

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Ein Protein für neue Aufgaben:
die cytosolische PH-Domäne des Cytohesin-1
als Paratop und als Substrat für Translokationen**

von Hartmut Volker Rohde
aus Aachen

2001

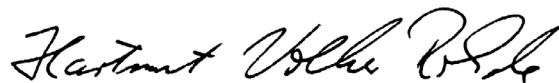
Erklärung

Diese Dissertation wurde im Sinne von §13 Absatz 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Professor Dr. Ernst-Ludwig Winnacker betreut.

Ehrenwörtliche Versicherung

Hiermit versichere ich ehrenwörtlich, daß ich diese Dissertation selbständig und ohne unerlaubte Hilfe anfertigte.

München, am 12.11.2001



(Hartmut Volker Rohde)

Dissertation eingereicht am 12.11.2001

Mündliche Prüfung am 20.12.2001

Erster Gutachter Professor Dr. Ernst-Ludwig Winnacker

Zweiter Gutachter Professor Dr. Patrick Cramer

Teile dieser Arbeit wurden oder werden veröffentlicht:

Rohde H.V., Bruhn H. & Steipe B. (2001), Reengineering of Cytosolic PH domains for Translocation, *Nova Acta Leopoldina Supp.* 16, 133-134.

Rohde H.V., Brunner M. & Steipe B. (2001), Translocation of PH domain variants into the periplasm of *E. coli* and yeast mitochondria, *Biol. Chem. Supp.* 382, S167.

Rohde H.V. & Steipe B. (2002), Protein-engineering of the PH domain of Cytohesin-1: Optimization of periplasmic expression in *E. coli* through circular permutation, *Biol. Chem.*, angenommen.

Die vorliegende Arbeit wurde in der Zeit vom Mai 1997 bis Dezember 2001 unter der Anleitung von Herrn Professor Dr. Boris Steipe am Genzentrum der Ludwig-Maximilians-Universität München durchgeführt.

Ich bedanke mich bei...

Herrn Professor Dr. Boris Steipe für das interessante Dissertationsthema, die Diskussionen und wertvollen wissenschaftlichen Anregungen.

Herrn Professor Dr. Ernst-Ludwig Winnacker für die hervorragenden Arbeitsbedingungen am Genzentrum und die Förderung meiner Arbeit bis zur Anfertigung des Erstgutachtens.

Herrn Professor Dr. Patrick Cramer für die Übernahme des Zweitgutachtens.

Herrn Professor Dr. Rudolf Grosschedl für die Fortdauer der hervorragenden Arbeitsbedingungen am Institut.

Herrn Professor Dr. Waldemar Kolanus für die Überlassung von Materialien, auf denen die Untersuchungen aufbauten.

Herrn Professor Dr. Michael Brunner, Frau Dr. Maxi Endres und Herrn Professor Dr. Dr. Walter Neupert für ihr Interesse und die Förderung meiner Arbeit.

Herrn Reinhold Härtel, Herrn Professor Dr. Karl-Peter Hopfner und Herrn Dr. Ingo Korndörfer für Ihre Unterstützung bei der Lösung von Fragestellungen am Computer.

Frau Dr. Heike Bruhn für die einführende Betreuung meiner Arbeit.

meinen Freunden und Kollegen Jan Barnikow, Dr. Ralf Leuchter, Dr. Jens Wiehler und Peter Wirtz für die stete Diskussionsbereitschaft und die abschließende Durchsicht meiner Arbeit.

Frau Sigi Kastenmüller für ihre außerordentliche Hilfsbereitschaft und

allen Mitgliedern des Genzentrums für das gute und anregende Betriebsklima.

Verzeichnis der verwendeten Abkürzungen:

A	Adeninbase von DNA
ADP	Adenosin-5'-diphosphat
Amp	Ampizillin
AP	alkalische Phosphatase
APS	Ammoniumpersulfonsäure
ARF	ADP-Ribosylierungsfaktoren
aTC	anhydro-Tetrazyklin
ATP	Adenosin-5'-triphosphat
bp	Basenpaare
BCIP	5-Brom-4-Chlor-3-Indolylphosphat
BSA	Rinderserum-Albumin
bzw.	beziehungsweise
C	Cytosinbase von DNA
CD	Circular-Dichroismus
CD 3	<i>cluster of differentiation 3</i>
CDR	Komplementaritäts-bestimmende Region
CP	Circularpermutante
C-terminal	carboxyterminal
Cy-1	Cytohesin-1
D	Denaturierungsmittel
d.h.	das heißt
Da	Dalton (Gramm pro Mol)
dATP	Desoxyadenosin-5'-triphosphat
dCTP	Desoxycytidin-5'-triphosphat
dGTP	Desoxyguanosin-5'-triphosphat
DMF	Dimethylformamid
DNA	Desoxyribonukleinsäure
DNase	Desoxyribonuklease
dNTP	Desoxyribonukleosid-5'-triphosphate
dsDNA	doppelsträngige DNA
dTTP	Desoxythymidin-5'-triphosphat
E	entfaltete Proteipopulation
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraessigsäure
<i>et al.</i>	<i>et alii</i>
<i>etc.</i>	<i>et cetera</i>
EtOH	Ethanol
F _{ab}	Antigen-bindendes Antikörperfragment
FN3-Domäne	Fibronectin Typ III Domäne
FRET	<i>Fluorecence Resonance Energy Transfer</i>
F _V -Fragment	Heterodimer aus V _H -Domäne und V _L -Domäne
G	Guaninbase von DNA

g3	Protein von Gen 3 des M13-Bakteriophagen
GdmCl	Guanidiniumchlorid
GDP	Guanosin-5'-diphosphat
GEF	Guaninnukleotid-Austauschfaktor
Gl.	Gleichung
GmbH	Gesellschaft mit beschränkter Haftung
Grp1	genereller Rezeptor für 3-Phosphoinositide
GTP	Guanosin-5'-triphosphat
h	Stunde
His-tag	C-terminale Proteinfusion von sechs Histidinresten
hPSTI	<i>human pancreatic secretory trypsin inhibitor</i>
Hsp	Hitzeschockprotein
HuCal	<i>human combinatorial antibody libraries</i>
IG-Domäne	Immunglobulin-Domäne
IPTG	Isopropyl-β-D-thiogalaktosid
Kan	Kanamycin
kDa	kilo Dalton
lac	Laktose-Operon
LB	Luria-Bertani
LMU	Ludwig-Maximilians-Universität
Lsg.	Lösung
M	molar
min	Minuten
ml	Mililiter
mM	milimolar
MOMP	<i>major outer membrane protein</i> von <i>Legionella pneumophila</i>
MPP	<i>matrix processing peptidase</i>
mRNA	<i>messenger-Ribonukleinsäure</i>
MS	Massenspektrum
mt-Hsp70	mitochondriales Hitzeschockprotein 70
μl	Mikroliter
μM	mikromolar
N	nativ gefaltete Proteinpopulation
NADH	Nikotinadenindinukleotid (reduzierte Form)
NBT	4-Nitrotetrazoliumchloridblau
Ni-NTA	Nickel-Nitrilotriacetat
nm	Nanometer
nM	nanomolar
NMR	Kernresonanzspektroskopie
NTA	Nitrilotriacetat
N-terminal	aminoterminal
OD	optische Dichte
ompA	<i>outer membrane protein A</i>
ori	Replikationsursprung
PAGE	Polyacrylamid-Gelelektrophorese

PBS-Puffer	<i>phosphat-buffered saline</i>
PCR	Polymerase-Kettenreaktion
PDB	<i>Brookhaven Protein Data Base</i> für Proteinstrukturen
PEG	Polyethylenglykol 6000
pfu	<i>plaque forming units</i>
pH	negativer dekadischer Logarithmus der H ⁺ -Konzentration
PH-Domäne	Pleckstrin-homologe Domäne
PIP3	Phosphatidylinositol-(3,4,5)-trisphosphat
PLC	Phospholipase C
pM	picomolar
pSu9	Signalpeptid der Untereinheit 9 (mitochondriale ATPase)
R	allgemeine Gaskonstante
RBS	Ribosomen-Bindestelle
r _{cf}	relative Zentrifugalkraft
rmsd	<i>root mean square deviation</i>
RNA	Ribonukleinsäure
RNase	Ribonuklease
RT	Raumtemperatur
s.	siehe
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
scF _V -Fragment	linkerverknüpftes (<i>single chain</i>) F _V -Fragment
SDS	Natriumdodecyl-Sulfat
SDS-PAGE	SDS-Polyacrylamid-Gelelektrophorese
Sec	Protein des allgemeinen sekretorischen Transportwegs
SRP	bakterielles <i>signal recognition particle</i>
ssDNA	einzelsträngige DNA
SU	Untereinheit
SV	Säulenvolumen
T	Thyminbase von DNA
Taq	<i>Thermus aquaticus</i>
TEMED	N, N, N', N'-Tetramethylethylendiamin
<i>tet</i>	Tetrazyklin
TIM	Translocase der inneren Membran
TOM	Translocase der äußeren Membran
Tris	Tris-(Hydroxymethyl)-aminomethan
U	enzymatische Einheit
Upm	Umdrehungen pro Minute
UV	ultra-violett
vgl.	vergleiche
V _H -Domäne	schwere, variable Immunglobulin-Domäne
vis	Licht im sichtbaren Bereich
V _L -Domäne	leichte, variable Immunglobulin-Domäne
w/v	Gewicht pro Volumen

Inhaltsverzeichnis

1.	EINLEITUNG	1
1.1.	Die Bedeutung von Proteinstrukturen	1
1.2.	Die Fragestellungen dieser Arbeit	4
1.3.	Funktion und Struktur des humanen Cytohesin-1 und seiner Pleckstrin-homologen Domäne	5
1.4.	Die molekulare Erkennung	8
1.4.1.	Die PH-Domäne als Gerüst für ein Paratop	9
1.4.2.	Immunglobulindomänen und weitere Proteine als Gerüststrukturen für Paratope	11
1.5.	Die Verfahren zur Selektion bindender Proteine aus einer Bibliothek	13
1.5.1.	Das <i>phage-display</i> -Verfahren	13
1.5.2.	Weitere Selektionsverfahren für Proteine	16
1.5.3.	Selektion von affinen Nukleinsäuren und Kleinmolekülen	18
1.6.	Translokationsoptimierung der PH-Domäne	19
1.6.1.	Das allgemeine Translokationssystem in <i>Escherichia coli</i>	19
1.6.2.	Der mitochondriale Import	21
2.	MATERIAL	23
2.1.	Bakterien und Phagen	23
2.1.1.	<i>Escherichia coli</i> K12-Stämme	23
2.1.2.	Bakteriophage	25
2.2.	DNA	25
2.2.1.	Vektoren	25
2.2.2.	Oligodesoxynukleotide	26
2.3.	Kommerziell erhältliche Enzyme und Proteine	30
2.4.	Standards und Kits	30
2.5.	Chemikalien	31
2.6.	Geräte	32

2.7.	Andere Materialien	33
3.	METHODEN	34
3.1.	Plasmid-DNA in <i>Escherichia coli</i>	34
3.1.1.	Kultivierung und Aufbewahrung von <i>Escherichia coli</i>	34
3.1.2.	Elektroporation von Plasmid-DNA in elektrokompetente <i>E. coli</i>	35
3.1.3.	Transformation von <i>E. coli</i> mit Plasmid-DNA	37
3.1.4.	Plasmid-DNA-Präparation	38
3.2.	Restriktionsendonukleaseverdau von DNA	39
3.2.1.	Extraktion der DNA aus Agarosegelen	41
3.2.2.	Ligation von DNA-Fragmenten	41
3.3.	Einführung gezielter Mutationen	41
3.3.1.	Präparation von Einzelstrang-DNA	41
3.3.2.	Phosphorylierung von Oligodesoxynukleotiden	42
3.3.3.	Ortsspezifische Mutagenese	43
3.3.4.	Polymerase-Kettenreaktion (PCR) zur Einführung von gezielten Mutationen	44
3.3.5.	Reassemblierungs-PCR und <i>QuikChange</i> zur Einführung von Punktmutationen	46
3.4.	Zufallsmutagenese über Polymerase-Kettenreaktion	48
3.4.1.	Mutagene PCR	48
3.4.2.	Randomisierung durch Einsatz randomisierter Oligodesoxynukleotide	49
3.4.3.	Herstellung einer in drei Segmenten randomisierten Bibliothek von PH-Domänen	51
3.5.	DNA-Sequenzierung	52
3.6.	Proteinexpressionen und die jeweiligen Zellaufschlüsse	53
3.6.1.	Grundregeln für die Genexpression von Proteinen in <i>E. coli</i>	53
3.6.2.	Periplasmatische Proteinexpression unter standardisierten Bedingungen	55
3.6.3.	Cytoplasmatische Proteinexpression unter standardisierten Bedingungen	56
3.6.4.	Cytoplasmatische Expression von unlöslichem Protein in <i>inclusion-bodies</i>	57
3.7.	Proteinchemische Aufreinigung und Dialyse	58
3.7.1.	Aufreinigung von Proteinen mit His-tag an Ni-NTA-Matrix	59
3.7.2.	Dialyse von Proteinen	60
3.8.	SDS-Polyacrylamid-Gelelektrophorese von Proteinen	61
3.8.1.	Gelelektrophoretische Auftrennung von Proteinen	61
3.8.2.	Silberfärbung von Proteingelen	62
3.8.3.	Coomassie-Färbung	63
3.8.4.	Western-Blots	63

3.9.	Das <i>phage-display</i>-Verfahren	66
3.9.1.	Bestimmung des Phagentiters	66
3.9.2.	Herstellung und Reinigung neuer Bakteriophagen	67
3.9.3.	Immobilisierung von Zielsubstanzen	67
3.9.4.	<i>Panning</i> von Phagen-Bibliotheken	68
3.10.	Spektroskopische Methoden	70
3.10.1.	Absorptionsspektroskopie	70
3.10.2.	Fluoreszenzspektroskopie von Proteinen	71
3.10.3.	Denaturierungs-Übergangskurven	71
3.10.4.	CD-Spektroskopie	72
3.10.5.	Massenspektrometrie	73
3.11.	Computer-Analysen	73
4.	ERGEBNISSE	74
4.1.	Die PH-Domäne als Gerüst für randomisierte Loops	74
4.1.1.	Identifizierung der Loops	74
4.1.2.	Randomisierung mit „Mutator“	75
4.1.3.	Synthese einer in drei Loops randomisierten Bibliothek von PH-Domänen	76
4.1.4.	Validierung der Bibliothek	86
4.1.5.	Untersuchung der thermodynamischen Stabilität von PH-Domänen	90
4.1.6.	Stabilität von randomisierten PH-Domänen	94
4.2.	Die Bibliothek randomisierter PH-Domänen im <i>phage-display</i>-Verfahren	96
4.2.1.	Ausgewählte Zielsubstanzen für das <i>phage-display</i> -Verfahren	96
4.2.2.	Selektionsergebnisse nach dem <i>phage-display</i> -Verfahren	97
4.2.3.	Alternative Selektionsverfahren zum <i>phage-display</i>	101
4.3.	Optimierung der PH-Domäne zur Translokation in das Periplasma von <i>E. coli</i>	102
4.3.1.	Austausch spezifischer Sequenzelemente und Destabilisierung der PH-Domäne zur Verbesserung der Translokation?	104
4.3.2.	Modellierung eines strukturierten Linkers für Circularpermutanten	110
4.3.3.	Proteinstabilität im Hinblick auf die Optimierung der Translokation	116
4.3.4.	Cytoplasmatische Expression im Hinblick auf die Translokations-optimierung	120
4.3.5.	Untersuchung der periplasmatischen Expression	123
4.3.6.	Prozessierungsprodukte nach periplasmatischer Expression	126
4.3.7.	Weitere Ergebnisse zur Translokation der Circularpermutante CP 2.6.	129
4.4.	Die Expression der PH-Domäne auf dem M13-Phagen	132
4.4.1.	Circularpermutante CP 3.6. in ersten <i>phage-display</i> -Versuchen	133

4.5.	Der Import der PH-Domäne in Mitochondrien	134
4.5.1.	Grundlagen	134
4.5.2.	Ergebnisse des mitochondrialen Imports von PH-Domänen	136
5.	DISKUSSION	141
5.1.	Wege zur funktionellen Veränderung von Proteinen	141
5.2.	Die Randomisierung der PH-Domäne	142
5.2.1.	Statistische Betrachtung von Randomisierungen	142
5.2.2.	Randomisierung mit „Mutator“ und ihre Umsetzung zu einer Bibliothek randomisierter PH-Domänen	144
5.3.	<i>Phage-display</i> mit der PH-Domäne	148
5.4.	Die Verbesserung der Translokation	153
5.4.1.	Hypothesenüberprüfung zur Translokationsoptimierung	156
5.4.2.	Verbesserte Translokation bei den Circularpermutanten CP 2.6 und CP 3.6.	157
5.4.3.	Interpretationen für die guten Translokationseigenschaften	160
5.4.4.	Fazit für die Translokation	165
5.5.	Die Aussagen aus dem mitochondrialen Import der PH-Domäne	166
5.6.	Vergleich der Transporte	171
6.	ZUSAMMENFASSUNG	173
6.1.	Die PH-Domäne als Paratop	173
6.2.	Die PH-Domäne als Substrat für Translokationen	174
7.	LITERATURVERZEICHNIS	176
8.	ANHANG	
8.1.	Mathematische Betrachtung von Randomisierungen	

1. Einleitung

1.1. Die Bedeutung von Proteinstrukturen

Die Suche nach dem Aufbau und der Funktionsweise lebendiger Organismen beschäftigt die Menschheit von alters her. Dabei bestechen Lebewesen durch ihre außerordentliche Komplexität. So ist beispielsweise der Mensch im Kontext seiner Umwelt und die einzelnen Körperzellen des Menschen sind im Kontext zu den anderen Zellen des Gewebes bzw. des restlichen menschlichen Körpers zu sehen. Die Kompartimente und molekularen Komponenten innerhalb einer Zelle interagieren vielschichtig miteinander. Zudem handelt es sich bei Lebewesen um dynamische Systeme, d.h. es liegen kaum (thermodynamische) Gleichgewichtszustände vor; vielmehr befinden sich die Bestandteile eines Organismus in einem Fließgleichgewicht, welches sich sogar zeitlich noch ändert.

Interessanterweise besteht eine Methode, Einblick in dieses vermeintliche Chaos zu erhalten, in der Aufklärung und Interpretation von statischen Strukturen einzelner Komponenten, die man z.B. durch histologische Untersuchungen auf der zellulären Ebene oder durch Proteinstrukturaufklärung auf molekularer Ebene erhält. Dies scheint zuerst einmal ein Widerspruch in sich zu sein. Worin liegt also die Bedeutung von Strukturaufklärungen im Hinblick auf das Naturverständnis?

Die Struktur eines Moleküls kann nicht nur Aufschluß über molekulare Interaktionen geben, sondern auch über die jeweiligen Funktionen dieser Interaktionen. So konnten Michel, Deisenhofer und Huber (Deisenhofer *et al.*, 1984: Nobelpreis 1988) die Struktur des membranständigen Photokomplexes von *Rhodospseudomonas viridis* aufklären. Dabei wurde über die eigentliche Struktur hinaus noch die Information gewonnen, wie das Photosystem im allgemeinen und seine einzelnen Komponenten im besonderen funktionieren. Um die jeweilige Funktion von Proteinen zu begreifen, ist es wünschenswert, jeder Aminosäuresequenz direkt eine detaillierte Struktur zuzuordnen bzw. vorherzusagen zu können. Erstaunlicherweise findet man, daß die Struktur der Proteine ausschließlich durch die Aminosäuresequenz determiniert ist. Man sollte also nach einem reichen Erfahrungsschatz gelöster Strukturen die Prinzipien, die dem Weg von der Aminosäuresequenz zur Proteinstruktur zu Grunde liegen,

verstehen können und dann in der Lage sein, beliebigen Aminosäuresequenzen die jeweilige Struktur vorhersagen zu können.

Bislang konnte man dieses sogenannte „Proteinfaltungsproblem“ (Anfinsen, 1973: Nobelpreis 1972) allerdings nicht lösen. Dies liegt unter anderem daran, daß nicht allgemeingültig geklärt ist, welche physikalischen Eigenschaften korrekt gefaltete Proteine gegenüber anderen Faltungsvarianten auszeichnen. Ein häufig (so auch in dieser Arbeit) verwendeter Ansatz ist die These, daß sich gerade die thermodynamisch stabilste Konformation eines Proteins als Struktur ausbildet, und sich Proteinstrukturen somit über die geringste freie Faltungsenthalpie aller Konformationen definieren (Goldenberg, 1988). Dieser thermodynamische Ansatz für die Strukturvorhersage setzt allerdings voraus, daß alle inter- und intramolekularen Interaktionen berücksichtigt werden und daß sich alle Faltungszustände reversibel ineinander überführen lassen. Schon bei löslichen Proteinen müßten Wechselwirkungen (und damit Enthalpieniedrigungen) mit anderen Proteinen und anderen Molekülen aus dem biologischen Kontext des Proteins in die Betrachtungen eingehen; bei Transmembranproteinen besteht *in vivo* das „Lösungsmittel“ aus der Lipidmembran, der wäßrigen Grenzphase sowie beispielsweise dem Cytosol. Es ist somit mehrphasig, was eine thermodynamische Betrachtung zur Strukturvorhersage „*in vitro*“ außerordentlich verkompliziert, wenn nicht sogar unmöglich macht. Thermodynamisch weniger stabile Proteinkonformationen könnten sich als Proteinstrukturen etablieren, wenn kinetisch bedingte, irreversible Faltungsprozesse die Struktur vorgeben. Welche physikalischen Parameter für die Strukturgebung genau ausschlaggebend sind, dürfte kaum feststellbar sein. Entscheidend ist in der Natur lediglich, daß die Aminosäuresequenz (meistens) zu einer eindeutigen Struktur führt: ein aktuelles Beispiel für stabile Fehlfaltungen sind die Prionenerkrankungen, die sich auf Aggregate von fehlgefalteten Proteinen zurückführen lassen (Dobson, 1999).

Die in dieser Arbeit untersuchte Pleckstrin-homologe Domäne (PH-Domäne: s. Kapitel 1.3.) läßt sich *in vitro* durch Zusatz von Denaturierungsmittel entfalten und durch Entzug des Denaturierungsmittels vollständig, d.h. reversibel, in seine ursprüngliche Konformation zurückfalten. Aus derartigen Experimenten läßt sich die freie Faltungsenthalpie bestimmen - und dennoch kann man *ab initio* nicht die Struktur aus der Aminosäuresequenz ermitteln, was das nächste Dilemma bei der Lösung des Proteinfaltungsproblems verdeutlicht: für jede Konformation müßte eine charakteristische physikalische Größe wie die freie Faltungsenthalpie berechnet werden. Der Vergleich der freien Faltungsenthalpien aller Konformationen sollte dann Aufschluß geben, welches die günstigste Konformation ist und welche damit die Struktur repräsentiert. Die Zahl der miteinander zu vergleichenden Konformationen ist aber nahezu unbegrenzt.

Vernachlässigt man alle Freiheitsgrade der Aminosäureseitenketten, wird die Struktur einer Aminosäure in einer Polypeptidkette durch die beiden Rotationswinkel Ψ und Φ festgelegt. Sie geben die Ablenkung um die Achse zwischen C_α -Atom und Carbonylgruppe (Ψ) und die Ablenkung um die Achse zwischen C_α -Atom und Aminogruppe (Φ) wieder. Beide Winkel liegen in Proteinstrukturen nicht gleichverteilt vor (Ramakrishnan & Ramachandran, 1965) und sind für die Paulingschen Idealstrukturen eindeutig definiert: die ideale α -Helix zeichnet sich durch $\Psi = 120^\circ$ und $\Phi = 120^\circ$ sowie das ideale β -Faltblatt durch $\Psi = 300^\circ$ und $\Phi = 60^\circ$ aus. Innerhalb eines Proteins können beide Winkel für jede einzelne Peptidbindung fast jeden Wert einnehmen. Beschränkt man beide Winkel jeweils auf m Möglichkeiten, setzt die Unabhängigkeit beider Winkel voneinander voraus und vernachlässigt die Konformationen der Seitenketten, lassen sich für ein Peptid mit n Aminosäuren m^{2n-2} Konformationen angeben: bei nur zehn Winkelmöglichkeiten und einem Peptid aus 101 Aminosäuren hieße dies, daß 10^{100} Strukturen hinsichtlich ihrer physikalischen Eigenschaften miteinander verglichen werden müßten.

Ausgangspunkt für die beobachteten, definierten Proteinstrukturen sind nicht alle möglichen Konformationen, die in einem chemischen Gleichgewicht zueinander stehen und aus denen sich die stabilste Konformation als größte Population herausbildet. Dies widerspräche Untersuchungen an Modellproteinen, die sich in Sekundenbruchteilen korrekt falten (Jackson, 1998). Es deutet vielmehr darauf hin, daß die Proteinfaltung ein kinetisch limitierter Prozeß ist und Faltungswege vorgegeben sind, die sich aus der Aminosäuresequenz ableiten (Levinthal, 1968). Dasselbe zeigen auch kinetische NMR-Studien, die eine erweiterte, detaillierte Sichtweise für die Proteinfaltung einfordern (Baldwin, 1995): ausgehend vom entfalteten, strukturell nicht definierten Protein falten sich zunächst verschiedene Peptidsequenzen voneinander unabhängig zu lokalen, intermediären Strukturen. Diese finden dann zu einer globalen Proteinstruktur zusammen. Dabei beschreiten nicht alle Proteine genau denselben „Faltungsweg durch die Energielandschaft“; vielmehr ist nur das Ziel (das gefaltete Protein) für alle Proteinmoleküle identisch. Der Faltungsmechanismus verläuft also weder synchron noch über (notwendigerweise) definierte Faltungsintermediate.

Strukturvorhersagen für Proteine basieren heutzutage unter anderem auf Sequenzidentitäten zu homologen Proteinen bekannter Struktur anstelle einer *ab initio*-Rechnung und dem Vergleich aller möglichen Konformere. Ein Beispiel dafür ist das Modellprotein dieser Arbeit. Die Struktur der PH-Domäne des humanen Cytohesin-1 (Kolanus *et al.*, 1996) wurde bislang noch nicht gelöst, läßt sich aber durch Sequenzvergleich mit der strukturell gelösten PH-Domäne des Grp1 (Lietzke *et al.*, 2000) modellieren (Guex & Peitsch, 1997).

1.2. Die Fragestellungen dieser Arbeit

Die Fragestellungen für diese Arbeit decken sich mit den generellen Problemstellungen, die der Strukturaufklärung von Proteinen zu Grunde liegen:

- (i) Ist jede Aminosäure der PH-Domäne für die Ausbildung der Proteinstruktur essentiell oder lassen sich einige Aminosäurereste beliebig gegen andere austauschen, ohne dabei die globale Struktur der PH-Domäne zu verändern?
- (ii) Sollte sich die globale Struktur der PH-Domäne durch beliebigen Austausch von Aminosäureresten in bestimmten Regionen nicht ändern, lassen sich dann Mutanten der PH-Domäne herstellen, die neuartige, intermolekulare Wechselwirkungen mit anderen Substanzen eingehen können?
- (iii) Welche strukturellen und biochemischen Eigenschaften von Proteinen beeinflussen deren Transportprozesse sowohl vom Cytoplasma in das Periplasma von *Escherichia coli* als auch in die mitochondriale Matrix von *Saccharomyces cerevisiae*?

Bei allen drei Fragestellungen dieser Arbeit geht es grundsätzlich um die strukturellen und biochemischen Eigenschaften von Proteinen in Abhängigkeit von ihrer Aminosäuresequenz am Beispiel der PH-Domäne und nicht darum, die Funktionen der natürlichen PH-Domäne *in vivo* zu ergründen. Letzteres ist bereits ausgiebig geschehen, wie das folgende Kapitel zeigt.

1.3. Funktion und Struktur des humanen Cytohesin-1 und seiner Pleckstrin-homologen Domäne

Cytohesin-1 ist ein humanes, intrazellulär lokalisiertes Protein, das in die Regulierung der Zelladhäsion involviert ist (Kolanus *et al.*, 1996). Es wurde bislang hauptsächlich in den Zellen des hämatopoetischen Systems nachgewiesen. Es besitzt ein Molekulargewicht von 47 kDa und ist aus mehreren Domänen aufgebaut, von denen eine die sogenannte Pleckstrin-homologe Domäne (PH-Domäne) ist.

Einer kurzen aminoterminalen „*coiled-coil*“-Domäne des Cytohesin-1 konnte bislang noch keine eindeutige Funktion zugeordnet werden. Es folgt eine ungefähr 200 Aminosäuren lange SEC7-Domäne, über die die Interaktion mit CD18 im Cytoplasma stattfindet (Kolanus *et al.*, 1996) und die den Guaninnukleotid-Austausch an ADP-Ribosylierungsfaktoren (ARFs) katalysiert. Somit läßt sich Cytohesin-1 den Guaninnukleotid-Austauschfaktoren zurechnen (GEFs, *guanine nucleotide exchange factors*) (Donaldson & Jackson, 2000; Meacci *et al.*, 1997). ARF-Proteine können in einer aktivierten Form mit einem GTP oder in einer inaktiven Form mit einem GDP vorliegen. GEFs sind für die Aktivierung (also den Austausch von GDP gegen GTP) und GAPs (GTPase-Aktivierungsproteine) für die Inaktivierung (also die Hydrolyse des GTP) zuständig (Donaldson & Jackson, 2000; Moss & Vaughan, 1999). Die Struktur der SEC7-Domäne von Cytohesin-1 wurde durch NMR-Spektroskopie aufgeklärt (Betz *et al.*, 1998). Der SEC7-Domäne folgt im Cytohesin-1 die PH-Domäne. Sie wird in dieser Arbeit allerdings nicht als Modul innerhalb des Cytohesin-1 untersucht, sondern als selbständiges Protein.

Bei den PH-Domänen handelt es sich um eine große Strukturfamilie. Sie konnten bislang in ungefähr 100 unterschiedlichen Proteinen nachgewiesen werden; viele von ihnen sind Bestandteil der Signaltransduktion und Cytoskelettregulierung (Rebecchi & Scarlata, 1998). Die Übereinstimmung in der Aminosäuresequenz ist unter den PH-Domänen verhältnismäßig gering, doch sind ihre Strukturen und auch ihre Aufgaben *in vivo* sehr ähnlich. Sie treten als Untereinheiten vieler Proteine auf: so etwa bei den Proteinkinasen β ARK, Akt und Btk, bei den Regulatoren kleiner G-Proteine Ras-GAP, SOS, Vav und Db1 sowie bei den Lipid-assoziierten Proteinen PLC β , γ , δ und PI3-Kinase (Gibson *et al.*, 1994). Sie zeigen zu Phosphatidylinositolphosphaten der Membranen spezifische Affinitäten (Lemmon *et al.*, 1996; Pawson, 1995; Shaw, 1996). Darüber assoziieren die oben genannten Proteine an der Innenseite der Zellmembran.

Indirekt können PH-Domänen enthaltende Proteine auf Signale reagieren, indem sie an Phosphatidylinositolphosphaten binden, die selber als „*second messenger*“ fungieren. So binden die PH-Domänen des Btk und Grp1 (genereller Rezeptor für 3-Phosphoinositide) spezifisch an Phosphatidylinositol-(3,4,5)-trisphosphat (PIP3) (Ferguson *et al.*, 2000; Klarlund *et al.*, 1997), das erst nach Aktivierung von Phosphoinositidlipiden durch eine Phosphoinositid-3-kinase in der Zellwand nachweisbar ist (Stephens *et al.*, 1993). Die Membranrekrutierung der entsprechenden Proteine findet damit erst in Folge eines Signals statt (Gray *et al.*, 1999; Varnai *et al.*, 1999; Venkateswarlu *et al.*, 1998).

Die Aminosäuresequenz der PH-Domäne des Cytohesin-1 ist zu der des Grp1 zu über 92% identisch (s. Abb. 4.14). Insbesondere die Bindungsmotive für PIP3 sind nahezu identisch (Lietzke *et al.*, 2000). So dürfte auch die PH-Domäne des Cytohesin-1 in Abhängigkeit vom Gehalt an PIP3 an der Membran assoziieren. *In vitro* konnte gezeigt werden, daß zusätzlich eine polybasische c-Domäne des Cytohesin-1 nachhaltig zur Bindung des Cytohesin-1 an PIP3 beiträgt (Nagel *et al.*, 1998a; Nagel *et al.*, 1998b).

Der Name „Pleckstrin-homologe Domänen“ (PH-Domänen) stammt von dem Protein Pleckstrin, das zwei derartige Domänenstrukturen aufweist. Pleckstrin wurde als Substrat der menschlichen Tyrosinkinase C in Thrombozyten gefunden (Rendu *et al.*, 1992). Von Proteindomänen, die sehr ähnliche Strukturen und Funktionen aufweisen, kann vermutet werden, daß sie sich von einem gemeinsamen Vorfahren ableiten. Deshalb werden sie als „Pleckstrin-homologe“ Domänen bezeichnet. Derartige PH-Domänen wurden bislang noch nie isoliert, sondern immer als Module in Fusion mit anderen Proteinstrukturen gefunden. PH-Domänen können sich in ihrer Aminosäuresequenz erheblich voneinander unterscheiden und besitzen zudem hypervariable Loop-längen (s. Tab. 1.1), weshalb diese Strukturfamilie erst spät identifiziert und definiert wurde (Haslam *et al.*, 1993; Mayer *et al.*, 1993). In welcher Vielzahl PH-Domänen in der Natur auftreten, konnte erst wahrgenommen werden, seitdem die Bioinformatik die entsprechenden Algorithmen zum gleichzeitigen Vergleich mehrerer Aminosäuresequenzen liefert und seitdem viele Strukturen von PH-Domänen vorliegen (Gibson *et al.*, 1994; Rost, 1996).

Tab. 1.1: Strukturhomologievergleich von PH-Domänen.

PDB#	N-Terminus	α Phos	β 1	L 1	β 2	L 2
1DYN	MKTSGNQDEIL		VIRKGLTI	NNIGIMKGG	SK EYWFVLT	AE
Cy-1	MAN		PDREGWLLKL	GGGRVKT	WK RRFVILT	DN
1DRO	GSGTGAGEG		HE GYVTRKHE	WDSTTKKASNRS	WD KVYMAA	KAG
1MPH			MEGFLNRKHEW	EAHNKKASSRS	WH NVYCVIN	NQ
1DBH	GQCCNE		FIMEGTLTR	VGAKH	ERHIFLF	DG
1PLS	MEPKR		IREGYLVKK	GSVFNT	WK PMWVLL	ED
1BAK	GSHMGKD		CIMH GYMSKM	GNPFKTQ	WQ RRYFYLF	PN
1MAI	GLQDD	PDLQALLK	GSQLLKV	KSSSW	RR ERFYKLQ	EDCK

PDB#	β 3	L 3	β 4	L 4	β 5	L 5
1DYN	NLSWY K	DDEEK	E KKYMLS	VDN	LKLRD	VEKGFMSK
Cy-1	ALYYF E	YTTDK	E PRGIIP	LEN	LSIRE	VEDSKKP
1DRO	RISFY K	DQKGYKSNPELTFRGE	PSY	DLQNA	AIEI	ASDYTKKK
1MPH	EMGFY K	DAKSAASGIPYHSEV	PVS	LKE	AICEVA	LDYKSKK
1DBH	LMICC K	SNHGQPRLPASNAEY	R LKEKFF	MRKV	QIND	KDDTNEYK
1PLS	GIEFY K	KKSDN	S PKGMIP	LKG	STLTS	PCQDFGKRM
1BAK	RLEWR	GEGEAPQ	SL LTT	MEE	IQSVEET	QIKERK
1MAI	TIWQ	ESRKVMRSPES	QL FS	IED	IQEVRM	GHRTEGLEKRDIPED

PDB#	β 6	L 6	β 7	L7	α 1
1DYN	HIFALFN	TEQRNVYKDY	RQ LELAC	ET	QEEVDSWKASFLRA
Cy-1	NVFEL	YIPDNKDQVIKAAKTEADGQVVEGNH	TV YRISA	PT	PEEKEEWIKAIKAA
1DRO	HVLRVK	LANGAL	FLLQA	HD	DTEMSQWVTSLKAQ
1MPH	HVFCLR	LSDGNE	YLFQA	KD	DEEMNTWIIQAIISSA
1DBH	HAFEII	LKDEN	S VIFSA	KS	AEEKNNWMAALISL
1PLS	FVFKITT	TKQ	QD HFFQA	AF	LEERDAWVRDINKA
1BAK	CLLLKI	RGG	KQ FILQC	DS	DPELVQWKKEKRDA
1MAI	RCFSIVF	KDQRN	T LDLIA	PS	PADAQHVVQGLRKI

Die bis zum Jahr 2000 gelösten Strukturen von PH-Domänen wurden überlagert und mit der Struktur des Dynamins (1DYN) verglichen. Die Loopregionen L 1, L 3, L 5 und L 6 sind bei den verschiedenen PH-Domänen längenvariabel. Die Aminosäuren der bislang strukturell nicht analysierten PH-Domäne des Cytohesin-1 (Cy-1) wurden durch Sequenzvergleich mit den anderen PH-Domänen den Strukturelementen zugeordnet. Auf diese Weise konnten die Loopregionen identifiziert werden. Sowohl die positionell übereinstimmenden Aminosäuren als auch die dazu im Cytohesin-1 ähnlichen Aminosäuren sind fett markiert. Mit L X sind die teilweise längenvariablen Loopregionen, mit β X die konservierten β -Faltblattstrukturen (hellgrau) gekennzeichnet. Die C-terminale α -Helix aller PH-Domänen ist mit α 1 (dunkelgrau) und die N-terminale α -Helix der Phospholipase C (1MAI) ist mit α Phos (dunkelgrau) beschriftet. Die Abkürzungen für die Proteine entsprechen dem jeweiligen PDB-Eintrag.

Die Strukturfamilie der PH-Domänen zeichnet sich durch eine charakteristische Abfolge von β -Faltblättern und durch die gemeinsame C-terminale α -Helix aus, die sich mit hoher Übereinstimmung überlagern lassen (s. Tab. 1.1 und Rebecchi & Scarlata, 1998). Zwischen den antiparallelen β -Strängen sind bei den PH-Domänen vier ausgeprägte, strukturell weniger definierte Bereiche (Loops) eingebunden. Von diesen konnten bis zu drei (Loop 1, 3 und 6) als spezifische Bindungsstellen für Phosphatidylinositolphosphate identifiziert werden (Lietzke *et al.*, 2000). Seit kurzem ist die Struktur der PH-Domäne des Grp1 [genereller Rezeptor für 3-Phosphoinositide] (Lietzke *et al.*, 2000) gelöst. Aufgrund der hohen Sequenzidentität zu der Aminosäuresequenz des Cytohesin-1 läßt sich die Struktur der PH-Domäne des Cytohesin-1 mit großer Wahrscheinlichkeit vorhersagen (s. Abb. 1.1). Über die ca. 120 Aminosäuren der beiden PH-Domänen sind 10 Punktmutationen verteilt und der Loop L1 ist bei Cytohesin-1 um die Aminosäure G(12) vergrößert (s. Tab. 4.4).

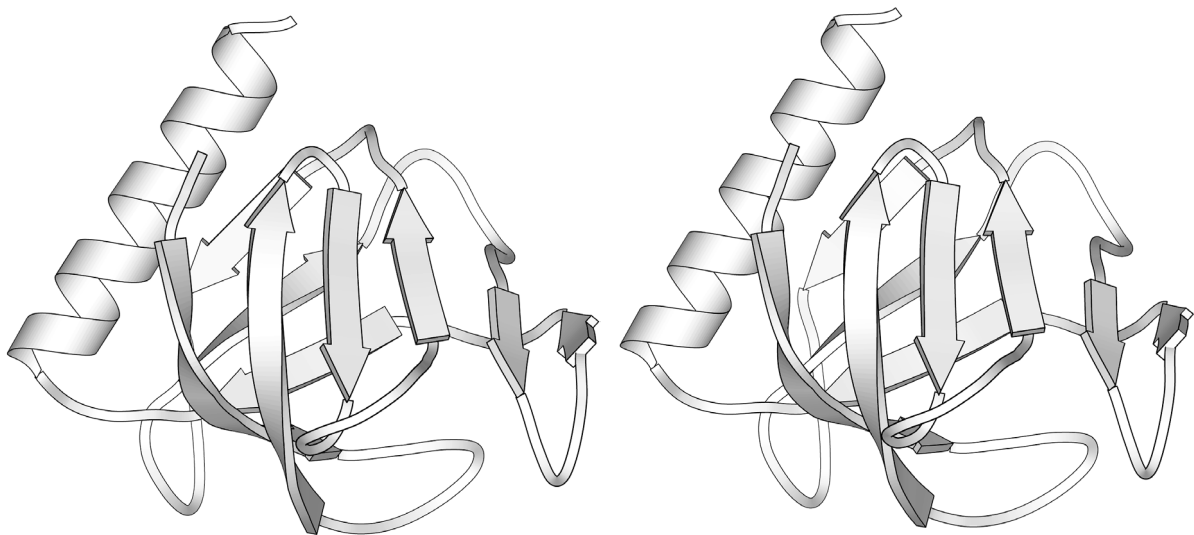


Abb. 1.1: Strukturmodell der PH-Domäne des Cytohesin-1.

Molscript Raster 3-D Stereodarstellung. Die Proteinstruktur wurde nach der homologen PH-Domäne des Grp1 (PDB: 1FGZ nach Lietzke *et al.*, 2000) modelliert (SWISS-Modell nach Guex & Peitsch, 1997).

1.4. Die molekulare Erkennung

Die Funktion von Proteinen beruht meistens auf stabilen oder regulierbaren Interaktionen zu anderen Proteinen und Molekülen. Erst das veränderbare Zusammenspiel vieler Moleküle vermittelt die Bandbreite von Mechanismen

und Reaktionen, die die Komplexität in der belebten Natur auszeichnet. Derartige Interaktionen können auch in einem mikroevoativen Prozeß *in vivo* erzeugt werden: das Immunsystem mit seinen Antikörpern ist das Paradigma dafür, wie Antikörper mit hoher Affinität zu Antigenen selektiert werden, um Abwehrreaktionen gegen Antigene bzw. Träger von Antigenen wie Viren einzuleiten. Antikörper setzen sich aus Immunglobulindomänen (IG-Domänen) zusammen, die wiederum aus hypervariablen Loops und konstanten Regionen bestehen. Die hypervariablen Loops dienen als Paratope und binden spezifisch an die Epitope der Antigene. Dagegen stellen die konstanten Regionen der Immunglobuline in erster Linie ein stabiles Gerüst für die hypervariablen Loops dar (Padlan, 1996).

Ein intaktes Immunsystem ermöglicht es, über eine Vielzahl von Mutations-, Rekombinations- und Selektionsereignissen hochspezifisch bindende Antikörper herzustellen. Dies ist nicht nur für die immunologische Grundlagenforschung interessant, sondern ermöglicht auch die Konstruktion von Immunreagenzien mit maßgeschneiderten Eigenschaften, was insbesondere für die therapeutischen und diagnostischen Anwendungen in Forschung und Medizin intensiv angestrebt wird. Antikörper werden auch heute noch von und in Versuchstieren selektiert bzw. exprimiert. Um dies zu umgehen und eine großtechnische Produktion zu ermöglichen, werden Antikörper auf Antikörperfragmente reduziert und *in vitro* selektiert (s. Kapitel 1.4.2.). Alternativ sucht man nach Molekülen, die die Rolle von Antikörpern übernehmen können - und vielleicht sogar „günstigere“ Eigenschaften aufweisen als diese. Dieser Ansatz wird mit der PH-Domäne verfolgt.

1.4.1. Die PH-Domäne als Gerüst für ein Paratop

Bei den Immunglobulindomänen dienen die hypervariablen Loops als epitopbindende Paratope und die konstanten Regionen als Gerüst bzw. Verankerung für diese Loops. PH-Domänen lassen sich formal ähnlich unterteilen. Auch sie besitzen Loops, die sich als Paratope anbieten und die sich in einem definierten Kontext antiparalleler β -Faltblätter befinden. Diese sollten als Gerüst für veränderte Loops dienen können. Eignen sich diese Loops aber auch für die Ausbildung von Interaktionen zu „Antigenen“? - Die Selektion auf Interaktion zu Antigenen setzt voraus, daß die Loops randomisiert werden; das heißt, anstelle der natürlichen Peptidsequenz werden (weitgehend) willkürlich andere Peptidsequenzen eingebaut, die aber die globale Struktur der PH-Domäne nicht zerstören dürfen. Die natürlichen Aminosäuren dürfen für die Struktur der PH-Domäne also nicht notwendig sein. Dies zu überprüfen,

war die erste Aufgabe dieser Arbeit (s. Kapitel 1.2.). Die Gesamtheit aller Mutanten mit unterschiedlichen Aminosäuresequenzen in den Loopbereichen werden als „Bibliothek“ bezeichnet. Die Längenvariabilität der gewählten Loops bei den verschiedenen PH-Domänen (s. Tab. 1.1) deutet auf eine Eignung dieser Loops zur Randomisierung hin (Ku & Schultz, 1995). Die ausgewählten Loops 1, 3 und 6 weisen zudem alle in eine Richtung und könnten gemeinsam ein diskontinuierliches Paratop darstellen, das ein Antigen bindet (s. Abb. 1.2). Einen Beweis für die Möglichkeit einer Bindung liefert bereits die Natur: Aminosäurereste dieser drei Loops interagieren bei Grp1 mit Inositol(1,3,4,5)tetrphosphat, wie die Kristallstruktur zeigt (Lietzke *et al.*, 2000). Mitglieder der PH-Domänen-Superfamilie erkennen zudem auch Proteinliganden (Blomberg *et al.*, 1999). Ein Beispiel dafür ist die Phosphotyrosin-bindende Domäne (PTB) des Insulin-Rezeptor-Substrat-1 (PDB: IRS1 nach Zhou *et al.*, 1995). Die Interaktion erfolgt hier allerdings über den Loop 6 und die Aminosäuren vom β 4-Faltblatt, während die Loops 1 und 3 nicht in die Bindung involviert sind.

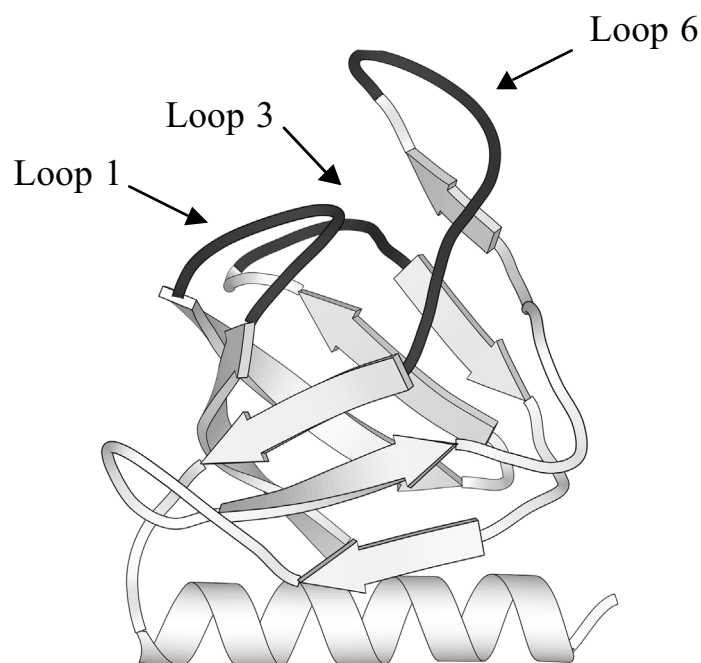


Abb. 1.2: Die zur Randomisierung ausgewählten Loops der PH-Domäne
Die dunkel gezeichneten Peptidsequenzen wurden randomisiert.

Für die praktische Anwendung von Gerüstproteinen eignen sich insbesondere Proteine, die als Monomer vorliegen, klein, robust, einfach zu verändern sind, und sich auch effizient (wenn möglich in *E. coli*) exprimieren lassen (Nord *et al.*, 1997; Steipe *et al.*, 1997). Die PH-Domäne des Cytohesin-1 genügt diesen

Grundanforderungen, da sie als Monomer nicht zur Aggregation neigt, mit über -40 kJ/mol eine hohe thermodynamische Stabilität aufweist und sich nach Denaturierung auch wieder vollständig in den nativen Zustand zurückfalten läßt (s. Kapitel 4.1.5.). Dazu besitzt sie kein *cis*-Prolin, wie sich der modellierten Struktur entnehmen läßt, und nach entsprechender Mutagenese auch kein Cystin oder Cystein, wodurch sie oxidationsunempfindlich ist und sich keine „falschen“ Disulfidbrücken ausbilden können (s. Tab. 4.4). Mit einer Länge von 110 Aminosäuren und einem Molekulargewicht von 15 kDa ist sie zudem verhältnismäßig klein und läßt sich mit über 30 % des Gesamtproteinanteils in *E. coli* löslich überexprimieren. Als humane Proteindomäne sollte auch der Einsatz von Varianten bei humanen Therapien aus immunologischer Sicht unbedenklich sein. Damit eignet sich die PH-Domäne des Cytohesin-1 von vornherein zur großtechnischen Expression in *E. coli* und der Anwendung als Therapeutikum. Diese Eigenschaften stellen die PH-Domäne grundsätzlich als ein besonders geeignetes Gerüstprotein für eine Paratop-Bibliothek dar (Steipe *et al.*, 1997).

1.4.2. Immunglobulindomänen und weitere Proteine als Gerüststrukturen für Paratope

Das Immunsystem ist ein hocheffizientes, kompliziertes Selektionssystem zur Abwehr von sehr unterschiedlichen Fremdsubstanzen. Als Werkzeug für die Erkennung und Markierung von Fremdsubstanzen dienen die beiden variablen IG-Domänen V_H und V_L der Antikörper. Die V_H -Domäne ist dabei Bestandteil der schweren und die V_L -Domäne Bestandteil der leichten Kette von Antikörpern. In der Natur werden jeweils drei variable Loops (CDR-Regionen: *complementarity determining regions*) dieser Domänen durch ein kompliziertes Verfahren aus Mutationen und Rekombinationen der entsprechenden Gene randomisiert und anschließend selektiert. Das Immunsystem stellt zwar das Paradigma für die Selektion bindender Moleküle dar, doch läßt sich dieses natürliche und komplexe Verfahren nur partiell *in vitro* umsetzen. Einige Strukturelemente der Antikörper erübrigen sich bei *in vitro*-Selektionen. Für die reine molekulare Erkennung kann man die Antikörper auf die beiden variablen IG-Domänen V_H und V_L reduzieren. Diese lassen sich als Heterodimer exprimieren (F_V -Fragment). Alternativ verbrückt man zur Expression beide Domänen mit einem synthetischen Linkerkonstrukt zu einem *single-chain-F_V*-Fragment (*scF_V*-Fragment). Derartige Konstrukte lassen sich als Gerüstproteine in den CDR-Regionen randomisieren und bei *in vitro* Selektionsverfahren einsetzen (Hoogenboom *et al.*, 1991).

Auch von anderen Proteinen hat man zeigen können, daß sie sich für die spezifische molekulare Erkennung anderer Moleküle eignen (Übersichtsartikel: Skerra, 2000). So wurden weitere Mitglieder der IG-Superfamilie hinsichtlich ihrer Tauglichkeit als Gerüst untersucht: Bei der Fibronectin Typ III-Domäne (FN3) wurden zwei Loops randomisiert und in einer Bibliothek von 10^8 Mutanten erfolgreich eingesetzt (Koide *et al.*, 1998); das gleiche ist auch mit der variable Domäne des menschlichen, cytotoxischen T-Lymphocyten assoziierten Proteins-4 möglich (Nuttall *et al.*, 1999).

Außerhalb der IG-Superfamilie bieten sich Proteaseinhibitoren als stabile, kleine Domänenstrukturen an. Auch sie sind von der Natur auf molekulare Erkennung ausgelegt. Ein prominentes Beispiel ist der Inhibitor des menschlichen Pankreastrypsins (*human pancreatic secretory trypsin inhibitor*: hPSTI), bei dem ein Loop randomisiert wurde (Szardenings *et al.*, 1995). Der Einsatz dieser Bibliothek erfolgte unter anderem erfolgreich gegen die Zielsubstanz MOMP (*major outer membrane protein* von *Legionella pneumophila*) (Decker *et al.*, 2000), gegen die auch Mutanten von PH-Domänen gesucht wurden.

Bei der Helix-Bündel-Gruppe von Gerüstproteinen besteht das Gerüst lediglich aus miteinander verknüpften α -Helices. Die sogenannte Z-Domäne leitet sich dabei von Protein A (*Staphylococcus aureus*) ab und besteht nur aus zwei Helices. Von dieser Domäne wurden Bibliotheken von $4 \cdot 10^7$ Mutanten hergestellt und erfolgreich eingesetzt (Nord *et al.*, 1997). Beim Einsatz des cytoplasmatischen Vierhelices-Bündel-Proteins Myohemerythrin wurden nach Randomisierung zweier Loops lediglich instabile und aggregierende Mutanten erhalten, weshalb die Experimentatoren auf das periplasmatische Vierhelices-Bündel-Protein Cytochrom b_{562} auswichen, dort zwei Loops randomisierten und bindende Mutanten erhielten (Ku & Schultz, 1995).

Die Lipocaline besitzen die Struktur eines β -Barrels: es besteht aus acht antiparallelen β -Faltblättern, die sich um eine zentrale Achse winden. Sie werden durch Loops verbunden, über die Lipocaline in der Natur Bindungen zu Kleinmolekülen eingehen können. Nach Randomisierung von vier Loops zu einer Bibliothek von $3.7 \cdot 10^8$ Mutanten konnten Proteine gegen Fluoreszein gefunden werden, die Dissoziationskonstanten von weniger als 35.2 nM aufweisen (Beste *et al.*, 1999). Diese Bibliothek von Lipocalin-Mutanten (genannt „Anticaline“) eignet sich ausgezeichnet zur Bindung von Kleinmolekülen (Weiss & Lowman, 2000).

Auch aktive Zentren von Enzymen lassen sich randomisieren und *in vitro* selektieren. Ein prominentes Beispiel dafür ist die humane, cytosolische Glutathion-S-Transferase. Sie bindet *in vivo* Glutathion. Das aktive Zentrum wurde zu $5 \cdot 10^6$ Mutanten randomisiert, und es ließen sich Mutanten mit neuen

Bindungseigenschaften zu Kleinmolekülen selektieren (Widersten & Mannervik, 1995). Ein weiteres eingesetztes Enzym ist das cytosolische Thioredoxin A. Bei diesem Protein wurde ein Loop randomisiert und eine Bibliothek von $1.8 \cdot 10^8$ Mutanten erfolgreich in einem Flagellen-Selektionsverfahren eingesetzt (Lu *et al.*, 1995).

Durch Fusion randomisierter Peptidsequenzen an Phagenproteine, lassen sich auch direkt bindende Peptidsequenzen selektieren (Wright *et al.*, 1995) (Kay *et al.*, 1993). Diese besitzen allerdings (zumindest ohne Liganden) nur ausnahmsweise eine definierte Struktur.

1.5. Die Verfahren zur Selektion bindender Proteine aus einer Bibliothek

Die zweite Fragestellung dieser Arbeit (s. Kapitel 1.2.) gilt der Suche nach Mutanten der PH-Domäne, die neuartige intermolekulare Wechselwirkungen zu anderen Zielsubstanzen („*targets*“) eingehen können. Derartige Mutanten werden mit Hilfe von Selektionsverfahren aus einer Bibliothek randomisierter Proteine gewonnen.

1.5.1. Das *phage-display*-Verfahren

Dieses Verfahren diente in der vorliegenden Arbeit der Selektion von Mutanten der PH-Domäne. Von den in Kapitel 1.4.2. vorgestellten Gerüstproteinen und den daraus ursprünglich erzeugten Bibliotheken wurde nur die Bibliothek des Thioredoxin A nicht der Selektion durch das *phage-display*-Verfahren sondern einem Flagellen-Selektionsverfahren ausgesetzt (Lu *et al.*, 1995). Allein schon dieser Umstand verdeutlicht die Wichtigkeit des *phage-display*-Verfahrens zur Selektion bindender Mutanten aus einer Bibliothek randomisierter Proteine oder Peptide gegen Zielsubstanzen (Übersichtsartikel: Smith, 1991; Wells & Lowman, 1992; Winter *et al.*, 1994). Im *phage-display*-Verfahren wird das randomisierte Protein oder Peptid genetisch mit dem g3-Protein des Bakteriophagen M13 fusioniert und eine Signalsequenz zur periplasmatischen Translokation vor dieses Konstrukt kloniert (Kay *et al.*, 1996). Coexprimiert man dieses Konstrukt mit einem M13-Helferphagen in *E. coli*, werden durch den Helferphagen die für den Phagen notwendigen Proteine komplementiert. Die Proteine der Phagenhülle einschließlich der g3-Fusion assemblieren in der inneren Membran von *E. coli* (s. Abb. 1.3A) und umhüllen die einzelsträngige

DNA. Während bei der Vermehrung von M13-Phagen *in vivo* die gesamte einzelsträngige Phagen-DNA von der Phagenhülle umschlossen wird, verpackt man (durch geschickte Wahl und Insertion von Selektionsmarkern) im *phage-display*-Verfahren die einzelsträngige DNA für die Protein-g3-Fusion und für den Helferphagen getrennt. Phagen mit der DNA für die Protein-g3-Fusion können zwar *E. coli* Zellen infizieren, jedoch fehlt die DNA für die restlichen Phagenproteine. Die DNA für die Protein-g3-Fusion kann auch doppelsträngig als Plasmid in *E. coli* vorliegen und wird dann Phagemid genannt.

Der artifizielle Phage präsentiert das gewünschte Protein auf der Oberfläche und beinhaltet die genetische Information für die Expression dieses Proteins. Im Phagen liegen also der Phänotyp (das Protein) und der Genotyp (die dafür kodierende DNA-Sequenz) gekoppelt vor. Während einer Generationszeit werden von einer einzigen *E. coli*-Zelle ca. 100-1000 Phagen in das Medium sezerniert (Model & Russel, 1988). Die Phagen lassen sich über Zentrifugations- und Fällungsschritte von den *E. coli*-Zellen trennen. Um aus dieser Bibliothek unterschiedlicher Phagen diejenigen mit erhöhten Affinitätseigenschaften zu Zielsubstanzen herauszulesen, werden die Phagen einer *in vitro*-Selektion („*panning*“) unterzogen (s. Abb. 1.3B und Kapitel 3.9.4.). Sie werden zu der immobilisierten Zielsubstanz gegeben, gegen die ein affines Protein aus der Bibliothek gesucht wird, und je nachdem, wie gut das repräsentierte Protein an der Zielsubstanz bindet, läßt sich der Phage nur schlechter wieder von der Zielsubstanz herunterwaschen. Je stringenter und intensiver gewaschen wird, um so eher werden nur noch hochaffine Proteinmutanten mit ihren Phagen an der Zielsubstanz binden. Dabei reichen selbst geringe Affinitätsunterschiede aus, da eine statistisch aussagekräftige Zahl von Phagen (ca. 10^{10}) eingesetzt wird und somit bei einer Bibliotheksgröße von 10^7 individuellen Klonen auch zu Anfang der ersten Selektionsrunde jede individuelle Mutante im statistischen Mittel noch durch 1000 Phagen vertreten ist.

Die besser bindenden Phagen werden durch Änderung der Lösungsmittleigenschaften mit Triethylamin (Hawkins *et al.*, 1992) oder durch Konkurrenz mit der gelösten Zielsubstanz von der immobilisierten Zielsubstanz eluiert. Mit den Phagen werden dann *E. coli*-Zellen infiziert. Aus ihnen läßt sich die entsprechende Phagemid-DNA präparieren und damit indirekt die Peptidsequenz der bindenden Proteine bestimmen. Nach einer Superinfektion mit Helferphagen lassen sich die angereicherten Phagen aber auch amplifizieren und dann in weiteren Selektionsrunden einsetzen.

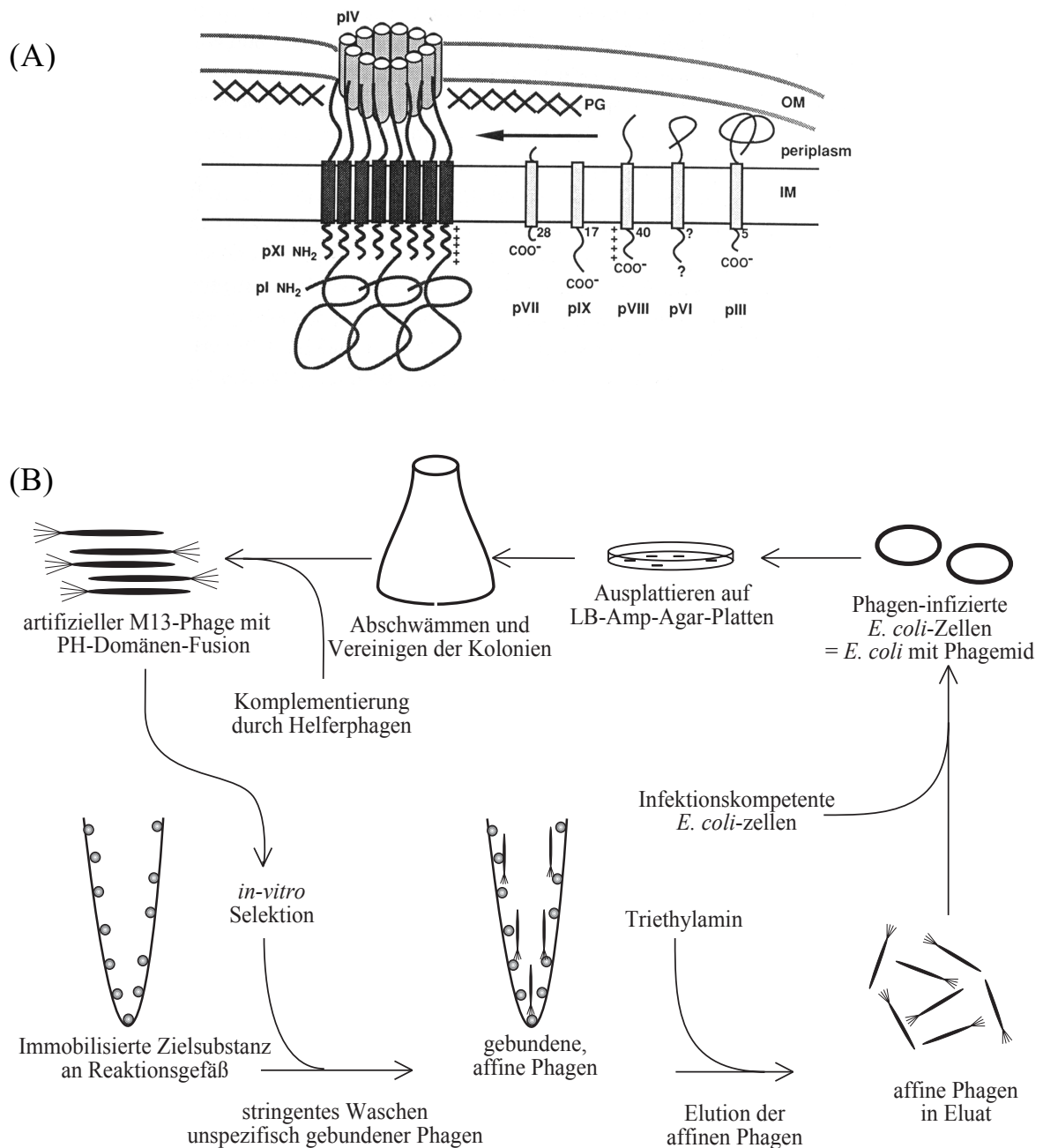


Abb. 1.3: Das phage-display-Verfahren

(A): Schematische Assemblierung von M13-Phagen

Die Proteine der Phagenhülle werden als Membranproteine exprimiert und assemblieren in der inneren Membran von *E. coli* (Abb. entnommen aus Kay *et al.*, 1996). g3-fusionierte Proteine werden zunächst in das Periplasma transloziert.

(B): Selektion über das phage-display-Verfahren

Zur Selektion von Proteinen werden die Phagen mit den fusionierten Proteinen an der immobilisierten Zielsubstanz gebunden und die schwächer bindenden Phagen durch stringentes Waschen entfernt. Die verbleibenden Phagen werden daraufhin eluiert. Mit ihnen lassen sich *E. coli*-Zellen infizieren und neue Phagen herstellen, die einer weiteren Selektionsrunde unterzogen werden können.

Für das *phage-display*-Verfahren existieren einige erweiterte Anwendungsmöglichkeiten. Das g3-Protein besteht aus drei Domänen, die über Linker miteinander verbrückt sind. Die Interaktion eines Proteins mit einer Bibliothek von möglichen Liganden kann man analysieren, indem das g3-Protein durch Deletion eines Linkers in zwei Proteine geteilt wird. Die N-terminale Domäne wird mit dem Protein und die C-terminale Domäne mit der Peptid-Bibliothek fusioniert. Nur bei Bindung einer Peptid-Mutante an das Protein wird ein für die Infektion von *E. coli*-Zellen notwendiges, komplettes g3-Protein auf der Phagenoberfläche dargestellt (Duenas & Borrebaeck, 1994; Spada *et al.*, 1997). Zur Selektion auf besonders stabile Mutanten (Jung *et al.*, 1999) und zur Suche nach Chaperonen (Bothmann & Plückthun, 1998) wurden Varianten des *phage-display*-Verfahrens eingesetzt.

Beim *phage-display*-Verfahren ist die Bibliothek randomisierter Proteine auf der Phagemid-DNA kodiert. Dies setzt die anfängliche, einmalige Transformation von Phagemid-DNA in *E. coli*-Zellen voraus. Damit ist der Größe und Komplexität von *phage-display*-Bibliotheken eine experimentelle Grenze gesetzt. Im Labormaßstab liegt sie bei der auch in dieser Arbeit erreichten Größenordnung von 10^8 Klonen, im industriellen Maßstab existiert sogar eine Bibliothek von scF_V-Fragmenten mit $2 \cdot 10^9$ Klonen (HuCal: *human combinatorial antibody libraries*; Knappik *et al.*, 2000).

Anstelle des Bakteriophagen M13 können auch andere Phagensysteme wie der T7-Phage im *phage-display* eingesetzt werden (Houshmand *et al.*, 1999).

1.5.2. Weitere Selektionsverfahren für Proteine

Das Ribosomen-Display-Verfahren wird zellfrei durchgeführt und setzt somit keine Transformation von DNA voraus (Hanes *et al.*, 2000; Hanes & Plückthun, 1997). Nach Randomisierung einer DNA-Sequenz wird diese *in vitro* transkribiert und die entsprechende mRNA *in vitro* translatiert. (Die Vorgehensweise entspricht der *in vitro* Translation im Rahmen der Importversuche in Mitochondrien: s. Kapitel 4.5.1.). Beim Ribosomen-Display wird verhindert, daß die Ribosomenuntereinheiten am *stop*-Codon dissoziieren. Man erhält eine über die Ribosomen gekoppelte Fusion der mRNA mit dem entsprechenden Protein. Dieses Konstrukt läßt sich analog dem Selektionsverfahren beim *phage-display* einsetzen. Neben der nicht durch Transformation limitierten Größe der Bibliothek von ca. 10^{12} Mutanten bietet dieses System den großen Vorteil, daß die Proteine keinem Transportprozeß unterzogen werden und sich die Bedingungen während der Expression und

Selektion durch Wahl der Temperatur- und Redox-Verhältnisse zum Teil einstellen lassen (Jermutus *et al.*, 2001).

Die Kopplung von Phänotyp (Protein) mit dem Genotyp (DNA) läßt sich auch über DNA-bindende Proteine erreichen. Cytoplasmatisch exprimierte Proteine wurden mit dem *lac*-Repressor fusioniert, was die Bindung des Fusionsproteins am *lac*^{OP} des kodierenden Plasmids ermöglicht. Nach Zellaufschluß ließ sich ein derartiges Konstrukt analog dem *phage-display*-Verfahren einer *in vitro* Selektion unterziehen. Anstelle der Infektion von *E. coli* wurde die Plasmid-DNA direkt transformiert (Cull *et al.*, 1992; Gates *et al.*, 1996). Die Größe der Bibliothek ist wie beim *phage-display*-Verfahren durch die Transformation von Plasmid-DNA in *E. coli*-Zellen limitiert. Ein analoges Verfahren wurde im Rahmen dieser Arbeit vergeblich untersucht, indem über eine Fusion der PH-Domäne mit dem *tet*-Repressor eine stabile, den Zellaufschluß überdauernde Bindung zum *tet*-Operator auf der Plasmid-DNA geknüpft werden sollte.

Die Kopplung von Phänotyp und Genotyp ist bei *in vivo*-Experimenten systemimmanent. Zelluläre Proteine und ihre kodierenden Gene liegen gemeinsam in der Zelle vor. Entscheidend sind bei der Untersuchung von Bibliotheken *in vivo* generell die Transformationseffizienz geeigneter Zellen und die Auswahl eines geeigneten Selektionssystems. Ein übliches Verfahren ist das Hefe-2-Hybrid Verfahren in *Saccharomyces cerevisiae*, mit dem Protein-Protein-Interaktionen herausgefunden werden können (Young, 1998). Da man systembedingt diese Interaktionen im Zellkern analysiert, werden die Proteine einem Transportprozeß in dieses Kompartiment unterzogen. Ein neuartiges Selektionssystem wird derzeit von Bettina Kellersch und Waldemar Kolanus (Genzentrum der LMU München) entwickelt. In diesem System wird eine Proteinbibliothek transient in Jurkat LT-Zellen transfiziert. Inhibiert eine Mutante der Bibliothek den Interleukin-2 Signalweg, verhindert sie damit die Expression der viralen HSV-Thymidinkinase (Li *et al.*, 1997) hinter dem Interleukin-2 Promotor (Zeitlmann *et al.*, 1998). In Verbindung mit Aciclovir oder Ganciclovir wirkt die HSV-Thymidinkinase lethal. Somit lassen nur Mutanten, die diesen Signalweg im Cytosol blockieren, ihre Trägerzelle überleben.

Lassen sich Proteine einer Bibliothek auf der Zelloberfläche exprimieren, können diese Zellen direkt über Affinität selektiert werden (analog dem *phage-display*-Verfahren). Ein Beispiel dafür ist die extrazelluläre Expression des randomisierten Thioredoxins A als Insertion im Flagellenprotein Flagellin bei *E. coli*, das die Präsentation des Proteins auf der Zelloberfläche und eine anschließende *in vitro*-Selektion ermöglicht (Lu *et al.*, 1995). Als zweites Beispiel sei das randomisierte Cystin-Knoten-Protein EETI-II erwähnt, das

fusioniert mit dem Iga_p-Autotransporter auf der Zelloberfläche von *E. coli* dargestellt und selektiert werden kann (Wentzel *et al.*, 1999).

1.5.3. Selektion von affinen Nukleinsäuren und Kleinmolekülen

Die einfachste Form einer Kopplung von Genotyp und Phänotyp ist die Identität von beiden. Einzelsträngige DNA bildet stabile Tertiärstrukturen aus und kann Affinitätseigenschaften aufweisen. Nach Randomisierung können mit dem SELEX-Verfahren (*systematic evolution of ligands by exponential enrichment*) bindende einzelsträngige DNA-Moleküle angereichert und mit Hilfe einer Polymerase-Kettenreaktion (PCR) amplifiziert werden. Ein Beispiel ist die Selektion einer einzelsträngigen, affinen DNA gegen Mikrogefäße des Rattenhirns (Blank *et al.*, 2001).

Mit der reversen Transkriptase ist nicht nur den RNA-Viren die Möglichkeit gegeben, die Information ihrer Transkripte (RNA) in genetische Information (DNA) „zurückzuschreiben“. Sie ermöglicht auch das Konzept der RNA-Aptamere (Famulok & Mayer, 1999), die ebenfalls im SELEX-Verfahren selektiert werden: gerade RNA kann eine definierte Struktur ausbilden, die Katalyse- und Affinitätseigenschaften aufweist. Zur Selektion von RNA-Aptameren wird über randomisierte Oligodesoxynukleotide in einer PCR-Reaktion randomisierte DNA hergestellt, mit einer RNA-Polymerase transkribiert und die RNA in einem Selektionsverfahren eingesetzt. Die bindenden RNA-Moleküle lassen sich mit der reversen Transkriptase in die entsprechende DNA-Sequenz rückübersetzen und sich wieder per PCR-Reaktion amplifizieren. Derartige RNA-Aptamere können im Gegensatz zum DNA-Pendant in eukaryontischen Zellen transkribiert werden und entfalten dort ihre Aktivität (Famulok *et al.*, 2000). Der Nukleaseempfindlichkeit von RNA kann durch Einsatz derivatisierter Nukleotidtriphosphate entgegengewirkt werden.

Anstelle der makromolekularen Proteine und Nukleinsäuren werden auch chemisch synthetisierte Bibliotheken von kleinen Molekülen in Selektionsverfahren eingesetzt. Ein Beispiel ist das Pyridonderivat SUN B8155, das ähnlich mit dem Calcitoninrezeptor interagiert, wie dies Calcitonin selber tut (Katayama *et al.*, 2001).

1.6. Translokationsoptimierung der PH-Domäne

Die Assemblierung der M13-Phagen findet in der inneren Membran der *E. coli* Zellen statt. Zur erfolgreichen Repräsentierung der PH-Domäne auf der Phagenoberfläche muß die PH-Domäne als Teil des Fusionsproteins mit dem g3-Protein in das Periplasma transloziert werden, während das g3-Protein zunächst in der Membran verbleibt (s. Abb. 1.3A). Die PH-Domäne des Cytohesin-1 ist hauptsächlich im Cytoplasma von hämatopoetischen Zellen lokalisiert und wird dort nicht in ein anderes Kompartiment überführt. Während die cytoplasmatische Expression in *E. coli* mit sehr guten Expressionsausbeuten verläuft, kann die PH-Domäne mit nur schlechter Ausbeute und mit unerwarteten Prozessierungsprodukten im Bereich des Signalpeptids in das Periplasma von *E. coli* transloziert werden (s. Kapitel 4.3.7. und 4.3.8.). Dies könnte ein Grund gewesen sein, weshalb erste Versuche, die PH-Domäne im *phage-display*-System zu verwenden, nicht zu bindenden Mutanten führten. Eine translokationsoptimierte PH-Domäne sollte sich besser auf der Phagenoberfläche repräsentieren und im *phage-display*-Verfahren einsetzen lassen.

Welche strukturellen und welche biochemischen Eigenschaften von Proteinen aber beeinflussen die Transportprozesse vom Cytoplasma in das Periplasma bzw. die periplasmatische Expression bei *Escherichia coli*? Dies ist die dritte Frage der allgemeinen Fragestellungen dieser Arbeit (s. Kapitel 1.2.). Ihr wird mit Hilfe gezielter Mutationen an der PH-Domäne und deren Auswirkungen auf die periplasmatische Expression nachgegangen. Inwiefern sich die Ergebnisse für Transportvorgänge verallgemeinern lassen, wird anhand des Imports in die mitochondriale Matrix von *Saccharomyces cerevisiae* untersucht.

1.6.1. Das allgemeine Translokationssystem in *Escherichia coli*

E. coli-Zellen sind nur aus zwei Kompartimenten aufgebaut. Das Cytoplasma wird durch die innere Membran vom Periplasma abgegrenzt. Auf der anderen Seite des Periplasmas bildet die Zellwand eine Barriere gegen das Außenmedium. *E. coli* verfügt nicht über einen Translationsapparat im Periplasma, weshalb alle dort benötigten Proteine durch die innere Membran transloziert werden müssen. Markiert sind diese Proteine durch N-terminale Signalpeptide. Diese Konstrukte werden als „Präproteine“ und nach der Prozessierung der Signalsequenz als „reife“ Proteine bezeichnet. Redox-Proteine der Atmungskette beinhalten Co-Faktoren, die aus dem Cytoplasma in das Periplasma mittransportiert werden. Diese Proteine müssen gefaltet durch die innere

Membran transportiert werden. Ein Charakteristikum dieser Proteine ist ein „*Twin-Arginine-Motif*“ im Signalpeptid (Bogsch *et al.*, 1998; Weiner *et al.*, 1998). Der dazugehörige Transportmechanismus konnte bislang noch nicht geklärt werden. Die meisten periplasmatischen Proteine werden über den allgemeinen sekretorischen Transportmechanismus durch die innere Membran transloziert. Zur rekombinanten, periplasmatischen Expression der PH-Domäne wurde in dieser Arbeit das modifizierte Signalpeptid des ompA-Protein gewählt („*outer membrane protein A*“ auch „*outer membrane protein II*“ von *E. coli*; Beck & Bremer, 1980). Dieses Protein interagiert mit dem SecB-Protein des allgemeinen sekretorischen Transportmechanismus (Kumamoto, 1989), weshalb auch der Transport der PH-Domäne über diesen Transportweg erfolgen sollte.

Der allgemeine Transportmechanismus ist bei *E. coli* gut charakterisiert (Übersichtsartikel: Danese & Silhavy, 1998; Economou, 1999; Economou, 2000; Wickner & Leonard, 1996) und wird hier nur grob skizziert:

Bei Eukaryonten sind während der Translation von sekretorischen Proteinen die Ribosomen an der Oberfläche des endoplasmatischen Reticulums (ER) lokalisiert und das Protein wird noch während der Translation in das ER hineinsynthetisiert (co-translational). Bei *E. coli* erfolgt der Transport durch die innere Membran post-translational, also nach der abgeschlossenen Proteinsynthese (Randall, 1983; Zimmermann & Wickner, 1983). Vor der Translokation befindet sich das Präprotein in einem translokationskompetenten Zustand (Economou, 2000; Kumamoto & Nault, 1989). Dieser wird häufig durch Chaperone wie SRP (*bacterial signal recognition particle*: Lührink *et al.*, 1992) und den Triggerfaktor (Lill *et al.*, 1988) stabilisiert, die am Signalpeptid komplexieren. Zusätzlich lassen sich Interaktionen von Präproteinen mit den Chaperonen SecB (Kim & Kendall, 1998; Randall *et al.*, 1997), GroEL/S (Kandror *et al.*, 1995) oder DnaJ/KGrpE (Hendrick *et al.*, 1983) nachweisen. Die Auswirkungen der Chaperone auf die Faltung des Proteins hinter dem Signalpeptid und die Konsequenzen für den translokationskompetenten Zustand lassen sich dabei nicht einschätzen. Allgemeingültige Aussagen zur Substratspezifität von Chaperonen lassen sich im Fall von SecB treffen (Randall *et al.*, 1997): dieses bei der Translokation häufig involvierte Chaperon kann nur mit Präproteinen interagieren, die eine Mindestgröße von 150 Aminosäuren besitzen - wodurch die Interaktion im Fall der PH-Domäne ausgeschlossen werden kann. Außerdem benötigt das SecB-Chaperon zur Interaktion mit Präproteinen eine Häufung basischer und aromatischer Aminosäuren als Bindungsmotiv (Knoblauch *et al.*, 1999; Topping & Randall, 1994). Welche Voraussetzungen Proteine bieten müssen, um von den verschiedenen Chaperonen komplexiert zu werden, ist ansonsten weitgehend unklar.

Eine Schlüsselrolle für den Transport von Proteinen nimmt das SecA-Protein ein. Es ist im Cytoplasma, als peripheres Membranprotein und als integrales Membranprotein in *E. coli* vorhanden. Selbstregulatorisch bindet es zudem an seiner eigenen mRNA und verhindert damit die weitere Translation (Schmidt & Oliver, 1989). Das lösliche SecA-Protein bindet am Präprotein-Komplex und geleitet ihn zum Translokationsapparat (Fekkes *et al.*, 1998), wo der gesamte Komplex mit dem cytosolisch ausgerichteten Anteil des Sec-Membrankomplexes interagiert (SecY/E/G: Duong & Wickner, 1997). ATP-Hydrolyse führt bei SecA zu einer Konformationsänderung, so daß es zu einem integralen Membranprotein wird. Dabei führt es das Signalpeptid in den Sec(Y/E/G)-Komplex ein und dissoziiert anschließend in das Cytosol (Economou & Wickner, 1994). Durch weitere, mehrfache Anlagerung von SecA-Molekülen werden unter ATP-Hydrolyse jeweils ca. 20 Aminosäurereste durch die Membran gedrückt, was durch den Protonengradienten zwischen beiden Kompartimenten unterstützt wird (Schiebel *et al.*, 1991).

1.6.2. Der mitochondriale Import

Zur Untersuchung und Optimierung der periplasmatischen Expression von PH-Domänen in *Escherichia coli* wurden in dieser Arbeit verschiedene Mutanten hergestellt. Läßt sich die deutlich verbesserte periplasmatische Expression bzw. Translokation einiger Mutanten auch bei anderen Transportsystemen beobachten? Zur Klärung dieser Frage wurden einige Mutanten im mitochondrialen Importsystem eingesetzt. Dies geschieht im Rahmen einer Kooperation mit Michael Brunner, Maxi Endres und Walter Neupert vom Institut für Physiologische Chemie der LMU München. Die Klonierungen für diese Untersuchungen erfolgten als Teil dieser Arbeit, die Versuchsplanungen wurden gemeinsam abgesprochen und die Ergebnisse zusammen diskutiert. Die Durchführung der Versuche oblagen und obliegen Maxi Endres. Sie werden deshalb nicht im Methodenteil dieser Arbeit berücksichtigt.

Der mitochondriale Import ist ein ausgezeichnet charakterisierter Transportvorgang von Proteinen, da er aus historischer Sicht sehr interessant ist und sich zudem gut untersuchen läßt. Nach der Endosymbiontentheorie leiten sich in eukaryontischen Zellen Mitochondrien von Purpurbakterien und Chloroplasten von endocytierten Cyanobakterien ab. Dies erklärt die eigene, zirkuläre DNA und die eigene Transkriptions- sowie Translationsmaschinerie dieser Organellen. Die Proteine der mitochondrialen Matrix werden nur teilweise im Mitochondrium synthetisiert. Ein Großteil der mitochondrialen Proteine liegt im Nukleus kodiert vor. Dies wird durch den Transfer von mitochondrialen

Genen in das nukleäre Genom erklärt. Die entsprechenden Proteine müssen aus dem Cytosol der Zelle in die Mitochondrien importiert werden. Eines dieser Proteine ist bei der Hefe *Neurospora crassa* (i.G. zu *Saccharomyces cerevisiae*) die Untereinheit 9 der F₀-ATPase, die mit Hilfe des N-terminalen Signalpeptids pSu9 in die mitochondriale Matrix transportiert wird (Schmidt *et al.*, 1983). (Das Signalpeptid pSu9 wird in dieser Arbeit als Signalpeptid für den Import der PH-Domänen in die mitochondriale Matrix verwendet). Dazu durchquert das Präprotein die äußere mitochondriale Membran über den TOM-Komplex (*Translocase Outer Membrane*), gelangt in den Intermembranraum und direkt über den TIM-Komplex (*Translocase Inner Membrane*) durch die innere Membran in die mitochondriale Matrix. Strukturell und funktionell sind TOM- und TIM-Komplex gut charakterisiert (Übersichtsartikel: Herrmann & Neupert, 2000; Rassow *et al.*, 1999; Rehling *et al.*, 2001). Die Assemblierung der ATPase in der inneren Membran erfolgt aus der Matrix heraus (Tzagoloff & Meagher, 1972). Analog dem SecA-Protein bei der Translokation von *E. coli* (s. Kapitel 1.5.2.) ist das mitochondriale Matrixchaperon Hitzeschockprotein 70 (mt-Hsp70) daran beteiligt, das Präprotein in die Matrix zu schleusen (Ungermann *et al.*, 1994). Die Modellvorstellung zur Translokation von Präproteinen (Gaume *et al.*, 1998; Neupert *et al.*, 1990) besagt, daß zuerst das Signalpeptid durch den TIM-Komplex in die Matrix diffundiert und durch mt-Hsp70 unter ATP-Hydrolyse arretiert wird. Das restliche Protein *P* liegt dabei noch (partiell) gefaltet im Intermembranraum oder im Cytosol vor und verhindert die weitere Diffusion. *P* befindet sich in einem thermodynamischen Gleichgewichtszustand zwischen gefaltetem und entfaltetem Zustand (s. Kapitel 4.1.4.1.). Befindet es sich im entfaltetem Zustand, diffundiert ein entfalteter Anteil von *P* in die Matrix und wird dort erneut durch mt-Hsp70 arretiert. Dies verhindert die Rückdiffusion von *P*. Die hier beschriebene, vereinfachte Darstellung dieses "Brownschen Motors" berücksichtigt nicht weitere essentielle Proteine des TIM-Komplexes und die Bedeutung des Membranpotentials für den Importvorgang (Ungermann *et al.*, 1996). Die beiden mitochondrialen Membranen unterscheiden sich gerade hinsichtlich ihrer Transporteigenschaften erheblich voneinander. Die äußere Membran ist für kleine Moleküle und Ionen permeabel, während der Transport an der inneren Membran besonders stark reguliert ist. An ihr wird über die Atmungskette ein elektrochemischer Gradient aufgebaut, der zu einem Protonengradienten führt. Dieser wird gemäß der chemiosmotischen Theorie durch die mitochondriale F₀F₁-ATPase (hier als ATP-Synthetase) in ATP umgesetzt (Mitchell, 1979).

Historisch fällt damit den mitochondrialen Transportvorgängen eine Schlüssel-funktion im erweiterten Verständnis um die Endosymbiontentheorie und die chemiosmotische Theorie zu.

2. Material

2.1. Bakterien und Phagen

2.1.1. *Escherichia coli* K12-Stämme

BL 21 (DE3)

hsdS gal (lcIts857 ind1 Sam7 nin5 lacUV5-T7 gene 1)F⁺dcm⁺ ompT lon

Der Stamm wird zur Überexpression plasmidkodierter Gene unter dem starken Promoter des T7-Bakteriophagen verwendet. Das Gen für die T7-RNA-Polymerase ist in das Genom des Stamms integriert und befindet sich dort hinter dem *lac*-Promotor/Operator. Durch Zugabe von IPTG wird die Expression der T7-RNA-Polymerase induziert, die das gewünschte Gen unter dem T7-Promotor transkribiert. Alle BL 21-Stämme sind bezüglich der Proteasen *OmpT* und *Lon* defizient und lassen sich auch für Genexpressionen einsetzen, die nicht unter dem T7-Promotor stattfinden.

CJ 236

F[']cat [pCJ 105 (M13^SCm^r)]/dut1 ung1 thi1 relA1 spoT1 mcrA

Der Stamm dient der Herstellung einzelsträngiger DNA. Diese Bakterien bauen bei der Replikation anstelle von Thymidin Uracil in die DNA ein.

JM 83

F⁻ ara Δ(lac-proAB) rpsL (Str^r) [Φ80 dlacΔ(lacZ)M15] thi

Der Stamm wird zur Expression und anschließenden Translokation von Proteinen in das Periplasma verwendet. Da Bakterien dieses Stammes nicht *recA*-defizient sind, kann das Resistenzgen eines Plasmids in das Genom rekombinieren. Zur Expression in diesem Stamm müssen deshalb frisch transformierte Klone verwendet werden.

JM 109

e14⁻ (McrA⁻) recA1 endA1 gyrA96 thi⁻¹ hsdR17(rK⁻ mK⁺) supE44 relA1 D(lac⁻ proAB) [F' traD36 proAB lacIqZDM15]

Dieser Stamm kann DNA nicht rekombinieren (*recA*) und ist DNase defizient (*endA1*). Er eignet sich somit zum Präparieren von DNA und für Klonierungsexperimente.

SCS 110

rpsL (Strr) thr leu endA thi⁻¹ lacY galK galT ara tonA tsx dam dcm supE44D (lac⁻ proAB) [F' traD36 proAB lacIqZDM15]

Diesem Stamm fehlen die DNA-methylierenden Enzyme *dam*- und *dcm*-Methylase. Die Herstellung unmethylierter Plasmid-DNA ist für die Restriktion mit methylierungssensitiven Enzymen notwendig.

TG 1

supE thi-1 D(lac⁻proAB) D(mcrB-hsdSM)5 (rK⁻ mK⁻) [F' traD36 proAB lacIqZDM15]

Aus diesem Stamm lassen sich elektrokompente Zellen mit besonders hoher Transformationskompetenz herstellen. Er eignet sich deshalb hervorragend zur Herstellung von plasmidkodierten Bibliotheken (z. B. für *phage-display* Bibliotheken).

XL 1-blue

recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI^qZDM15 Tn10 (Tetr)]

Dieser Stamm kann DNA nicht rekombinieren (*recA*) und ist DNase defizient (*endA1*). Er eignet sich somit besonders zum Präparieren von DNA und für Klonierungsexperimente.

Die TG 1-Zellen wurden von Stratagene[®] GmbH (Heidelberg) und alle weiteren *E. coli*-Stämme aus der Sammlung der Arbeitsgruppe bezogen.

2.1.2. Bakteriophage

M13K07

II' (M13mp1), *ori* p15A, *kan*⁺ (Tn903)

M13-Helferphage dient der Herstellung einzelsträngiger DNA und der Komplementierung von Phagen für das *phage-display*-Verfahren (Vieira & Messing, 1987, vertrieben von Bio-Rad Laboratories GmbH, München).

2.2. DNA

2.2.1. Vektoren

Folgende grundsätzlich verschiedene Vektoren wurden in dieser Arbeit verwendet:

pASK75 (Skerra, 1994)

high-copy-Plasmid (Plasmid mit hoher Kopienzahl) zur periplasmatischen Expression unter der Kontrolle des *tet*-Promotors/-Operators. Es besitzt das *tet*-Repressor-Gen in einem dicistronischen Operon mit der β -Lactamase (Ampizillin-Resistenz). Zur regulären Replikation dient ein *colE1*-Replikationsursprung, zur Darstellung einzelsträngiger DNA ein *f1*-Replikationsursprung.

pRSet5d (Schöpfer, 1993)

high-copy-Plasmid zur cytoplasmatischen Expression unter der Kontrolle des T7-Promotors mit der β -Lactamase (Ampizillin-Resistenz). Zur regulären Replikation dient ein *colE1*-Replikationsursprung, zur Darstellung einzelsträngiger DNA ein *f1* Replikationsursprung. Dieser Vektor wurde zur Herstellung unlöslicher Proteinaggregate (*inclusion-bodies*) unterschiedlicher Mutanten der PH-Domäne eingesetzt.

pJG4-5 (Invitrogen™)

high-copy-Plasmid zur Fusion von Proteinen mit der B42-Aktivierungsdomäne von *Saccharomyces cerevisiae* unter dem GAL1-Promotor. Der Vektor besitzt Ampizillin-Resistenz und das TRP-Gen zur Plasmidselektion in Hefe auf Tryptophan-Mangelmedium. Die Replikation in *E. coli* erfolgt über den *colE1* Replikationsursprung. Dieser Vektor dient der Untersuchung der PH-Domänen im Hefe-2-Hybrid-System.

pGEM (Promega™)

high-copy-Plasmid zur *in vitro*-Transkription. Das klonierte Gen läßt sich unter dem SP6-Promotor transkribieren (Butler & Chamberlin, 1982). Dieser Vektor dient der *in vitro*-Expression unterschiedlicher Mutanten der PH-Domäne für Importstudien in Mitochondrien von *S. cerevisiae*.

pACYC184 (Fermentas™)

Dieses Plasmid besitzt anstelle eines *colE1*- einen *rep*-Replikationsursprung. Damit läßt es sich als zweites Plasmid neben den anderen hier aufgeführten Vektoren in *E. coli* etablieren und verleiht *E. coli* Resistenz gegen Chloramphenicol und Tetrazyklin.

2.2.2. Oligodesoxynukleotide

Oligodesoxynukleotide wurden von den Firmen Interactiva (Ulm) und MWG Biotech AG (Ebersberg) sowie von der Arbeitsgruppe von Georg Arnold (Genzentrum der LMU München) nach dem Phosphoramiditverfahren (Sinha *et al.*, 1984) hergestellt. Sie wurden entschützt und entsalzt geliefert.

Sequenzier-Oligodesoxynukleotide

<i>hind-rev (antisense):</i>	5'	TTCACTTCAC	AGGTCAAGC	3'
<i>ompA-for (sense):</i>	5'	GCTATCGCGA	TTGCAGTG	3'
<i>tetpromotor-s (sense):</i>	5'	TTGATAGAGT	TATTTTACC	3'
<i>PHAF2-3AS (antisense):</i>	5'	TTGGAGTCCT	CCACTTCC	3'
<i>PT_CY1G3770AS (antisense):</i>	5'	CCTGTAGCAT	TCCACAGACA	GCCCTC 3'
<i>AH21SAS (antisense):</i>	5'	CTATGCGGCC	CCATTCA	3'
<i>g3pN2-Do-Test s (sense):</i>	5'	CGGGCACTGT	TACTCAAG	3'
<i>jg798.s (sense):</i>	5'	GCTGAGTGGA	GATGCC	3'
<i>T7 (antisense):</i>	5'	TAATACGACT	CACTATAGGG	3'

Oligodesoxynukleotide zur Herstellung einer in drei Loops randomisierten Bibliothek von PH-Domänen

Die randomisierten Desoxynukleotide Adenin (A), Cytosin (C), Guanosin (G) und Thymin (T) sind zuerst mit Kleinbuchstaben, dann mit Umlauten und Ziffern angegeben. Ihre Zusammensetzung (in %) ist folgend tabellarisch aufgeführt.

Zur Randomisierung des ersten Loops:

Loop1rms (sense): 5' TGGCTACTTA AGabcdefgh ijklmnopqr stuvwxyzäA
 AGAGGCGCTG GTTCATTCTGA 3'

mit:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
A	16	06	00	17	12	00	17	12	00	17	12	00	27	22	01	21	05	00	63
C	59	06	52	04	12	04	04	12	04	04	12	04	62	05	52	09	14	49	17
G	10	08	41	72	74	19	72	74	19	72	74	19	07	68	36	59	04	46	19
T	15	80	07	07	02	77	07	02	77	07	02	77	04	05	11	11	77	05	01

	t	u	v	w	x	y	z	ä
A	64	01	62	16	01	06	01	01
C	12	15	10	55	06	01	03	07
G	16	73	21	11	38	05	88	85
T	08	11	07	18	55	88	08	07

Zur Randomisierung des dritten Loops:

Loop2rms (sense): 5' CTTTGAGTAT ACCabcdefg hijklmnopq rstuATCATC
 CCTTTAGAGA ATCTGAG 3'

mit

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
A	62	16	01	20	76	15	63	64	01	15	71	01	08	14	01	27	22	01	17	12	00
C	10	55	06	10	09	56	17	12	15	15	13	21	73	74	33	62	05	52	04	12	04
G	21	11	38	69	11	02	19	16	73	69	09	77	14	07	57	07	68	36	72	74	19
T	07	18	55	01	04	27	01	08	11	01	07	01	05	05	09	04	05	11	07	02	77

Zur Randomisierung des sechsten Loops:

Loop3rms (sense): 5' GGCCGCGAAG ACCabcdefg hijklmnopq rstuvwxyzä
 öü12345678 90TACCGGAT CAGCGCGCCG ACG 3'

mit:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t
A	15	71	01	16	13	00	20	76	15	17	12	00	27	22	01	21	05	00	21	05
C	15	13	21	09	62	01	10	09	56	04	12	04	62	05	52	09	14	49	09	14
G	69	09	77	65	12	45	69	11	02	72	74	19	07	68	36	59	04	46	59	04
T	01	07	01	10	13	54	01	04	27	07	02	77	04	05	11	11	77	05	11	77

	u	v	w	x	y	z	ä	ö	ü	1	2	3	4	5	6	7	8	9	0
A	00	15	71	01	17	12	00	62	62	00	16	73	02	62	16	01	21	05	00
C	49	15	13	21	04	12	04	17	12	01	67	07	54	10	55	06	09	14	49
G	46	69	09	77	72	74	19	20	19	22	12	15	18	21	11	38	59	04	46
T	05	01	07	01	07	02	77	01	07	77	05	05	26	07	18	55	11	77	05

5'-Korrektur der Oligodesoxynukleotide *Loop1rms* und *Loop2rms*:

T3-loop1-ohne.s: 5' GCCAACCGGT TGATTAACCC TCACTAAAGT
 GGCTACTTAA G 3'

T3-loop2-ohne.s 5' GCCAACCGGT TGATTAACCC TCACTAAAGC
 TTTGAGTATA CC 3'

TT3-loop1-ohne.s 5' ATTAACCCTC ACTAAAGTTG GCTACTTAAG 3'

TT3-loop2-ohne.s 5' ATTAACCCTC ACTAAAGTCT TTGAGTATAC C 3'

Oligodesoxynukleotid für den Dimerisierungstest:

loop2-ohne rms.s: 5' CTTTGAGTATACCA 3'

Oligodesoxynukleotide zur Synthese der Circularpermutanten:

PCR-Einführung neuer N- und C-Termini:

<i>PHCir1.s:</i>	5'	GCACCATGGC	GACGCCGAG	GAGAAG	3'
<i>PHCir2.s:</i>	5'	GCACCATGGC	GTCTGACCAA	GTGATCAAGG	3'
<i>PHCir0alt.as:</i>	5'	CACAGGTCAA	GCTTGAATTC	TTATTAGTGA	
		TGGTGATGGT	GATGTGCGGC	CGC	3'
<i>PHCir3alt.as:</i>	5'	GGTGATGTGC	GGCCGAGAC	CCGCGCTGA	TCCGG 3'
<i>PHCir4alt.as:</i>	5'	GGTGATGTGC	GGCCGCCGA	GACCCAGCTT	
		TATTGTCGGG	GA		3'
<i>PHCir5alt.as:</i>	5'	GGTGATGTGC	GGCCGCTTTG	TTGTCACCTG	
		AGGTGATTGC	TGCAATCCAC	TCCTCCTC	3'
<i>OhnLi.s:</i>	5'	AAGGTAGACC	ATGGCGCGA	AGACCGAGGC	3'

Gensynthese der Linkerkonstrukte:

<i>PHCirLiA.s:</i>	5'	CTGCTGAAAG	GTCCTGGCT	AC	3'
<i>PHCirLiA.as:</i>	5'	TTAAGTAGCC	AGGAACCTTT		3'
<i>PHCirLiB.s:</i>	5'	TCAGGTGACA	ACAAAGGTCT	GCAGGCT	3'
<i>PHCirLiB):</i>	5'	CAGCAGAGCC	TGCAGACCTT	TGTTGTCACC	3'
<i>PHCirLiC.s:</i>	5'	CATGGCGGGT	CTGCAGGCT		3'
<i>PHCirLiC.as:</i>	5'	CAGCAGAGCC	TGCAGACCCG	C	3'
<i>PHCirLiD.s:</i>	5'	TCGAGCGGCG	GTGGCAGCGG	CGGCGGC	3'
<i>PHCirLiD.as:</i>	5'	GCCGCTGCCA	CCGCCG		3'
<i>PHCirLiE.s:</i>	5'	TCGAGCGGCG	GTGGCAGCGG	CGGTGGTTCG	
		GGTGGCGGC			3'
<i>PHCirLiE.as:</i>	5'	ACCCGAACCA	CCGCCGCTGC	CACCCCGC	3'
<i>PHCirLiF.s:</i>	5'	GACCGTGAAG	GCTGGCTAC		3'
<i>PHCirLiF.as:</i>	5'	TTAAGTAGCC	AGCCTTCACG	GTCGCCGCC	3'

Oligodesoxynukleotide zur Einführung von Punktmutationen

<i>PHmuKRR_QNV.as:</i>	5'	CAGTCAGAAT	GAACCAAACG	TTCTGCCAGG	
		TCTTTACCCT	G		3'
<i>PHmuKRR_QNV.s:</i>	5'	CAGGGTAAAG	ACCTGGCAGA	ACGTTTGTT	
		CATTCTGACT	G		3'
<i>PHmuL_V.as:</i>	5'	CGAGCTTAAG	aacCCAGCCT	TCTCGG	3'
<i>PHmuL_F.as:</i>	5'	CGAGCTTAAG	AAACCAGCCT	TCTCGG	3'
<i>PHmuT_A.s:</i>	5'	ATCAGCGCGC	CGGCTCCGGA	GGAGAAGGAG	GAG 3'
<i>PHmuTEE_AQQ.s:</i>	5'	ATCAGCGCGC	CGGCTCCGCA	GCAGAAGGAG	GAGTGG 3'
<i>PHompOmpNar.as:</i>	5'	AGCCTTCTCG	GTCTTTTGGC	GCCGCCATGG	CTACGG 3'
<i>MiLi.s:</i>	5'	AAGGTAGACC	ATGGCGTCTG	ACCAAGTGAT	
		CAAGCCAAT	CCAGACCGAG	AAGG	3'

Klonierung der PH-Domäne in unterschiedliche Expressionssysteme

phage-display-Verfahren:

<i>Phage-amber-NheI.as:</i>	5'	ACTTTCAACA	GTCTGTGCGC	TAGCGGCCCC	ATTCAG 3'
<i>PH-pDUCK Not.as:</i>	5'	GTCATGTGCG	GCCGCACCGT	CCCT	3'

Selektion in Jurkat-Zellen:

PHAFLII.as: 5' CTGCCACCTC CGAGCTTAAG TAGCCAGCCT TCTC 3'
PHXHOI.as: 5' CGTAGAAAGG GTCCCTCGAG ATGGCTGCTT TAATGC 3'

Hefe-2-Hybrid:

PJGstop.s: 5' GCCATCTCGA GGGACGGTTA ATGAGGCGAG
AAGCTTTGG 3'
JG1350.as: 5' CAGGTCTGAAG CGGCCGC 3'
RB707 PHBg12.as: 5' GGTTAGATCT TCATTGTTCT GCGGCCTGGT GATGG 3'

2.3. Kommerziell erhältliche Enzyme und Proteine

Maus <i>anti</i> -g3p (pIII) Antikörper	MoBiTec GmbH, Göttingen
dam-Methylase	New England Biolabs GmbH, Schwalbach
Peroxidase-konjugierter	Jackson Immuno Res., West Grove, USA
Kaninchen <i>anti</i> -Maus Antikörper	
Restriktionsendonukleasen	New England Biolabs GmbH, Schwalbach
	MBI Fermentas, St. Leon-Rot
RNase A	Roth GmbH, Karlsruhe
T4-DNA-Ligase	New England Biolabs GmbH, Schwalbach
T4-DNA-Polymerase	New England Biolabs GmbH, Schwalbach
T4-Polynukleotidkinase	New England Biolabs GmbH, Schwalbach
<i>Taq</i> -DNA-Polymerase	MBI Fermentas, St. Leon-Rot
<i>Vent</i> -DNA-Polymerase	New England Biolabs GmbH, Schwalbach

2.4. Standards und Kits

DNA-Längenstandards	MBI Fermentas, St. Leon-Rot
Peroxid-Substratlg. SuperSignal [®]	Pierce, Rockford, IL, USA
Proteinmarker	Bio-Rad Laboratories GmbH, München
Qiagen [®] Mini Plasmid-Kit	Qiagen, Düsseldorf
Qiagen [®] Midi Plasmid-Kit	Qiagen, Düsseldorf
QIAquick Gel Extraction Kit	Qiagen, Düsseldorf
QIAquick PCR Purification Kit	Qiagen, Düsseldorf

2.5. Chemikalien

Adenosin-5'-triphosphat (ATP)	Boehringer Mannheim GmbH, Mannheim
Agarose, <i>elektrophorese grade</i>	Roth GmbH, Karlsruhe
Ammoniumpersulfonsäure (APS)	Merck AG, Darmstadt
Ampizillin (Amp)	Roth GmbH, Karlsruhe
anhydro-Tetrazyklin-hydrochlorid	Acros Organics, Geel, Belgien
Bacto-Trypton	Difco Laboratories, Detroit, USA
Bacto-Yeast	Difco Laboratories, Detroit, USA
Borsäure	Roth GmbH, Karlsruhe
Bromphenolblau	Merck AG, Darmstadt
Chloramphenicol, kristallin	Roth GmbH, Karlsruhe
2'-Desoxynukleotid-5'-triphosphate	Boehringer Mannheim GmbH, Mannheim
di-Kaliumhydrogenphosphat	Merck AG, Darmstadt
N,N-Dimethylformamid	Merck AG, Darmstadt
5,5'-Dithiobis-(2-Nitrobenzoesäure) (DTNB)	Merck AG, Darmstadt
Essigsäure, p.a.	Fluka Chemie, Neu-Ulm
Ethanol, p.a. (EtOH)	Merck AG, Darmstadt
Ethidiumbromid (EtBr)	Merck AG, Darmstadt
Ethylendiamintetraacetat (EDTA)	Roth GmbH, Karlsruhe
Formaldehyd, 37 %, p.a.	Roth GmbH, Karlsruhe
Glyzerin (87 %)	Riedel de Haën, Seelze
Glycin, p.a.	Riedel de Haën, Seelze
Guanidiniumchlorid (GdmCl)	ICN-Biomedicals GmbH, Meckenheim
Harnstoff, p.a.	Roth GmbH, Karlsruhe
Imidazol	Roth GmbH, Karlsruhe
<i>i</i> - β -D-thiogalactosid (IPTG)	Roth GmbH, Karlsruhe
Kalium-dihydrogenphosphat	Merck AG, Darmstadt
Kaliumazetat	Merck AG, Darmstadt
Kaliumchlorid	Merck AG, Darmstadt
Kanamyzin-disulfat	Merck AG, Darmstadt
Magnesiumchlorid	Merck AG, Darmstadt

Magnesiumsulfat-heptahydrat	Merck AG, Darmstadt
Mangan-(II)-chlorid-dihydrat	Merck AG, Darmstadt
2-Mercaptoethanol	Roth GmbH, Karlsruhe
Natriumazetat	Merck AG, Darmstadt
Natriumazid	Merck AG, Darmstadt
Natriumchlorid	Roth GmbH, Karlsruhe
Natriumdodecylsulfat (SDS)	Roth GmbH, Karlsruhe
Natriumhydrogensulfat	Merck AG, Darmstadt
Natriumhydroxid	Merck AG, Darmstadt
Nickelsulfat-hexahydrat	Sigma, St. Louis, U.S.A.
Paraffin, dünnflüssig	Merck AG, Darmstadt
Polyethylenglykol 6000 (PEG)	Roth GmbH, Karlsruhe
Roti [®] -Phenol	Roth GmbH, Karlsruhe
2-Propanol, 99 % zur Synthese	Roth GmbH, Karlsruhe
Silbernitrat, Rotipur [®]	Roth GmbH, Karlsruhe
Tetrazyklin Hydrochlorid (Tet)	Merck AG, Darmstadt
N, N, N', N'-Tetramethylethylen- diamin (TEMED)	Roth GmbH, Karlsruhe
Tris-(Hydroxymethyl)- aminomethan (Tris)	Roth GmbH, Karlsruhe
Xylencyanol FF	Merck AG, Darmstadt

2.6. Geräte

Absorptionsspektrophotometer Uvikon 941	Kontron Instruments GmbH, Neufahrn
Brutschrank für Agarplatten	Binder Labortechnik GmbH, Tuttlingen
Brutschrank für Schüttelkulturen	Adolf Kühner AG, Birsfelden, Schweiz
CD-Spektrometer J710	Jasco GmbH, Groß-Umstadt
Eagle Eye TM II	Stratagene GmbH, Heidelberg
Elektrophoresekammern	Bio-Rad Laboratories GmbH, München

Fluoreszenzspektrophotometer Hitachi F4500	Jouan Meßtechnik GmbH, München
Kühlzentrifuge Universal 16R	Hettich-Zentrifugen, Tuttlingen
Kühlzentrifuge Z233MK	Hermle Labortechnik GmbH, Wehingen
Kühlzentr. Sorvall RC-3B/-5B/-5C	Du Pont GmbH, Bad Homburg
Kugelmühle	Retsch
Pipetten Gilson Pipetman	Abimed, Langenfeld
Refraktometer, Abbésches	Carl Zeiss, Jena
ThermoBlock Typ 5436	Eppendorf GmbH, Hamburg
Thermocycler Varius V45	Landgraf, Langenhagen
Überkopfrotor	Anfertigung der Werkstatt (Genzentrum)
Wasserbad DC1	Gebrüder Haake GmbH, Karlsruhe

2.7. Andere Materialien

Blot-Membran Nitrozellulose	Roth GmbH, Karlsruhe
Dialyse Membran Spectra Por 1 (MWCO: 6-8 kDa) 14,6 mm	Spectrum Medical Ind., Houston, USA
Eppendorf-Reaktionsgefäße	Eppendorf GmbH, Hamburg
Film (Kodak Scient. Imaging Film, BIOMAX™ MR)	Kodak, Rochester, New York, USA
<i>Gel drying film</i>	Promega, Madison, Wisconsin, USA
Glasperlen Ø 0.1-0.2 mm	Bernd Euler, Frankfurt
Magnetrührfische	Roth GmbH, Karlsruhe
Nickel-NTA-Agarose	Qiagen, Hilden
NUNC-Maxisorb-Tubes	NUNC, Paris
Plastikverbrauchsmaterialien	Peske, Aindling-Pichl
Quarzglas-Küvetten	Hellma GmbH & Co., Müllheim/Baden
Sterilfilter Millex GV	Millipore S.A., Molsheim, Frankreich
Zentrifugenflaschen 11 / GS3 / GSA / SS34	Nalgene Company, Rochester, N.Y., USA

3. Methoden

Die Vorschriften orientieren sich überwiegend an den von Sambrook *et al.* publizierten Protokollen (Sambrook *et al.*, 1989). Medien werden mit entionisiertem, Puffer mit zusätzlich ultrafiltriertem Wasser angesetzt. Mischungsverhältnisse sind für flüssige Substanzen als Volumen pro Volumen (v/v), für Lösungen fester Stoffe als Masse pro Volumen (w/v) angegeben. Molekularbiologische Arbeiten werden mit sterilen Medien und Gefäßen durchgeführt. Nach dem Einsatz von Bakterien werden alle kontaminierten Gefäße und Flüssigkeiten autoklaviert.

3.1. Plasmid-DNA in *Escherichia coli*

3.1.1. Kultivierung und Aufbewahrung von *Escherichia coli*

Einzelkolonien werden durch einen Verdünnungsausstrich auf Luria-Bertani-Kulturschalen (LB-Platten, je nach Stamm und Plasmid mit den entsprechenden Antibiotika versetzt) nach Inkubation für 12-15 h bei 37°C erhalten. Die Platten können, mit Parafilm verschlossen, etwa einen Monat bei 4°C gelagert werden. Kulturschalen werden generell auf dem Deckel liegend aufbewahrt. Für Flüssigkulturen von *E. coli* wird LB-Medium je nach Selektion (z.B. auf plasmidtragende Zellen) Antibiotikum zugesetzt. Nach Animpfen mit einer Einzelkolonie wird, soweit nicht anders angegeben, für ca. 12-15 h mit 200 Upm bei 37°C geschüttelt. Die Zelldichte wird über die optische Dichte bei 600 nm (OD_{600}) verfolgt. Zur Bestimmung dient ein Absorptionsspektrophotometer (Uvikon 941). Da es sich bei der Messung der Zellzahl in erster Linie um Streueffekte handelt, diese Geräte aber nur zur Messung von Absorptions-Phänomenen konzipiert sind, muß für jedes Gerät eine Eichkurve erstellt werden, mit deren Hilfe sich von der optischen Dichte auf die Zellzahl pro Volumen zurückschließen läßt.

LB-Medium

10 g	Bacto Trypton
5 g	Bacto Yeast Extract
5 g	NaCl
1.3 ml	2 M NaOH

wird ad 1 l mit entionsiertem Wasser aufgefüllt, autoklaviert und bei RT gelagert; für Kulturschalen + 1.5 % Bacto Agar

Antibiotika-Stammlösungen:

Ampizillin (Amp)	100 mg/ml	in H ₂ O
Tetrazyklin (Tet)	50 mg/ml	in Ethanol
Chloramphenicol (Cam)	30 mg/ml	in Ethanol
Kanamyzin (Kan)	70 mg/ml	in H ₂ O

Die sterilfiltrierten, aliquotierten Stammlösungen werden bei -20°C gelagert und 1/1000 verdünnt eingesetzt.

Zur längerfristigen Aufbewahrung von *E. coli*-Stämmen werden Glyzerinkulturen angelegt. Hierzu werden 700 µl einer Übernachtskultur mit 300 µl sterilem 87 % Glyzerin in einem 2 ml Eppendorf-Reaktionsgefäß gemischt und bei -80°C gelagert. Glyzerinkulturen können über mehrere Jahre aufbewahrt werden.

3.1.2. Elektroporation von Plasmid-DNA in elektrokompetente *E. coli*

Herstellung elektroporationskompetenter *E. coli*

Mit 20 ml Vorkultur des gewünschten *Escherichia coli* K12-Stammes werden 2 l LB-Medium in einem 5 l-Erlenmeyer-Kolben angeimpft und bei 37°C, 100 Upm bis zu einer OD₆₀₀ = 0.5 bis 0.6 angezogen. Die Kultur wird dann im Kühlraum auf Eis gestellt und mindestens 1 h leicht geschüttelt (40 Upm). Die Kultur wird in einem auf 0°C gekühlten Rotor bei 0°C in GSA-Zentrifugenbechern 10 min bei 3 000 *rcf* zentrifugiert. Das Pellet wird in 100 ml eiskaltem Wasser mit einer Pipette resuspendiert und erneut 10 min bei 3 000 *rcf* zentrifugiert. Das Pellet wird in 30 ml eiskaltem GYT-Medium aufgenommen und wiederum 10 min bei 3 000 *rcf* und 0°C zentrifugiert. Das Pellet (ca. 2 ml) wird mit 2 ml GYT-Medium gründlich resuspendiert und die Suspension als Aliquots à 30 µl in flüssigem Stickstoff schockgefroren und bei -70°C gelagert.

Die Kompetenz der Zellen sollte bei 10^9 Klonen / μg DNA liegen; d.h. nach Elektroporation von 0.01 ng pUC19-Plasmid-DNA sollte man 10 000 Klone erhalten.

Die Angabe der Elektrokompetenz bezieht sich immer auf sehr geringe DNA-Mengen (< 1 ng), die für eine Elektroporation eingesetzt werden!

Für die Herstellung von Bibliotheken wurden *E. coli* TG 1 der Firma Stratagene™ eingesetzt, die eine Kompetenz von 10^{10} Klonen / μg pUC19 erzielen.

GYT-Medium:

2.5 g Trypton
1.25 g Hefe-Extrakt
100 ml Glyzerin

wird ad 1 l mit entionisiertem Wasser aufgefüllt, autoklaviert und bei RT gelagert.

Elektroporation kompetenter Zellen mit Plasmid-DNA

Ein 30 μl -Aliquot elektrokompentener *E. coli* wird auf Eis aufgetaut, mit maximal 20 μl salzfreier DNA (maximal 20 ng) versetzt und in eiskalte Elektroporationsküvetten (Abstand der Kondensatorplatten: 2 mm) pipettiert. Die Elektroporation erfolgt mit einer Spannung von 1.25 kV pro mm, wobei die Zeitkonstante ca. 4 ms betragen sollte. Anschließend werden die Zellen sofort mit 1 ml SOB-Medium verdünnt und in ein Reagenzglas überführt. Nach 1 h Inkubation (37°C, 200 Upm) werden die *E. coli*-Zellen mit Antibiotikum (je nach Resistenz durch das elektroporierte Plasmid) unter Selektionsdruck genommen. Dies kann auf LB-Platten oder in Flüssigkultur erfolgen.

SOB-Medium:

20 g Bacto Trypton
5 g Bacto Yeast Extract
7.8 mM NaCl
2.5 mM KCl

pH 7.0 (NaOH), wird ad 1 l mit Wasser aufgefüllt und autoklaviert. Dies ergibt SOC-Medium. Vor Gebrauch werden die steriltfiltrierte Glukose- und MgCl_2 -Lösungen zugegeben:

20 mM Glukose
10 mM MgCl_2 .

Zur Herstellung einer großen Zahl Klone ist die Elektroporation optimal. Das Verfahren wurde in der vorliegenden Arbeit zur Herstellung von Bibliotheken verwendet.

Dazu werden ca. 2 µg ligierte DNA (s. Kapitel 3.2.2.) in 10 Einzel-elektroporationen eingesetzt und daraus insgesamt ca. 10⁷ Klone erzielt.

3.1.3. Transformation von *E. coli* mit Plasmid-DNA

Herstellung transformationskompetenter Zellen (nach Tang *et al.*, 1994)

200 ml LB-Medium werden in einem 2 l-Erlenmeyerkolben mit 20 ml Vorkultur des gewünschten *E. coli*-Stammes angeimpft. Die Kultur wird bei 200 Upm und 37°C bis zu einer OD₆₀₀ von 0.9 geschüttelt, durch Schwenken in einem Eiswasserbad abgekühlt und in Zentrifugenbecher überführt. Alle weiteren Schritte erfolgen unter strikter Eiskühlung.

Die Bakteriensuspension wird 15 min bei 3 000 *rcf* und 4°C zentrifugiert. Das Pellet wird in 40 ml Tfb I-Lösung resuspendiert und die Suspension in 50 ml-Falkongefäße überführt. Nach 15 minütiger Zentrifugation bei 4°C, 2 000 *rcf* wird das Pellet in 8 ml Tfb II-Lösung aufgenommen, zu je 50 µl aliquotiert, sofort in flüssigem Stickstoff eingefroren und bei -70°C gelagert.

Die Kompetenz der Zellen sollte bei 10⁶ Klonen / µg DNA liegen.

Tfb I

100 mM KCl
50 mM MnCl₂
30 mM Kaliumazetat
10 mM CaCl₂
15 % Glycerin
wird mit Wasser aufgefüllt und sterilfiltriert.

Tfb II

75 mM CaCl₂
10 mM KCl
10 mM NaMOPS, pH 7.0
15 % Glycerin
wird mit Wasser aufgefüllt und sterilfiltriert.

Transformation kompetenter Zellen mit Plasmid DNA (nach Pope & Kent, 1996)

Ein Aliquot kompetenter Zellen wird auf Eis aufgetaut und mit ca. 1 ng Plasmid-DNA versetzt. Der Ansatz wird auf eine bei 37°C vorgewärmte LB-Kulturschale mit geeignetem Antibiotikum plattiert und diese bei 37°C für 12-15 h inkubiert. Bei einer Kompetenz von 10^6 Klonen / μg DNA sollten nach ca. 12 h 1000 Kolonien sichtbar sein.

Dieses Verfahren wurde bevorzugt für Klonierungsexperimente (in *E. coli* XL 1-blue) oder zur Herstellung von Einzelklonen zwecks Proteinexpression verwendet.

3.1.4. Plasmid-DNA-Präparation

Alkalische Lyse (nach Birnboim, 1979)

2 ml einer Übernacht-Flüssigkultur eines plasmidtragenden *E. coli*-Stammes werden in Eppendorf-Reaktionsgefäße überführt und 2 min mit 18 000 *rcf* zentrifugiert. Das Zellsediment wird in 100 μl Lösung I resuspendiert, mit 200 μl Lösung II versetzt, durch Invertieren gemischt und 5 min bei RT inkubiert. Anschließend wird durch Zugabe von 150 μl Lösung III neutralisiert. Nach 5-minütiger Inkubation bei RT wird mit 400 μl TE-gesättigter Phenol/Chloroform-Lösung kurz gevortext und 10 min bei 4°C, 18 000 *rcf* zentrifugiert.

Die obere, wäßrige Phase enthält die Plasmid-DNA. Sie wird in ein neues 2 ml-Reaktionsgefäß überführt, zur Extraktion von Phenolrückständen zweimal mit je 500 μl Chloroform gevortext und nach je 1 min mit 18 000 *rcf* Zentrifugation die plasmidhaltige, wäßrige Phase in ein weiteres Reaktionsgefäß pipettiert. Nach Zugabe des 2.5-fachen Volumen (1 ml) an 4°C kaltem absolutem Ethanol und anschließendem Durchmischen wird die gefällte Plasmid-DNA durch Zentrifugation bei 4°C mit 18 000 *rcf* pelletiert. Der Überstand wird verworfen, das Pellet mit 0.5 ml 70 % Ethanol (RT) gewaschen, 10 min bei 4°C zentrifugiert und bei RT vollständig getrocknet. Zur Entfernung der RNA wird das Pellet in 20 μl TE/RNaseA gelöst. Die DNA wird bei -20°C gelagert. Bei *high-copy*-Plasmiden ist eine Ausbeute von ca. 3-5 μg Plasmid-DNA pro ml Bakterienkultur zu erwarten.

Lösung I:

50 mM Tris/HCl, pH 8.0
10 mM EDTA

Lösung II:

0.2 M NaOH
1 % SDS

Lösung III:

60 ml 5 M Kaliumazetat
11.5 ml Eisessig
28.5 ml Wasser
damit beträgt die Konzentration von K^+ 3 M und von Azetat 5 M.

TE:

10 mM Tris/HCl, pH 8.0
1 mM EDTA

Plasmid-Schnellpräparation

Hierfür dient das unter Kapitel 3.1.4.1. beschriebene Verfahren, allerdings wird bereits nach Zugabe von Lsg. III zentrifugiert, der Phenol-Extraktionsschritt übersprungen und das Pellet mit Ethanol gefällt sowie gewaschen. Nach Trocknung wird es in TE/RNase aufgenommen. Die isolierte DNA ist zwar durch Proteine verunreinigt, doch beeinträchtigt dies die analytische Restriktionsverdau meistens nicht. Diese kostengünstige Präparation sollte nur mit den DNase-defizienten *E. coli*-Stämmen XL 1-blue oder JM 109 erfolgen.

Qiagen-Plasmid-DNA-Präparation

(Präparation gemäß der Anweisung des Herstellers)

3.2. Restriktionsendonukleaseverdau von DNA

Die Restriktionsenzyme werden unter den vom Hersteller empfohlenen Bedingungen verwendet. Muß DNA mit unterschiedlichen Enzymen restringiert werden, für die der Hersteller keinen gemeinsamen Reaktionspuffer empfiehlt, wird zunächst mit dem Enzym geschnitten, das eine niedrigere Salzkonzentra-

tion benötigt. Nach Hitzeinaktivierung, meist 20 min bei 65°C, werden die Pufferbedingungen angeglichen und mit dem zweiten Enzym restringiert.

Analytisch:

Zu etwa 1 µg Plasmid-DNA wird 1 µl des empfohlenen 10 x Reaktionspuffers und 1 µl einer BSA-Lösung (1 mg/ml, Endkonzentration 100 µg/ml) gegeben. Der Ansatz wird mit sterilem Wasser auf 10 µl aufgefüllt und dann 1-2 U der entsprechenden Restriktionsendonukleasen dazupipettiert. Dieser Ansatz wird mindestens 1 h bei der für das jeweilige Enzym optimalen Temperatur inkubiert.

Durch Elektrophorese in Agarosegelen kann man DNA-Fragmente unterschiedlicher Größe aufgrund der verschiedenen Mobilität einfach trennen. Zur Analyse von DNA-Fragmenten in der Größenordnung von 400-6000 bp werden in der Regel 1 %-ige Agarosegele, bei kleineren Fragmenten 1.4-1.7 %-ige verwendet.

Zur Herstellung von Agarosegelen wird trockene Agarose im entsprechenden Volumen 1 x TBE-Puffer in einem Mikrowellenherd aufgeköcht, bis eine homogene, viskose Lösung entsteht. Zu der heißen Lösung wird ein 1/10 000 Volumen Ethidiumbromid-Lösung. (10 mg/ml) gegeben, um nach der Elektrophorese die DNA durch die Fluoreszenz des interkalierenden Farbstoffes im UV-Licht (254 oder 366 nm) sichtbar zu machen. Nach Abkühlen auf unter 50°C wird das Gel in die entsprechende Apparatur gegossen und ein Taschenformer eingeschoben. Die Proben werden mit 6 x Blaumarker gemischt und dann in die Taschen des erstarrten Gels aufgetragen. Neben den Proben wird ein Längenstandard aufgetragen, mit dessen Hilfe die Größen der einzelnen Fragmente durch Vergleich ermittelt werden können. Über diesen Standard lassen sich auch die DNA-Konzentration der Proben abschätzen, indem die Intensitäten der Banden mit der Intensität einer Bande des Standards verglichen werden.

Gele der Größe 10 x 15 cm werden bei 80-100 V, Gele der Größe 6 x 8 cm bei ca. 50 V in 1 x TBE-Puffer gefahren. Die Laufzeiten betragen zwischen 30 und 60 min. Die Gele werden mit der digitalen Kamera Eagle Eye II dokumentiert.

1 x TBE-Puffer:

89 mM Tris
89 mM Borsäure
2 mM EDTA

6 x Blaumarker:

- 1.5 mg/ml Bromphenolblau
- 1.5 mg/ml Xylencyanol
- 50 % Glyzerin, anschließend wird mit Wasser auf das Gesamtvolumen aufgefüllt.

3.2.1. Extraktion der DNA aus Agarosegelen

Die Extraktion der DNA aus Gelbanden wird mit Hilfe des QIAquick *Gel Extraction Kits* nach den Angaben des Herstellers durchgeführt.

3.2.2. Ligation von DNA-Fragmenten

Zur Umklonierung werden das gewünschte Vektorplasmid und Insert mit den gleichen Enzymen gespalten, die Fragmente präparativ isoliert und aus dem Agarosegel extrahiert. Die gewünschte Kombination von Vektor und Insert wird in einem molaren Verhältnis von etwa 1:10 bis 1:1 gemischt (Endvolumen 20 µl). Nach Zugabe von 1/10 Volumen 10 x Ligasepuffer wird auf RT abgekühlt und ca. 2 U T4-DNA-Ligase dazugegeben. Der Ansatz wird 1 h bei 20°C inkubiert. 5 µl davon werden anschließend in kompetente *E. coli* transformiert. Als Kontrolle dient ein Ligationsansatz mit Vektor-DNA ohne Insert.

3.3. Einführung gezielter Mutationen

3.3.1. Präparation von Einzelstrang-DNA

(nach Sambrook *et al.*, 1989)

1 ml 2 x YT-Medium (mit 30 µg/ml Chloramphenicol und Antibiotikum zur Selektion auf das Plasmid) wird mit einer Einzelkolonie von *E. coli* CJ 236 angeimpft und bei 37°C mit 200 Upm geschüttelt. Sobald in der Kultur eine leichte Trübung erkennbar wird, wird sie zu 20 ml gleichen Mediums in einen 100 ml Erlenmeyerkolben gegeben und mit dem Phagen M13K07 (Endkonzentration 2×10^7 pfu/ml) transfiziert. Nach zweistündigem Schütteln bei 37°C

und 250 Upm wird Kanamyzin (Endkonzentration 70 µg/ml) zur Selektion auf Phageninfektion zugesetzt und weiter über Nacht inkubiert (37°C, 250 Upm).

Die Übernachtskultur wird zweimal 5 min bei 4°C und 12 000 *rcf* zentrifugiert und das Zellpellet jeweils verworfen. Die Phagenpartikel werden aus dem Überstand durch Zugabe von 135 µl 20 % PEG 6000 in 2.5 M NaCl-Lösung pro ml Überstand gefällt. Nach 15 min Inkubation bei RT und 5 min Zentrifugation bei 12 000 *rcf* und 4°C wird das Pellet durch Vortexen in 500 µl TE aufgenommen und in Eppendorf-Reaktionsgefäße überführt. Nach Aufschluß der Phagenpartikel durch 1 Vol. TE-gesättigten Phenols wird bei 4°C mit 18 000 *rcf* zentrifugiert. Der Überstand wird in ein neues Reaktionsgefäß überführt und die phenolische Phase erneut mit 200 µl TE-Puffer extrahiert. Die Überstände werden in einem Corex-Glas vereinigt und die Einzelstrang-DNA nach Zugabe von 0.1 Vol. 8 M LiCl-Lösung mit 2.5 Vol. EtOH 1 h bei -20°C gefällt. Nach Zentrifugation (20 min, 4°C, 12 000 *rcf*) wird das Pellet mit 70 % EtOH gewaschen, nochmals zentrifugiert (10 min) und bei RT getrocknet. Die DNA wird in 25-50 µl TE-Puffer aufgenommen und bei -20°C gelagert.

Durch Elektrophorese eines Aliquots im Agarosegel kann die Reinheit und durch Extinktionsmessung bei 260 nm die Ausbeute abgeschätzt werden ($OD_{260} = 1$ entspricht etwa 33 µg/µl Einzelstrang-DNA). Es sind etwa 1 µg DNA pro ml Übernachtskultur zu erwarten.

2 x YT-Medium:

16 g/l Bacto Trypton

10 g/l Bacto Yeast Extrakt

5 g/l NaCl

wird mit NaOH auf pH 7.4 eingestellt, mit Wasser ad 1 l aufgefüllt und autoklaviert.

PEG-Lösung:

20 % PEG 6000

2.5 M NaCl

wird mit Wasser aufgefüllt und sterilfiltriert.

3.3.2. Phosphorylierung von Oligodesoxynukleotiden

Oligodesoxynukleotide werden in sterilem Wasser zu einer Endkonzentration von 100 pmol/µl gelöst. Zu 200 pmol eines Oligodesoxynukleotids werden

10 µl 3 x Kinase-Mix sowie 4.5 U T4-Polynukleotidkinase gegeben, der Ansatz wird mit Wasser auf 30 µl aufgefüllt. Nach einer Inkubationszeit von 45 min bei 37°C wird die Kinase durch zehnminütige Behandlung bei 65°C inaktiviert. Die Lösung wird mit 70 µl Wasser aufgefüllt (Endkonzentration der phosphorylierten Oligodesoxynukleotide: 2 pmol/µl) und bei -20°C gelagert.

3 x Kinase Mix:

300 mM	Tris/HCl, pH 8.0
60 mM	DTT
30 mM	MgCl ₂
3 mM	ATP

wird zu 10 µl aliquotiert und ist bei -20°C lagerfähig.

3.3.3. Ortsspezifische Mutagenese

(nach Kunkel, 1985)

Für die ortsspezifische Mutagenese werden 5'-phosphorylierte Mutagenese-Oligodesoxynukleotide eingesetzt. 0.2 pmol Einzelstrang-DNA werden mit 2 pmol Mutagenese-Oligodesoxynukleotiden und 1 µl 10 x *annealing*-Puffer gemischt und mit Wasser auf 10 µl aufgefüllt. Der Ansatz wird 5 min auf 95°C erhitzt, um die Sekundärstrukturen der DNA zu lösen, anschließend für 15 min bei der spezifischen Hybridisierungstemperatur inkubiert und danach sofort auf Eis gestellt. Zur *in vitro*-Polymerisation werden dem Ansatz 1 µl 10 x Synthese-Puffer, 1 µl T4-DNA-Polymerase (3 U) und 1 µl T4-DNA-Ligase (6 U) zugegeben. Nach vorsichtigem Schütteln wird 5 min auf Eis und 90 min bei 37°C inkubiert.

Vor der Transformation in *E. coli* kann die jetzt doppelsträngige Plasmid-DNA *dam*-methyliert werden. Dazu werden zu dem Reaktionsansatz 1 µl 15 x *dam*-Methylase-Puffer, 0.5 µl S-Adenosyl-methionin-Lösung und 1 µl *dam*-Methylase (5 U) gegeben und 1 h bei 37°C inkubiert. Das Enzym wird anschließend 15 min bei 65°C inaktiviert. Der Mutageneseansatz wird mit TE-Puffer auf ein Volumen von 50 µl aufgefüllt und kann bei -20°C gelagert werden. Zur Transformation in *E. coli* werden 10 µl eingesetzt.

10 x annealing-Puffer:

500 mM	NaCl
200 mM	Tris/HCl, pH 7.4
20 mM	MgCl ₂

10 x Synthese-Puffer:

100 mM	Tris/HCl, pH 7.9
50 mM	MgCl ₂
20 mM	DTT
10 mM	ATP
5 mM	jedes dNTPs

15 x dam-Methylase Puffer:

300 mM	Tris/HCl, pH 7.5
255 mM	EDTA
75 mM	2-Mercaptoethanol

S-Adenosyl-methionin-Lösung:

2.5 mM	S-Adenosyl-methionin
--------	----------------------

Alle (wäßrigen) Lösungen werden sterilfiltriert und bei -20°C aufbewahrt.

Die Hybridisierungstemperatur T_M wird nach der „2 + 4-Regel“ berechnet:

$$T_M = 2 (n_A + n_T) + 4 (n_G + n_C)$$

mit n_i der Anzahl der entsprechenden Basen im Bereich der Hybridisierung.

3.3.4. Polymerase-Kettenreaktion (PCR) zur Einführung von gezielten Mutationen

Zur gezielten Mutation von Plasmid-DNA über eine PCR wird ein mutierendes Oligodesoxynukleotid (z.B. *PHmuT_A.s* s.u.) und ein Gegenstrang-Oligodesoxynukleotid verwendet, die an je einem Strang der als Templat vorliegenden Plasmid-DNA hybridisieren. Nach der Hybridisierung (*annealing*) erfolgt eine 3'-Verlängerung der Oligodesoxynukleotide (*extension*) mit Hilfe der Desoxynukleotidtriphosphate (dNTP) und einer thermostabilen DNA-Polymerase (*Vent*-Polymerase bzw. *Pfu*-Polymerase mit Korrekturleseaktivität oder *Taq*-Polymerase ohne Korrekturleseaktivität). Da die verlängerten Oligodes-

oxynukleotide nach der Schmelze (*denaturation*) ebenfalls als Templat für eine weitere Hybridisierungs- und Verlängerungsreaktion dienen, erfolgt eine exponentielle Amplifikation der DNA. Nach einer abschließenden Verlängerung liegt doppelsträngige DNA (dsDNA) vor, die nach Reinigung mit geeigneten sequenzspezifischen Endonukleasen geschnitten und anschließend kloniert werden kann.

Ein mutierendes Oligodesoxynukleotid sollte exemplarisch folgendermaßen geplant sein:

PHmuT_A.s: 5' ATCAG**CGCGC** CGgctCCgGA GGAGAAGGAG GAG 3'

- fett**: Restriktionsschnittstelle für spätere Klonierung (hier: *Bss*HII)
klein: gegenüber Templat geänderte Desoxynukleotide
kursiv: gleiche Sequenz wie der sense-Strang des Templats, zur Hybridisierung
3'-Ende: hieran findet nach Hybridisierung am antisense-Strang des Templats die Verlängerung des Oligodesoxynukleotids statt.

Für eine Standard-PCR werden folgende Substanzen eingesetzt:

10 x PCR-Reaktionspuffer	5.0 µl
dNTP (2 mM)	5.0 µl
MgCl ₂ (25 mM)	5.0 µl
Templat: Plasmid-DNA (0.5 ng/µl)	2.0 µl
<i>sense</i> Oligodesoxynukleotid (10 pmol/µl)	2.0 µl
<i>antisense</i> Oligodesoxynukleotid (10 pmol/µl)	2.0 µl
<i>Taq</i> -Polymerase (1 U / µl)	1.0 µl
steriles Wasser	ad 50.0 µl

Der Einsatz anderer DNA-Polymerasen (*Vent*-Polymerase, *Pfu*-Polymerase) erfordert häufig eine Modifizierung der PCR-Bedingungen (Mg²⁺-Konzentration, Temperaturverlauf); dafür ist aufgrund der 3'→5'-Korrekturleseaktivität dieser DNA-Polymerasen die Fehlerrate für den Einbau von dNTPs geringer. Qualitativ und quantitativ lassen sich PCR-Reaktion meistens zusätzlich verbessern, indem anstelle des gesamten Plasmids nur das Fragment des Templats zugegeben wird, das amplifiziert werden soll. Die Amplifikation der

Templat-DNA erfolgt in einem PCR-Thermoblock; dort wird die dsDNA zunächst aufgeschmolzen, die Oligodesoxynukleotide bei der spezifischen *annealing*-Temperatur gebunden und durch die Polymerase zur dsDNA aufgefüllt. Zum Schutz vor Verdunstung werden alle PCR-Ansätze mit flüssigem Paraffin überschichtet. Folgendes Programm-Muster wird verwendet:

PCR-Programm:

60 s bei 94°C (*denaturation*)

25 Amplifikationszyklen mit jeweils

30 s bei 94°C (*denaturation*)

60 s bei T_m (*annealing*)

90 s bei 72°C (*extension*)

Ende: 360 s bei 72°C

Nach Ende der PCR-Reaktion werden die Ansätze durch analytische Gelelektrophorese überprüft. Die amplifizierte DNA kann mit Hilfe des Qiagen PCR *Purification* Kits nach Anleitung der Hersteller gereinigt werden, um verbleibende dNTPs und Polymerase zu entfernen.

3.3.5. Reassemblierungs-PCR und *QuikChange* zur Einführung von Punktmutationen

Liegt keine geeignete Restriktionsschnittstelle in unmittelbarer Nähe zur Mutation vor, wird alternativ indirekt das 5'-Ende durch Reassemblierungs-PCR (*reassembly* PCR) erweitert oder die *QuikChange*-Methode (nach Stratagene) angewendet.

Reassemblierungs-PCR (*reassembly* PCR)

Für eine Reassemblierungs-PCR wird in einem ersten Schritt neben dem wie im vorherigen Kapitel aufgeführten PCR-Produkt noch ein weiteres PCR-Produkt hergestellt. Dazu verwendet man ein genau inverses Oligodesoxynukleotid zu dem mutierenden Oligodesoxynukleotid und ein weiteres Gegenstrang-Oligodesoxynukleotid. Man erhält damit zwei PCR-Produkte, die nur im Bereich der Mutation identisch sind, ansonsten aber die Sequenz des Templats vor bzw. nach der Mutationsstelle aufweisen. Nach der Aufreinigung der PCR-Fragmente

läßt sich in einer weiteren PCR-Reaktion mit beiden PCR-Produkten des ersten Schritts als Template und den verwendeten Gegen-Oligodesoxynukleotiden ein PCR-Produkt herstellen, in dessen Mitte sich die gewünschte Mutation befindet. Dieses PCR-Produkt wird mit Hilfe von sequenzspezifischen Endonukleasen geschnitten und in einen Vektor kloniert. Dieses Verfahren erlaubt also die Einführung von Mutationen in Plasmid-DNA unabhängig von benachbarten Endonuklease-Schnittstellen.

***QuikChange*-Methode**

Bei diesem Verfahren zur Einführung von Punktmutationen dient das ungeschnittene Plasmid als Templat. Es werden zwei inverse Oligodesoxynukleotide eingesetzt, die mittig jeweils die mutierte DNA-Sequenz besitzen. Bei der folgenden (nicht-exponentiellen) *QuikChange*-PCR erhält man als DNA-Produkt ein zirkuläres Plasmid mit der eingeführten Mutation und mit zwei Einzelstrangbrüchen. Mit *DpnI* werden selektiv alle dam-methylierten 5' GATC 3'-Sequenzen des eingesetzten Templats restringiert und die DNA in *E. coli* elektroporiert.

Für eine *QuikChange*-PCR werden folgende Substanzen eingesetzt:

10 x Pfu-turbo-Puffer	5.0 µl
dNTP (2 mM)	5.0 µl
Templat: Plasmid-DNA (4 kB) (20 ng/µl)	1.0 µl
<i>sense</i> Oligodesoxynukleotid (10 pmol/µl)	1.0 µl
<i>antisense</i> Oligodesoxynukleotid (10 pmol/µl)	1.0 µl
<i>Pfu</i> -Polymerase (2.5 U/µl)	1.0 µl
steriles Wasser	ad 50 µl

PCR-Programm für eine *QuikChange*-PCR:

60 s bei 94°C (*denaturation*)

30 Amplifikationszyklen mit jeweils

60 s bei 94°C (*denaturation*)

60 s bei 55°C (*annealing*)

480 s bei 68°C (*extension*)

Ende: 360 s bei 72°C

3.4. Zufallsmutagenese über Polymerase-Kettenreaktion

3.4.1. Mutagene PCR

(nach Cadwell & Joyce, 1992)

Zur erwünschten Einführung von Replikationsfehlern bei der PCR wird die Lesegenauigkeit der Taq-Polymerase durch Zugabe von Mn^{2+} erniedrigt. Nach der PCR erhält man ein PCR-Produkt mit zufällig verteilten Basenaustauschen, wobei sich die Zahl der Basenaustausche durch die Mn^{2+} -Konzentration vorherbestimmen läßt: sie beträgt bei dem unten angeführten Ansatz theoretisch 0.66 % pro Nukleotidposition (Cadwell & Joyce, 1992). Bei äquimolarem Einsatz von dNTP werden Pyridinbasen statistisch häufiger gegen Purinbasen ausgetauscht als umgekehrt. Zur Vermeidung einer Drift wird deshalb eine höhere Konzentration an Pyridinnukleotidtriphosphaten in der PCR eingesetzt.

mPCR-Ansatz zur ungerichteten Mutagenese der PH-Domäne

10 x <i>m</i> Reaktionspuffer	10.0 µl
dCTP und dTTP	1.0 mM
dGTP und dATP	0.2 mM
Templat:ptoPHG3-Vektor bzw. ptoCP3.5G3-Vektor	20 ng
ompA for-Oligodesoxynukleotid	3 µM
CyG3770as-Oligodesoxynukleotid	3 µM
MnCl ₂	0.5 mM
Taq-Polymerase (1 U / µl)	5.0 µl
steriles Wasser	ad 100 µl

Der 10 x *m*Reaktionspuffer hat die folgende Zusammensetzung:

0.5 M	KCl
0.1 M	Tris-HCl
0.1 %	Gelatine
70 mM	MgCl ₂ mit NaOH auf pH 8.3

PCR-Programm für die *m*PCR:

60 s bei 94°C (*denaturation*)

30 Amplifikationszyklen mit

60 s bei 94°C (*denaturation*)

60 s bei 45°C (*annealing*)

60 s bei 72°C (*extension*)

Ende: 360 s bei 72°C

Die Klonierung der *m*PCR-Produkte erfolgt über die Schnittstellen *NcoI* und *NotI* in den Vektor pTomPHG3.

3.4.2. Randomisierung durch Einsatz randomisierter Oligodesoxynukleotide

Für die zufällige Anordnung von Aminosäuren eines ganzen Peptidsegments wird die dazugehörige DNA-Sequenz randomisiert. Die Randomisierung erfolgt dabei während der Synthese der Oligodesoxynukleotide, bei der anstelle der reinen Desoxynukleotide bei jedem Syntheseschritt Gemische von Desoxynukleotiden eingesetzt werden (s. Kapitel 2.2.2.). Die Konstruktion eines randomisierten Oligodesoxynukleotids ist in Abb. 4.1 dargestellt.

Um die Synthese von Heteroduplex-DNA zu vermeiden, wurde das Verfahren der „asymmetrischen PCR“ entwickelt. Dabei wird in einer ersten PCR durch 50-fachen molaren Überschuss des randomisierten Oligodesoxynukleotids gegenüber dem *antisense*-Oligodesoxynukleotids bevorzugt Einzelstrang-DNA hergestellt, die dann in einer weiteren Reaktion durch weitere Zugabe des *antisense*-Oligodesoxynukleotids zu dsDNA aufgefüllt wird (s. Kapitel 4.1.2.1.).

Ist das 5'-Ende des randomisierten Oligodesoxynukleotids zu kurz gewählt, läßt sich nach einer asymmetrischen PCR das PCR-Produkt nicht restringieren. Anstelle einer (situationsbedingt nicht möglichen) Neubeschaffung von 5'-verlängerten Oligodesoxynukleotiden wurde die „asymmetrische Verlängerungs-PCR“ entwickelt. Dazu wird eine reguläre PCR (50 µl-Ansatz) mit äquimolarem Einsatz des randomisierten *sense*-Oligodesoxynukleotids *Loop1rms sense* bzw. *Loop2rms sense* und dem nicht randomisierten *antisense*-Oligodesoxynukleotid durchgeführt (25 Zyklen). Das PCR-Produkt liegt dann als Hetero-

duplex-DNA mit zu kurzem 5'-Ende vor (stromaufwärts der Klonierungsschnittstelle). Das PCR-Produkt wird mit Hilfe des QiaSpin PCR-Kits gereinigt, in 50 µl kationischem EB-Puffer eluiert. Die Hälfte des Eluats dient für die „asymmetrische Verlängerungs-PCR“ als Templat. Dabei wird über 12 Amplifikationszyklen unter Standard-PCR-Bedingungen mit den entsprechenden Verlängerungs-Oligodesoxynukleotiden *TT3-loop1-ohne (sense)* bzw. *TT3-loop2-ohne (sense)* (s. Kapitel 2.2.2.) und ohne Gegen-Oligodesoxynukleotid einzelsträngige, randomisierte, 5'-verlängerte DNA hergestellt. Sie wird in einer anschließenden Auffüllreaktion mit nicht randomisiertem Oligodesoxynukleotiden komplementär ergänzt (s. asymmetrische PCR).

asymmetrische PCR (für Loop 6):

1. PCR zur Herstellung von einzelsträngiger, randomisierter DNA

10 x Reaktionspuffer	5.0 µl
dNTP (2 mM)	5.0 µl
MgCl ₂ (25 mM)	5.0 µl
Templat: Plasmid-DNA (4 ng/µl)	1.0 µl
<i>Loop3rms sense</i> (10 pmol/µl) (randomisiertes Oligodesoxynukleotid)	5.0 µl
<i>antisense</i> Oligodesoxynukleotid (1 pmol/µl) (nicht randomisiertes Oligodesoxynukleotid)	1.0 µl
<i>Taq</i> -Polymerase (1 U / µl)	2.0 µl
steriles Wasser	ad 50.0 µl

PCR-Programm für die asymmetrische PCR:

60 s bei 94°C (*denaturation*)

30 Amplifikationszyklen mit

30 s bei 94°C (*denaturation*)

60 s bei 61°C (*annealing*)

90 s bei 72°C (*extension*)

2. PCR zum Auffüllen des komplementären Gegenstranges aus der 1. PCR:

erster PCR-Ansatz	50 µl
10 x Reaktionspuffer	5.0 µl
dNTP (2 mM)	5.0 µl
MgCl ₂ (25 mM)	5.0 µl
<i>antisense</i> Oligodesoxynukleotid (100 pmol/µl)	0.5 µl
<i>Taq</i> -Polymerase (1 U / µl)	1.0 µl
steriles Wasser	ad 100 µl

PCR-Programm mit nur **einer** Amplifikation:

30 s bei 94°C	(<i>denaturation</i>)
120 s bei 61°C	(<i>annealing</i>)
500 s bei 72°C	(<i>extension</i>)

Zur Herstellung einer Bibliothek, die in drei Loopstrukturen randomisiert ist, wird über eine Reihe von Dimerisierungs-, Ligations- und Restriktionsschritten ein DNA-Fragment hergestellt, das in allen drei gewünschten DNA-Abschnitten unabhängig voneinander randomisiert vorliegt.

3.4.3. Herstellung einer in drei Segmenten randomisierten Bibliothek von PH-Domänen

(s. Kapitel 4.1.3. Das Klonierungsschema ist in Abb. 4.7 wiedergegeben.)

- Zuerst werden die jeweils in einem Segment randomisierten PCR-Fragmente hergestellt (s. Kapitel 3.4.2.): rms3 wird über asymmetrische PCR mit anschließender Auffüllreaktion, rms1 sowie rms2 werden über PCR mit anschließender asymmetrischer Verlängerung und Auffüllreaktion zu Homoduplex-DNA aufgebaut. (Die randomisierten Oligodesoxynukleotide werden hier nur rms1, rms2 und rms3 genannt).
- Das PCR-Fragment von rms 1 wird mit *AflII*, das von rms2 mit *AccI* und das von rms3 mit *BpiI* restringiert und über ein Agarosegel aufgereinigt.
- Die restringierten PCR-Fragmente von rms1 und rms2 werden durch Ligation dimerisiert (*AflII* und *AccI* besitzen palindromische Schnittstellen); anschließend wird die Ligase hitzeinaktiviert.

- Das rms 1-Dimer wird mit *AccI* und das rms 2-Dimer mit *BpiI* restringiert und die restringierten Dimere werden über ein Agarosegel aufgereinigt.
- Das restringierte rms2-Dimers wird mit *AccI* monomerisiert und *AccI* anschließend hitzeinaktiviert.
- Das restringierte Dimer von rms1, das monomere rms2-Fragment und das restringierte rms3-Fragment werden ligiert und anschließend wird die Ligase hitzeinaktiviert.
- Der gesamte Ligationsansatz wird mit *AflIII* und *XhoI* restringiert und das in drei Segmenten unabhängig randomisierte Fragment von 326 bp über ein Agarosegel aufgereinigt.
- Dieses Fragment wird in den Vektor plomPHG3 kloniert (dies betraf $5 \cdot 10^7$ Klone der in dieser Arbeit verwendeten Bibliothek von randomisierten PH-Domänen).
- Bei 10^7 weiteren Klonen der Bibliothek wurde das unabhängig randomisierte Fragment einer weiteren PCR-Amplifikation unterzogen: dazu wird auf die Restriktion des dreifach randomisierten DNA-Fragments mit *XhoI* verzichtet und das Fragment nach *AflIII*-Restriktion über ein Agarosegel aufgereinigt. Ein *AflIII*-restringiertes (nicht randomisiertes) DNA-Fragment wird vor das randomisierte Fragment ligiert, das Ligationsprodukt mit 20 PCR-Zyklen asymmetrisch amplifiziert und anschließend aufgefüllt (s.o.). Auch dieses PCR-Produkt wird *AflIII* und *XhoI* restringiert, aufgereinigt und in den Vektor plomPHG3 kloniert.

Über den Vektor plomPHG3 wird die PH-Domäne N-terminal mit der ompA^M-Signalpeptid und C-terminal über ein *amber*-Codon mit dem g3-Protein fusioniert exprimiert. Als Operator/Promotor dient das lac^{o/p}-System.

3.5. DNA-Sequenzierung

Für die Sequenzierung von Plasmid-DNA sollte das Plasmid hochrein vorliegen (z.B. nach Aufreinigung mit dem Qiagen[®] Mini Plasmid-Kit: s. Kapitel 3.1.4.). Für die Sequenzierreaktion wird ein Oligodesoxynukleotid eingesetzt, das mit deutlicher Präferenz 30-50 Basenpaare vor der zu untersuchenden DNA-Sequenz hybridisiert.

Die Sequenzreaktionen wurden im Labor von Helmut Blum am Genzentrum und von der Firma Toplab, Martinsried, mit *Cycle-Sequencing*-Kits und fluoreszenzmarkierten Terminatoren nach der Kettenabbruch-Methode durch

Didesoxynukleotide durchgeführt. Die Detektion und Sequenzanalytik erfolgte jeweils mit einem automatischen DNA Sequenzer ABI-373 A.

3.6. Proteinexpressionen und die jeweiligen Zellaufschlüsse

3.6.1. Grundregeln für die Genexpression von Proteinen in *E. coli*

Die Expressionen von eukaryontischen Genen in *E. coli* werden unter Berücksichtigung folgender Aspekte durchgeführt

- *E. coli* ist ein Prokaryont und verfügt nicht über eukaryontische posttranslationale Modifizierungen und Prozessierungen (Glykosylierungen, prosthetische Gruppen *etc.*).
- Der Kontext während und nach der Proteinfaltung unterscheidet sich in allen Organismen. Dies gilt z.B. für Chaperone aber auch für alle anderen Protein-Protein-Wechselwirkungen.
- Das Redoxpotential des Cytoplasmas ist reduktiv, während im Periplasma oxidative Bedingungen herrschen. Strukturelle Disulfidbrücken sind nur im Periplasma stabil. Die Zellmembran trennt das reduktive Cytoplasma vom oxidativen Periplasma, welches mit der Zellwand gegen das Medium abgeschlossen ist.
- Viele Proteine wirken bei Überexpression toxisch, weshalb das gewünschte Protein erst ab dem Zeitpunkt der Induktion und nicht schon zuvor überexprimiert werden soll.
- Wird *E. coli* nur als „Proteinfabrik“ eingesetzt, benötigt man einen starken Promotor wie den T7-Promotor im entsprechenden Bakterienstamm BL 21(DE3). Ist das Protein zudem toxisch, wird der Stamm BL 21(DE3)pLysS verwendet, der T7-Lysozym coexprimiert, das ein natürlicher Inhibitor der T7-RNA-Polymerase ist (Huang *et al.*, 1999). Die Expression nach Induktion in diesem System ist in der Regel so stark, daß das Protein in der Zelle aggregiert und dann als unlösliches Aggregat (*inclusion-bodies*) vorliegt. Zur regulierbaren, gegenüber dem T7-System verringerten Expression wurde in dieser Arbeit das *tet*-System aus pASK75 (Skerra, 1994) eingesetzt. Auch das *lac*^{o/p}-System, das auf den historischen

Arbeiten von Francois Jacob und Jaques Monod beruht, fand Anwendung in dieser Arbeit.

- Der genetische Code ist zwar universell, doch unterscheidet sich die Häufigkeit der einzelnen Codons und tRNAs bei Prokaryonten und Eukaryonten erheblich (Nakamura & Tabata, 1997). Bei der Expression eukaryontischer Gene in *E. coli* kann beispielsweise das seltene Arginin-Codon AGG die Proteinausbeute limitieren.
- Splicing- und Terminationsprozesse müssen an das *E. coli*-System angeglichen werden.

Im Rahmen dieser Arbeit wurden unter verschiedenen Gesichtspunkten Proteine in *E. coli* hergestellt. Jede Fragestellung erfordert ein anderes Versuchsprotokoll und teilweise auch ein anderes Vektorkonstrukt. Drei grundlegend unterschiedliche Methoden werden detailliert aufgeführt:

- (i) Periplasmatische Expression unter standardisierten Bedingungen (s. Kapitel 3.6.2.)
- (ii) Cytoplasmatische Expression unter standardisierten Bedingungen (s. Kapitel 3.6.3.)
- (iii) Cytoplasmatische Expression als *inclusion-bodies* (s. Kapitel 3.6.4.)

Allen Expressionen sind die folgenden Vorgehensweisen gemeinsam:

- Vorkulturen werden entweder mit einem Einzelklon nach Transformation oder aber aus dem Glyzerinstock eines Einzelklons angeimpft. Sie werden in LB-Medium mit dem zur Plasmidselektion notwendigen Antibiotikum angesetzt und über Nacht bei 37°C unter Schütteln inkubiert, so daß sie sich folgendentags in der stationären Wachstumsphase befinden.
- Die Vorkulturen werden mit antibiotikumhaltigem Medium im Verhältnis 1/5 bis 1/100 zu den Hauptkulturen verdünnt und bei 37°C unter Schütteln inkubiert. Nach ca. einer Stunde befinden sich die Hauptkulturen in der exponentiellen Wachstumsphase.
- Gegen Ende dieser Phase erfolgt die Expressionsinduktion des plasmid-kodierten Proteins mit dem promotor- bzw. repressorspezifischen Induktionsmittel. Entsprechend der Fragestellung und der Proteineigenschaften

werden insbesondere die Expressionstemperatur und -dauer, aber auch die Schüttelintensität (Sauerstoffversorgung) gewählt.

- Nach der Expression werden die *E. coli* abzentrifugiert und das Protein mit unterschiedlichen Aufschlußverfahren aus den Zellen gewonnen.

3.6.2. Periplasmatische Proteinexpression unter standardisierten Bedingungen

- Das proteinkodierende (mit ompA^M-Signalsequenz) Plasmid wird in *E. coli* JM83 transformiert und ausplattiert.
- 20 ml einer LB-Amp-Vorkultur werden mit einem Einzelklon angeimpft und über Nacht bei 37°C unter Schütteln inkubiert.
- Die Vorkultur wird mit 250 ml LB-Amp-Medium in einem 500-ml-Erlenmeyerkolben verdünnt und bei 37°C/180 Upm bis zu einer OD₆₀₀ von 0.9 bis 1.0 (ca. 1.5 h) inkubiert.
- Induziert wird mit 25 µl anhydro-Tetrazyklinlösung (2 mg aTC auf 1 ml DMF; Endkonzentration: 0.43 nM) und dann 4 h bei 25°C/180 Upm inkubiert.
- Die Zellen werden durch 15-minütige Zentrifugation in GS3-Zentrifugengefäßen bei 4500 *rcf*/4°C vom Medium getrennt und der Überstand dekantiert.
- Das Bakterienpellet wird vorsichtig in 5 ml eiskaltem Periplasma-Aufschlußpuffer suspendiert, in ein SS34-Zentrifugengefäß überführt und 1.5 h auf Eis gerührt. Das EDTA des Aufschlußpuffers komplexiert dabei die Mg²⁺- und Ca²⁺-Ionen der Zellwand, so daß sie labil und permeabel wird.
- Die entstandenen Sphäroblasten werden scharf abzentrifugiert (40 min, 0°C, 50 000 *rcf*), so daß die Proteine des Periplasmas durch die poröse Zellwand gepreßt werden. Aus dem Überstand können die Proteine des Periplasmas anschließend über Säulenchromatographie gereinigt werden (dazu: 300 µl Ni-NTA-Agarose und 1 ml Elutionspuffer: s. Kapitel 3.7.1.)

Periplasma-Aufschlußpuffer:

1 M	NaCl
50 mM	Tris/HCl, pH 8.0
1 mM	EDTA

3.6.3. Cytoplasmatische Proteinexpression unter standardisierten Bedingungen

- Das proteinkodierende (ohne ompA^M-Signalsequenz) Plasmid wird in *E. coli* BL 21(DE3) oder BL 21 transformiert und ausplattiert.
- 2 ml einer LB-Amp-Vorkultur werden mit einem Einzelklon angeimpft und über Nacht bei 37°C unter Schütteln inkubiert.
- 10 ml LB-Amp-Medium in einem 20-ml-Erlenmeyerkolben werden mit der Vorkultur angeimpft und bei 37°C/200 Upm bis zu einer OD₆₀₀ von 0.9 bis 1.0 (ca. 1.5 h) inkubiert.
- Induziert wird mit 1 µl aTC (2 mg aTC auf 1 ml DMF; Endkonzentration: 0.43 nM) und dann 7 h bei 20°C/200 Upm inkubiert.
- Die Zellen werden durch zehnmünütige Zentrifugation in SS34-Zentrifugengefäßen bei 7000 *rcf*/4°C vom Medium getrennt, der Überstand dekantiert und das Pellet in 500 µl PBS resuspendiert.

denaturierender Gesamtzellaufschluß

- Zu der Bakteriensuspension werden 1.5 ml 6 M GdmCl/PBS (pH 7.4/NaOH) pipettiert, gemischt und der Ansatz auf Eis gekühlt.
- Die lysierten Zellen werden sonifiziert (Ultraschallbehandlung: 2 min / *duty cycle* 30 / *outputcontrol* 3), so daß RNA und DNA mechanisch zerkleinert werden und damit die Suspension dünnflüssig wird.
- Nach zehnmünütiger Zentrifugation bei 14 000 *rcf*/4°C können die Proteine aus dem Überstand durch Säulenchromatographie gereinigt werden (dazu: 300 µl Ni-NTA-Agarose und 1 ml Elutionspuffer: s. Kapitel 3.7.1.).

nativer Zellaufschluß

Um Proteine direkt löslich aus *E. coli* zu gewinnen, muß die Zellwand und die Zellmembran mechanisch (bis zu 50 ml Kulturvolumen mit Hilfe der Kugelmühle, ab 500 ml Kulturvolumen mit der *french press*, sonst: Ultraschall) oder enzymatisch zerstört werden (Zugabe von T7-Lysozym, oder Freisetzung cytoplasmatisch exprimierten T7-Lysozyms in BL 21(DE3)pLysS durch mehrfaches Einfrieren und Auftauen, was zur Schädigung der Zellmembran führt).

Exemplarisch für die Gewinnung kleiner Mengen löslichen Proteins aus *E. coli* ist die Expression und Aufarbeitung unter chemisch nativen Bedingungen:

- Die in 500 µl PBS resuspendierten *E. coli*-Zellen (s.o.) werden in ein 1.5 ml-Eppendorf-Reaktionsgefäß überführt mit einer Spatelspitze Glasperlen versetzt (100 mg mit Ø 0.1-0.2 mm) und 20 min bei maximaler Schwingungsfrequenz mit einer Kugelmühle aufgeschlossen.
- Der Ansatz wird 10 min bei 18400 *rcf*/4°C zentrifugiert, der Überstand dekantiert und das Pellet in weiteren 500 µl PBS resuspendiert.
- Dieser Vorgang wird noch zweimal wiederholt und die Überstände werden vereinigt. Die löslichen Proteine können anschließend chromatographisch aufgetrennt werden.

Um nachfolgend zu analysieren, wieviel unlösliches Protein in den Zellen vorhanden war, bzw. wieviel Protein durch die mechanische Belastung durch die Kugelmühle präzipitiert ist, wird das zu letzt verbliebene Pellet mit 8 M Harnstoff/PBS anstelle von PBS resuspendiert, erneut mit der Kugelmühle behandelt und anschließend zentrifugiert. Auch diese Vorgänge werden noch zweimal wiederholt die Überstände vereinigt und analysiert.

3.6.4. Cytoplasmatische Expression von unlöslichem Protein in *inclusion-bodies*

Um den Einfluß der ompA^M-Signalsequenz auf die Proteinstabilität zu untersuchen, wurde die Mutante CP 2.6. mit ompA^M-Signalsequenz als unprozessiertes Protein hergestellt. Dies geschieht durch cytoplasmatische Expression in *inclusion-bodies* hinter dem T7-Promoter in pRSET-Vektoren (Schöpfer, 1993).

- Das Plasmid (Das Gen für das Präprotein wird unter dem T7-Promotor exprimiert) wird in *E. coli* BL 21(DE3) transformiert und ausplattiert.
- 100 ml einer LB-Amp-Vorkultur werden mit einem Einzelklon angeimpft und über Nacht bei 37°C unter Schütteln inkubiert.
- Die Vorkultur wird mit 1.5 l LB-Amp-Medium in einem 5 l-Erlenmeyerkolben verdünnt und bei 37°C/120 Upm bis zu einer OD₆₀₀ von 0.8 (ca. 1.5 h) inkubiert.
- Induziert wird mit 750 µl IPTG-Lösung (1 M in Wasser) und dann 3.5 h bei 40°C/120 Upm inkubiert.
- Die Zellen werden durch zehnminütige Zentrifugation in GS3-Zentrifugengefäßen bei 4000 *rcf*/4°C vom Medium getrennt, der Überstand dekantiert, das Pellet in 25 ml PBS resuspendiert und auf Eis gekühlt.

- In zwei Zyklen wird die Zellsuspension mit der *french press* bei maximal 1050 bar aufgeschlossen.
- Das Zelllysat wird bei 30°C mit 16 U DNaseI inkubiert und anschließend im SS34-Rotor bei 20 000 *rcf*/4°C zentrifugiert.

Die *inclusion-bodies* befinden sich ungelöst im Pellet, weshalb der Überstand verworfen und das Pellet gereinigt wird (vgl. Freund *et al.*, 1998):

- Das Pellet wird mit 10 ml IB-Puffer I, dann mit 10 ml IB-Puffer II gewaschen und anschließend in Wasser resuspendiert.
- Nach zehnminütiger Zentrifugation bei 20 000 *rcf*/4°C wird der Überstand verworfen. Das Pellet wird in 3 ml 6 M GdmCl/PBS (pH 7.4/NaOH) aufgenommen und erneut 10 min bei 20 000 *rcf*/4°C zentrifugiert. Der Überstand wird dekantiert und das Pellet wiederum in 2 ml 6 M GdmCl-Lösung resuspendiert und zentrifugiert. Die Überstände werden vereinigt; hieraus wird das Protein aufgereinigt und anschließend renaturiert. Für diesen cytoplasmatisch exprimierten und denaturierend vorliegenden Ansatz werden 4 ml Ni-NTA-Agarose benötigt und diese mit 6 ml Elutionspuffer eluiert (s. Kapitel 3.7.1.).

IB-Puffer I:

50 mM	EDTA
10 mM	Tris/HCl pH 8
1 %	Triton-X-100

IB-Puffer II:

500 mM	GdmCl
100 mM	Tris/HCl pH 7.5
20 mM	EDTA

Für diese Arbeit wurden viele weitere Protokolle zur Expression und Aufreinigung verschiedener Proteine ausgearbeitet; die meisten leiten sich von den in den vorigen Kapiteln genannten ab.

3.7. Proteinchemische Aufreinigung und Dialyse

Proteine lassen sich über ihre Eigenschaften, wie Größe, Form, Ladung oder Affinitäten aufreinigen. Zur Trennung rekombinant exprimierter Proteine werden üblicherweise die Proteine mit zusätzlichen Aminosäuren N- oder

C-terminal fusioniert, so daß sich die (chemischen) Eigenschaften des so modifizierten Proteins deutlich von denen der abzutrennenden Proteine unterscheiden. Das in dieser Arbeit wichtigste Verfahren ist die Aufreinigung über einen C-terminalen His₆-tag (sechs Histidinreste werden an den C-Terminus des Proteins fusioniert). Das chromatographische Verfahren beruht auf der Affinität von Histidin zu Nickelionen, die über Nitrilotriazetat (NTA) an einer Sepharosematrix gebunden sind (Hochuli *et al.*, 1987). Die Elution erfolgt durch Konkurrenz mit Imidazol, welches die Histidinreste verdrängt. Meistens kann das gewünschte Protein bis zur Homogenität in einem Schritt aufgereinigt werden.

Die Kapazität der verwendeten Ni-NTA-Agarose beträgt für ein 15 kDa-Protein etwa 5 bis 10 mg Protein/ml (s. Qiagen Handbuch „The QIAexpressionist™“). Sie ist vom verwendeten Auftragspuffer und der Proteinkonzentration abhängig. So verringert sich die Kapazität beim Auftrag von Proteinlösungen mit 4 M GdmCl ca. um den Faktor 10 gegenüber Proteinlösungen in PBS.

3.7.1. Aufreinigung von Proteinen mit His-tag an Ni-NTA-Matrix

(SV: Säulenvolumen an Ni-NTA-Matrix, über das die Aufreinigung von Proteinen mit His-tag erfolgt.)

Von Auftrag, Durchlauf, Waschlösung und Eluat werden Aliquots behalten.

- Die Ni²⁺-NTA-Agarose-Säulen (beladen mit Ni²⁺) werden mit 4 SV Aufschlußpuffer äquilibriert und anschließend wird der partikelfreie Zellaufschluß aufgetragen.
- Nach vollständigem Durchlauf des Auftrags wird zweimal mit je 4 SV Waschpuffer gewaschen.
- Eluiert wird fraktionierend mit 4 SV Elutionspuffer.

Aufschlußpuffer:

50 mM Tris
1 M NaCl
mit HCl auf pH 8.0

Waschpuffer:

50 mM Tris
1 M NaCl
35 mM Imidazol
mit HCl auf pH 8.0

Elutionspuffer:

50 mM Tris
1 M NaCl
300 mM Imidazol
mit HCl auf pH 8.0

Bei der Aufreinigung denaturierter Proteine enthalten Auftragspuffer und Waschpuffer zusätzlich 4 M GdmCl. In diesem Fall wird das Protein auf der Säule nach dem Waschen mit 4 SV PBS zurückgefaltet und dann im nativen Zustand eluiert.

3.7.2. Dialyse von Proteinen

Proteinlösungen werden zweimal gegen das hundertfache Volumen des gewünschten Puffers (meistens PBS) jeweils entweder 3 h bei RT oder bei 4°C über Nacht dialysiert. Hierzu werden Dialyseschläuche verwendet (Spectra Por 1 aus regenerierter Cellulose mit einer Ausschlußgröße von 6-8 kDa), die zuvor zweimal in TE-Puffer aufgekocht wurden und bis zur Verwendung bei 4°C in TE-Puffer steril aufbewahrt werden. Nach der Dialyse werden die Proteine sterilfiltriert und nach Zugabe von 1/100 Vol. 2 % NaN₃-Lösung bei 4°C oder -20°C aufbewahrt.

PBS:

4 mM KH₂PO₄
16 mM Na₂HPO₄
115 mM NaCl
(ohne zusätzliche Einstellung des pH-Wertes)

3.8. SDS-Polyacrylamid-Gelelektrophorese von Proteinen

3.8.1. Gelelektrophoretische Auftrennung von Proteinen

Proteine können mit der SDS-PAGE (SDS-Polyacrylamid-Gelelektrophorese) in Polyacrylamid-Gelen nach Denaturierung mit SDS (Natriumdodecyl-Sulfat) und Reduktion nach ihrem Molekulargewicht und ihrer Aminosäurezusammensetzung aufgetrennt werden. Die PH-Domänen (ca. 15 kDa) lassen sich gut in 15 %-igen Gelen unter den Pufferbedingungen von Fling und Gregerson (Fling & Gregerson, 1986) nachweisen.

Herstellung der Gele (SDS-PAGE-Gele)

Das folgende Rezept genügt zum Gießen von vier Gelen. Für das 15 %-ige Trenngel werden 20 ml Ultrapure Protogel (30 % Acrylamid, 0.8 % Bisacrylamid), 10 ml 4 x *lower-tris*-Puffer und 10 ml Wasser in einem Becherglas gemischt, mit 20 µl TEMED und 260 µl 10 % APS versetzt und die Gelkammern bis ca. 2 cm unter dem Rand gefüllt. Um Eintrocknung während der Polymerisation zu vermeiden, wird mit etwa 1 ml 50 % 2-Propanol überschichtet, das nach erfolgter Polymerisation mit Wasser herausgespült wird. Für das 5 %-ige Sammelgel werden 1.9 ml Ultrapure Protogel, 3 ml 4 x *upper-tris*-Puffer und 7 ml Wasser mit 20 µl TEMED und 40 µl 10 % APS-Lösung angesetzt und auf das Trenngel gegossen. Nach dem Einsetzen des Taschenformers und erfolgter Polymerisation können die Gele bei 4°C in feuchten Zellulosetüchern eingewickelt ca. 4 Wochen gelagert werden.

4 x *lower-tris*-Puffer:

3 M Tris
0.4 % SDS
mit HCl auf pH 8.85

4 x *upper-tris*-Puffer:

0.5 M Tris
0.4 % SDS
mit HCl auf pH 6.8

SDS-PAGE

Zur Elektrophorese werden die Gele in die Vertikal-Gelelektrophoresekammer eingespannt und die Pufferkammern mit Laufpuffer gefüllt. Die Proteinproben werden mit 1/4 Vol. an 5 x SDS-Auftragspuffer versetzt und zur Denaturierung 5 min auf 95°C erhitzt. Vor dem Auftragen der Proben werden die Taschen mit Laufpuffer gespült. Die Elektrophorese erfolgt bei 100-150 Volt und 35 mA (bei 1.5 mm dicken Gelen). Zur Visualisierung der Proteinbanden werden die Gele mit Silber oder Coomassie gefärbt und alternativ einer Immunfärbung unterzogen (s. folgende Kapitel)

5 x SDS-Auftragspuffer:

50 % Glyzerin
7.5 % SDS
250 mM Tris
0.5 % Bromphenolblau
12.5 % 2-Mercaptoethanol
mit HCl auf pH 8.0

Fling & Gregerson Laufpuffer:

50 mM Tris
190 mM Glycin
0.1 % SDS
(ohne zusätzliche Einstellung des pH-Wertes)

3.8.2. Silberfärbung von Proteingelen

Die Silberfärbung von SDS-PAGE-Gelen ist sehr empfindlich, im Idealfall können Einzelbanden von 1 ng Polypeptid nachgewiesen werden. Das Gel wird mit den unten angegebenen Lösungen behandelt (Blum *et al.*, 1987). Nach erfolgter Färbung kann es zwischen zwei feuchten Polyazetat-Folien (*gel drying film*) getrocknet werden. Die Farbintensität der Proteinbanden nach Silberfärbung sind nicht proportional zur eingesetzten Proteinmenge.

SDS-PAGEs werden nach elektrophoretischer Auftrennung von Proteinen wie folgt unter leichtem Schwenken entwickelt:

- 5' Fixierlösung

50 %	Methanol
12 %	Essigsäure
0.02 %	Formaldehyd
- 10' Waschen

50 %	Ethanol
------	---------
- 10' Waschen

30 %	Ethanol
------	---------
- 1' *sensitizer*

0.2 mg/ml	Na ₂ S ₂ O ₃
-----------	---
- 3 x 30'' Waschen

	Wasser
--	--------
- 20' Farbe-Lösung

1 mg/ml	AgNO ₃
0.03 %	Formaldehyd
- 2 x 30'' Waschen

	Wasser
--	--------
- 5-20' Entwickler

40 pg/ml	Na ₂ S ₂ O ₃
6 %	Na ₂ CO ₃
0.02 %	Formaldehyd
- 30'-1 h Stop-Lösung

5 %	Essigsäure
-----	------------
- 1 h Trocken-Lösung

50 %	Ethanol
5 %	Methanol
5 %	Glyzerin

3.8.3. Coomassie-Färbung

Die Coomassie-Färbung ist weniger sensitiv; dafür ist die Färbungsintensität der eingesetzten Menge an Protein proportional. Für die Färbung wird das Gel zunächst in 0.125 % Coomassie Brilliant Blue R-250, 50 % Ethanol und 7 % Essigsäure überführt und 30 min unter Schütteln inkubiert. Zur Entfärbung wird es zunächst 10 min in 10 % Ethanol / 7 % Essigsäure und danach in 7 % Essigsäure geschüttelt.

3.8.4. Western-Blots

Die durch SDS-PAGE aufgetrennten Proteine können im elektrischen Feld auf eine Nitrozellulose-Membran transferiert werden. Auf dieser können mit (primären) Antikörpern über eine Immunfärbung Epitope von Proteinen bzw. Proteine selber spezifisch nachgewiesen werden. Entsprechende epitop-

spezifische Nachweise wurden mit der PH-Domäne des Cytohesin-1 (Ratten IgM-Antikörper 6D5-1 von Elisabeth Kremmer, GSF, München) mit dem (His)₆-tag (*scF_V*-AP-Konjugat aus *E. coli*, Lindner *et al.*, 1997) und mit dem g3-Protein des M13-Bakteriophagen (muriner *anti-g3p*, MoBiTec GmbH, Göttingen) durchgeführt.

Zum Transfer der Proteine aus einem SDS-PAGE-Gel wird das Gel luftblasenfrei auf eine kurz in Wasser inkubierte Nitrozellulose-Folie gelegt, auf beiden Seiten mit drei Lagen feuchtem *Whatman*-Filterpapier belegt und in einen mit Transferpuffer gefüllten *Blotting*-Tank überführt.

Transferpuffer:

150 mM Glycin
10 % Methanol
(ohne zusätzliche Einstellung des pH-Wertes)

Nach dem Transfer (14 h bei 25 V oder 7 h bei 50 V mit 200 mA und 1.5 W) wird die Nitrozellulosemembran je nach Nachweisreaktion wie folgt behandelt:

Nachweis von g3-Protein

- Um unspezifische Proteinbindungsstellen zu blockieren, wird die Membran mindestens 30 min mit RotiBlock-Lösung in TBS inkubiert.
- Sie wird zweimal zehnminütig mit TBS / 0.05 % Tween gewaschen und dann 1.5 h mit 10 ml TBS/0.05 % Tween/0.5 % Milchpulver und 2 µl murinem *anti-g3*-Protein Antikörper (als primärem Antikörper) inkubiert.
- Sie wird viermal zehnminütig mit 200 ml TBS/0.1 % Tween gewaschen und dann 2 h mit 10 ml TBS/0.05 % Tween/0.5 % Milchpulver und 2 µl peroxidasekonjugiertem Kaninchen *anti-Maus* Antikörper (als sekundärem Antikörper) inkubiert.
- Viermal zehnminütig wird die Membran mit 200 ml TBS/0.1 % Tween gewaschen und abschließend mit dem Peroxidase-Substrat SuperSignal[®] West Pico und einem Film zur Detektion der Chemilumineszenz ausgewertet.

TBS:

150 mM NaCl
10 mM Tris
mit HCl auf pH 7.4

Nachweis des C-terminalen (His)₆-tag

Bei diesem Western-Blot werden die Proteine auf eine zuvor mit Methanol behandelte PVDF-Membran anstelle auf eine Nitrozellulose-Membran transferiert.

- Um unspezifische Proteinbindungsstellen zu blockieren, wird die Membran mindestens 30 min mit RotiBlock-Lösung in hTBST inkubiert.
- Sie wird eine Minute mit hTBST gewaschen und dann 30 min mit 10 ml hTBST/500 µl RotiBlock-Lösung und 200 µl *anti-His-tag scFv*-Lysat inkubiert (Dieses Lysat wird rekombinant aus *E. coli* gewonnen: Lindner *et al.*, 1997).
- Dreimal wird die Membran zehnminütig mit hTBST gewaschen und anschließend kurz in AP-Puffer inkubiert.
- Der Blot wird mit 5 ml AP-Puffer/16 µl BCIP-Stammlösung/33 µl NBT-Stammlösung entwickelt, bis die Proteine mit (His)₆-tag deutlich zu sehen sind. Die Färbung wird durch Zugabe von Wasser beendet.

hTBST-Puffer:

500 mM NaCl
50 mM Tris pH 7.0
0.2 % Tween 20

AP-Puffer:

100 mM Tris pH 9.5
100 mM NaCl
5 mM MgCl₂

BCIP-Stammlösung:

5 % in DMF

NBT-Stammlösung:

5 % in 70 % DMF

3.9. Das *phage-display*-Verfahren

Als Grundlage zur Durchführung des *phage-displays* dienen das Laborhandbuch „*Phage Display of Peptides and Proteins*“ (Kay *et al.*, 1996) und persönliche Mitteilungen von Elke Prohaska (Universität Regensburg, Institut für medizinische Mikrobiologie und Hygiene). Das Prinzip ist in Kapitel 1.5.1. erläutert.

3.9.1. Bestimmung des Phagentiters

E. coli Stämme, die mit F' gekennzeichnet sind, besitzen das F-Gen zur Ausbildung von F-Pili (s. Kapitel 2.1.1.). Über diese erfolgt die Infektion durch M13-Bakteriophagen bzw. durch ähnliche Phagenkonstrukte. *E. coli* mit F-Pilus lassen sich in der exponentiellen Wachstumsphase hoch effizient infizieren. Die Anzahl von Phagen in einer Lösung wird indirekt über die Anzahl der Bakterien ermittelt, die sich durch die Phagen nach Verdünnung infizieren lassen und die damit die durch die Phagen vermittelte Antibiotikumresistenz erlangen (Anzahl der Phagen in *pfu: plaque forming units*).

Zur Herstellung phagen-kompetenter *E. coli* werden TG 1-Zellen (alternativ XL 1-blue oder JM 109) aus einem Glyzerinstock bis zu einer OD₆₀₀ von ca. 0.5 in 10 ml 2 x YT-Medium ohne Antibiotikum inkubiert; dies entspricht ungefähr einer Zellzahl von $5 \cdot 10^9$ Zellen pro 10 ml Medium). Je nach Aufgabenstellung wird die Lösung, deren Phagentiter bestimmt werden soll, seriell verdünnt, oder aber der geeignete Maßstab an *E. coli* Kultur gewählt (10 ml), um ganze Banken in neue *E. coli* zu übertragen. Für die Infektion wird die Phagenlösung zur Bakterienkultur gegeben und 1 h bei 37°C ohne Schütteln inkubiert. Die infizierten Bakterien werden dann auf LB-Platten mit dem phagen-spezifischen Antibiotikum (hier: Ampizillin) ausplattiert und über Nacht bei 37°C inkubiert. Nach ca. 12 h läßt sich anhand der Zahl der Klone und der jeweiligen Verdünnung der Phagentiter in der Stammlösung zurückrechnen.

3.9.2. Herstellung und Reinigung neuer Bakteriophagen

Zum Herstellen neuer Phagen werden die LB-Platten mit einer Lösung aus 30 % Glycerin/ LB-Medium quantitativ abgespült und als Glycerinstock bei -80°C gelagert. Eine der Komplexität der Bibliothek entsprechende Menge des Glycerinstocks (ca. $100\ \mu\text{l}$) wird zum Animpfen einer Vorkultur von 20 ml 2 x YT-Medium/Amp/0.5 % Glukose in einem 50-ml-Falkongefäß verwendet (ca. 10 % der Gesamtzellzahl ist die Lebendzellzahl in eine Woche alten Glycerinstocks). Bei einer OD_{600} von ca. 0.3 werden 10^{11} pfu Helferphagen M13K07 hinzupipettiert und die Kultur 45 min bei 30°C ohne Schütteln, dann 45 min mit Schütteln inkubiert. Zum Entfernen der Glukose und der nicht infektiösen Helferphagen wird die Kultur bei 2600 *rcf* für 10 min zentrifugiert und der Überstand verworfen. Das Pellet wird in frischem 2 xYT-Medium/Ampizillin/Kanamycin resuspendiert und bis zu einer OD_{600} von ca. 0.9 bei 37°C unter Schütteln inkubiert. Im Fall der PH-Domänen-Bibliotheek im lac^{op} -System wird mit 0.5 mM IPTG bei 30°C über 6 bis 10 h induziert.

Zur Aufreinigung der Phagen wird die *E. coli*-Kultur mit 2 600 *rcf* für 10 min und der anfallende Überstand erneut 10 min bei 16 000 *rcf* zentrifugiert. Zum Überstand wird 1/5 des Volumens PEG/NaCl-Lösung gegeben, die Lösung geschüttelt und 1 h auf Eis gelegt. Die präzipitierten Phagen werden 30 min bei 16 000 *rcf* und 4°C abzentrifugiert und das Pellet in 40 ml Wasser resuspendiert. Zur Abtrennung unlöslicher Substanzen werden die Phagen 10 min bei 5 000 *rcf* zentrifugiert und der Überstand mit den Phagen erneut einer Fällung mit 8 ml PEG/NaCl-Lösung unterzogen. Nach einstündiger Fällung auf Eis wird 30 min bei 5 000 *rcf*/ 4°C zentrifugiert, der Überstand verworfen und das Pellet in 1 ml Wasser oder PBS gelöst.

Von dieser Phagenlösung wird der Titer (s. Kapitel 3.9.1.) bestimmt.

(Medien und Lösungen: s. Kapitel 3.3.1.)

3.9.3. Immobilisierung von Zielsubstanzen

Als Zielsubstanzen wurden in dieser Arbeit die ϵ -Domäne von CD3, das 25 kDa-Protein MOMP (*major outer membrane protein* von *Legionella pneumophila*) und das an BSA gekoppelte Hapten Kreatinin verwendet (s. Kapitel 4.2.1.).

Die Proteine werden in Immobilisierungslösung verdünnt und an 4-ml NUNC-Maxisorb-Röhrchen gemäß den Anweisungen des Herstellers immobilisiert:

Immobilisierungslösung:

300 µg Protein (MOMP, CD3ε oder BSA-Kreatinin gelöst in PBS)
400 µl 5 x Carbonat-Puffer (0.5 M NaHCO₃ pH 9.5)
mit Wasser ad 2 ml.

Die Proteinlösung wird jeweils in ein 4-ml NUNC-Maxisorb-Röhrchen gegeben, mit einem Deckel verschlossen und mit Hilfe eines Überkopfrtors bei 4°C 2.5 h inkubiert. Zur Blockierung der restlichen Bindungsfläche wird die Proteinlösung verworfen, das Röhrchen 3 x 10 min bei Raumtemperatur mit PBS, dann 2.5 h bei 4°C mit RotiBlock/PBS-Lösung über Kopf geschüttelt und abschließend 3 x 5 min mit PBS-Lösung gewaschen. Die Lösungen werden nach jeder Inkubation verworfen. Die Proteine sind nun über hydrophobe Wechselwirkungen an der Oberfläche der NUNC-Maxisorb-Röhrchen immobilisiert („coated“) und die restlichen Bindungsstellen mit RotiBlock abgesättigt. Die beschichteten NUNC-Maxisorb-Röhrchen stehen direkt für das *phage-display* zur Verfügung, lassen sich aber auch über einige Wochen verschlossen bei -20°C lagern.

3.9.4. Panning von Phagen-Bibliotheken

Ca. 10^{13} Phagen in 2 ml PBS/RotiBlock-Lösung werden in die beschichteten NUNC-Maxisorb-Röhrchen gegeben, die Röhrchen verschlossen und 2 h bei 4°C im Überkopfrtor inkubiert. Die Phagenlösung wird nach der Inkubation verworfen und die gebundenen Phagen mit unterschiedlicher Stringenz von der immobilisierten Substanz gewaschen. Es sollte immer so stringent gewaschen werden, daß 10^3 bis 10^5 Phagen eluiert werden können. Nach der ersten *panning*-Runde befinden sich in der Phagen-Lösung nur sehr wenige Phagen, die über die Mutante der PH-Domäne Affinität zur immobilisierten Zielsubstanz aufweisen, während mit steigender Zahl von *panning*-Runden Phagen mit größerer Affinität zur immobilisierten Zielsubstanz angereichert werden und sich somit die Stringenz erhöht. Bei der ersten *panning*-Runde wird sechsmal mit 4 ml PBS und sechsmal mit 4 ml PBS/0.1 % Tween-20 gewaschen. Dabei werden nach der Phagen-Inkubation die Waschlösungen bei Raumtemperatur in die Röhrchen gegeben, diese verschlossen, einmal invertiert und die Waschlösung verworfen. Die Flüssigkeitsreste werden durch Ausklopfen auf Zellulosetücher weitgehend entfernt.

Elution mit Triethylamin

Zur Elution wird nach dem letzten Waschen 1 ml 100 mM Triethylamin/Wasser in die Röhren pipettiert und 30 min bei Raumtemperatur im Überkopffrotor eluiert. Durch den pH-Wechsel auf ca. pH 10 und durch das sterisch große Triethylammonium-Ion als Gegenion werden die Bindungen der Phagenproteine an die immobilisierten Substanzen kompetitiert bzw. aufgelöst. Die Elutionslösung wird quantitativ in ein Eppendorfreaktionsgefäß überführt und dort mit 500 µl 1 M Tris pH 7.4 Lösung neutralisiert.

Elution durch Kompetitierung

Dazu wird nach dem letzten Waschen die freie Zielsubstanz, mit der die Röhren beschichtet worden waren, in PBS gelöst zum Ansatz gegeben und 1 h bei Raumtemperatur im Überkopffrotor inkubiert. An der Zielsubstanz bindende Phagen, die sich kompetitieren lassen, werden so von den immobilisierten Zielsubstanzen gelöst und zur weiteren Infektion von *E. coli*-Zellen eingesetzt (s. Kapitel 3.9.2.). Auf diese Weise selektiert man auf eine hohe k_{off} -Geschwindigkeitskonstante.

Auf der Suche nach bindenden Mutanten der PH-Domänen gegen das Hapten Kreatinin wurde Kreatinin chemisch über einen Linker an BSA-gekoppelt. Dieses BSA-gekoppelte Kreatinin wurde als Substanz an die NUNC-Maxisorb-Röhren immobilisiert. Um auszuschließen, daß man beim *phage-display* Mutanten der PH-Domäne gegen BSA oder den Linker erhält, präeluiert man zuerst mit BSA-Linker ohne Kreatinin durch Competition und dann mit Kreatinin oder mit Triethylamin.

Elution durch Infektion

Die über die Zielsubstanz und die daran bindende PH-Domäne immobilisierten Phagen lassen sich auch „biologisch“ eluieren, indem 1 ml einer infektiös-kompetenten *E. coli*-Kultur nach dem letzten Waschen in die Röhren gegeben wird und diese 30 min bei Raumtemperatur im Überkopffrotor infiziert wird (nach persönlicher Mitteilung von Elke Prohaska).

Zur Überprüfung der Waschschritte und der Elution werden die Phagentiter der letzten Waschlösung und der neutralisierten Elutionslösung bestimmt. Die *panning*-Runde muß so häufig wiederholt werden, bis Waschbedingungen gefunden werden, die bei der Elution die gewünschte Zahl von Phagen freisetzen. Ein erfolgreiches Waschen läßt sich also nur *ex eventu* nach Bestimmung der Phagentiter nachweisen (s. Kapitel 3.9.1.).

3.10. Spektroskopische Methoden

3.10.1. Absorptionsspektroskopie

Für alle absorptionsspektroskopischen Untersuchungen wird ein Uvikon 941 UV/Vis-Spektrophotometer verwendet.

Konzentrationsbestimmung von DNA

Die Konzentrationsbestimmung von DNA-Proben erfolgt nach Verdünnung durch Bestimmung der OD_{260} in Quarzglas-Mikroküvetten (100 μ l Füllvolumen). Für Doppelstrang-DNA entspricht eine Absorptionseinheit 50 μ g/ml, für Einzelstrang-DNA 40 μ g/ml.

Absorptionsspektren von Proteinen

Absorptionsspektren von Proteinlösungen werden durch Messung der Extinktion im Wellenlängenbereich von 250 bis 350 nm in 1 ml Quarzglas-Küvetten bestimmt. Als Referenz dient eine identische Lösung ohne Protein (Dialysepuffer).

Konzentrationsbestimmung von Proteinen

Die Konzentrationsbestimmung erfolgt (für reine Proteine) durch Messung der Absorption bei 280 nm in 1 ml Quarzglas-Küvetten nach Denaturierung in konzentrierter GdmCl-Lösung (5 bis 6 M) gegen eine Referenz mit dem gleichen Lösungsmittel ohne Protein. In vollständig denaturierten Proteinen besitzen alle Chromophore die gleiche chemische Umgebung. Den molaren Absorptionskoeffizienten von Proteinen bei 280 nm (ϵ_{280}) ergibt folgende Gleichung (Gill & Hippel, 1989):

$$\epsilon_{280} / [l \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}] = 5690 \cdot n_W + 1280 \cdot n_Y + 120 \cdot n_{\text{SSBr}}$$

mit n_W Anzahl der Tryptophane,
 n_Y Anzahl der Tyrosine und
 n_{SSBr} Anzahl der Disulfidbrücken im Molekül.

(ϵ_{280} wird also fast nur durch die Anzahl der Tryptophanreste im Proteinmolekül bestimmt.)

Eluate, die durch Ni-NTA-Aufreinigung der löslichen Proteinfraction des cytoplasmatischen Aufschlusses erhalten wurden, sind häufig durch ein Protein mit dem apparenten Molekulargewicht von 26 kDa kontaminiert. Hierbei handelt es sich um WHP, ein Tryptophan-loses aber Histidin-reiches *E. coli*-Protein, das als IMAC-Kontamination bekannt ist (Wulfing *et al.*, 1994). Der molare Extinktionskoeffizient des denaturierten WHP beträgt $\epsilon = 5480 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$, während die PH-Domänen einen molaren Extinktionskoeffizienten von $\epsilon = 29160 \text{ l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ aufweisen. Das Molekulargewicht der Wildtyp PH-Domäne mit His-tag beträgt $15\,215 \text{ g mol}^{-1}$, das des WHP $20\,852 \text{ g mol}^{-1}$. Findet man in einem Coomassie-gefärbten SDS-Gel nach der Auftragung der gereinigten, cytoplasmatisch exprimierten Wildtyp PH-Domäne (bei 15 kDa) eine Bande bei 26 kDa von WHP mit 1/10 der Intensität (verglichen mit der der PH-Domäne), so bedeutet dies, daß in der Probe zwar die PH-Domäne mit 10 % WHP verunreinigt ist, doch wird die spektroskopisch bestimmte Konzentration der PH-Domäne dadurch kaum beeinflusst, da das WHP nur mit unter 2 % zur Absorption der Probe beiträgt.

3.10.2. Fluoreszenzspektroskopie von Proteinen

Für die Aufnahme von Fluoreszenz-Spektren werden alle Puffer partikelfrei filtriert (\varnothing : 2 μm) und entgast. Als Geräteeinstellungen werden, wenn nicht anders angegeben, die Exzitationswellenlänge von 282 nm, die Exzitations-Schlitzbreite von 2.5 nm, ein Emissionswellenlängenbereich von 290-450 nm und eine Emissions-Schlitzbreite von 5 nm gewählt.

3.10.3. Denaturierungs-Übergangskurven

Zur Messung von Übergangskurven wird eine konzentrierte Proteinlösung zu gleichen Endkonzentrationen (ca. 2 μM) in PBS-Puffer und PBS-Puffer mit Denaturierungsmittel (6 M GdmCl) verdünnt oder dialysiert.

Zum Herstellen dieser Lösungen wird die von der Ni-NTA-Agarose eluierte Proteinlösung in zwei gleiche Volumina geteilt, in je ein 50-ml Falkongefäß überführt, je eine Dialysemembran über die Öffnung gespannt und mit je einem durchbohrten Falkondeckel verschlossen. Die Falkongefäße werden deckelabwärts in die Dialyselösung (PBS bzw. PBS/6 M GdmCl pH 7.4) gehängt und über mindestens 6 h äquilibriert. Die Dialyselösungen werden ausgetauscht und die Proteinlösungen gegen frische Lösungen erneut äquilibriert. Auf diese

Weise läßt sich das präzipitationsanfällige Aufkonzentrieren der Protein-Stammlösung zur späteren Verdünnung in PBS/GdmCl-Puffer vermeiden.

Aus diesen beiden Stammlösungen werden durch Mischen Lösungen für die Übergangskurve mit steigender Konzentration des Denaturierungsmittels, [D] hier [GdmCl], hergestellt (Volumina je 1 ml). Zur Einstellung des Faltungsgleichgewichts wird über Nacht bei 20°C inkubiert.

Die Proteinlösungen werden ohne Bläschenbildung und Erwärmung in die Rührküvette pipettiert und dort 20 sec äquilibriert. Die Küvette wird innerhalb einer Meßreihe im Strahlengang des Gerätes belassen und nicht gespült. Es wird mit der Probe der niedrigsten [D] begonnen. Die Wellenlänge für die Exzitation beträgt 282 nm, die Emission wird bei 327 nm gemessen. Das Signal wird über 20 s gemittelt. Vermessene Proben werden möglichst vollständig abpipettiert und danach die Lösung mit der jeweils nächsthöheren [D] eingefüllt. Nach jeder Messung wird der Brechungsindex n_{D20} der Meßlösung mit einem Refraktormeter ermittelt. Aus diesem Wert wird die genaue Konzentration [D] (hier: [GdmCl]) nach folgender Gleichung berechnet (Pace, 1986):

$$[\text{GdmCl}] = 57.147 \Delta n_{D20} + 38.68 (\Delta n_{D20})^2 - 91.60 (\Delta n_{D20})^3$$

3.10.4. CD-Spektroskopie

CD-Spektren werden bei 20°C mit einer Lesegeschwindigkeit von 50 nm/min und einer Dämpfung von 1 sec 20-fach akkumuliert. Die Proteinkonzentration wird so eingestellt, daß die Absorption im Meßbereich nicht über 1 steigt. Sie kann außerdem aus der A_{280} , die man durch die HT → OD-Konversion am Spektrometer erhält, nach der Messung bestimmt werden. Alle Proben werden vor der Messung partikelfrei filtriert (ϕ : 0.2 μm).

Im fernen UV-Bereich (200-260 nm) wird wegen der starken Absorption von Chlorid-Ionen ein salzarmer Puffer (4 mM KH_2PO_4 , 16 mM Na_2HPO_4 , 10 mM NaCl, pH 7.4) bei einer Schichtdicke von 0.2 cm verwendet. Für den nahen UV-Bereich können 0.5- oder 1 cm-Küvetten und PBS als Puffer verwendet werden. Das gemessene Signal wird wie folgt in $[Q]_{\text{MRW}}$ (*mean residual weight*, das Signal pro Peptidbindung, Einheit $\frac{\text{deg} \cdot \text{cm}^2}{\text{dmol}}$) umgerechnet:

$$[\Theta]_{\text{MRW}} = \frac{x \cdot 10^{23} \text{ deg}}{y \frac{\text{mol}}{1000 \text{ cm}^3} \cdot n_{\text{AS}} \cdot z \text{ cm} \cdot 10} \quad ,$$

wobei x das Signal in mdeg, y die Proteinkonzentration in M, n_{AS} die Anzahl der Aminosäuren in der Polypeptidkette und z die Schichtdicke in cm bezeichnet.

3.10.5. Massenspektrometrie

Um periplasmatische Prozessierungsprodukte von Mutanten der PH-Domäne zu analysieren, wurden massenspektrometrische Untersuchungen vorgenommen. Zur Probenvorbereitung werden die Proteine mit einer ungefähren Konzentration von 30 μM gegen 1 %-ige Essigsäure in Wasser dialysiert und anschließend mit Methanol (1:1) gemischt. 5 μl werden in das Quadrupol-Massenspektrometer Finnigan TSQ 7000 zur Messung injiziert. Die Datenaufnahme erfolgt mit den Standard-Geräteeinstellungen. Die Auswertung und die Berechnung der molaren Massen werden durch zugehörige Programmsysteme unterstützt.

3.11. Computer-Analysen

Kristallstrukturen von Proteinen werden mit dem Programmpaket *InsightII* 2.3. (Biosym Technologies Inc.), Molscript (Avatar Software AB, Stockholm) oder RasMac visualisiert. DNA- und Proteinsequenz-Analysen, wie die Erstellung der Restriktionsanalysen, Planung von Oligodesoxynukleotiden für ortsspezifische Mutagenesen und Sequenzvergleiche, werden mit Hilfe des *sequence analysis package* der Genetics Computer Group, University of Wisconsin (Devereux *et al.*, 1984) durchgeführt. Denaturierungskurven und Spektren werden mit dem *Kaleidagraph*-Programmpaket (Abelbeck Software) ausgewertet.

4. Ergebnisse

PH-Domänen sind *in vivo* im Cytosol eukaryontischer Zellen lokalisiert und verhelfen als Module dort Proteinen zu einer Membranassoziation, indem sie über bis zu drei Loops bestimmte Lipide binden. Fast alle diese Eigenschaften werden in dieser Arbeit verändert: in Kapitel 4.1 wird deutlich, daß man die PH-Domäne des Cytohesin-1 als selbständiges Protein rekombinant in *E. coli* exprimieren und die Aminosäuresequenz dreier Loops verändern kann, ohne daß sie ihre Exprimierbarkeit (und Struktur) einbüßt. In Kapitel 4.2. wird nach Randomisierung dieser Loops versucht, ihr über das *phage-display*-Verfahren andere Bindungspartner als Lipide zu vermitteln. In Kapitel 4.3. wird ihr beigebracht, in ungewohnte Kompartimente vorzudringen; sie wird nämlich in das Periplasma von *E. coli* transloziert. Einzig das letzte Kapitel dieses Ergebnisteils (Kapitel 4.4.) beschäftigt sich mit einer „naturrelevanten Aufgabenstellung“: die PH-Domäne wird in Mitochondrien von *Saccharomyces cerevisiae* importiert. Diesem Transportprozeß mußten sich im Laufe der Evolution schon viele mitochondriale Proteine stellen, als ihr mitochondrial kodiertes Gen in den Nukleus transferiert wurde. Seitdem werden sie im Cytosol exprimiert und (post-translational) in das Mitochondrium importiert.

4.1. Die PH-Domäne als Gerüst für randomisierte Loops

4.1.1. Identifizierung der Loops

Die Strukturfamilie der PH-Domänen zeichnet sich durch eine definierte Abfolge von β -Faltblättern und durch eine C-terminale α -Helix aus (s. Kapitel 1.3.), die sich mit hoher Übereinstimmung überlagern lassen (Rebecchi & Scarlata, 1998). Zwischen diesen Elementen der Sekundärstruktur befinden sich Loops. Diese zeigen bei verschiedenen Strukturen unterschiedlicher PH-Domänen eine hohe Variabilität in der Länge wie auch in der Aminosäuresequenz. Nach der Überlagerung der bekannten Strukturen von PH-Domänen mit dem Computerprogramm LSQMAN [Uppsala Software Factory] (Kleywegt & Jones, 1998) wurden die längen- und sequenzvariablen Loops 1, 3 und 6 charakterisiert und durch Sequenzvergleich bei der PH-Domäne des Cytohesin-1 identifiziert (s. Tab. 1.1). Eine genaue Struktur-

analyse bzw. -vorhersage ist erst seit der Veröffentlichung der PH-Domänenstruktur von Grp1 möglich (Lietzke *et al.*, 2000), die zu der des Cytohesin-1 zu über 92 % identisch ist (Peptidsequenz).

4.1.2. Randomisierung mit „Mutator“

Die Loops 1, 3 und 6 wurden in ihrer Aminosäuresequenz randomisiert, indem die entsprechende DNA-Sequenz zufällig mutiert wurde. Dies erfolgte mit entsprechenden Oligodesoxynukleotiden, die nach dem Phosphoramididverfahren synthetisiert wurden (Sinha *et al.*, 1984). Bei den randomisierten Basenpositionen wurde anstelle des reinen Phosphoramidid-Reagenz einer Base eine Lösung mit allen vier Phosphoramidid-Reagentien jeweils unterschiedlicher Konzentration eingesetzt. Dadurch erhält man Oligodesoxynukleotide, die je nach Konzentrationswahl an jeder Position eine vorgegebene Verteilung der vier Basen aufweisen.

Üblicherweise werden zur Randomisierung einer Basenposition die vier Basen (oder Basenkombinationen) in Form ihrer Phosphoramidid-Reagentien äquimolar eingesetzt. Mit diesem experimentellen Ansatz gelangt man zu einer näherungsweise Gleichverteilung aller Aminosäuren an jeder Position und kann die Zahl der *stop*-Codons minimieren. Wird die Bibliothek eines Proteins auf diese Weise hergestellt und mit *E. coli*-Zellen exprimiert, können bis zu fünf Aminosäuren unabhängig voneinander randomisiert werden mit der (statistischen) Garantie, daß jede Kombination der Aminosäuren ($20^5 = 3.2 \cdot 10^6$) einmal in der Bibliothek vorhanden ist (s. Anhang).

Die Bibliothek von PH-Domänen soll an 29 Aminosäurepositionen randomisiert werden. Damit gibt es $20^{29} = 5.4 \cdot 10^{37}$ mögliche Kombinationen und genauso viele Mutanten. Da für das *phage-display*-Verfahren nicht deutlich mehr als 10^8 Mutanten im Labormaßstab erzeugt werden können, wäre rein rechnerisch maximal nur jede $5.4 \cdot 10^{29}$ -te mögliche Mutante in der Bibliothek enthalten. Ferner ist fraglich, ein wie großer Anteil der vollständig randomisierten 10^8 Mutanten sich überhaupt exprimieren läßt: vor dieser Arbeit wurden noch keine PH-Domänen in drei Loops randomisiert und die Exprimierbarkeit untersucht.

Deshalb wurde auf eine (sonst übliche) Randomisierung durch äquimolaren Einsatz von Basen verzichtet und das von Boris Steipe und Heike Bruhn entwickelte Computerprogramm „Mutator“ zur Berechnung der Basenverhältnisse in randomisierten Oligodesoxynukleotiden eingesetzt und in randomisierten Oligodesoxynukleotide verwirklicht (Steipe & Bruhn, in Vorbereitung). Grundlage für „Mutator“ ist die Beobachtung in der Natur, daß physikoche-

misch ähnliche Aminosäuren auch ähnliche Codons besitzen. Wird zur Randomisierung eines Codons in hoher Konzentration jeweils die originale Base und werden in geringer Konzentration die drei anderen Basen eingesetzt, wird nach Randomisierung ein Codon der Wildtyp- oder einer physikochemisch ähnlichen Aminosäure bevorzugt, obwohl dabei keine Aminosäure an einer Position ausgeschlossen wird. Weitgehend ausschließen lassen sich dagegen die *stop*-Codons *ochre* und *opal*, indem „Mutator“ die Adeninkonzentration an jeder dritten Codon-Basenposition sehr klein wählt. „Mutator“ berücksichtigt dabei stets die unterschiedlichen Reaktionsgeschwindigkeiten bei der Kopplungsreaktion aller Basen (bzw. derer Phosphoramidid-Reagenzien). Mit Hilfe von „Mutator“ kann zu jedem Mischungsverhältnis von Basen innerhalb eines Codons die daraus resultierende Aminosäureverteilung in der Bibliothek an dieser Position berechnet werden, bzw. umgekehrt eine gewichtete Aminosäureverteilung an jeder Position durch ein berechnetes Mischungsverhältnis an Basen vorgegeben werden. Durch die bevorzugte Expression der Wildtyp-Aminosäure an jeder randomisierten Position bleiben bei der Mehrzahl der Mutanten mögliche lokale Wechselwirkungen aus dem Wildtyp erhalten, die für die Stabilität der Mutante notwendig sein könnten - zumal wenn die Struktur des Proteins bislang nicht vorliegt und über lokale Wechselwirkungen nur spekuliert werden darf, wie dies bei der PH-Domäne der Fall ist.

Die Oligodesoxynukleotide zur Randomisierung der Loop 1, 3 und 6 mit den jeweiligen Mischungsverhältnissen an jeder Basenposition sind in Kapitel 2.2.2. aufgeführt. Die randomisierten Oligodesoxynukleotide wurden in der Arbeitsgruppe von Georg Arnold (Genzentrum der LMU München) synthetisiert. Der nicht-äquimolare Einsatz der Basen setzte eine manuell bedienbare Synthesemaschine voraus. Da das entsprechende Gerät (Ecosyn D300, Eppendorf) zugunsten einer moderneren Maschine ersetzt wurde, war eine wiederholte Synthese unter „Mutator“-Vorgaben nicht möglich.

4.1.3. Synthese einer in drei Loops randomisierten Bibliothek von PH-Domänen

Die eigentliche Randomisierung einer Bibliothek erfolgt mit der Synthese der Oligodesoxynukleotide, was im vorherigen Kapitel ausgeführt wurde. Wie aber läßt sich aus einem randomisierten Oligodesoxynukleotid klonierbare (doppelsträngige) DNA gewinnen und wie können drei Segmente (für die drei Loopregionen) voneinander unabhängig zu doppelsträngiger DNA umgesetzt und dann doch zusammen kloniert werden? Das Ziel ist dabei sehr hoch gesteckt: die Bibliothek randomisierter PH-Domänen soll mindestens 10^7 Mutanten umfassen.

Die obigen Fragen lassen sich in einem Absatz beantworten und werden anschließend in diesem Kapitel ausführlich abgehandelt:

Der Einsatz randomisierter Oligodesoxynukleotide in einer laborüblichen PCR-Reaktion führt zwar zu doppelsträngiger DNA, doch sind die randomisierten Positionen beider Stränge nicht komplementär zueinander und die DNA „loopt“ zu Heteroduplex-DNA aus. Heteroduplex-DNA läßt sich (zumindest bei den hier verwendeten randomisierten Oligodesoxynukleotiden) nicht ausreichend restringieren und anschließend ligieren, was in dieser Arbeit gezeigt werden konnte. Zur Vermeidung von Heteroduplex-DNA wurde die „asymmetrische“ PCR entwickelt. Nach deren Anwendung konnte allerdings nur das DNA-Segment von Loop 6 quantitativ restringiert werden. Die beiden DNA-Segmente der Loops 1 und 3 ließen sich nach Anwendung der asymmetrischen PCR dagegen nicht quantitativ restringieren, da der 5'-Bereich zur Restriktion nicht ausreichend lang gewählt war, was entsprechende Untersuchungen offenbaren. Deshalb wurde eine „Verlängerungs-PCR“ entwickelt, die es ermöglicht, ein 5'-verlängertes, randomisiertes Homoduplex-DNA-Fragment zu synthetisieren, das sich quantitativ restringieren läßt. Nach diesen zwei etablierten PCR-Methoden lagen drei randomisierte, klonierbare Fragmente vor, die im Rahmen einer vier-Fragment-Ligation mit einem Vektor zu einer Bibliothek zu klonieren wären. Dazu wurden die drei randomisierten PCR-Fragmente in einem Assemblierungsverfahren zu einem in drei Segmenten randomisierten DNA-Fragmente ligiert und dann in den gewünschten Vektor kloniert. Diese Verfahren führten zu einer Bibliothek von $6 \cdot 10^7$ unabhängig voneinander randomisierten PH-Domänen.

Die schematische Darstellung eines randomisierten Oligodesoxynukleotids ist in Abb. 4.1 wiedergegeben; sie wird auch für die folgenden Abbildungen verwendet.

Heteroduplex-DNA nach Einsatz randomisierter Oligodesoxynukleotide

Heteroduplex-DNA entsteht beim äquimolaren Einsatz randomisierter Oligodesoxynukleotide im Verhältnis zu nicht-randomisierten Gegen-Oligodesoxynukleotiden in laborüblichen PCR-Reaktionen. Dies gilt für die Hälfte aller PCR-Produkte - vorausgesetzt daß sich bei jeder PCR-Runde die Zahl der PCR-Produkte verdoppelt (s. Abb. 4.2). Die Denaturierung und anschließende Hybridisierung randomisierter DNA führt ausschließlich zu Heteroduplex-DNA, da sich statistisch nie zwei völlig komplementäre DNA-Stränge „finden“ werden. Der Einsatz zweier randomisierter Oligodesoxynukleotide wird auch

unter idealen, laborüblichen PCR-Bedingungen immer Heteroduplex-DNA in mindestens einem der beiden randomisierten Segmente zur Folge haben.

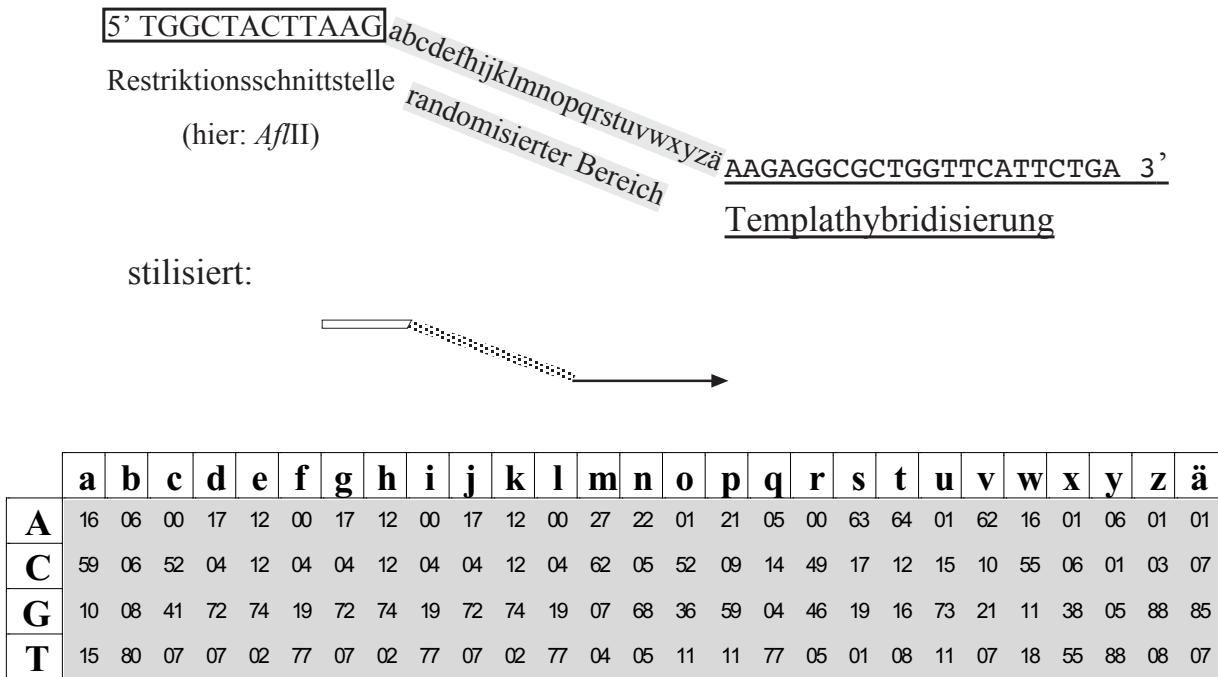


Abb. 4.1: Aufbau des Oligodesoxynukleotids Loop1rms (*sense*) zur Randomisierung des ersten Loops der PH-Domäne.

Die randomisierten Positionen sind mit Kleinbuchstaben markiert, wobei die prozentuale Zusammensetzung an Basen für jede Position tabellarisch aufgeführt ist. Beim Einsatz eines Oligodesoxynukleotides wie Loop1rms (*sense*) in einer Polymerase-Kettenreaktion (PCR) hybridisiert der unterstrichene 3'-Bereich am Templat. Der grau-markierte Bereich beinhaltet die randomisierten Positionen. Der umrahmte Bereich des Oligodesoxynukleotides dient der späteren Klonierung über ein Restriktionsenzym (hier *AflII*).

Heteroduplex-DNA wird nach Einsatz randomisierte Oligodesoxynukleotide beobachtet, wenn sich das entsprechende PCR-Produkt nur unzureichend klonieren läßt. Die Klonierung kann dabei aus zweierlei Gründen versagen: liegt die Restriktionsschnittstelle vor dem randomisierten Bereich nicht als Homoduplex-DNA vor, erkennt das Restriktionsenzym nicht die DNA-Sequenz. Hybridisiert der 5'-Bereich mit dem Gegenstrang, kann dort möglicherweise das Enzym restringieren; dadurch aber verkleinert sich der komplementäre 5'-Bereich und denaturiert, so daß die anschließende Ligation z.B. mit Vektor-DNA nicht erfolgen kann.

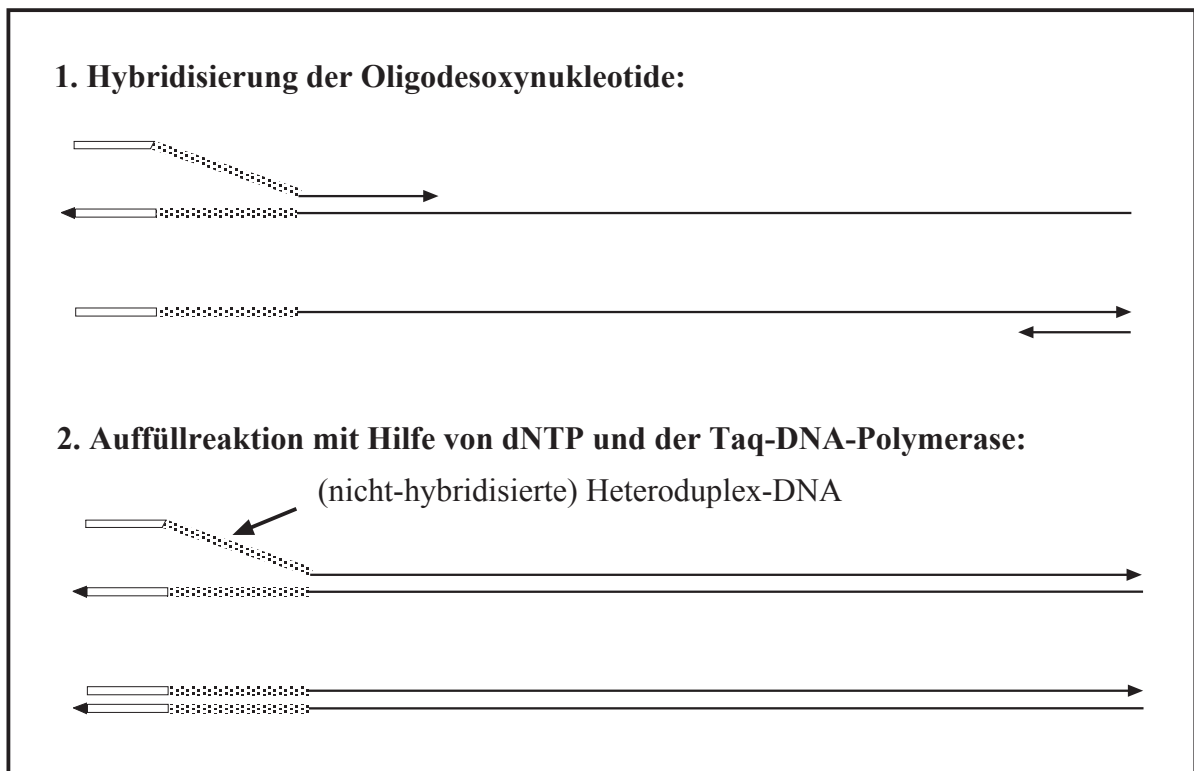


Abb. 4.2: Entstehung von Heteroduplex-DNA.

Im Bereich der Randomisierung und späteren Restriktionsschnittstelle entsteht bei einer PCR-Reaktion unter äquimolarem Einsatz des randomisierten Oligodesoxynukleotids und des nicht-randomisierten Gegen-Oligodesoxynukleotids Heteroduplex-DNA. (Stilisierung s. Abb. 4.1).

Die Existenz von Heteroduplex-DNA läßt sich in seltenen Fällen auch nach gelelektrophoretischer Auftrennung im Agarosegel nachweisen (s. Abb. 4.3): Ausgang dieser Beobachtung ist Heteroduplex-DNA aus einer PCR-Reaktion mit einem randomisierten Oligodesoxynukleotid, die zuvor als eine Bande aus einem Agarosegel extrahiert worden war. Nach erneuter gelelektrophoretischer Auftrennung des gereinigten PCR-Produktes treten weitere Banden in der Höhe der mehrfachen Größen des ursprünglichen PCR-Produktes auf. Dabei hybridisieren die zur Restriktion vorgesehene 5'-Bereich nicht intramolekular mit dem „eigenen“ Gegenstrang, sondern mit dem Gegenstrang eines anderen Heteroduplex-DNA-Moleküls und lassen sich dann als Dimer nachweisen. In gleicher Weise erfolgen noch weitere Hybridisierungen zu höheren Oligomeren, wie in der Abbildung dargestellt ist.

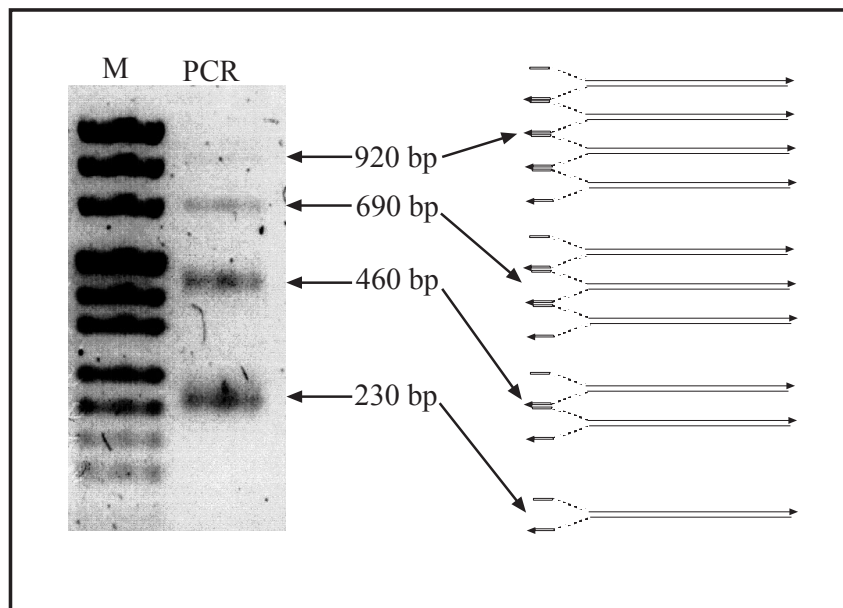


Abb. 4.3: Oligomere durch intermolekulare Hybridisierung randomisierter Heteroduplex-DNA.

Gelelektrophoretische Auftrennung (1.2 % Agarose/TBE) eines zuvor gereinigten DNA-Fragments. Das DNA-Fragment entstammt einer PCR-Reaktion unter äquimolarem Einsatz eines randomisierten und eines nicht-randomisierten Oligodesoxynukleotids. Nach der PCR-Reaktion konnte es als 230 bp-Bande aus einem Agarosegel extrahiert werden (QIAquick Gel Extraction Kit). Dabei wird das DNA-beinhaltende Agarosestück in einem chaotropen Puffer auf 50°C erwärmt, bis es gelöst ist, über eine Silicamatrix aus der Lösung extrahiert und mit Wasser von der Silicamatrix eluiert. Ohne weitere Behandlung erzeugt dieses 230 bp DNA-Fragment nach erneuter Auftrennung im Agarosegel (mit Ethidiumbromid) neben der erwarteten DNA-Bandengröße noch DNA-Fragmente mit dem zwei-, drei- und vierfachen Molekulargewicht. Derartige Oligomerisierungen wurden bislang noch nie bei nicht-randomisierten PCR-Fragmenten beschrieben. Die Erklärung ist rechts schematisch dargestellt: eine intermolekulare Hybridisierung führt zur Oligomerisierung (Stilisierung: s. Abb. 4.1).

Das Verfahren der asymmetrischen Polymerase-Kettenreaktion

Der Einsatz äquimolarer Mengen randomisierten Oligodesoxynukleotids und nicht-randomisierten Gegenprimers in PCR-Reaktionen führte bei den Klonierungsversuchen von einer Einzel-Loop-Bibliothek zu jeweils unter 1000 Klone. Das entwickelte Verfahren der asymmetrischen PCR (s.u. und Kapitel 3.4.2.) ermöglichte dagegen bei Loop 6 die Klonierung von über 10^6 Mutanten unter vergleichbaren Transformationsbedingungen wie bei der Heteroduplex-DNA. Heteroduplex-DNA wird dabei vermieden, indem in einem ersten Schritt einzelsträngige, randomisierte DNA hergestellt wird (3'-Verlängerung des randomisierten Oligodesoxynukleotids) und in einem folgenden Schritt mit einem nicht-randomisierten Gegen-Oligodesoxynukleotid der komplementäre Strang „basentreu“ aufgefüllt wird.

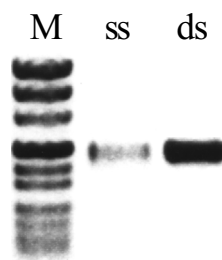
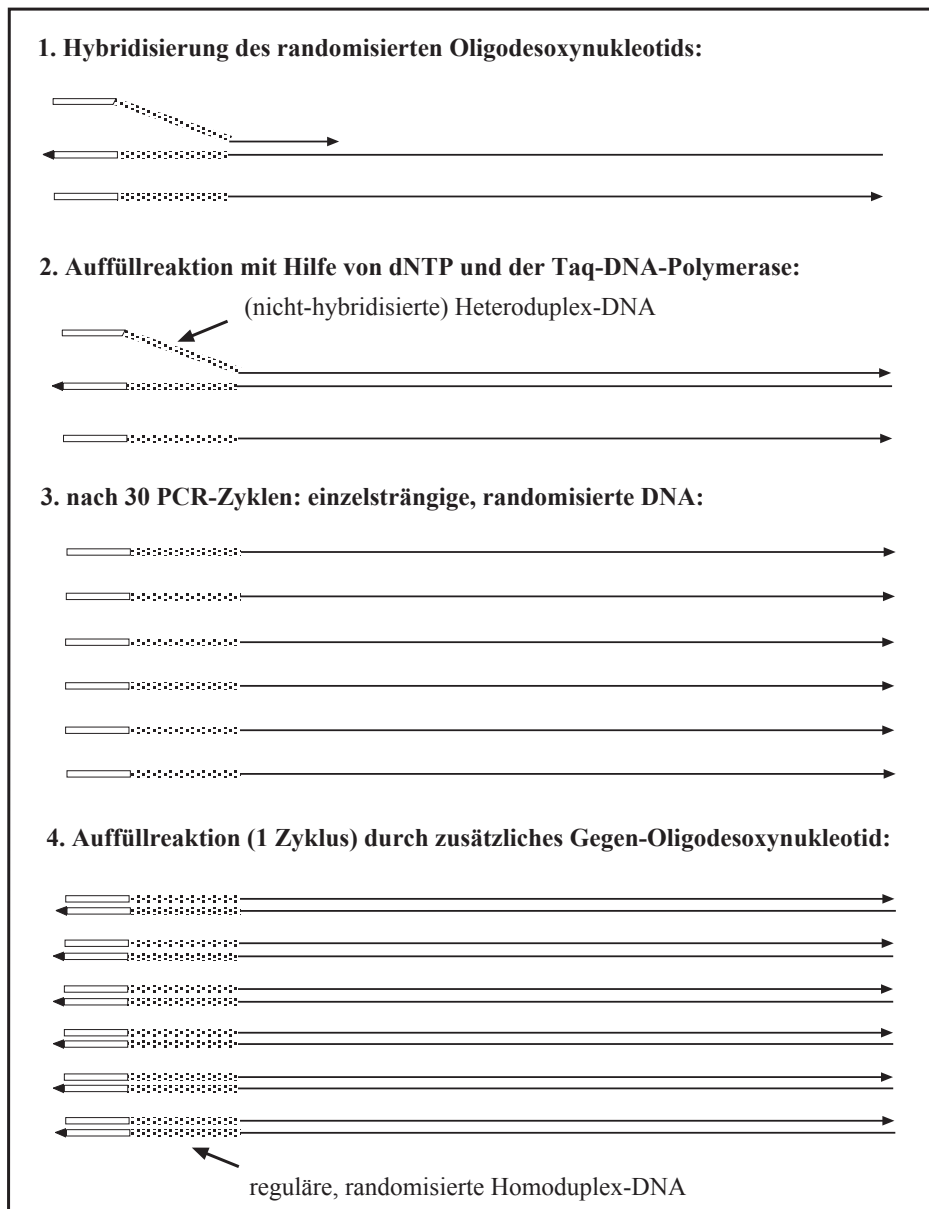


Abb. 4.4: Asymmetrische PCR.

Zur Herstellung randomisierter Homoduplex-DNA wurde die asymmetrische PCR mit anschließender Auffüllreaktion entwickelt. (A) Reaktionsschema mit Stilisierung nach Abb. 4.1 und (B) gelelektrophoretische Auftrennung von bevorzugt ssDNA vor der Auffüllreaktion (ss: 5 µl Auftrag) und nach der Auffüllreaktion über einen Zyklus zu dsDNA (ds: 5 µl Auftrag) am Beispiel von Loop2rms. Erst durch die Auffüllreaktion wird der Großteil an dsDNA synthetisiert.

Bei der asymmetrischen PCR wird in einer ersten PCR-Reaktion über 30 Zyklen das randomisierte Oligodesoxynukleotid im 50-fachen Überschuß gegenüber dem nicht randomisierten eingesetzt; es entsteht das verlängerte, randomisierte Oligodesoxynukleotid als Hauptprodukt (einzelsträngige DNA). Für einen weiteren Zyklus wird dem ersten PCR-Ansatz das nicht-randomisierte Oligodesoxynukleotid äquimolar zum eingesetzten randomisierten zugesetzt; es hybridisiert am Einzelstrang und dieser wird mit der *Taq*-DNA-Polymerase zu regulärer Homoduplex-DNA aufgefüllt (s. Abb. 4.4). Auf diese Weise erhält man randomisierte Homoduplex-PCR-Fragmente, die sich restringieren und klonieren lassen. Dieses Verfahren fand bei Loop 6 mit dem Oligodesoxynukleotid Loop3rms Anwendung (s. Kapitel 2.2.2.).

Das Verfahren der asymmetrischen Verlängerungs-Polymerase-Kettenreaktion

Während sich das DNA-Segment von Loop 6 nach dem Verfahren der asymmetrischen PCR randomisieren und klonieren ließ, ist dies mit demselben Verfahren bei den Loops 1 und 3 nicht direkt möglich. Mit Hilfe eines in dieser Arbeit entwickelten Dimerisierungstests konnte nachgewiesen werden, daß die 5'-Enden (vor den randomisierten Bereichen) nicht effizient restringiert wurden, weil sie trotz gegenteiliger Angabe des Vertreibers von Restriktionsenzymen zu kurz gewählt waren (s. Abb. 4.5).

Eine Neusynthese der randomisierten Oligodesoxynukleotide *Loop1rms.s* und *Loop2rms.s* mit verlängertem 5'-Ende schied aus technischen Gründen aus, weshalb eine weitere PCR-Methode zur Verlängerung der Nukleotidsequenz vor den Schnittstellen der Endonukleasen *AflIII* und *AccI* entwickelt werden mußte (s. Abb. 4.6 und Kapitel 3.4.2.).

Bei der asymmetrischen Verlängerungs-PCR werden in einer konventionellen PCR-Reaktion das randomisierte Oligodesoxynukleotid sowie das nicht-randomisierte Oligodesoxynukleotid äquimolar eingesetzt, so daß randomisierte Heteroduplex-DNA entsteht (s. Abb. 4.3) und diese gereinigt. 1/10 dieser Heteroduplex-DNA dient in einem weiteren PCR-Schritt (gleiches Volumen wie bei der ersten PCR) als Templat. Zur Verlängerung von *Loop1rms.s* wird das Oligodesoxynukleotid *T3-Loop1 ohne.s*, von *Loop2rms.s* das Oligodesoxynukleotid *T3-Loop2 ohne.s* (s. Abb. 4.5) und jeweils kein Gegen-Oligodesoxynukleotid zugesetzt. Nach 12 PCR-Zyklen liegt 5'-verlängerte, randomisierte, einzelsträngige DNA vor, die mit nicht-randomisierten Gegen-Oligodesoxynukleotid zu restringierbarer, „verlängerter“ Homoduplex-DNA aufgefüllt wird.

4. Ergebnisse

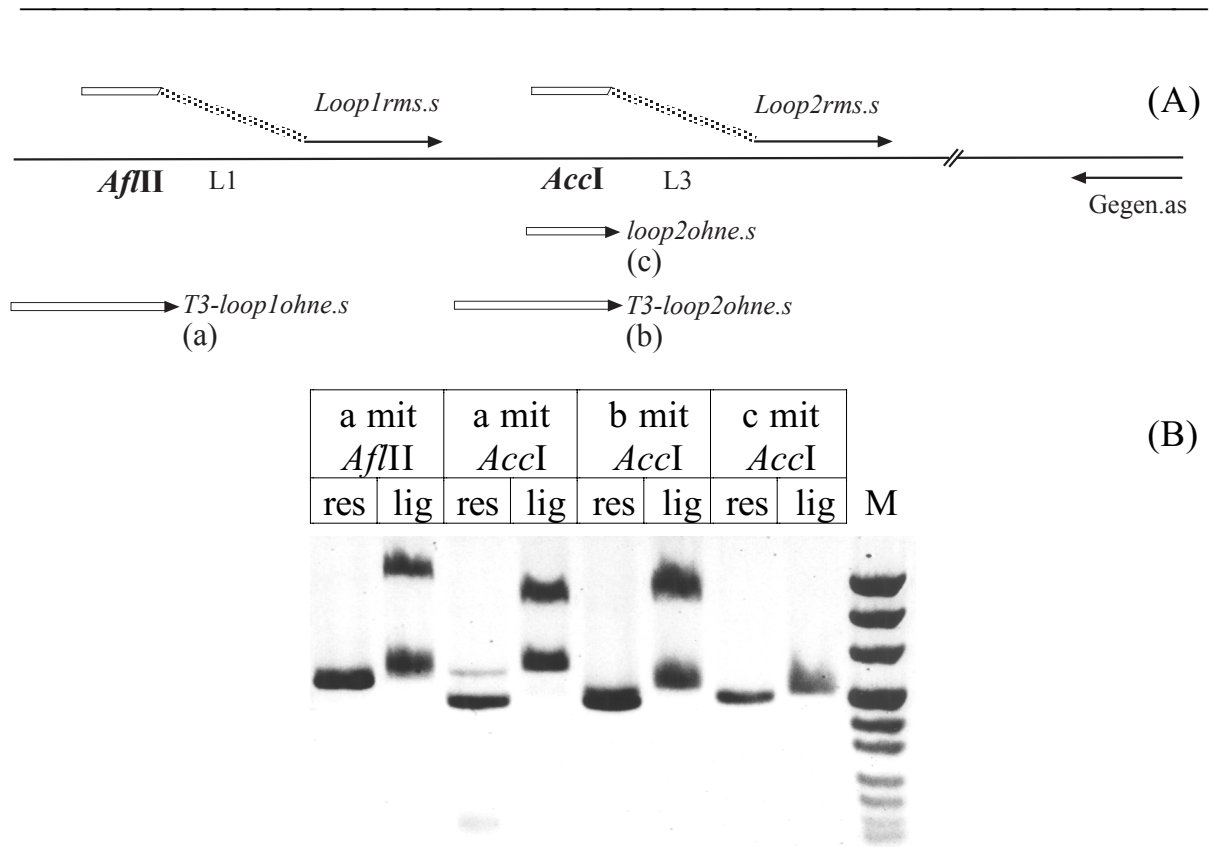


Abb. 4.5: Dimerisierungstest zum Nachweis der Restriktion (*AccI* / *AflII*).

Für diesen Test wurden über PCR-Reaktionen DNA-Fragmente (ohne Randomisierung der Loops L1 und L3) gemäß (A) hergestellt. Diese wurden alternativ mit den Endonukleasen *AflII* und *AccI* restringiert, die Endonukleasen hitzeinaktiviert, der Restriktionsansatz re-ligiert und über ein Agarosegel (1.2 %, EtBr) analysiert. In (B) aufgetragen sind jeweils der restringierte Ansatz vor (res) und nach der Religation (lig). Man erkennt bei den PCR-Produkten mit den Oligodesoxynukleotiden a und b im religierten Ansatz jeweils das zum originalen PCR-Produkt religierte DNA-Fragment und in ungefähr gleicher Ausbeute das Dimerprodukt bei der doppelten Größe. Die Existenz der Dimere beweist die vorherige Restriktion der PCR-Produkte. Produkt c dagegen ließ sich nicht mit *AccI* restringieren und demzufolge der restringierte Ansatz auch nicht dimerisieren. Dabei entspricht das Oligodesoxynukleotid loop 2 ohne.s dem nicht-randomisierten 5'-Ende von *Loop2rms.s*, was den erfolglosen Klonierungsversuch von Loop 3 (mit *Loop2rms.s*) nach der asymmetrischer PCR erklärt.

Bei der asymmetrischen Verlängerungs-PCR werden in einer konventionellen PCR-Reaktion das randomisierte Oligodesoxynukleotid sowie das nicht-randomisierte Oligodesoxynukleotid äquimolar eingesetzt, so daß randomisierte Heteroduplex-DNA entsteht (s. Abb. 4.3) und diese gereinigt. 1/10 dieser Heteroduplex-DNA dient in einem weiteren PCR-Schritt (gleiches Volumen wie bei der ersten PCR) als Templat. Zur Verlängerung von *Loop1rms.s* wird das Oligodesoxynukleotid *T3-Loop1 ohne.s*, von *Loop2rms.s* das Oligodesoxynukleotid *T3-Loop2 ohne.s* (s. Abb. 4.5) und jeweils kein Gegen-Oligodesoxynukleotid zugesetzt. Nach 12 PCR-Zyklen liegt 5'-verlängerte, randomisierte, einzelsträngige DNA vor, die mit nicht-randomisierten Gegen-Oligodesoxynukleotid zu restringierbarer, „verlängerter“ Homoduplex-DNA aufgefüllt wird.

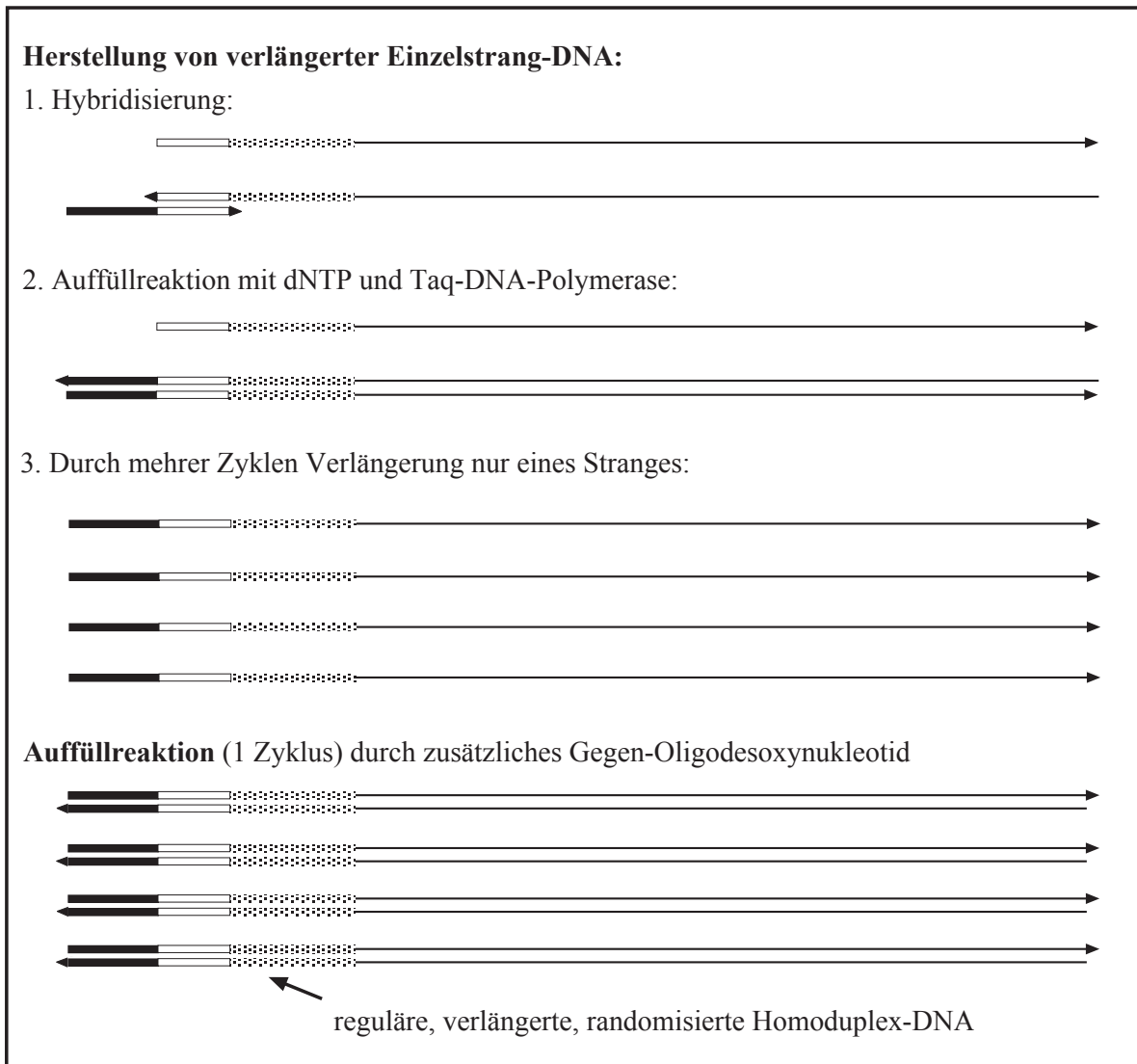


Abb. 4.6: Asymmetrische Verlängerungs-PCR.

Mit dieser Methode wurden die Loops 1 und 3 randomisiert. Es wird 5'-verlängerte, randomisierte Homoduplex-DNA synthetisiert, die sich restringieren und klonieren läßt. Die Stilisierung erfolgt gemäß Abb. 4.1.

Assemblierung der drei getrennt randomisierten PCR-Fragmente

Die durch asymmetrische PCR bzw. asymmetrische Verlängerungs-PCR erzeugten DNA-Fragmente kodieren jeweils für einen randomisierten Loop. Mit dem Dimerisierungstest (s. Abb. 4.5) läßt sich auch ohne Klonierung zeigen, ob sich die randomisierten DNA-Segmente ligieren lassen.

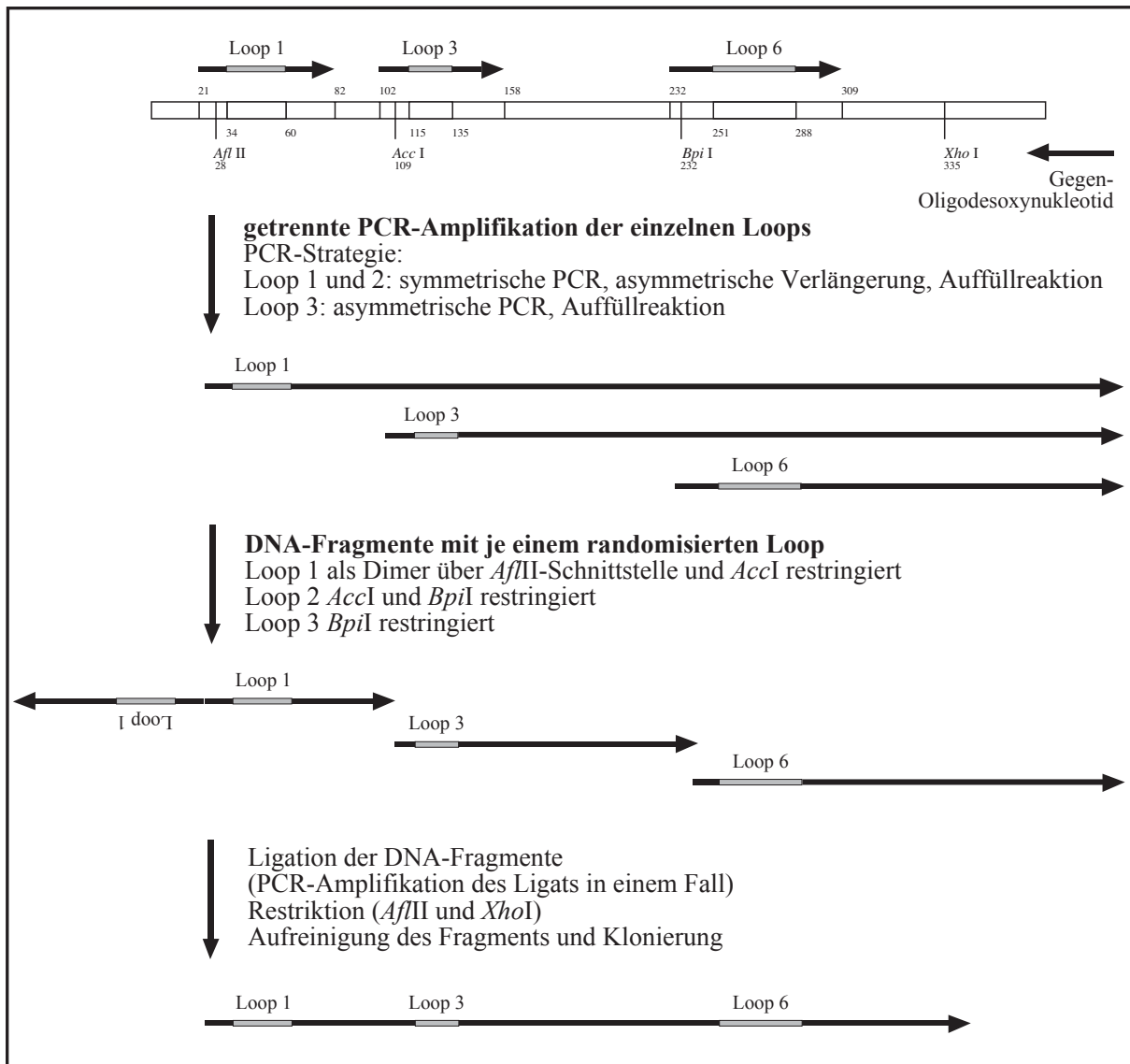


Abb. 4.7: Klonierung der in drei Loops unabhängig randomisierten PH-Domäne.

Zur Klonierung einer in drei Loops randomisierten Bibliothek von PH-Domänen wurden zunächst die Loops individuell mit obigen Methoden randomisiert. Ein Ligationsansatz mit drei randomisierten Fragmenten und Vektor (vier-Fragment-Ligation) zur Herstellung einer großen Bibliothek von über 10^7 Mutanten ist nicht durchführbar. Deshalb wurde aus den randomisierten Fragmenten über diverse Restriktions- und Ligationsschritte ein Fragment ligiert und dieses in einer zwei-Fragment-Ligation in den *phage-display*-Vektor plomPHg3 kloniert (s. Abb. 4.7). Bei 1/6 aller Mutanten der erzeugten Bibliothek wurde das in drei Segmenten randomisierte Fragment über PCR mit 12

Zyklen amplifiziert und dann kloniert. Unvollständig restringierte DNA wurde jeweils entfernt, indem wie beim Dimerisierungstest nach der Restriktion die Fragmente dimerisiert und als solche nach gelelektrophoretischer Auftrennung isoliert wurden. Durch Elektroporation in TG 1-*E. coli*-Zellen konnte eine Bibliotheksgröße von $6 \cdot 10^7$ unabhängigen Klonen erzielt werden.

Durch die getrennte Herstellung der randomisierten Segmente und den Verzicht (zumindest bei 5/6 der Bibliothek: s.o.) auf eine nachträgliche PCR-Amplifikation mit vielen Zyklen kann ausgeschlossen werden, daß individuelle, randomisierte DNA verstärkt amplifiziert bzw. angereichert und damit die Komplexität der Bibliothek erniedrigt wird.

4.1.4. Validierung der Bibliothek

Eine Beurteilung der Güte einer Bibliothek kann nur *ex eventu* erfolgen: eine Bibliothek von Peptidsequenzen ist „gut“, wenn sie diejenigen Peptidsequenzen enthält, welche zur Lösung der vorgegebenen Zielsetzung führen (z.B. Bindung an einem anderen Protein). Die erfolgreiche Anwendung einer Bibliothek setzt allerdings voraus, daß die Methode zur Selektion der gewünschten Mutanten geeignet gewählt ist. Folgende Kriterien zur Überprüfung einer Bibliothek sind wichtig:

- (a) die Größe der Bibliothek: die Zahl unabhängiger Klone,
- (b) die Expression und Faltung der randomisierten Proteine,
- (c) die Diversität der Bibliothek durch die Zahl von randomisierten Aminosäuren.

ad (a)

Die Größe der Bibliothek beträgt $6 \cdot 10^7$ Klone mit unabhängig voneinander randomisierten Loops.

ad (b)

Aus der Bibliothek wurden willkürlich 50 Klone gewählt und cytoplasmatisch exprimiert. Bei über 60 % der Klone ließen sich die jeweilige lösliche PH-Domänen über Western-Blot Analyse mit einem *anti*-His-tag *scF_v*-Fragment nachweisen (s. Abb. 4.8: Lindner *et al.*, 1997). Bei den restlichen untersuchten Klonen lagen Fehler in der DNA-Sequenz vor (*stop*-Codons, Verschiebungen im Leserahmen). Nicht-exprimierte PH-Domänen mit korrekt randomisierter DNA-Sequenz wurden nicht nachgewiesen (s. Tab. 4.1).

ad (c)

34 der 50 untersuchten Mutanten wurden sequenziert, 29 von ihnen exprimieren lösliche PH-Domänen (s. Tab. 4.1). Dabei sind die nicht-exprimierenden Klone 8, 15, 20 und 29 im Leserahmen verschoben, was sich auf Fehler während der PCR-Reaktionen zurückführen läßt und die Mutante 17 besitzt außerhalb der randomisierten Segmente ein *stop*-Codon. Deletionen in den Loops führen nicht notwendigerweise zu nicht-exprimierbaren Proteinen: die Mutante 28 weist sieben und die Mutante 30 sogar neun deletierte Aminosäurepositionen in Loop 6 auf; trotzdem lassen sie sich cytoplasmatisch exprimieren.

Die 29 sequenzierten und exprimierbaren Mutanten besitzen gegenüber dem Wildtyp durchschnittlich sieben Punktmutationen in den drei randomisierten Loops. Die höchste erfaßte Zahl von Randomisierungen beträgt 13, die niedrigste 1. Eine Wildtypkontamination ohne Randomisierungen konnte nicht nachgewiesen werden.

Klon	pos	8	9	11	12	13	14	15	neg
Expression	+	-	+	+	+	+	+	-	-

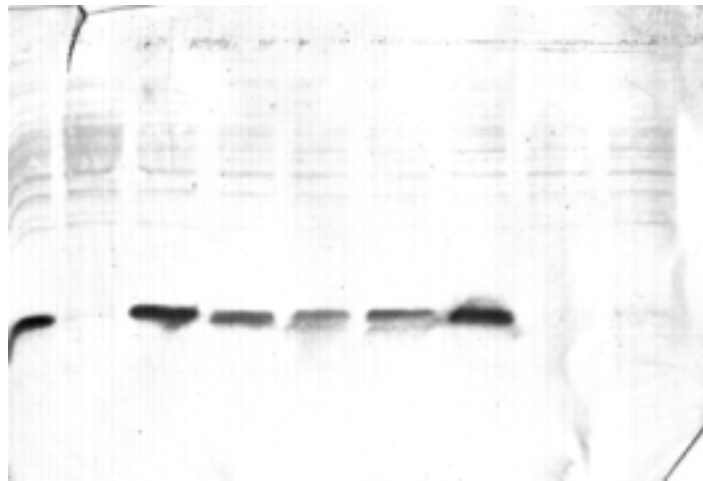


Abb. 4.8: Expressionsnachweis von randomisierten PH-Domänen aus Zellextrakt mit einem *anti*-His tag *scF_V*-Fragment im Western-Blot.

50 willkürlich ausgewählte, randomisierte Mutanten der PH-Domäne mit C-terminalem (His)₆-tag wurden cytoplasmatisch exprimiert (s. Kapitel 3.6.3.: nativer Zellaufschluß) und die Expression mittels eines *anti*-His tag *scF_V*-Fragment im Western-Blot kontrolliert (s. Kapitel 3.8.4.). Davon ließen sich 31 im Western-Blot nachweisen. Die mit "+" gekennzeichneten Mutanten zeigen einen C-terminalen His-tag, die mit "-" gekennzeichneten Mutanten weisen keinen His-tag auf, da sie entweder nicht exprimiert werden oder aber nicht löslich vorliegen; die Aminosäuresequenzen der im Blot aufgeführten Mutanten sind in Abb. 4.9 angegeben. "pos" ist die cytoplasmatisch exprimierte Wildtyp PH-Domäne und "neg" eine Probe ohne exprimierte PH-Domäne.

Tab. 4.1: Aminosäuresequenzen und cytoplasmatische Expressionen von 34 zufällig ausgewählten, randomisierten PH-Domänen.

Klon#	Exp	Loop1	Loop3	Loop6	rms#
1	+	R-L	H	L	4
3	+	DK-V	T	Y	5
4	+	Q	P		2
6	+	MN	M		3
8	-	FGRGA.*			
9	+	A-K-N	A	A-HL	7
11	+	SHS	M-H		5
12	+	VR-RT-M	H	A-R-Q-KN	11
13	+	KCLR	N	R-MGVHQ-A	12
14	+	SGM-N	A	R-R-S	8
15	-	APM-K	S		
16	+	SQS-K	IVQ	RA-YS	11
17	-	SMNM	I-S	A-VRR-NL	12
18	+	V			1
19	+	D-H-C	Q-TK	G-P-ARGRV	13
20	-	SS-A*			
23	+	D-ME	IA	SPA-A-AN	11
25	+	R-SL	H-E-T	G-RF-K-K	11
27	+	W-DI	NQ-S	N-SL-AC-I	12
28	+	A-A-Y	V-QK-S	N.....Y	9+
29	-	A-GHY	PGYH-FRE.*		
30	+	ND-L	R-S	NR.....	7+
31	+	E-N			2
32	+	D-S-L	S-H		5
33	+	MN	M		3
34	+	D-M	N		3
35	+	S	L	V-K-R-LQ-M	8
36	+	S-QGQ	P	HL-D-N	9
38	+	A	NN-Q	RL-RVK	9
41	+	D-R-N		DS-RS-QR-LI	11
45	+	NL		G-K-Q-A	6
47	+	SD			2
49	+		YQQ		3
50	+	I-M-N	H	R-L-S	7
WT	+	KLGGGRVKTWK....	TTDKEPRGI....	TEADGQVVEGNHTVY.*	

Grau unterlegt sind die randomisierten Bereiche von Loop 1, 3 und 6. Unten ist die Aminosäuresequenz des Wildtyps in den Loopregionen angegeben. Stimmen bei den Mutanten die Aminosäurereste mit dem Wildtyp überein, ist dies mit "-" gekennzeichnet. Die Zahl von Aminosäuren, die im Bereich der ausgewählten Loops jeder Mutante vom Wildtyp abweichen, ist unter rms# angegeben. Mit "+" sind die cytoplasmatisch exprimierbaren PH-Domänen markiert. Die Mutanten 28 und 30 lassen sich trotz Deletion in Loop6 exprimieren. Die cytoplasmatische Herstellung von löslichen Protein erfolgte nach Kapitel 3.6.3. und der Nachweis über den C-terminalen His-tag im Western-Blot (s. Kapitel 3.8.4. und Abb. 4.8).

Die sequenzierten (und exprimierbaren) Mutanten sind in ihren Loops sehr unterschiedlich häufig mutiert:

- Loop 1 beinhaltet die Aminosäuren 11 bis 19. Von diesen acht Aminosäuren entsprechen durchschnittlich 3.0 nicht der jeweiligen Wildtyp-Sequenz. Nur Klon 49 ist in Loop 1 nicht randomisiert.
- Loop 3 reicht von Aminosäure 38 bis 44. Von den sieben Aminosäuren sind durchschnittlich 1.7 nicht mit der Wildtyp-Sequenz identisch, wobei vier PH-Domänen in Loop 3 nicht randomisiert vorliegen.
- Loop 6 ist zwischen den Aminosäuren 83 und 95 lokalisiert. Von den 12 Aminosäuren weichen durchschnittlich 2.9 von der Wildtyp-Sequenz ab, wobei allerdings 10 PH-Domänen in diesem Loop keine Randomisierung aufweisen.

Die durchschnittliche Randomisierung an jeder vorgesehenen Position über alle Loops beträgt damit 27 %. Bei der statistischen Betrachtung fällt auf, daß bei einigen Mutanten Loops gar nicht randomisiert sind. Dabei handelt es sich wahrscheinlich um Wildtyp-Sequenzen, die nicht durch die randomisierten Oligodesoxynukleotide sondern über das (Wildtyp-) Templat während der PCR-Reaktionen in die Bibliothek gelangt sind. Zum Vergleich mit der verwendeten Randomisierungsstrategie nach „Mutator“ ist es deshalb rechnerisch sinnvoll, alle nicht-randomisierten Loops zu vernachlässigen. Dann beträgt die durchschnittliche Wahrscheinlichkeit einer Randomisierung an jeder Position 32 % bzw. liegen durchschnittlich 9.3 randomisierte Aminosäuren pro Mutante vor. Bei der Planung der randomisierten Oligodesoxynukleotide mit Hilfe des Programms „Mutator“ wurde aber von durchschnittlich 16 randomisierten Aminosäuren pro Mutante der PH-Domäne ausgegangen (Steipe & Bruhn, in Vorbereitung: s. Kapitel 4.1.3.). Setzt man die korrekte Berechnung der Randomisierung mit „Mutator“ und die korrekte Umsetzung in randomisierte Oligodesoxynukleotide bei der automatisierten Synthese voraus, beruht die einzige verbleibende Erklärung auf der Anreicherung von „wildtyp-ähnlicher“ DNA während der unterschiedlichen PCR-Reaktionen. Am Templat hybridisiert präferenziell ein Oligodesoxynukleotid besser, dessen Sequenz weitgehend mit der des Templats übereinstimmt, als eines mit einer erheblich höheren Abweichung vom Wildtyp. Ausschließen ließe sich dies bei der Klonierung zukünftiger Bibliotheken, indem der Bereich des Templats, der bei den Oligodesoxynukleotiden randomisiert vorliegt, durch Restriktion entfernt wird; somit könnte über diesen Bereich keine Bevorzugung wildtyp-ähnlicher, randomisierter Oligodesoxynukleotide mehr erfolgen.

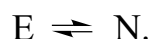
4.1.5. Untersuchung der thermodynamischen Stabilität von PH-Domänen

Die Expressionsausbeute eines Proteins wird quantitativ von seiner Stabilität im gefalteten Zustand beeinflusst. Ungefaltete und damit unstrukturierte Proteine werden in Zellen von Proteasen hydrolysiert oder aggregieren, weshalb die Stabilität ein Charakteristikum für die Haltbarkeit eines Proteins in der Zelle ist. So können Immunglobulin-V_L-Domänen erst ab einer thermodynamischen Stabilität von -12 kJ/mol cytoplasmatisch (reduziert) exprimiert werden (Ohage & Steipe, 1999).

Voraussetzung für die Randomisierung von Aminosäurepositionen in Proteinen ist der geringe Beitrag dieser Aminosäuren an der Gesamtstabilität. Dieser Beitrag kann nicht vorhergesagt werden, sollte aber im Fall der randomisierten PH-Domäne gering sein (s. Kapitel 1.4.1.). Bei der Randomisierung mit „Mutator“ werden die meisten Wildtyp-Aminosäuren beibehalten und nur wenige mutiert, so daß für die Stabilität essentiellen Aminosäuren in den meisten randomisierten Mutanten vorhanden sein sollten.

Die Stabilität eines Proteins läßt sich auf unterschiedliche Art und Weise definieren: es läßt sich die Proteasestabilität bestimmen, indem Degradationskinetiken in Gegenwart von Proteasen untersucht werden. Die thermische Stabilität läßt sich überprüfen, indem der Faltungszustand des Proteins bei unterschiedlichen Temperaturen analysiert wird. Neben der Einstellung eines thermodynamischen Gleichgewichts zwischen nativer und entfalteter Spezies bei unterschiedlichen Temperaturen aggregiert dabei häufig das entfaltete Protein, wodurch diese Spezies aus dem thermodynamischen Gleichgewicht gezogen wird. Um diesen kinetischen Effekt zu vermeiden, untersucht man die thermodynamische Stabilität mit Hilfe von Denaturierungsmitteln, so daß das entfaltete Protein weiterhin im thermodynamischen Gleichgewicht zum gefalteten Protein steht.

Der Faltungsvorgang eines Proteins kann im Idealfall durch ein reversibles, dynamisches Gleichgewicht zwischen einer gefalteten, „nativen“ (N) und einer entfalteten, denaturierten Spezies (E) beschrieben werden:



Um den Faltungszustand eines Proteins beurteilen zu können, benötigt man eine Sonde, mit deren Hilfe man Moleküle im Zustand „E“ von denen im Zustand „N“ unterscheiden kann. Im Fall der PH-Domänen eignet sich hierfür die intrinsische Fluoreszenz der aromatischen Aminosäuren Tryptophan und Tyrosin. Die PH-Domäne besitzt als Wildtyp dreizehn Tyrosine und vier Tryp-

tophane (s. Tab. 4.4). Das Tryptophan (109) ist dabei ein Aminosäurerest der α -Helix und orientiert sich in das hydrophobe Innere des Proteins. Es ist umgeben von den Aminosäureresten der β -Faltblätter 1 und 2 sowie den Aminosäureresten der α -Helix. Es besitzt damit eine stark strukturierte, unpolare Umgebung. Bei einer Anregung (Exzitation) mit $\lambda_{\text{ex}} = 282 \pm 2.5$ nm ist ein Fluoreszenzmaximum (Emission) bei $\lambda_{\text{em}} = 335$ nm zu beobachten (s. Abb. 4.9). Entfaltet man die PH-Domäne durch hohe Konzentrationen an Denaturierungsmittel wie beispielsweise Guanidiniumchlorid (GdmCl), wird das Trp(109) aus seiner unpolaren, strukturierten Umgebung in die polare, unstrukturierte Umgebung der wässrigen Lösung überführt. Dadurch verschiebt sich bei gleicher Exzitationswellenlänge das Emissionsmaximum in den längerwelligen Bereich zu $\lambda_{\text{em}} = 350$ nm (s. Abb. 4.9). Das Differenzspektrum des gefalteten zum ungefalteten Protein besitzt sein Maximum bei $\lambda_{\text{max}} = 327$ nm. Durch die Messung der Fluoreszenzintensität bei dieser Emissionswellenlänge läßt sich das Verhältnis zwischen entfalteten und gefalteten Spezies sehr empfindlich verfolgen.

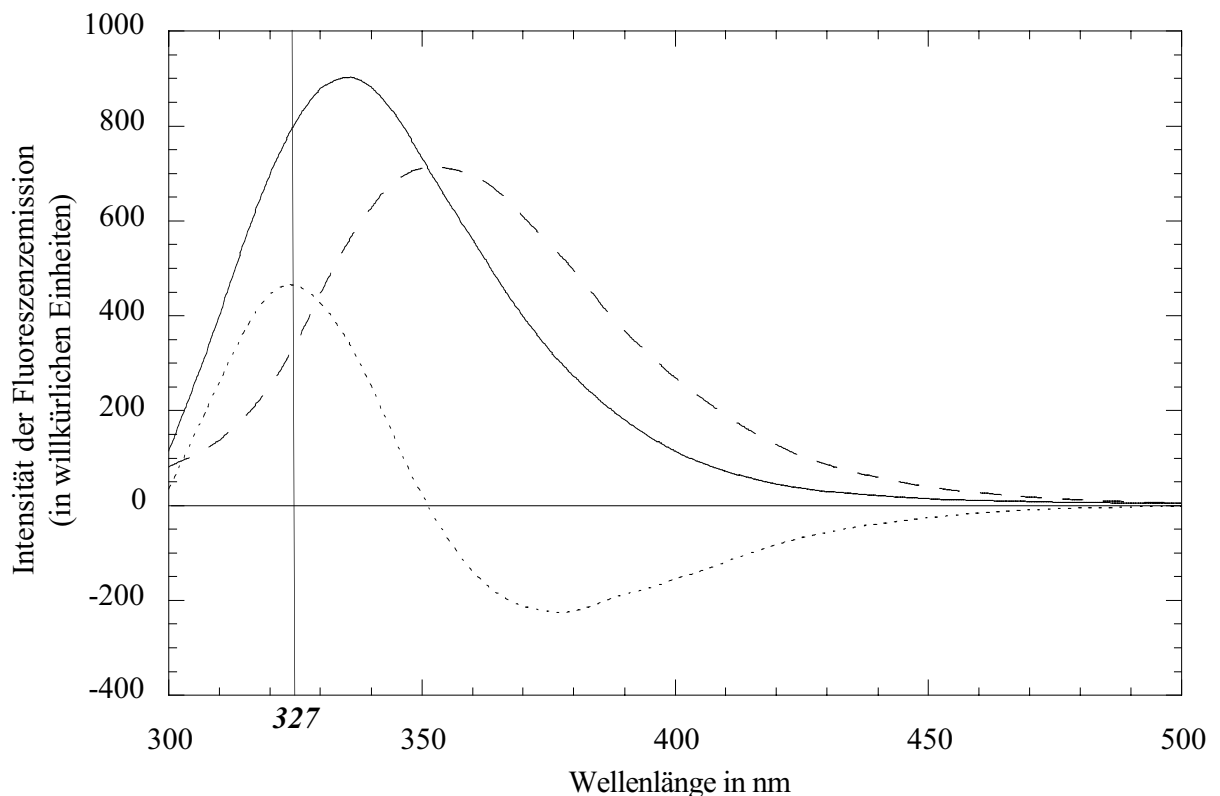


Abb. 4.9: Fluoreszenzemissionsspektren der Wildtyp PH-Domäne.

Durchgezogene Linie: gefaltetes Protein in PBS-Puffer, gestrichelte Linie: entfaltetes Protein in PBS/5M GdmCl, gepunktete Linie: Differenzspektrum (Emission des gefalteten Proteins abzüglich der Emission des ungefalteten Proteins). Anregungswellenlänge 280 ± 2.5 nm, Emissionsschlitzbreite 5 nm, Proteinkonzentration $5 \mu\text{M}$. Durch Entfaltung in 5 M GdmCl verschiebt sich die Tryptophan-Fluoreszenz in den längerwelligen Bereich. Bei 327 nm ist die Differenz der Emission zwischen gefaltetem und ungefaltetem Protein am größten.

Zur Bestimmung der Faltungsenergie eines Proteins werden sogenannte „Übergangskurven“ vermessen (Pace, 1990). Hierzu wird die Tryptophan-Fluoreszenz des Proteins in Gegenwart unterschiedlicher Konzentrationen an Denaturierungsmittel „[D]“ nach Anregung bei 280 nm bestimmt (s. Kapitel 3.10.3.). Bei Auftragung der Fluoreszenz gegen [D] zeigt sich, daß bei niedrigen Konzentrationen des Denaturierungsmittels überwiegend N, bei hohen überwiegend E existiert, was sich jeweils in linearen Basislinien äußert. Dazwischen liegt der Übergangsbereich, in dem E und N zu ähnlichen Anteilen vorliegen (s. Abb. 4.10A).

Zur Ermittlung der Faltungsstabilität (Freie Faltungsenthalpie) ΔG_F^P geht man folgendermaßen vor: Das Gleichgewicht



ist charakterisiert durch die Gleichgewichtskonstante

$$(2) \quad K = \frac{[N]}{[E]} .$$

K ist über die Gleichung

$$(3) \quad \Delta G_F = -RT \ln K$$

mit der Energiebilanz der Reaktion ΔG_F verknüpft, wobei R die allgemeine Gaskonstante (8.314 J/K·mol) und T die absolute Temperatur bezeichnet. ΔG_F ist abhängig von der jeweiligen Konzentration des Denaturierungsmittels. Gesucht ist die Faltungsstabilität ΔG_F^P im reinen Puffer ohne Denaturierungsmittel (Index „P“: PBS-Puffer).

Zu ΔG_F^P gelangt man, indem man für jeden Punkt im Übergangsbereich die Abstände a und b zu den jeweiligen Basislinien bestimmt (s. Abb. 4.10A). Daraus ergibt sich das Verhältnis von entfaltetem zu gefaltetem Protein K , woraus sich über Gleichung 3 ΔG_F ergibt. Man trägt nun die so ermittelten Werte für ΔG_F gegen [D] auf und konstruiert die Ausgleichsgerade

$$(4) \quad \Delta G_F = \Delta G_F^P + m \cdot [D].$$

ΔG_F^P , der Abschnitt auf der y-Achse, ist dann der extrapolierte Wert der Stabilität auf den reinen Puffer ohne Denaturierungsmittel. Da die Faltungsreaktion von links nach rechts formuliert ist (Gleichgewicht 1), ist $\Delta G_F^P < 0$. Die Steigung der Geraden wird als m -Wert oder „Kooperativität“ bezeichnet. Sie ist ein Maß für die Empfindlichkeit des Faltungsübergangs gegenüber dem Denaturierungsmittel.

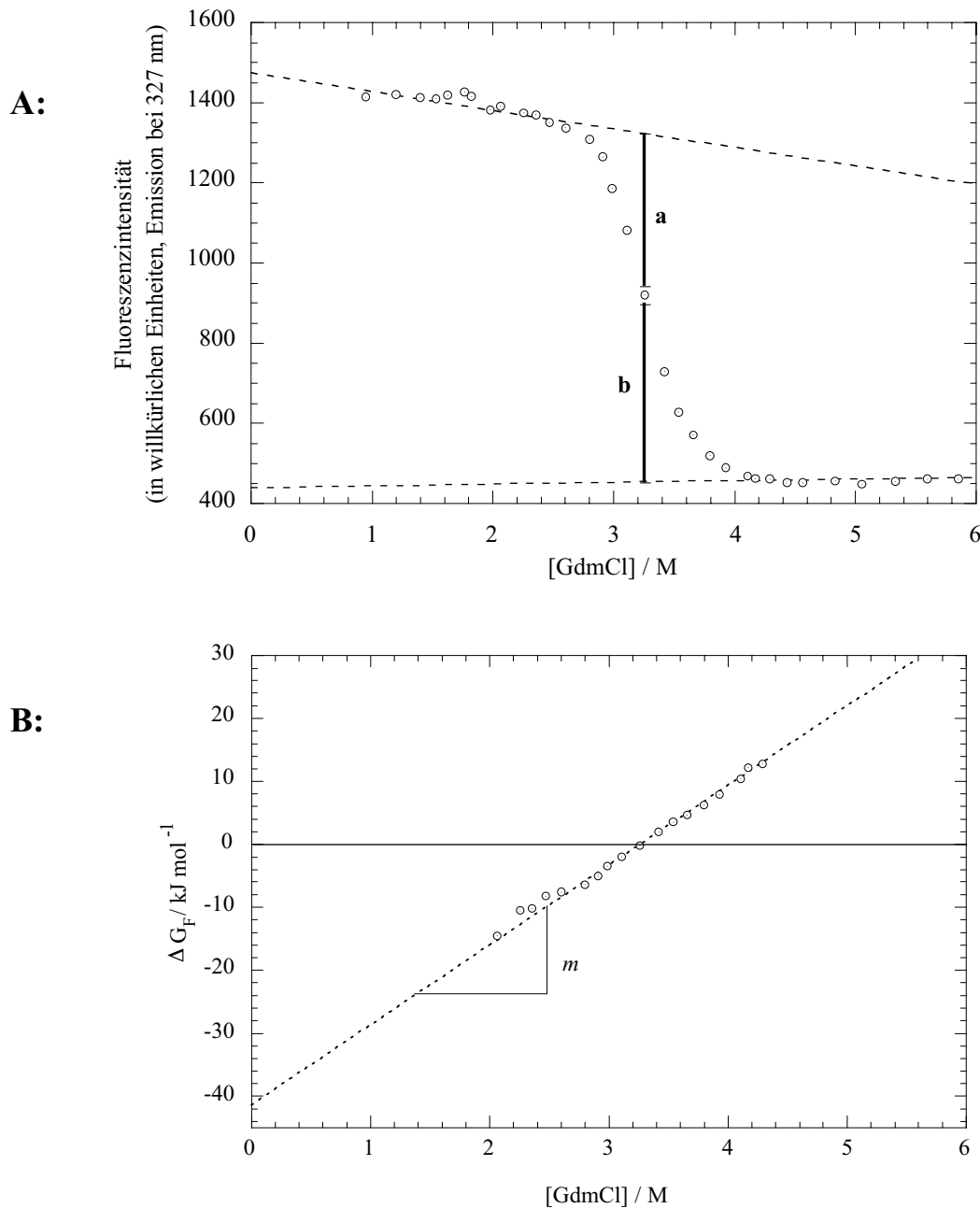


Abb. 4.10.: Übergangskurven der Wildtyp PH-Domäne.

A: Die Fluoreszenz ($\lambda_{\text{ex}} = 282 \text{ nm}$, $\lambda_{\text{em}} = 327 \text{ nm}$) der Wildtyp PH-Domäne nimmt mit steigender Konzentration des Denaturierungsmittels GdmCl in einem sigmoidalen Verlauf ab. Die Basislinien für den nativen (N) und entfalteten Zustand (E) sind als gestrichelte Linien dargestellt. Im Übergangsbereich kann aus den Abständen der jeweiligen Meßpunkte zu den Basislinien die Gleichgewichtskonstante K berechnet werden:

$$\frac{b}{a+b} = \frac{[\text{E}]}{[\text{N}]+[\text{E}]} = \frac{1}{1+K}$$

B: Die berechneten Werte für K lassen sich mit $\Delta G = -RT \ln K$ als Energien beschreiben und erneut gegen $[\text{GdmCl}]$ auftragen. Aus der Extrapolation gegen Null ergibt sich die freie Faltungsenthalpie des Proteins in PBS. Die Steigung der Gerade liefert den m -Wert, ein Maß für die Kooperativität des Übergangs.

Die Gleichungen (2) bis (4) können mit den Geradengleichungen der Basislinien des nativen ($g_N[D] = a_N \cdot [D] + b_N$) bzw. des entfalteten Proteins ($g_E[D] = a_E \cdot [D] + b_E$) zu einer einzigen Gleichung kombiniert werden. Sie beschreibt die Fluoreszenz in Abhängigkeit von der Denaturierungsmittel-Konzentration:

$$(5) \quad Fl.([D]) = a_N \cdot [D] + b_N + \frac{(a_E \cdot [D] + b_E) \pm (a_N \cdot [D] + b_N)}{\exp\left(\frac{\Delta G_F^P \pm m \cdot [D]}{RT}\right) + 1}$$

Durch Anpassen der Meßwerte an Gleichung (5) erhält man so in einem Schritt die Parameter ΔG_F^P , m -Wert, Steigungen sowie die Achsenabschnitte für die Basislinien. Diese Methode besitzt den Vorteil, daß jeder Meßpunkt zur Ermittlung des Ergebnisses beiträgt, während bei der Extrapolationsmethode wie in Abb. 4.10B durch ungenaue Auswahl von Basislinien und Übergangsbereich erhebliche Fehler entstehen können.

Bei der Formulierung der Faltungsreaktion als dynamisches Gleichgewicht zwischen N und E wird vorausgesetzt, daß die Reaktion *reversibel* verläuft. Um dies experimentell zu überprüfen, wurden parallel drei Übergangskurven angesetzt. Eine wird aus nativem, die zweite aus denaturiertem Ausgangsmaterial und die dritte Kurve aus PBS- und GdmCl-Stammlösungen pipettiert, die beide das Protein bereits enthalten. Diese Methode hat den Vorteil, daß die Proteinkonzentration unabhängig von Pipettierfehlern beim Mischen der Stammlösungen ist (s. Kapitel 3.10.3.) Für eine reversible Faltungsreaktion erwartet man in allen drei Fällen den gleichen Kurvenverlauf. Dies ist für die Wildtyp PH-Domäne der Fall (persönliche Mitteilung von Heike Bruhn). Da alle Mutanten von dieser PH-Domäne abstammen, wurde für sie generell eine reversible Faltung nach einem Zweizustandsmodell angenommen.

4.1.6. Stabilität von randomisierten PH-Domänen

Zehn willkürlich ausgewählte, randomisierte PH-Domänen (s. Tab. 4.1) wurden cytoplasmatisch exprimiert und gereinigt. Von ihnen wurden mit GdmCl fluoreszenzspektroskopisch Übergangskurven aufgenommen und daraus die thermodynamische Stabilität bestimmt (s. Tab. 4.2).

Tab. 4.2: Thermodynamische Stabilität randomisierter PH-Domänen.

Vermessene PH-Mutanten	Zahl der Mutationen	ΔG_F^P (kJ/mol)	m (kJ/mol·M)	$\Delta\Delta G_F^P$ (kJ/mol)
1	4	-35,2	-13,6	+2,0
3	5	-36,0	-14,0	+1,2
9	7	-31,4	-11,5	+5,8
11	5	-36,2	-13,8	+1,0
23	11	-29,0	-11,8	+8,2
32	5	-30,1	-11,1	+7,1
33	3	-33,5	-11,5	+3,7
35	8	-27,8	-11,8	+9,4
47	2	-26,9	-10,5	+10,3
50	7	-10,7	-7,6	+26,5
WT	0	-37,2	-11,8	0

Die thermodynamische Stabilität von randomisierten PH-Domänen (s. Tab 4.1) wird durch die freie Faltungsenthalpie ΔG_F^F wiedergegeben. Die Differenz $\Delta\Delta G_F^F$ der Mutanten zum Wildtyp verdeutlicht den Verlust an Stabilität durch die Mutationen. Die Kooperativität m wird durch die Mutationen nicht einheitlich beeinflusst. Die Übergangskurven wurden in GdmCl bei 20°C im Rahmen eines Praktikums vermessen.

Alle vermessenen Mutanten werden durch willkürliche Punktmutationen in den ausgewählten Loops gegenüber dem Wildtyp destabilisiert. Eine Korrelation zwischen der Zahl von Mutationen und der Destabilisierung oder auch der Kooperativität konnte nicht nachgewiesen werden (s. Tab. 4.2). Diese Untersuchungen bestätigen das Konzept von „Mutator“, wonach von den 29 Aminosäuren dreier Loops nur durchschnittlich 7 (vorgegeben: 16) Aminosäuren mutiert werden, während die restlichen Aminosäuren dieser Loops beibehalten werden und sich stabilisierend auf das Gerüst der PH-Domäne auswirken können. Bei einer rigorosen Randomisierung z.B. mit der Codonwahl NNB (s. Kapitel 5.2.2.) wäre zu befürchten, daß sich ein Großteil der Mutanten nicht exprimieren läßt, wenn schon wenige Mutationen zu einer deutlichen Destabilisierung beitragen, wie es Tab. 4.2 zeigt.

Die Randomisierung der PH-Domäne an sich und der Nachweis, daß sich die Mutanten einer derartigen Bibliothek stabil exprimieren lassen, sind Grundvoraussetzung für den Einsatz in Selektionsverfahren. Dabei wird innerhalb der Bibliothek nach Mutanten gesucht, die neuartige Eigenschaften aufweisen. In dieser Arbeit diente dazu das *phage-display*.

4.2. Die Bibliothek randomisierter PH-Domänen im *phage-display*-Verfahren

Das Prinzip des *phage-display* wird in Kapitel 1.5.1. eingehend erläutert. Mit diesem Verfahren wurde aus der Bibliothek randomisierter PH-Domänen nach Mutanten gesucht, die an Zielsubstanzen binden. Welche Zielsubstanzen sich dafür besonders eignen, läßt sich erst *ex eventu* feststellen, weshalb testweise sehr unterschiedliche Zielsubstanzen ausgewählt wurden.

4.2.1. Ausgewählte Zielsubstanzen für das *phage-display*-Verfahren

Die Zielsubstanzen wurden an NUNC-Maxisorb-Röhrchen durch Denaturierung immobilisiert (s. Kapitel 3.9.3.). Folgende drei wurden verwendet:

- Bei CD3 (CD: „*cluster of differentiation*“) handelt es sich um ein Marker-Antigen für T-Zellen. Von diesem Protein wurde die ϵ -Domäne (12 kD) für das *phage-display*-Verfahren ausgewählt. Eine bindende PH-Domäne ließe sich in der Krebstherapie und -diagnostik anwenden. Gegen CD3 ϵ gibt es bereits ein bindendes scF_V-Fragment (Mack *et al.*, 1995).
- Das MOMP-Fragment („*major outer membrane protein*“ von *Legionella pneumophila*) besitzt eine Größe von 25 kD. Gegen diese Protein-Domäne wurde schon über *phage-display* mit einer randomisierten hpSTI-Mutante (hpsti: „*human pancreatic secretory trypsin inhibitor*“) erfolgreich eine Bindung nachgewiesen und charakterisiert (Decker *et al.*, 2000: s. Kapitel 1.4.2.). Mit einer bindenden Mutante könnten sich *Legionella pneumophila* im Trinkwasser nachweisen lassen.
- Kreatinin ist eine niedermolekulare Substanz („Hapten“), die im menschlichen Körper als Phosphat- und Energiedepot dient. Über den mit einer bindenden Mutante bestimmbaren Kreatiningehalt ließe sich eine ärztliche Differentialdiagnose am Patienten durchführen. Kreatinin wurde über einen Linker an BSA gekoppelt und so als Zielsubstanz eingesetzt.

4.2.2. Selektionsergebnisse nach dem *phage-display*-Verfahren

Über zunehmende Zahl von Selektionsrunden ("panning") gegen Zielsubstanzen sollten sich affine PH-Domänen anreichern. Auch vor der Bestimmung von Gleichgewichts- und Geschwindigkeitsbindungskonstanten einzelner Mutanten zu Zielsubstanzen kann präferentiell der Verlauf des *phage-display*-Verfahrens über die Selektionsrunden mitverfolgt werden:

Affinitätserhöhung über mehrere Selektionsrunden

Beim ideal-verlaufenden *phage-display*-Verfahren werden in der ersten Selektionsrunde aus allen Phagen mit den randomisierten PH-Domänen die nicht an der Zielsubstanz bindenden Phagen entfernt, während die bindenden z.B. mit Triethylamin eluiert und nur diese nach Amplifikation in einer weiteren Selektionsrunde gegen die Zielsubstanz eingesetzt werden (s. Kapitel 3.9.4.). Damit erhöht sich die durchschnittliche Affinität der Bibliothek von PH-Domänen über jede Selektionsrunde, was sich im ansteigenden Quotienten der Phagentiter von Eluat zu letzter Waschlösung über die zunehmende Zahl von Selektionsrunden ausdrückt; dieser Quotient gibt an, wie gut die Phagen noch durch Waschen von der Zielsubstanz gelöst werden. Ein erheblicher Anstieg dieses Quotienten über die Selektionsrunden konnte bei allen drei Zielsubstanzen beobachtet werden (s. Abb. 4.11).

Mit fortschreitenden Selektionsrunden wurde stringenter gewaschen, indem die Konzentration an Detergenz (Tween-20) von anfänglich 0.1 % bis auf 1 % und die Zahl der Waschschrte erhöht wurde. Jede Selektionsrunde wurde solange wiederholt, bis Waschbedingungen gefunden waren, nach denen jeweils 10^3 bis 10^5 infektiöse Phagen mit Triethylamin eluiert werden konnten (s. Kapitel 3.9.4.). Erst diese eluierten Phagen wurden nach Amplifikation in einer weiteren Selektionsrunde eingesetzt.

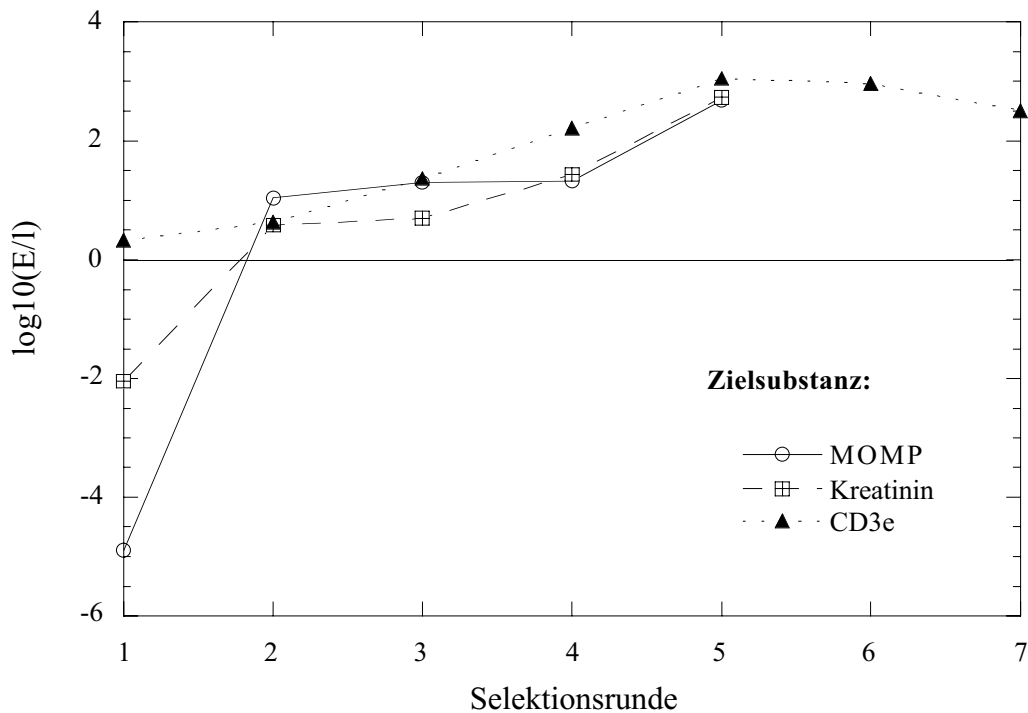


Abb. 4.11: Anreicherung von affinen Phagen über zunehmende Selektionsrunden.

Der Quotient der Phagentiter vom Eluat (E) zu dem der letzten Waschlösung (I) nimmt von Selektionsrunde zu Selektionsrunde zu (halblogarithmische Darstellung). Nach fünf Selektionsrunden lassen sich bei allen drei Zielsubstanzen ca. 1000-mal mehr Phagen im Eluat als in der vorherigen (letzten) Waschlösung nachweisen.

Charakterisierung der Bibliothek durch Restriktionsanalyse

Zur Charakterisierung der Bibliotheken nach jeder Selektionsrunde wurde die Phagemid-DNA der gesamten eluierten Phagen präpariert und mit den Restriktionsenzymen *XbaI* und *NotI* behandelt. Die *XbaI*-Schnittstelle befindet sich auf dem Plasmid *ploPHmg3* im Promotorbereich der PH-Domäne und die *NotI*-Schnittstelle inmitten der DNA-Sequenz, die für den Linker zwischen PH-Domäne und g3-Protein kodiert. Nach Restriktion mit *XbaI* und *NotI* sollte deshalb unabhängig von der Randomisierung immer ein 447 bp-Fragment erhalten werden, sofern die jeweilige Bibliothek intakte PH-Domänen-g3-Fusionen (in detektierbarem Ausmaß) enthält. Bei den Zielsubstanzen CD3e und Kreatinin wurde nach der vierten Selektionsrunde kein entsprechendes Fragment mehr beobachtet, während bei der Selektion gegen MOMP eine Bande der Fragmentgröße 447 bp erhalten blieb (s. Abb. 4.12). Bei allen drei Zielsubstanzen entstanden mit steigender Zahl von Selektionsrunden nach Restriktionsverdau weitere Banden mit kleineren Fragmentgrößen. Sie weisen

auf Mutanten mit deletierten PH-Domänen hin, die im folgenden als Deletionsmutanten bezeichnet werden. Diese nahmen bei allen Zielsubstanzen mit zunehmender Zahl an Selektionsrunden zu, während die Zahl der Phagen mit Genen für PH-Domänen abnahm.

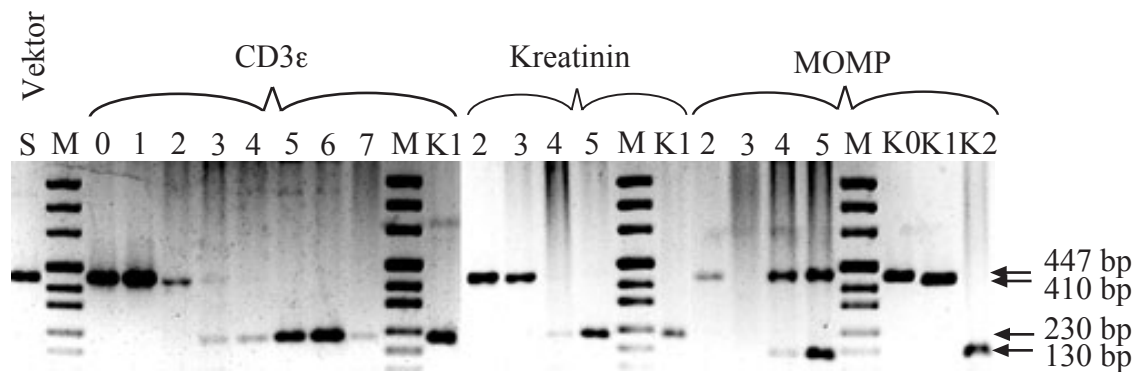


Abb. 4.12: Restriktionsanalyse der Phagemide von Einzelklonen und Bibliotheken eluierter Phagen nach den jeweiligen Selektionsrunden.

Die Phagemide der eluierten Phagen nach den jeweiligen Selektionsrunden 0 bis 7 gegen die unterschiedlichen Zielsubstanzen CD3 ϵ , Kreatinin und MOMP wurden präpariert und mit den Endonukleasen *Xba*I und *Nof*I restringiert. Beim nicht-randomisierten Vektor (S) aber auch nach Randomisierung sollte ein DNA-Fragment von 447 bp im Gel (1.2 % Agarose) zu beobachten sein. Mit Selektionsrunde 0 wird die nicht-selektierte Bank bezeichnet. Die restringierte DNA der ersten Selektionsrunde gegen Kreatinin und MOMP wurde in diesem Gel nicht aufgetragen, zeigte aber in anderen Gelen ausschließlich die korrekte Fragmentgröße.

Bei CD3 ϵ und Kreatinin als Zielsubstanzen konnte nach der vierten Selektionsrunde nicht mehr das erwartete DNA-Fragment von 447 bp nachgewiesen werden, während bei MOMP als Zielsubstanz auch noch nach fünf Selektionsrunden eine Bande der erwarteten Fragmentgröße sichtbar war. Zusätzlich wurde die DNA von Einzelklonen (K0, K1 und K2) nach der jeweils letzten Selektionsrunde präpariert. Bei der Zielsubstanz MOMP wurde einmal die korrekte (447 bp), einmal eine geringfügig verkleinerte (ca. 410 bp) und einmal eine erheblich verkleinerte Fragmentgröße (130 bp) nach Restriktionsverdau beobachtet, während bei CD3 ϵ und Kreatinin jeweils nur eine Deletionsmutante mit 230 bp Fragmentgröße nachgewiesen wurde.

Diesem Ergebnis zufolge werden (randomisierte) PH-Domänen bei allen Zielsubstanzen mit fortschreitenden Selektionsrunden im *phage-display*-Verfahren abgereichert.

Sequenzanalyse der angereicherten Mutanten

Nach der fünften Selektionsrunde konnte nur gegen die Zielsubstanz MOMP in der Restriktionsanalyse der gemeinsam eluierten Phagen neben der Deletionsbande bei 130 bp auch eine „stabile“ Bande mit korrekter Fragmentgröße beobachtet werden (Abb. 4.12). Zur Analyse von Einzelklonen wurde diese stabile Bande kloniert, die DNA-Sequenz von acht Einzelklonen bestimmt und in-silicio translatiert. Die Klone 1 bis 6 sind identisch und alle Klone (ein-

schließlich 7 und 8) weisen ein *amber*-(*stop*)-Codon im randomisierten Bereich des ersten Loops an Aminosäureposition 11 auf. Dieses *amber*-Codon kodiert für Glutamin, da für die Amplifikation der Phagenbibliotheken TG 1 *E. coli*-Zellen verwendet wurden und diese *supE*-supprimiert sind (s. Kapitel 2.1.1.). Die Klone 1 bis 6, 7 und 8 stimmen bis auf das *amber*-Codon nicht in den randomisierten Aminosäuren der Loopregionen miteinander überein (s. Tab. 4.3). In Klon 7 sind zusätzlich sechs Aminosäuren des Loop 3 deletiert.

Das Ziel des *phage-display*-Verfahrens bestand in der Selektion und anschließenden Charakterisierung von bindenden Mutanten aus der in dieser Arbeit klonierten Bibliothek randomisierter PH-Domänen. Dieses Ziel konnte für keine der Zielsubstanzen MOMP, CD3 ϵ und Kreatinin befriedigend erreicht werden. Einzig bei MOMP wurden mit TG 1 *E. coli*-Zellen Mutanten der PH-Domäne angereichert, die allesamt ein *amber*-Codon aufweisen. Da dies die einzige Übereinstimmung bei drei unterschiedlichen, selektierten Klonen ist, beruht die Selektion entweder auf der Bevorzugung eines Glutamin (bei *supE*-Stämmen) oder aber auf der Bevorzugung eines häufigen Translationsabbruchs an diesem Codon (s. Kapitel 5.3.)

Tab. 4.3: Aminosäuresequenzen von Einzelklonen nach der fünften Selektionsrunde gegen die Zielsubstanz MOMP.

		Loop 1			Loop 3	
K.1-6	-----	*-----m--	-h-----	-----	-v-----	-----
K.7	-----	*-s-s-rp--	-----	-----	d-e d-s--	-----
K.8	-----	*s-s--t--	-----	-----	a-----	-----
WT	NPDREGWLLK	LGGGRVKTWK	RRWFILTDNA	LYYFEYTTDK	EPRGIIPLEN	LSIREVEDSK
			Loop 6			
K.6	-----	-----	--gsh-rrap	a-rk-----	-----	-----
K.7	-----	-----	-----r-----	-----	-----
K.8	-----	--r-----r--	-----	--ga-a-k	rh-a-----	-----
WT	KPNVFELYIP	DNKDQVIKAA	KTEADGQVVE	GNHTVYRISA	PTPEEKEEWI	KAIKAAISRD ...

Nach der fünften Selektionsrunde gegen MOMP wurde die Bibliothek (Phagemide) *Xba*I und *Not*I restringiert und das erwartete Fragment mit einer Größe von 447 bp kloniert. Die DNA-Sequenz von acht Einzelklonen zeigt sechs identische Klone (K.1-6). Nach in silicio-Translation ergibt sich für alle sequenzierten Klone ein *amber*-(*stop*)-Codon (*) im randomisierten Loop 1 an Position 11, während die sonstigen randomisierten Aminosäurepositionen nicht übereinstimmen. Klon K.7 weist zudem sechs deletierte Aminosäuren in Loop 6 auf.

Bei allen untersuchten Zielsubstanzen führte die Selektion über das *phage-display*-Verfahren zwar zu einer Anreicherung affiner Phagen (s. Abb. 4.11), doch nahm gleichzeitig die Zahl der korrekt randomisierten PH-Domänen im Zuge fortgesetzter Selektionsrunden zugunsten von Deletionsmutanten ab (s. Abb. 4.12). Mit der Bibliothek randomisierter PH-Domänen wurde eine Vielzahl experimenteller Variationen des *phage-display*-Verfahrens untersucht, wobei insbesondere die Elutionstechnik variiert wurde (s. Kapitel 3.9.4.). Die entsprechenden Ergebnisse unterschieden sich allerdings nicht deutlich von den in dieser Arbeit vorgestellten, weshalb das *phage-display*-Verfahren zur Selektion bindender PH-Domänen eingestellt wurde.

Die präferentielle Anreicherung von Deletionsmutanten bzw. Abreicherung von Punktmutanten der PH-Domäne im *phage-display*-Verfahren kann systemimmanent sein. Vor der Phagenassemblierung werden nämlich die (g3-fusionierten) PH-Domänen in das Periplasma transloziert und bleiben über das g3-Protein an der inneren Membran verankert. Die Wildtyp PH-Domäne selbst läßt sich aber nur sehr schlecht in das Periplasma translozieren. Dies veranlaßte einerseits dazu, nach translokationsoptimierten Mutanten der PH-Domäne zu suchen, wovon Kapitel 3.2. handelt, und andererseits dazu, die Bibliothek in anderen Selektionsverfahren einzusetzen, was im folgenden Kapitel ausgeführt ist.

4.2.3. Alternative Selektionsverfahren zum *phage-display*

Während das *phage-display*-Verfahren die Translokation des auf der Phagenoberfläche exprimierten Proteins in das Periplasma impliziert, entfällt jeglicher Transportvorgang, wenn die Bibliothek über ein DNA-bindendes Protein (weitgehend) irreversibel an die dafür kodierende Plasmid-DNA binden kann und sich dieses Protein-Plasmid-Konjugat undissoziiert isolieren läßt. Man kann es dann analog den Phagen im *phage-display*-Verfahren in Selektionsrunden einsetzen. Ein derartiges Verfahren konnte bereits mit dem lac-System erfolgreich durchgeführt werden (Cull *et al.*, 1992). Analog dazu wurde in dieser Arbeit eine Fusion der PH-Domäne mit dem *tet*-Repressor-Protein kloniert und die Bindung zum *tet*-Operator auf dem Plasmid untersucht (Sänger *et al.*, 2000). Dabei konnte nach Zellaufschluß unter nativen Bedingungen aber weder die Plasmid-DNA über die Isolierung des Proteins noch das Fusionsprotein über die Isolierung von Plasmid-DNA gewonnen werden. Eine Klonierung der Bibliothek von PH-Domänen als Fusionsprotein mit dem *tet*-Repressor versprach deshalb keinen Erfolg und weitere Versuche mit diesem *tet*-Fusionsverfahren wurden eingestellt.

Die *in vivo*-Selektion von PH-Domänen in menschlichen T-Zellen (Jurkat LT-Zellen) zur Unterbrechung von Signaltransduktionswegen verzichtet auf jegliche Transportvorgänge oder isolierbare Protein-Plasmid-Konjugate. Die Schwierigkeiten beim *phage-display* oder beim *tet*-Fusionsverfahren (s.o.) treten somit definitiv nicht auf. Ein entsprechendes Selektionsverfahren wird in der Arbeitsgruppe von Waldemar Kolanus etabliert (Genzentrum der LMU München; s. Kapitel 1.5.2.). Dazu wird u.a. eine Bibliothek von $3.2 \cdot 10^6$ plasmidkodierten Klonen eingesetzt, die im Rahmen dieser Arbeit kloniert wurde.

Bereits gut etabliert ist dagegen die Selektion mit dem Hefe-2-Hybrid-Verfahren im Zellkern von *Saccharomyces cerevisiae*, was wiederum einen Transportvorgang voraussetzt. Die Bibliothek von PH-Domänen wurde dazu dem Arbeitskreis von Ralf Baumeister zur Verfügung gestellt. Der gleiche experimentelle Ansatz wird mit einem in dieser Arbeit klonierten Vektorsystem (pJG4-5; s. Kapitel 2.2.1.) in der Arbeitsgruppe von Roger Brent weiterverfolgt (Molecular Science Institute, Berkeley, California; s. Kapitel 1.5.2.).

Ergebnisse aus diesen Untersuchungen stehen noch aus.

4.3. Optimierung der PH-Domäne zur Translokation in das Periplasma von *E. coli*

Eine Ursache für die vergebliche Suche nach bindenden Mutanten der PH-Domänen über das *phage-display*-Verfahren kann in der mäßigen Translokationseffizienz der PH-Domäne in das Periplasma von *E. coli* begründet liegen.

In diesem Zusammenhang wurden unterschiedliche Theorien aufgestellt, welche Eigenschaften die Translokation von Proteinen im allgemeinen und die der PH-Domäne im besonderen limitieren. Dazu wurden die seiner Zeit bekannten Strukturen von PH-Domänen mit dem Computerprogramm LSQMAN überlagert und mit dem Computerprogramm *InsightII* 2.3. visualisiert. Ein anschließender Sequenzvergleich mit der PH-Domäne des Cytohesin-1 erlaubte Rückschlüsse auf deren Struktur (s. Kapitel 4.1.1.). Aus diesen strukturellen Erkenntnissen heraus wurden verschiedene Mutanten hergestellt, die den unterschiedlichen Theorien zufolge verbesserte Translokation zeigen sollten (s. Tab. 4.4). Diese Mutanten wurden hinsichtlich ihrer cytoplasmatischen sowie periplasmatischen Ausbeute, ihrer Stabilität, ihrer Prozessierungsmuster nach periplasmatischer Expression und ihrer Expression auf dem M13-Phagen untersucht.

Tab. 4.4: Aminosäure-Sequenzvergleich der Mutanten zur Translokationsoptimierung sowie der originalen PH-Domänen von Cytohesin-1 und Grp1.

	1	8	21	51	71								
WT	Npdregwllk	lgggrvktwk	rwfildtna	lyyfeytttk	prgiiplen	lsirevedsk	kpnvfelyip	dnk					
orig.wt	-----c	-----	-----c	-----	-----	-----	-----c	-----					
Grp1	-----	-----	-----c	-----	-----	-----pr	-----c	-----n- sh-					
KRR-QNV	N-----	-----	-----q v	-----	-----	-----	-----	-----					
N-Ter	Apk-----	-----	-----	-----	-----	-----	-----	-----					
WTmiLi	Sdqvika-	-----	-----	-----	-----	-----	-----	-----					
L(8)F	N-----f	-----	-----	-----	-----	-----	-----	-----					
L(8)V	N-----v	-----	-----	-----	-----	-----	-----	-----					
T(102)A	N-----	-----	-----	-----	-----	-----	-----	-----					
TEE-AQQ	N-----	-----	-----	-----	-----	-----	-----	-----					
C-Ter	N-----	-----	-----	-----	-----	-----	-----	-----					
C/N-Ter	Glqallkgs-	-----	-----	-----	-----	-----	-----	-----					
CP 1.6.	lqallkgs-	-----	-----	-----	-----	-----	-----	-----agsaaaahhhhh*					
CP 2.6.	-----	-----	-----	-----	-----	-----	-----	-----agsaaaahhhhh*					
CP.ohnLi	-----	-----	-----	-----	-----	-----	-----	-----agsaaaahhhhh*					
CP 3.6.	-----	-----	-----	-----	-----	-----	-----	-----agsaaaahhhhh*					
CP 1.7.	lqallkgs-	-----	-----	-----	-----	-----	-----	-----					
CP 2.7.	-----	-----	-----	-----	-----	-----	-----	-----					
CP 3.7.	-----	-----	-----	-----	-----	-----	-----	-----					
	α Phos	β 1	L1	β 2	L2	β 3	L3	β 4	L4	β 5	L5	β 6	L6
	Termini CP X.6.												
	74	81	101	102	111								
WT	dqvikaak	kteadgqvve	gnhtvyrisa	p	tpeekeewi	kaikaaisrd	gaaahhhhh*						
orig.wt	-----c	-----r	-----	-----	-----c	-----	-----						
Grp1	-----c	-----r	-----v	-----	s-----m	-s--s--	-----						
KRR-QNV	-----	-----	-----	-----	-----	-----	-----*						
N-Ter	-----	-----	-----	-----	-----	-----	-----*						
WTmiLi	-----	-----	-----	-----	-----	-----	-----*						
L(8)F	-----	-----	-----	-----	-----	-----	-----*						
L(8)V	-----	-----	-----	-----	-----	-----	-----*						
T(102)A	-----	-----	-----	-----	a-----	-----	-----*						
TEE-AQQ	-----	-----	-----	-----	a-qq-----	-----	-----*						
C-Ter	-----	-----	-----	-----	a--tsgdnka	aahhhhh*	-----*						
C/N-Ter	-----	-----	-----	-----	a--tsgdnka	aahhhhh*	-----*						
CP 1.6.	S-----	-----	-----	-----	-----a--tsgdnkg	-----	-----						
CP 2.6.	S-----	-----	-----	-----	-----sg	gsgggg	-----						
CP.ohnLi	S-----	A-----	-----	-----	-----sg	gsgggg	-----						
CP 3.6.	S-----	-----	-----	-----	-----sg	gsgggsgggg	-----						
CP 1.7.	-----	-----	-----	-----	T-----a--tsgdnkg	-----	-----						
CP 2.7.	-----	-----	-----	-----	T-----sg	gsgggg	-----						
CP 3.7.	-----	-----	-----	-----	T-----sg	gsgggsgggg	-----						
	β 11	Li1	β 12	Li2	β 7	L7	α 1						
	Termini CP X.7.						Linker (CP)						

Bei der ursprünglichen PH-Domäne des Cytohesin-1 (orig. wt) wurden alle Cystein-Reste sowie Arginin 87 substituiert (WT). Anhand der PH-Domäne von Grp1 (Lietzke *et al.*, 2000) läßt sich die Struktur von WT modellieren (Guex & Peitsch, 1997). Auf WT beziehen sich die Mutanten dieser Untersuchung. Die β (-Faltblätter), α (-Helices) und L(oops) sind gekennzeichnet. Zum WT identische Aminosäuren werden mit "-" bezeichnet; die erste Aminosäure einer jeden Mutante ist mit einem Großbuchstaben dargestellt. Substituierte Aminosäuren (und WT) sind mit Kleinbuchstaben wiedergegeben und der Translationsstop jeder Mutante mit "*" markiert. Die Termini der Circularpermutanten liegen in L6 und L7. Bei ihnen wird die Peptidsequenz am Ende der zweiten Zeile mit der ersten Zeile fortgesetzt.

4.3.1. Austausch spezifischer Sequenzelemente und Destabilisierung der PH-Domäne zur Verbesserung der Translokation?

Die Proteine des allgemeinen sekretorischen Translokationssystems von *E. coli* sind gut charakterisiert und der Einfluß einiger Chaperone auf die Translokation von Präproteinen ist bekannt. Während der Translokation über den allgemeinen sekretorischen Transportmechanismus liegt das Präprotein entfaltet vor. Zuvor befindet sich das Protein in einem „translokationskompetenten Zustand“ (s. Kapitel 1.6.1). Dieser Begriff beschreibt das zu transportierende Protein zwar physiologisch nicht aber strukturell.

Welche physikochemischen Eigenschaften Präproteine auszeichnen, die sich besonders effektiv translozieren lassen, ist bislang nicht geklärt, und wie ein Protein verändert werden sollte, um es besser zu translozieren, ist bislang nicht systematisch erfaßt.

Unter der Annahme, die PH-Domäne läge vor der Translokation entfaltet vor, sollte nur die Aminosäuresequenz den Transportvorgang beeinflussen. Deshalb wurden Mutanten kloniert, bei denen spezifische Sequenzelemente ausgetauscht sind:

- Bei Membranproteinen dienen Cluster positiv geladener Aminosäuren häufig als *stop*-Transfersignal (von Heijne & Gavel, 1988). Die Wildtyp PH-Domäne besitzt ein derartiges Cluster KRR(20,21,22) im β 2-Faltblatt, was in der Mutante KRR(20,21,22)QNV gegen ungeladene Aminosäuren ausgewechselt wurde.
- Die Aminosäuren hinter dem Signalpeptid wurden durch den N-Terminus des originalen reifen *ompA*-Proteins („*outer membrane protein A*“ von *Enterobacter aerogenes*: Braun & Cole, 1983) ersetzt zu NP(1,2)APT (diese Mutante heißt N-ter). Ein ähnlicher Ansatz ermöglichte die periplasmatische Expression der *chicken muscle triphosphate isomerase* in *E. coli* (Summers & Knowles, 1989).

Liegt dagegen die Wildtyp PH-Domäne als Präprotein zuerst gefaltet im Cytoplasma vor, muß sie zur Translokation entfaltet werden. In ihrer Faltung destabilisierte Mutanten könnten sich unter diesen Umständen besser in das Periplasma translozieren lassen als der Wildtyp. Deshalb wurden destabilisierende Mutationen in die Wildtyp PH-Domäne eingeführt:

- Das hydrophobe Proteininnere befindet sich in einem geordneten Zustand geringer Entropie, der Proteinen maßgeblich ihre Stabilität verleiht („hydrophober Effekt“). Durch Austausch eines in das Proteininnere

orientierten Leucinrests gegen das sterisch kleinere Valin [in der Mutante L(8)V] sollte es in einen weniger geordneten Zustand versetzt und damit destabilisiert werden, wie dies auch beim T4-Lysozym beobachtet wurde (Xu *et al.*, 1998). Die Mutation gegen das sterisch voluminösere Phenylalanin [in der Mutante L(8)F] sollte dagegen zu einer sterischen Abstoßung im dichtgepackten Proteininneren führen, was ebenfalls bei T4-Lysozym zu einer Destabilisierung führte (Eriksson *et al.*, 1993).

- Helices weisen häufig eine charakteristische Verteilung von Aminosäuren auf, die dieses sekundäre Strukturelement stabilisieren. So befindet sich auch im helicalen N-Terminus der Wildtyp PH-Domäne eine Wasserstoffbrückenbindung zwischen der Hydroxylgruppe des T(102) und der Carboxylatgruppe des E(104) („N-cap-Struktur“). Mit der Mutation T(102)A wird sie unterbunden. Beim menschlichen Wachstumshormon führte eine vergleichbare Mutation zu einer Destabilisierung von 4 kJ/mol (Zhukovsky *et al.*, 1994).
- Eine weitere Stabilisierung von Helices wird durch geladene Aminosäuren erreicht: durch die gleichorientierte Ausrichtung der Peptidbindungen in einer α -Helix entsteht ein Dipol, wobei der N-Terminus der α -Helix positiv und der C-Terminus negativ polarisiert ist. Dieser Dipol wird in der Natur durch entgegengesetzt geladene Aminosäuren ausgeglichen. Die Mutation saurer Aminosäurereste aus dem helicalen N-Terminus und basischer Aminosäurereste aus dem C-Terminus führen somit zu Destabilisierungen, wie dies bei Protein G gezeigt werden konnte (Strop *et al.*, 2000). Deshalb wurden in der Mutante TEE(102,104,105)AQQ neben der Mutation T(102)A noch die zusätzlichen Mutationen E(104)Q und E(105)Q verwirklicht. Durch Substitution einer positiven (K114) und Einführung einer negativen Ladung (D117) im helicalen C-Terminus sollte bei der Mutante C-ter ebenfalls der Helixdipol geschwächt werden. Die zusätzliche N-terminale Einführung einer weiteren α -Helix (aus der Phospholipase C: PDB: 1MAI; Ferguson *et al.*, 1995) sollte keinen zusätzlichen Einfluß auf die Gesamtstabilität der Mutante C/N-ter gegenüber C-ter nehmen, dafür aber die Sequenz des N-Terminus zusätzlich verändern (als Änderung sequenzspezifischer Elemente).

Die Aminosäuresequenzen der in diesem Kapitel aufgeführten Mutanten sind in Tab. 4.4 zusammengefaßt.

Änderung der Topologie durch Circularpermutation zur verbesserten Translokation?

Die Abhängigkeit der Translokation von der Abfolge der sekundären Strukturelemente unter Beibehaltung der globalen Proteinstruktur wurde in Circularpermutanten (CP) untersucht (s. Abb. 4.13). Dazu wurde in verschiedenen Varianten der N- und der C-Terminus der Wildtyp PH-Domäne mit drei unterschiedlichen Linkern neu fusioniert und dafür an zwei anderen Positionen neue N- bzw. C-Termini in die PH-Domäne eingeführt (s. Abb. 4.14).

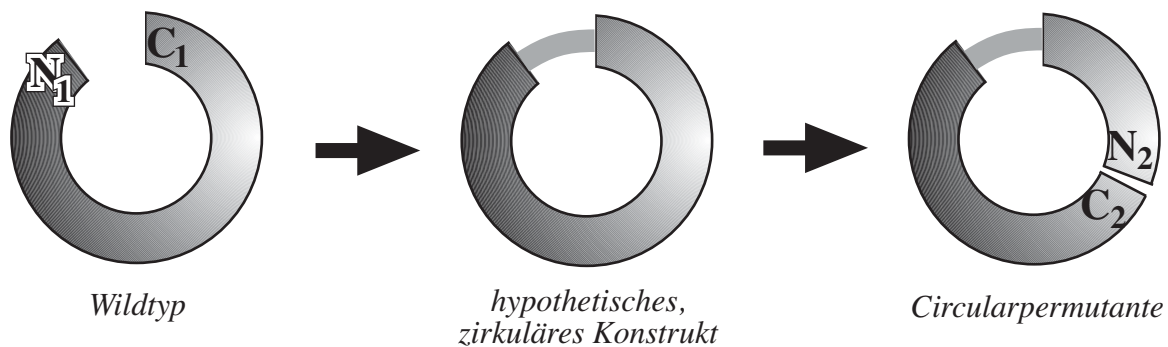


Abb. 4.13: Schematische Erläuterung einer Circularpermutante.

Die Proteine sind als Ringe dargestellt. Die Ringform des Modells bleibt erhalten, auch wenn die ursprünglichen Termini N₁ und C₁ über einen Linker miteinander verbunden und neue Termini N₂ und C₂ eingeführt werden.

Circularpermutanten sind für die Untersuchung bzw. Optimierung der Translokation von Präproteinen in das Periplasma von *E. coli* sehr interessant - unabhängig davon, ob Präproteine vor der Translokation in einem gefalteten oder ungefalteten Zustand vorliegen: Liegen sie ungefaltet vor, d.h. ist nur die Aminosäuresequenz für die Translokation von Bedeutung, sind sie interessante Versuchsobjekte, weil sie verglichen mit dem Wildtyp andere Termini besitzen. Gerade dem N-Terminus kann für die Translokation eine besondere Rolle zufallen: so zeigte eine statistische Auswertung periplasmatischer Proteine bei *gram*-negativen Bakterien für die ersten 18 Aminosäuren hinter der Prozessierungsstelle eine Häufung von negativ geladenen und neutralen Aminosäuren (Kajava *et al.*, 2000). Die neu eingeführten Termini bei Circularpermutanten könnten somit zu völlig neuen Translokationseigenschaften führen.

Liegen dagegen Präproteine vor der Translokation gefaltet vor, sind die Circularpermutanten um so interessantere Versuchsobjekte, weil die Entfaltung

an der Membran über den allgemeinen sekretorischen Transportmechanismus anderen physikalischen Eigenschaften gehorchen könnte, als dies bei der Bestimmung der thermodynamischen Stabilität von Proteinen *in cuvette* mit Denaturierungsmittel der Fall ist (vgl. Kapitel 4.1.5.).

Circularpermutanten: Nomenklatur und Konstruktion

Von der PH-Domäne des Cytohesin-1 wurden sechs Circularpermutanten (CP) hergestellt, die sich in der Art des Linkers und in der Position der neuen N- und C-Termini voneinander unterscheiden (Aminosäuresequenz und Sekundärstrukturelemente: s. Tab. 4.4). Die Circularpermutante CP X.Y. ist folgendermaßen charakterisiert (s. Abb. 4.14):

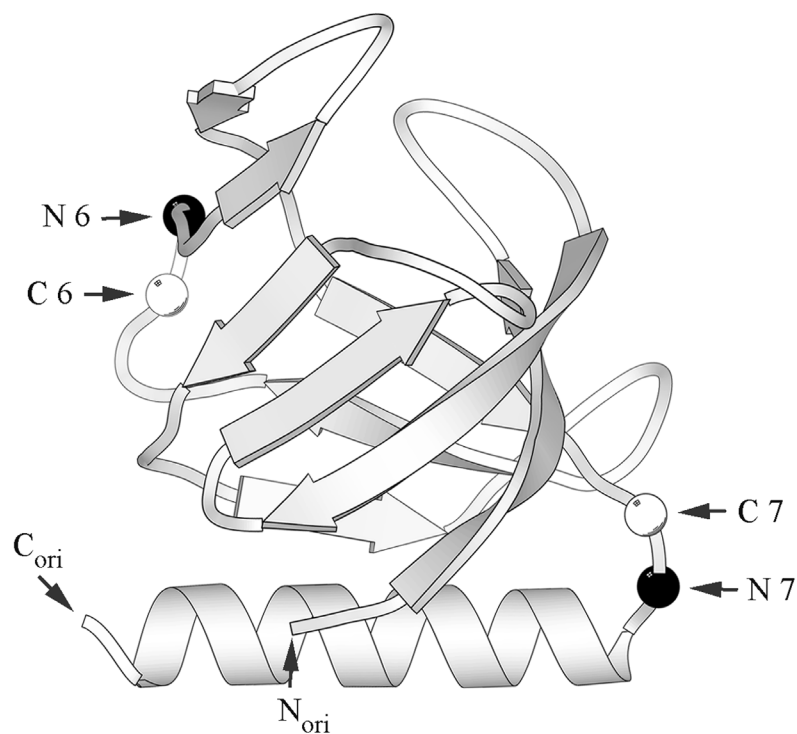


Abb. 4.14: Struktureller Aufbau der circularpermutierten PH-Domänen.

Die originalen N_{ori} - und C_{ori} -Termini werden mit unterschiedlichen Linkern fusioniert und stattdessen die neuen Termini N 6 und C 6 in Loop 6 bzw. N 7 und C 7 vor der α -Helix eingeführt. Diese Proteinstruktur wurde nach der homologen PH-Domäne des Grp1 [PDB: 1FGZ (Lietzke *et al.*, 2000)] modelliert (SWISS-Modell nach Guex & Peitsch, 1997).

für

- $X = 1$ besitzt sie einen Linker zwischen den originalen Termini, der sich aus der zusätzlichen α -Helix der Phospholipase C und einem Peptid zusammensetzt, das mit Hilfe des Computerprogramm "One2all-gap" ermittelt wurde (s. Kapitel 4.3.4.),
- $X = 2$ besitzt sie einen kurzen, unstrukturierten Linker und
- $X = 3$ besitzt sie einen langen, unstrukturierten Linker.

Mit dem Y ist die Position der neuen N- und C-Termini festgelegt:

- $Y = 6$ bedeutet neue N- und C-Termini in Loop 6 und
- $Y = 7$ bedeutet neue N- und C-Termini in Loop 7, also unmittelbar vor der α -Helix der Wildtyp PH-Domäne.

Zur Konstruktion einer Circularpermutante wird die Wildtyp PH-Domäne hypothetisch zweigeteilt: der ursprünglich N-terminale Teil wird mit der restlichen PH-Domäne vertauscht und ein Linker zwischen beide Fragmente der PH-Domäne eingeführt.

Die Circularpermutante CP 2.6. wurde kloniert, indem aus dem Vektor ptoPH (Wildtyp PH-Domäne mit ompA^M-Signalpeptid) zunächst über getrennte PCR-Reaktionen zwei vektorielle Zwischenprodukte hergestellt wurden (s. Abb. 4.15):

- der Vektor „Produkt 4“ besitzt hinter dem Signalpeptiden die Basen, die für den C-terminalen Teil der Wildtyp PH-Domäne kodieren und
- der Vektor „Produkt 2“ besitzt hinter dem Signalpeptidgen die regulären Basen der Wildtyp PH-Domäne, die jedoch in Loop 6 mit einem His-tag und *stop*-Codon enden.

In Produkt 2 wird über eine drei-Fragment-Ligation zwischen Signalpeptid und N-Terminus der ursprüngliche C-Terminus aus Produkt 4 und dahinter ein Linkerkonstrukt kloniert. Das Linkerkonstrukt wird über eine Gensynthese hergestellt (s. Abb. 4.16).

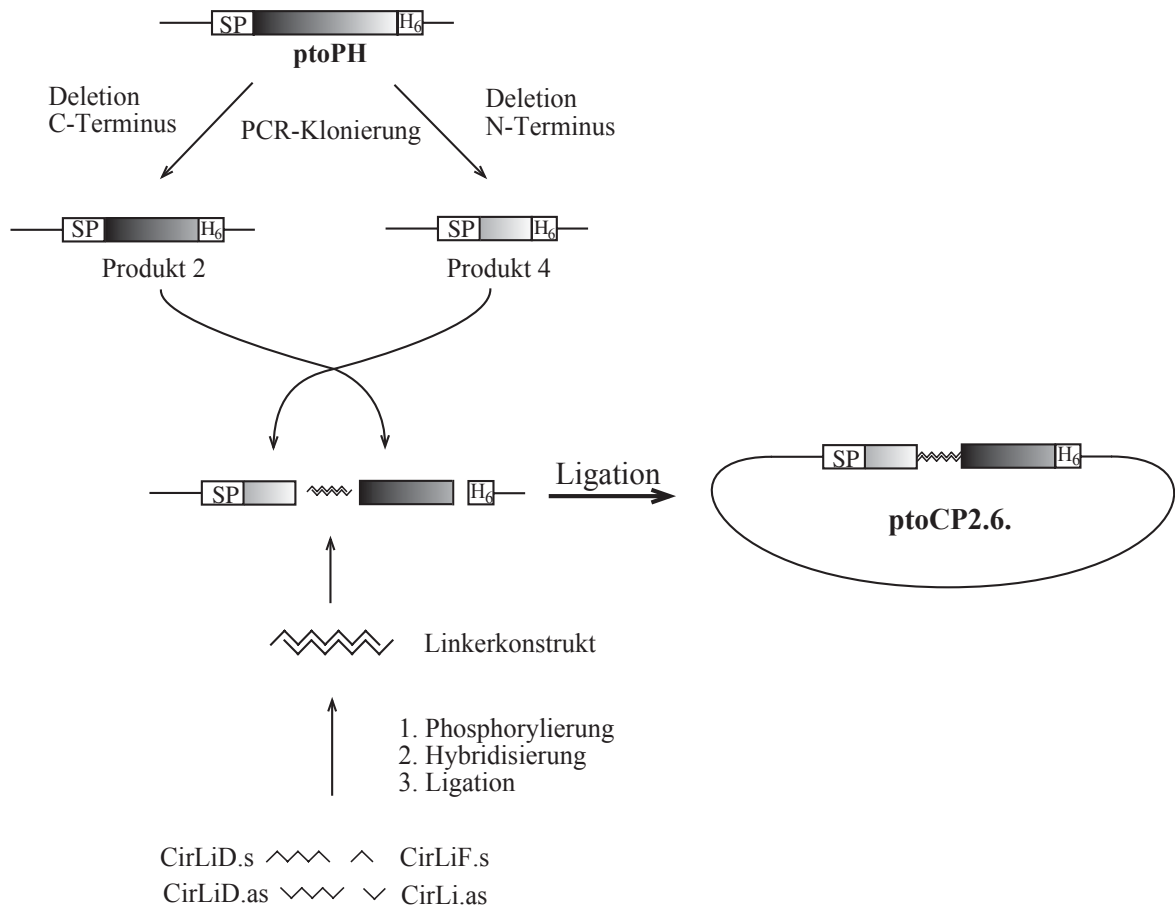


Abb. 4.15: Schema zur genetischen Konstruktion von Circularpermutanten.

"SP" ist das modifizierte Signalpeptid des "outer membrane protein A" (*E. coli*). "H₆" ist C-terminale His₆-tag. ptoPH ist Ausgangs- und ptoCP2.6. Endvektor zur jeweiligen periplasmatischen Expression der Konstrukte. CirLi... sind Oligodesoxynukleotide zur Linkersynthese nach Abb. 4.16.

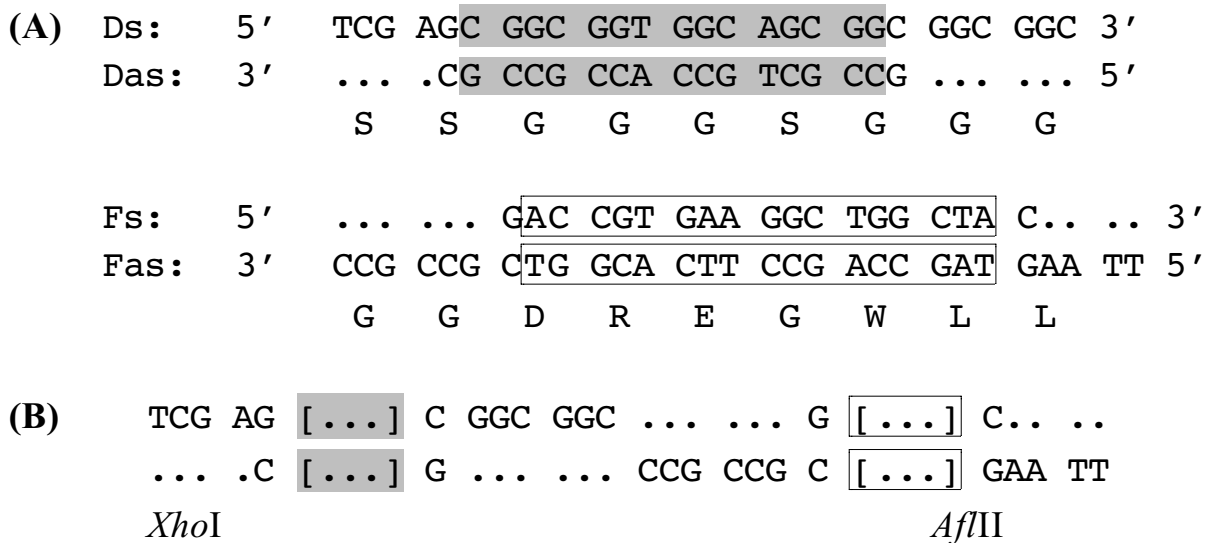


Abb. 4.16: (A) getrennte Hybridisierung zweier Oligodesoxynukleotide Ds/Das und Fs/Fas zu dsDNA und (B) anschließende Ligation zum vollständigen, klonierbaren Linker für die Circularpermutante CP 2.6..

Die Oligodesoxynukleotide wurden phosphoryliert und dann jeweils paarweise à 10 pmol hybridisiert. Je 1/10 beider Ansätze wurden mit 10 ng Produkt 2-Fragment (*AflIII/HindIII* restringiert) und 100 ng Produkt 4-Vektor (*XhoI/HindIII* restringiert) in einer Drei-Fragment-Ligation eingesetzt, um die Circularpermutante CP 2.6. zu klonieren. Bei den anderen Circularpermutanten wurde analog verfahren.

4.3.2. Modellierung eines strukturierten Linkers für Circularpermutanten

Der Faltungsweg von Circularpermutanten unterscheidet sich von dem des Wildtyps, weshalb die gleiche Struktur am Ende des Faltungsweges nicht selbstverständlich ist. Untersuchungen dazu gibt es an der SH3-Domäne des α -Spectrin (Viguera *et al.*, 1996). Ein strukturierter Linker zwischen den originalen Termini könnte den Faltungsweg "ebenen", indem er zur Struktur- bildung des gesamten circularpermutierten Proteins beiträgt.

Der Konstruktion eines strukturierten Linkers lag erneut der Strukturvergleich von PH-Domänen zugrunde (s. Tab. 1.1). Der Sequenzvergleich mit der PH-Domäne des Cytohesin-1 ermöglichte es, Raumkoordinaten für den N- und C-Terminus dieser PH-Domäne festzulegen. Beim Strukturvergleich der PH-Domänen fiel die PH-Domäne der Phospholipase C δ -1 auf, die N-terminal

eine weitere α -Helix besitzt (PLC δ PDB: 1MAI; Ferguson *et al.*, 1995). Der Abstand vom N-Terminus der zusätzlichen bis zum C-Terminus der C-terminalen Helix beträgt bei dieser PH-Domäne lediglich 13.2 Å zwischen den C $_{\alpha}$ -Atomen, die Entfernung des ersten Aminosäurerestes vom ersten β -Faltblatt bis zum Ende der C-terminalen α -Helix dagegen ca. 17.2 Å, was mit Hilfe des Computerprogramms InsightII 2.3. ermittelt werden konnte (s. Abb. 4.17). Durch Fusion der zusätzlichen α -Helix mit der PH-Domäne des Cytohesin-1 sollten sich der N- und C-Terminus auf 13.2 Å nähern und sollten sich die Strukturdaten von der PH-Domäne der Phospholipase C auf die der PH-Domäne des Cytohesin-1 übertragen lassen.

Ferner findet zwischen beiden Helices der Phospholipase C eine elektrostatische Wechselwirkung statt: der Abstand zwischen der Carboxylatgruppe des Asp (18) beträgt zum N2-Atom des His (120) 4.11 Å. Dieselbe Carboxylatgruppe des Aspartats interagiert zudem ionisch mit der ϵ -Aminofunktion von Lys (127), wobei der Abstand in der Kristallstruktur 3.52 Å beträgt (s. Abb. 4.17). In der Literatur werden die Abstände von interagierenden Ionenpaaren mit ungefähr 4 Å angegeben (Barlow & Thornton, 1983). Sie tragen zur Faltungstabilität von Proteinen bei und könnten möglicherweise auch eine circularpermutierte PH-Domäne stabilisieren. Aufgrund sterischer Überlappungen mit dem weiteren Linkerkonstrukt (s.u.) wurde allerdings schließlich auf die Übertragung dieser ionischen Wechselwirkungen auf die PH-Domäne des Cytohesin-1 verzichtet (s. Tab. 4.6).

Grundsätzlich wurde für die Modellierung eines strukturierten Linkers angenommen, daß die PH-Domänen des Cytohesin-1 und der Phospholipase C im Bereich des β 1-Faltblattes und der α 1-Helix im Peptidgerüst übereinstimmen. Ein konstruierter Linker zwischen den beiden Helices der Phospholipase C sollte sich dann zusammen mit der α (Phos)-Helix auf die PH-Domäne des Cytohesin-1 übertragen lassen und könnte dann zur Stabilisierung in den Circularpermutante CP 1.X. beitragen.

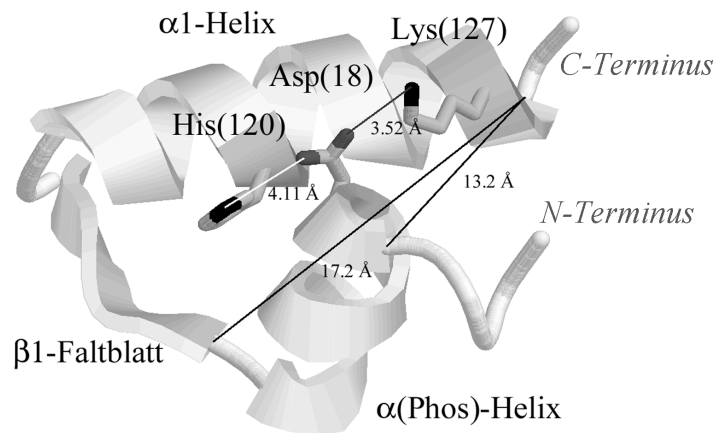


Abb. 4.17: Die ionischen Wechselwirkungen und der Abstand von N- und C-Terminus bei der PH-Domäne der Phospholipase C δ -1 (PDB: 1MAI).

Asp(18) der N-terminalen α (Phos)-Helix interagiert mit His(120) und Lys(127) der C-terminalen α 1-Helix. Der N-Terminus des β 1-Faltblattes ist zum C-Terminus der C-terminalen α 1-Helix 17.2 Å entfernt, während der Abstand vom N-Terminus der α (Phos)-Helix zum C-Terminus der α 1-Helix lediglich 13.2 Å beträgt. Die Übertragung der α (Phos)-Helix auf die PH-Domäne des Cytohesin-1 gestattet die Festlegung von Raumkoordinaten für die PH-Domäne des Cytohesin-1 zur Konstruktion eines Linkers zwischen beiden Termini bei den Circularpermutanten CP 1.X. mit konstruiertem Linker. Dabei nähern sich beide Termini einander auf 13.2 Å an. Die ionischen Wechselwirkungen beider Helices ließen sich nicht für die Circularpermutanten CP 1.X. auf die PH-Domäne des Cytohesin-1 übertragen. Auf die Darstellung der weiteren Strukturelemente zwischen dem β 1-Faltblatt und der α 1-Helix wurde in dieser Abbildung verzichtet.

Für die Suche nach einer Peptidsequenz, die die Lücke zwischen beiden α -Helices der Phospholipase C schließt, wurde das Computerprogramm „one2all gap“ verwendet (ein in der Arbeitsgruppe entwickeltes, unveröffentlichtes Struktur-Vergleichs-Programm). Mit dessen Hilfe wird eine Strukturdatenbank nach Peptiden aus bevorzugt nicht-homologen Proteinen durchsucht, deren jeweils vier ersten und letzten Aminosäuren sich möglichst gut mit den Koordinaten der vier Aminosäuren vor und hinter der Lücke („Ankersequenzen“) überlagern lassen (geringer rmsd-Wert: *root mean square deviation*; gibt das Maß der Abweichung nach Überlagerung an). Die Peptidlänge wird bei dieser Recherche jeweils vorgegeben; sie setzt sich aus den vier ersten und letzten Aminosäuren, sowie dem verbindenden Peptid mit p Aminosäuren zusammen (Die Gesamtlänge der Peptide beträgt demnach $2 \cdot 4 + p$ Aminosäuren). Überlagert man die Peptide einer Länge und stellt sie gemeinsam mit dem Computerprogramm *InsightII 2.3* dar, findet man für jeweils vorgegebene

Peptidlängen ein unterschiedliches Maß an struktureller Übereinstimmung (*clustering*). Dies spiegelt den Sachverhalt wieder, daß sich in der Evolution bestimmte Peptidlängen und -strukturen für die Strukturbildung von Proteinen besonders bewährt haben und deshalb häufiger in Proteinstrukturen angetroffen werden als andere. Bei der Wahl der beiden Ankersequenzen in der Phospholipase C konvergieren gerade 14-er Peptide strukturell am stärksten untereinander und von diesen zeigt das Peptid der β -Lactamase (PDB: 1MBL) ab der Aminosäure D(124) bis zu L(137) die höchste strukturelle Übereinstimmung mit den beiden Ankersequenzen der Phospholipase C: sie beträgt $rmsd = 0.72 \text{ \AA}$. In Tab. 4.5 sind die 20 Aminosäuresequenzen der strukturell ähnlichsten 14-er Peptide angegeben. Die jeweiligen rmsd-Werte beziehen sich dabei auf die Überlagerung mit der β -Lactamase. Die dazugehörigen Proteine sind zwar teilweise homolog, doch unterscheiden sich ihre Aminosäuresequenzen in den ausgewählten Peptidsequenzen deutlich voneinander.

Der Aminosäurevergleich an jeder Position in Tab. 4.5 zeigt, daß nicht nur die Strukturen der 14-er Peptide, sondern auch die jeweiligen Aminosäuren an jeder Position konvergieren. Statistisch ausgewertet ist dies in Tab. 4.6A, woraus sich eine Consensus-Sequenz ableiten läßt.

Mit der Wahl einer konvergenten Linkerlänge und einer konvergenten Zusammensetzung der Aminosäurereste an jeder Position erhält man eine thermodynamisch stabilere Linkerkonstruktion, als wenn man die Linkerlänge und insbesondere die Reihenfolge der Aminosäuren willkürlich wählt, wie dies an einer V_L -Domäne gezeigt werden konnte (Niggemann & Steipe, 2000). Deshalb wurde die ermittelte Consensus-Sequenz für das Linker-Hexapeptid unter zusätzlicher Berücksichtigung sterischer Einflüsse als Linkerkonstruktion übernommen und mit der α (Phos)-Helix auf die PH-Domäne des Cytohesin-1 übertragen. Bei der Linkerkonstruktion mit Hilfe von „*one2all gap*“, wie sie hier durchgeführt wurde, entfiel der Einbau der gewünschten ionischen Wechselwirkung zwischen beiden Helices. Der Aminosäurevergleich der entscheidenden Peptidsequenzen ist in Tab. 4.6B aufgeführt.

Die Konstruktion des Linkers erfolgte nach logischen und statistischen Auswahlkriterien. Nur im günstigsten Fall kann man allerdings davon ausgehen, daß das entworfene Hexapeptid tatsächlich die Struktur der β -Lactamase ab R(129) besitzt. In jedem Fall ist aber die konstruierte Sequenz mit der erwünschten Struktur lokal kompatibel.

Tab. 4.5: Die 20 Peptidsequenzen mit der höchsten strukturellen Übereinstimmungen zur Peptidsequenz der β -Lactamase (PDB: 1MBL) ab Aminosäure 124.

PDB#	AS#	0123	456789	0123	rmsd/Å	Proteinfunktion
1MBL	124	DASL	RYSDNA	AQNL	0.00	β -Lactamase
1ALQ	124	EASM	TYSNT	ANNK	0.15	circ.permut. β -Lactamase
1BTL	124	SAAI	TMSDNT	AANL	0.21	β -Lactamase
1CEM	288	VDYS	WFGDQR	AKAN	0.30	circ.permut. β -Lactamase
TF4	270	LLAK	ETGKQK	YIDD	0.74	Cellulase
1CLC	369	EMWE	TLGDEE	YLRD	0.80	Cellulase
1JFR	227	LKRF	IDS DTR	YEQF	0.83	β -Glucanase
1TF4	222	WLYK	ATGDDS	YLAK	0.89	Cellulase
1CLC	472	VANK	ISPNND	YVNA	0.96	Cellulase
1KFS	713	IMAH	LSRDKG	LLTA	1.08	DNA-Exonuclease
1NSG	2086	RKYM	KSGNVK	DLTQ	1.14	Immunosupressor
1CEM	223	VYAQ	YTG DTR	WNQV	1.20	Cellulase
2MHR	33	FDCI	RDNSAP	NLAT	1.31	Myohemerythrin
5CSM	177	TKLI	KSKDVE	GIMK	1.32	Chorismat Mutase
1NSG	2036	RLYF	GERNVK	GMFE	1.33	Immunosupressor
2A0	699	ESNL	TAQDKK	GIVE	1.34	Arc B Phosphotransferase
1IIB	81	SLLY	GKVDGL	GVLK	1.35	Phosphotransferase
1CPQ	25	TGVA	KAFDAE	AAKV	1.37	Cytochrom c
1PUD	332	HHLI	RAGEIL	GAML	1.41	tRNA-G Transglycosylase
1YCS	365	HNAV	CAGHTE	IVKF	1.42	P53BP2

Der rmsd-Wert (*root mean square deviation*) gibt das Maß der Abweichung nach Überlagerung an. Die den Ankersequenzen entsprechenden Tetrapeptide sind grau unterlegt.

Tab. 4.6: Prozentuale Häufigkeit von Aminosäureresten an jeder Position nach „one2all gap“ Analyse und die daraus abgeleitete Peptidsequenz für den „strukturierten“ Linker.

(A)

AS#	Anker in C-Terminus				Hexapeptidlinker						Anker in N-Terminus			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
AS	E/V	A/L	A	I	T	S	G	D	N	K	A/Y/G	L	N	K
%	15	20	25	20	20	20	40	65	20	20	25	25	20	20
Cons	V	A	A	I	T	S	G	D	N	K	G	L	N	K

(B)

Protein		Anker in C-Terminus	Hexapeptidlinker	Anker in N-Terminus	
Consensus		V A A I	T S G D N K	G L N K	
	1MBL	D A S L	R Y S D N A	A Q N L	
1MAI	<i>H W</i>	V Q G L	R K I * P	D L Q A	L L K G S
	<i>E W</i>	I K A I	K A A *		P D R E G
Linkerkonstr.	E W	I A A I	T S G D N K	<u>G</u> L Q A	L L K G S

(A) statistische Auswertung von Tab. 4.5 sowie die daraus abgeleitete Consensus-Sequenz für die Struktur des 14-er Peptids aus der β -Lactamase und (B) Vergleich der Peptidsequenzen, die für die Konstruktion des „strukturierten“ Linkers relevant sind, mit der daraus gefolgerten Peptidsequenz für das Linkerkonstrukt.

Die homologen Aminosäuren der beiden PH-Domänen von 1MAI (Phospholipase C- δ 1) und des Cytohesin-1 sind fett eingezeichnet. Die kursiv markierten Aminosäuren der Phospholipase C erzeugen untereinander ionische Wechselwirkungen. „*“ gibt den C-Terminus der jeweiligen PH-Domäne an. Zur Ermittlung des Linkerkonstrukts wurden die aufgeführten Aminosäuren jeder Position entsprechend berücksichtigt und aus sterischen Überlegungen heraus ein Glycinrest an den Ankeranfang des N-Terminus festgelegt (unterstrichen).

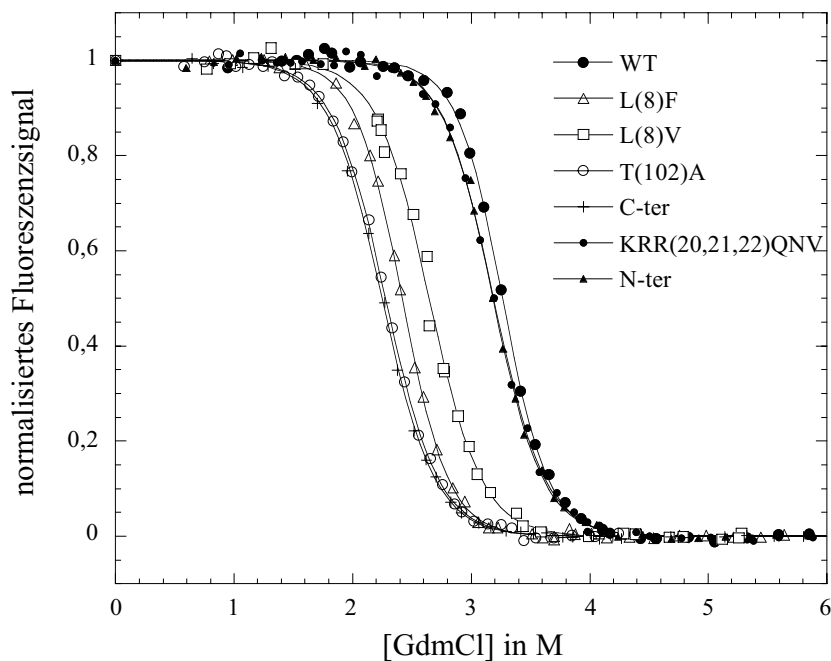
4.3.3. Proteinstabilität im Hinblick auf die Optimierung der Translokation

Sollten die Mutanten der PH-Domäne als Präproteine vor der Translokation gefaltet im Cytoplasma vorliegen, müssen sie zur Translokation entfaltet werden. Dazu bedürfte es bei destabilisierten Mutanten weniger Energieaufwand, was zu einer verbesserten Translokation führen könnte. Von der Proteinstabilität ist zusätzlich auch die Proteasesensitivität und die Aggregationsneigung abhängig, die sich beide in den Expressionsausbeuten niederschlagen können.

Die generellen Voraussetzungen zur Bestimmung der thermodynamischen Stabilität, das Meßverfahren *etc.* sind in Kapitel 4.1.5. ausführlich behandelt. Zur Untersuchung der thermodynamischen Stabilität wurden die Proteine cytoplasmatisch überexprimiert (s. Kapitel 3.6.3.), gereinigt (s. Kapitel 3.7.1.), dialysiert und Proteinlösungen mit unterschiedlichen Konzentrationen an GdmCl angesetzt (s. Kapitel 3.10.3.).

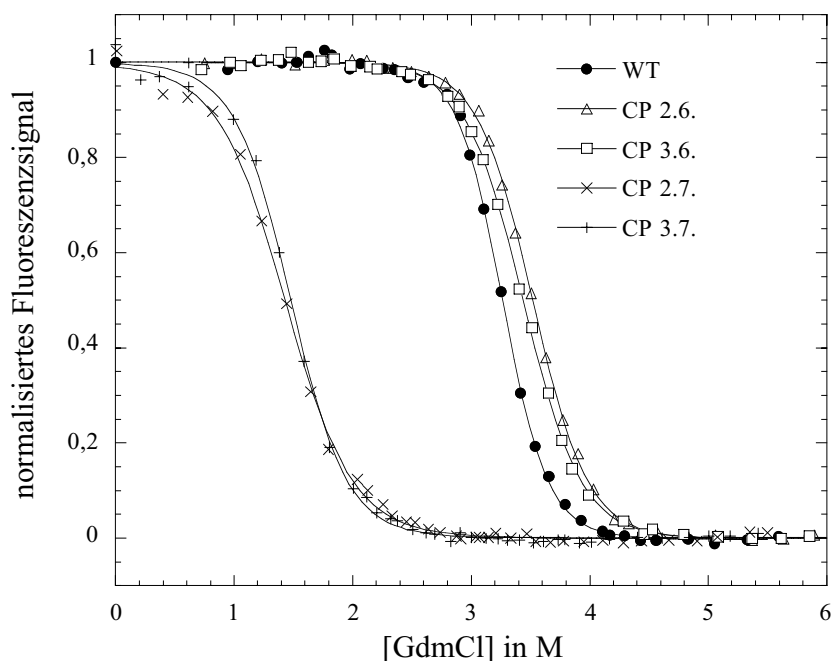
Die Mutanten TEE(102,104,105)AQQ und C/N-ter präzipitieren nach der Aufreinigung und ließen sich nicht in ausreichend hoher Konzentration löslich aufbereiten. Die Präzipitation zeugt von einer erheblichen Destabilisierung: das Protein denaturiert schon in der Lösung ohne Denaturierungsmittel, aggregiert dort mit anderen denaturierten Proteinmolekülen und fällt dann (irreversibel) aus. Alle anderen Mutanten ließen sich (löslich) zur fluoreszenzspektroskopischen Bestimmung von Übergangskurven einsetzen.

Die Circularpermutanten mit modelliertem Linker (CP 1.6. und CP 1.7.) zeigen ein Fluoreszenzverhalten in Abhängigkeit von der Konzentration an Denaturierungsmittel. Daher läßt sich die Faltungsstabilität nicht nach $E \rightleftharpoons N$ auswerten. Anstelle eines sigmoidalen beobachtet man bei CP 1.6. für geringe Konzentrationen an Denaturierungsmittel einen linearen und ab 2 M GdmCl einen mehrstufigen Kurvenverlauf der Emissionsintensität bei 327 nm (s. Abb. 4.19). Der Kurvenverlauf von CP 1.7. verlief ähnlich, ließ sich aber nicht reproduzieren. Beide Übergänge sind somit nicht kooperativ, was mit der Bildung weiterer, stabiler Faltungsintermediate erklärt werden kann. Von allen weiteren Mutanten konnten fluoreszenzspektroskopisch Übergangskurven bestimmt werden, die dem Zweizustandsmodell genügen (s. Abb. 4.18). Von ihnen lassen sich die Freie Faltungsenthalpie und die Kooperativität der Faltung auswerten (s. Tab. 4.7).



(A)

PH-Domänen mit destabilisierenden [L(8)F, L(8)V, T(102)A, C-ter] und mit sequenzspezifischen Mutationen [KRR(20,21,22)QNV und N-ter]



(B)

Circularpermutierte PH-Domänen

Abb. 4.18: Übergangskurven zur Bestimmung der thermodynamischen Stabilität.

A: PH-Domänen mit destabilisierenden oder sequenzspezifischen Mutationen

B: Circularpermutanten.

Die Fluoreszenzmessung erfolgte jeweils bei $\lambda_{\text{ex}} = 282 \text{ nm}$ und $\lambda_{\text{em}} = 327 \text{ nm}$. Die Übergangskurven wurden anschließend normalisiert, so daß das normalisierte Fluoreszenzsignal dem Gehalt an gefaltetem Protein in Abhängigkeit von der GdmCl-Konzentration entspricht.

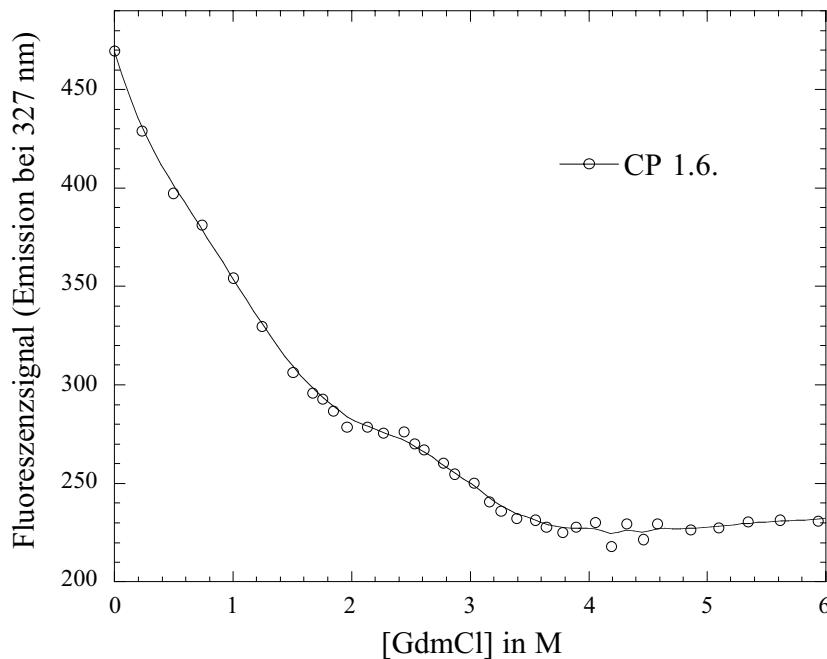


Abb. 4.19: Circularpermutante CP 1.6. mit modelliertem Linker.

Sie gehorcht nicht dem Zweizustandsmodell.

Tab. 4.7: Freie Faltungsenthalpie ΔG_F^P (kJ/mol) und Kooperativität m der Mutanten zur Translokationsoptimierung.

Vermessene Mutanten der PH-Domäne	ΔG_F^P (kJ/mol)	m (kJ/mol·M)	$\Delta\Delta G$ (kJ/mol)
WT	- 41.4	12.6	0
L(8)F	- 30.0	12.5	11.4
L(8)V	- 29.1	11.0	13.3
T(102)A	- 24.7	10.5	16.7
C-ter	- 24.3	10.6	17.1
KRR(20,21,22)QNV	- 35.6	11.1	5.8
N-ter	- 36.7	11.4	4.7
CP 2.6.	- 36.8	10.5	4.6
CP 3.6.	- 34.1	10.0	7.0
CP 2.7.	- 12.0	8.1	29.4
CP 3.7.	- 15.0	9.8	26.4

Ermittelt aus den Übergangskurven von Abb. 4.18: Freie Faltungsenthalpie ΔG_F^P und Kooperativität m nach dem in Kapitel 4.1.5. beschriebenen Verfahren. Alle Mutanten zeigen gegenüber dem Wildtyp eine Destabilisierung. Die Differenz zum Wildtyp ist mit $\Delta\Delta G$ angegeben.

In Kenntnis allgemeiner, stabilisierender Einflüsse auf Proteine durch den hydrophoben Effekt, die helicale N-cap-Struktur und den Ausgleich des Helixdipols wurden Mutationen in die Wildtyp PH-Domäne eingeführt, die diese stabilisierenden Einflüsse bei ihr unterbinden sollten (s. Kapitel 4.3.1). Wie sich Tab. 4.7 entnehmen läßt, werden die im Proteininneren mutierten Varianten L(8)F und L(8)V, sowie die N-cap-mutierte Variante T(102)A den in sie gesetzten Erwartungen gerecht und sind gegenüber dem Wildtyp um jeweils mehr als 10 kJ/mol destabilisiert. Die im Helixdipol zusätzlich mutierte Variante TEE(102,104,105)AQQ ist sogar so instabil, daß sich von ihr keine Übergangskurve aufnehmen ließ. Bei der dipoldestabilisierten Variante C-ter führt die Mutation zu einem gut löslichen Protein mit einer gegenüber dem Wildtyp um 17.1 kJ/mol reduzierten Stabilität, während die zusätzliche Fusion einer weiteren α -Helix in der Mutante C/N-ter eine so große Destabilisierung bewirkt, daß sich von ihr keine Übergangskurve bestimmen ließ.

Der Einfluß sequenzspezifischer Mutationen bei den Varianten KRR(20,21,22)QNV und N-ter auf die thermodynamische Stabilität ist wie erwartet mit jeweils ca. 5 kJ/mol gering. Die Entfernung lösungsmittelorientierter positiver Ladungen aus dem β 2-Faltblatt hat somit eine genauso geringe Auswirkung auf die Gesamtstabilität der Wildtyp PH-Domäne wie das Angleichen der ersten beiden Aminosäuren aus dem β 1-Faltblatt an den N-Terminus des reifen ompA-Proteins.

Die Circularpermutanten lassen sich bezüglich der Stabilitätsmessung in drei Gruppen einteilen:

- Die Struktur und die Faltungseigenschaften derer mit konstruiertem Linker (CP 1.6. und CP 1.7.) sind ungeklärt und ihre thermodynamische Stabilität läßt sich nicht nach dem Zweizustandsmodell bestimmen.
- Die Circularpermutanten mit dem neu eingeführten N-Terminus unmittelbar vor der α -Helix (CP 2.7. und CP 3.7.) liegen zwar gefaltet vor, sind aber sehr instabil.
- Die Circularpermutanten mit unspezifischen Linkern und dem neuen N-Terminus in Loop 6 (CP 2.6. und CP 3.6.) sind dagegen ähnlich stabil wie der Wildtyp.

Die Kooperativität der Faltung erreicht für den Wildtyp den höchsten Wert und für die Circularpermutanten die niedrigsten Werte. Dies könnte ein Hinweis auf den veränderten Faltungsweg bei Circularpermutanten sein. Zwischen diesen Extrema liegen die *m*-Werte der restlichen Varianten.

4.3.4. Cytoplasmatische Expression im Hinblick auf die Translokationsoptimierung

Die rekombinante Expression eukaryontischer Proteine in *E. coli* wird von vielen Faktoren beeinflusst. (s. Kapitel 3.6.1.). Die periplasmatische Expression der gleichen Proteine (fusioniert mit einem Signalpeptid) wird zusätzlich durch die Effizienz des Transportvorganges bestimmt. Um Aussagen über den Transportvorgang anhand von periplasmatischen Expressionsausbeuten treffen zu können, bedarf es deshalb zumindest des Vergleichs mit der quantitativen cytoplasmatischen Expression der jeweiligen Mutante. Allerdings unterscheiden sich beide Expressionen nicht nur im Transportvorgang voneinander: die verschiedene Ausstattung der Kompartimente mit Proteasen und Chaperonen sorgt zusätzlich für unterschiedliche Expressionsausbeuten in beiden Kompartimenten.

Der Expressionsvergleich von Mutanten setzt ein standardisiertes Expressionsverfahren voraus; es wurde also nicht versucht, die maximale cytoplasmatische Expressionsausbeute zu erzielen. Die Proteine wurden unter dem *tet*-Promotor in BL21-Zellen exprimiert; diese *E. coli*-Zellen wurden wie in Kapitel 3.6.3. beschrieben kultiviert. Der native Zellaufschluß erfolgte mittels einer Kugelmühle, die Extraktion der löslichen Proteine mit PBS und anschließend die Extraktion unlöslicher sowie durch die Kugelmühle präzipitierter Proteine mit 8 M Harnstoff in PBS (s. Kapitel 3.6.3.: nativer Zellaufschluß). Sowohl die lösliche als auch die unlösliche Fraktion einer jeden Mutante wurde mit SDS-PAGE chromatographisch untersucht (s. Abb. 4.20). Zum quantitativen Vergleich der cytoplasmatischen Expressionsausbeute wurden die Zellen nach Überexpression denaturierend aufgeschlossen, die Proteine über den His₆-tag an einer Ni-NTA-Säule gereinigt und zurückgefaltet (s. Kapitel 3.7.1.). Über Absorption des Eluats bei 280 nm wurde die Proteinausbeute bestimmt (s. Kapitel 3.10.1. und Abb. 4.21). Danach lassen sich grob in drei Klassen von Mutanten unterteilen:

- die Mutante KRR(20,21,22)QNV und die destabilisierten Mutanten T(102)A und C-ter werden quantitativ ähnlich wie der Wildtyp exprimiert,
- die Mutanten L(8)F, TEE(102,104,105)AQQ, CP 1.6., CP 2.6., CP 3.6. [auch jeweils noch mit der zusätzlichen Mutation KRR(20,21,22)QNV] werden cytoplasmatisch mit ungefähr einem Drittel an Ausbeute des Wildtyps exprimiert,
- die Mutanten N-ter, L(8)V, C/N-ter, CP 1.7., CP 2.7. und CP 3.7. werden verglichen mit dem Wildtyp zu weniger als einem Sechstel exprimiert.

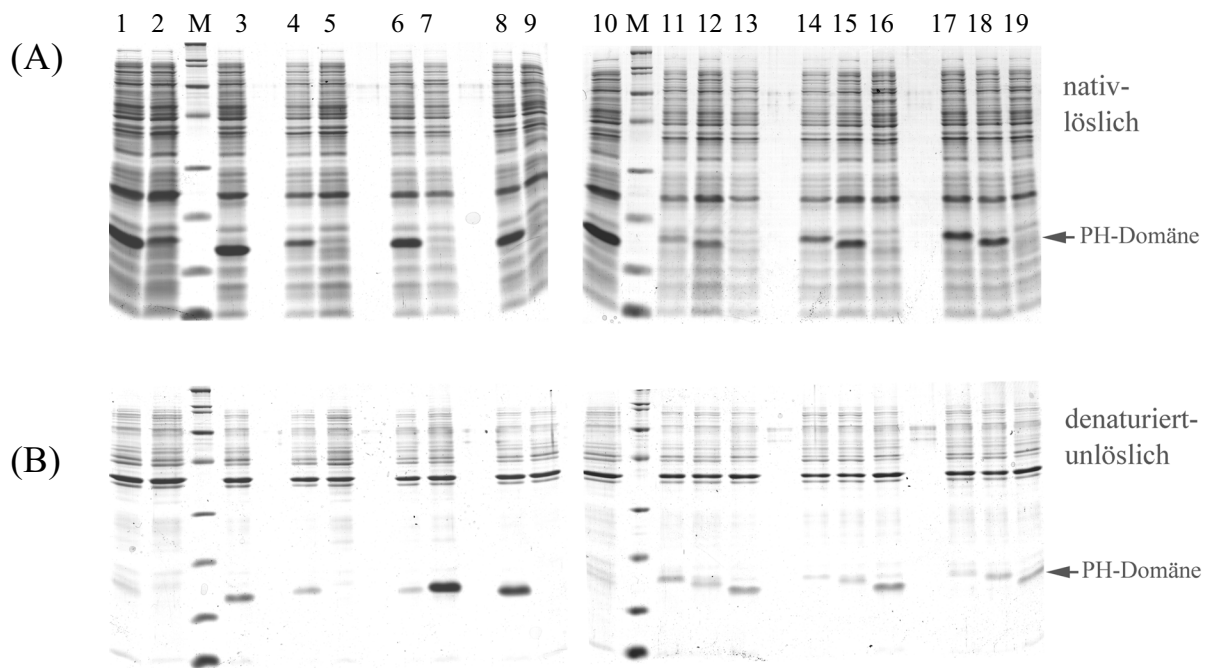


Abb. 4.20: Gelchromatographische Auswertung (15 %-iges SDS-PAGE, Coomassie-Färbung) cytoplasmatisch exprimierter Mutanten der PH-Domäne zur Translokationsoptimierung im Zellaufschluß aus BL21-Zellen.

(A) lösliche Proteine nach Aufschluß mit Zelmühle. Der jeweilige Auftrag entspricht einem Bakterien-Kulturvolumen von 30 μ l.

(B) nach nativem Aufschluß mit der Zelmühle präzipitiert vorliegende Proteine durch Aufnahme des Zellpellets mit 8 M Harnstoff. Der Auftrag entspricht einem Bakterien-Kulturvolumen von 30 μ l.

Die Aufträge: 1: WT; 2: N-ter; 3: KRR(20,21,22)QNV; 4: L(8)F; 5: L(8)V; 6: T(102)A; 7: TEE(102,104,105)AQQ; 8: C-ter; 9: C/N-ter; 10: WT; 11: CP 1.6.; 12: CP 1.6.(KRR->QNV); 13: CP 1.7.; 14: CP 2.6.; 15: CP 2.6.(KRR->QNV); 16: CP 2.7.; 17: CP 3.6.; 18: C.P. 3.6.(KRR->QNV); 19: CP 3.7.; M: Proteinstandard, *broad range* Marker (Biorad): 6.5 kDa, 14.5 kDa, 21.5 kDa, 31 kDa, 45 kDa, 66 kDa, 97.4 kDa, 116 kDa, 200 kDa.

Die Mutanten CP.X.6.(KRR->QNV) sind Circularpermutanten mit der zusätzlichen Mutation KRR(20,21,22)QNV. Die Substitution positiver Ladungen führt zu einer höheren gelelektrophoretischen Beweglichkeit.

Dabei lassen sich die Mutanten TEE(102,104,105)AQQ, CP 1.7., CP 2.7. und CP 3.7. fast nur denaturiert im Cytoplasma nachweisen, was auf eine Aggregation zu *inclusion-bodies* oder auf Denaturierung während des mechanischen Zellaufschlusses mittels Zelmühle hinweist. L(8)V und C-ter liegen sowohl im löslichen als auch im unlöslichen Aufschluß zu ungefähr gleichen Teilen vor, während die anderen Mutanten fast ausschließlich in der löslichen Proteinfraction zu finden sind (s. Abb. 4.20). Damit besteht ein direkter

Zusammenhang zur jeweiligen Stabilität der Mutanten (s. vorheriges Kapitel): besonders instabile Proteine lassen sich mit der Kugelmühle nicht löslich extrahieren. Die Ausbeute cytoplasmatisch exprimierter Mutanten korreliert nicht mit der Stabilität: so ist die Mutante C-ter um 17.1 kJ/mol gegenüber dem Wildtyp destabilisiert, läßt sich aber quantitativ ähnlich wie der Wildtyp exprimieren, während die Circularpermutante CP 2.6. ähnlich stabil wie der Wildtyp ist, aber nur mit einem Drittel seiner Ausbeute exprimiert wird.

Die Mutation KRR(20,21,22)QNV besitzt weder beim Wildtyp noch bei den Circularpermutanten einen deutlichen Einfluß auf die Expressionsausbeute unter standardisierten Expressionsbedingungen, obwohl bei dieser Mutation das seltene Arginin-Codon (AGG) [von R21] entfernt wurde. Es wirkt sich somit beim Wildtyp nicht expressionslimitierend aus.

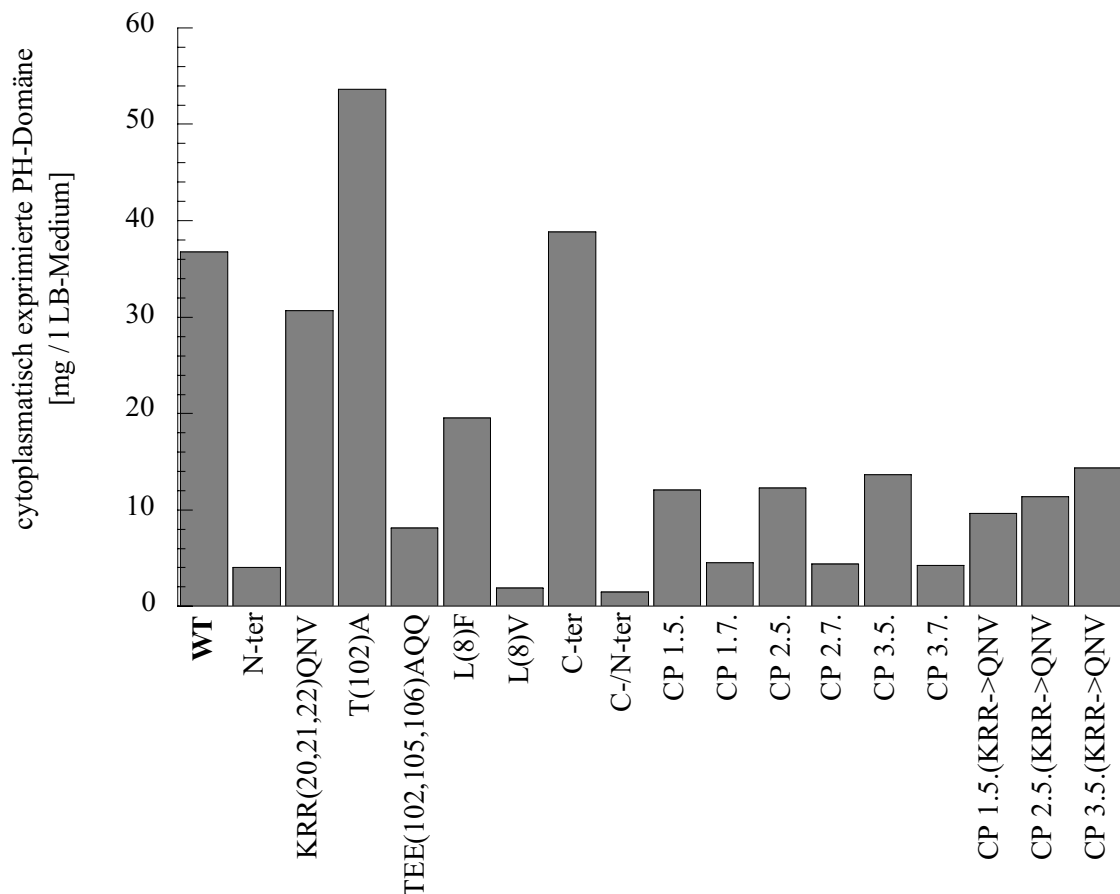


Abb. 4.21: Proteinausbeute cytoplasmatisch exprimierter Mutanten der PH-Domäne zur Translokationsoptimierung.

Die Proteine wurden nach Kapitel 3.6.3. exprimiert, nach Kapitel 3.7.1. über den His-tag und die Ausbeute über die OD₂₈₀ nach Kapitel 3.10.1. bestimmt.

Die Mutanten CP.X.6.(KRR->QNV) sind Circularpermutanten mit der zusätzlichen Mutation KRR(20,21,22)QNV.

4.3.5. Untersuchung der periplasmatischen Expression

Die geringe periplasmatische Expression der Wildtyp PH-Domäne wird in dieser Arbeit als mögliche Ursache für die nicht erfolgreiche Selektion der Bibliothek randomisierter PH-Domänen im *phage-display*-Verfahren angesehen. Die Verbesserung der periplasmatischen Expression durch Mutationen des Wildtyps sollte dann die erfolgreiche Anwendung dieses Selektionsverfahrens auf eine Bibliothek randomisierter PH-Domänen ermöglichen.

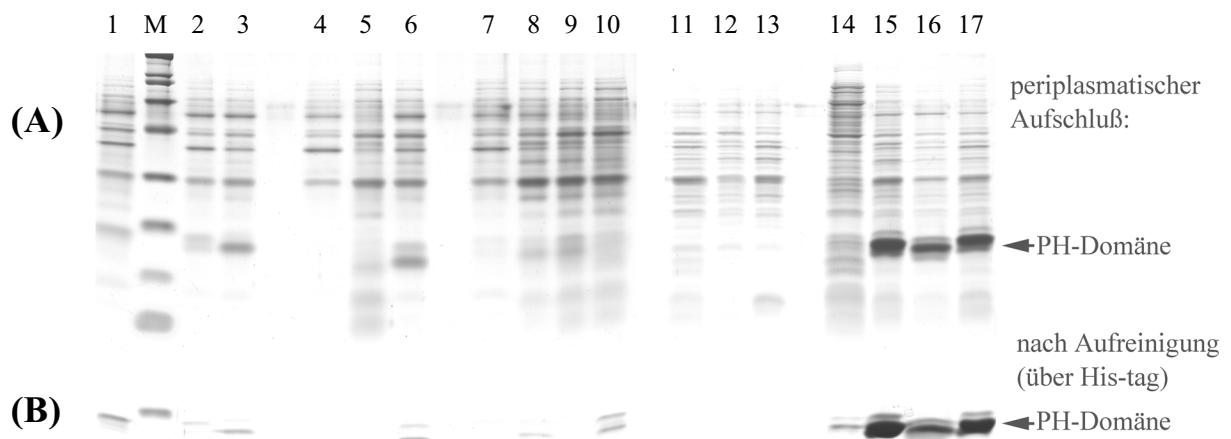


Abb. 4.22: Gelchromatographische Auswertung (15 %-iges SDS-PAGE, Coomassie-Färbung) periplasmatisch exprimierter Mutanten zur Translokationsoptimierung mit $ompA^M$ -Signalpeptid aus JM83-Zellen.

(A) Periplasmatischer Zellaufschluß nach periplasmatischer Expression der Mutanten und (B) die über den His-tag aufgereinigten Proteine dieser Expression. Deutlich wird die vielfach verstärkte periplasmatische Expression der Circularpermutanten CP 2.6., CP.ohnLi und CP 3.6. (Aufträge 15-17) gegenüber den restlichen Mutanten. Die Aufträge: 1: WT; 2: N-ter; 3: WT.miLi; 4: C-ter; 5: C/N-ter; 6: KRR(20,21,22)QNV; 7: L(8)F; 8: L(8)V; 9: T(102)A; 10: TEE(102,104,105)AQQ; 11: CP 1.7.; 12: CP 2.7.; 13: CP 3.7.; 14: CP 1.6.; 15: CP 2.6.; 16: CP.ohnLi; 17: CP 3.6.; M: Proteinstandard: 6.5 kDa, 14.5 kDa, 21.5 kDa, 31 kDa, 45 kDa, 66 kDa, 97.4 kDa, 116 kDa, 200 kDa.. (Erklärungen zu CP.ohnLi und WT.miLi s. Kapitel 4.3.9.).

4. Ergebnisse

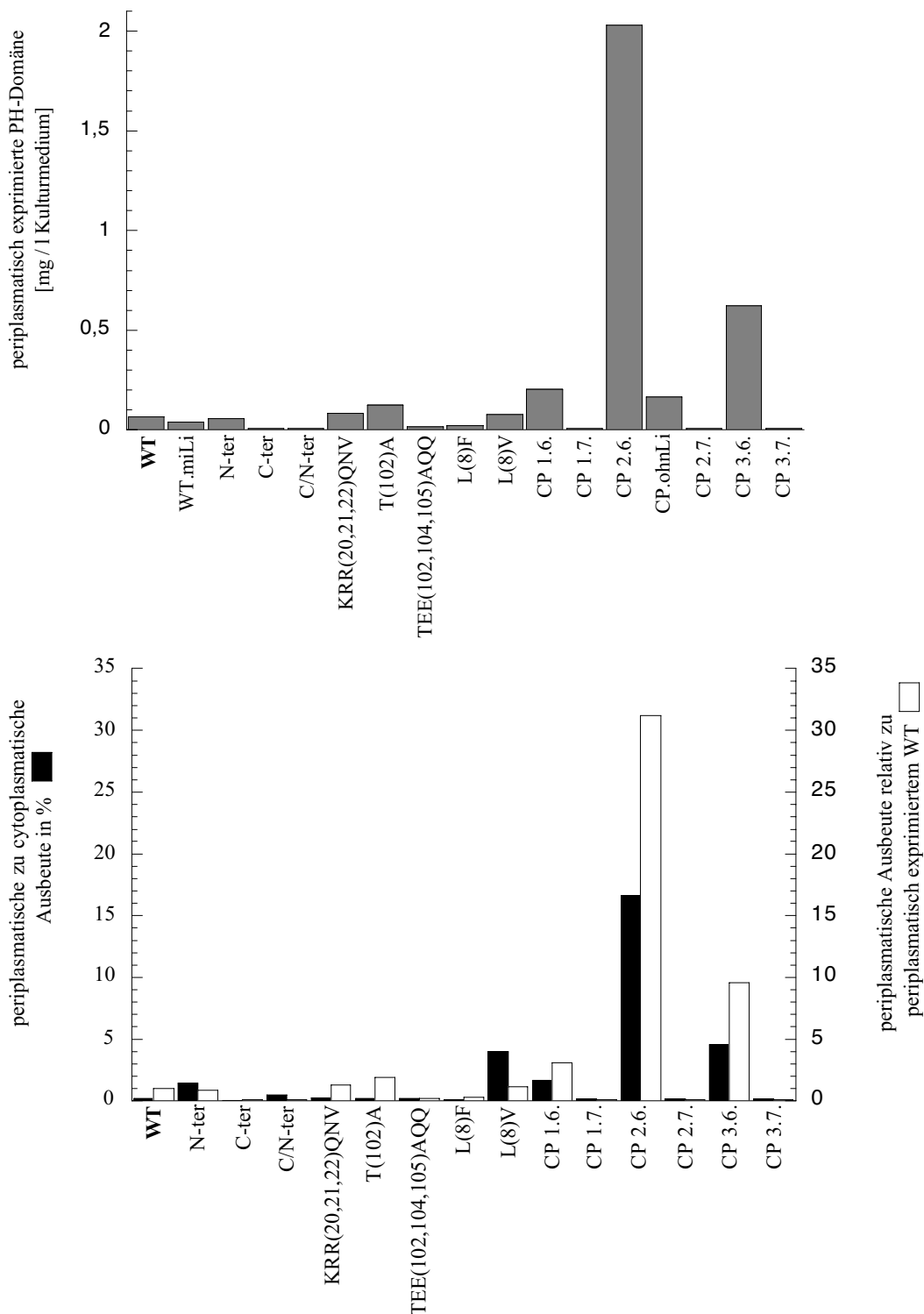


Abb. 4.23: (A) Absolute und (B) relative periplasmatische Expressionsausbeuten der verschiedenen PH-Domänen zur Translokationsoptimierung.

In (B) ist das Verhältnis der periplasmatischen Expression zur jeweiligen cytoplasmatischen Expression und das Verhältnis zum periplasmatisch exprimierten Wildtyp dargestellt. (Erklärungen zu CP.ohnLi und WT.miLi s. Kapitel 4.3.9.)

Zur Untersuchung der periplasmatischen Expression wurden die Mutanten mit dem modifizierten Signalpeptid des ompA-Proteins (ompA^M) genetisch fusioniert und unter standardisierten Bedingungen in das Periplasma von JM 83 *E. coli* Zellen transloziert (s. Kapitel 3.6.2.). Nach periplasmatischem Zellaufschluß wurden die Proteinproben gelchromatographisch untersucht (s. Abb. 4.22), über den His-tag gereinigt (s. Kapitel 3.7.1.) und die Ausbeute an gereinigter, periplasmatisch exprimierter PH-Domäne über den Bradford-Test durch Absorption bei $\lambda = 595 \text{ nm}$ bestimmt (s. Abb. 4.23A).

Bei der periplasmatischen Expression zeigen die sequenzspezifisch mutierten und die meisten destabilisierten Varianten ähnlich schlechte Expressionsausbeuten wie der Wildtyp ($65 \mu\text{g}$ pro Liter LB-Kulturmedium): nur die Mutanten C-ter und C/N-ter lassen sich aufgrund ihrer geringen Ausbeuten nicht mehr nachweisen (s. Abb. 23A). Während sich die Circularpermutanten CP. X.7. unter den gewählten Versuchsbedingungen ebenfalls nicht im Periplasma quantifizieren lassen, zeigen die Mutanten CP 2.6. mit über 2 mg , CP 3.6. mit über 0.6 mg und CP.ohnLi mit 0.2 mg Protein pro Liter Kulturvolumen eine deutlich gesteigerte periplasmatische Expression (s. Abb. 4.22). Die periplasmatische Expression der Mutante CP 1.6. führt zur Zellyse, was anhand der cytoplasmatischen Proteine im Hintergrund in Abb. 4.22A sichtbar wird und was beim periplasmatischen Aufschluß deutlich wurde, als Nukleinsäuren die Viskosität der Probe deutlich erhöhten. Verglichen mit der cytoplasmatischen Expression (ohne Signalpeptid) ließ sich die Circularpermutante CP 2.6. nach Fusion mit dem ompA^M -Signalpeptid zu 16% im Periplasma nachweisen, während sich der Wildtyp gerade einmal zu 0.2% im Periplasma exprimieren läßt. Die periplasmatische Expression von CP 2.6. ist über 30-fach höher als die des Wildtyps (s. Abb. 4.23B).

Ein Zusammenhang zwischen cytoplasmatischer und periplasmatischer Expri-mierbarkeit bei den verschiedenen Mutanten läßt sich nicht erkennen (s. Abb. 23B). Die an der α -Helix mutierte Variante C-ter läßt sich mit dem Wildtyp vergleichbar gut cytoplasmatisch exprimieren, ist thermodynamisch dem Wildtyp gegenüber deutlich destabilisiert (s. Tab. 4.7) und läßt sich nach Fusion mit dem ompA^M -Signalpeptid nicht im Periplasma nachweisen, während sich die Mutante N-ter nur zu 10% gegenüber dem Wildtyp cytoplasmatisch, aber ähnlich wie der Wildtyp periplasmatisch exprimieren läßt.

Mögliche Ursachen für die überragende periplasmatische Expression von CP 2.6. werden in Kapitel 5.4. eingehend diskutiert und die Ergebnisse ergänzender Experimente zur Klärung dieses Phänomens in Kapitel 4.3.9. vorgestellt.

4.3.6. Prozessierungsprodukte nach periplasmatischer Expression

Nicht nur die geringe periplasmatische Ausbeute bei Expression der Wildtyp PH-Domäne ist ein Indiz für Schwierigkeiten bei der Translokation: Schon im SDS-Gel erkennt man unter anderem für den aufgereinigten Wildtyp mehrere Banden, die auf unterschiedliche Prozessierungsprodukte der Präproteine hinweisen (s. Abb. 4.22B).

Signalpeptide werden nach der Translokation über den allgemeinen Translokationsmechanismus prozessiert (Economou, 1999). Die dazu notwendige Signalpeptidase I prozessiert das gefaltete Protein im Periplasma (Chatterjee *et al.*, 1995). Zur Untersuchung dieser Prozessierung wurden einige Mutanten der PH-Domäne periplasmatisch exprimiert (s. Kapitel 3.6.2.), gereinigt (s. Kapitel 3.7.1) und über Massenspektroskopie vermessen (s. Kapitel 3.10.5.). Dabei ließen sich die Massen der einzelnen Prozessierungsprodukte bis auf 2 Da genau bestimmen, die molaren Verhältnisse der unterschiedlichen Prozessierungsprodukte zueinander jedoch nur semiquantitativ abschätzen (s. Tab. 4.8).

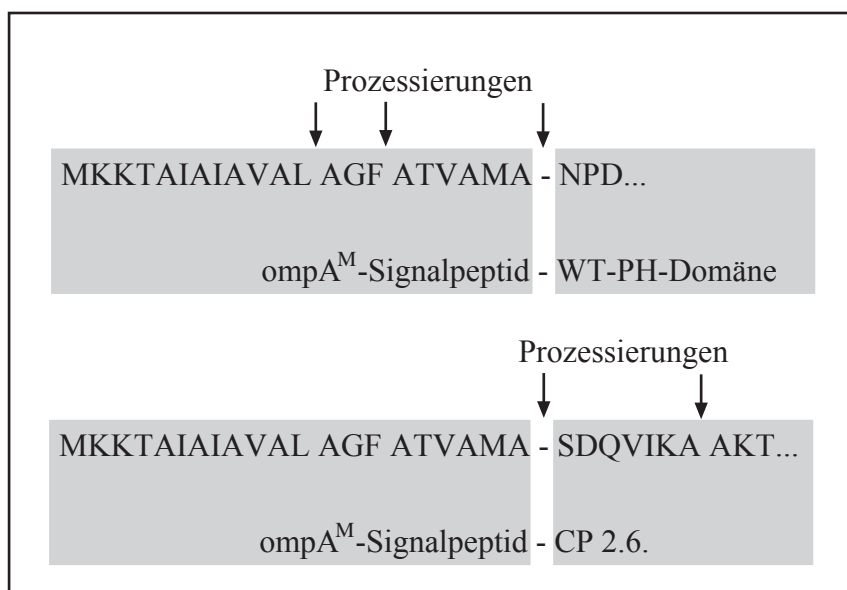


Abb. 4.24: Prozessierungen der periplasmatisch exprimierten Wildtyp PH-Domäne und der Circularpermutante CP 2.6. fusioniert mit dem modifizierten Signalpeptid *ompA^M* nach massenspektroskopischer Analyse.

Die erwartete Prozessierung zum reifen Protein erfolgt bei P-1. Bei der Wildtyp PH-Domäne lassen sich zudem im Bereich des Signalpeptids die Prozessierungsprodukte P+6 sowie P+9 und bei CP 2.6. das Prozessierungsprodukt P-8 innerhalb der Circularpermutante nachweisen.

Tab. 4.8: Semiquantitative Auswertung der Prozessierungsprodukte von Mutanten zur Translokationsoptimierung nach massenspektroskopischer Analyse.

Protein	P-1	P+6	P+9	P-8	0
CP 2.6.	90 %	-	-	10 %	-
L(8)V	80 %	5 %	15 %	-	-
L(8)F	10 %	-	-	-	90 %
N-ter	15 %	30 %	25 %	-	30 %
WT	10 %	5 %	10 %	-	75 %

Die Prozessierungsprodukte sind wie in Abb. 4.24 bezeichnet. „0“ ist nicht-prozessiertes Fusionsprotein aus ompA^M-Signalpeptid und PH-Domäne.

Alle massenspektroskopisch vermessenen Mutanten zeigen das korrekt prozessierte Protein [P-1]. Dazu kommen bei der Circularpermutante CP 2.6. ein weiteres Prozessierungsprodukt [P-8] im N-Terminus des reifen Proteins und bei Wildtyp, destabilisierten und sequenzspezifisch mutierten Varianten die Prozessierungsprodukte [P+6] sowie [P+9] innerhalb des ompA^M-Signalpeptids (s. Abb. 4.24).

Die massenspektroskopische Analyse wurde einmalig am Max-Planck-Institut für Biochemie (Martinsried) vorgenommen. Gerade das Auftreten nicht-prozessierter Fusionsproteine könnte auf partielle Zellyse während des Periplasmaaufschlusses oder auf die geringe Aktivität der Signalpeptidase I (Chatterjee *et al.*, 1995) in Folge von Translokationsunterbrechungen (s. Translokationsunterbrechung beim Import in Mitochondrien Kapitel 4.5.2.) zurückgeführt werden. Die Mutation L(8)V wirkt sich positiv auf die korrekte Prozessierung aus. Gleichzeitig verringert diese Mutation die cytoplasmatische Expressionsausbeute, während die periplasmatische Expressionsausbeute dem Wildtyp entspricht. Zur weitergehende Interpretation der massenspektroskopischen Daten sollten diese zunächst durch weitere Messungen bestätigt werden. Die quantitative Auswertung der massenspektroskopischen Untersuchungen verdeutlicht aber, daß die vermessenen Circularpermutanten im Gegensatz zu den anderen Mutanten nach periplasmatischer Expression keine zusätzlichen Prozessierungen im Signalpeptid aufweisen. Dafür wird der N-Terminus der Circularpermutante prozessiert.

Vorhersage von Prozessierungsstellen

Zur Identifizierung von Signalpeptiden und den dazugehörigen Prozessierungsstellen wurden die *ompA^M*-fusionierte Wildtyp PH-Domäne und CP 2.6. über den SignalP V1.1-Server (<http://www.cbs.dtu.dk/services/SignalP/>; Nielsen *et al.*, 1997) ausgewertet (s. Abb. 4.25). Das dazugehörige Verfahren beruht auf der Analyse durch ein neuronales Netzwerk, welches die eingegebenen Peptidsequenzen mit den bekannten Prozessierungsstellen anderer Proteine vergleicht (hier bei *gram*-negativen Bakterien). Dabei geht nur die Primärstruktur in die Statistik ein. Der C-Wert gibt die Wahrscheinlichkeit an, mit der die betreffende Aminosäure die erste Aminosäure des prozessierten Proteins darstellt.

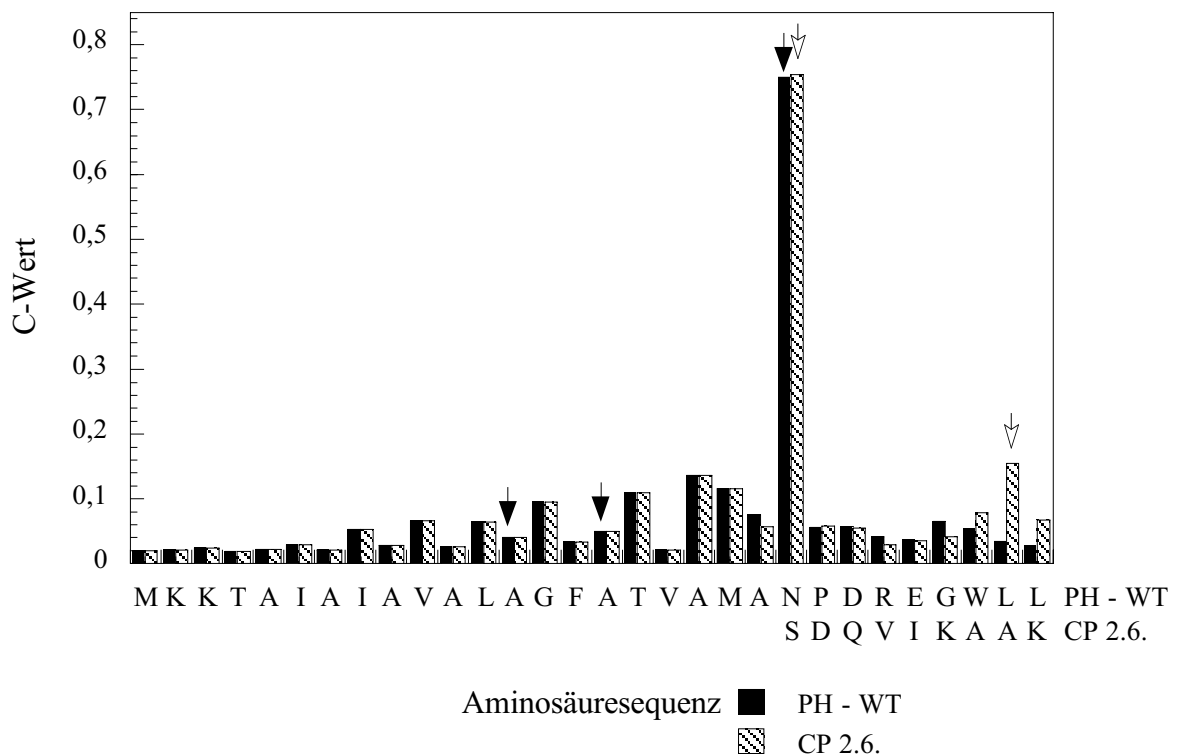


Abb. 4.25: Prozessierungsanalyse der 30 N-terminalen Aminosäuren von Wildtyp PH-Domäne und der Circularpermutante CP 2.6. mit dem *ompA^M*-Signalpeptid fusioniert.

Über den SignalP V1.1-Server (<http://www.cbs.dtu.dk/services/SignalP/>) wurde nach Prozessierungsstellen für Signalpeptidasen bei *gram*-negativen Bakterien gesucht. Mit gefüllten Pfeilen sind die über Massenspektroskopie nachgewiesenen Prozessierungen des Wildtyps und mit den ungefüllten Pfeilen die nachgewiesenen Prozessierungen der Circularpermutante CP 2.6. eingezeichnet.

Für Aminosäure 22 wird für beide Proteine ein C-Wert von ca. 0.75 berechnet (s. Abb. 4.25). Die Gegenwahrscheinlichkeit beträgt damit für beide Peptidsequenzen 0.25 und besagt, daß nur 25 % aller Prozessierungen vor einer der anderen Aminosäuren [als vor N(22)] stattfinden sollten. Zwischen den Aminosäuren 21 und 22 liegt demzufolge die erwartete Prozessierungsstelle durch die Signalpeptidase I. Diese Angabe stimmt mit der Versuchsplanung überein, da die gewählte modifizierte Signalsequenz des *outer membrane protein A* (ompA^M) von *E. coli* bis einschließlich zu Aminosäure 21 reicht und das prozessierte Protein nach der Translokation mit der jeweiligen Aminosäure 22 [P-1] N-terminal beginnen soll. Während sich die weiteren massenspektroskopisch detektierten Prozessierungsprodukte des Wildtyps [P+6] und [P+9] nicht mit erhöhten C-Werten in Einklang bringen lassen, zeigt die Position 29 bei der Circularpermutante den zweithöchsten C-Wert von 0.16 (s. Abb. 4.25). Das entsprechend prozessierte Protein läßt sich im Massenspektrum als Prozessierungsprodukt [P-8] mit einem prozentualen Anteil von ca. 10 % nachweisen (s. Abb. 4.24).

Die periplasmatisch exprimierte Circularpermutante CP 2.6. zeigt als einzige massenspektroskopisch untersuchte Mutante genau die Prozessierungsstellen, die man nach Vergleich mit natürlich-periplasmatisch in *gram*-negativen Bakterien exprimierten Proteinen vorhersagen konnte. Alle anderen massenspektroskopisch vermessenen Mutanten weisen zusätzliche Prozessierungen auf, wie sie bei *gram*-negativen Bakterien sonst nicht beobachtet werden, wenn auch die destabilisierte Mutante L(8)V überwiegend „natürlich“ prozessiert wird. Während die Circularpermutante also vom allgemeinen sekretorischen Apparat bzw. der Signalpeptidase I als „natürliches“ Substrat angenommen wird, ist die Translokation bzw. die Prozessierung der anderen Mutanten erschwert.

4.3.7. Weitere Ergebnisse zur Translokation der Circularpermutante CP 2.6.

Die Circularpermutante CP 2.6. wird verglichen mit dem Wildtyp 30-fach besser periplasmatisch exprimiert; zudem wird sie (im Gegensatz zum Wildtyp) vom allgemeinen sekretorischen Apparat *gram*-negativer Bakterien als „natürliches“ Substrat erkannt, was sich im Prozessierungsmuster äußert. Damit ist das Ziel einer Optimierung der periplasmatischen Expression bzw. Translokation der Wildtyp PH-Domäne erreicht; es ist aber nicht geklärt, auf welchen Eigenschaften der Circularpermutante die verbesserte periplasmatische Expression bzw. Translokation beruhen: dem können sowohl struktur- als auch

sequenzspezifische Ursachen zugrunde liegen. Zwei weitere, im folgenden aufgeführte Untersuchungen schränken die Erklärungsmöglichkeiten ein.

Abhängigkeit der thermodynamischen Stabilität vom Signalpeptid

CP 2.6. wurde einschließlich des Signalpeptides (ompA^M) als cytoplasmatisches Protein-Aggregat (*inclusion-bodies*) exprimiert (s. Kapitel 3.6.4.) und eine Übergangskurve zur Bestimmung der thermodynamischen Stabilität des gereinigten Proteins vermessen. Die Übergangskurve der Circularpermutante mit Signalpeptid zeigt eine um 3 kJ/mol höhere thermodynamische Stabilität als die der Circularpermutante ohne Signalpeptid (s. Abb. 4.26). Dies verdeutlicht, daß das Signalpeptid im Fall dieser Circularpermutante keine Destabilisierungsfunktion besitzt, wie dies für das Präprotein der β -Lactamase postuliert wurde (Laminet *et al.*, 1990) und in Funktionsanalysen bestätigt werden konnte (McCarthy *et al.*, 1998): bei 37°C liegt *in vitro* translatierte β -Lactamase gefaltet und funktionell vor und das dazugehörige Präproteine ist inaktiv, während bei 30°C beide Proteine funktionell sind.

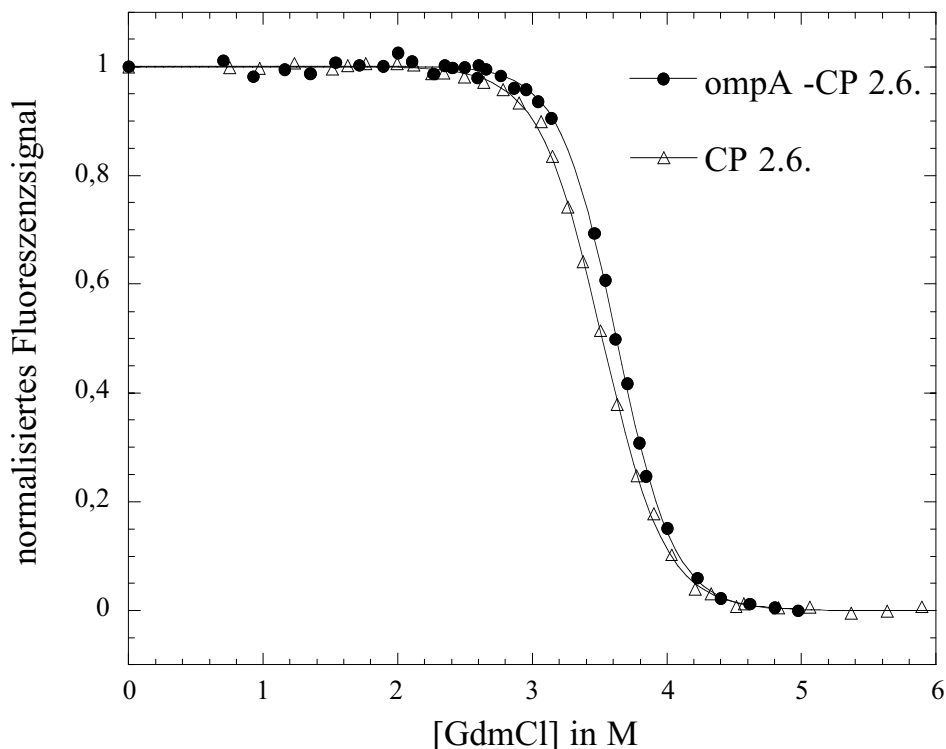


Abb. 4.26: Übergangskurven der Circularpermutante CP 2.6. in Abhängigkeit von der Fusion mit dem Signalpeptid ompA^M .

Zwei Schlußfolgerungen können aus diesem Experiment gezogen werden: die Faltung von Präproteinen vor der Translokation im Cytoplasma ist aus energetischer Sicht möglich und die gute periplasmatische Expression der Circularpermutante CP 2.6. beruht nicht auf der direkten Destabilisierung durch das Signalpeptid. Allerdings mag durch das Signalpeptid die Mutante im Cytosol indirekt destabilisiert werden, wenn dadurch zusätzliche Chaperone an ihr binden.

Einfluß einer N-terminalen, unstrukturierten Peptidsequenz auf die Translokation der Circularpermutante CP 2.6. und auf den Wildtyp

Für die Primärsequenz der mit ompA^M-Signalpeptid fusionierten Circularpermutante CP 2.6. wird eine weitere Prozessierungsstelle durch Signalpeptidase I vor Position [P-8] vorhergesagt, die sich auch nach periplasmatischer Expression massenspektroskopisch nachweisen läßt (s. Kapitel 4.3.8.). Die sieben Aminosäuren SDQVIKA zwischen [P-1] und [P-8] befinden sich bei der modellierten Struktur der Wildtyp PH-Domäne in Loop 6, sind dort nicht Bestandteil eines sekundären Strukturelements und gehen in der modellierten Struktur keine weiteren intramolekularen Wechselwirkungen mit weiteren Aminosäureresten ein. Ihnen sollte deshalb als N-Terminus der Circularpermutanten CP X.6. keine besondere strukturelle oder thermodynamisch stabilisierende Bedeutung zukommen. Die Bedeutung dieser sieben unstrukturierten Aminosäuren für die Translokation aber ist ungewiß. Sie könnten der Circularpermutante zu ihren hervorragenden Translokationseigenschaften verhelfen. Um eine mögliche Abhängigkeit der Translokation von ihnen zu studieren, wurden in der Mutante CP.ohnLi diese sieben Aminosäuren entfernt, so daß hinter dem Signalpeptid direkt die (strukturierte) Proteindomäne der Circularpermutante CP 2.6. beginnt. In einer weiteren Mutante (WT.miLi) wurden diese sieben Aminosäuren als Linker zwischen Signalpeptid und Wildtyp PH-Domäne kloniert (die Aminosäuresequenzen beider Mutanten sind in Tab. 4.4 aufgeführt).

Nach periplasmatischer Expression läßt sich zwar eine deutliche Verringerung der Ausbeute durch Entfernen der sieben N-terminalen Aminosäuren (CP.ohnLi) gegenüber CP 2.6. nachweisen, doch läßt sich diese Mutante weiterhin besser als die anderen nicht circularpermutierten PH-Domänen periplasmatisch exprimieren (s. Abb. 4.22 und Abb. 4.23). Die Wildtyp PH-Domäne läßt sich nach Insertion dieser sieben Aminosäuren (WT.miLi) sogar schlechter periplasmatisch exprimieren. Damit zeigen sich diese sieben Aminosäuren nicht für die verbesserte Translokation der Circularpermutante CP 2.6. verantwortlich.

4.4. Die Expression der PH-Domäne auf dem M13-Phagen

Im *phage-display* System wird das auf der Oberfläche des M13-Phagen zu repräsentierende Protein als Fusion mit einem Hüllprotein (hier das g3-Protein) des Phagen exprimiert. Dies setzt die Translokation durch die innere Membran von *E. coli* voraus (s. Kapitel 1.5.1.). Die Translokationsoptimierung der PH-Domäne sollte ihrer verbesserten Expression auf der Phagenoberfläche dienen. Da kein Aktivitätsnachweis der PH-Domäne auf der Phagenoberfläche verfügbar ist, wurde ihre Expressionsausbeute als g3-Proteinfusion auf Phagen quantifiziert. Die Phagen wurden gereinigt, der Proteinbestandteil denaturiert und über ein 7.5 %-iges SDS-Gel elektrophoretisch aufgetrennt. Das Gel wurde auf eine Nitrozellulose-Membran geblottet und die g3-Proteinfusion neben dem g3-Protein mit einem murinen *anti-g3*-Antikörper als Primärantikörper und einer Kaninchen *anti-Maus*-Antikörper-Peroxidase-Fusion als sekundärer Antikörper detektiert (s. Kapitel 3.8.4.).

Die Fusionierung des g3-Proteins mit den Circularpermutanten führt zu einer deutlich höheren Expression der Circularpermutanten verglichen mit dem Wildtyp (s. Abb. 4.27). Allerdings lassen sich Proteinmengen aus Western-Blots (s. Kapitel 3.8.4.) nur semiquantitativ abschätzen.

WT M 1.6. 2.6. 3.6.

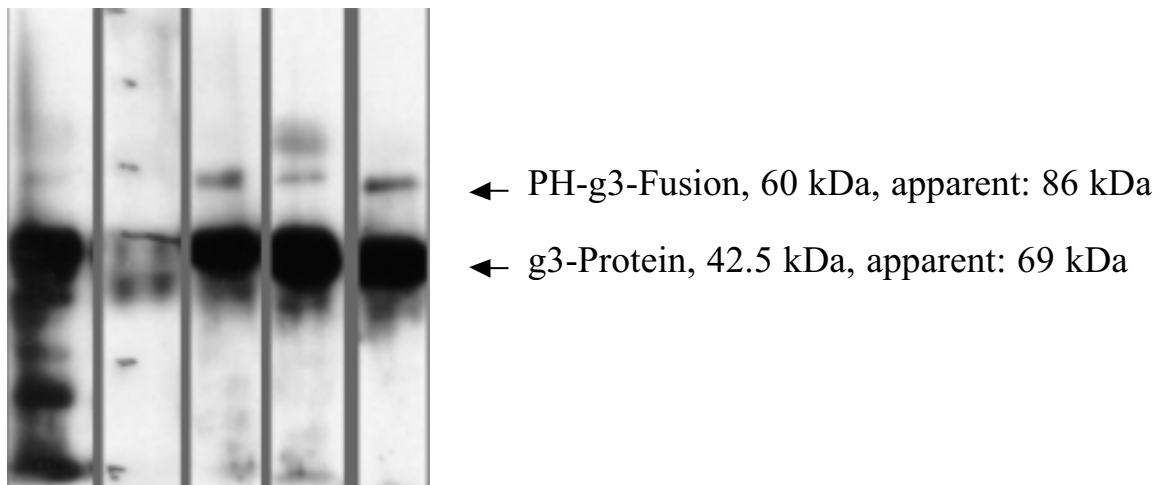


Abb. 4.27: Ausschnitte eines Western-Blots von g3-Protein fusionierten PH-Domänen auf M13-Phagen (*anti-g3*-Protein-Antikörper)

Im 7.5 %-igen SDS-Gel läßt sich dem g3-Protein eine apparente Masse von 69 kDa und dem Fusionsprotein von 86 kDa zuordnen. (M: Proteinstandard; *broad range* Marker von Biorad).

4.4.1. Circularpermutante CP 3.6. in ersten *phage-display*-Versuchen

Die Selektion der in drei Loops randomisierten Wildtyp PH-Domäne mit dem *phage-display*-Verfahren führte unabhängig von der gewählten Zielsubstanz zu Deletionsmutanten und Mutanten mit *amber*-(*stop*-)Codons (s. Tab 4.3.). Eine mögliche Ursache dafür kann in der Topologie der Wildtyp PH-Domäne liegen, die zu einer schlechten Translokation führt.

Um das Auftreten von Deletionsmutanten bei randomisierten Wildtyp PH-Domänen und randomisierten Circularpermutanten zu untersuchen, wurden vier weitere g3-Protein fusionierte Bibliotheken mit einer Komplexität von jeweils über $5 \cdot 10^6$ Mutanten hergestellt. Für zwei Bibliotheken wurde aus der *phage-display* Bibliothek von Kapitel 4.1.3. ein randomisiertes DNA-Fragment von Loop 1 und 3 nach Restriktion isoliert in das PH-g3-Konstrukt bzw. in das CP 3.6.-g3-Konstrukt kloniert. In zwei weiteren Bibliotheken wurden beide Konstrukte einer mutagenen PCR-Reaktion unterzogen (s. Kapitel 3.4.1.) und das über die gesamte PH-Domäne bzw. CP 3.6. randomisierte PCR-Fragment in die Konstrukte kloniert. Diese vier Bibliotheken wurden in einem *phage-display*-Verfahren eingesetzt, bei dem auf eine *in vitro*-Selektion durch Affinität zu Zielsubstanzen („*panning*“) verzichtet und damit nur auf Wachstum selektiert wurde. Statt eines *panning* wurden die hergestellten Phagen verdünnt und mit ihnen kompetente *E. coli*-Zellen infiziert (s. Kapitel 1.5.1.). Eine Restriktions-analyse der Bibliotheken (wie in Abb. 4.12) zeigte, daß nach zwei “Selektionsrunden” bei allen vier Bibliotheken nur noch 20 % und nach drei “Selektionsrunden“ nur noch weniger als 5 % der ursprünglich eingesetzten (randomisierten) Konstrukte vorhanden sind. Eine unterschiedliche Anreicherung der Wildtyp PH-Domäne oder der Circularpermutante CP 3.6. erfolgte nicht. Auf eine Fortführung dieser Experimente durch Sequenzierung von Einzelklonen (wie in Kapitel 4.2.2.) wurde verzichtet. Dieses Experiment grenzt die möglichen Ursachen für die Erzeugung von Deletionsmutanten beim *phage-display*-Verfahren mit der randomisierten Wildtyp PH-Domäne ein: die Schwierigkeiten bei der Translokation führen nicht direkt zu Deletionsmutanten und Phagen mit deletierten PH-Domänen besitzen gegenüber Phagen mit regulären (Wildtyp oder circularpermutierten) PH-Domänen einen Wachstumsvorteil.

Im Rahmen von *phage-display*-Verfahren mit *panning* sollten sich die Circularpermutanten dennoch eher bewähren als der Wildtyp, da sie sich besser auf der Phagenoberfläche darstellen lassen (s. Abb. 4.27). Somit tragen mehr Phagen jeweils mindestens eine Kopie der jeweiligen Mutante und bei einer *in vitro*-Selektion können statistisch eher Aviditätseffekte auftreten, wenn mehrere

Kopien einer schwach bindenden Mutante einem (individuellen) Phagen zu hoher Affinität an einer Zielsubstanz verhelfen. Bei einer weiteren Ursachenforschung für die Deletionsmutanten sollte man ein bewährtes, vergleichbares *phage-display*-Protein verwenden, im obigen Experiment einsetzen und auf Deletionsmutanten hin untersuchen. Erst dies zeigt, ob die Erzeugung von Deletionsmutanten von der Wahl des randomisierten Proteins oder aber z.B. von der Art der Randomisierung abhängig ist.

4.5. Der Import der PH-Domäne in Mitochondrien

Die Mutanten der PH-Domäne werden sehr unterschiedlich effizient in das Periplasma von *E. coli* transloziert bzw. periplasmatisch exprimiert. Ob sie sich auch bei anderen Transportvorgängen voneinander unterscheiden, wurde anhand des Imports einiger Mutanten in die mitochondriale Matrix von *Saccharomyces cerevisiae* untersucht.

Die vektoriellen Konstrukte für die Import-Experimente wurden im Rahmen dieser Arbeit kloniert und am Lehrstuhl für Physiologische Chemie der LMU München eingesetzt.

4.5.1. Grundlagen

Für den Transport von Proteinen aus dem Cytoplasma in die mitochondriale Matrix der Hefen *Neurospora crassa* und *Saccharomyces cerevisiae* gibt es mindestens zwei Transportmechanismen. Mit der Wahl des Signalpeptides pSu9 von der mitochondrialen F_0 -ATPase (Untereinheit 9) ist der Transportweg über den TOM-Komplex durch die äußere Membran und den TIM-23-Komplex durch die innere Membran in die mitochondriale Matrix festgelegt (s. Kapitel 1.6.2. und Gaume *et al.*, 1998). Das Signalpeptid pSu9 besteht aus 66 Aminosäuren (s. Abb. 4.28). Durch ein zweites *start*-Codon in Position 15 entstehen bei der *in vitro* Translation zwei Proteine unterschiedlicher Länge, wobei nur das Vollängentranslationsprodukt zu einem mitochondrialen Signalpeptid führt. Das Signalpeptid kann von der mitochondrialen Peptidase MPP (*matrix processing peptidase*) hinter Position 35 und hinter Position 66 prozessiert werden. Proteine, die bei 35 (hinter *MPP1*) prozessiert sind, werden als Intermediärprodukte und Proteine, die nach 66 (hinter *MPP2*) prozessiert sind, als reife Proteine bezeichnet.

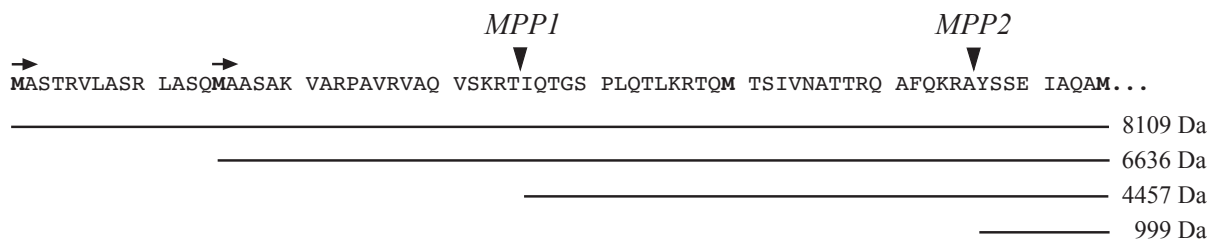


Abb. 4.28: Aminosäuresequenz und Prozessierungsstellen des mitochondrialen Signalpeptids pSu9.

Die beiden Translationsstarts sind mit einem Pfeil und die Prozessierungsstellen *MPP1* und *MPP2* für die mitochondriale Peptidase MPP mit einem Keil gekennzeichnet. Die radioaktive Markierung erfolgt durch Inkorporation von [³⁵S] an den fett gekennzeichneten Methioninresten. Zusätzlich sind die beiden Translations- sowie die beiden Prozessierungsprodukte mit jeweils einer Linie wiedergegeben und die molekularen Massen dieser virtuellen Produkte hinzugefügt. Die Aminosäuren 1 bis 74 (pSu9) wurden an die PH-Domäne fusioniert. Sie beinhalten das eigentliche Signalpeptid (1 bis 66).

Nach *in vitro* Transkription mit SP6-RNA-Polymerase und *in vitro* Translation mit Kaninchen Reticulocyten-Lysat (Pelham & Jackson, 1976) unter [³⁵S]-Methionin-Inkorporation liegen die Fusionsproteine als Vollängenprodukte vor. Der Import der Fusionsproteine erfolgt, indem das Fusionsprotein im Reticulocyten-Lysat unter Zugabe von ATP, NADH, Phosphokreatinin *etc.* (vgl. Gaume *et al.*, 1998) bei einer definierten Temperatur (hier 25°C) über eine bestimmte Zeit zu isolierten Mitochondrien (Daum *et al.*, 1982) gegeben wird. Nach der Inkubation wird das Membranpotential durch Zugabe von Valinomycin zerstört, was zu einer Arretierung des Imports führt. Jeder Ansatz wird in drei Aliquots geteilt. Ein Aliquot wird mit eiskaltem Puffer isotonisch verdünnt und auf Eis aufbewahrt. Ein zweites Aliquot wird mit isotonischem Puffer verdünnt und mit ProteinaseK auf Eis behandelt. Die ProteinaseK hydrolysiert dabei die Proteinbestandteile, die nicht durch die äußere Membran der Mitochondrien geschützt werden. Das dritte Aliquot wird mit ProteinaseK in einem hypotonischen Puffer auf Eis inkubiert. Dabei schwillt die Matrix, die äußere Membran platzt und der Intermembranraum wird für die ProteinaseK zugänglich. Mitochondrien ohne äußere Membran werden als Mitoplasten bezeichnet. Die Mitochondrien bzw. Mitoplasten werden daraufhin durch Zentrifugation reisoliert und die Proteine über ein denaturierendes SDS-Gel aufgetrennt. Die *in vitro* translatierten Proteine und ihre Prozessierungsprodukte lassen sich mittels Autoradiographie auf einem Western-Blot nachweisen. Ähnliche Versuche sind bereits veröffentlicht (Gaume *et al.*, 1998). Diese Experimente wurden von Maxi Endres durchgeführt und sind deshalb nicht im Methodenteil dieser Arbeit behandelt.

4.5.2. Ergebnisse des mitochondrialen Imports von PH-Domänen

Alle vier untersuchten Mutanten der PH-Domäne wurden in die mitochondriale Matrix importiert und lassen sich dort als reife Proteine (Prozessierung an *MPP2*) im Autoradiogramm nachweisen (16 kDa) (s. Abb. 4.29). Man beobachtet stets bei den unbehandelten Mitochondrien das unprozessierte Protein mit regulärem *start*-Codon (24.5 kDa) und das unprozessierte Protein mit zusätzlichem *start*-Codon (23 kDa). Die Wildtyp PH-Domäne zeigte das gleiche Importverhalten bzw. Bandenmuster wie die destabilisierte Mutante L(8)V: Neben dem reifen Protein erschien im Autoradiogramm nach Inkubation der Mitoplasten mit ProteinaseK ein Fragment (a) mit der ungefähren apparenten Masse von 10 kDa. Bei der Circularpermutante CP 2.7. zeigte sich nach derselben Behandlung neben dem reifen Protein ein Fragment (b) mit der apparenten Masse 12 kDa. Ein derartiges Fragment ließ sich für die Circularpermutante CP 2.6. in Mitoplasten nicht nachweisen. Bei ihr erschien stattdessen bei den unbehandelten Mitochondrien ein Intermediärprodukt (c) der apparenten Masse 20 kDa.

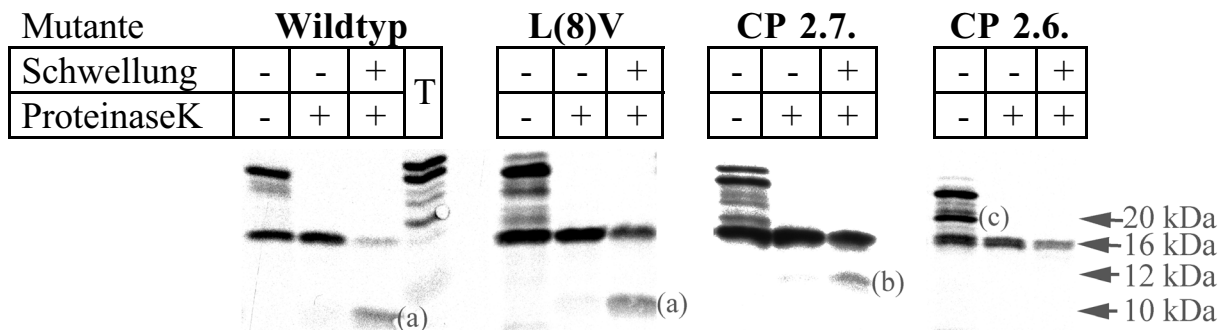


Abb. 4.29: Import der Wildtyp PH-Domäne, der destabilisierten Mutante L(8)V, sowie der Circularpermutanten CP 2.7 und CP 2.6. als Fusionsprotein mit dem pSu9-Signalpeptid in Mitochondrien.

Die Proteine wurden bei 25°C für 5 min mit Mitochondrien inkubiert. Jeweils drei Aliquots wurden untersucht, indem die Mitochondrien unbehandelt, mit ProteinaseK behandelt und nach Schwellung als Mitoplasten mit ProteinaseK inkubiert in einem SDS-PAGE aufgetrennt und mittels Autoradiogramms detektiert wurden. T ist das *in vitro* translatierte Fusionsprotein aus pSu9-Signalpeptid und der Wildtyp PH-Domäne ohne Inkubation mit Mitochondrien. (a) und (b) sind Fragmente der jeweiligen Mutante nach Inkubation der Mitoplasten mit ProteinaseK. Bei CP 2.6. wird ein derartiges Fragment nicht detektiert, dafür aber ein Intermediärprodukt im Intermembranraum nachgewiesen (c).

ProteinaseK-Behandlung von Mitochondrien nach dem Import verursachte bei keiner Mutante Fragmente im Intermembranraum oder in der Matrix, woraus geschlossen wird, daß der Import der Konstrukte in den Intermembranraum durch den TOM-Komplex ohne Unterbrechung erfolgt.

Neben dem jeweils regulär importierten und durch die MPP regulär prozessierten Präprotein (*MPP2*-Prozessierung) erscheinen im Autoradiogramm weitere Banden (a, b und c in Abb. 4.29), die beweisen, daß der Import durch die innere Membran bei allen untersuchten Mutanten mit jeweils einer Unterbrechung verläuft. Dafür verantwortlich ist eine importunterbrechende Peptidsequenz, die sich wie folgt durch die untersuchten Mutanten eingrenzen und definieren läßt. Bei der Wildtyp PH-Domäne [und bei der Mutante L(8)V] befindet sich diese Sequenz ungefähr im mittleren Bereich der Aminosäuresequenz. Gelangt diese Sequenz der PH-Domäne unmittelbar vor die innere Membran, wird der Import unterbrochen und in der Matrix prozessiert die MPP an *MPP2* das Signalpeptid. Erst nach ProteinaseK-Behandlung der Mitoplasten wird der in den Intermembranraum ragende Anteil der Mutanten (der C-Terminus ab der importunterbrechenden Peptidsequenz) proteolysiert. Der N-terminale Anteil dieser Konstrukte gelangt dann in die Matrix, wo er als Fragment (a) im Autoradiogramm nachgewiesen wird. Die gleiche Situation liegt bei der Circularpermutante CP 2.7. vor. Das beobachtete Fragment (b) ist nach ProteinaseK-Behandlung der Mitoplasten größer als das beim Wildtyp - die Sequenz befindet sich folglich am C-Terminus dieser Mutante und dieser reicht bis vor den Loop 7. Damit ist das importunterbrechende Peptid eingegrenzt zwischen dem mittleren Aminosäuresequenzbereich der PH-Domäne bis zu Loop 7.

Auch bei der Circularpermutante CP 2.6. beobachtet man eine Importunterbrechung. Allerdings verharrt das Protein nicht in der inneren Membran, wenn das importunterbrechende Peptid den weiteren Import blockiert. In dieser Situation ragt nämlich nur der N-Terminus des Signalpeptids in die Matrix und wird dort von der MPP an *MPP1* prozessiert. Das restliche mit CP 2.6. fusionierte Signalpeptid verhindert nicht mehr die Rückdiffusion und CP 2.6. gelangt als Intermediärprodukt in den Intermembranraum - dort verbleibt es, da das *MPP1*-prozessierte Signalpeptid nicht mehr funktional ist. Dieser Vorgang ist nur möglich, weil sich die importunterbrechende Peptidsequenz bei CP 2.6. im N-terminalen Bereich der Mutante befindet - und dieser beginnt mit Loop 6.

Faßt man die Beobachtungen an allen vier Mutanten der PH-Domäne zusammen, liegt das importunterbrechende Peptid genau zwischen dem

N-Terminus von CP 2.6. und dem C-Terminus von CP 2.7. bzw. zwischen den Loops 6 und 7. Es ist damit sicher in der Aminosäuresequenz

74 100
-DQVIKAAKTEADGQVVEGNHTVYRISA-

enthalten.

In Abb. 4.30A sind die im Autoradiogramm (s. Abb. 4.29) beobachteten Fragmente der Mutanten erklärt. Die Auswirkungen der Importunterbrechung durch das Peptid D(74) bis A(100) auf den mitochondrialen Import der einzelnen Mutanten sind in Abb. 4.30B skizziert.

4. Ergebnisse

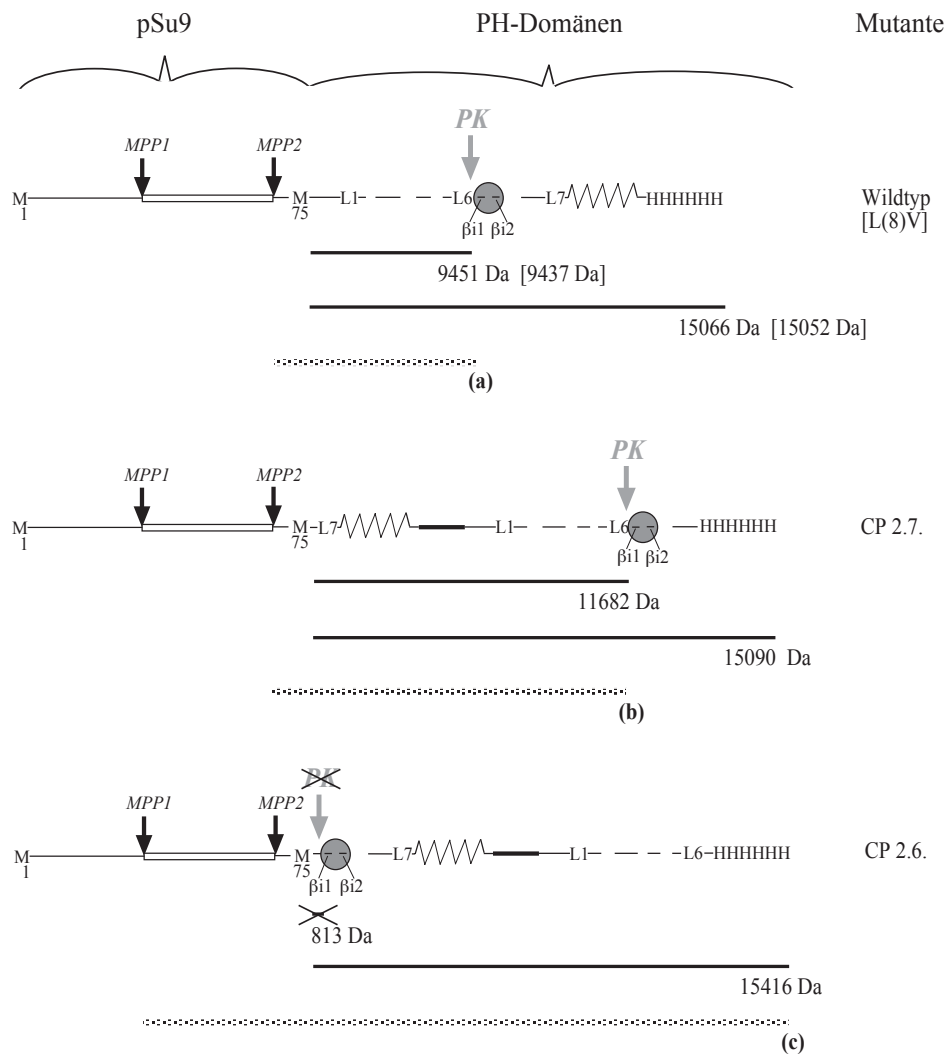


Abb. 4.30A: Prozessierung des C-Terminus von pSu9-fusionierten PH-Domänen durch ProteinaseK an der inneren mitochondrialen Membran.

Durch die Unterbrechung des Imports ragt die Peptidsequenz verschiedener PH-Domänen nach Loop 6 (hier berechnet ab Aminosäure 80 der PH-Domäne) aus der inneren Membran in den Intermembranraum hinein und lässt sich nach Deletion der äußeren Membran durch Schwellung mit ProteinaseK (*PK*) proteolysieren. Über Autoradiographie beobachtet man jeweils neben dem reifen Protein beim Wildtyp das Fragment (a) und bei CP 2.7. das Fragment (b) in der Matrix (s. Abb. 4.29). Ein derartiges Fragment findet man bei CP 2.6. nicht; stattdessen lässt sich bei CP 2.6. im Intermembranraum ein *MPP1*-prozessiertes Intermediärprodukt nachweisen (c). Die beobachteten Banden setzen sich aus dem jeweiligen Fragment der PH-Domäne (schwarze Linie unter der stilisierten PH-Domäne mit Angabe der Molmasse) bzw. der vollständigen PH-Domäne und dem durch *MPP* prozessierten Anteil von pSu9 zusammen (s. Abb. 4.28). Zur Stilisierung der PH-Domänen sind die β -Faltblätter mit dünnen Strichen, die α -Helix mit einer Sägezahnlinie und der Linker mit einer dicken Linie eingezeichnet. Zur Orientierung sind zudem der N-Terminus von pSu9 (M1), der N-Terminus der fusionierten PH-Domäne (M75) und die Loops 1, 6 und 7 (L1, L6, L7) und der His-tag beschriftet. Im Bereich der beiden β -Faltblätter β i1 und β i2 findet die Unterbrechung des Imports statt, die durch eine graue Kreisfläche gekennzeichnet ist.

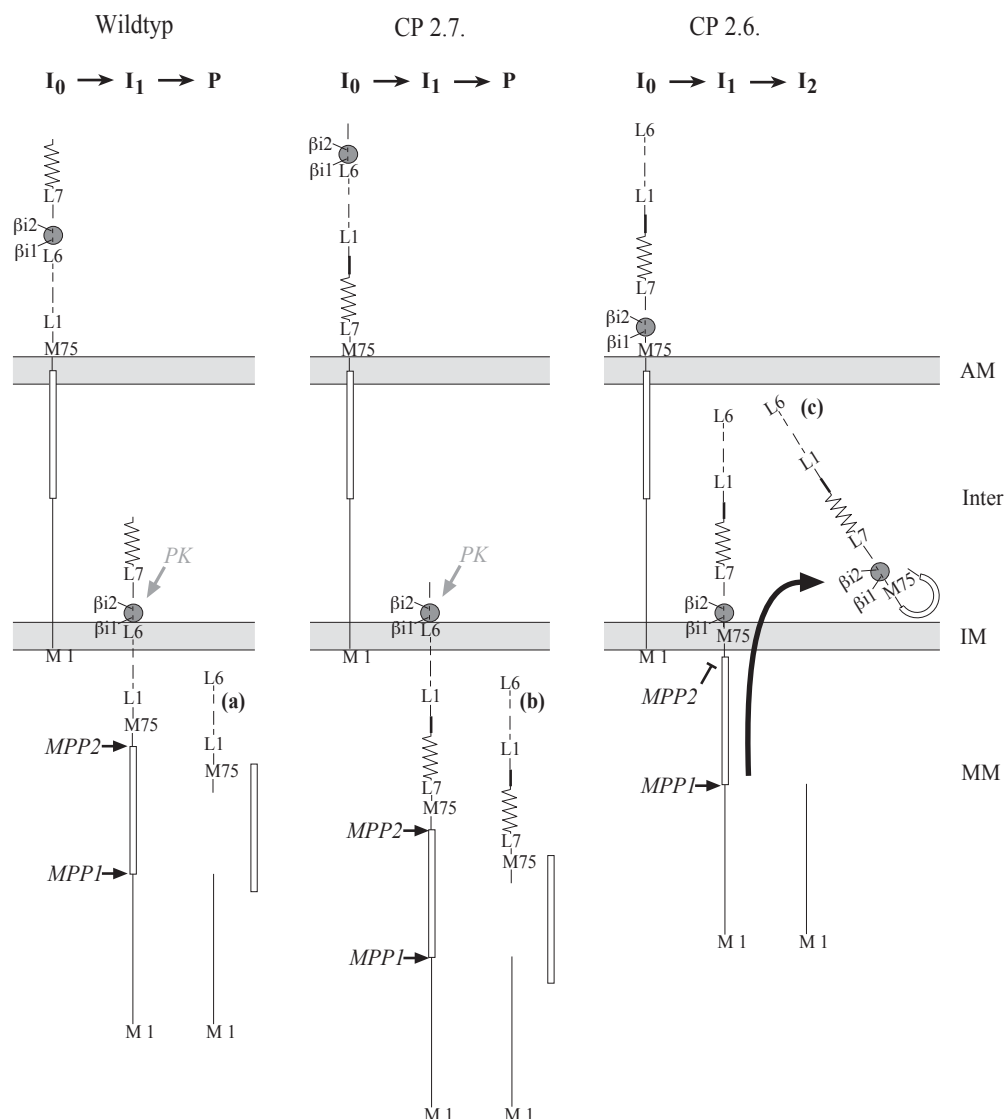


Abb. 4.30B: Import der Wildtyp PH-Domäne sowie der Circularpermutanten CP 2.7. und CP 2.6. als Fusionsprotein mit dem pSu9-Signalpeptid in Mitochondrien.

Im Zustand I_0 durchspannt das pSu9-Signalpeptid den Intermembranraum (Inter). Dabei ist der N-Terminus schon in der mitochondrialen Matrix (MM), während sich die verschiedenen, fusionierten Mutanten der PH-Domäne noch außerhalb der Mitochondrien befinden. Die Mutanten sind wie unter Abb. 4.30B stilisiert. Der Import wird im Zustand I_1 im Bereich der β -Faltblätter $\beta 1$ und $\beta 2$ (graue Kreisfläche) bei allen Mutanten unterbrochen. Das Signalpeptid wird bei Wildtyp und CP 2.7. an *MPP1* und *MPP2* prozessiert und die beiden Proteine verharren in diesem metastabilen Transportzustand. Nach Deletion der Außenmembran (AM) und nachfolgender Prozessierung (P) mit ProteinaseK (*PK*) wird der jeweilige in den Intermembranraum ragende C-Terminus prozessiert. In der Matrix lassen sich dann die Fragmente (a) und (b) nachweisen. Bei CP 2.6. dagegen wird die Prozessierung an *MPP2* zum reifen Protein verhindert, indem die Prozessierungsstelle durch die innere Membran (IM) bzw. dort assoziierte Proteine geschützt wird. Das *MPP1* prozessierte Protein diffundiert in einem Folgeschritt (I_2) zurück in den Intermembranraum und lässt sich dort als stabiles Intermediärprodukt (c) nachweisen.

5. Diskussion

5.1. Wege zur funktionellen Veränderung von Proteinen

Zur Veränderung von Proteinen in ihrer Sequenz, Struktur und Eigenschaft durch Mutagenese gibt es methodisch zwei prinzipiell unterschiedliche Ansätze:

- Beim rationalen Design werden dem Protein gezielt Änderungen beigefügt. Ideen dafür können auf statistischen Beobachtungen beruhen, wie dies z.B. mit Hilfe von Strukturvergleichs-Programmen geschieht, die in der Arbeitsgruppe entwickelt wurden (*one2all* und *one2all gap*, beide unveröffentlicht). Die funktionelle und/oder strukturelle Abhängigkeit eines Proteins von einer bestimmten Aminosäureposition läßt sich mit einer weiteren rationalen Vorgehensweise analysieren, indem diese Aminosäure mutiert und die daraus resultierenden Auswirkungen auf das Protein untersucht werden.
- Der andere häufig gewählte Ansatz zur Veränderung von Proteinen besteht darin, das Protein zufällig zu mutieren (randomisieren), damit eine Bibliothek zu erhalten und aus dieser genau das Protein herauszuselektieren, das den gewünschten Eigenschaften entspricht („*survival of the fittest*“).

Diese beiden methodischen Ansätze zur Veränderungen von Proteinen werden in der Fachwelt ähnlich dogmatisch unterschieden, wie auch die biblische Schöpfungsgeschichte (entspräche dem rationalen Proteindesign) und die Darwinsche Evolutionstheorie (entspräche den evolutiven Verfahren) zu einem „Glaubenskrieg“ führten.

Während aber aus heutiger Sicht die Evolutionstheorie fast überall anerkannt und die Schöpfungsgeschichte eher als Parabel denn als wörtliche Entstehung der Arten gedeutet wird, ist eine Anwendung der beiden methodischen Ansätze zur künstlichen Veränderung von Proteinen durchaus möglich und beide wurden in dieser Arbeit eingesetzt:

- Zur Verbesserung der Translokation wurden gezielt Mutanten der PH-Domäne mit diversen neuen, physikochemischen Eigenschaften hergestellt und einzeln biochemisch untersucht; es fanden rationale Verfahren Anwendung.

- Zur Etablierung der PH-Domäne als Paratop wurde sie in drei Loops randomisiert, die so geschaffene Bibliothek validiert und einem evolutiven Selektionsverfahren unterzogen - dem *phage-display*-Verfahren.

Der rationale methodische Ansatz erbrachte den gewünschten Erfolg, indem eine Circularpermutante mit über 30-fach verbesserter periplasmatischer Ausbeute in *E. coli* exprimiert werden konnte. Dagegen lieferte der evolutive Ansatz bislang noch keine greifbaren Ergebnisse, da noch keine bindenden Mutanten aus der Bibliothek von PH-Domänen selektiert werden konnten. Im übertragenen Sinn war also in dieser Arbeit bislang die „Schöpfungsmethode“ der „evolutiven Methode“ überlegen.

In Zukunft kann die Circularpermutante randomisiert und im *phage-display*-Verfahren eingesetzt werden. Mit ihrer verbesserten Translokationseigenschaft könnte sie sich für dieses Selektionsverfahren sehr viel besser eignen als der Wildtyp und dann auch dem evolutiven Verfahren zum Erfolg verhelfen, indem bindende (Circularper-)Mutanten aus einer Bibliothek selektiert werden. Beide prinzipiellen Ansätze zur funktionellen Änderung der PH-Domäne waren dann für dieses Ergebnis notwendig.

5.2. Die Randomisierung der PH-Domäne

5.2.1. Statistische Betrachtung von Randomisierungen

Möchte man an fünf Positionen einer Polypeptidkette unabhängig voneinander jeweils alle 20 Aminosäuren einsetzen, gibt es genau $N = 20^5 = 3.2 \cdot 10^6$ mögliche unterschiedliche Polypeptidketten. Bei dieser Größenordnung scheidet eine individuelle Herstellung aller Mutanten aus, weshalb auf einen Randomisierungsansatz zurückgegriffen wird. Es stellen sich die Fragen

- wie viele Mutanten müssen in einer Bibliothek hergestellt werden, um mit hoher Wahrscheinlichkeit sicherzustellen, daß jede der $3.2 \cdot 10^6$ möglichen Mutanten mindestens einmal in ihr vorhanden ist?
- wie viele Mutanten müssen durch Randomisierung für eine Bibliothek erzeugt werden, um mit hoher Wahrscheinlichkeit sicher zu stellen, daß bis auf wenige alle der $3.2 \cdot 10^6$ möglichen Mutanten mindestens einmal in ihr vorhanden sind?

Bei der Berechnung wurde vorausgesetzt, daß alle zwanzig verschiedenen Aminosäuren an allen fünf Positionen gleichverteilt vorliegen. Dies ist durch Einsatz von dNTP in definierten Verhältnissen nicht möglich, ließe sich aber möglicherweise über den Einsatz von Tridesoxynukleotiden bei der Oligodesoxynukleotidsynthese erreichen (Knappik *et al.*, 2000).

Die Rechnung ist im Anhang dieser Arbeit wiedergegeben (s. Kapitel 8.1.). Sie besagt, daß bei einer erzeugten Bibliotheksgröße von $6 \cdot 10^7$ (wie in dieser Arbeit erzielt) mit einer Wahrscheinlichkeit von mehr als 97.7 % alle der $3.2 \cdot 10^6$ möglichen Peptidketten in der Bibliothek mindestens einmal vertreten sind und mit einer Wahrscheinlichkeit von mehr als 99.97 % alle Peptidketten vorhanden sind und dabei höchstens eine einzige fehlt.

Die zusätzliche Randomisierung einer sechsten Position führt zu $6.4 \cdot 10^7$ möglichen Mutanten, die sich nicht mit einer Bibliotheksgröße von $6 \cdot 10^7$ Klonen „vollständig“ repräsentieren ließen. Der Bibliotheksgröße ist beim Einsatz im *phage-display*-Verfahren durch die notwendige anfängliche Elektroporation in *E. coli*-Zellen eine Grenze von ca. 10^8 Klonen gesetzt (s. Kapitel 1.5.1.). Sie gilt nicht für das zellfreie Ribosomen-Display Verfahren, in dem aus ca. 10^{12} Ribosomen-Konstrukten selektiert werden kann (s. Kapitel 1.5.2.). Acht Positionen ließen sich dann vollständig randomisieren und unter idealen technischen Bedingungen sollten sich mit hoher Wahrscheinlichkeit dann fast alle der $2.6 \cdot 10^{10}$ möglichen Mutanten in dieser Bibliothek befinden.

Wenn bei den meisten Randomisierungen dennoch mehr als fünf bzw. acht Positionen vollständig randomisiert werden, gehen die Experimentatoren davon aus, daß die meisten randomisierten Aminosäuren eine spätere Bindung ohnehin nicht maßgeblich beeinflussen. Zudem läßt sich im vorhinein nicht abschätzen, welche genauen Positionen zur Erzeugung von Affinität (zu einer bestimmten Zielsubstanz) variiert werden sollten und welche nur einen geringen Einfluß auf spätere Bindungen besitzen.

In diesem Kapitel wurden bislang nur die Auswirkungen von Randomisierungen auf die Erzeugung von Affinitäten betrachtet. Eine weitere Konsequenz von Randomisierungen kann aber zusätzlich in der Aufhebung von Wechselwirkungen beruhen, wie sie im Wildtyp für die Proteinfaltung notwendig waren, oder der zufällige Einbau disruptiver Mutationen unterbindet direkt die korrekte Faltung der entsprechenden Mutante. Je geringer die Randomisierung ausfällt, desto höher ist die Wahrscheinlichkeit, strukturierte Mutanten zu erhalten, aus denen selektiert werden kann, wenn auch in dieser Arbeit besonders variable (bei verschiedenen PH-Domänen) Loops zur Randomisierung ausgewählt wurden (s. Kapitel 1.4.1.), die deshalb nicht essentiell zur Strukturbildung der PH-Domäne beitragen sollten.

Dieser Spagat zwischen starker Randomisierung zur Erzeugung von Affinitäten und geringer Randomisierung zur Bewahrung der strukturellen Eigenschaften wurde durch die Randomisierung mit Hilfe des Computerprogramms „Mutator“ gemeistert (Steipe & Bruhn, in Vorbereitung). Dabei werden sehr viele Positionen in ihren Aminosäureresten variiert, aber an jeder Position die Wildtyp- oder eine physikochemisch ähnliche Aminosäure in der Bibliothek bevorzugt. Beim Großteil der randomisierten Mutanten bleiben so noch die möglichen lokalen und nicht-lokalen Wechselwirkungen des Wildtyps erhalten und können zur Stabilität des gesamten Proteins beitragen. Dennoch sind an allen randomisierten Positionen der Bibliothek alle Aminosäuren vertreten und können zu einer Bindung nach Selektion führen. Dieses Verfahren eignet sich ganz besonders für Proteine, deren genaue Struktur nicht bekannt ist, weshalb sich die Wechselwirkungen in den randomisierten Bereichen nicht detailliert abschätzen lassen.

5.2.2. Randomisierung mit „Mutator“ und ihre Umsetzung zu einer Bibliothek randomisierter PH-Domänen

Zur Randomisierung von Aminosäuren in Proteinen werden die entsprechenden DNA-Codons variiert, indem bei der Oligodesoxynukleotidsynthese an jeder Position die Basen (bzw. deren Phosphoramidid-Reagentien) in unterschiedlichen Konzentrationsverhältnissen eingesetzt werden (s.u. und als Übersichtsartikel: Steipe, 1999). Das Computerprogramm „Mutator“ berechnet zu jedem Mischungsverhältnis von Basen in den randomisierten Codons die daraus resultierende Aminosäureverteilung in der Bibliothek an dieser Position. Umgekehrt kann man auch eine gewünschte Gewichtung der Aminosäureverteilung an einer Position (bzw. Codon) vorgeben und die dazu benötigten Konzentrationsverhältnisse an Basen zur Oligodesoxynukleotidsynthese berechnen lassen. Somit läßt sich für jede individuelle Aminosäureposition eine individuelle Verteilung an Aminosäureresten bei der Oligodesoxynukleotidsynthese einstellen.

Der Planung von Oligodesoxynukleotiden mit „Mutator“ und ihrer Synthese folgt die Klonierung zu einer entsprechenden Bibliothek. Die drei randomisierten Loops befinden sich zwar bei der PH-Domäne in unmittelbarer räumlicher Nähe und bilden ein diskontinuierliches Paratop aus (s. Abb. 1.2), doch sind sie auf der DNA so weit voneinander entfernt, daß sie mit jeweils eigenen Oligodesoxynukleotiden randomisiert wurden. Erst die Entwicklung von speziellen PCR-Techniken („asymmetrische PCR“ und „asymmetrische Verlängerungs-PCR“, s. Kapitel 4.1.3.) erlaubte die Herstellung getrennter, randomisierter, isolierter und ligierbarer DNA-Segmente für die jeweiligen Loopregionen.

Diese wurden in einem Assemblierungsverfahren zu einem Fragment ligiert und dann gemeinsam kloniert. Durch die getrennte Randomisierung der DNA-Segmente und den Verzicht auf eine abschließende Amplifikation des in drei Segmenten randomisierten Fragments (zumindest bei über 80 % der Klone aus der Bibliothek) wurden die Loops unabhängig voneinander randomisiert, so daß nicht mit einem Verlust an Diversität durch Überrepräsentierung einzelner Mutanten in der Bibliothek zu rechnen ist. Durch Optimierung der Randomisierungsstrategie und der Verwendung hochkompetenter *E. coli*-Zellen für die Elektroporation konnte eine Bibliothek von $6 \cdot 10^7$ randomisierten PH-Domänen hergestellt werden. Diese Größenordnung führte beim Einsatz anderer Bibliotheken zu erfolgreichen Selektionen: bei der Glutathion-S-Transferase genügte dazu $5 \cdot 10^6$ (Widersten & Mannervik, 1995) Klone, während die Bibliothek von „Anticalinen“ sogar $3.7 \cdot 10^8$ Klone beträgt (Beste *et al.*, 1999). Im kommerziellen Rahmen gibt es auch größere Antikörper-Bibliotheken für *phage-display*-Verfahren wie HuCal (*human combinatorial antibody library*: Knappik *et al.*, 2000) mit $2 \cdot 10^9$ Klonen. Zumindest die Zahl randomisierter Mutanten sollte bei der PH-Domäne somit für eine erfolgreiche Selektion ausreichen.

Mit den durch „Mutator“ berechneten Daten und der in Kapitel 4.1.3. aufgeführten Klonierungsstrategie wurden 29 Aminosäurepositionen in drei Loops der PH-Domäne randomisiert. Für die Oligodesoxynukleotidsynthese (s. Kapitel 2.2.2.) wurden durchschnittlich 16 randomisierte (Codons) Aminosäuren über drei Loops vorgegeben, d.h. durchschnittlich sollten etwa 55 % der Aminosäuren im randomisierten Bereich gegenüber dem Wildtyp verändert sein. Die statistische Auswertung von 29 randomisierten und exprimierbaren Mutanten zeigte bei 17 % jeweils aller drei Loops keine Randomisierung und bei den restlichen 83 % der Loops eine durchschnittliche Randomisierung von 32 % der Aminosäurepositionen. Insgesamt wurden durchschnittlich sieben Aminosäuren der Wildtyp PH-Domäne durch andere Aminosäuren ersetzt (s. Kapitel 4.1.4.). Die Abweichung von den erwarteten Werten beruhen wahrscheinlich auf der Amplifikation von Wildtyp (Templat-) DNA bei den PCR-Reaktionen und auf der präferentiellen Bevorzugung randomisierter Oligodesoxynukleotide mit „wildtyp-ähnlicher“ DNA-Sequenz ebenfalls bei den PCR-Reaktionen. Durch geringfügige Änderung der PCR-Technik sollten sich diese Abweichungen in Zukunft beheben lassen, indem die als Templat eingesetzte DNA nicht mehr die zu randomisierenden Segmente enthält. Dann sollte die PH-Domäne erwartungsgemäß randomisiert werden können. Die Festlegung auf ein bestimmtes Randomisierungsausmaß war für die Oligodesoxynukleotidsynthese notwendig. Sie basierte auf den Vermutungen, daß einige Mutationen in den drei Loops die Struktur und Exprimierbarkeit der

PH-Domäne nur geringfügig beeinträchtigen und trotzdem zu selektierbaren Mutanten führen.

Die Validierung der Bibliothek zeigt (s. Kapitel 4.1.4.), daß sich 31 der 50 willkürlich ausgewählten Mutanten aus der Bibliothek löslich cytoplasmatisch exprimieren lassen. Bei 19 Mutanten wurde kein lösliches Protein nachgewiesen. Vier davon sequenzierte Mutanten weisen Verschiebungen im Leserahmen gegenüber dem Wildtyp auf, die sich auf die PCR-Amplifikationen zurückführen lassen. Damit konnte nachgewiesen werden, daß alle untersuchten, randomisierten Mutanten ohne genetische Defekte auch exprimiert werden. Die Stabilität der PH-Domäne wird durch zufällig eingeführte Mutation in den Loopregionen um ca. 1 bis 10 kJ/mol reduziert, wobei sich auch eine Mutante (Klon 50) mit der Freien Faltungsenthalpie von -11 kJ/mol (also um über 26 kJ/mol gegenüber dem Wildtyp reduziert) gefaltet cytoplasmatisch exprimieren läßt (s. Kapitel 4.1.6.). Die Randomisierung führt zwar zu einer Destabilisierung der untersuchten Mutanten, erlaubt aber deren Expression. Die Stabilität wird damit den Erwartungen an randomisierte Proteine gerecht und das Paratopgerüst der PH-Domäne erweist sich als sehr robust.

Die erzeugte Bibliothek randomisierter PH-Domänen besitzt nach der Klonierung eine Komplexität von $6 \cdot 10^7$ Klonen, von denen sich über 60 % löslich exprimieren lassen. Ein Vergleich dieser Angabe mit anderen Bibliotheken (s. Kapitel 1.4.2.) ist mangels Literaturdaten nicht möglich, da dort stets nur die Größe, nie aber die detaillierten Ergebnisse der Validierung aufgeführt sind.

Da sich auch Mutanten mit deutlich verringerter Stabilität im Cytoplasma exprimieren lassen (Klon 50), könnte man versuchen, die Bibliothek stärker zu randomisieren, um sie anschließend in Selektionsverfahren einzusetzen. Eine solche stärkere Randomisierung könnte in einer Modifizierung der Klonierungsstrategie liegen (s.o.) oder aber in Design und Synthese neuer Oligodesoxynukleotide, für die ein höheres Randomisierungsausmaß gewählt wird. Da sich bei „Mutator“ die Randomisierung jedes Codons individuell vorgeben läßt, wäre eine vergleichsweise starke Randomisierung in Loop 6 angebracht, bei dem sich sogar 9 bzw. 7 Aminosäuren entfernen lassen, ohne daß die PH-Domäne ihre cytoplasmatische Exprimierbarkeit verliert (Klon 28 bzw. 30, s. Kapitel 4.1.4.).

Die „maßgeschneiderte“ und regulierbare Randomisierung von Codons durch „Mutator“ erlaubt eine „vorsichtige“ Randomisierung, bei der teilweise Wildtyp-Aminosäuren im randomisierten Bereich erhalten bleiben und das Protein stabilisieren können. Soll ein Codon bzw. die entsprechende Aminosäure rigoros randomisiert werden, bietet sich bei der Oligodesoxynukleotidsynthese der Einsatz äquimolarer Mengen an Basen an (bzw. deren Phosphoramidid-

Reagentien). Der einfachste Ansatz wäre ein äquimolarer Einsatz aller vier Basen. Diese Mischung bei der Synthese wird als „N“ bezeichnet. Eine Bibliothek mit einem derartig randomisierten (N)(N)(N)-Codon besitzt an dieser Codonposition alle Codon-Varianten in gleichem Verhältnis (Oliphant *et al.*, 1986). Die Gleichverteilung aller Codon-Varianten führt allerdings zu einer späteren Überrepräsentierung der Aminosäuren, die stärker degeneriert sind: so existieren für Serin sechs unterschiedliche Codons gegenüber Methionin, für das nur ein Codon kodiert. Außerdem beträgt die Wahrscheinlichkeit für ein *stop*-Codon $3/64 = 4.7\%$, was bei Randomisierung mehrerer Codons zu häufigen Kettenabbrüchen auf Peptidebene führt. Wird anstelle des äquimolaren Einsatzes aller vier Basen an der dritten Position auf den Einsatz von A verzichtet (also nur C, G und T, genannt „B“), entfallen die beiden *stop*-Codons *ochre* und *opal*. Gleichzeitig ist die Verteilung der Aminosäuren ausgeglichener. Eine häufige Randomisierungsstrategie ist deshalb die Codonwahl (N)(N)(B) (Kay *et al.*, 1993). Auch die noch weiter einschränkenden Codonwahlen (N)(N)(S) und (N)(N)(K) werden eingesetzt [mit „S“ für G und C sowie „K“ für G und T] (Beste *et al.*, 1999; Devlin *et al.*, 1990). Beim Einsatz derartiger Randomisierungen von Codons kann nur durch die Wahl der randomisierten Aminosäureposition direkt Einfluß auf die Randomisierung eines Proteins genommen werden. Die Randomisierung einer (unerwartet) strukturell bedeutsamen Aminosäure verursacht bei diesen Strategien den Funktionsverlust der gesamten Bibliothek, während sich dies bei „Mutator“ nur im Ausfall weniger Mutanten äußert. Dies kann sich gerade bei der Randomisierung strukturell unaufgeklärter Proteine, wie es die PH-Domäne war, entscheidend auswirken.

Bei den genannten Verfahren zur Synthese randomisierter Oligodesoxynukleotide (einschließlich „Mutator“) kommen die Basen an jeder Basenposition des Codons in einem definierten Verhältnis zum Einsatz. Indem anstelle der Basen die kodierenden Desoxytrinukleotide in der randomisierenden Synthese verwendet werden, kann direkt das Verhältnis der Aminosäuren eingestellt werden. Dieses Verfahren wurde unter anderem zur Randomisierung der HuCAL-(*scF_V*-)Bibliothek eingesetzt (s. Kapitel 1.5.1.; Knappik *et al.*, 2000). Dieses Verfahren ist allerdings sehr kostenintensiv und läßt sich nur manuell umsetzen: für die Synthese eines randomisierten Codons, müssen 20 Desoxytrinukleotid-Lösungen im zuvor festgelegten Verhältnis zueinander eingesetzt werden, während nach „Mutator“ nur die üblichen vier Basen-Phosphoramidid-Reagentien im festgelegten Verhältnis zueinander zum Einsatz kommen, was sich mit einem (vor drei Jahren) gebräuchlichen Synthesegerät (Ecosyn D300, Eppendorf) automatisch umsetzen ließ.

5.3. *Phage-display* mit der PH-Domäne

Die Bibliothek randomisierter PH-Domänen wurde im *phage-display*-Verfahren zur Selektion gegen die Zielsubstanzen MOMP, Kreatinin und CD3 ϵ eingesetzt (s. Kapitel 4.2.1.). Dabei fand bei allen drei Zielsubstanzen eine Anreicherung von Vollängen-PH-Domänen über die zunehmende Zahl von Selektionsrunden statt, was aus Restriktionsverdauen nach den unterschiedlichen Selektionsrunden deutlich wird (s. Abb. 4.12). Lediglich bei MOMP zeigte sich auch nach der 5. Selektionsrunde noch ein DNA-Restriktionsfragment der erwarteten Größe. Nach Klonierung dieses DNA-Fragments weisen sechs von acht untersuchten Klonen dieselbe Sequenz auf. Diese sechs Klone und die beiden weiteren Klone besitzen jeweils nur ein gemeinsames Merkmal, nämlich ein *amber*-Codon in Position 11; weitere gemeinsame randomisierte Aminosäurepositionen sind nicht vorhanden (s. Tab. 4.3). Die Selektion war folglich konvergiert. Das randomisierte Codon des Oligodesoxynukleotids *Loop1rms* (s. Kapitel 2.2.2.) für Position 11 besitzt eine prozentuale Basenverteilung (A/C/G/T) von (16/59/10/15)(6/6/8/80)(0/52/41/7) [unterstrichen sind die Basen, die zu einem *amber*-Codon führen]. Die Wahrscheinlichkeit für das *amber*-Codon (TAG) beträgt damit $0.15 \cdot 0.6 \cdot 0.41 = 0.37\%$. Da die Selektion in TG 1-Zellen stattfand und dieser *E. coli*-Stamm den Genotyp *supE* besitzt (s. Kapitel 2.1.1.), sollte bei (TAG) anstelle eines Translationsarrests ein Glutamin in die Polypeptidkette translatiert und die Peptidsynthese fortgesetzt werden. Andererseits fungiert bei Überexpression in *supE*-Stämmen das *amber*-Codon auch als *stop*-Codon. So werden laut „Gebrauchsanweisung“ für *phage-display*-Bibliotheken *scF_V*-Antikörperfragmente, die über ein *amber*-Codon an das g3-Protein fusioniert sind, ohne Induktion auf der Phagenoberfläche exprimiert, bei Überexpression (durch Induktion mit IPTG) aber ohne fusioniertes g3-Protein in das Periplasma transloziert (daraus gereinigt und charakterisiert) (Viti *et al.*, 2000). Da zur Herstellung von PH-Domänen-fusionierten Phagen die PH-Domänen-g3-Fusion durch IPTG-Induktion überexprimiert wird (s. Kapitel 3.9.2.), wird dieses *amber*-Codon auch als *stop*-Codon im *supE-E. coli*-Stamm TG 1 erkannt. Die Wahrscheinlichkeit für ein reguläres Glutamin-Codon (CAA oder CAG) an Position 11 beträgt insgesamt 1.45 %. Diese Codons sind damit statistisch viermal häufiger in der Bibliothek anzutreffen als das *amber*-Codon, so daß man bei einer Selektion gegen MOMP zu Glutamin an Position 11 auch reguläre Codons erwartet hätte. Ein *ochre*- oder *opal-stop*-Codon wird an Position 11 bei der Randomisierung ausgeschlossen, indem auf die Adeninbase in der dritten Codonposition verzichtet

wird (s.o.). Die Selektion der *phage-display*-Bibliothek konvergierte somit auch bei MOMP nicht zu bindenden Mutanten der PH-Domäne sondern zu Deletionsmutanten und zu einem *stop*-Codon in Position 11. Demzufolge besitzen diese Nonsens-Mutanten gegenüber korrekt randomisierte PH-Domänen einen Selektionsvorteil.

Die durchschnittliche (scheinbare) Affinitätsvergrößerung der phagenfusio-nierten Bibliothek von PH-Domänen über zunehmende Selektionsrundenzahl gegen alle untersuchten Zielsubstanzen (bestimmt über den Phagentiter der letzten Waschlösung in Relation zu dem des Eluats: s. Abb. 4.11) läßt sich als die Anreicherung von Deletionsmutanten mit deutlichem Faltungsdefekt deuten. Durch die Repräsentierung der hydrophoben, sonst in das Proteininnere orientierten Aminosäurereste treten möglicherweise unspezifische Bindungen zu der hydrophoben, wenn auch mit „RotiBlock“ abgesättigten Oberfläche der NUNC-Röhrchen auf.

In der Literatur werden nur verhältnismäßig selten neue, erfolgreiche *phage-display*-Selektionen mit anderen randomisierten Gerüstproteinen veröffentlicht. Dies könnte im seltenen Auffinden bindender Mutanten und der aufwendigen Optimierung der Bibliotheken begründet liegen: So wurde das Bilin-bindende Lipocalin 1994 als Paratopgerüst vorgeschlagen (Schmidt & Skerra, 1994) und 1998 die Optimierungen für die periplasmatische Expression vorgestellt (Schmidt *et al.*, 1998). Seitdem konnten lediglich zwei Gruppen bindender Lipocalinmutanten publiziert werden: 1999 „Anticaline“ gegen Fluoreszein (Beste *et al.*, 1999) sowie 2000 gegen Digoxigenin (Schlehuber *et al.*, 2000) jeweils selektiert über das *phage-display*-Verfahren.

Welche Ursachen kommen für die Selektion von Nonsens-Mutanten aus der Bibliothek in Frage?

Aus praktischen Gesichtspunkten wurde die Bibliothek randomisierter PH-Domänen in den Vektor p10PHmg3 kloniert, in dem die PH-Domäne über ein *amber*-Codon mit dem g3-Protein fusioniert ist und unter dem *lac*-Promotor bei IPTG-Induktion exprimiert wird. Die Existenz dieses *amber*-Codons als Grund für die Selektion von Nonsens-Mutanten in *supE-E. coli*-Stämmen ist allerdings auszuschließen, da Vorversuche über Western-Blot-Analyse belegten, daß zwar die Wahl des *lac*-Promotors gegenüber dem *tet*-Promotor nach jeweiliger Induktion zu einer (geringfügigen) Expressionsverbesserung der PH-Domäne auf der Phagenoberfläche führt, die Existenz oder Abwesenheit eines *amber*-Codons zwischen PH-Domäne und g3-Protein sich aber nicht auf die Expression der PH-Domäne auf der Phagenoberfläche auswirkt. Eine

Aussage über die Zahl der Translationsabbrüche an diesem Codon und die Auswirkung eines weiteren *amber*-Codons (s.o.) läßt sich allerdings aus diesen Vorversuchen nicht ableiten.

Eine zweite Ursache kann in der Auswahl der Versuchsbedingungen liegen. Die drei untersuchten Zielsubstanzen könnten für die Selektion bindender PH-Domänen ungeeignet sein. Allerdings wurden bereits gegen MOMP eine bindende hPSTI-Mutante (Decker *et al.*, 2000) und gegen CD3 ϵ ein bindendes *scF_V*-Fragment über jeweilige *phage-display*-Verfahren selektiert (Mack *et al.*, 1995, s. Kapitel 4.2.1.), weshalb sich diese Zielsubstanzen für *in vitro* Selektionen generell zu eignen scheinen; diese Argumentation gilt aber nur für den Einsatz strukturähnlicher Bibliotheken. Wie erheblich die Zielsubstanzen durch die Wahl einer Bibliothek eingeschränkt sind, wird aus der Anwendung der „Anticaline“ im *phage-display*-Verfahren deutlich; die „Anticaline“ leiten sich von dem Bilinbindenden Lipocalin der *Pieris brassicae* ab, wobei es sich bei Bilin um Biliverdin IX γ , einer prosthetische Hämgruppe handelt. Bindende Mutanten der „Anticaline“ wurden bislang nur gegen die Haptene Fluoreszein (Beste *et al.*, 1999) und Digoxigenin (Schlehuber *et al.*, 2000) erhalten und die sind als zyklische Kohlenwasserstoffsysteme dem heterozyklischen Biliverdin IX γ strukturell ähnlich; gegen Proteindomänen wurden dagegen noch keine bindenden Anticalinmutaten nachgewiesen (Weiss & Lowman, 2000). Die PH-Domänen binden *in vivo* an Phosphatidylinositolphosphaten (Ferguson *et al.*, 2000). Sollten auch sie nach Randomisierung ein genauso geringes Spektrum erfolgversprechender Zielsubstanzen besitzen, wie dies bei den Anticalinen der Fall ist, verwunderte es nicht, wenn sich bislang noch keine bindenden Mutanten der PH-Domäne gegen die drei Zielsubstanzen selektieren ließen und sich die Bibliothek randomisierter PH-Domänen bislang nur für die Selektion „phosphatidylinositolphosphat-ähnlicher“ Zielsubstanzen eignen würde, was Gegenstand weiterer Untersuchungen sein sollte. Eine grundsätzliche Verbesserung der Affinität zu komplexen Epitopen von Proteinen könnte in einer flexibleren Gestaltung der Loops bestehen, indem zusätzliche Glycinreste zwischen Basisprotein und Loops kloniert werden, so daß sich die Loops besser den Epitopen von Zielsubstanzen anlagern. Bislang ist die Flexibilität der Loops zumindest bei der PH-Domäne des Grp1 gering. Dies wird aus einem Strukturvergleich des Grp1 mit und ohne Inositol-(1,3,4,5)-tetraphosphat deutlich (Lietzke *et al.*, 2000): die Bindung führt nicht zu einer deutlichen Konformationsänderung innerhalb der PH-Domäne. Antikörper dagegen verändern ihre Konformation nach Antigenkontakt häufig deutlich (MacCallum *et al.*, 1996), wie es beispielsweise der H3-Loop des F_{ab}-Fragment 14/9 nach Bindung eines Heptapeptids zeigt (PDB: 1HIM und 1HIL; Schulze-Gahmen *et al.*, 1993). Wie deutlich sich die Flexibilität der Loops auf die Affinität und

Spezifität von Antikörperbindungen auswirkt, zeigen die vergleichenden Untersuchungen an Antikörpern der ersten und zweiten Immunantwort (Manivel *et al.*, 2000). Loops von Antikörpern der ersten Immunantwort sind strukturell sehr flexibel, was eine schwache Bindung zu einer Vielzahl von Antigenen erlaubt. Dagegen weisen Antikörper der zweiten Immunantwort eine geringere Flexibilität ihrer Paratope auf, was zu einer höheren Substratspezifität und -affinität führt. Die adäquate Flexibilisierung der randomisierten Loops bei der PH-Domäne könnte in ähnlicher Weise wie bei den Antikörpern zu häufigeren Bindung an Zielsubstanzen beitragen und zu bindenden Mutanten beim *phage-display*-Verfahren führen. Die lokalen und nicht-lokalen Wechselwirkungen der Wildtyp-Loops mit dem Proteingerüst ließen sich allerdings nach Einführung von Glycinresten kaum mehr aufrecht erhalten, so daß man die PH-Domäne dann gleichzeitig auch vollständig in ihren Loops randomisieren sollte. Ob der Großteil aller Mutanten dann allerdings noch exprimierbar ist, müßte experimentell zuerst nachgewiesen werden.

Auch die Randomisierung *per se* steht als mögliche Ursache für die Selektion von Nonsense-Mutanten zur Disposition: wie in Kapitel 5.2.2. dargestellt, sollte sich eine stärkere Randomisierung der Loops verwirklichen lassen, ohne dabei die Exprimierbarkeit der (meisten) randomisierten Mutanten nachhaltig zu beeinflussen. Wie sich dies auf die Selektion bindender Mutanten auswirkt, sei dahingestellt. Vor der Klonierung einer selektierbaren (großen) Bibliothek, sollte allerdings zunächst ein sinnvolles Ausmaß an Randomisierung experimentell abgesichert werden, indem die cytoplasmatische Expression von Mutanten analysiert wird.

Eine letzte Ursache für die Selektion von Nonsense-Mutanten könnte im gewählten Selektionsverfahren bestehen. Um dies zu untersuchen, sollten entweder Voraussetzungen für den verbesserten Einsatz der PH-Domäne im *phage-display*-Verfahren geschaffen oder aber das Selektionsverfahren sollte ersetzt werden. Beide Ansätze wurden weiterverfolgt.

Die Präsentation der PH-Domäne auf der Phagenoberfläche durch Fusion mit dem g3-Protein setzt die Translokation der PH-Domäne durch die innere Membran in das Periplasma von *E. coli* voraus (s. Kapitel 1.5.1.). Der Western-Blot Nachweis der Wildtyp PH-Domäne (als Fusion mit dem g3-Protein) auf dem Phagen zeigt, daß verglichen mit dem g3-Protein nur wenige Kopien auf der Phagenoberfläche repräsentiert werden, was sich in einer vergleichsweise dünnen Bande äußert (s. Abb. 4.27). Die Anzahl von g3-Proteinen eines M13-Bakteriophagen beträgt drei bis fünf Kopien (Rasched & Oberer, 1986; Specthrie *et al.*, 1992). Hieraus folgt, daß nur ein geringer Anteil von Phagen überhaupt eine Kopie der PH-Domäne auf der Oberfläche

trägt. Ineffizienter Transport der PH-Domäne in das Periplasma könnte für einen so großen Selektionsdruck in *E. coli* sorgen, daß sich aus der Bibliothek von PH-Domänen-g3-Fusionen nur Nonsense-Mutanten selektieren lassen. Ziel müßte es also sein, diese Transportschwierigkeiten zu beseitigen, wovon Kapitel 5.4. handelt. Aber auch wenn der ineffiziente Transport keinen negativen Selektionsdruck im *phage-display*-Verfahren ausüben sollte, beeinflußt er die Selektion doch indirekt negativ: Phagen mit zwei oder mehr Kopien ihrer PH-Domänen Mutante auf der Oberfläche werden nur äußerst selten exprimiert und somit entfallen Aviditätseffekte, mit denen schwach bindende Mutanten aus der Bibliothek selektiert werden könnten. Als ein Beispiel für den Aviditätseffekt sei die *phage-display*-Selektion einer Bibliothek von Hexapeptiden (fusionsiert mit dem g3-Protein) gegen einen *anti*- β -Endorphin Antikörper als Zielsubstanz angeführt (Cwirla *et al.*, 1990): dabei wurden Hexapeptide selektiert, deren Affinitäten lediglich 0.35 bis 8.3 μ M gegen den Antikörper betragen (verglichen mit einem bekannten Liganden, der eine Affinität von 7.1 nM zeigt), sich dafür aber vergleichsweise gut auf der Phagenoberfläche exprimieren ließen. Schwach bindende Mutanten der PH-Domäne könnte man folglich über das *phage-display*-Verfahren selektieren, wenn sie sich mit mehreren Kopien auf der Phagenoberfläche darstellen ließen.

Zur besseren Expression der PH-Domäne auf der Phagenoberfläche wurde nach Mutanten gesucht, die besser in das Periplasma von *E. coli* transloziert werden (s. Kapitel 5.4.). Gefunden wurden dabei Circularpermutanten (s. Kapitel 4.3.2.), die sich bis zu 30-fach besser periplasmatisch exprimieren und sich auch besser auf der Phagenoberfläche darstellen lassen (s. Abb. 4.27). Nach Randomisierung dieser Mutante wurde mit ihr und mit dem randomisierten Wildtyp ein *phage-display*-Verfahren ohne *panning* durchgeführt, indem die Phagen nach ihrer Expression und Reinigung lediglich verdünnt zur Infektion neuer *E. coli*-Zellen eingesetzt wurden (s. Kapitel 4.4.1.). Die Restriktionsanalysen nach den Selektionsrunden zeigten bei der Selektion mit der Circularpermutante dieselbe Tendenz zu Deletionsmutanten wie dies beim Wildtyp beobachtet wurde. Damit trägt die verbesserte Translokationseffizienz der Circularpermutante gegenüber dem Wildtyp nicht dazu bei, Deletionsmutanten zu vermeiden. Da allerdings keine vergleichbaren Untersuchungen mit anderen Bibliotheken vorliegen, sollte dieses Ergebnis nicht zu hoch bewertet werden. Die bessere Expression auf der Phagenoberfläche qualifiziert dennoch diese Mutante gegenüber dem Wildtyp für kommende *phage-display*-Selektionen, da somit mehr Phagen eine Mutante repräsentieren und eher Aviditätseffekte auftreten können.

Mögliche *phage-display*-spezifische Ursachen für die Selektion von Nonsense-Mutanten werden durch Einsatz der Bibliothek bei *in vivo* Selektionssystemen

umgangen. Die Selektion über das Hefe-2-Hybrid-System ist dabei im Zellkern lokalisiert, was wiederum einen Transportprozeß in sich birgt. Auf jegliche Transportprozesse wird dagegen verzichtet, wenn die Bibliothek in *E. coli* zur metabolischen oder inhibitorischen Selektion verwendet wird, oder aber wenn sie im Cytoplasma von Zellkulturen eingesetzt wird. Letzteres geschieht im Arbeitskreis von Waldemar Kolanus (Genzentrum der LMU München) und ermöglicht es, die PH-Domäne an ihrem „Entdeckungsort“ zu selektieren – in (Jurkat L)T-Zellen (s. Kapitel 1.5.2.).

5.4. Die Verbesserung der Translokation

Die PH-Domäne läßt sich nur schlecht in das Periplasma translozieren, was als mögliche Ursache für ihre geringe Expression auf der Phagenoberfläche und für die Selektion von Nonsens- Mutanten beim *phage-display*-Verfahren angesehen wird (s. Kapitel 5.3.). Wie aber läßt sich die periplasmatische Expression bzw. die Translokation eines Proteins verbessern? Der allgemeinen Frage, worin sich (gut) periplasmatisch exprimierbare Proteine von nicht oder schlecht periplasmatisch exprimierbaren Proteinen unterscheiden, wurde in der Literatur bislang noch nicht systematisch nachgegangen. Deshalb wurden dazu Hypothesen aufgestellt und Mutanten der PH-Domäne entworfen, die die Hypothesen bestätigen oder widerlegen (Mutanten: s. Tab. 4.4):

- spezifische Sequenzelemente der PH-Domäne verhindern eine effektive Translokation (s. Kapitel 4.3.1.),
- die Stabilität der gesamten PH-Domäne oder ihrer (nicht analysierten) Subdomänen ist zu groß, um sie für die Translokation zu entfalten (s. Kapitel 4.3.1.) und
- die Reihenfolge der Strukturelemente (Topologie) ist bei der Wildtyp PH-Domäne ungünstig (s. Kapitel 4.3.2.).

Die untersuchten Kriterien sind primär die periplasmatische Expression (s. Kapitel 4.3.7.) und die dabei erfolgte Prozessierung der jeweiligen Präproteine (s. Kapitel 4.3.8.). Ergänzend wurde die cytoplasmatische Expression (s. Kapitel 4.3.6.), die thermodynamische Stabilität (s. Kapitel 4.3.5.) und die Expression auf der Phagenoberfläche (s. Kapitel 4.4.) bestimmt.

Gezielten Mutationen von Proteinen zur Verbesserung der Translokation wurde bislang weniger Aufmerksamkeit geschenkt als der Coexpression anderer (periplasmatischer) Proteine, die zu einer verbesserten periplasmatischen Expression verhelfen. Zur Etablierung der Lipocaline als Proteingerüst für das *phage-display*-Verfahren erwies sich die Überexpression der *E. coli* eigenen DsbC (Disulfid Oxidoreduktase C) als hilfreich (Schmidt *et al.*, 1998). Sie sorgt im Periplasma für die korrekte Verknüpfung von Disulfidbrücken (Shevchik *et al.*, 1994). Dieser Ansatz entfällt für die PH-Domäne, da in ihr keine Cysteine vorliegen und somit hierin auch nicht die Ursache für die schlechte periplasmatische Exprimierbarkeit liegen kann. Problematischer könnten bei der Faltung dagegen sechs Prolinreste sein, die über die PH-Domäne verteilt vorliegen (s. Tab. 4.4). Die Strukturanalyse zeigt, daß sich alle in der *trans*-Konfiguration befinden und nicht in der (für die freie Aminosäure) um ca. 2 kJ/mol energiereicheren *cis*-Konfiguration. *Cis*-Prolin-Konfigurationen treten bei vielen Proteinen auf und können die Stabilität des gesamten Proteins erhöhen. Läge eine *cis*-Konfiguration vor, könnte die periplasmatische Expression der PH-Domäne durch Coexpression einer Prolyl-isomerase wie *FkpA* verbessert werden. Dieses Chaperon katalysiert den Konfigurationsaustausch und wird im Periplasma von *E. coli* nach Hitzeschock exprimiert (Arie *et al.*, 2001). Zur verbesserten Expression eines schlecht funktionell auf Phagen exprimierten Antikörperfragments wurde anstelle der Coexpression eines bekannten Chaperons eine genomische DNA-Bibliothek mit dem Phagen coexprimiert (Bothmann & Plückthun, 1998). Der Phage kodiert dabei für die Mutanten der genomischen DNA-Bibliothek. Über einen Bindungsassay ließ sich damit das periplasmatisch exprimierte Skp von *E. coli* als Chaperon identifizieren. Um einen derartigen Ansatz auch mit der PH-Domäne durchzuführen, benötigte man für sie ebenfalls einen funktionellen Nachweis bzw. eine Selektionsmethode im Periplasma oder auf der Phagenoberfläche. Beides wäre nicht direkt sondern bestenfalls über die Fusion der PH-Domäne mit einem in das Periplasma translozierten, resistenzvermittelnden Proteins wie der β -Lactamase möglich gewesen.

Lipocalinen, *cis*-Prolin-haltigen Proteinen und dem oben erwähnten Antikörperfragment verhilft die Coexpression von periplasmatischen Chaperonen zur korrekten Faltung im Periplasma. Schlechte periplasmatische Expressionsausbeuten wären primär jeweils über Proteolyse un- oder fehlgefalteter Proteine im Periplasma zu erklären. Die PH-Domäne aber weist hervorragende Faltungseigenschaften aus, was bei der Aufnahme von Übergangskurven deutlich wird (s. Kapitel 4.1.5.). Somit sollte die Faltung auch im Periplasma nicht problematisch sein. Alle Vorversuche deuten viel eher auf ein Problem bei der Translokation hin. Dieses äußert sich in der geringen periplasmatischen

Expression und der nachfolgenden aberranten Prozessierung durch die Signalpeptidase.

Dieses Translokationsproblem könnte darauf beruhen,

- daß die PH-Domäne nach der Translation und vor der Translokation nicht in einen „geeigneten translokationskompetenten“ Zustand überführt wird oder aber
- daß die intrinsischen Eigenschaften der PH-Domäne die Translokation behindern.

Der „geeignete translokationskompetente“ Zustand (Economou, 2000; Kumamoto & Nault, 1989) ist eine phänomenologische Beschreibung und gibt nicht an, in welchem strukturellen Zustand sich das Präprotein vor der Translokation befindet. Da viele Proteine und Proteinkomplexe wie SRP, GroEL/S, Triggerfaktor, DnaJ und SecB als notwendige Chaperone für die Translokation bekannt sind (s. Kapitel 1.6.1.), wird allgemein vermutet, daß sich Präproteine vor der Translokation in einem ungefalteten (translokationskompetenten) Zustand befinden und dann über den allgemeinen sekretorischen Transportmechanismus in das Periplasma transloziert werden. Ein dazu von allen einheitlich benötigtes cytoplasmatisches Chaperon für die Destabilisierung von Präproteinen konnte allerdings bislang noch nicht nachgewiesen werden, was sicherlich auch mit der Substratspezifität von Chaperonen wie DnaJ (Rüdiger *et al.*, 2001) zusammenhängt. Ginge man davon aus, daß es zu jedem periplasmatisch exprimierten Protein bei *E. coli* mindestens ein spezifisches cytoplasmatisches Chaperon gibt, das erst den translokationskompetenten Zustand ermöglicht, sollten sich *E. coli*-periplasmafremde Proteine nicht rekombinant im Periplasma exprimieren lassen, da für sie kein geeignetes Chaperon vorliegt, bzw. sie nicht so evolutiv verändert wurden, daß ein cytoplasmatisches Chaperon von *E. coli* binden könnte. Beispielsweise das bilinbindende Lipocalin (Schmidt & Skerra, 1994), das als Gerüstprotein für die „Anticaline“ diente, aber auch Antikörperfragmente (Skerra & Plückthun, 1991) lassen sich sehr wohl periplasmatisch exprimieren. Beides sind *in vivo* eukaryontische Proteine, die cotranslational transloziert werden, so daß sie in der Evolution nicht der Problematik eines translokationskompetenten Zustandes ausgesetzt wurden und sich deshalb auch nicht besser in *E. coli* translozieren lassen sollten. Wenn sich die PH-Domäne besonders schlecht periplasmatisch exprimieren läßt, sollte man deshalb weniger von fehlenden Chaperonen ausgehen, als intrinsische Eigenschaften der PH-Domäne dafür zur Verantwortung zu

ziehen. Diese wurden durch die gewählten Mutationen näher untersucht (s. Kapitel 4.3. und folgende) und werden im folgenden diskutiert.

5.4.1. Hypothesenüberprüfung zur Translokationsoptimierung

Im vorherigen Kapitel wurde die PH-Domäne mit Proteinen verglichen, die sich (z.T. nach Coexpression periplasmatischer Chaperone) besonders effizient im Periplasma von *E. coli* exprimieren lassen. „Realistischer“ wäre der Vergleich der PH-Domäne mit Proteinen, die „Translokations-Schwierigkeiten“ besitzen. Typische Beispiele dafür sind Membranproteine, die naturgemäß in der Membran „stecken bleiben“. Bei *E. coli* zeichnen sie sich durch hydrophobe Aminosäurebereiche im Intermembranbereich aus (Sääf, 1998), die von positiven Clustern an Aminosäureresten flankiert werden. Letztere wirken als *stop*-Transfersignale (von Heijne & Gavel, 1988). Die Mutation hydrophober Bereiche in cytoplasmatischen Proteinen, führt bei Unkenntnis der Struktur zur Reduzierung des „hydrophoben Effekts“ und dies wiederum zur Verhinderung der Proteinfaltung, weshalb derartige Mutationen nicht in das Kalkül gezogen wurden. Positive Ladungscluster aber, wie sie bei der PH-Domäne in der Aminosäuresequenz KRR(20,21,22) vorliegen, könnten als *stop*-Transfersignal dienen und lassen sich (zumindest im Falle der PH-Domäne) mutieren, ohne daß die Proteinstabilität deutlich reduziert wird. Die Mutation KRR(20,21,22)QNV verhalf der entsprechenden Variante allerdings nicht zu einer verbesserten Translokation und auch die Circularpermutanten CP 2.6. und CP 3.6. werden nicht durch diese Mutation in ihren Translokationseigenschaften verändert, so daß die Annahme einer *stop*-Transfersequenz hinfällig ist. Genauso wurde keine Verbesserung für die sequenzspezifisch mutierte Variante N-ter gefunden, deren N-Terminus der N-terminalen Sequenz des reifen *outer membrane proteins A* angeglichen wurde.

Sollte die Wildtyp PH-Domäne als Präprotein vor der Translokation zuvor im Cytoplasma gefaltet sein, müßte sie zur Translokation entfaltet werden und *E. coli* müßte dazu den Energiebetrag aufbringen, der auch über Übergangskurven bestimmt werden kann (s. Kap. 4.1.5.). Destabilisierte Mutanten werden dann „energiegünstiger“ und vielleicht auch effizienter transloziert. Zur Untersuchung dieser Hypothese wurde die Wildtyp PH-Domäne durch gezielte Mutation der α -Helix und des Proteininneren destabilisiert, was allerdings bei keiner Mutante zu einer deutlichen Verbesserung der periplasmatischen Expression führte. Die thermodynamische Stabilität der PH-Domäne verursacht somit nicht die schlechte Translokation.

Die Auswirkungen von Circularpermutationen (s. Kapitel 4.3.2.) auf die Translokation kann vielschichtig sein: durch die Einführung neuer Termini beginnt die Translokation der PH-Domäne mit einer Aminosäuresequenz aus Loop 6 (bei den Varianten CP X.6.) oder mit der Aminosäuresequenz der (beim Wildtyp C-terminalen) α -Helix. Die Bedeutung des N-Terminus für die Translokationseffizienz ist statistisch erwiesen (Kajava *et al.*, 2000: überwiegend negative und ungeladene Aminosäuren in den 18 N-terminalen Aminosäurepositionen reifer, in *gram*-negativen Bakterien periplasmatisch exprimierter Proteinen). Somit werden grundlegend neue Voraussetzungen für die Translokation geschaffen. Sollten Präproteine (zumindest die Mutanten der PH-Domäne) gefaltet vorliegen, könnten sich die Circularpermutanten in ihrer sukzessiven Entfaltung vom Wildtyp unterscheiden, was zu anderen Translokationseffizienzen führen könnte. Zumindest bei den Circularpermutanten CP X.6. wird genau dies beobachtet, indem sie um ein Vielfaches besser periplasmatisch exprimiert werden als der Wildtyp (CP 2.6. mit 30-facher periplasmatischer Expressionssteigerung, CP 3.6. mit 10-facher: s. Kapitel 4.3.7.) und im Gegensatz zum Wildtyp sowie den anderen massenspektroskopisch untersuchten Mutanten genau das Prozessierungsmuster (nach Translokation und Prozessierung mit der periplasmatischen Signalpeptidase I) aufweisen, das man es nach statistischer Analyse von natürlichen, in *gram*-negativen Bakterien periplasmatisch exprimierten Proteinen erwartet (s. Kapitel 4.3.8.).

5.4.2. Verbesserte Translokation bei den Circularpermutanten CP 2.6 und CP 3.6.

Mit der periplasmatischen Expression der Circularpermutanten (CP 2.6. und CP 3.6.) ist es bislang erstmalig gelungen, ein Protein durch Circularpermutation in seiner Funktion (hier die Translokationseigenschaft) zu verbessern. In der Vergangenheit wurden Circularpermutanten nur unter dem Aspekt des „Proteinfaltungsproblems“ (Anfinsen, 1973) und der Membranorientierung betrachtet (*Outer membrane Protein A* von *E. coli*: Köblik & Kramer, 1995) (transmembrane Untereinheit des Glucosetransporters IICB^{Glc}: Beutler *et al.*, 2000). Die erste veröffentlichte Circularpermutation betrifft den Trypsin-Inhibitor aus Rinderpankreas, dessen Termini chemisch miteinander verbunden wurden und der anschließend im Protein gespalten wurde. Die so hergestellte Circularpermutante ließ sich denaturieren und dann wieder in die ursprüngliche Struktur rückfalten (Goldenberg & Creighton, 1983). (Goldenberg & Creighton, 1983). Genetisch konstruierte Circularpermutanten wurden von der Phosphoribosyl-Anthranilsäure-Isomerase (Luger *et al.*, 1989), der Aspartat-

Transcarbamoylase (Zhang & Schachman, 1996), der SH3-Domäne des α -Spectrins (Viguera *et al.*, 1996), dem Chymotrypsin-Inhibitor 2 (Otzen & Fersht, 1998) und weiteren Proteinen hergestellt. Anhand der Circularpermutanten dieser Proteine konnte gezeigt werden, daß für die Bildung einer globalen Proteinstruktur und die daraus abgeleitete Aktivität alleine die Folge der Aminosäuren verantwortlich ist (Anfinsen, 1973), während die Positionen der Termini wenig Einfluß auf die globale Struktur besitzen. Während bei obigen Circularpermutanten die Termini gezielt eingefügt wurden, konnte mit Hilfe eines Randomisierungsverfahrens beim Grünfluoreszierenden Protein (GFP) gezeigt werden, daß sich neue Termini lediglich in einigen Loops einführen lassen und zu funktionellem, d.h. grün fluoreszierendem, Protein führen (Topell *et al.*, 1999). Dasselbe Randomisierungsverfahren auf die Disulfid-Oxidoreduktase DsbA angewandt, erbrachte dagegen die Erkenntnis, daß sich nur in 30 % der Proteinstruktur keine Termini einführen lassen, während dies bei 70 % möglich ist und funktionelles Protein erhalten wird (Hennecke *et al.*, 1999). DsbA war (bis *dato*) das einzige circularpermutierte Protein, das periplasmatisch in *E. coli* exprimiert wurde (s. Kapitel 5.4.4.).

Wie auch beim DsbA gibt es unter den circularpermutierten PH-Domänen Mutanten, die nicht zur (verbesserten) periplasmatischen Expression führen, sich aber wohl cytoplasmatisch exprimieren und gereinigt untersuchen lassen; dies gilt für die Circularpermutanten CP X.7., deren N-Terminus sich vor der α -Helix befindet. Allerdings sind diese Mutanten deutlich gegenüber dem Wildtyp destabilisiert, weshalb sie im Periplasma rasch proteolysiert werden könnten, auch wenn sie zuvor sehr effizient transloziert sein sollten. Zusätzlich beeinflußt die Wahl des Linkers zwischen den originalen Termini deutlich die Eigenschaften der Circularpermutanten: unspezifische Linker, wie sie in CP 2.6. (kurzer Linker) und CP 3.6. (langer Linker) verwirklicht wurden, führen zu ähnlich thermodynamisch stabilen Mutanten wie der Wildtyp und zur hervorragenden periplasmatischen Expression dieser Mutanten. Dagegen führt ein modellierter Linker bei der periplasmatischen Expression der Mutante CP 1.6. zur Zelllyse und zu weiteren Faltungsintermediaten im aufgereinigten Protein, was die Übergangskurve veranschaulicht (s. Abb. 4.19). Verursacht wurde diese nicht-kooperative Faltung durch die Einführung einer α -helicalen Aminosäuresequenz aus der Phospholipase C (PDB: 1MAI; Ferguson *et al.*, 1995) und einem zusätzlichen (auf statistischen Erkenntnissen beruhenden) modellierten Hexapeptid als Linker zwischen beiden Helices (s. Kapitel 4.3.4.). Die nicht-kooperative Faltung kann dabei auf der Erzeugung stabiler Faltungsintermediate durch die Einführung eines weiteren sekundären Strukturelements beruhen, oder aber die Destabilisierung der α -Helix (von Cytohesin-1) führt bei der Circularpermutante zu Faltungsintermediaten. Diese Destabilisierung wurde

jedenfalls in der (nicht circularpermutierten) Variante C-ter beobachtet, bei der die α -Helix des Cytohesin-1 genauso mutiert wurde wie in den Circularpermutanten CP 1.X.. Sie läßt sich mit der Destabilisierung des helicalen Dipols durch Substitution einer positiven (K114) und Einführung einer negativen Ladung (D117) erklären (s. Kapitel 4.3.1. und Tab. 4.4). Die nicht-kooperative Faltung von CP 1.6. besitzt gegenüber CP 2.6. hinsichtlich der cytoplasmatischen Expressionsausbeute keine Auswirkungen (s. Tab. 4.21), was für eine geringe Proteasesensitivität der intermediären Strukturen spricht.

Die Circularpermutanten auf der Phagenoberfläche

Das pragmatische Ziel der Translokationsoptimierung war die verbesserte Expression der PH-Domäne auf der Oberfläche des M13-Phagen, was im Rahmen der Meßmöglichkeiten anhand eines Western-Blots von denaturierten Phagen auch erreicht werden konnte: die Circularpermutanten CP 1.6., CP 2.6. und CP 3.6. werden besser auf der Phagenoberfläche exprimiert als der Wildtyp (s. Abb. 4.27). Da aber für die PH-Domäne auf der Phagenoberfläche kein quantitativer Nachweis existiert, läßt sich diese Verbesserung nicht so quantifizieren, wie dies z.B. für das translokationsoptimierte *scF_V*-Fragment des Antikörpers 4-4-20 mit seiner Fluoreszein-Bindungseigenschaft möglich war (Bothmann & Plückthun, 1998), bei dem durch Coexpression des periplasmatischen Chaperons Skp eine vierfach höhere periplasmatische Expressionsausbeute bestimmt werden konnte als wenn auf die Coexpression verzichtet wurde (s. Kapitel 5.4.).

Das *phage-display*-Verfahren mit dem g3-Protein fusionierten Peptid setzt die Translokation der PH-Domäne durch die innere Membran und die Insertion des g3-Proteins in die innere Membran von *E. coli* voraus (s. Kapitel 1.5.1.); ob für die g3-fusionierte und die nicht-fusionierte PH-Domäne die gleichen biochemischen Voraussetzungen bei der Translokation gegeben sind, ist nicht sichergestellt. So interagiert das SecB-Protein als Chaperon für die Translokation erst mit Substratproteinen, die größer als 150 Aminosäuren sind (Randall *et al.*, 1997). Das Fusionsprotein könnte sich somit im Cytosol komplexieren lassen, während die nicht-fusionierte PH-Domäne nicht durch SecB gebunden werden kann.

5.4.3. Interpretationen für die guten Translokationseigenschaften

Für die deutlich verbesserte Translokation der Circularpermutanten CP 2.6 und CP 3.6. gegenüber der Wildtyp PH-Domäne in das Periplasma von *E. coli* gibt es vier grundlegende Erklärungsmöglichkeiten, die im Rahmen dieser Arbeit untersucht wurden:

- I) Die beiden Circularpermutanten lassen sich wesentlich besser translatieren als der Wildtyp, so daß bei ihnen erheblich mehr Protein für die Translokation zur Verfügung steht oder die Wildtyp PH-Domäne wird im Periplasma proteolysiert.
- II) Das ompA^M-Signalpeptid destabilisiert die beiden Circularpermutanten vor der Translokation stärker als beim Wildtyp.
- III) Durch die Topologieänderung gelangen translokationsbegünstigende Peptidsequenzen an den N-Terminus des prozessierten Proteins, bzw. es werden gerade die N-terminalen basischen Aminosäuren des Wildtyps aus dem Bereich der 18 ersten translokationskritischen entfernt (Kajava *et al.*, 2000), die bislang die Translokation behinderten.
- IV) Die veränderte Topologie führt zu einer Verbesserung der Translokation, da die PH-Domänen im translokationskompetenten Zustand gefaltet vorliegen und die Entfaltung bei den Circularpermutanten CP 2.6. und CP 3.6. günstiger als beim Wildtyp verläuft. Dabei wird die Entfaltung nicht durch eine ungerichtete energetische Barriere, sondern durch die maximale vektorielle (gerichtete) Kraft limitiert, die für die sukzessive am N-Terminus beginnende Denaturierung benötigt wird („Kräftemodell“).

ad I.

Der Vergleich von cytoplasmatischer und periplasmatischer Ausbeute der einzelnen Mutanten zeigt, daß nur die besonders instabilen Circularpermutanten in beiden Fällen gleich (schlecht) exprimiert werden (s. Abb. 4.23B). Die cytoplasmatische Ausbeute der beiden Circularpermutanten CP 2.6. und CP 3.6. beträgt jeweils ungefähr 30 % zu der des Wildtyps, während die periplasmatische Expressionssausbeute bei CP 2.6. 30-fach höher liegt als die des Wildtyps. Der Einfluß der Translation auf die periplasmatische Expression ist folglich bei den Mutanten der PH-Domäne nicht ersichtlich. Die ähnliche thermodynamische Stabilität der Proteine, die nahezu identische Aminosäuresequenz und globale Struktur sollten zudem zu einer ähnlichen Proteasesensitivität im

Periplasma führen, so daß nicht von einer proteolytischen Hydrolyse des Wildtyps auszugehen ist.

ad II.

Die Circularpermutante CP 2.6. konnte (nach Expression über *inclusion bodies*) mit fusioniertem ompA^M-Signalpeptid isoliert und die thermodynamische Stabilität untersucht werden (s. Abb. 4.26): das Signalpeptid besitzt *in vitro* keinen destabilisierenden Einfluß auf sie. Bei der *in vivo* periplasmatisch in *E coli* exprimierten β -Lactamase dagegen konnte der destabilisierende Einfluß des Signalpeptids auf die Proteinstabilität nachgewiesen werden: *in vitro*-Experimente an der β -Lactamase mit Signalpeptid bei 37°C zeigen, daß das Protein denaturiert vorliegt, ohne Signalpeptid dagegen nativ und funktionell. Bei 30°C sind beide Proteine stabil, woraus auf die Destabilisierung der β -Lactamase durch das Signalpeptid geschlossen wird (McCarthy *et al.*, 1998). Welche Eigenschaften (erweiterte) Signalpeptide neben der Markierung und der möglichen Destabilisierung von Präproteinen besitzen können, wird am Beispiel der α -lytischen, extrazellulären Serinprotease von *Lysobacter enzymogenes* deutlich: das Prä-Pro-Protein wird in das Periplasma transloziert und dort das Signalpeptid prozessiert. Das verbleibende Proprotein faltet sich dort, woraufhin die Pro-Region autokatalytisch entfernt wird. Die so aktivierte Protease wird sekretiert (Silen & Agard, 1989). Die aktivierte Protease befindet sich dabei in einem metastabilen Zustand. Wird sie nämlich nur als Präprotein periplasmatisch exprimiert, stellt sich eine deutlich stabilere Konformation ein, die sich auch durch Rückfaltung erzielen läßt (Anderson *et al.*, 1999). Derartige zusätzliche Eigenschaften, wie bei der β -Lactamase oder der Serinprotease, zeigt das ompA^M-Signalpeptid bei der Circularpermutante nicht.

ad III.

Auch die ersten 18 Aminosäuren der Circularpermutante beinhalten noch zwei basische Aminosäurereste, während der Wildtyp in diesem Bereich vier basische Aminosäurereste aufweist. Damit entspricht die Aminosäureverteilung bei der Circularpermutante eher natürlich periplasmatisch exprimierten Proteinen (Kajava *et al.*, 2000), was sich in der verbesserten Translokation auswirken könnte. Untersucht wurde diese sequenzspezifische Möglichkeit, indem die ersten sieben Aminosäuren von CP 2.6. in der Mutante CP.ohnLi deletiert wurden (s. Kapitel 4.3.9.) und ihr die periplasmatische Expression bestimmt wurde: sie ist zwar gegenüber CP 2.6. deutlich verringert, doch läßt sie sich weiterhin mehr als dreifach besser periplasmatisch exprimieren als der Wildtyp (s. Abb. 4.22 und 4.23). Eine Insertion dieser sieben Aminosäuren zwischen

Signalpeptid und der Wildtyp PH-Domäne bei der Mutante WT.miLi führt dagegen nicht zu einer Steigerung der periplasmatischen Expression gegenüber dem Wildtyp. Das bedeutet für die Translokation, daß diesen sieben Aminosäuren keine besondere Bedeutung als unstrukturierter Linker zukommt und daß die Aminosäuresequenz des N-Terminus nicht entscheidend zur Translokationseffizienz beiträgt. Geht man von einem translokationskompetenten Zustand aus, bei dem das Signalpeptid durch Chaperone wie SRP und den Triggerfaktor komplexiert wird und weitere Chaperone mit Hilfe der ersten an der PH-Domäne binden (s. Kapitel 1.6.1.), sollte sich ebenfalls der Abstand zwischen Signalpeptid und Protein deutlicher auswirken.

ad IV.

Es gibt Präproteine im translokationskompetenten Zustand, die durch Chaperone destabilisiert werden. Dies veranlaßt zu der (üblichen) Meinung, daß alle Präproteine vor ihrer Translokation entfaltet vorliegen. Die Beteiligung und Nicht-Beteiligung unterschiedlicher Chaperonen an der Translokation bei unterschiedlichen Präproteinen zeugt jedoch schon von der Heterogenität der Chaperonkomplexierung (s. Kapitel 1.6.1.). Die Wechselwirkung von Chaperonen mit Proteinen läßt sich häufig auf Peptidmotive der Substratproteine zurückführen. Die Chaperone DnaK (Rüdiger *et al.*, 1997) und DnaJ (Rüdiger *et al.*, 2001) binden an kurzen Peptidsequenzen, in denen insbesondere aromatische und hydrophobe, und der Triggerfaktor an welchen, in denen aromatische sowie positiv-geladene Aminosäurereste (Patzelt *et al.*, 2001) enthalten sind. Sollten die Bindungen von Chaperonen nur von Peptidmotiven im ungefalteten Zustand abhängig sein, besitzt die Position dieser Peptide im Protein keinen Einfluß auf die Komplexierung, weshalb sich die Circularpermutanten und der Wildtyp nicht hinsichtlich der Komplexierung durch diese Chaperone voneinander unterscheiden sollten. Sofern die Punktmutationen (z.B. zur Destabilisierung der PH-Domäne) nicht zufällig Bindungsmotive für diese Chaperone erzeugen oder zerstören, dürften auch sie sich nicht auf eine Komplexierung auswirken.

Die Translokation erfolgt, indem sich SecA-Protein an das Signalpeptid anlagert, der gesamte Komplex zum Membrankomplex (SecY/E/G) transportiert wird und dort das SecA-Protein unter ATP-Hydrolyse das Protein in den Membrankomplex zieht (s. Kapitel 1.6.1.). Liegen die Mutanten im translokationskompetenten Zustand gefaltet vor, müssen sie durch die ATP-Hydrolyse entfaltet werden. Für diese Entfaltung ist die durch Übergangskurven ermittelte Energie notwendig. Die Translokation wird dabei durch eine protonenbewegende Kraft (*proton-motive force*) unterstützt (Arkowitz & Wickner, 1994; Manting & Driessen, 2000) und erfolgt schrittweise durch die innere Membran;

dabei werden bei jedem Schritt ca. 20 Aminosäurereste durch die Membran gezogen (Schiebel *et al.*, 1991). Die Denaturierung verläuft nicht global-ungerichtet (wie im Zweizustandsmodell gefordert über eine energetische Barriere) sondern vektoriell. Limitierend sollte sich dann die maximale Kraft (F_{Denat}) auswirken, die während der Translokation zur Entfaltung des Proteins aufgebracht werden muß. Vergleichen läßt sich dies mit einem verknäuelten Faden, dessen Ende schon in einem Nadelöhr steckt und nun durch dieses Nadelöhr gezogen wird. Die Gesamtenergie für das Auflösen des Knäuels ist weniger bedeutsam, als der benötigte maximale Kraftaufwand in einer Situation, in der „es hakt“, um weiteren Faden durch das Nadelöhr zu ziehen.

Die benötigte Kraft zur Denaturierung von Proteinen läßt sich durch Einzelmolekül-Kraft-Spektroskopie bestimmen. Dabei werden die beiden Termini auseinandergezogen und zu jedem Abstand beider Termini die benötigte Kraft bestimmt (Dabei gibt die Fläche unter dieser Kurve die Faltungsenergie an). Dieses Meßverfahren wurde schon erfolgreich auf die Immunglobulindomänen des Titins und Fibronectin Typ III-Domäne (FN3) angewandt (Rief *et al.*, 2000): die Titin-Domäne ist anfangs sehr zugestabil. Ist das erste β -Faltblatt von der Domäne gelöst, zeigt das restliche Proteine keinerlei Stabilität mehr und entfaltet spontan. Da die Kraft eine vektorielle Größe ist, ist die Richtung entscheidend, in die gezogen wird. Dies wird bei der maximalen Kraft zur Denaturierung einer kurzen, doppelsträngigen DNA besonders deutlich (s. Abb. 5.1). Für das Denaturieren von DNA durch Ziehen an beiden 5'-Enden mit der Kraft F_p wird eine sehr viel höhere (maximale) Kraft benötigt, als wenn senkrecht zur Helix beide Stränge mit der Kraft F_s auseinandergezogen werden und sich sukzessive die Basenpaarbindungen lösen. Entsprechende Experimente sind in Vorbereitung und lassen diese Aussagen zu (persönliche Mitteilung von Hermann Gaub, LMU München, Institut für angewandte Physik).

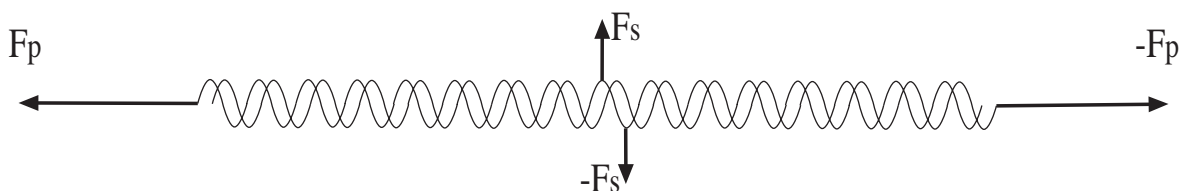


Abb. 5.1: Maximale, benötigte Kraft zur Denaturierung von doppelsträngiger DNA

Für die Denaturierung doppelsträngiger DNA wird bei einer Zugrichtung von den 5'-Enden aus über eine kurze Strecke eine sehr viel höhere Kraft F_p benötigt, als beim Zug senkrecht zur DNA mit der Kraft F_s über eine längere Strecke. Die benötigte Energie zur Denaturierung der DNA ist in beiden Fällen identisch und wird durch die Fläche unterhalb der Kraft-gegen-Strecke-Kurve wiedergegeben.

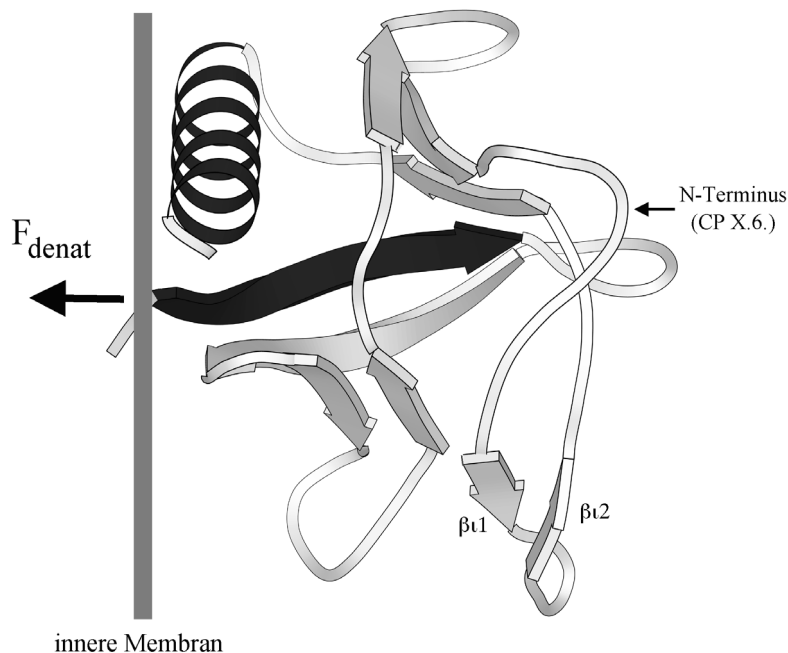


Abb. 5.2: Kräftemodell zur Translokation: die Topologie der Wildtyp PH-Domäne.

Bei der Translokation wird das Protein mit der Kraft F_{denat} sukzessive denaturiert und über den Sec(Y/E/G)-Komplex in das Periplasma gezogen. Das $\beta 1$ -Faltblatt (schwarz) bei der Wildtyp PH-Domäne ist senkrecht zur inneren Membran ausgerichtet und interagiert über 11 Aminosäuren mit zwei Faltblättern und der α -Helix (schwarz). Letztere liegt wie ein Keil zwischen innerer Membran und dem restlichen Protein. Dagegen besteht das N-terminale $\beta 1$ -Faltblatt der Circularpermutanten CP X.6. nur aus drei Aminosäuren, die mit zwei Faltblättern interagieren. Danach folgt ein „turn“, der einen Richtungswechsel der Kraft zur Folge hat.

Bei der Translokation von Proteinen über den allgemeinen sekretorischen Transportmechanismus von *E. coli* wird das Signalpeptid mit SecA zuerst an der Membran ausgerichtet und dann sukzessive in die Membran gezogen. Die Ausgangssituation ist im „Kräftemodell“ für Translokationen wiedergegeben (s. Abb. 5.2). Die Denaturierung des (hypothetisch) strukturierten Proteins erfolgt durch den Zug mit der Kraft F_{denat} . Bei der Wildtyp PH-Domäne verläuft das 11 Aminosäuren umfassende $\beta 1$ -Faltblatt senkrecht zur Zellwand und interagiert mit den zu $\beta 1$ antiparallelen β -Faltblättern 2 und 7 sowie mit der α -Helix. Diese Strukturelemente stellen somit gegen den Zug am $\beta 1$ -Faltblatt ein starkes „Widerlager“ dar. Dabei liegt die α -Helix wie ein Keil zwischen der Zellmembran und der restlichen PH-Domäne. Bei den Circularpermutanten CP X.6. ist das $\beta 1$ -Faltblatt das erste sekundäre Strukturelement, welches aber

nur eine Länge von drei Aminosäuren besitzt und dort mit den Faltblättern β_2 und 4 interagiert. Die Ausrichtung ist im Modell dabei nicht antiparallel. Es schließt sich direkt ein „turn“ (Loop Li1) an, der einen Richtungswechsel der Denaturierungskraft zur Folge hat, betrachtet man den Vorgang mit festgelegten Koordinaten für die PH-Domäne. Das Widerlager für den Zug am β_1 -Faltblatt ist folglich im Vergleich zum Wildtyp besonders gering. In Bezug auf das DNA-Modell (s. Abb. 5.1) könnten sich die beiden Mutanten in der Richtung der aufgewendeten Kraft zu Anfang der Translokation oder/und durch die Zahl von Interaktionen voneinander unterscheiden, was im DNA-Modell durch die Zahl der Basenpaarungen wiedergegeben wird und der Größe des Widerlagers entspricht. Ist jeweils der N-Terminus vom Wildtyp oder von der Circularpermutante gelöst, zerfällt das restliche Protein. Ähnliches wird bei der Denaturierung von Titin im Kraftfeld beobachtet (Rief *et al.*, 2000).

Die Untersuchung der Mutanten über Einzelmolekül-Kraft-Spektroskopie könnte ein weiteres Indiz für das „Kräftemodell“ erbringen. Grundvoraussetzung für die Erklärung der Translokation durch das „Kräftemodell“ ist ein nativ-ähnlich strukturierter, translokationskompetenter Zustand von Wildtyp und Circularpermutante. Das „Kräftemodell“ widerspricht scheinbar den Beobachtungen aus Importexperimenten in Hefe-Mitochondrien, bei denen sich gerade die Substruktur des β_1/β_2 -Faltblattes als so stabil erweist, daß sie bei Wildtyp und den Circularpermutanten CP 2.6. sowie CP 2.7. für eine Importunterbrechung sorgt (s. Kapitel 5.5.). Vergleichend diskutiert werden beide Transportphänomene in Kapitel 5.6..

5.4.4. Fazit für die Translokation

Der Unterschied zwischen den Circularpermutanten und dem Wildtyp hinsichtlich der Translokation kann in der Struktur begründet liegen, was mit Hilfe des „Kräftemodells“ erklärt werden kann; allerdings könnte sich auch der translokationskompetente Zustand der Mutanten deutlich voneinander unterscheiden, indem verschiedene Chaperone mit den unterschiedlichen Mutanten interagieren. Ausschließen läßt sich auch nicht, daß durch Änderung der Topologie eine (im Fall CP X.6.) für die Translokation besonders entscheidende Aminosäuresequenz an eine für die Translokation besonders wichtige Position des prozessierten Proteins gelangt, wie ihn der N-Terminus darstellt (Kajava *et al.*, 2000).

Die Untersuchungen des allgemeinen sekretorischen Transportmechanismus fokussieren bislang fast ausschließlich die Bestandteile des Transportapparates

(s. Kapitel 1.6.1.). Als Substrate wurden bislang lediglich periplasmatisch exprimierte Proteine oder integrale Membranproteine verwendet, um Interaktionen z.B. mit Chaperonen bezüglich des Transportmechanismus aufzuklären. Die Auswirkungen von gezielten Mutationen an diesen Substratproteinen hinsichtlich der Translokation bzw. hinsichtlich der periplasmatischen Expression wurden bislang noch nie näher untersucht. Mit der Wahl einer in der Natur cytosolisch exprimierten Proteindomäne wurde im Fall der PH-Domäne ein Protein als Substrat verwendet, das im Laufe der Evolution nie auf verbesserte Translokation selektiert wurde. Diese positive Selektion hätte in der Optimierung der Aminosäuresequenz genauso liegen können, wie in der Änderung struktureller Eigenschaften oder in der Coexpression von Chaperonen. Deshalb eignet sich die PH-Domäne mit ihren Mutanten ganz besonders als Modells substrat zur Untersuchung von Transportvorgängen.

Die Auswirkungen von Circularpermutationen auf die Translokation wurden indirekt bereits im Fall der periplasmatisch exprimierten Disulfid Oxidoreduktase DsbA von *E. coli* untersucht (Hennecke *et al.*, 1999: s. Kapitel 5.4.2.). Dabei stellten die Autoren fest, daß sich in 70 % der Proteinbereiche neue Termini einführen lassen. In den restlichen 30 % sei dies nicht möglich, da die entsprechenden Circularpermutanten zu instabil seien. Nach obiger Diskussion könnte man dieses Ergebnis auch interpretieren, indem die Topologien dieser Mutanten eine Translokation in das Periplasma nicht oder nur unzureichend ermöglichen. Auch diese Versuche könnten somit zeigen, daß die Topologie von Proteinen für die Translokation in das Periplasma von *E. coli* außerordentlich bedeutsam ist.

5.5. Die Aussagen aus dem mitochondrialen Import der PH-Domäne

Zum Vergleich mit der periplasmatischen Translokation in *E. coli* wurden einige Mutanten der PH-Domäne mit dem mitochondrialen Signalpeptid pSu9 versehen und im mitochondrialen Importsystem untersucht. Dabei zeigte sich, daß der Import durch die innere mitochondriale Membran beim Wildtyp, bei der Mutante L(8)V und bei der Circularpermutante CP 2.7. unterbrochen wird, was nach ProteinaseK-Prozessierung aus der Western-Blot-Analyse der Mitoplasten anhand von Prozessierungsfragmenten deutlich wird (s Abb. 4.29). Bei der Circularpermutante CP 2.6. dagegen beobachtet man ein derartiges Fragment nicht, dafür aber das im Signalpeptid (*MPP1*-) prozessierte Präpro-

tein (Intermediat) im mitochondrialen Intermembranraum. Die Unterbrechung läßt sich eindeutig auf die Peptidsequenz

74 100
 –DQVIKAAKTEADGQVVEGNHTVYRISA–

zurückführen, die durch die Circularpermutanten definiert ist (s. Kapitel 4.5.1.). Die direkt aus den Fragmenten abgeleitete Schlußfolgerung für den Importverlauf bei den unterschiedlichen Varianten ist in Abb. 4.30B wiedergegeben.

Mit dieser Peptidsequenz aus der PH-Domäne konnte erstmalig beim mitochondrialen Import eine importunterbrechende Sequenz nachgewiesen und identifiziert werden. Sie ermöglichte den stabilen Transport von CP 2.6. als Intermediat in den Intermembranraum. Bei Cytochrom b_2 , einem im mitochondrialen Intermembranraum lokalisierten Protein, wurde einer 20 Aminosäuren umfassenden, hydrophoben Peptidsequenz vor dem reifen Protein zunächst auch die Rolle als *stop*-Importsignal zugeordnet. Es konnte jedoch gezeigt werden, daß diese Aminosäuresequenz bis in die mitochondriale Matrix importiert wird, dort das Signalpeptid vor ihr prozessiert wird und sie dann als Signalsequenz für die reverse Translokation fungiert (Hartl *et al.*, 1987). Nach der heutigen Modellvorstellung zur Translokation von Präproteinen (s. Kapitel 1.6.2.; Gaume *et al.*, 1998; Neupert *et al.*, 1990) könnte diese Aminosäuresequenz die Arretierung durch das mt-Hsp70-Protein in der mitochondrialen Matrix erschweren, weshalb nach der Prozessierung des Signalpeptids (wie im Fall von CP 2.6.) das Protein durch den TIM-Komplex zurückdiffundiert. Allerdings ließen sich diese Versuchsergebnisse auch so interpretieren, wie dies bei der PH-Domäne geschah: C-terminal zu der hydrophoben Peptidsequenz befindet sich bei Cytochrom b_2 eine importunterbrechende Peptidsequenz und den hydrophoben Aminosäuren steht lediglich die Rolle zu, wie sie im späteren Intermediat von CP 2.6. der Anteil des pSu9-Signalpeptids inne hat. Während beim Cytochrom b_2 diese beiden Interpretationen möglich sind und damit weder dem hydrophoben noch dem dazu C-terminalen Sequenzbereich eine eindeutige Funktion zugewiesen werden kann, erweist sich beim Wildtyp der PH-Domäne und bei CP 2.7. die oben angegebene Sequenz eindeutig als importunterbrechende Sequenz. Die Importunterbrechung ist dabei unabhängig von den Aminosäuren der restlichen PH-Domäne, denn während der Importunterbrechung ragt bei der Mutante CP 2.7. genau diese Peptidsequenz „metastabil“ aus der inneren Membran in den Intermembranraum und kann damit keine weiteren Interaktionen zum restlichen Protein eingehen. Dieses befindet sich dann im TIM-Komplex und in der Matrix.

Zur weiteren funktionellen Charakterisierung dieser Peptidsequenz wird nur sie zwischen das pSu9-Signalpeptid und ein mitochondrial gut importiertes Protein (Dihydrofolatreduktase: DHFR) kloniert, so daß dieses Protein nun in den Intermembranraum importiert werden sollte. Ergebnisse aus diesem Experiment stehen noch aus. Sie sollten die Befunde aus den Untersuchungen mit der PH-Domäne bestätigen und könnten diese Peptidsequenz dazu etablieren, (weitgehend) beliebige Proteine gezielt in den Intermembranraum zu transportieren, wie dies z.B. beim oben erwähnten Cytochrom b_2 geschieht.

Weshalb wird aber der Import durch diese Sequenz unterbrochen?

Die 27 Aminosäuren der importunterbrechenden Peptidsequenz interagieren teilweise in der Wildtyp PH-Domäne miteinander (s. Abb. 5.3). Sollte dieser Strukturbereich als stabile Subdomäne im Intermembranraum die weitere Translokation verhindern, trügen nur die Aminosäuren V(76) bis T(94) zur Stabilität bei.

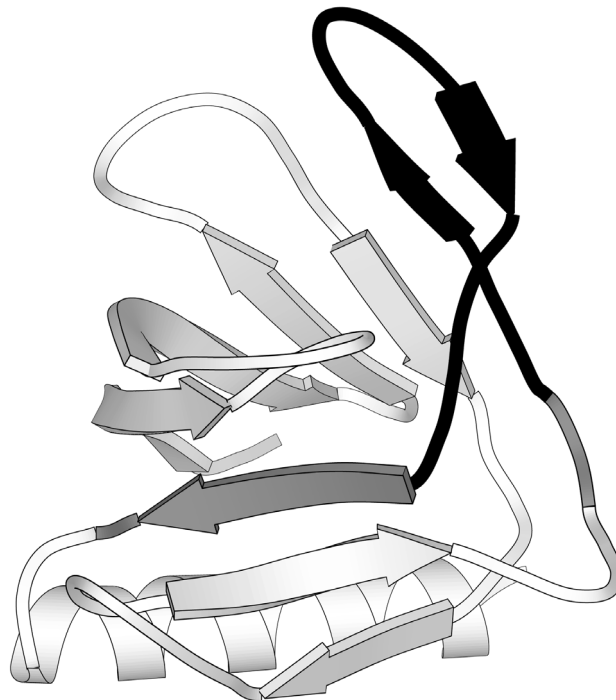


Abb. 5.3: Das importunterbrechende Peptid innerhalb der PH-Domäne.

Die Peptidsequenz zwischen dem C-Terminus von CP 2.6. und dem N-Terminus von CP 2.7. beinhaltet das importunterbrechende Peptid (grau markiert). Der schwarz markierte Bereich innerhalb dieses Peptides könnte davon eine stabile Subdomäne ausbilden, die den Import verhindert bzw. unterbricht.

Im Hinblick auf das Modell zur Translokation von Präproteinen (Gaume *et al.*, 1998; Neupert *et al.*, 1990) sollte die thermodynamische Stabilität der Subdomäne für die Unterbrechung des Imports verantwortlich sein: nach diesem Modell diffundiert ungefaltetes Protein schrittweise durch den TIM-Komplex in die Matrix, wobei dort mt-Hsp70 am Protein bindet und die Rückdiffusion verhindert. Ein Zug aus der Matrix findet dabei nicht statt. Der Import wird demzufolge an einer Peptidsequenz so lange unterbrochen, wie sie gefaltet vorliegt. Erst die Entfaltung (im Rahmen des chemischen Gleichgewichts) erlaubt ihr die weitere Diffusion in die Matrix.

Zur Bestimmung der thermodynamischen Stabilität (als Subdomäne) könnten in Zukunft Übergangskurven des importunterbrechenden Peptids aufgenommen werden (s. Kapitel 4.1.5.). Als Meßsonde wären dazu mehrere Systeme denkbar: die Fusion eines Tryptophan an den C-Terminus und eines Cysteins an den N-Terminus sollte im gefalteten Zustand zu einer Fluoreszenzlöschung und bei Entfaltung zu Fluoreszenz führen (vgl. Fluoreszenzübergangskurven von IG-Domänen). Im Rahmen von FRET-Messungen (Fluoreszenz Resonanz Energie Transfer) könnten die Termini mit je einem geeigneten fluoreszierenden Donor bzw. Akzeptor gekoppelt werden und der Abstand beider voneinander in Abhängigkeit von der Konzentration an Denaturierungsmittel über die Fluoreszenzübertragung bestimmt werden. Auch die direkte Strukturanalyse über CD-Spektroskopie oder Infrarotspektroskopie käme in Erwägung.

In Analogie zum „Kräftemodell“ für die Translokation von PH-Domänen in das Periplasma von *E. coli* (s. Abb. 5.2) wäre zudem die Kraft interessant, die zur Denaturierung dieser Subdomäne aufzuwenden ist. Dies ließe sich mittels Einzelmolekül-Kraft-Spektroskopie analysieren.

Die hohe Stabilität eines 19-er Peptids (ohne Disulfidbrücken) in Lösung ist erstaunlich, da bei so kleinen Peptiden kaum die Möglichkeit besteht, über ein hydrophobes Proteininneres und den daraus abgeleiteten „hydrophoben Effekt“ zu einer stabilen Struktur zu gelangen. Im Fall der importunterbrechenden Sequenz wäre als stabilisierendes Element die Chelatisierung eines mehrwertigen Kations über die Aminosäuren E83, D85 und Q87 möglich, die bei der PH-Domäne im Loop Li1 in eine Richtung weisen. Der Glutamat- und der Aspartatrest wurden deshalb zu Serinresten mutiert. Ob sich allerdings dadurch die Importeigenschaften der PH-Domäne ändern, wird derzeit noch untersucht.

Anstelle der Stabilität könnte auch ein Peptidmotiv aus der Peptidsequenz für die Unterbrechung des Imports verantwortlich sein. So könnten Aminosäurereste spezifisch am TIM-Komplex binden und den weiteren Import verhindern. Zur Analyse könnten Varianten von Peptidmotiven hergestellt und durch Import in Mitochondrien analysiert werden. Allerdings wird auch die Struktur

eines Proteins/Peptids durch die Aminosäuresequenz festgelegt (Anfinsen, 1973). Ob das Peptidmotiv oder aber seine Struktur für die Importunterbrechung verantwortlich ist, ließe sich folglich derartigen Analysen nicht entnehmen.

Eine Struktur derselben thermodynamischen Stabilität aber mit einer anderen Peptidsequenz erhielte man durch Einsatz der enantiomeren Peptidsequenz. Diese erhält man durch chemische Synthese des Peptids aus D-Aminosäuren. Diese müßten N-terminal mit den L-Aminosäuren eines Signalpeptids und C-terminal mit einem Nachweisreagenz gekoppelt werden. Sollte das Nachweisreagenz (im Gegensatz zu dem gleichen Konstrukt mit ausschließlich L-Aminosäuren) in die Matrix importiert werden, wäre ein Peptidmotiv und nicht die Stabilität einer Subdomäne für den Import verantwortlich.

Hydrophile, hydrophobe, kleine und positiv sowie negativ geladene Aminosäuren sind innerhalb der importunterbrechenden Peptidsequenz sehr gleichmäßig verteilt, was allgemein für *stop*-Transfersequenzen nicht beobachtet wird (von Heijne & Gavel, 1988). Diese Verteilung könnte der Sequenz zu guter Löslichkeit und autarken Faltungseigenschaften verhelfen, so daß sich eine eigenständig faltende Subdomäne ausprägen könnte. Ihre Faltung im Intermembranraum ist dabei unabhängig vom restlichen Protein, das sich währenddessen sogar schon/noch in anderen Kompartimenten aufhalten kann; zudem ist der Faltungsweg einer Subdomäne (auf der Zeitskala) naturgemäß kürzer als der ganzer Proteine. Im Gegensatz zu Subdomänen können sich ganze Proteine somit während des Imports im Intermembranraum nicht rückfalten und damit den Import unterbrechen. Dies erklärt, weshalb der Stabilität von Proteinen beim Import keine große Bedeutung zukommt und der Import der instabilen Circularpermutante CP 2.7. (-12 kJ/mol) in gleicher Weise an der inneren mitochondrialen Membran unterbrochen wie der des vergleichsweise sehr stabilen Wildtyps (-41 kJ/mol). Daß die thermodynamische Stabilität von Proteinen die Translokation wenig beeinflusst, war allerdings auch schon vor dieser Arbeit bekannt: das stabile Cytochrom b_2 wird besser in Mitochondrien importiert als die instabile Dehydrofolatreduktase (Gaume *et al.*, 1998).

Sollte sich die Importunterbrechung tatsächlich auf die Existenz einer stabilen Subdomäne zurückführen lassen, ist in diesem Zusammenhang die cytoplasmatische Expression der Mutante 30 nach der Randomisierung (s. Tab. 4.1) in *E. coli* interessant: bei ihr sind neun Aminosäuren [Q(87) bis V(95)] aus dem zentralen Bereich der importunterbrechenden Peptidsequenz V(76) bis T(94) deletiert. Trotzdem läßt sich diese Mutante löslich exprimieren, sollte also eine stabile Struktur besitzen. Die stabile Subdomäne wäre somit für die Faltung der Wildtyp PH-Domäne nicht notwendig.

5.6. Vergleich der Transporte

Im Rahmen dieser Arbeit wurden zwei Transportprozesse anhand von Mutanten der PH-Domäne studiert. Bei beiden fokussiert sich die Erklärung für die jeweiligen Beobachtungen auf die Stabilität derselben Peptidsequenz. Allerdings wird ihr im „Kräftemodell“ für die Translokation in das Periplasma von *E. coli* eine besonders günstige Entfaltung und beim Import in Mitochondrien von *S. cerevisiae* eine besonders große Stabilität zugeschrieben. Wie läßt sich dieser scheinbare Widerspruch erklären?

Die Circularpermutanten CP 2.6. und CP 3.6. lassen sich hervorragend im Periplasma exprimieren, während destabilisierende und sequenzspezifisch modifizierende Mutationen gegenüber dem (schlecht exprimierten) Wildtyp nicht zu einer deutlich verbesserten Expression führen. Die plausibelste Erklärung dieses Phänomens ist im „Kräftemodell“ dargestellt (s. Abb. 5.3). Es basiert auf der Annahme, daß die PH-Domänen im translokationskompetenten Zustand strukturiert vorliegen. Beim Wildtyp interagiert das β 1-Faltblatt als N-Terminus über einen langen Bereich von elf Aminosäuren mit den antiparallelen β -Faltblättern 2 und 7. Die aufzuwendende Zugkraft durch den sekretorischen Apparat ist bei dieser starken Interaktion des N-Terminus erheblich größer als dies bei den Circularpermutanten CP 2.6. und CP 3.6. der Fall ist. Bei ihnen interagiert das β 1-Faltblatt als erstes sekundäres Strukturelement mit dem (nicht dazu antiparallelen) β 2-Faltblatt über nur drei Aminosäuren und mit dem β 4-Faltblatt (s. Kapitel 5.4.3.). Es schließt sich der Loop Li1 als „turn“ an, so daß sich die Zugrichtung ändert. Damit läßt sich der N-Terminus der Circularpermutanten CP 2.6. und CP 3.6. mit sehr viel geringerem Kraftaufwand von der restlichen Struktur lösen, als dies beim Wildtyp der Fall ist. Ist jeweils der N-Terminus vom globulären Protein gelöst, sollte das restliche Protein mit wenig weiterem Kraftaufwand zerfallen.

Aus den mitochondrialen Importversuchen kann man schlußfolgern: die Struktur um β 1/Li1/ β 2 ist so stabil und von den umgebenen Peptidsequenz so unabhängig, daß sie eine Subdomäne ausbildet, die den Import durch den TIM-Komplex unterbricht. Der Import durch den TOM-Komplex mit einem Durchmesser von ca. 22 Å (Hill *et al.*, 1998) setzt zumindest eine Auflösung der globalen Struktur der jeweiligen PH-Domäne voraus. Da beide Importvorgänge miteinander gekoppelt sind, dürfte eine Faltung zur globalen Proteinstruktur im Intermembranraum ausgeschlossen werden können, während die Ausbildung einer Subdomäne während des Importvorgangs im Intermembran-

raum möglich ist. Dies ist ein grundlegender Unterschied zum Kräftenmodell, das die globale Struktur der PH-Domänen vor der Translokation voraussetzt.

Die Transportmechanismen werden in dieser Arbeit jeweils „aus der Sicht“ des Substrates analysiert. Dabei erweist sich die PH-Domäne mit ihren Mutanten als sehr vielseitiges Versuchsobjekt, um die unterschiedlichen Voraussetzungen für Transportvorgänge zu beleuchten. Beim Vergleich beider Transportprozesse aus der Sicht des Substrates ist zu berücksichtigen, daß sich der mitochondriale Import und die Translokation in das Periplasma fundamental evolutiv voneinander unterscheiden: nach der Endosymbiontentheorie leiten sich Mitochondrien von endozytierten Purpurbakterien ab (s. Kapitel 1.6.2.). Demzufolge verläuft der Import von Präproteinen in genau der entgegengesetzten Richtung wie die Translokation bei *E. coli*. Eine funktionelle Ähnlichkeit bekannter molekularer Transportmechanismen sollte sich folglich immer über Konvergenz, nicht aber über Homologie evolutiv erklären. Die Begünstigung unterschiedlicher Substrateigenschaften bei beiden Transportprozessen verwundert unter diesen Umständen nicht.

6. Zusammenfassung

6.1. Die PH-Domäne als Paratop

Die Pleckstrin-homologe (PH-) Domäne des humanen Cytohesin-1 besteht aus einem Proteingerüst sowie vier längeren Loops. Von diesen weisen drei in eine Richtung und bilden eine komplexe, flexible Oberflächenstruktur aus. Sollte man diese Oberflächenstruktur durch Mutation der Loops als Bindungstasche (Paratop) für Epitope von Schlüsselmolekülen etablieren können, wäre ein breiter Einsatz der PH-Domäne als Wirkstoff oder spezifisches Nachweisreagenz interessant, zumal sie sich in *E. coli* mit hohen Ausbeuten cytoplasmatisch löslich exprimieren läßt.

In dieser Arbeit konnte gezeigt werden, daß sich die drei Loops verändern lassen, ohne daß die PH-Domäne ihre Struktur verliert; von daher eignet sich die PH-Domäne als Proteingerüst. Sie wurde insgesamt in 29 Aminosäurepositionen mit einem neuartigen Verfahren gewichtet randomisiert, indem an jeder Position die Wildtyp-Aminosäure bevorzugt wird. In Anbetracht der Zahl randomisierter Positionen sollte damit gegenüber einer ungewichteten Randomisierung kein Verlust an Komplexität für die Bibliothek zu befürchten sein, durch den möglichen Erhalt lokaler und nicht lokaler Wechselwirkungen aber die Zahl stabiler (damit exprimierbarer und selektierbarer) Mutanten deutlich erhöht werden. Die Randomisierung erfolgte dabei mit drei Oligodesoxynukleotiden, die in den randomisierten Positionen jeweils eine definierte Basenverteilung aufweisen. Zur Klonierung einer Bibliothek wurden sie im dazu entwickelten Verfahren der „asymmetrischen PCR-Reaktion“ eingesetzt und daraufhin zu einem in drei Segmenten randomisierten DNA-Fragment assembliert. Mit dieser Strategie konnten $6 \cdot 10^7$ Mutanten erzeugt werden. (Aus deutlich kleineren Bibliotheken anderer Proteine ließen sich bereits bindende Mutanten isolieren.) Die randomisierten Mutanten der PH-Domäne wurden im *phage-display*-Verfahren zur Selektion gegen drei Zielsubstanzen eingesetzt. Danach konnten ausschließlich Deletionsmutanten und Mutanten mit *stop*-Codons nachgewiesen werden, die keine Expression von PH-Domänen erlauben. Zurückgeführt wird dieses Ergebnis auf die schlechten Transporteigenschaften der PH-Domäne bei der Translokation in das Periplasma von *E. coli*, weshalb nicht auf bindende Paratope aus der Bibliothek selektiert werden konnte. Nach Verbesserung der Translokationseigenschaften von

PH-Domänen sollte sich das *phage-display*-Verfahren zur Selektion bindender Mutanten fortsetzen lassen.

6.2. Die PH-Domäne als Substrat für Translokationen

Die im *phage-display*-Verfahren eingesetzten M13-Bakteriophagen assemblieren in der inneren Membran von *E. coli*. Dies setzt die Translokation der mit dem g3-Protein fusionierten PH-Domäne in das Periplasma voraus.

Die geringe periplasmatische Expression bei mehrheitlich aberranten Prozessierungen im Bereich des Signalpeptids und die geringe Darstellung auf der Phagenoberfläche veranlaßten zur Translokationsoptimierung der PH-Domäne. Während der allgemeine sekretorische Transportmechanismus von *E. coli* durch die beteiligten Membranproteine strukturell und funktionell gut verstanden ist, sind die Eigenschaften und Voraussetzungen für die Translokation von Substratproteinen (mit Signalpeptid als Präprotein bezeichnet) bislang weniger gut charakterisiert. Der „translokationskompetente“ Zustand beschreibt die Präproteine nur phänomenologisch. Für die schlechte Translokation wurden mehrere biochemische und biophysikalische Eigenschaften der PH-Domäne in Betracht gezogen und verschiedene Mutanten hergestellt, die demzufolge eine verbesserte Translokationseigenschaft aufweisen sollten. Dabei erwies sich weder die Verringerung der thermodynamischen Stabilität noch das *engineering* ausgewählter, spezifischer Sequenzelemente als translokationsbegünstigend. Wird dagegen durch Einführung neuer N- und C-Termini sowie der Verbrückung der ursprünglichen Termini mit einem Linker die Topologie verändert, können bei zwei dieser sogenannten Circularpermutanten bis zu 30-fach höhere Expressionsausbeuten im Periplasma erzielt werden. Die Circularpermutation wurde damit erstmalig erfolgreich im rationalen Proteindesign angewendet.

Die vorliegenden Ergebnisse legen nahe, daß die Mutanten der PH-Domäne vor der Translokation in einem nativ-ähnlichen Zustand gefaltet vorliegen und zur Translokation entfaltet werden müssen. Das in dieser Arbeit vorgeschlagene „Kräftemodell“ erklärt die verbesserte Translokation der Circularpermutanten CP X.6. gegenüber dem Wildtyp. Danach ist die maximale Kraft zur Entfaltung des Proteins die translokationslimitierende Größe, was sich mit Hilfe von Einzelmolekül-Kraft-Spektroskopie weiter untersuchen ließe.

Wie sich die Mutationen an der PH-Domäne bei weiteren Transportprozessen auswirken, wurde beim mitochondrialen Import analysiert. Die untersuchten Mutanten zeigten unabhängig von ihrer thermodynamischen Stabilität und ihrer periplasmatischen Expression eine Unterbrechung des Imports. Ursache dafür

ist eine Peptidsequenz von 27 Aminosäuren, die sich mit Hilfe der Circularpermutanten eindeutig identifizieren lässt. Sie führt bei der Circularpermutante CP 2.6. zu einer stabilen Expression im Intermembranraum und beim Wildtyp sowie bei der Circularpermutante CP 2.7. zu einem Verharren in der inneren Membran. Bei Mitochondrien konnte zuvor noch nie eine importunterbrechende Peptidsequenz nachgewiesen werden. Sie sollte sich zur stabilen Expression von Proteinen im Intermembranraum einsetzen lassen.

In der (modellierten) Raumstruktur der PH-Domäne interagieren 19 der 27 Aminosäuren in einem Faltblatt/*turn*/Faltblatt-Motiv. Sie könnten als stabile Subdomäne den Import unterbrechen. Diese Interpretation ergänzt ein Modell zur Translokation von Präproteinen, wonach das Präprotein vom Intermembranraum schrittweise durch die innere Membran (bzw. den TIM-Komplex) in die Matrix diffundiert und dort arretiert wird. Dadurch wird die Rückdiffusion verhindert. Die Unterbrechung des weiteren Imports währt solange, bis aufgrund des thermodynamischen Gleichgewichts die Peptidsequenz vor der Membran entfaltet vorliegt und dann in die Matrix diffundieren kann. Ergänzende Experimente zum mitochondrialen Import sind in Vorbereitung.

In dieser Arbeit konnte die PH-Domäne mit ihren Mutanten somit als Substrat für die Untersuchung von Transportprozessen etabliert werden. Die zukünftige Anwendung dieser Mutanten auf weitere Transportsysteme liegt dabei auf der Hand.

Die Bibliothek randomisierter PH-Domäne wird in Kooperation mit anderen Arbeitskreisen zur Selektion spezifisch bindender und inhibierender Mutanten eingesetzt.

7. Literaturverzeichnis

- Anderson, D.E., Peters, R.J., Wilk, B. & Agard, D.A.** (1999). α -lytic protease precursor: characterization of a structured folding intermediate. *Biochemistry* 38, 4728-4735.
- Anfinsen, C.B.** (1973). Principles that govern the folding of protein chains. *Science* 181, 223-230.
- Arie, J.P., Sassoon, N. & Betton, J.M.** (2001). Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of *Escherichia coli*. *Mol Microbiol* 39, 199-210.
- Arkowitz, R.A. & Wickner, W.** (1994). SecD and SecE are required for the proton electrochemical gradient stimulation of preprotein translocation. *EMBO J* 13, 954-963.
- Baldwin, R.L.** (1995). The nature of protein folding pathways: the classical versus the new view. *J Biomol NMR* 5, 103-109.
- Barlow, D.J. & Thornton, J.M.** (1983). Ion-pairs in proteins. *J Mol Biol* 168, 867-885.
- Beck, E. & Bremer, E.** (1980). Nucleotide sequence of the gene ompA coding the outer membrane protein II of *Escherichia coli* K-12. *Nucleic Acids Res* 8, 3011-3027.
- Beste, G., Schmidt, F.S., Stibora, T. & Skerra, A.** (1999). Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold. *Proc Natl Acad Sci U S A* 96, 1898-1903.
- Betz, S.F., Schnuchel, A., Wang, H., Olejniczak, E.T., Meadows, R.P., Lipsky, B.P., Harris, E.A., Staunton, D.E. & Fesik, S.W.** (1998). Solution structure of the cytohesin-1 (B2-1) Sec7 domain and its interaction with the GTPase ADP ribosylation factor 1. *Proc Natl Acad Sci U S A* 95, 7909-7914.
- Beutler, R., Ruggiero, F. & Erni, B.** (2000). Folding and activity of circularly permuted forms of a polytopic membrane protein. *Proc Natl Acad Sci U S A* 97, 1477-1482.
- Birnboim, C.A.K.** (1979). Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7, 1513-1523.

- Blank, M., Weinschenk, T., Priemer, M. & Schluesener, H.** (2001). Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels. Selective targeting of endothelial regulatory protein pigen. *J Biol Chem* 276, 16464-16468.
- Blomberg, N., Baraldi, E., Nilges, M. & Saraste, M.** (1999). The PH superfold: a structural scaffold for multiple functions. *Trends Biochem Sci* 24, 441-445.
- Blum, H., Beier, H. & Gross, H.J.** (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamid gels. *Electrophoresis* 8, 93-99.
- Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C. & Palmer, T.** (1998). An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* 273, 18003-18006.
- Bothmann, H. & Plückthun, A.** (1998). Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nat. Biotechnol* 16, 376-380.
- Braun, G. & Cole, S.T.** (1983). Molecular characterization of the gene coding for major outer membrane protein OmpA from *Enterobacter aerogenes*. *Eur J Biochem* 137, 495-500.
- Butler, E.T. & Chamberlin, M.J.** (1982). Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *J Biol Chem* 257, 5772-5778.
- Cadwell, R.C. & Joyce, G.F.** (1992). Randomization of genes by PCR mutagenesis. *PCR Methods Appl* 2, 28-33.
- Chatterjee, S., Suciu, D., Dalbey, R.E., Kahn, P.C. & Inouye, M.** (1995). Determination of K_m and k_{cat} for signal peptidase I using a full length secretory precursor, pro-OmpA-nuclease A. *J Mol Biol* 245, 311-314.
- Cull, M.G., Miller, J.F. & Schatz, P.J.** (1992). Screening for receptor ligands using large libraries of peptides linked to the C terminus of the lac repressor. *Proc Natl Acad Sci U S A* 89, 1865-1869.
- Cwirla, S.E., Peters, E.A., Barrett, R.W. & Dower, W.J.** (1990). Peptides on phage: a vast library of peptides for identifying ligands. *Proc Natl Acad Sci U S A* 87, 6378-6382.
- Danese, P.N. & Silhavy, T.J.** (1998). Targeting and assembly of periplasmic and outer-membrane proteins in *Escherichia coli*. *Annu Rev Genet* 32, 59-94.

- Daum, G., Bohni, P.C. & Schatz, G.** (1982). Import of proteins into mitochondria. Cytochrome b2 and cytochrome c peroxidase are located in the intermembrane space of yeast mitochondria. *J Biol Chem* 257, 13028-13033.
- Decker, J., Weinberger, K., Prohaska, E., Hauck, S., Kosslinger, C., Wolf, H. & Hengerer, A.** (2000). Characterization of a human pancreatic secretory trypsin inhibitor mutant binding to *Legionella pneumophila* as determined by a quartz crystal microbalance. *J Immunol Methods* 233, 159-165.
- Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H.** (1984). X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from *Rhodospseudomonas viridis*. *J Mol Biol* 180, 385-398.
- Devereux, J., Haerberli, P. & Smithies, O.** (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12, 387-395.
- Devlin, J.J., Panganiban, L.C. & Devlin, P.E.** (1990). Random peptide libraries: a source of specific protein binding molecules. *Science* 249, 404-406.
- Dobson, C.M.** (1999). Protein misfolding, evolution and disease. *Trends Biochem Sci* 24, 329-332.
- Donaldson, J.G. & Jackson, C.L.** (2000). Regulators and effectors of the ARF GTPases. *Curr Opin Cell Biol* 12, 475-482.
- Duenas, M. and Borrebaeck, C.A.** (1994). Clonal selection and amplification of phage displayed antibodies by linking antigen recognition and phage replication. *Biotechnology (N Y)* 12, 999-1002.
- Duong, F. & Wickner, W.** (1997). Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. *EMBO J* 16, 2756-2768.
- Economou, A.** (1999). Following the leader: bacterial protein export through the Sec pathway. *Trends Microbiol* 7, 315-320.
- Economou, A.** (2000). Bacterial protein translocase: a unique molecular machine with an army of substrates. *FEBS Lett* 476, 18-21.
- Economou, A. & Wickner, W.** (1994). SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* 78, 835-843.
- Eriksson, A.E., Baase, W.A. & Matthews, B.W.** (1993). Similar hydrophobic replacements of Leu99 and Phe153 within the core of T4

lysozyme have different structural and thermodynamic consequences. *J Mol Biol* 229, 747-769.

Famulok, M. & Mayer, G. (1999). Aptamers as tools in molecular biology and immunology. *Curr Top Microbiol Immunol* 243, 123-136.

Famulok, M., Mayer, G. & Blind, M. (2000). Nucleic acid aptamers-from selection *in vitro* to applications *in vivo*. *Acc Chem Res* 33, 591-599.

Fekkes, P., de Wit, J.G., van der Wolk, J.P., Kimsey, H.H., Kumamoto, C.A. & Driessen, A.J. (1998). Preprotein transfer to the *Escherichia coli* translocase requires the cooperative binding of SecB and the signal sequence to SecA. *Mol Microbiol* 29, 1179-1190.

Ferguson, K.M., Kavran, J.M., Sankaran, V.G., Fournier, E., Isakoff, S.J., Skolnik, E.Y. & Lemmon, M.A. (2000). Structural basis for discrimination of 3-phosphoinositides by pleckstrin homology domains. *Mol Cell* 6, 373-384.

Ferguson, K.M., Lemmon, M.A., Schlessinger, J. & Sigler, P.B. (1995). Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. *Cell* 83, 1037-1046.

Fling, S.P. & Gregerson, D.S. (1986). Peptide and protein molecular weight determination by electrophoresis using a high-molarity tris buffer system without urea. *Anal Biochem* 155, 83-88.

Freund, C., Gehrig, P., Baici, A., Holak, T.A. & Plückthun, A. (1998). Parallel pathways in the folding of a short-term denatured *scF_V* fragment of an antibody. *Fold Des* 3, 39-49.

Gates, C.M., Stemmer, W.P., Kaptein, R. & Schatz, P.J. (1996). Affinity selective isolation of ligands from peptide libraries through display on a lac repressor "headpiece dimer". *J Mol Biol* 255, 373-386.

Gaume, B., Klaus, C., Ungermann, C., Guiard, B., Neupert, W. & Brunner, M. (1998). Unfolding of preproteins upon import into mitochondria. *EMBO J* 17, 6497-6507.

Gibson, T.J., Hyvonen, M., Musacchio, A., Saraste, M. & Birney, E. (1994). PH domain: the first anniversary. *Trends Biochem Sci* 19, 349-353.

Gill, S.C. & Hoppel, P.H.v. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319-326.

Goldenberg, D.P. (1988). Genetic studies of protein stability and mechanisms of folding. *Annu Rev Biophys Biophys Chem* 17, 481-507.

Goldenberg, D.P. & Creighton, T.E. (1983). Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J Mol Biol* 165, 407-413.

Gray, A., Van Der Kaay, J. & Downes, C.P. (1999). The pleckstrin homology domains of protein kinase B and GRP1 (general receptor for phosphoinositides-1) are sensitive and selective probes for the cellular detection of phosphatidylinositol 3,4-bisphosphate and/or phosphatidylinositol 3,4,5-trisphosphate *in vivo*. *Biochem J* 344, 929-936.

Guex, N. & Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714-2723.

Hanes, J., Jermutus, L. & Plückthun, A. (2000). Selecting and evolving funct. proteins *in vitro* by ribosome display. *Methods Enzymol* 328, 404-430.

Hanes, J. & Plückthun, A. (1997). *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* 94, 4937-4942.

Hartl, F.U., Ostermann, J., Guiard, B. & Neupert, W. (1987). Successive translocation into and out of the mitochondrial matrix: targeting of proteins to the intermembrane space by a bipartite signal peptide. *Cell* 51, 1027-1037.

Haslam, R.J., Koide, H.B. & Hemmings, B.A. (1993). Pleckstrin domain homology. *Nature* 363, 309-310.

Hawkins, R.E., Russell, S.J. & Winter, G. (1992). Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *J Mol Biol* 226, 889-896.

Hendrick, J.P., Langer, T., Davis, T.A., Hartl, F.U. & Wiedmann, M. (1983). Control of folding and membrane translocation by binding of the chaperone DnaJ to nascent polypeptides. *Proc Natl Acad Sci U S A* 90, 10216-10220.

Hennecke, J., Sebbel, P. & Glockshuber, R. (1999). Random circular permutation of DsbA reveals segments that are essential for protein folding and stability. *J Mol Biol* 286, 1197-1215.

Herrmann, J.M. & Neupert, W. (2000). Protein transport into mitochondria. *Curr Opin Microbiol* 3, 210-214.

Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R. & Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature* 395, 516-521.

- Hochuli, E., Dobeli, H. & Schacher, A.** (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J Chromatogr* 411, 177-184.
- Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P. & Winter, G.** (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (F_{ab}) heavy and light chains. *Nucleic Acids Res* 19, 4133-4137.
- Houshmand, H., Froman, G. & Magnusson, G.** (1999). Use of bacteriophage T7 displayed peptides for determination of monoclonal antibody specificity and biosensor analysis of the binding reaction. *Anal Biochem* 268, 363-370.
- Huang, J., Villemain, J., Padilla, R. & Sousa, R.** (1999). Mechanisms by which T7 lysozyme specifically regulates T7 RNA polymerase during different phases of transcription. *J Mol Biol* 293, 457-475.
- Jackson, S.E.** (1998). How do small single-domain proteins fold? *Fold Des* 1998;3(4):R81-91 3, R81-R91.
- Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. & Plückthun, A.** (2001). Tailoring in vitro evolution for protein affinity or stability. *Proc Natl Acad Sci U S A* 98, 75-80.
- Jung, S., Honegger, A. & Plückthun, A.** (1999). Selection for improved protein stability by phage display. *J Mol Biol* 294, 163-180.
- Kajava, A.V., Zolov, S.N., Kalinin, A.E. & Nesmeyanova, M.A.** (2000). The net charge of the first 18 residues of the mature sequence affects protein translocation across the cytoplasmic membrane of *gram*-negative bacteria. *J Bacteriol* 182, 2163-2169.
- Kandror, O., Sherman, M., Rhode, M. & Goldberg, A.L.** (1995). Trigger factor is involved in GroEL-dependent protein degradation in *Escherichia coli* and promotes binding of GroEL to unfolded proteins. *EMBO J* 14, 6021-6027.
- Katayama, T., Furuya, M., Yamaichi, K., Konishi, K., Sugiura, N., Murafuji, H., Magota, K., Saito, M., Tanaka, S. & Oikawa, S.** (2001). Discovery of a non-peptide small molecule that selectively mimics the biological actions of calcitonin. *Biochim Biophys Acta* 1526, 183-190.
- Kay, B.K., Adey, N.B., He, Y.S., Manfredi, J.P., Mataragnon, A.H. & Fowlkes, D.M.** (1993). An M13 phage library displaying random 38-amino-acid peptides as a source of novel sequences with affinity to selected targets. *Gene* 128, 59-65.

Kay, B.K., Winter, J. & McCafferty, J. (1996). Phage Display of Peptides and Proteins - A Laboratory Manual. Academic Press.

Kim, J. & Kendall, D.A. (1998). Identification of a sequence motif that confers SecB dependence on a SecB-independent secretory protein *in vivo*. *J Bacteriol* 180, 1396-1401.

Klarlund, J.K., Guilherme, A., Holik, J.J., Virbasius, J.V., Chawla, A. & Czech, M.P. (1997). Signaling by phosphoinositide-3,4,5-trisphosphate through proteins containing pleckstrin and Sec7 homology domains. *Science* 275, 1927-1930.

Kleywegt, G.J. & Jones, T.A. (1998). Databases in protein crystallography. *Acta Crystallogr D Biol Crystallogr* 54, 1119-1131.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellenhofer, G., Hoess, A., Wolle, J., Plückthun, A. & Virnekas, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* 296, 57-86.

Knoblauch, N.T., Rudiger, S., Schonfeld, H.J., Driessen, A.J., Schneider-Mergener, J. & Bukau, B. (1999). Substrate specificity of the SecB chaperone. *J Biol Chem* 1999 Nov 26;274(48):34219-25 274, 34219-34225.

Köbnik, R. & Kramer, L. (1995). Membrane assembly of circularly permuted variants of the E. coli OmpA. *J Mol Biol* 250, 617-626.

Koide, A., Bailey, C.W., Huang, X. & Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. *J Mol Biol* 284, 1141-1151.

Kolanus, W., Nagel, W., Schiller, B., Zeitlmann, L., Godar, S., Stockinger, H. & Seed, B. (1996). Alpha L beta 2 integrin/LFA-1 binding to ICAM-1 induced by cytohesin-1, a cytoplasmic regulatory molecule. *Cell* 86, 233-242.

Ku, J. & Schultz, P.G. (1995). Alternate protein frameworks for molecular recognition. *Proc Natl Acad Sci U S A* 92, 6552-6556.

Kumamoto, C.A. (1989). Escherichia coli SecB protein associates with exported protein precursors *in vivo*. *Proc Natl Acad Sci U S A* 86, 5320-5324.

Kumamoto, C.A. & Nault, A.K. (1989). Characterization of the *Escherichia coli* protein-export gene secB. *Gene* 75, 167-175.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A* 82, 488-492.

- Laminet, A.A., Ziegelhoffer, T., Georgopoulos, C. & Plückthun, A.** (1990). The *Escherichia coli* heat shock proteins GroEL and GroES modulate the folding of the β -lactamase precursor. *EMBO J* 9, 2315-2319.
- Lemmon, M.A., Ferguson, K.M. & Schlessinger, J.** (1996). PH domains: diverse sequences with a common fold recruit signaling molecules to the cell surface. *Cell* 85, 621-624.
- Levinthal, C.** (1968). Are there pathways in protein folding? *J. Chim. Phys.* 65, 44-45.
- Li, L.P., Schlag, P.M. & Blankenstein, T.** (1997). Transient expression of SV 40 large T antigen by Cre/LoxP-mediated site-specific deletion in primary human tumor cells. *Hum Gene Ther* 8, 1695-1700.
- Lietzke, S.E., Bose, S., Cronin, T., Klarlund, J., Chawla, A., Czech, M.P. & Lambright, D.G.** (2000). Structural basis of 3-phosphoinositide recognition by pleckstrin homology domains. *Mol Cell* 6, 385-394.
- Lill, R., Crooke, E., Guthrie, B. & Wickner, W.** (1988). The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane. *Cell* 54, 1013-1018.
- Lindner, P., Bauer, K., Krebber, A., Nieba, L., Kremmer, E., Krebber, C., Honegger, A., Klinger, B., Mocikat, R. & Plückthun, A.** (1997). Specific detection of his-tagged proteins with recombinant anti-His tag scF_v-phosphatase or scF_v-phage fusions. *Biotechniques* 22, 140-149.
- Lu, Z., Murray, K.S., Van Cleave, V., LaVallie, E.R., Stahl, M.L. & McCoy, J.M.** (1995). Expression of thioredoxin random peptide libraries on the *Escherichia coli* cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interactions. *Biotechnology (N Y)* 13, 366-372.
- Luger, K., Hommel, U., Herold, M., Hofsteenge, J. & Kirschner, K.** (1989). Correct folding of circularly permuted variants of a β/α barrel enzyme *in vivo*. *Science* 243, 206-210.
- Luirink, J., High, S., Wood, H., Giner, A., Tollervey, D. & Dobberstein, B.** (1992). Signal-sequence recognition by an *Escherichia coli* ribonucleoprotein complex. *Nature* 359, 741-743.
- MacCallum, R.M., Martin, A.C. & Thornton, J.M.** (1996). Antibody-antigen interactions: contact analysis and binding site topography. *J Mol Biol* 262, 732-745.

- Mack, M., Riethmuller, G. & Kufer, P.** (1995). A small bispecific antibody construct expressed as a functional single-chain molecule with high tumor cell cytotoxicity. *Proc Natl Acad Sci U S A* 92, 7021-7025.
- Manivel, V., Sahoo, N.C., Salunke, D.M. & Rao, K.V.** (2000). Maturation of an antibody response is governed by modulations in flexibility of the antigen-combining site. *Immunity* 13, 611-620.
- Manting, E.H. & Driessen, A.J.** (2000). *Escherichia coli* translocase: the unravelling of a molecular machine. *Mol Microbiol* 37, 226-238.
- Mayer, B.J., Ren, R., Clark, K.L. & Baltimore, D.** (1993). A putative modular domain present in diverse signaling proteins. *Cell* 73, 629-630.
- McCarthy, D., Kramer, G. & Hardesty, B.** (1998). Reactivation of thermally inactivated pre- β -lactamase by DnaK, DnaJ, and GrpE. *Protein Sci* 7, 1164-1171.
- Meacci, E., Tsai, S.C., Adamik, R., Moss, J. & Vaughan, M.** (1997). Cytohesin-1, a cytosolic guanine nucleotide-exchange protein for ADP-ribosylation factor. *Proc Natl Acad Sci U S A* 94, 1745-1748.
- Mitchell, P.** (1979). Keilin's respiratory chain concept and its chemiosmotic consequences. *Science* 206, 1148-1159.
- Model, P. & Russel, M.** (1988). Filamentous bacteriophage. "The bacteriophages" (R. Calendar, ed.), Plenum New York. 2, 375-456.
- Moss, J. & Vaughan, M.** (1999). Activation of toxin ADP-ribosyltransferases by eukaryotic ADP-ribosylation factors. *Mol Cell Biochem* 193, 153-157.
- Naas, J. & Schmid, H.L.** (1972). Mathematisches Wörterbuch mit Einbeziehung der Theoretischen Physik, 3. Auflage. Akademie-Verlag / B.G. Teubner, Berlin / Stuttgart.
- Nagel, W., Schilcher, P., Zeitlmann, L. & Kolanus, W.** (1998a). The PH domain and the polybasic c domain of cytohesin-1 cooperate specifically in plasma membrane association and cellular function. *Mol Biol Cell* 9, 1981-1994.
- Nagel, W., Zeitlmann, L., Schilcher, P., Geiger, C., Kolanus, J. & Kolanus, W.** (1998b). Phosphoinositide 3-OH kinase activates the β 2 integrin adhesion pathway and induces membrane recruitment of cytohesin-1. *J Biol Chem* 273, 14853-14861.

- Nakamura, Y. & Tabata, S.** (1997). Codon-anticodon assignment and detection of codon usage trends in seven microbial genomes. *Microb Comp Genomics* 2, 299-312.
- Neupert, W., Hartl, F.U., Craig, E.A. & Pfanner, N.** (1990). How do polypeptides cross the mitochondrial membranes? *Cell* 63, 447-450.
- Nielsen, H., Engelbrecht, J., S., B. & von Heijne, G.** (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10, 1-6.
- Niggemann, M. & Steipe, B.** (2000). Exploring local and non-local interactions for protein stability by structural motif engineering. *J Mol Biol* 296, 181-195.
- Nord, K., Gunneriusson, E., Ringdahl, J., Stahl, S., Uhlen, M. & Nygren, P.A.** (1997). Binding proteins selected from combinatorial libraries of an α -helical bacterial receptor domain. *Nat Biotechnol* 15, 772-777.
- Nuttall, S.D., Rousch, M.J., Irving, R.A., Hufton, S.E., Hoogenboom, H.R. & Hudson, P.J.** (1999). Design and expression of soluble CTLA-4 variable domain as a scaffold for the display of functional polypeptides. *Proteins* 36, 217-227.
- Ohage, E. & Steipe, B.** (1999). Intrabody construction and expression. I. The critical role of V_L domain stability. *J Mol Biol* 291, 1119-1128.
- Oliphant, A.R., Nussbaum, A.L. & Struhl, K.** (1986). Cloning of randomsequence oligodeoxynucleotides. *Gene* 44, 177-183.
- Otzen, D.E. & Fersht, A.R.** (1998). Folding of circular and permuted chymotrypsin inhibitor 2: retention of the folding nucleus. *Biochemistry* 37, 8139-8146.
- Pace, C.N.** (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131, 266-280.
- Pace, C.N.** (1990). Measuring and increasing protein stability. *Trends Biotechnol* 8, 93-98.
- Padlan, E.A.** (1996). X-ray crystallography of antibodies. *Adv Protein Chem* 49, 57-133.
- Patzelt, H., Rüdiger, S., Deuerling, E., Dong, L., Schneider-Mergener, J. & Bukau, B.** (2001). Substrate Specificity of Trigger Factor. *Nova Acta Leopoldina Supplementum* 16, 157-158.

- Pawson, T.** (1995). Protein modules and signalling networks. *Nature* 373, 573-580.
- Pelham, H.R. & Jackson, R.J.** (1976). An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem* 67, 247-256.
- Pope, B. & Kent, H.M.** (1996). High efficiency 5 min transformation of *Escherichia coli*. *Nucleic Acids Res* 24, 536-537.
- Ramakrishnan, C. & Ramachandran, G.N.** (1965). Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys J* 5, 909-933.
- Randall, L.L.** (1983). Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. *Cell* 33, 231-240.
- Randall, L.L., Topping, T.B., Hardy, S.J., Pavlov, M.Y., Freistroffer, D.V. & Ehrenberg, M.** (1997). Binding of SecB to ribosome-bound polypeptides has the same characteristics as binding to full-length, denatured proteins. *Proc Natl Acad Sci U S A* 94, 802-807.
- Rasched, I. & Oberer, E.** (1986). Ff coliphages: structural and functional relationships. *Microbiol Rev* 50, 401-427.
- Rassow, J., Dekker, P.J., van Wilpe, S., Meijer, M. & Soll, J.** (1999). The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J Mol Biol* 286, 105-120.
- Rebecchi, M.J. & Scarlata, S.** (1998). Pleckstrin homology domains: a common fold with diverse functions. *Annu Rev Biophys Biomol Struct* 27, 503-528.
- Rehling, P., Wiedemann, N., Pfanner, N. & Truscott, K.N.** (2001). The mitochondrial import machinery for preproteins. *Crit Rev Biochem Mol Biol* 36, 291-336.
- Rendu, F., Eldor, A., Grelac, F., Bachelot, C., Gazit, A., Gilon, C., Levy-Toledano, S. & Levitzki, A.** (1992). Inhibition of platelet activation by tyrosine kinase inhibitors. *Biochem Pharmacol* 44, 881-888.
- Rief, M., Gautel, M. & Gaub, H.E.** (2000). Unfolding forces of titin and fibronectin domains directly measured by AFM. *Adv Exp Med Biol* 481, 129-141.
- Rost, B.** (1996). PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol* 266, 525-539.

- Rüdiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau B.** (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J* 16, 1501-1507.
- Rüdiger, S., Schneider-Mergener, J. & Bukau, B.** (2001). Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J* 20, 1042-1050.
- Sääf, A., Wallin, E., von Heijne, G.** (1998). Stop-transfer function of pseudo-random amino acid segments during translocation across prokaryotic and eukaryotic membranes. *Eur J Biochem* 251, 821-829.
- Sambrook, J., Fritsch, E.F., Maniatis, T.** (1989). *Molecular Cloning, a Laboratory Manual*. 2 edit. Cold Spring Harbor Laboratory Press, New York.
- Sänger, W., Orth, P., Kisker, C., Hillen, W. & Hinrichs, W.** (2000). The Tetracycline Repressor-A Paradigm for a Biological Switch. *Angew Chem Int Ed Engl* 39, 2042-2052.
- Schiebel, E., Driessen, A.J., Hartl, F.U. & Wickner, W.** (1991). $\Delta \mu H^+$ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* 64, 927-939.
- Schlehuber, S., Beste, G. & Skerra, A.** (2000). A novel type of receptor protein, based on the lipocalin scaffold, with specificity for digoxigenin. *J Mol Biol* 297, 1105-1120.
- Schmidt, A.M., Bloss, I. & Skerra, A.** (1998). Improved folding of aporetinol-binding protein in the periplasm of *Escherichia coli*: positive influences of dsbC coexpression and of an amino acid exchange in the vitamin A binding site. *Protein Eng* 11, 601-607.
- Schmidt, B., Hennig, B., Kohler, H. & Neupert, W.** (1983). Transport of the precursor to neurospora ATPase subunit 9 into yeast mitochondria. Implications on the diversity of the transport mechanism. *J Biol Chem* 1983 Apr 25;258(8):4687-9 258, 4687-4689.
- Schmidt, F.S. & Skerra, A.** (1994). The bilin-binding protein of *Pieris brassicae*. cDNA sequence and regulation of expression reveal distinct features of this insect pigment protein. *Eur J Biochem* 219, 855-863.
- Schmidt, M.G. & Oliver, D.B.** (1989). SecA protein autogenously represses its own translation during normal protein secretion in *Escherichia coli*. *J Bacteriol* 171, 643-649.

- Schöpfer, R.** (1993). The pRSET family of T7 promoter expression vectors for *Escherichia coli*. *Gene* 124, 83-85.
- Schulze-Gahmen, U., Rini, J.M. & Wilson, I.A.** (1993). Detailed analysis of the free and bound conformations of an antibody. X-ray structures of Fab 17/9 and three different F_{ab}-peptide complexes. *J Mol Biol* 1993 Dec 20;234(4):1098-118 234, 1098-1118.
- Shaw, G.** (1996). The pleckstrin homology domain: an intriguing multifunctional protein module. *Bioessays* 18, 35-46.
- Shevchik, V.E., Condemine, G. & Robert-Baudouy, J.** (1994). Characterization of DsbC, a periplasmic protein of *Erwinia chrysanthemi* and *Escherichia coli* with disulfide isomerase activity. *EMBO J* 13, 2007-2012.
- Silen, J.L. & Agard, D.A.** (1989). The α -lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*. *Nature* 341, 462-464.
- Sinha, N.D., Biernat, J., McManus, J. & Koster, H.** (1984). Polymer support oligonucleotide synthesis XVIII: use of β -cyanoethyl-N,N-dialkyl-amino-/N-morpholino phosphoramidite of deoxynucleosides for the synthesis of DNA fragments simplifying deprotection and isolation of the final product. *Nucleic Acids Res* 12, 4539-4557.
- Skerra, A.** (1994). Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene* 15, 131-135.
- Skerra, A.** (2000). Engineered protein scaffolds for molecular recognition. *J Mol Recognit* 13, 167-187.
- Skerra, A. & Plückthun, A.** (1991). Secretion and *in vivo* folding of the F_{ab} fragment of the antibody McPC603 in *Escherichia coli*: influence of disulphides and *cis*-prolines. *Protein Eng* 4, 971-979.
- Smith, G.P.** (1991). Surface presentation of protein epitopes using bacteriophage expression systems. *Curr Opin Biotechnol* 2, 668-673.
- Spada, S., Krebber, C. & Plückthun, A.** (1997). Selectively infective phages (SIP). *Biol Chem* 378, 445-456.
- Specthrie, L., Bullitt, E., Horiuchi, K., Model, P., Russel, M. & Makowski, L.** (1992). Construction of a microphage variant of filamentous bacteriophage. *J Mol Biol* 228, 720-724.

- Steipe, B.** (1999). Evolutionary approaches to protein engineering. *Curr Top Microbiol Immunol* 243, 55-86.
- Steipe, B., Bruhn, H., Funk, M. & Henkel, T.** (1997). Novel synthetic protein structural templates for the generation, screening and evolution of functional molecular surfaces. International Patent Publication *WO 97/45538*.
- Stephens, L.R., Jackson, T.R. & Hawkins, P.T.** (1993). Agonist-stimulated synthesis of phosphatidylinositol(3,4,5)-trisphosphate: a new intracellular signalling system? *Biochim Biophys Acta* 1179, 27-75.
- Strop, P., Marinescu, A.M. & Mayo, S.L.** (2000). Structure of a protein G helix variant suggests the importance of helix propensity and helix dipole interactions in protein design. *Protein Sci* 9, 1391-1394.
- Summers, R.G. & Knowles, J.R.** (1989). Illicit secretion of a cytoplasmic protein into the periplasm of *Escherichia coli* requires a signal peptide plus a portion of the cognate secreted protein. Demarcation of the critical region of the mature protein. *J Biol Chem* 264, 20074-20081.
- Szardenings, M., Vasel, B., Hecht, H.J., Collins, J. & Schomburg, D.** (1995). Highly effective protease inhibitors from variants of human pancreatic secretory trypsin inhibitor (hPSTI): an assessment of 3-D structure-based protein design. *Protein Eng* 8, 45-52.
- Tang, X., Nakata, Y., Li, H.O., Zhang, M., Gao, H., Fujita, A., Sakatsume, O., Ohta, T. & Yokoyama, K.** (1994). The optimization of preparations of competent cells for transformation of *E. coli*. *Nucleic Acids Res* 22, 2857-2858.
- Topell, S., Hennecke, J. & Glockshuber, R.** (1999). Circularly permuted variants of the green fluorescent protein. *FEBS Lett* 457, 283-289.
- Topping, T.B. & Randall, L.L.** (1994). Determination of the binding frame within a physiological light for the chaperone SecB. *Protein Sci* 3, 730-736.
- Tzagoloff, A. & Meagher, P.** (1972). Assembly of the mitochondrial membrane system. VI. Mitochondrial synthesis of subunit proteins of the rutamycin-sensitive adenosine triphosphatase. *J Biol Chem* 247, 594-603.
- Ungermann, C., Guiard, B., Neupert, W. & Cyr, D.M.** (1996). The Δ psi- and Hsp70/MIM44-dependent reaction cycle driving early steps of protein import into mitochondria. *EMBO J* 15, 735-744.

- Ungermann, C., Neupert, W. & Cyr, D.M.** (1994). The role of Hsp70 in conferring unidirectionality on protein translocation into mitochondria. *Science* 266, 1250-1253.
- Varnai, P., Rother, K.I. & Balla, T.** (1999). Phosphatidylinositol 3-kinase-dependent membrane association of the Bruton's tyrosine kinase pleckstrin homology domain visualized in single living cells. *J Biol Chem* 274, 10983-10989.
- Venkateswarlu, K., Gunn-Moore, F., Oatey, P.B., Tavare, J.M. & Cullen, P.J.** (1998). Nerve growth factor- and epidermal growth factor-stimulated translocation of the ADP-ribosylation factor-exchange factor GRP1 to the plasma membrane of PC12 cells requires activation of phosphatidylinositol 3-kinase and the GRP1 pleckstrin homology domain. *Biochem J* 335, 139-146.
- Vieira, J. & Messing, J.** (1987). Production of single-stranded plasmid DNA. *Methods Enzymol* 153, 3-11.
- Viguera, A.R., Serrano, L. & Wilmanns, M.** (1996). Different folding transition states may result in the same native structure. *Nat Struct Biol* 3, 874-880.
- Viti, F., Nilsson, F., Demartis, S., Huber, A. & Neri, D.** (2000). Design and use of phage display libraries for the selection of antibodies and enzymes. *Methods Enzymol* 326, 480-505.
- von Heijne, G. & Gavel, Y.** (1988). Topogenic signals in integral membrane proteins. *Eur J Biochem* 174, 671-678.
- Weiner, Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A. & Turner, R.J.** (1998). A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 93, 93-101.
- Weiss, G.A. and Lowman, H.B.** (2000). Anticalins versus antibodies: made-to-order binding proteins for small molecules. *Chem Biol* 7, R177-184.
- Wells, J.A. & Lowman, H.B.** (1992). Rapid evolution of peptide and protein binding properties in vitro. *Curr Opin Biotechnol* 3, 355-362.
- Wentzel, A., Christmann, A., Kratzner, R. & Kolmar, H.** (1999). Sequence requirements of the GPNG β -turn of the *Ecballium elaterium* trypsin inhibitor II explored by combinatorial library screening. *J Biol Chem* 274, 21037-21043.
- Wickner, W. & Leonard, M.R.** (1996). *Escherichia coli* preprotein translocase. *J Biol Chem* 271, 29514-29516.

- Widersten, M. & Mannervik, B.** (1995). Glutathione transferases with novel active sites isolated by phage display from a library of random mutants. *J Mol Biol* 250, 115-122.
- Winter, G., Griffiths, A.D., Hawkins, R.E. & Hoogenboom, H.R.** (1994). Making antibodies by phage display technology. *Annu Rev Immunol* 12, 433-455.
- Wright, R.M., Gram, H., Vattay, A., Byrne, S., Lake, P. & Dottavio, D.** (1995). Binding epitope of somatostatin defined by phage-displayed peptide libraries. *Biotechnology (N Y)* 13, 165-169.
- Wulfing, C., Lombardero, J. & Plückthun, A.** (1994). An *Escherichia coli* protein consisting of a domain homologous to FK506-binding proteins (FKBP) and a new metal binding motif. *J Biol Chem* 269, 2895-2901.
- Xu, J., Baase, W.A., Baldwin, E. & Matthews, B.W.** (1998). The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci* 7, 158-177.
- Young, K.H.** (1998). Yeast two-hybrid: so many interactions, (in) so little time... *Biol Reprod* 58, 302-311.
- Zeitlmann, L., Knorr, T., Knoll, M., Romeo, C., Sirim, P. & Kolanus, W.** (1998). T cell activation induced by novel gain-of-function mutants of Syk and ZAP-70. *J Biol Chem* 273, 15445-15452.
- Zhang, P. & Schachman, H.K.** (1996). *In vivo* formation of allosteric aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains: implications for protein folding and assembly. *Protein Sci* 5, 1290-1300.
- Zhou, M.M., Ravichandran, K.S., Olejniczak, E.F., Petros, A.M., Meadows, R.P., Sattler, M., Harlan, J.E., Wade, W.S., Burakoff, S.J. & Fesik, S.W.** (1995). Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature* 378, 584-592.
- Zhukovsky, E.A., Mulkerrin, M.G. & Presta, L.G.** (1994). Contribution to global protein stabilization of the N-capping box in human growth hormone. *Biochemistry* 33, 9856-9864.
- Zimmermann, R. & Wickner, W.** (1983). Energetics and intermediates of the assembly of Protein OmpA into the outer membrane of *Escherichia coli*. *J Biol Chem* 258, 3920-3925.

8. Anhang

8.1. Mathematische Betrachtung von Randomisierungen

Möchte man an fünf Positionen einer Polypeptidkette jeweils alle 20 Aminosäuren einsetzen, gibt es genau

$$N = 20^5 = 3.2 \cdot 10^6$$

unterschiedliche Polypeptidketten. Bei dieser Zahl scheidet eine getrennte Herstellung aller Mutanten aus, weshalb auf einen Randomisierungsansatz zurückgegriffen werden muß, wie dies auch in dieser Arbeit vorgenommen wurde. Es wird dazu eine Bibliothek von

$$s = 6 \cdot 10^7 = 18.75 \cdot N$$

hergestellt. Es stellen sich die Fragen,

- wie groß ist die Wahrscheinlichkeit, daß in einer Bibliothek von $s = 6 \cdot 10^7$ randomisierten Mutanten ausnahmslos alle N möglichen Mutanten mindestens einmal repräsentiert sind?
- wie vergrößert sich diese Wahrscheinlichkeit, wenn einige Mutante in der Bibliothek fehlen dürfen?

Bei der Behandlung dieser Fragen wird davon ausgegangen, daß keine Mutante von vornherein bevorzugt wird, daß also das Auftreten aller Mutanten gleichwahrscheinlich ist („Gleichverteilung“). Dann kann nämlich die Definition der Wahrscheinlichkeit von Laplace durch

$$\text{Anzahl der günstigen Fälle} / \text{Anzahl der möglichen Fälle}$$

benutzt werden.

Die Anzahl der möglichen Fälle beträgt bekanntlich N^s . Mathematisches Hilfsmittel zur Berechnung der Anzahl der günstigen Fälle ist der „Polynomische Satz“:

$$(x_1 + x_2 + \dots + x_N)^s = \sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!} \cdot x_1^{\varphi_1} \cdot x_2^{\varphi_2} \cdot \dots \cdot x_N^{\varphi_N}$$

dabei sind N und s beliebige natürliche Zahlen, wobei hier an die oben genannten gedacht ist. x_1, \dots, x_N sind unbestimmte, also völlig beliebige Größen. Das Symbol

$$\sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s}}$$

bedeutet hier und im folgenden, daß alle hinter diesem Symbol stehenden Summanden aufsummiert werden sollen, für die die nicht negativen Zahlen $\varphi_1, \dots, \varphi_n$ den Nebenbedingungen (hier $\varphi_1 + \dots + \varphi_N = 0$) genügen. (Im Fall $N = 2$ geht dieser Satz in den bekannteren „Binomischen Satz“ über (Naas & Schmid, 1972).

Den mathematischen Inhalt des „Polynomischen Satzes“ kann man folgendermaßen interpretieren: Bei gliedweisem Ausmultiplizieren der s (gleichen) „Polynome“ $(x_1 + x_2 + \dots + x_N)$ gibt der Faktor von $x_1^{\varphi_1} \cdot x_2^{\varphi_2} \cdot \dots \cdot x_N^{\varphi_N}$ genau die Anzahl der Möglichkeiten wieder, daß φ_1 -mal der Faktor x_1 , φ_2 -mal der Faktor x_2 , ... , φ_N -mal der Faktor x_N zusammentreffen. Mit dieser Interpretation läßt sich der „Polynomische Satz“ auf die gegebene Problemstellung anwenden: dazu denke man sich die N Mutanten über den Index i durchnummeriert. Die gliedweise Multiplikation entspricht dann der Erzeugung der Mutanten. Die Anzahl der Möglichkeiten, daß φ_1 -mal die erste, φ_2 -mal die zweite, ... , φ_N -mal die N -te Mutante erzeugt wird, ist also gleich dem Faktor von $x_1^{\varphi_1} \cdot x_2^{\varphi_2} \cdot \dots \cdot x_N^{\varphi_N}$ im „Polynomischen Satz“ - die Wahrscheinlichkeit für dieses Ereignis nach Laplace also

$$\frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!} \cdot \frac{1}{N^s}$$

Die Wahrscheinlichkeit W_N , daß alle Mutanten mindestens einmal erzeugt werden, ist dann

$$W_N = \frac{1}{N^s} \cdot \sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s \\ \varphi_i \geq 1 \text{ für alle } i}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!} = \frac{1}{N^s} \cdot \sum_{\substack{\varphi_1, \dots, \varphi_N > 0 \\ \varphi_1 + \dots + \varphi_N = s}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!}$$

Dies ist die formale Antwort auf die erste Frage. Zur Verdeutlichung erfolgt die Berechnung der Gegenwahrscheinlichkeit P (wie groß ist die Wahrscheinlichkeit, daß mindestens eine Mutante i nicht hergestellt wurde):

$$P\left(\bigcup_{i=1}^N A_i\right) = \frac{1}{N^s} \cdot \sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s \\ \varphi_i = 0 \text{ für ein } i}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!},$$

wenn man das Ereignis $\varphi_i = 0$ mit A_i bezeichnet.

(Die zweite dieser beiden Darstellungen ist kürzer, vielleicht prägnanter, die erste bringt die Nebenbedingung deutlicher zum Ausdruck und erleichtert unten die Bildung der Gegenwahrscheinlichkeit).

Diese Darstellungen beantworten prinzipiell die erste Frage. Dabei wurden bisher keine speziellen Eigenschaften von N und s herangezogen: selbst im einfachen Fall $s < N$ liefern die Formeln das richtige Ergebnis: in diesem Fall gibt es kein N -tupel von Zahlen $\varphi_i \geq 1$, deren Summe den Wert s besitzt, die Summe ist also leer und hat daher definitionsgemäß den Wert 0 .

In weniger einfachen Fällen wie dem hier vorliegenden, wenn

(*) N groß und s wesentlich größer

ist, kann man mit diesen Formeln jedoch nichts anfangen, da N^s und die auftretenden Summanden (letztere mit „wenigen“ Ausnahmen) außerordentlich große Zahlen sind.

Um in der Situation (*) konkrete Ergebnisse zu erzielen, wird die „Gegenwahrscheinlichkeit“ $1 - W_N$ betrachtet. Dies ist die Wahrscheinlichkeit für das Vorliegen des komplementären Ereignisses, im Sinne von Laplace also

Anzahl der ungünstigen : Anzahl der möglichen Fälle.

Das Gegenteil der zu W_N gehörigen Nebenbedingung „ $\varphi_i \geq 1$ für alle i “ ist „es gibt ein i mit $\varphi_i = 0$ “. Daher lautet die gesuchte Gegenwahrscheinlichkeit

$$(**) \quad 1 - W_N = \frac{1}{N^s} \cdot \sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s \\ \text{es gibt ein } i \\ \text{mit } \varphi_i = 0}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!}.$$

Um einen Überblick zu gewinnen, wird diese Gegenwahrscheinlichkeit durch einen einfachen Ausdruck nach oben abgeschätzt.

Dazu sei zunächst bemerkt, daß die Nebenbedingung „Es gibt ein i mit $\varphi_i = 0$ “ nicht ausschließt, daß auch zwei oder mehr Indizes diese Eigenschaft besitzen.

Trotzdem kommt der betreffende Summand in (**) nur einmal vor. Ersetzt man aber die Summe aus (**) durch

$$(***) \quad \sum_{i=1}^N \left\{ \sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s \\ \varphi_i = 0}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!} \right\},$$

dann tritt ein solcher Summand offenbar in dieser Summe so oft auf, wie zu ihm Indizes mit dem Wert 0 gehören. Da alle Summanden positiv sind, ist die Summe (***) größer als die in (**). Die Summe läßt sich leicht explizit auswerten: eine Vertauschung der Summationsbuchstaben φ_i und φ_N zeigt, daß alle Summen in der geschweiften Klammer denselben Wert

$$\sum_{\substack{\varphi_1, \dots, \varphi_{N-1} \geq 0 \\ \varphi_1 + \dots + \varphi_{N-1} = s}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_{N-1}!} = (N-1)^s$$

besitzen, letzteres nach dem „Polynomischen Satz“ mit $N-1$ anstelle von N und $x_1 = x_2 = \dots = x_{N-1} = 1$. Damit ergibt sich

$$1 - W_N \leq \frac{1}{N^s} N \cdot (N-1)^s = N \cdot \left(1 - \frac{1}{N}\right)^s$$

oder

$$W_N \geq 1 - N \cdot \left(1 - \frac{1}{N}\right)^s$$

Bemerkung:

Auch zur Herleitung dieser Ungleichung wurden keine speziellen Eigenschaften von N und s benutzt. Liegt aber (*) nicht vor, so könnte die linke Seite negativ, also die Aussage unbrauchbar werden.

Die Behandlung der zweiten Fragestellung verläuft analog:

Sei n beliebig mit $2 \leq n < N$ und W_n die Wahrscheinlichkeit, daß mindestens n Mutanten erzeugt werden. Dann ergibt sich die Gegenwahrscheinlichkeit $1 - W_n$, also die Wahrscheinlichkeit, daß mindestens $(N - n + 1)$ Mutanten nicht erzeugt werden, zu

$$\begin{aligned}
1 - W_n &= \frac{1}{N^s} \sum_{\substack{\varphi_1, \dots, \varphi_N \geq 0 \\ \varphi_1 + \dots + \varphi_N = s \\ \varphi_i = 0 \text{ für mindestens} \\ N-n+1 \text{ Werte von } i}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_N!} \\
&\leq \frac{1}{N^s} \cdot \binom{N}{N-n+1} \cdot \sum_{\substack{\varphi_1, \dots, \varphi_{n-1} \geq 0 \\ \varphi_1 + \dots + \varphi_{n-1} = s}} \frac{s!}{\varphi_1! \cdot \varphi_2! \cdot \dots \cdot \varphi_{n-1}!} = \frac{1}{N^s} \cdot \binom{N}{N-n+1} \cdot (n-1)^s \\
&= \binom{N}{n-1} \cdot \left(\frac{n-1}{N}\right)^s
\end{aligned}$$

$$1 - \binom{N}{n-1} \cdot \left(\frac{n-1}{N}\right)^s \leq W_n$$

Bemerkung:

Wie oben ist auch hier festzustellen, daß keine besonderen Eigenschaften der Zahlen s , N , n zur Herleitung benötigt werden, daß aber negative untere Schranken (wie offensichtlich für $s < n$ mit der Wahrscheinlichkeit 0) nicht auszuschließen sind.

Abschließend soll das anfangs angegebene Beispiel

$$N = 20^5 = 3.2 \cdot 10^6, \quad s = 6 \cdot 10^7$$

mit

$$n = N - 1$$

numerisch behandelt werden. Abkürzend sei $x := \frac{s}{N} = 18.75$.

Bei den angegebenen Schranken treten folgende Größen auf

$$\begin{aligned}
N \cdot \left(1 - \frac{1}{N}\right)^s &= N \cdot \left(1 - \frac{x}{s}\right)^s, \\
\frac{N \cdot (N-1)}{2} \cdot \left(1 - \frac{2}{N}\right)^s &= \frac{N \cdot (N-1)}{2} \cdot \left(1 - \frac{2x}{s}\right)^s \approx \frac{1}{2} \cdot N^2 \cdot \left(1 - \frac{2x}{s}\right)^s.
\end{aligned}$$

In der Wahrscheinlichkeitsrechnung ist es üblich, für große s

$$\left(1 - \frac{y}{s}\right)^s \text{ durch } e^{-y}$$

zu ersetzen. Numerische Betrachtungen zeigen, daß der hierdurch bewirkte Fehler für $y = x$ und $y = 2x$ vernachlässigbar klein ist). Mit dieser Näherung ergibt sich

$$N \cdot \left(1 - \frac{x}{s}\right)^s \approx N \cdot e^{-x} = 2.3021 \cdot 10^{-2}$$

$$\frac{1}{2} \cdot N^2 \cdot \left(1 - \frac{2x}{s}\right)^s \approx \frac{1}{2} \cdot N^2 \cdot e^{-2x} = 2.6499 \cdot 10^{-4}$$

und damit

$$W_N \geq 97.7 \% \text{ und } W_{N-1} \geq 99.974\%$$

Lebenslauf

Persönliche Daten

Name: Hartmut Volker Rohde
Geburtsdatum: 20. Mai 1969
Geburtsort: Aachen
Familienstand: ledig
Eltern: Professor Dr. Hanns-Walter Rohde
Sigrid Ingeborg Rohde geb. Jeschke
Staatsangehörigkeit: deutsch

Schulbildung & Wehrpflicht

1975 – 79 Grundschule Erzbergerallee Aachen
1979 – 88 Viktoriaschule Aachen
Juni 1988 Abitur
1988 – 89 Wehrdienst

Studium

1989 - 97 Studium der Diplomstudiengänge Biologie und Chemie an der Rheinisch-Westfälischen Technischen Hochschule Aachen.
1995 Diplom in Chemie mit dem Resultat „sehr gut“. Die Arbeit wurde am Deutschen Wollforschungsinstitut der RWTH Aachen unter Betreuung von Herrn Professor Dr. Hartwig Höcker angefertigt, und das Ergebnis patentiert.
1997 Diplom in Biologie mit dem Resultat „Auszeichnung“. Hierfür wurde die Springorum-Medaille der RWTH verliehen. Die Arbeit wurde am Mikrobiologischen Institut der RWTH Aachen unter Betreuung von Herrn Professor Dr. Klaus Wolf angefertigt.
ab Mai 1997 Doktorand am Genzentrum der Ludwig-Maximilians-Universität München bei Professor Dr. Ernst-Ludwig Winnacker im Arbeitskreis von Professor Dr. Boris Steipe.