 und 4 und des Salmonellentoxins SpvB. Bei pART3 ist die vollständige cDNA-Sequenz angegeben. Bei den anderen ARTs ist lediglich der Teil der Sequenz gezeigt, der die kodierende Region für die katalytische Domäne enthält. Rechts ist die Anzahl der Nukleotide des gezeigten Ausschnitts (schwarz), links die Position der Aminosäuren im Bezug auf das Gesamtprotein angegeben (hellblau). Die Lage der verwendeten Primer (blau) und die Schnittstellen der Restriktionsenzyme, die für die Klonierung benutzt wurden (rot), sind eingezeichnet. Das HYE bzw. RSE-Motiv ist rot, Start- und Stopcodons sind grün markiert. Zudem ist bei der Sequenz des Salmonellentoxins SpvB der Polyprolinabschnitt N-terminal der katalytischen Domäne gelb hervorgehoben. Genauere Informationen über die Primer sind in Abbildung 44 angegeben.

C Vektorkarten

- **pCR2.1**

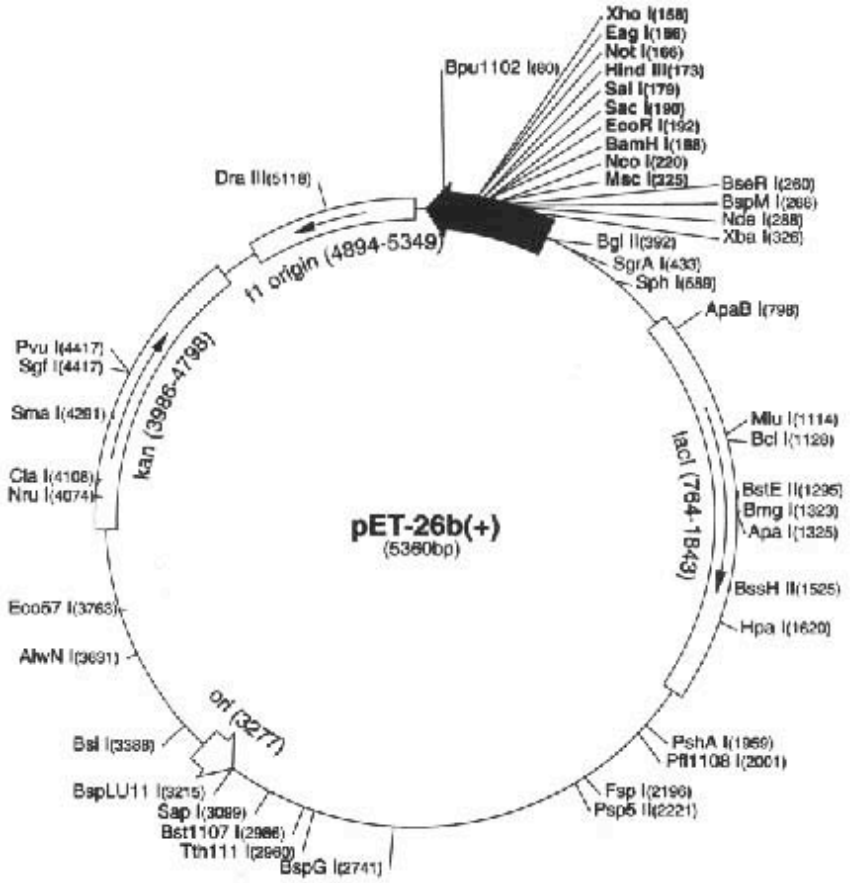


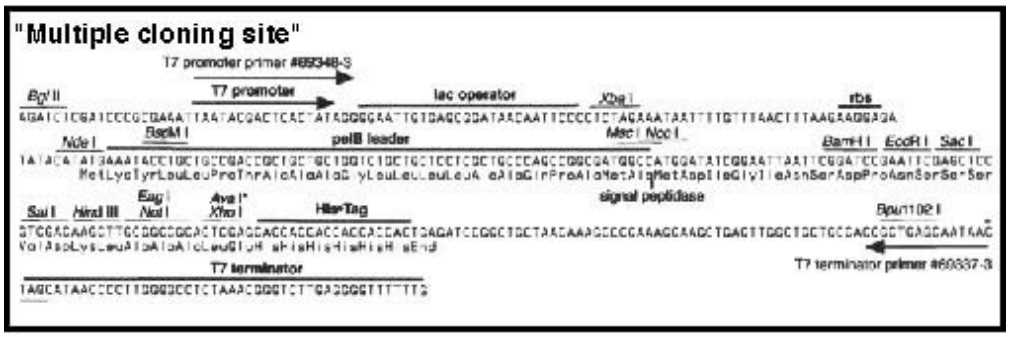
Kommentar zu pCR2.1:

3929 Nucleotide	**Basen**
lacZα Gen	1-545
M13 "reverse priming site"	205-221
T7-Promotor	326-381
M13 (-20) "forward priming site"	389-404
f1 origin	546-983
Kanamycin-Resistenz orf	1317-2111
Ampicillin-Resistenz orf	2129-2989
pUC origin	3134-3807

orf = offenes Leseraster

- **pET26b**



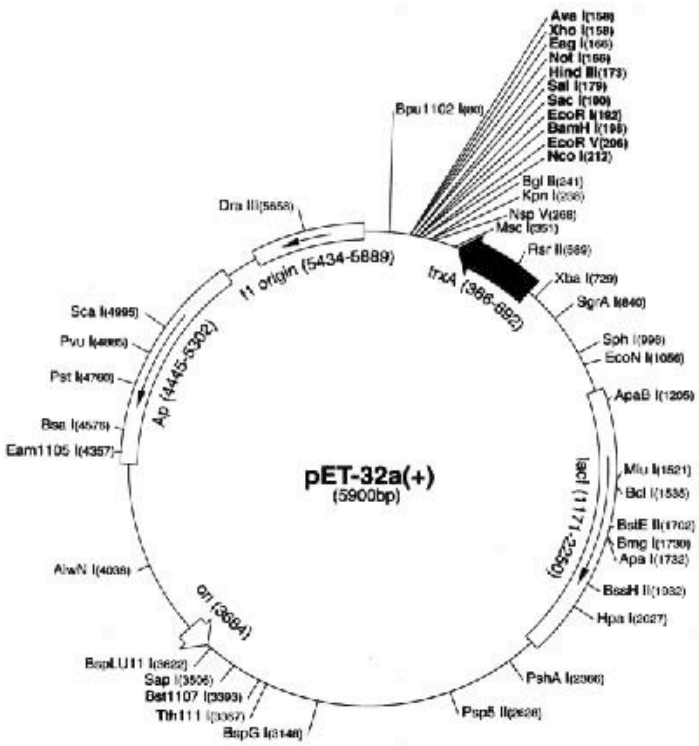


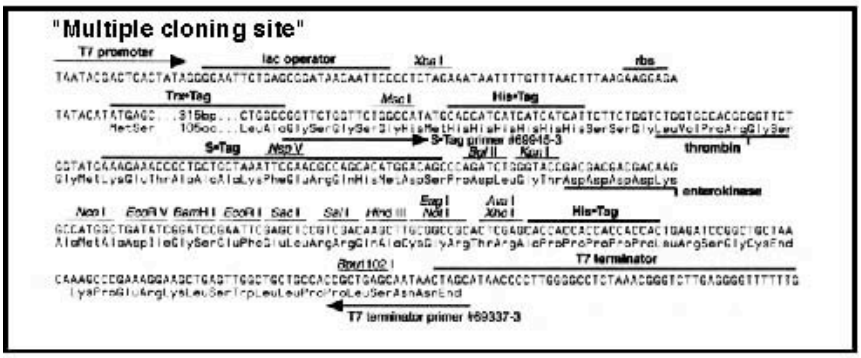
Kommentar zu pET26b:

5360 Nucleotide	**Basen**
T7-Promotor	361-377
T7-Transkriptionsstart	360
pelB cds	224-289
Multiple cloning site (NcoI-XhoI)	158-225
HisTag cds	140-157
T7-Terminator	26-72
lacI cds	764-1843
pBR322 origin	3277
Kanamycin-Resistenz orf	3986-4798
f1 origin	4894-5349

cds = kodierende Sequenz / orf = offenes Leseraster

- **pET32a**





Kommentar zu pET32a:

5900 Nucleotide	**Basen**
T7-Promotor	764-780
T7-Transkriptionsstart	763
Trx Tag cds	366-692
His Tag cds	327-344
S Tag cds	249-293
Multiple cloning site (NcoI-XhoI)	158-217
HisTag cds	140-157
T7-Terminator	26-72
lacI cds	1171-2250
pBR322 origin	3684
Ampicillin-Resistenz orf	4445-5302
f1 origin	5434-5889

cds = kodierende Sequenz / orf = offenes Leseraster

Abbildung 46: Vektorkarten (pCR2.1, pET26b und pET32a)

Die Vektorkarten wurden aus den Produktinformationen der Hersteller kopiert und für diese Abbildung modifiziert. pCR2.1: Invitrogen (www.invitrogen.com/content.cfm?pageid=3449), pET26b: Novagen (www.emdbiosciences.com/product/TB071) , pET32a: Novagen (www.emdbiosciences.com/product/TB122).

D pART-Subgruppen-Alignments

Subgruppe 1

```
                                                        ß1
1GS0   HHHHHHHHTTTTT  SEEEEE EEEEEEETTHHHHS --TT S EEEEEEE  GGGHHHH
1a26   -----------------------EEEEEETTHHHHHGGGGGSS EEEEEEE  GGGHHHH
1a26   -----------------------IFRIEREGESQRYKPFKQLHNRQLLWHGSRTTNFAGI
hs1    -----------------------IFKIEREGECQRYKPFKQLHNRRLLWHGSRTTNFAGI
mm1    -----------------------IFKIEREGESQRYKPFRQLHNRRLLWHGSRTTNFAGI
hs2    VISQYLQSTHAPTHSDYTMTLLDLFEVEKDGEKEAFR--EDLHNRMLLWHGSRMSNWVGI
mm2    VISQYLQSTHAPTHKDYTMTLLDVFEVEKEGEKEAFR--EDLPNRMLLWHGSRLSNWVGI
hs3    VIQTYLEQT---GSNHRCPTLQHIWKVNQEGEEDRFQAHSKLGNRKLLWHGTNMAVVAAI
mm3    AIQTYLKQT---GNSYRCPDLRHVWKVNREGEGDRFQAHSKLGNRRLLWHGTNVAVVAAI
hs4    --------------SKSPVDVLQIFRVGRVNETTEFL--SKLGNVRPLLHGSPVQNIVGI
mm4    --------------SEQPVDILQIFRVGRVNEATEFL--SKLGNVRLLFHGSPVRNILGI
cons                          ::.: : .*   :    .* *   * **:      .*
pred   -----------------------EEEEEECCHHHHHHHHHCCCEEEEEECCHHHHHHH

                                 ß2     a2                 ß3
1GS0   HHH S---    SSS GGG--TTTSSSEEEESSHHHHHGGG   SS--S EEEEEEEEEE
1a26   HHH S---    TTTTGGG--TTT SSEEEESSHHHHHTTS   SS--S EEEEEEEEEE
1a26   LSQGL---RIAPPEAPVTG--YMFGKGIYFADMVSKSANYCHTSQA--DPIGLILLGEVA
hs1    LSQGL---RIAPPEAPVTG--YMFGKGIYFADMVSKSANYCHTSQG--DPIGLILLGEVA
mm1    LSQGL---RIAPPEAPVTG--YMFGKGIYFADMVSKSANYCHTSQG--DPIGLILLGEVA
hs2    LSHGL---RIAHPEAPITG--YMFGKGIYFADMSSKSANYCFASRL--KNTGLLLLSEVA
mm2    LSHGL---RVAPPEAPITG--YMFGKGIYFADMSSKSANYCFASRL--KNTGLLLLSEVA
hs3    LTSGL---RIMPHSG------GRVGKGIYFASENSKSAGYVIGMKCGAHHVGYMFLGEVA
mm3    LTSGL---RIMPHSG------GRVGKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGEVA
hs4    LCRGLLLPKVVEDRGVQRTDVGNLGSGIYFSDSLSTSIKYSHPGET--DGTRLLLICDVA
mm4    LSRGLLLPKVAEDRGVQRTDVGNLGSGIYFSDSLSTSIKYAHAGET--DGSRLLVVCDVA
cons   *  **   ::    .        .*.***:.  *.*  *    .   .    :.: :**
pred   HHHCC---CCCCCCCCCCC--CCCCCCEEECCCHHHHHHHHCCCCC--CCEEEEEEEEEE

                            ß4
1GS0     SEEEESS  TTGGGGTTT SEEEE  SEE  GGG EEET--T--EEE  S  EE -
1a26     SEEEESS    SS-  TT  EEEE BSEEE TTT EEET--T--EEE    EEE -S
1a26   LGNMYELKNASHITK-LPKGKHSVKGLGKTAPDPTATTTLD--G--VEVPLGNGIST-GI
hs1    LGNMYELKHASHISK-LPKGKHSVKGLGKTTPDPSANISLD--G--VDVPLGTGISS-GV
mm1    LGNMYELKHASHISK-LPKGKHSVKGLGKTTPDPSASITLE--G--VEVPLGTGIPS-GV
hs2    LGQCNELLEANPKAEGLLQGKHSTKGLGKMAPSSAHFVTLN--G--STVPLGPASDT-GI
mm2    LGQCNELLEANPKAQGLLRGKHSTKGMGKMAPSPAHFITLN--G--STVPLGPASDT-GI
hs3    LGREHHINTDNPSLKSPPPGFDSVIARGHTEPDPTQDTELELDGQQVVVPQGQPVPCPEF
mm3    LGKEHHITIDDPSLKSPPSGFDSVIARGQTEPDPAQDIELELDGQPVVVPQGPPVQCPSF
hs4    LGKCMDLHEKDFPLTEAPPGYDSVHGVSQTASVTTD------------------------
mm4    LGKCVNLFKKDFSLTEAPPGYDSVHGVSETTSVPTD------------------------
cons   **.  .:   .        * .*. . ..  . .:
pred   CCCEEEECCCCCCCC-CCCCCCEEEECEEEEECCCCCEEEC--C--EEECCCCEEEC-CC

                 ß5          ß6
1GS0     SSS   SB EEEES GGGEEE EEEEEEEEE-    ------
1a26   --- S SBSB EEEES GGGEEE EEEEEEEEE     ------
1a26   ---NDTCLLYNEYIVYDVAQVNLKYLLKLKFNYKTS--------
hs1    ---NDTSLLYNEYIVYDIAQVNLKYLLKLKFNFKTSLW------
mm1    ---NDTCLLYNEYIVYDIAQVNLKYLLKLKFNFKTSLW------
hs2    LNPDGYTLNYNEYIVYNPNQVRMRYLLKVQFNF-LQLW------
mm2    LNPEGYTLNYNEFIVYSPNQVRMRYLLKIQFNF-LQLW------
hs3    ---SSSTFSQSEYLIYQESQCRLRYLLEVHL-------------
mm3    ---KSSSFSQSEYLIYKESQCRLRYLLEIHL-------------
hs4    -------FEDDEFVVYKTNQVKMKYIIKFSMPGDQIKDFHPSDH
mm4    -------FQDDEFVVYKTNQVKMKYIVKFCTPGDQIKEFHPHEN
cons         :  .*:::*.  * .::*:::.
pred   ---CCCCEEECEEEECCCCCEEEEEEEEEEEEECCC--------
```

Subgruppe 2

```
                                        ß1
1a26   EEEEEETTHHHHHGGGGGS--------S EEEEEEE  GGGHHHHHHH S    TTTTGG
1a26   IFRIEREGESQRYKPFKQL--------HNRQLLWHGSRTTNFAGILSQGLRIAPPEAPVT
hs5    IQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSP--FINAIIHKGFD---ERHAYI
mm5    IQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSP--FINAIIHKGFD---ERHAYI
hs5    IQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSP--FVNAIIHKGFD---ERHAYI
mm5    IQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSP--FVNAIIHKGFD---ERHAYI
cons   **** **** **: **:********** **********--*:*********---******
pred   CCEEECHHHHHHHHHHHHHHHHHHCCCCCEEEEEECCC--HHHHHHHCCCC---CCCCCC
conf   903206679999999998999873798753778744 70--03788844888---466778

               ß2      α2                              ß3
1a26   GTTT SSEEEESSHHHHHTTS---------   SSS---- EEEEEEEEEE  SEEEESS
1a26   GYMFGKGIYFADMVSKSANYC---------HTSQAD----PIGLILLGEVALGNMYELKN
hs5    GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFS
mm5    GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFS
hs5    GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFS
mm5    GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPIHKDRSCYICHRQLLFCRVTLGKSFLQFS
cons   ********************************** ************:***************
pred   CCCCCCCEEEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEEEEEEECCCCEEECCC
conf   881238555412442200000256655445543344443333168998860176121157

                    ß4                                          ß5
1a26       SS  TT  EEEE BSEEE TTT EEETTEEE    EEE S  S SBSB EEEES G
1a26   ASHITKLPKGKHSVKGLGKTAPDPTATTTLDGVEVPLGNGISTGINDTCLLYNEYIVYDV
hs5    TMKMAHAPPGHHSVIGR------------------PSVNG---------LAYAEYVIYRG
mm5    TMKMAHAPPGHHSVIGR------------------PSVNG---------LAYAEYVIYRG
hs5    AMKMAHSPPGHHSVTGR------------------PSVNG---------LALAEYVIYRG
mm5    AMKMAHSPPGHHSVTGR------------------PSVNG---------LALAEYVIYRG
cons   :*****:******* **------------------*****---------** ********
pred   CCCCCCCCCCCCCCEEC------------------CCCCC---------CCCCEEEEECC
conf   74478888788752546------------------75788---------89870588768

            ß6
1a26   GGEEE EEEEEEEEE   ------------
1a26   AQVNLKYLLKLKFNYKTS------------
hs5    EQAYPEYLITYQIMKPEAPSQTATAAEQKT
mm5    EQAYPEYLITYQIMKPEAPSQTATAAEQKT
hs5    EQAYPEYLITYQIMRPEGMVDG--------
mm5    EQAYPEYLITYQIVRPEGMVDG--------
cons   *************::**.  :
pred   CEEEEEEEEEEEEECCCCCCCCCCCCCCCC
conf   652068989989822458876644110059
```

Subgruppe 3

```
                                          ß1
1a26   EEEEEETTHHHHHGGGGG--------SS EEEEEEE  GGGHHHHHHH S    TTTT
1a26   IFRIEREGESQRYKPFKQ--------LHNRQLLWHGSRTTNFAGILSQGLRIAPPEAP
hs7    IERIQNAFLWQSYQVKKRQMDIKNDHKNNERLLFHGTDADSVPYVNQHGFNR--SCAG
hs8    IERIQNPDLWNSYQAKKKTMDAKNGQTMNEKQLFHGTDAGSVPHVNRNGFNR--SYAG
mm8    IERIQNPALWRRYQAYKKGMDEKNGNVRNEKHLFHGTEASSLPQLNSNGFNR--SYAG
hs9    VEKIDNEVLMAAFQRKKKMMEEKLHRQPVSHRLFQQVPYQFCNVVCRVGFQR--MYST
mm9    VEQIDNKVLLAAFQEKKKMMEERTPKGSGSQRLFQQVPHQFCNTVCRVGFHR--MYST
hs10   VERVSHPLLQQQYELYRERLLQRCERRPVEQVLYHGTTAPAVPDICAHGFNR--SFCG
mm10   VERVSHPLLQQQYQLHRERLMQSCQQRPVEQVLYHGTSESAVLDICAHGFNR--SFCG
cons   :*::.:  *   ::  :. :     .   .: *:: .       :   **:*--  .
pred   CCEECCHHHHHHHHHHHHHHHHHHCCCCCEEEEEECCCHHHHHHHHHHCCCC--CCCC
conf   91202878999999999999998728984306776058978999999826887--7578

                ß2    α2                ß3
1a26   GGGTTT SSEEEESSHHHHHTTS  - SSS EEEEEEEEEE  SEEEESS ------
1a26   VTGYMFGKGIYFADMVSKSANYCHT-SQADPIGLILLGEVALGNMYELKNA------S
hs7    KNAVSYGKGTYFAVDASYSAKDTYSKPDSNGRKHMYVVRVLTGVFTKGRAGLVTPPPK
hs8    KNAVAYGKGTYFAVNANYSANDTYSRPDANGRKHVYYVRVLTGIYTHGNHSLIVPPSK
mm8    KNATAYGKGTYFAVKASYSACDTYSRPDTNGRKYMYYVRVLTGNYTNGNASLIVPPSR
hs9    PCDPKYGAGIYFTKNLKNLAEKAKKISAADKLIYVFEAEVLTGFFCQGHPLNIVPPPL
mm9    SYNPVYGAGIYFTKSLKNLADKVKKTSSTDKLIYVFEAEVLTGSFCQGNSSNIIPPPL
hs10   RNATVYGKGVYFAKRASLSVQDRYSPPNADGHKAVFVARVLTGDYGQGRRGLRAPPLR
mm10   RNGTLYGQGVYFAKRASLSVLDRYSPPNAEGYKAVFVAQVLTGDYGQGSRGLKAPPLR
cons        ** * **:   .  . .  . . ::    :: ..**** : :*      **
pred   CCCCCCCCEEEECCCCCCHHHCCCCCCCCCCCEEEEEEEEECCCCCCCCCCCCCCCCC
conf   6777125404531432101101035647887479998753103424787130116767

                ß4                                          ß5
1a26     SS  TT  EEEE BSEEE TTT EEETTEEE    EEE S  S SBSB EEEES G
1a26   HITKLPKGKHSVKGLGKTAPDPTATTTLDGVEVPLGNGISTGINDTCLLYNEYIVYDV
hs7    NPHNPTDLFDSVTN-------------------------------NTRSPKLFVVFFD
hs8    NPQNPTDLYDTVTD-------------------------------NVHHPSLFVAFYD
mm8    DPQNAADLYDTVTD-------------------------------NDKNPSIFVVFYD
hs9    SP-GAIDGHDSVVD-------------------------------NVSSPETFVIFSG
mm9    SP-GALDVNDSVVD-------------------------------NVSSPETIVVFNG
hs10   GPGHVLLRYDSAVD-------------------------------CICQPSIFVIFHD
mm10   VSG-QVLRYDSAVD-------------------------------CLQQPRIFVIFHD
cons    .       *:..:-------------------------------    *  :* * .
pred   CCCCCCCCCCCCCC-------------------------------CCCCCCEEEEECC
conf   77888876666577-------------------------------8888833788769

            ß6
1a26   GGEEE EEEEEEEEE   ---------------------------------
1a26   AQVNLKYLLKLKFNYKTS---------------------------------
hs7    NQAYPEYLITFTA-------------------------------------
hs8    YQAYPEYLITFRK-------------------------------------
mm8    NQTYPEYLITFRQ-------------------------------------
hs9    MQAIPQYLWTCTQEY-------VQSQDYSSGPMRPFAQHPWRGFASGSPVD
mm9    MQAMPLYLWTCTQDR-TFSQHPMWSQGYSSGPGMVSSLQSWEWVLNGSSV-
hs10   TQALPTHLITCEHVPRASPDDPSGLPGRSPDT-------------------
mm10   TQALPTHLITCKNILPGTP--------------------------------
cons    *: * :* *
pred   CEECEEEEEEEEC-------------------------------------
conf   6733289777749-------------------------------------
```

Subgruppe 4

```
1a26  EEE  TTSHHHHHHHHHHHHT  STTTTEEEEE EEEEEEETTHHHHHGGGGGSS------
1a26  IKVVDKDSEEAKIIKQYVKNTHAATHNAYDLKVVEIFRIEREGESQRYKPFKQLH------
hs11  LIPLHNQTHEYNEVANLFGKTMDRNR------IKRIQRIQNLDLWEFFCRKKAQLKKKRGV
mm11  LVSLQNQTHEYNEVASLFGKTMDRNR------IKRIQRIQNLDLWEFFCRKKAQLKKKRGV
hs12  KITLSSSSEEYQKVWNLFNRTLPFYF------VQKIERVQNLALWEVYQWQKGQMQKQNGG
mm12  KITLSSSSEEYQKVWNIFNRTLPFYF------VQKIERIQNMGLWEVYQWQKCQMQKQNGG
hs13  LSEIHHLHPEYVRVSEHFKASMKNFK------IEKIKKIENSELLDKFTWKKSQMKEE---
mm13  FIELNNQDEEYAKISEQFKASMKQFK------IVTIKRIWNQKLWDTFERKKQKMKNK---
hs14  QVPVSAEDKSYRIIYNLFHKTVPEFKYR----ILQILRVQNQFLWEKYKRKKEYMNRKMFG
mm14  QVPVSAEDKSYRIIYNLFHKTVPEFKYR----ILQILRVQNQFLWEKYKRKKEYMNRKMSG
cons     :     .*  : . *  ::       ----:  * :: *  * : :  :*  ::.:
pred  CCCCCCCCHHHHHHHHHHHHCCCCCE------EEEEEEECCHHHHHHHHHHHHHHHHHCC
conf  972677884789999998611798--33----68988704877889999999999987168

                ß1                                  ß2    α2
1a26  ----- EEEEEEE  GGGHHHHHHH S    TTTTGGGTTT SSEEEESSHHHHHTTS
1a26  -----NRQLLWHGSRTTNFAGILSQGLRIAPPEAPVTGYMFGKGIYFADMVSKSANYCHTS
hs11  PQ--INEQMLFHGTSSEFVEAICIHNFDWR--INGIHGAVFGKGTYFARDAAYSSRFCKDD
mm11  PQ--INEQMLFHGTSSEFVEAICIHNFDWR--INGVHGAVFGKGTYFARDAAYSSRFCKDD
hs12  KA--VDERQLFHGTSAIFVDAICQQNFDWR--VCGVHGTSYGKGSYFARDAAYSHHYSKSD
mm12  KE--VDERQLFHGTSANFVDAICQQNFDWR--VCGLHGTSYGKGSYFARDAAYSHHYSKSD
hs13  ------GKLLFYATSRAYVESICSNNFDSF--LHETHENKYGKGIYFAKDAIYSHKNCPYD
mm13  -----TEMFLFHAVGRIHMDYICKNNFEWI--LHGNREIRYGKGNYFTKEAMYSHKSCSYD
hs14  RDRIINERHLFHGTSQDVVDGICKHNFDPR--VCGKHATMFGQGSYFAKKASYSHNFSKKS
mm14  RDRIINERHLFHGTSQDVVDGICKHNFDPR--VCGKHATMFGQGSYFAKKASYSHNFSKKS
cons          **:...   :: ** :**:  --:   :   :*:* **::.* ** . .  .
pred  CC--CCCEEEECCCCHHHHHHHHHCCCCC--CCCCCCCCCCCCEEECCCHHHHHHCCCCC
conf  97--211532106887899999998178988--678778722560211001345540466 66

                                ß3
1a26  SSS ------------------EEEEEEEEEE  SEEEESS    SS ------- TT  -
1a26  QADP------------------IGLILLGEVALGNMYELKNASHITKL-------PKGKH-
hs11  IKHGNTFQIHGVSLQQRHLFRTYKSMFLARVLIGDYINGDSKYMRPPSKDGSY--VNLYD-
mm11  IKHGNTFQIHGVSLQQRHLFRTYKSMFLARVLIGDYINGDSKYMRPPSKDGSY--VNLYD-
hs12  TQ--------------------THTMFLARVLVGEFVRGNASFVRPPAKEGWS--NAFYD-
mm12  TH--------------------SHMMFLARVLVGDFVRGSTSFVRPPAKEGQS--NAFYD-
hs13  AK--------------------NVVMFVAQVLVGKFTEGNITYTSPPP---------QFD-
mm13  SR--------------------GTVMFVARVLVGSVIEGNMTLSSPPA---------LYD-
hs14  SKG-------------------VHFMFLAKVLTGRYTMGSHGMRRPPPVNPGSVTSDLYD-
mm14  SKG-------------------VHFMFLAKVLTGRYTMGSHGMRRPPPVNPGSVTSDLYD-
cons   :                       **:*:** *    *.     **.          :*-
pred  CCCCCCCCCCCCCCCCCCCCCCCCEEEEEEECCCCCCCCCCCCEECCCCCCCC--CCCCC-
conf  566654432222222121013651278852103563258855266689878 75--34457-

       ß4                                        ß5          ß6
1a26  EEEE BSEEE TTT EEETTEEE    EEE S  S SBSB EEEES GGGEEE EEEEEEE
1a26  SVKGLGKTAPDPTATTTLDGVEVPLGNGISTGINDTCLLYNEYIVYDVAQVNLKYLLKLKF
hs11  SCVD------------------------------DTWNPKIFVVFDANQIYPEYLIDFH-
mm11  SCVD------------------------------DTWNPKIFVVFDANQIYPEYLIDFH-
hs12  SCVN------------------------------SVSDPSIFVIFEKHQVYPEYVIQYTT
mm12  SCVN------------------------------SMSDPTIFVVFEKHQVYPEYLIQYST
hs13  SCVD------------------------------TRSNPSVFVIFQKDQVYPQYVIEYTE
mm13  SCVD------------------------------TRLNPSVFVIFRKEQIYPEYVIEYME
hs14  SCVD------------------------------NFFEPQIFVIFNDDQSYPYFVIQYEE
mm14  SCVD------------------------------NFFEPQIFVIFNDDQSYPYFVIQYEE
cons  ***:------------------------------   :* :**:*  .* ** ::*::
pred  CCCC------------------------------CCCCCEEEEEECCCEEEEEEEEEEC-
conf  7567------------------------------865870589884857413787 7758-

1a26  EE   -----------------
1a26  NYKTS-----------------
hs11  ----------------------
mm11  ----------------------
hs12  SSKPSVTPSILLALGSLFSSRQ
mm12  SSKPPASPSIFVALGNLFTSRQ
hs13  DKACV------------IS---
mm13  LEKEK---------GCIIS---
hs14  VSNTVSI---------------
mm14  VSNTVSI---------------
cons
pred  ----------------------
conf  ----------------------
```

Subgruppe 5

```
                                                          ß1
1a26   HHHHHHHHHHT  STTTTEEEEE EEEEEEE-TTHH-HHHGGGGGSS EEEEEEE  GGGH
1a26   AKIIKQYVKNTHAATHNAYDLKVVEIFRIER-EGES-QRYKPFKQLHNRQLLWHGSRTTNF
hs15   FEKIQKLTGAPHTPV-PAPDF----LFEIEYFDPAN-AKFYETKGERDLIYAFHGSRLENF
mm15   FEKIQQLTGAPHTPV-PTPDF----LFEIEYFDPAN-SRFYETKGERDLIYAFHGSRLENF
hs16   QLKFM---HTPHQ-------------FLLLSSPPAKESNFRAAKKLFGSTFAFHGSHIENW
mm16   QLKFM---HTPHQ-------------FLLLSSPPAKESNFRAAKKLFGSTFAFHGSHIENW
hs17   -LKFM---HTSHQ-------------FLLLSSPPAKEARFRTAKKLYGSTFAFHGSHIENW
mm17   -LKFM---HTSHQ-------------FLLLSSPPAKEARFRTAKKLYGSTFAFHGSHIENW
cons     *:     :.*   -     ---- * :    **: :.*  :*   .  :**:**::**:
pred1  CCHHHHHHCCCCCCC-CCCCC----EEEEEEECCCC-HHHHHHHCCCCCEEEECCCCCHHH
conf1  911432006798877-88410----47787605323-57897511578467760762145 7
pred2  CCCCC---CCCEE-------------EEEECCCCHHHHHHHHHHCCCCCEEEEECCCHHHH
conf2  98765---77314-------------77840782111211011057884588626625667

                                 ß2    α2
1a26   HHHHHH S    T-TTTGGGTTT SSEEEESSHHHHHTTS   SSS --------------
1a26   AGILSQGLRIAPP-EAPVTGYMFGKGIYFADMVSKSANYCHTSQADP--------------
hs15   HSIIHNGLHCH---LNKT--SLFGEGTYLTSDLSLALIYSPHGHGWQ--------------
mm15   HSIIHNGLHCH---LNKT--SLFGEGTYLTSDLSLALIYSPHGHGWQ--------------
hs16   HSILRNGLVVASNTRLQLHGAMYGSGIYLSPMSSISFGYSGMNKKQKVSA-KDEPASSSKS
mm16   HSILRNGLVVASNTRLQLHGAMYGSGIYLSPMSSISFGYSGMNKKQKVSS-KDEPASSSKS
hs17   HSILRNGLVNASYTKLQLHGAAYGKGIYLSPISSISFGYSGMGKGQHRMPSKDELVQRYNR
mm17   HSILRNGLVNASYTKLQLHGAAYGKGIYLSPISSISFGYSGMGKGQHRMPSKDELVQRYNR
cons   ***::***        :   : :*.* **:.  *::: **  .:  :
pred1  HHHHHHHHEEE---ECCC--EEEEEEEEECHHHHHHHHCCCCCCCCC--------------
conf1  99875201024---2051--78863477332230131014899875 4--------------
pred2  HHHHHHCCCCCCCCCCCCCCCCCCCCEEEECCCHHHHHHHHCCCCCCCCCH-HCCCCCCCCC
conf2  77773173417788533456510050575350333321000367871110-0000000034

                          ß3                            ß4
1a26   ----------------EEEEEEEEEE  SEEEESS    SS  TT  EEEE BSEEE TTT
1a26   ----------------IGLILLGEVALGNMYELKNASHITKLPKGKHSVKGLGKTAPDPTA
hs15   HSLLGPILSCVAVCEVIDHPDVKCQTKKKDSKE------------IDRRRARIKH------
mm15   HSLLGPILSCVAVCEVIDHPDVKCQIKKKDSKE------------IDRSRARIKH------
hs16   SNT-SQSQKKGQQSQFLQSRNLKCIALCEVITS------------SDLHKHGE--------
mm16   SNA-SQSQKKGQQSQFLQSRNLKCIALCEVITS------------PDLHKHGE--------
hs17   MNTIPQTRSIQ--SRFLQSRNLNCIALCEVITS------------KDLQKHGN--------
mm17   MNTIPQTRSIQ--SRFLQSRNLNCIALCEVITS------------KDLQKHGN--------
cons    .      .    ...::  :::*    :  ..             *  :
pred1  HHHHHHHHHHHHHHHHCCCCCCCEEECCCCCCC------------CCCCEEEEEE------
conf1  013322243311001207876301100102400------------1442113440------
pred2  CCC-CCCCCCCCCCCCCCCCCEEEEEEECCCC------------CCCEECCC--------
conf2  222-4434677320101477610344332017 8------------66254351--------

                                   ß5          ß6
1a26    EEETTEEE    EEE S  S SBSB EEEES GGGEEE EEEEEEEEE   ---------
1a26   TTTLDGVEVPLGNGISTGINDTCLLYNEYIVYDVAQVNLKYLLKLKFNYKTS---------
hs15   ------------------SEGGDIPPKYFVVTNNQLLRVKYLLVYSQKPPKRAPSQLSWFS
mm15   ------------------SEGGEIPPKYFVVTNNQLLRVKYLLVYSQKQPKRASSQLSWLS
hs16   ----------------------------IWVVPNTDHVCTRFFFVYEDGQVGDAN-------
mm16   ----------------------------IWVVPNTDHVCTRFFFVYEDGQVGDAN-------
hs17   ----------------------------IWVCPVSDHVCTRFFFVYEDGQVGDAN-------
mm17   ----------------------------IWVCPVSDHVCTRFFFVYEDGQVGDAN-------
cons                               :* . .: : .::::**.:     *
pred1  ------------------CCCCCCCCCEEEECCCCEEEEEEEEEECCCCCCCCCCCCCEEE
conf1  ------------------689988840488707837888888876367775775778863012
pred2  ----------------------------EEEEECCCCCCCEEEEEEECCCCCCCC-------
conf2  ----------------------------68961676424026898607811345 2-------

1a26   ------------------------------------
1a26   ------------------------------------
hs15   SHWFTVMISLYLLLLLIVSVINSSAFQHFWNRAKR
mm15   SHWFVIMMSLYLLLLLIVSVTNSSAFHHFWNRVKR
hs16   -----INTQEGGIHKEILRVIGDQTATG-------
mm16   -----INTQEGGIHKEILRVIGNQTATG-------
hs17   -----INTQDPKIQKEIMRVIGTQVYTN-------
mm17   -----INTQDPKIQKEIMRVIGTQVYTN-------
cons       :  .   :   *: * . ..
pred1  EECCHHHHHHHHHHHHHHHCCCHHHHHHHHHCC
conf1  10000235667776665421143578989987439
pred2  -----CCCCCCHHHHHHHHHHCCCCCC-------
conf2  -----4345733678888875000036 9-------
```

Abbildung 47: Aminosäuresequenz-Alignments der humanen pART-Subgruppen 1-5

Den hier gezeigten Alignments der pART-Subgruppen liegen die Ergebnisse der PSI-BLAST-, T-Coffee-, PSIPRED- und GeneTHREADER-Analysen zu Grunde. Oberhalb der Alignments sind die bekannten Sekundärstrukturmotive aus der 3D-Struktur der Hühner-PARP (1a26) angegeben (grün). Die Sekundärstrukturmotive sind durch die folgenden Buchstaben abgekürzt: E = β-Strang, H = α-Helix, G = 3(10)-Helix, I = Pi-Helix, B = Rest in einer isolierten β-Brücke, T = durch Wasserstoffbrücken-Bindungen stabilisierte Schleife, S = Schleife. Die Lage der charakteristischen Sekundärstrukturmotive ist hervorgehoben (türkis bzw. in der Aminosäuresequenz unterstrichen). Für die Subgruppe 1 sind zusätzlich die Sekundärstukturmotive der 3D-Struktur der murinen pART2 (PARP-2, 1GS0) aufgeführt. Positionen, mit identischen Aminosäurenresten bei allen pARTs einer Subgruppe sind durch Sterne, Ähnlichkeiten durch Doppelpunkte und Punkte gekennzeichnet. Das charakteristische HYE-Motiv bzw. seine Varianten sind rot markiert. Die Positionen der Introns sind in das Alignment projiziert (Grau = Phase-0-Intron, blau = Phase-1-Intron, gelb = Phase-2-Intron). Für jeweils ein Mitglied jeder Subgruppe bzw. im Falle von Subgruppe 5 für zwei Mitglieder ist die mit PSIPRED erhaltene Vorhersage der Sekundärstruktur unterhalb des Alignments dargestellt (Blau /Subgruppe 1: hpART1, Subgruppe 2: hpART5, Subgruppe 3: hpART7, Subgruppe 4: hpART11, Subgruppe 5: hpART15 und hpART16). Die vorhergesagten Sekundärstrukturmotive sind durch die folgenden Buchstaben abgekürzt: E = β-Strang, H = α-Helix, C = „coil“ (Bereich ohne definierte Sekundärstruktur). Der Zuverlässigkeitsgrad der Vorhersage (confidence, höchste Sicherheit = 9) ist ebenfalls angegeben (gelb). hs = homo sapiens, mm = mus musculus.

9 Danksagung

Mein besonderer Dank gilt Prof. Dr. F. Koch-Nolte für die hervorragende Betreuung, die Anregungen und die ständige Bereitschaft zur Diskussion in allen Phasen der Arbeit.

Prof. Dr. F. Koch Nolte, Pedro Reche (DNAX Research Institute) und Fernando Bazan (DNAX Research Institute, Palo Alto, Kalifornien, USA) danke ich für ihre Vorarbeiten zum bioinformatischen Teil der Arbeit. Michael Rhen und Dilek Teczan Merdol (Karolinska Institut, Stockholm, Schweden) danke ich für die Zusendung der SpvB-Matrizen und ihre Kooperation beim SpvB-Projekt.

Katrin Bartels, Roman Girisch, Maren Kühn und Marion Nissen danke ich für die gute Einarbeitung in die angewandten Techniken und die praktische Unterstützung beim laborexperimentellen Teil dieser Arbeit. Im Speziellen danke ich Maren Kühn für die Durchführung der „Northern-Blots“ (Abb.21) und ihre Hilfe beim Chromosomen-„Mapping“ (Abb.32), Roman Girisch für die Hilfe bei der Etablierung der Proteinexpression und die Durchführung der Enzymassays (Abb.28, 29 und 38-41) und Katrin Bartels für die Durchführung des SpvB-Aktin-Enzymassays (Abb.42).

Dr. Gustavo Glowacki und Marion Nissen danke ich für die stetige Hilfsbereitschaft, ihren Rat bei laborexperimentellen Problemen und ihre Freundschaft. PD Dr. Friedrich Haag und Dr. Stefan Rothenburg danke ich für die Anregungen und Diskussionen während des Entstehens dieser Arbeit.

Bei allen Mitgliedern des Forschungslabors und des Diagnostiklabors des Institutes für Immunologie des Universitätsklinikums Hamburg Eppendorf bedanke ich mich für die kollegiale Arbeitsatmosphäre und eine inspirierende Zeit im Labor.

Der „Studienstiftung des deutschen Volkes“ danke ich für die finanzielle Unterstützung.

Meiner Familie und meinen Freunden danke ich für ihre Unterstützung und ihr Verständnis für die durch diese Arbeit entstandenen Entbehrungen.

10 Lebenslauf

Name: Otto

Vorname: Helge

Geburtsdatum: 06.01.1975

Adresse: Charlottenstr.19, 20257 Hamburg

e-mail: helge.otto@t-online.de

Geburtsort: Kassel

Familienstand: ledig

Ausbildung

1981-1985: Grundschule Lohfelden

1985-1991: Offene Schule Waldau, Kassel

1991-1994: Gymnasium Herderschule, Kassel

1994-2001: Studium der Humanmedizin, Universität Hamburg

03/2002-08/2003: Arzt im Praktikum bei Dr. D. Braumann, II.Medizin, Abt. für Hämatologie / Onkologie, AK Altona, Hamburg

seit 11/2003: Assistenzarzt bei Prof. Dr. F. Hagenmüller, I.Medizin, Abt. für Gastroenterologie, AK Altona, Hamburg

Stipendien

1994-2001 Stipendiat der „Studienstiftung des Deutschen Volkes“

Dissertation

seit 1998: Thema „Molekulare Klonierung und Charakterisierung neuer ADP-Ribosyltransferasen bei *Homo sapiens* und *Salmonella typhimurium*“ unter der Leitung von Prof. Dr. med. F. Koch-Nolte, Inst. für Immunologie, Universitätsklinik Hamburg-Eppendorf

11 Publikationen

Teile dieser Arbeit wurden in folgenden wissenschaftlichen Zeitschriften publiziert bzw. zur Publikation eingereicht:

- Helge Otto, Dilek Tezcan-Merdol, Roman Girisch, Friedrich Haag, Mikael Rhen and Friedrich Koch-Nolte (2000) "The SpvB gene-product of the *Salmonella enterica* virulence plasmid is a mono(ADP-ribosyl)-transferase" *Molecular Microbiology* **37:** 1106-1115

- Helge Otto, Pedro A. Reche, Fernando Bazan, Katharina Dittmar, Friedrich Haag, Friedrich Koch-Nolte (2005) "In silico characterization of the family of PARP-like poly-(ADP-ribosyl)transferases (pARTs)" (*submitted for publication 05/2005*)

Teile dieser Arbeit wurden bei folgenden Kongressen in Form von Postern und Abstracts publiziert:

- Helge Otto, Marion Nissen, Pedro Reche, Fernando Bazan, Friedrich Haag, Friedrich Koch-Nolte (1999) "PLIRT1, a novel PARP-like intracellular ADP-ribosyltransferase" Abstract / Poster, Herbsttagung Gesellschaft für Biochemie und Molekularbiologie e.V., Hamburg 1999

- Helge Otto, Marion Nissen, Sohila Nazem Fernando Bazan, Friedrich Haag, Friedrich Koch-Nolte (2000) "A family of human genes encoding putative PARP-like intracellular ADP-ribosyltransferases" Abstract / Poster, Herbsttagung Gesellschaft für Biochemie und Molekularbiologie e.V., München 2000. *Biological Chemistry* **381**: 234

- Helge Otto, Marion Nissen, Gustavo Glowacki, Katrin Bartels, Pedro Reche, Fernando Bazan, Friedrich Haag, Friedrich Koch-Nolte (2001) „Common and distinctive features of the gene families encoding PARP-like intracellular poly-ADP-ribosyltransferases and RT6-like extracellular mono-ADP-ribosyltransferases" Abstract / Poster, ADPR 2001, 13th International Symposium on ADP-Ribosylation Reactions, New York City 2001

12 Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Hamburg, den 20.06.2005

Helge Otto