

EVOLUTION VON DOUBLE PSI BETA-BARREL PROTEINEN  
UND ASPARTATPROTEASEN

DISSERTATION

der Fakultät für Chemie und Pharmazie  
der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines Doktors  
der Naturwissenschaften

2006

vorgelegt von

MICHAEL HULKO



Diese Arbeit wurde am Max-Planck-Institut für Entwicklungsbiologie in Tübingen in der Abteilung für Proteinevolution unter der Leitung von Andrei Lupas im Zeitraum von Januar 2003 bis Juni 2006 durchgeführt.

Die Betreuung der Arbeit in der der Fakultät für Chemie und Pharmazie der Eberhard-Karls-Universität Tübingen wurde von Prof. Gabriele Dodt übernommen.

Tag der mündlichen Prüfung: 02.11.2006

Dekan: Prof. Dr. L. Wesemann  
1. Berichterstatter: Prof. Dr. A. Lupas  
2. Berichterstatterin: Prof. Dr. G. Dodt  
3. Berichterstatter: Prof. Dr. R. Sterner

# Inhaltsverzeichnis

1	Proteinevolution .....	1
2	Zusammenfassung und Zielsetzung .....	5
3	Abkürzungen und Fachausdrücke .....	7
4	Evolution der Aspartatproteasen .....	9
4.1	Einleitung .....	9
4.2	Materialien und Methoden .....	14
4.3	Ergebnisse .....	20
4.4	Diskussion .....	30
5	Evolution der Double Psi Beta-Barrel .....	34
5.1	Einleitung .....	34
5.2	Materialien und Methoden .....	37
5.3	Ergebnisse .....	39
5.4	Diskussion .....	47
6	Prokaryontische Signaltransduktion.....	50
6.1	Einleitung .....	50
6.2	Materialien und Methoden .....	53
6.3	Ergebnisse .....	63
6.4	Diskussion .....	75
7	Anhang .....	81
7.1	Exprimierte Proteinkonstrukte .....	81
	HIV-1 Protease WT, HIV-1 Protease D25N .....	81
	HIV-1 Protease Q2C, D25N, C95A, L97C .....	81
	Trypsinogen S200A.....	82
	ClpA .....	82
	ClpA $\Delta$ N.....	84
	HPR-ClpA Fusion .....	85
	Saccharopepsin-ClpA Fusion .....	86
	GFP.....	88
	GFP-HPRBS.....	88
	GFP-ssrA .....	89
	Trap-GroEL .....	89
	PhS018.....	90
	Af1504.....	91
	Af1502.....	91
	Af HAMP .....	92
	Af HAMP A291V .....	92

	Af HAMP A291C.....	92
	Af HAMP A291F .....	93
	Af1503 TM2CC, Af1503 TM2CC D88A, Af1503 TM2CC D92A.....	93
	Af1503 EP .....	94
	Af HAMP-DHP, Af HAMP A291V -DHP .....	95
	Af HAMP-EnvZ, Af HAMP A291V-EnvZ .....	95
	Af HAMP-cTsr, HAMP A291V-cTsr .....	96
	Af HAMP-NarX, HAMP A291V-NarX .....	97
	Tar .....	98
	Tar AfHAMP.....	99
7.2	Bemerkung zur Proteinreinigung .....	100
7.3	Vektorkarten .....	101
	Vektorkarte pET30 .....	101
	Vektorkarte pGex 4T-1 .....	102
7.4	Reinheitskontrollen .....	102
7.5	Genotypen verwendeter E. coli Stämme .....	103
7.6	Rezepturen.....	103
	Polyacrylamid-Gele.....	103
	LB-Medium .....	103
	M9 Minimalmedium.....	104
7.7	Materialienachweise .....	104
7.8	Internetnachweise .....	105
8	Danksagung .....	106
9	Literatur .....	107



# 1 Proteinevolution

Proteine sind aus zwanzig verschiedenen natürlichen Aminosäurebausteinen zusammengesetzt. Es sind kettenförmige, polymere Moleküle. Mit der Aufklärung der DNA-Doppelhelixstruktur durch James Watson und Francis Crick 1953 (Watson and Crick, 1953), der Entschlüsselung des genetischen Codes durch Marshall Nirenberg und Heinrich Matthaei 1961 (Nirenberg and Matthaei, 1961) und der Entwicklung von Techniken zur Proteinsequenzierung durch Frederick Sanger 1951 (Sanger and Tuppy, 1951) sowie zur DNA-Sequenzierung 1977 durch Frederick Sanger (Sanger et al., 1977) und Allan Maxam und Walter Gilbert (Maxam and Gilbert, 1977) wurden Proteinsequenzen zugänglich. Seit dem Einsatz der Röntgenkristallographie zur Aufklärung von Proteinstrukturen von John Kendrew, der 1957 die Struktur von Myoglobin (Kendrew, 1950; Kendrew and Parrish, 1957) ermittelte, sind mehr und mehr dreidimensionale Proteinstrukturen verfügbar. Zuerst herrschte Enttäuschung über die scheinbar unregelmäßige Struktur der Proteine im Vergleich zur DNA-Doppelhelix. Später erkannte man jedoch die Notwendigkeit der Strukturkomplexität. Sie erlaubt den Proteinen, als Träger der biologischen Funktion, eine enorme Vielfalt an katalytischen Aktivitäten zu verwirklichen. Die Aminosäuresequenz, die dreidimensionale räumliche Anordnung der Seitenketten der Aminosäuren und die Proteinaktivität sind nämlich eng miteinander verbunden.

Eine Domäne ist die kleinste Einheit der Proteinstrukturen, da sie die kleinste, selbständig faltende, kompakte Struktureinheit bezeichnet. Durch Proteinvergleiche wurde deutlich, dass etwa zwei Drittel aller Proteine aus Einzellern und etwa 80% der Proteine der Metazoen aus mehreren Domänen zusammengesetzt sind (Apic et al., 2001). Deshalb wird zunächst die Domäne auch als evolutionäre Einheit angesehen. Von über 50% der bekannten Proteinsequenzen konnte bereits mindestens eine homologe Sequenz ermittelt werden (Chothia et al., 2003). Das lässt auf ein hohes Ausmaß an Domänenduplikation in der Proteinevolution schließen. Die Duplikation einer Domäne erlaubt eine schnelle Generierung neuer Funktionen durch Veränderung einer bereits existierenden Domäne, während die Domänenkopie vitale Funktionen aufrecht erhält. Proteine entwickeln sich auf der Ebene ihrer Gene. Deshalb können alle bekannten genetischen Mechanismen zur Veränderung von DNA als Grundlage für Proteinentwicklung dienen. Zu diesen Mechanismen zählen: Punktmutationen, Genduplikation, Deletionen, Genfusionen, Insertionen, Rekombinationen und zirkuläre Permutationen.

Nahe Proteinverwandtschaften lassen sich auf Sequenzebene feststellen. Ein hoher Anteil identischer Aminosäuren ist ein sicheres Indiz für Proteinverwandtschaft, da zufällige Sequenzidentitäten durch die extrem große Menge theoretisch möglicher Sequenzen ausgeschlossen werden können. Verwandte Proteine, die aus einem gemeinsamen Vorfahren divergent entstanden sind, bezeichnet man als homolog. Sequenzanalysemethoden wie BLAST und PSI-BLAST (Altschul et al., 1997) erlauben eine schnelle Identifikation von

verwandten Sequenzen. Mit feineren Sequenzanalysemethoden wie HMMer (Grundy, 1998) oder HHsenser (Soding et al., 2005) lassen sich auch noch divergentere Sequenzen als verwandt identifizieren. Für diesen Grad der Sequenzverwandtschaft prägten Da-Fei Feng und Russell Doolittle den Begriff „Twilight Zone“ (Feng and Doolittle, 1996). Entfernt verwandte Proteine können in ihrer Sequenz so divergent sein, dass sich ihre Verwandtschaft durch Sequenzvergleiche alleine nicht mehr feststellen lässt. Dann hilft der Vergleich der Proteinstrukturen, die in der Regel stärker konserviert sind als die Sequenzen. Sehr weit entfernte Proteinverwandtschaften lassen sich nur noch durch den Vergleich weniger markanter Struktur-, Sequenz- oder Funktionsgemeinschaften identifizieren und befinden sich in der „Midnight Zone“ (Orengo and Thornton, 2005). Verwandte Proteine können unterschiedliche Reaktionstypen katalysieren und zu verschiedenen Enzymklassen gehören. Selbst dann sind meist Teile des aktiven Zentrums oder Zwischenstufen des Reaktionsmechanismus noch konserviert (Todd et al., 2001). Findet man strukturelle oder funktionelle Ähnlichkeiten ohne zugrunde liegende Verwandtschaft, dann bezeichnet man Proteine in dieser Eigenschaft als analog und nimmt eine konvergente Evolution an. Zum Beispiel entstand die katalytische Triade von Serinproteasen mehrfach in verschiedenen Proteinstrukturen und in unterschiedlicher Sequenzabfolge der beteiligten Reste Serin, Histidin und Aspartat.

Es gibt zwei prominente und zueinander ähnliche Klassifikationssysteme, die Proteine aufgrund ihrer Evolution klassifizieren: SCOP (Murzin et al., 1995) und CATH (Orengo et al., 1997). Beide Systeme verwenden dieselben Bezeichnungen für die verschiedenen Verwandtschaftsgrade. Bei SCOP wird die Klassifikation manuell vorgenommen, während sie bei CATH automatisiert erfolgt. Proteine werden auf der untersten Hierarchieebene der Familie zusammengefasst, wenn die Verwandtschaft auf Sequenzebene erkennbar ist. Miteinander verwandte Familien werden in Superfamilien gruppiert. Die Verwandtschaft stützt sich innerhalb der Superfamilien nicht mehr allein auf Sequenzeigenschaften, sondern ist nur noch unter Einbeziehung von strukturellen oder funktionellen Eigenschaften feststellbar. Superfamilien mit gleicher Topologie, d.h. gleicher relativer Lage und Verknüpfung der Sekundärstrukturelemente, werden in Faltungen gruppiert. Zwischen verschiedenen Superfamilien herrscht in der Regel keine verwandtschaftliche Beziehung mehr. Das heißt dann, dass verschiedene Superfamilien die gleiche Faltung konvergent entwickelten. Die oberste Hierarchieebene ist die Klasse, die angibt welche Sekundärstrukturelemente eine Faltung hauptsächlich enthält. In CATH gibt es zwischen Klasse und Faltung die Ebene Architektur, welche die Proteinform beschreibt, ohne die topologische Verknüpfung der Sekundärstrukturelemente zu berücksichtigen. Prinzipiell sind verschiedene Faltungen nicht miteinander verwandt, da die Sequenzen und Strukturen meist sehr verschieden sind. Es gibt jedoch Beispiele, anhand derer man nachweisen kann, dass bei divergenter Evolution ein Faltungswechsel stattgefunden hat. Dies wird von genetischen



Mechanismen wie zirkulärer Permutation, Insertionen und Deletionen verursacht, welche die Topologie eines Proteins verändern können (Grishin, 2001).

Mit der Aufklärung von immer mehr Proteinstrukturen wurde deutlich, dass der hohen Sequenz- und Funktionsvielfalt nur eine relative geringe Strukturvielfalt zugrunde liegt. Es wurden zum Zeitpunkt der Arbeit etwa 70.000 bis 80.000 Domänen klassifiziert, die in etwa 1500 bis 2000 Superfamilien mit hoher Strukturähnlichkeit und etwa 1000 Faltungen mit ähnlicher Strukturtopologie fallen (Tabelle 1).

**Tabelle 1 CATH und SCOP Statistik.** Die Tabelle zeigt die Anzahl der Einträge jeder Hierarchiestufe der beiden Klassifikationssysteme CATH (Stand vom 08.05.06) und SCOP (Version 1.69 vom Juli 2005).

	Anzahl der Vertreter der Hierarchiestufen					
	<i>Klasse</i>	<i>Architektur</i>	<i>Faltung</i>	<i>Superfamilie</i>	<i>Familie</i>	<i>Domäne</i>
CATH	4	40	1110	2147	7841	86151
SCOP	11		945	1539	2845	70859

Man kann eine stark unterschiedliche Popularität von Faltungen beobachten. Fast 40% aller Familien besitzen eine von 10 Faltungen, was weniger als 1% aller Faltungen ausmacht. Und etwa 85% der Faltungen haben Vertreter aus nur einer Familie (Pearl et al., 2003). Christine Orengo bezeichnete die am stärksten besetzten Faltungen wie  $\beta$ -trefoil, Jelly roll, Immunoglobulin-like, TIM barrel, Ferredoxin-like, Updown bundle, OB-fold, UB-roll, Globin-like und Doubly wound als *superfolds* (Orengo et al., 1994). Etwa 200 Superfamilien werden in Organismen aus allen drei Organismenreichen Eukaryonten, Prokaryonten und Archaea, gefunden. Weil einige davon extrem häufig besetzt sind, betrifft diese gleichmäßige Verteilung einen Großteil aller einzelnen Domänen (etwa 50% in Bakterien). Auffällig viele Domänen sind an Reaktionen der Proteinbiosynthese beteiligt. Aber man findet ebenso Domänen aus den Bereichen Regulation und Metabolismus. Vorwiegende Proteinaktivitäten vor der Entwicklung von Organismen könnten demnach DNA-Replikation und Proteinreproduktion gewesen sein. Betrachtet man Domänen in ihrem Proteinkontext, so sind weniger als 15% der Domänenkombinationen eines Genoms in anderen Genomen wiederzufinden (Lee et al., 2005). Daraus kann man schließen, dass die maßgebliche Domänenevolution vor der Trennung in die drei Organismenreiche stattgefunden hat. Und die Kombination von Domänen hat anschließend bei der Entwicklung der Arten eingesetzt.

Bei genauer Analyse der Sekundärstrukturanteile von Proteinen fand man einen stets hohen Anteil von sogenannten Supersekundärstruktureinheiten wie  $\alpha\alpha$ - oder  $\beta\beta$ -hairpins oder  $\beta\alpha\beta$ -Elemente (Soding and Lupas, 2003). *Superfolds* sind besonders reich an Supersekundärstrukturelementen (Salem et al., 1999), wobei der Updown bundle *superfold* zu 90% aus diesen Elementen besteht. Außerdem entdeckt man wiederkehrende Peptidfragmente

ähnlicher Sequenz, Struktur und sogar Funktion in einer Vielzahl unterschiedlicher nicht-verwandter Proteine. Diese Ähnlichkeiten deuten auf einen gemeinsamen Ursprung dieser Fragmente hin. Deshalb wurde vorgeschlagen, diese immer wiederkehrenden Peptideinheiten, anstelle der Domäne, als evolutionäre Einheit zu betrachten. Andrei Lupas und Kollegen prägten für diese Peptide den Begriff „antecedent peptides“ (Lupas et al., 2001). Diese ursprünglichen Peptideinheiten mussten nicht notwendigerweise selbständig falten können. Es ist denkbar, dass sie Sekundärstrukturen nur in Gegenwart von RNA oder DNA ausbildeten. Es konnte gezeigt werden, dass solche kleineren Peptidstrukturen die kombinatorische Komplexität prebiotischer Synthesereaktionen nicht überschreiten (Keller et al., 1994) und dass die Kombination von Peptid-Subdomänen die Entstehung von Proteindomänen stark beschleunigt (Bogarad and Deem, 1999; Cui et al., 2002), was als Hinweis auf die besonders erfolgreiche Verwendung dieser Einheiten für den Aufbau von Proteindomänen gedeutet werden kann. Für das häufige Auftreten der *superfolds* sind zwei Erklärungen vorgeschlagen worden. Da sie besonders reich an Supersekundärstrukturelementen sind, ist es denkbar, dass sie schneller als andere Faltungen zu einer kompakten Proteinstruktur geführt haben und damit schneller eine neue Aktivität entwickelten, die schließlich zu einem Selektionsvorteil führte. Dies würde dann für eine konvergente Evolution einiger *superfolds* sprechen (Soding and Lupas, 2003). Ebenso ist vorstellbar, dass *superfolds* Mutationen besser als andere Faltungen tolerieren. Ein denkbarer Grund dafür könnte sein, dass die Supersekundärstrukturelemente gegeneinander verschiebbar sind und deshalb eine erhöhte strukturelle Plastizität aufweisen (Orengo and Thornton, 2005).

## 2 Zusammenfassung und Zielsetzung

Ziel dieser Arbeit war es, die Evolution der Double psi beta-barrel Proteine und Aspartatproteasen näher zu untersuchen. Dabei gab es drei Untersuchungsschwerpunkte, die in drei Kapiteln beschrieben werden.

Im ersten Schwerpunkt ging es um die Faltungen Double psi beta-barrel und Acid proteases. Die Vertreter dieser Faltungen sind beta-*barrel* Proteine. Zu den Double psi beta-barrel Proteinen gehören N-terminale Domänen von AAA-ATPasen wie Vat oder Cdc48. Zu den Acid proteases gehören die Aspartatproteasen Pepsin und HIV-1 Protease. Obwohl die Strukturen dieser Faltungen bereits als sehr ähnlich erkannt und beschrieben wurden, werden sie bisher als analoge Strukturen mit konvergenter Evolution beschrieben. Dies stützte sich vor allem auf die fehlende Sequenzähnlichkeit und geringe Unterschiede der topologischen Verknüpfung ihrer  $\beta$ -Stränge. Ihre Verwandtschaft wurde jedoch bereits von Murray Coles und Kollegen erwogen. Neben der strukturellen Ähnlichkeit, fehlten aber weitere Argumente für diese Verwandtschaft. In dieser Arbeit sollten funktionelle Untersuchungen weitere Hinweise auf die Abstammung dieser Faltungen liefern. Dazu wurden Chaperonassays durchgeführt, welche Aggregationsverhinderung als funktionelle Gemeinsamkeit von Vertretern dieser Faltungen offenbarten. Die Ergebnisse bestärkten die Vermutung, die Vertreter dieser Faltungen als entfernte Verwandte zu betrachten. Damit wurde ein weiteres Beispiel geliefert, wie sich eine Faltung durch divergente Evolution ändern kann und dies zu verwandten Proteinfaltungen führt. Ein zusätzlicher Aspekt in diesem Schwerpunkt war die Frage, ob es möglich sei, Domänen dieser Faltungen in ihrem Domänenkontext gegeneinander auszutauschen. Dafür sollte eine Domäne einer Aspartatprotease mit einer AAA-Domäne fusioniert werden, um sie in den Domänenkontext von Double psi beta-barrel Domänen zu bringen. Langfristig sollte dies die Möglichkeit eröffnen, AAA-Domänen mit Substratspezifitäten der Aspartatproteasen zu entwickeln.

Im zweiten Schwerpunkt wurde die Verwandtschaft der beiden Proteinfaltungen Double psi beta-barrel und Swapped hairpin-barrel untersucht. Die Verwandtschaft wurde aufgrund von markanten Strukturähnlichkeiten bereits vermutet. Die Sequenzähnlichkeit ist jedoch sehr gering und hinterließ Unsicherheit. Die Verwandtschaft dieser Faltungen wäre ein weiteres Beispiel für einen Faltungswechsel durch divergente Evolution. Der Hypothese über die homologen Faltungen fehlte jedoch ein plausibler Evolutionsmechanismus, der die Entstehung durch bekannte genetische Mechanismen beschreiben würde. In dieser Arbeit sollten die Strukturen der Proteine PhS018 und Af1504 untersucht werden, um Hinweise auf die Evolution der Double psi beta-barrel und Swapped hairpin-barrel Faltungen zu erhalten. Durch Aufklärung der Struktur des Proteins PhS018 von *Pyrococcus horikoshii* konnte ein weitere, neue Faltung beschrieben werden: die Faltung RiSy like-barrel. Sie steht intermediär

zwischen Double psi beta-barrel und Swapped hairpin-barrel und kann die Entstehung dieser beiden Faltungen erklären. Ein weiterer Aspekt in diesem Schwerpunkt war, die Entstehung der Double psi beta-barrel aus einfachen Supersekundärstrukturmotiven zu untersuchen. Dafür konnte ein gemeinsames, strukturell hoch konserviertes  $\beta\alpha\beta$ -Element der Faltungen Double psi beta-barrel, Swapped hairpin-barrel und RiSy like-barrel beschrieben werden, welches eine ursprüngliche Peptideinheit darstellen könnte. Neben PhS018 wurde das potentielle Double psi beta-barrel Protein Af1504 von *Archaeoglobus fulgidus* untersucht. Die Aufklärung dieser Struktur war nicht möglich. Dafür wurde für dieses Protein eine mögliche Struktur vorhergesagt. Biochemische Daten wurden gesammelt, um die Vorhersage zu unterstützen.

Im dritten Schwerpunkt wurde die Funktion eines Operons aus *Archaeoglobus fulgidus* untersucht. In diesem Operon wird das Protein Af1504 zusammen mit den Proteinen Af1502, Af1503 und Af1505 kodiert. Die funktionelle Charakterisierung der Proteine des Operons zielte zunächst auf ein besseres Verständnis der Funktion von Af1504, was im vorigen Schwerpunkt bereits unter strukturellen Aspekten untersucht wurde. Zum Operon gehört das Transmembranprotein Af1503, für das bereits eine cytoplasmatische HAMP Domäne vorhergesagt wurde, welche in prokaryontischen Signaltransduktionsproteinen weit verbreitet vorkommt. Die Struktur dieser HAMP Domäne sollte aufgeklärt werden, was durch NMR-Spektroskopie auch gelang. Dies war die erste Struktur einer HAMP Domäne überhaupt. Die HAMP Domäne lag in einer *coiled coil* Konformation vor, die bisher noch nie bei einem nativen Protein beobachtet wurde: *complementary x-da* Konformation. Aus dieser Besonderheit konnte ein molekularer Signaltransduktionsmechanismus für prokaryontische Transmembranrezeptoren abgeleitet werden. Ein weiterer Aspekt in diesem Schwerpunkt war, nach möglichen Liganden für die extraplasmatische Domäne von Af1503 zu suchen. Als Ergebnis wurde Calcium als potentieller Ligand identifiziert und die Interaktion wurde charakterisiert. Abschließend wurde eine mögliche funktionelle Interaktion der Proteine des Operons und eine Funktion des Operons für das Archaeum *A. fulgidus* vorgeschlagen und diskutiert.

### 3 Abkürzungen und Fachausdrücke

AAA	ATPase associated with various cellular activities
<i>A. fulgidus</i>	<i>Archaeoglobus fulgidus</i>
BSA	Bovine serum albumin (Rinder Serumalbumin)
CATH	Class Architecture Topology Homologous superfamily
CD	Circular Dichroismus
DNA	Desoxyribonukleinsäure
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraacetat
FPLC	Fast protein liquid chromatographie
$\epsilon_{(280)}$	Extinktionskoeffizient für Lichtabsorption der Wellenlänge 280 nm
GFP	Grün fluoreszierendes Protein
GST	Glutathion-S-Transferase
HEPES	4-(2-Hydroxyethyl)-1-Piperazinethansulfonsäure
HIV	Menschliches (human) Immunschwäche Virus
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranosid
ITC	Isothermal titration calorimetry
kD	Kilodalton
LB	Luria-Bertani
MES	2-Morpholinoethansulfonsäure
MOPS	2-Morpholinopropansulfonsäure
NMR	Nuclear magnetic resonance
OD	Optische Dichte
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
Pdbid	Identifikator einer Proteinstruktur in der „RCSB Protein Data Base“
<i>P. horikoshii</i>	<i>Pyrococcus horikoshii</i>
pI	Isoelektrischer Punkt
PMSF	Phenylmethylsulfonylfluorid
RMSD	Root-mean square deviation
RNA	Ribonukleinsäure
SCOP	Structural classification of proteins
SDS-PAGE	Sodiumdodecylsulfat-Polyacrylamid Gelelektrophorese
Tris	Tris-(hydroxymethyl)-aminomethan
Vat	Valosine-containing protein-like ATPase from <i>Thermoplasma acidophilum</i>

Aminosäuren (A: Alanin, G: Glycin, S: Serin, C: Cystein, T: Threonin, H. Histidin, K: Lysin, R: Arginin, N: Asparagin, D Asparaginsäure, Q: Glutamin, E: Glutaminsäure, V: Valin, L: Leucin, I: Isoleucin, P: Prolin, F: Phenylalanin, Y: Tyrosin, W: Tryptophan, M: Methionin)

Fachausdrücke, die sich als englische Ausdrücke in der Wissenschaftssprache bereits etabliert haben, können als deutsche Ausdrücke nur unzureichend mit der gleichen Bedeutung wiedergeben werden. In diesen Fällen wurden im Text die englischen Ausdrücke in kursiver Schrift verwendet. An dieser Stelle werden sie erläutert.

*Superfold* ist einer der zehn am stärksten bevölkerten Faltungen

*Turn* ist ein Sekundärstrukturelement, das die Richtung der Aminosäurekette innerhalb weniger Aminosäurereste umkehrt.

*Hairpin* ist eine Sekundärstrukturanordnung, bei der zwei antiparallele Sekundärstrukturelemente über einen *turn* miteinander verbunden sind.

*Loop* ist ein meist ungeordneter Bereich einer Aminosäurekette, der Sekundärstrukturelemente über größere Entfernung miteinander verbindet.

*Barrel* beschreibt ein fass-artiges Aussehen eines Proteins

*Alignment* ist eine nach bestimmten Kriterien ausgerichtete Schreibweise von Proteinsequenzen

*Coiled coil* drückt eine umeinander gewundene Anordnung von  $\alpha$ -Helices aus

*Heptad periodicity* Aus sieben Aminosäuren bestehende Wiederholungseinheit, deren Positionen durch Buchstaben *a-g* bezeichnet werden und Positionen *a* und *d* von Aminosäuren mit hydrophoben Seitenketten besetzt werden.

*Knobs-into-holes* beschreibt eine mögliche geometrische Anordnung der hydrophoben Reste einer *coiled coil* Struktur

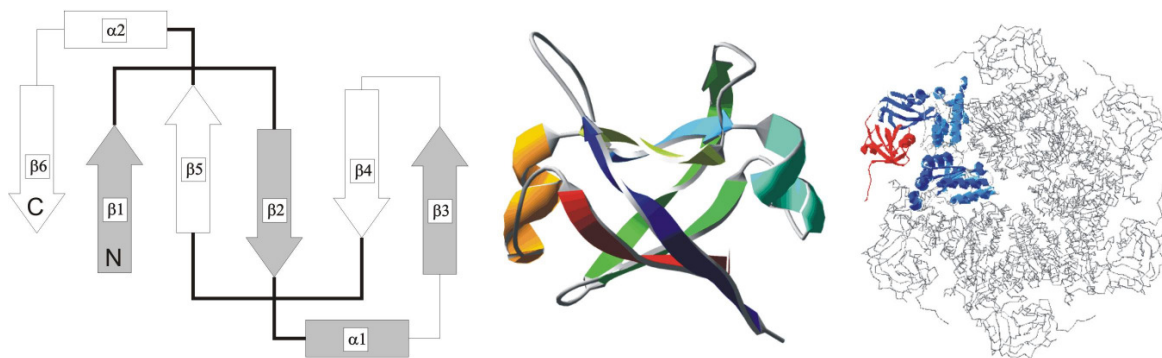
*Inclusion bodies* sind unlösliche Proteinaggregate von überexprimierten Protein

*Swarm-plate* Nährbodenplatte, mit der Schwimmverhalten von Bakterien untersucht wird

## 4 Evolution der Aspartatproteasen

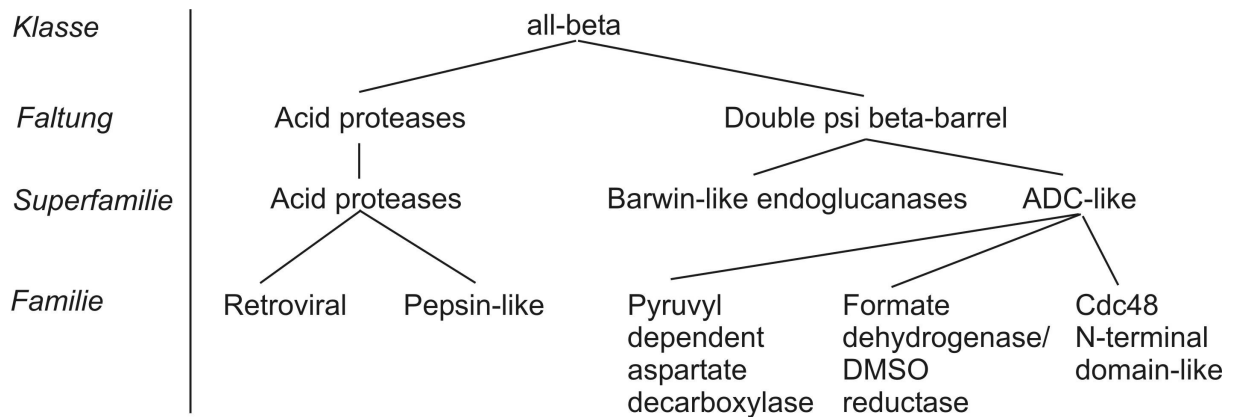
### 4.1 Einleitung

Double psi beta-barrel und Acid protease sind in SCOP als zwei Faltungen der Klasse all-beta eingetragen. Die Double psi beta-barrel Faltung beschreibt ein sechssträngiges  $\beta$ -barrel mit parallelen und antiparallelen  $\beta$ -Faltblatt Verknüpfungen (Abbildung 1, Castillo et al., 1999).



**Abbildung 1 Topologie und Struktur der Double psi beta-barrel Proteine.** (Links) Topologiediagramm der Double psi beta-barrel Faltung.  $\beta$ -Stränge sind als Pfeile und  $\alpha$ -Helices sind als quadratische Kästen dargestellt. Die pseudosymmetrischen Hälften sind durch ausgefüllte und leere Formen unterschieden. *loops* sind Linien und das charakteristische  $\Psi$ -Motiv ist fett hervorgehoben. N- und C-Terminus sind mit "N" bzw. mit "C" markiert. (Mitte) Vat-Nn (pdbid 1cz4) als Vertreter der Double psi beta-barrel Proteine. Dreidimensionale Bänderdarstellung der monomeren Domänenstruktur. Die Abfolge der Sekundärstrukturelemente vom N- zum C-Terminus ist von Blau über Grün und Gelb nach Rot farbkodiert. (Rechts) Eine Double psi beta-barrel Domäne in ihrem strukturellen Kontext einer hexameren AAA-ATPase (In dieser Abbildung ist das eukaryontische Vat Homologe p97 (pdbid 1r7r) dargestellt). Die monomere Einheit ist farbig hervorgehoben und die Double psi beta-barrel Domäne ist dabei rot dargestellt.

Das *barrel* besteht aus zwei pseudosymmetrischen, ineinander verzahnten  $\beta\beta\alpha\beta$ -Elementen. In der Double psi beta-barrel Domäne der AAA-ATPase Vat (N-terminale Domäne und deswegen als Vat-Nn bezeichnet) von *Thermoplasma acidophilum* ist die Pseudosymmetrie der  $\beta\beta\alpha\beta$ -Elemente noch auf Sequenzebene erkennbar (Golbik et al., 1999). Jeder  $\beta$ -Strang einer Hälfte wechselwirkt ausschließlich mit  $\beta$ -Strängen der jeweils anderen Hälfte. Der *loop* zwischen  $\beta 1$  und  $\beta 2$  bzw. zwischen  $\beta 4$  und  $\beta 5$  wird von  $\beta 5$  bzw.  $\beta 2$  durchkreuzt, was zu einer pseudo-verknoteten Struktur führt. In der Aufsicht sieht dieses markante und charakteristische Strukturmerkmal wie der griechische Buchstabe  $\Psi$  (psi) aus. Die Double psi beta-barrel Faltung enthält die Superfamilien „barwin-like endoglucanases“ und „ADC-like proteins“ (Abbildung 2). Zur Superfamilie der „ADC-like proteins“ gehören die Familien „Aspartat-decarboxylase“, „Formate/DMSO reductase (C-terminal domain)“ und „Cdc48 N-terminal domain-like proteins“.



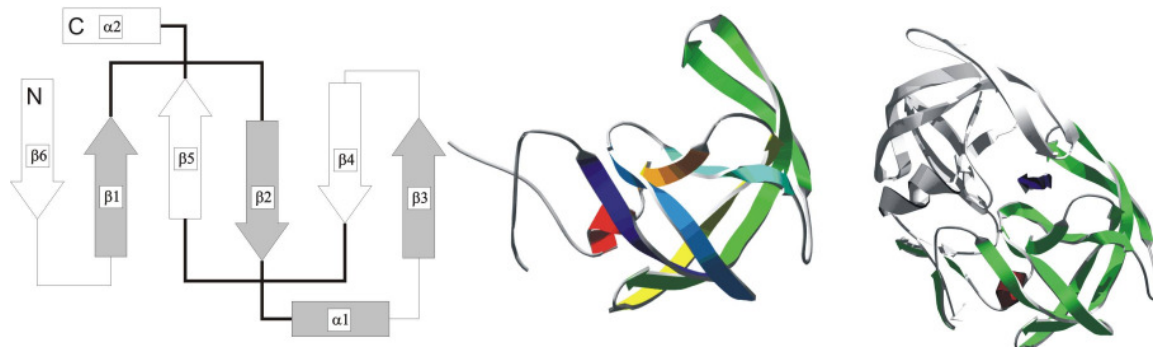
**Abbildung 2 SCOP Klassifizierung.** Links stehen die vier Hierarchieebenen der SCOP Klassifizierung. Rechts sind die Zweige der Acid proteases und Double psi beta-barrel Faltungen als Baumdiagramm abgebildet. (Nach SCOP Version 1.69 vom Juli 2005)

Vertreter der letztgenannten Familie sind beispielsweise N-terminale Domänen der AAA-ATPasen Nsf, p97, Pex1 und Vat, sowie die N-terminale Domäne von Ufd1, welches am Ubiquitin abhängigen Proteinabbau beteiligt ist. Mit Hilfe von NMR Untersuchungen von Ufd1 wurden zwei verschiedene Bindestellen für Ubiquitin gefunden (Park et al., 2005); je eine für Mono- und Polyubiquitin. Die Aminosäurereste der Bindungsstellen liegen jeweils in den *psi-loops* und den räumlich benachbarten  $\alpha$ -Helices. N-terminale Domänen von AAA-ATPasen sind an Protein-Protein-Wechselwirkungen beteiligt und dienen der Substraterkennung oder Regulierung der ATPase-Aktivität. Wechselwirkungen zwischen der N-terminalen Domäne von p97 und dem Adapterprotein p47 wurden in Kristallstrukturen des Komplexes untersucht (Dreveny et al., 2004). p47 bindet zwischen beide N-terminalen Domänen von p97. An der Bindung sind Reste der Double psi beta-barrel Domäne beteiligt, die im ersten *psi-loop* und in den  $\beta$ -Strängen zwischen beiden *psi-loops* liegen. In *in vitro* Versuchen mit den beiden N-terminalen Domänen von Vat konnte Chaperonaktivität nachgewiesen werden (Golbik et al., 1999). Die native Funktion der N-Domänen scheint aber die Regulation der Substratentfaltung zu sein (Gerega et al., 2005). Auch Vat-Nn, die N-terminale Domäne von Vat, zeigt Chaperonaktivität in Assays zur Aggregationsvermeidung (unveröffentlichte Daten von Sergej Djuranovich, Max-Planck-Institut für Entwicklungsbiologie, Tübingen).

Die Faltung Acid protease ist der Double psi beta-barrel Faltung sehr ähnlich. Es handelt sich ebenfalls um ein sechssträngiges  $\beta$ -barrel mit fast gleichen Strang-Verknüpfungen. Der wesentliche Unterschied zwischen beiden Faltungen ist, dass der C-terminale Strang  $\beta_6$  der Double psi beta-barrel bei den Acid proteases N-terminal liegt. Er wird aber ebenso als  $\beta_6$  bezeichnet, um die topologische Äquivalenz widerzuspiegeln (Abbildung 3). Es gibt noch weitere Eigenheiten der Acid protease Faltung. In keinem Vertreter ist Pseudosymmetrie auf Sequenzebene festzustellen. Die  $\alpha$ -Helix  $\alpha_1$  ist in den meisten Proteasen degeneriert. Nur bei Rous Sarcoma Virus Retropepsin ist eine rudimentäre



Helix mit einer knappen Umdrehung erkennbar. Bei der HIV-1 Protease ist an die Stelle der Helix ein antiparalleles zweisträngiges  $\beta$ -Faltblatt getreten, das in der Protease das aktive Zentrum nach Substratbindung von der Umgebung abschirmt (Abbildung 3).



**Abbildung 3 Topologie und Struktur der Aspartatproteasen.** (Links) Topologiediagramm der Acid protease Faltung;  $\beta$ -Stränge sind als Pfeile und  $\alpha$ -Helices sind als quadratische Kästen dargestellt. Die pseudosymmetrischen Hälften sind durch ausgefüllte und leere Formen unterschieden. *loops* sind Linien und das charakteristische  $\Psi$ -Motiv ist fett hervorgehoben. N- und C-Terminus sind mit "N" bzw. mit "C" markiert. (Mitte) HIV-1 Protease (pdbid 5hvp) als Vertreter der Aspartatproteasen. Dreidimensionale Bänderdarstellung der monomeren Domänenstruktur. Die Abfolge der Sekundärstrukturelemente vom N- zum C-Terminus ist von Blau über Grün und Gelb nach Rot farbkodiert. (Rechts) Die Domäne in ihrem strukturellen Kontext. HIV-1 Protease ist ein Dimer. Der blaue Pfeil stellt Substrat dar und markiert damit die Substratbindetasche. Die monomere Einheit ist farbig hervorgehoben.

Manche Proteasen besitzen zusätzliche N- und C-terminale  $\beta$ -Stränge, die ein viersträngiges  $\beta$ -Faltblatt mit einer weiteren Proteasedomäne bilden und darüber dimerisieren.

Pepsin und HIV-1 Protease sind Vertreter der Aspartatproteasen. Sie sind in SCOP in den beiden Familien Pepsin-like und Retroviral der Superfamilie Acid proteases zu finden (Abbildung 2). Die eukaryontischen Vertreter der Pepsin-like Familie, zu denen z.B. Pepsin, Gastricin, Renin und Cathepsin E gehören, bestehen aus zwei interagierenden Domänen einer Peptidkette. Diese beiden Domänen werden als homolog angesehen, die durch eine Genduplikation einer einzelnen Domäne mit anschließender Genfusion und Diversifikation entstanden sind (Tang et al., 1978). Aspartatproteasen von Retroviren, wie HIV-1 Protease oder Rous Sarcoma Virus Retropepsin, gehören zur retroviralen Familie und sind Homodimere einer einzelnen Domäne (Wlodawer et al., 1989). Von dieser Domäne nimmt man an, dass sie mit den Domänen der eukaryontischen Aspartatproteasen verwandt ist (Lin et al., 1994). Obwohl keine Sequenzähnlichkeit zwischen den viralen und den eukaryontischen Proteasedomänen nachgewiesen werden kann, bestehen Ähnlichkeiten in der Tertiärstruktur und der topologischen Anordnung der Sekundärstrukturelemente. Das aktive Zentrum befindet sich bei beiden Familien genau zwischen den *psi-loops*. Beide verwenden zwei katalytisch aktive Aspartatreste im aktiven Zentrum, wobei zwei interagierende Domänen je einen Aspartatrest zum aktiven Zentrum beisteuern. Die Aspartatreste sind

aktivierende Liganden eines Wassermoleküls, welches die Aminogruppe einer Peptidbindung nukleophil substituieren kann und damit zur Peptidkettenspaltung führt. Pepstatin ist ein gemeinsamer Inhibitor. Es ist ein nicht spaltbares Peptidanalogen, das mit hoher Affinität reversibel an das aktive Zentrum bindet. Die Substratspezifität kann jedoch sehr verschieden sein. Pepsin zeigt als Verdauungsprotease kaum Substratspezifität, während HIV-1 Protease als virale Prozessierungspeptidase Erkennungssequenzen von etwa acht Aminosäuren Länge hochspezifisch spaltet (Dunn et al., 1994).

Während die Phylogenie innerhalb beider Familien klar beschrieben ist (Dunn et al., 2002; Kageyama, 2002), findet man unterschiedliche Meinungen zur evolutionären Beziehung zwischen beiden Familien. Jordan Tang und Kollegen postulierten die Abstammung der eukaryontischen Proteasen aus den viralen, nachdem virale Proteasen im Genom des eukaryontischen Wirts zurückgeblieben sind (Tang et al., 1978). Im Gegensatz dazu, postulierten Mohana Rao und Kollegen die Abstammung der viralen Proteasen von den eukaryontischen (Rao et al., 1991). Demnach soll nach einer Gendeletion eine einzelne Domäne unter dem Selektionsdruck der beschränkten viralen Kodierungskapazität vorteilhaft gewesen sein. Unterstützt wird diese Meinung von der Idee, dass Proteasen hoher Substratspezifität von Proteasen geringer Substratspezifität abstammen (Neurath, 1984). Das würde nämlich ebenfalls bedeuten, dass die hochspezifischen viralen Prozessierungsproteasen von den relativ unspezifischen eukaryontischen Verdauungsproteasen abstammen. Als dritte Möglichkeit wird die gemeinsame Abstammung von einer ursprünglichen Protease diskutiert, wobei keine der beiden Familien aus der anderen entstanden sein soll, sondern sich divergent aus einem gemeinsamen Vorfahren entwickelten (Rao et al., 1991). Als Gegenargument wurde verwendet, dass keine Aspartatprotease in Prokaryonten existiere. Dies würde nämlich auf eine alte, vermeintlich ursprüngliche Aspartatprotease hindeuten, die schon vor der Trennung der Organismenreiche existierte. In dieser Arbeit wird gezeigt, dass Aspartatproteasen eine inhärente Chaperonaktivität besitzen. Chaperone kommen mit der ähnlichen Double psi beta-barrel Faltung in allen Organismenreichen vor und bieten sich damit als potentieller, gemeinsamer Vorgänger der beiden Aspartatprotease an. Dies wurde, ohne Kenntnis der inhärenten Chaperoneigenschaft der Aspartatproteasen, von Murray Coles und Kollegen postuliert (Coles et al., 1999).

Die Faltungen Double psi beta-barrel und Acid protease können topologisch durch eine zirkuläre Permutation ineinander überführt werden, indem der N-terminale  $\beta$ -Strang der Double psi beta-barrel an den C-Terminus transloziert wird bzw. der C-terminale  $\beta$ -Strang der Aspartatproteasen an den N-Terminus. Diese zirkuläre Permutation bietet einen möglichen genetischen Mechanismus, durch den es bei divergenter Evolution zu einem Faltungswechsel gekommen sein könnte. Beide Faltungen besitzen komplexe Struktur motive mit überkreuzenden, pseudoverknoteten  $\beta$ -Strängen, was ein charakteristisches Strukturmerkmal ist. Keine der beiden Faltungen gehört zu den evolutionär besonders günstigen und stark bevölkerten *superfolds*. Deswegen sind die komplexen *psi-loops* als analoges

Strukturmerkmal durch konvergente Evolution wenig plausibel. Wobei sich aus den *psi-loops* aber umgekehrt auch keine Verwandtschaft zwischen beiden Faltungen ableiten lässt. Gegen eine Verwandtschaft dieser Faltungen spricht auch die fehlende Sequenzähnlichkeit zwischen deren Vertretern. Dies veranlasste Richard Castillo und Kollegen, eine konvergente Entstehung von Aspartatproteasen und Double psi beta-barrel Proteinen zu postulieren (Castillo et al., 1999). Verwandtschaft zwischen Aspartatproteasen und Double psi beta-barrel Proteinen wurde hingegen von Murray Coles und Kollegen vorgeschlagen (Coles et al., 1999). Der letzte gemeinsame Vorfahre soll demnach eine symmetrische Double psi beta-barrel Domäne mit Chaperonaktivität gewesen sein. Von dieser Domäne ausgehend, entstanden Cdc48 N-terminal domain-like Proteine durch Fusion mit einer AAA-Domäne und anschließender Entwicklung regulatorischer Substraterkennungsfunktionen. Ein anderer Zweig beschreibt eine zirkuläre Permutation und eine neu entwickelte, proteolytische Aktivität und mündet im potentiellen Vorfahren der Aspartatproteasen. Eine weitere Verzweigung führt zu den viralen und über Genduplikation und Genfusion zu den eukaryontischen Proteasen.

Um die verwandtschaftliche Beziehung zwischen Aspartatproteasen und Double psi beta-barrel Proteinen genauer zu beleuchten, wurden funktionelle Studien zur Protein-Protein-Wechselwirkung dieser Proteine durchgeführt. Chaperone können entfaltete oder nicht nativ gefaltete Proteine binden und dadurch ihre Aggregation vermeiden oder ihre Rückfaltung unterstützen. Auch Proteasen binden entfaltete Proteine, spalten allerdings gebundenes Substrat in nachfolgenden hydrolytischen Schritten. Bindung entfalteter Peptide als Gemeinsamkeit führte zur Idee, dass Proteasen und Chaperone von einem gemeinsamen peptidbindenden Vorfahren abstammen könnten. Durch Einführung gezielter Mutationen konnte im Labor bereits ein Chaperon in eine Protease umgewandelt werden (Quemeneur et al., 1998). Dies verdeutlicht die Möglichkeit der Abstammung der Proteasen von Chaperonen. Im erwähnten Fall konnte das Chaperon Cyclophilin aufgrund seiner Fähigkeit, Peptide zu binden, in eine Serinprotease umgewandelt werden.

In dieser Arbeit wurde untersucht, ob proteolytisch inaktivierte Aspartatproteasen Chaperonaktivität besitzen und damit eine funktionelle Ähnlichkeit zu Cdc48 N-terminal domain-like Proteinen besitzen. Gemeinsamkeit auf funktioneller Ebene würde einen evolutionären Zusammenhang zwischen Double psi beta-barrel Proteinen und Aspartatproteasen bekräftigen, der bisher aufgrund topologischer Ähnlichkeit lediglich angenommen wurde. Dafür wurden Chaperonassays, die auf Aggregationsvermeidung beruhen, mit proteolytisch inaktivierten Aspartatproteasen Pepsin und HIV-1 Protease durchgeführt. Um zu überprüfen, ob die Idee von Chaperonen als Vorfahren von Proteasen generalisiert werden kann, wurden Trypsin als Serinprotease und Thermolysin als Metalloprotease ebenfalls auf ihre Chaperonaktivität getestet.

Darüber hinausgehend sollte geprüft werden, ob die Faltungen Acid protease und Double psi beta-barrel gegenseitig austauschbar sind und welche Rolle ihr Domänenkontext

spielt. Double psi beta-barrel Proteine findet man häufig N-terminal zu AAA-Domänen. Beispiele dafür sind Vat, p97 oder Cdc48. Deshalb wurden in dieser Arbeit Acid protease Domänen von HIV-1 Protease und Saccharopepsin mit der AAA+-Domäne des prokaryontischen ClpA-Proteins fusioniert. ClpA ist am Qualitätskontrollsystem der prokaryontischen Proteinbiosynthese beteiligt und ist gut charakterisiert (Sauer et al., 2004). Werden Proteine durch Leserasterfehler unkontrolliert und unterminiert synthetisiert, so werden sie mit einer Peptidmarkierung versehen, z.B. dem sogenannten „ssrA-tag“: AANDENYALA. Diese Markierung wird von ClpA erkannt, was zur Proteinfaltung und danach zum Proteinabbau durch die Protease ClpP führt. Hier sollte nun überprüft werden, ob die Substitution der nativen ClpA N-Domäne durch die HIV-1 Protease oder die C-Domäne von Saccharopepsin möglich ist und die Proteinfaltungsaktivität von ClpA erhalten bleibt. Ferner sollte geklärt werden, ob die Substratspezifität der chimären AAA+-ATPase vom ssrA-tag auf eine Schnittstellenerkennungssequenz der HIV-1 Protease umgeprägt wird.

## 4.2 Materialien und Methoden

**Klonierung der HIV-1 Proteasen** Felix Müller-Sarnowski (Max-von-Pettenkofer-Institut, München) stellte das Plasmid pBD2 zur Verfügung, welches die Gensequenz der wildtyp HIV-1 Protease trägt. Dieses Plasmid wurde als Templat für PCR verwendet. Über terminale Primer wurden Restriktionsschnittstellen für NdeI und HindIII eingeführt, über die das PCR Produkt in den Expressionsvektor pET30b (Novagen) ligiert wurde. Mutationen wurden mit entsprechenden Primern und Overlap-Extension PCR eingeführt. Das Konstrukt HPR-WT trägt die wildtyp Sequenz, HPR-M enthält die Mutation D25N und HPR-DS enthält die Mutationen Q2C, D25N, C95A und L97C. Das Expressionsplasmid wurde für die Expression in *E. coli* C41 (DE3) Zellen transformiert. Zur Expression wuchsen die Zellen in LB-Medium bei 37° C und wurden bei einer optischen Dichte von ~ 0,6 mit 1 mM IPTG induziert. Nach vier Stunden wurden die Zellen geerntet, in 50 mM Tris pH 8, 50 mM NaCl resuspendiert und lysiert. Unlösliche *inclusion bodies* wurden aus dem unlöslichen Teil der Lysesuspension vorgereinigt. Dafür wurden sie solange in 50 mM Tris, pH 8, 50 mM NaCl, 5 mM EDTA, 2 M Harnstoff und 0,5% Triton X-100 suspendiert und anschließend durch Zentrifugation vom Überstand wieder getrennt bis der Überstand klar blieb. Dann wurden sie in 50 mM Tris, pH 7,5 mit 8 M Harnstoff und 10 mM DTT aufgelöst. Es folgte die Reinigung des Proteins durch Kationenaustauschchromatographie (SP Sepharose FastFlow, 10 ml, GE Healthcare) unter denaturierenden Bedingungen in 25 mM Tris, pH 7,5 mit 8 M Harnstoff mit einem linearen NaCl Gradienten über 10 Säulenvolumen von 0-500 mM NaCl zur Elution. Fraktionen mit HIV-1 Protease wurden vereinigt und über Nacht gegen 30 mM Ameisensäure, pH 2,8 dialysiert. Die Rückfaltung wurde nach der Vorschrift von John Louis und Mitarbeitern durchgeführt (Louis et al., 2003). Dazu wurde Proteinlösung auf 0,3 mg/ml verdünnt und durch Zugabe von 10 mM Essigsäure pH 6 auf das sechsfache des Ausgangsvolumens

verdünnt und anschließend über Nacht bei Raumtemperatur gegen 20 mM Natriumphosphatpuffer pH 5,8 dialysiert. Zuletzt wurde die Proteinlösung durch Zentrifugation von Präzipitat befreit und konzentriert. Die Menge freier Thiolgruppen wurde nach der Methode von Ellman (Ellman, 1959) abgeschätzt, die Proteinmenge wurde über Absorption bei 280 nm bestimmt und die Reinheit wurde mit SDS-PAGE nachgewiesen.

**Herstellung von Pepsin und Pepsin C-Domäne** Schweinepepsin wurde als lyophilisiertes Pulver mit Laktose als Stabilisator gekauft. Das Pulver wurde in 50 mM MES pH 6,5, 100 mM NaCl gelöst und enthaltene Laktose wurde mit einer NICK oder PD10 Entsalzungssäule (GE Healthcare) entfernt. Die Proteinmenge wurde über Absorption bei 280 nm und dem berechneten Extinktionskoeffizienten ( $\epsilon_{(280\text{ nm})} = 51850\text{ M}^{-1} * \text{cm}^{-1}$ ) bestimmt. Die Pepsin C-Domäne wurde durch limitierten proteolytischen Abbau von Pepsin hergestellt. Dazu wurde Pepsin in 30 mM HEPES, pH 7,3, 100 mM NaCl gelöst, wie oben beschrieben von Laktose befreit und zwei Stunden bei Raumtemperatur inkubiert, um die pH empfindliche N-Domäne bei neutralem pH-Wert zu denaturieren. Anschließend wurde Proteinase K (0,1 % der Masse von Pepsin) zugegeben und weitere 30 Minuten bei Raumtemperatur inkubiert. Nach Zugabe von 1 mM PMSF wurde Proteinase K von der Pepsin C-Domäne durch Anionenaustauschchromatographie (MonoQ, 1ml, GE Healthcare) in 30 mM HEPES, pH 7,3, 100 mM NaCl mit linearem NaCl Gradienten über 20 Säulenvolumen von 100-600 mM NaCl getrennt und die vereinigten Fraktionen mit der Pepsin C-Domäne wurden über Nacht gegen 30 mM HEPES, pH 7,3, 100 mM KCl dialysiert. Die Proteinmenge wurde über Absorption bei 280 nm und dem berechneten Extinktionskoeffizienten ( $\epsilon_{(280\text{ nm})} = 27550\text{ M}^{-1} * \text{cm}^{-1}$ ) bestimmt und die Reinheit wurde mit SDS-PAGE nachgewiesen.

**Herstellung von Thermolysin** Lyophilisiertes Thermolysin wurde in 30 mM HEPES pH 7,3, 100 mM KCl, 2 mM 1,10-Phenanthrolin mit einer Endkonzentration von 10  $\mu\text{M}$  gelöst und über Nacht gegen den gleichen Puffer dialysiert.

**Herstellung von Trypsin S200A** pET3a-Plasmid, welches die Gensequenz von menschlichem Trypsinogen trägt, wurde von Diethard Mattanovich (Universität für Bodenkultur, Wien) zur Verfügung gestellt. Das aktive Serin 200 (übliche Positionsbezeichnung bezogen auf Chymotrypsin) wurde mit Hilfe des „Site-directed Mutagenesis Kit“ (Stratagene) gegen Alanin ausgetauscht und das Plasmid wurde in *E. coli* BL21 (DE3) Gold Zellen transformiert. Exprimiert wurde bei 37° C für vier Stunden, nachdem mit 1 mM IPTG bei einer OD von ~ 0,6 induziert wurde. Die Reinigung folgte gemäß der Vorschrift von Hubertus Hohenblum und Mitarbeitern (Hohenblum et al., 2004) mit einem zusätzlichen Anionenaustauschchromatographieschritt. Nach der Zelllyse wurden *inclusion bodies* solange in 50 mM Tris, pH 8, 5 mM EDTA und 0,5% Triton suspendiert und anschließend durch Zentrifugation vom Überstand wieder getrennt bis der Überstand klar

blieb. Dann wurden sie in 30 mM Tris, pH 8,6 mit 8 M Harnstoff, 1 mM EDTA und 10 mM DTT aufgelöst und für Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) unter denaturierenden Bedingungen in 30 mM Tris, pH 8,6, 1 mM EDTA mit 8 M Harnstoff und einem linearen NaCl Gradienten von 0-750 mM NaCl über 15 Säulenvolumen zur Elution verwendet. Trypsinogen ist ein cysteinreiches Protein und während der Rückfaltung können viele nicht-native Disulfidbrücken entstehen. Deshalb wurden Thiolgruppen mit 10 mM oxidiertem Glutathion (24 Stunden, 4° C) und durch resultierende gemischte Disulfide blockiert. Anschließend wurde die Proteinlösung gegen 100 mM Tris, pH 8,6, 8 M Harnstoff und 1 mM EDTA dialysiert, um oxidiertes Glutathion zu entfernen. Zur Rückfaltung wurde Proteinlösung mit 8 mg/ml 40-fach auf 0,2 mg/ml rasch verdünnt, indem die Lösung in 50 mM Tris pH 8,6, 50 mM CaCl<sub>2</sub>, 3 mM reduziertes Glutathion, 0,3 mM oxidiertes Glutathion getropft wurde. Die Lösung stand dann über Nacht bei 4° C, damit sich Disulfidbrücken formen konnten. Rückgefaltetes Trypsinogen wurde durch Zentrifugation von Präzipitat befreit, konzentriert und mit Enterokinase (Roche Diagnostics) nach Herstellerangaben prozessiert. Nach Zugabe von 1 mM PMSF wurde die Trypsinlösung auf einer Gelfiltrationssäule in 50 mM Tris, pH 8, 25 mM CaCl<sub>2</sub> (Superdex G75 26/60, GE Healthcare) von Nebenprodukten befreit. Die Proteinmenge wurde über Absorption bei 280 nm bestimmt und die Reinheit wurde mit SDS-PAGE nachgewiesen.

**Chaperonassays** Chaperonaktivität wurde über das Maß der Fähigkeit der Aggregationsverhinderung von Substratproteinen gemessen. Aggregation wurde über den Anstieg des Absorptionssignals bei 340 nm als Folge von Lichtstreuung an Proteinaggregaten in einem Lambda 25 Spektrophotometer (Perkin Elmer) bestimmt. Das Photometer war für Messung bei definierter Temperatur mit einer temperierbaren Küvettenhalterung ausgestattet. Aggregation wurde entweder chemisch oder thermisch induziert. Bei chemisch induzierter Aggregation wurde Substratprotein in 6 M Guanidiniumchlorid gelöst und Aggregation wurde durch 100-fache Verdünnung in den Assaypuffer ausgelöst. Bei thermisch induzierter Aggregation wurde die Aggregation durch Entfaltung der Substratproteine durch erhöhte Temperatur ausgelöst. Die Substratproteine Rhodanese und Luciferase wurden für chemisch induzierte Aggregationsassays bei 25° C eingesetzt. Endkonzentrationen waren 1 µM Rhodanese und 0,15 µM Luciferase. 0,3 µM Citratsynthase und 0,15 µM Luciferase wurden als Substratproteine für thermisch induzierte Aggregation bei 50° C bzw. 48° C verwendet. Bei thermisch induzierten Aggregationsassays mit Trypsin S200A als Testprotein wurde stets bei 43° C gearbeitet. Das konstante Absorptionssignal, welches aus der Aggregation des Substratproteins in Abwesenheit der potentiellen Chaperone resultierte, wurde als 100%-Wert der Aggregation festgesetzt und alle weiteren Messungen wurden darauf bezogen. Pufferbedingungen wurden an die zu testenden Proteine angepasst (Assays mit Pepsin: 50 mM MES, pH 6,5, 100 mM NaCl; mit Pepsin C-Domäne: 30 mM HEPES, pH 7,3, 100 mM KCl; mit Thermolysin 30 mM HEPES, pH 7,3, 100 mM KCl, 2 mM 1,10-Phenanthrolin; mit

Trypsin S200A und thermischer Denaturierung: 50 mM Tris, pH 8, 25 mM CaCl<sub>2</sub>; mit Trypsin S200A und chemischer Denaturierung: 30 mM HEPES, pH 7,3, 100 mM KCl, 25 mM CaCl<sub>2</sub>; mit HIV-1 Protease: 20 mM Natriumphosphat pH 5,8). Für jeden Testansatz wurde der Mittelwert aus mindestens drei Einzelmessungen gebildet. Um Aggregationsverminderung durch proteolytischen Abbau der denaturierten Substratproteine auszuschließen, wurde der Proteinanteil der Probelösungen nach abgelaufenem Chaperonassay mit 5% Trichloressigsäure gefällt. Das Präzipitat wurde durch Zentrifugation vom Überstand getrennt, in Laemmli-Puffer aufgenommen, neutralisiert und mit SDS-PAGE analysiert. Die Proteinmenge wurde über die Bandenintensität nach Comassie Brilliant Blue Färbung abgeschätzt und mit Kontrollansätzen verglichen. BSA diente in allen Assays als Kontrollprotein, um das Ausmaß unspezifischer Aggregationsverhinderung zu bestimmen. Beim Einsatz kompetitiver Inhibitoren, wurde dem Testansatz fünf Minuten vor Start der Messung Inhibitorstammlösung zugegeben. Abweichungen davon werden im Text explizit angegeben. Die Endkonzentrationen betragen 35 µM für Pepstatin (Stamm: 2 mg/ml in Ethanol) und 62 µM für Ac-Leu-Val-Phe-Aldehyd HIV-1 Protease Inhibitor (Stamm: 2 mg/ml in Ethanol).

**Komplexisolierung mit Gelfiltration** In 500 µl 50 mM MES pH 6,5, 100 mM NaCl mit und ohne Pepsin (33 µM) wurden 10 µl einer Ammoniumsulfatsuspension von Citratsynthase (Endkonzentration 3,6 µM) gegeben und 60 Minuten entweder bei Umgebungstemperatur oder bei 50° C inkubiert. Auftretender Niederschlag wurde durch Zentrifugation entfernt und der Überstand wurde auf eine Gelfiltrationssäule (Superose 6 10/300, GE Healthcare) in 50 mM MES, pH 6,5, 100 mM NaCl aufgetragen. Das Absorptionsprofil bei 280 nm wurde aufgezeichnet und Aliquots aus gesammelten Fraktionen wurden mit SDS-PAGE mit anschließender Silberfärbung analysiert.

**Herstellung von ClpA und ClpAΔN** Das ClpA-Gen wurde mit PCR aus genomischer DNA von *E. coli* amplifiziert. Über terminale Primer wurden Restriktionsschnittstellen für NdeI und XhoI eingeführt, über welche das PCR-Produkt in das Expressionsplasmid pET30b kloniert wurde. Das fertige Plasmid wurde als Templat für die Sequenz der AAA+-Domäne von ClpA verwendet. Die Gensequenz für die AAA+-Domäne kodierte für den Bereich G154-H758 und wird im folgenden ClpAΔN genannt. Sie wurde wie das ClpA-Gen in den Expressionsvektor pET30b kloniert. Die Expression und die Reinigung erfolgten für beide Proteine gleich. *E. coli* BL21 (DE3) Gold Zellen wurden mit dem Plasmid transformiert und wuchsen bei 37° C bis zu einer OD ~ 0,6. Dann wurden sie mit 1 mM IPTG induziert und nach vier Stunden geerntet. Nach Resuspendierung in 30 mM Tris, pH 7,5, 50 mM NaCl wurden die Zellen lysiert. Lösliches ClpA bzw. ClpAΔN im Überstand wurde mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 30 mM Tris, pH 7,5 mit einem linear ansteigendem NaCl-Gradienten über 10 Säulenvolumen von 50-525 mM NaCl

gereinigt. Fraktionen mit dem Zielprotein wurden vereinigt und auf Eis mit 40% Ammoniumsulfat (Endkonzentration) versetzt. Der Niederschlag wurde durch Zentrifugation vom Überstand getrennt, in 50 mM HEPES, pH 7,5, 300 mM KCl, 0,5 mM DTT und 10% Glycerol gelöst. Die Proteinmenge wurde über Absorption bei 280 nm bestimmt.

**Herstellung von HPR-ClpA und Saccharopepsin-ClpA** HPR-ClpA ist ein Fusionsprotein aus der HIV1-Protease Domäne und der AAA+-Domänen von ClpA. Für jede Domäne wurde ein PCR-Produkt hergestellt, welche passende Überhänge für die Überlappung mit dem anderen PCR-Produkt hatten. Als Templat dienten bereits in Vektoren ligierte Gensegmente. Mit Overlap-Extension PCR wurden die beiden einzelnen PCR-Produkte zusammengefügt. Über terminale NdeI und XhoI Schnittstellen, die über die terminalen Primer angefügt wurden, konnte das Gen in den pET30b Expressionsvektor ligiert werden. Das resultierende Produkt kodierte für P1-F99 von HIV-1 Protease in Fusion mit R143-H758 von ClpA und wird als HPR-ClpA bezeichnet. Die Klonierung von Saccharopepsin verlief analog. Als Templat für die C-terminale Saccharopepsindomäne diente die Gensequenz von Saccharopepsin in pTrxFus-Plasmid, welches von Rickey Yada (Department of Food Science, University of Guelph, Guelph, Kanada) zur Verfügung gestellt wurde. Analog zu HPR-ClpA wurde es über NdeI und XhoI Restriktionsschnittstellen in den pET30b Expressionsvektor ligiert. Das resultierende Produkt kodierte für A268-I406 von Saccharopepsin in Fusion mit R143-H758 von ClpA und wird als Saccharopepsin-ClpA bezeichnet. Die Expression und die Reinigung erfolgten für beide Proteine gleich. *E. coli* BL21 (DE3) Gold Zellen wurden mit dem Plasmid transformiert und wuchsen bei 37° C bis zu einer OD ~ 0,6. Dann wurden sie mit 1 mM IPTG induziert und nach vier Stunden geerntet. Beide Proteine wurden als *inclusion bodies* exprimiert. Nach Resuspendierung in 50 mM Tris, pH 8, 100 mM NaCl wurden die Zellen lysiert. *Inclusion bodies* wurden durch mehrmaliges Resuspendieren in 50 mM Tris, pH 8, 0,5 M Harnstoff, 5 mM EDTA, 1 mM DTT und 0,5% Triton, anschließender Zentrifugation und Abtrennung vom Überstand gewaschen. Die *inclusion bodies* wurden in 8 M Harnstoff, 50 mM Tris, pH 7,5, 10 mM DTT und 1 mM EDTA gelöst und mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 50 mM Tris, pH 7,5, 6 M Harnstoff, 1 mM DTT, 1mM EDTA mit einem linear ansteigendem KCl-Gradienten über 10 Säulenvolumen von 50-525 mM KCl gereinigt. Fraktionen mit dem Zielprotein wurden vereinigt zur Rückfaltung gegen 50 mM HEPES, pH 7,5, 300 mM KCl, 10% Glycerol, 100 µM ATP, 1 mM EDTA und 500 mM Arginin dialysiert. Anschließend wurde gegen 50 mM HEPES, pH 7,5, 300 mM KCl und 10% Glycerol dialysiert und die Proteinmenge über Absorption bei 280 nm bestimmt.

**Herstellung von GFP** Es wurden drei GFP Varianten hergestellt. Für natives GFP wurde der GFP enthaltende Clonetechnikvektor verwendet. Dieser enthielt einen der GFP-Sequenz vorangehenden, mit IPTG induzierbaren lac-Promoter. Von diesem Vektor wurde das GFP-



Gen mit PCR amplifiziert und über Primer mit terminalen HindIII und BamHI Schnittstellen versehen. Über diese wurde das PCR-Produkt in einen pET30-Vektor ligiert, der im gleichen Leseraster eine Schnittstellenerkennungssequenz der HIV-1 Protease (FNFPQITL) vor einem abschließenden Stopcodon kodierte. Das resultierende Konstrukt wird als GFP-HPRBS bezeichnet. Die dritte GFP-Variante wurde von Peter Zwickl (Max-Planck-Institut für Biochemie, München) zur Verfügung gestellt und war im Vektor pTrc99-His6-GFP-ssrA enthalten und wird hier als GFP-ssrA bezeichnet. Dieses Plasmid enthält einen tac-Promoter vor der GFP-Gensequenz und abschließend eine ssrA-Tag AANDENYALA-kodierende Sequenz. Für die Expression wurden *E. coli* BL21 (DE3) Gold Zellen mit Plasmiden mit GFP und GFP-HPRBS transformiert. GFP-ssrA tragendes Plasmid wurde in *E. coli* KY2266-Zellen transformiert, die ClpA-defizient sind, um Abbau von GFP-ssrA nach der Expression zu vermeiden. Die Expression und die Reinigung erfolgten stets auf die gleiche Weise. Zellen wuchsen bei 37° C bis zu einer OD ~ 0,6, wurden dann mit 1 mM IPTG induziert und wurden 16 weitere Stunden bei 20° C geschüttelt. Nach der Zellernte wurden sie in 20 mM Tris, pH 8, 150 mM NaCl suspendiert und lysiert. Die weitere Reinigung folgte einer Vorschrift von Alexander Yakhnin und Kollegen (Yakhnin et al., 1998). Dabei wurde dem Überstand des Lysats 100 mM Tris, pH 8 und 40% Ammoniumsulfat (Endkonzentrationen) zugefügt. Nach einer Stunde Inkubation auf Eis wurde zentrifugiert, das Präzipitat verworfen und mit dem Überstand weitergearbeitet. Dem wurde Ammoniumsulfat bis 80% und 25% Ethanol (Endkonzentrationen) zugeführt. Nach heftigem Schütteln wurde 5 Minuten zentrifugiert (4000 g), so dass sich eine obere, ethanolische, GFP-enhaltende Phase gebildet hatte. Die ethanolische Phase wurde mit 25% n-Butanol (Endkonzentration) versetzt, geschüttelt und zentrifugiert. Die nun untere wässrige Phase wurde mit 50% Chloroform (Endkonzentration) versetzt, geschüttelt und zentrifugiert. Die nun obere wässrige Phase enthielt GFP. Zur Erhöhung der Ausbeute wurde die Chloroformphase und die letzte Interphase mit 30 mM MOPS, pH 7, 50 mM NaCl ausgeschüttelt, zentrifugiert, von der Chloroform-Phase abgetrennt und mit der wässrigen Phase aus dem vorigen Schritt vereint. Zum Schluss wurde gegen 20 mM Tris, pH 8, 150 mM NaCl dialysiert und die Konzentration wurde über einen Bicinchoninsäureassay bestimmt.

**Herstellung von Trap-GroEL** Ein pET30 Expressionsplasmid mit Trap-GroEL in *E. coli* C41 (DE3) Zellen wurde von Astrid Ursinus (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) zur Verfügung gestellt. Trap-GroEL ist die D87K Mutante von *E. coli* GroEL, die Substrate zwar noch binden aber mangels ATPase-Aktivität nicht mehr freisetzen kann. Zur Expression wuchsen die Zellen bei 37° C bis zu einer OD ~ 0,6. Dann wurden sie mit 1 mM IPTG induziert und nach vier Stunden geerntet. Nach Resuspendierung in 30 mM MOPS, pH 7,0, 50 mM NaCl wurden die Zellen lysiert. Lösliches Trap-GroEL im Überstand wurde mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 30 mM MOPS, pH 7,0 mit einem linear ansteigenden NaCl-Gradienten über 15 Säulenvolumen von 50-750 mM

NaCl gereinigt. Fraktionen mit Trap-GroEL wurden vereinigt und auf Eis mit 50% und anschließend 65% Ammoniumsulfat (Endkonzentration) versetzt. Der Niederschlag wurde jeweils durch Zentrifugation vom Überstand getrennt. Das Präzipitat nach der zweiten Fällung wurde in 30 mM HEPES, pH 7,3, 100 mM KCl und 5 mM MgCl<sub>2</sub> gelöst. In diesem Puffer wurde Gelfiltration mit der gelösten Probe durchgeführt (Sephacryl S300, 26/60, GE Healthcare). Die Proteinmenge wurde über einen Bicinchoninsäureassay bestimmt.

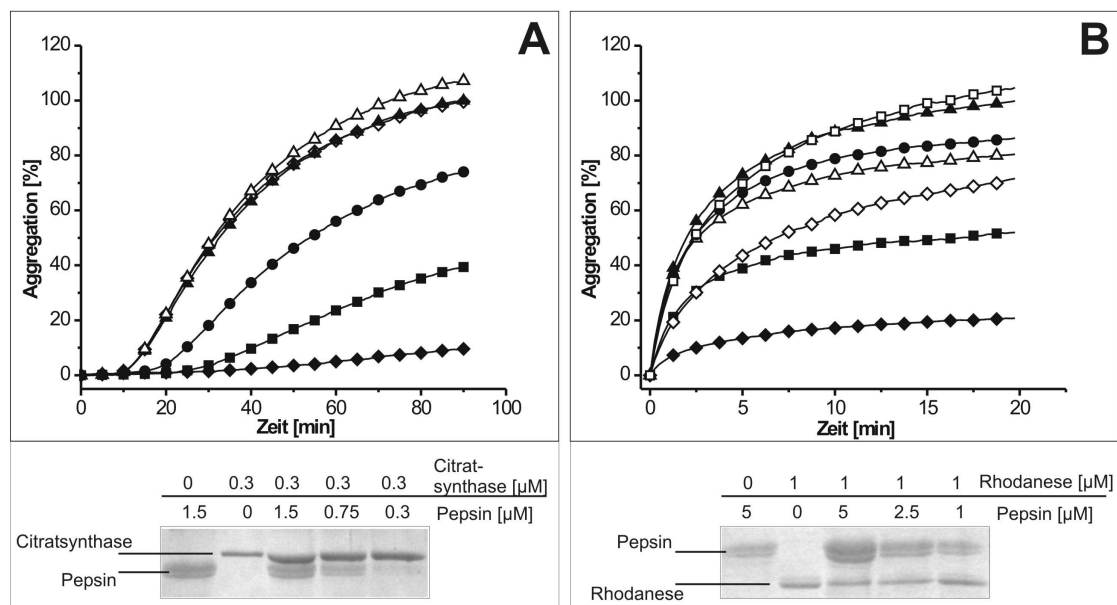
**GFP-Entfaltungssassays** Über den Verlust von Fluoreszenz wurde die GFP-Entfaltung durch ClpA-ATPasen ermittelt. Dazu wurde GFP mit ClpA gemischt und mit Puffer auf ein Gesamtvolumen von 200 µl gebracht. Die Reaktion wurde durch Zugabe von MgCl<sub>2</sub> und ATP gestartet und sofort in einem Cytofluor II Multi-well plate Reader von Perseptive Biosystems gemessen (Filter für die Fluoreszenzanregung: 485/20 nm; Filter für die Fluoreszenzemission: 530/30 nm). Phosphocreatin und Phosphocreatin Kinase wurden zugegeben, um das ATPase-Produkt ADP zu ATP zu regenerieren. Dies sollte Produkthemmung der ATPase vermeiden. Trap-GroEL wurde zugegeben, um entfaltetes GFP aus dem Entfaltungs-/Faltungs-Gleichgewicht durch irreversible Bindung zu entfernen. Die Zusammensetzung des Testansatzes war folgende (alles Endkonzentrationen): 0,5 µM ClpA-Protein (bezogen auf Hexamere von ClpA, ClpAΔN, HPR-ClpA oder Saccharopepsin-ClpA), 1 µM Trap-GroEL (als 14-mer), 10 µM Phosphocreatin, 0,0005 U Phosphocreatin Kinase, 0,5 µM GFP, 2 mM ATP, 20 mM MgCl<sub>2</sub> und als Puffer 50 mM HEPES, pH 7,5, 300 mM KCl, 0,5 mM DTT und 10% Glycerol.

**ATPase-Assays** Die Assays beruhen auf der Freisetzung von Phosphat aus ATP durch ATPasen. Das Phosphat wurde in einer anschließenden Farbreaktion mit Malachitgrün nachgewiesen. Der Puffer für die ATPase-Reaktion war 50 mM HEPES, pH 7,5, 300 mM KCl, 0,5 mM DTT und 10% Glycerol. Die Konzentration von ClpA-Proteinen war 0,5 µM ClpA-Protein (bezogen auf Hexamer von ClpA, ClpAΔN, HPR-ClpA oder Saccharopepsin-ClpA). Die Reaktion wurde durch Zugabe von 2 mM ATP und 20 mM MgCl<sub>2</sub> gestartet. Nach definierten Zeitpunkten wurden 12 µl Testansatz durch Zugabe von 8 µl 60 mM EDTA gestoppt und aufbewahrt. Zur Phosphatbestimmung wurde zu jeder Probe 700 µl Malachitgrün-Lösung (0,034 g/100ml Malachitgrün, 1,05 g/100ml Ammoniummolybdat in 1 M HCl und 0,1% Triton X-100) und nach 90 Sekunden 100 µl 34% Citratlösung zugegeben. Nach 20 Minuten Inkubation bei Raumtemperatur wurden die Proben photometrisch bei einer Wellenlänge von 640 nm in einem SLT Labinstruments Easy Reader EAR 400 AT gemessen.

### 4.3 Ergebnisse

**Aggregationsverhinderung durch Pepsin** Verhinderung der Aggregation von denaturierten Proteinen ist ein Charakteristikum von Chaperonen. Aggregation kann über das

Absorptionssignal, das durch lichtstreuende Aggregate im Strahlengang eines Photometers verursacht wird, gemessen werden. In Anwesenheit von Chaperonen ist das Absorptionssignal kleiner als in deren Abwesenheit, da Aggregatbildung durch Chaperone verlangsamt oder verhindert wird. Auf diese Weise wurden proteolytisch inaktivierte Proteasen auf ihre Chaperonaktivität untersucht. Bei pH 6,5 ist Pepsin proteolytisch vollständig inaktiv, weil beide katalytischen Aspartatreste deprotoniert vorliegen (Lin et al., 1992). Die Durchführung der Assays bei pH 6,5 ermöglichte, proteolytisch inaktives Pepsin mit minimaler Veränderung des aktiven Zentrums zu verwenden.

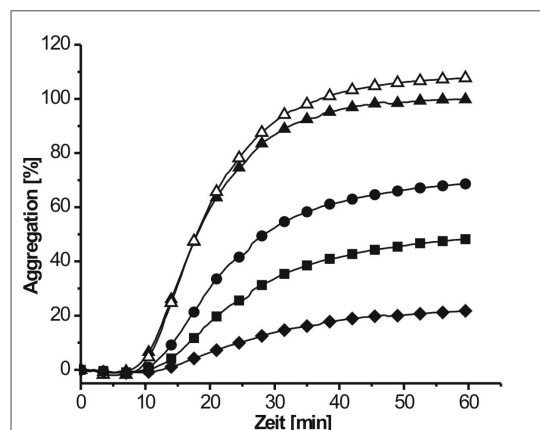


**Abbildung 4 Chaperonaktivität von Pepsin.** Aggregationsassays mit thermisch denaturierter Citratsynthase und chemisch denaturierter Rhodanese. (A, oben) Citratsynthase ( $\blacktriangle$ ) ohne Pepsin, ( $\bullet$ ) mit gleicher Menge Pepsin, ( $\blacksquare$ ) 2,5-facher Pepsinüberschuss, ( $\blacklozenge$ ) 5-facher Pepsinüberschuss, ( $\diamond$ ) 5-facher Pepsinüberschuss bei Anwesenheit von Pepstatin, ( $\triangle$ ) 5-facher BSA-Überschuss. (Unten) SDS-PAGE Analyse zeigt die Proteinzusammensetzung der Testlösungen zum Testende, um eventuelle proteolytische Aktivität auszuschließen. (B, oben) Rhodanese ( $\blacktriangle$ ) ohne Pepsin, ( $\bullet$ ) mit gleicher Menge Pepsin, ( $\blacksquare$ ) 2,5-facher Pepsinüberschuss, ( $\blacklozenge$ ) 5-facher Pepsinüberschuss, ( $\square$ ) 2,5-facher Pepsinüberschuss bei Anwesenheit von Pepstatin, ( $\diamond$ ) 5-facher Pepsinüberschuss bei Anwesenheit von Pepstatin, ( $\triangle$ ) 5-facher BSA-Überschuss. (Unten) SDS-PAGE Analyse zeigt die Proteinzusammensetzung der Testlösungen zum Testende, um eventuelle proteolytische Aktivität auszuschließen.

Anwesenheit von 1-, 2,5- und 5-fachen molaren Überschüssen von Pepsin reduzierte die thermisch induzierte Aggregation der Citratsynthase auf 75%, 40% und 10% (Abbildung 4). Die Kurven zeigen die typische Anfangsphase mit verzögerter Aggregation, solange die Proben noch erwärmt werden und Citratsynthase noch nicht denaturiert wurde. Nach 10 Minuten setzte die Aggregation der Citratsynthase ein. Bei Anwesenheit von Pepsin setzte die Aggregation noch später ein. Die Aggregationsvermeidung nahm mit zunehmender Pepsinkonzentration zu. Die Kontrollprobe mit 5-fachem molarem BSA-Überschuss zeigte

keine unspezifische Aggregationsvermeidung. Bei chemisch denaturierter Rhodanese reduzierte die Anwesenheit von 1-, 2,5- und 5-fachen molaren Überschüssen von Pepsin die Aggregation auf 85%, 50% und 20%. Hier setzte die Aggregation sofort bei Messbeginn ein, da die Substrate bereits denaturiert waren. Die Anwesenheit von Pepsin reduzierte die Geschwindigkeit als auch das Ausmaß der Aggregation. Ein 5-facher molarer Überschuss von BSA reduzierte Aggregation auf 80%. Dies offenbarte einen messbaren Anteil unspezifischer Aggregationsvermeidung. Dieser unspezifische Anteil war aber im Vergleich zur Restaggregation mit 5-fach molarem Pepsinüberschuss (20%) relativ gering.

Um proteolytische Restaktivität auszuschließen, wurde die Gesamtproteinmenge im Anschluss einer Messung gefällt und mit SDS-PAGE analysiert. Proteinmengen wurden über Bandenintensitäten abgeschätzt und mit Kontrollproben verglichen. Die Menge Citratsynthase war in Proben und Kontrollen gleich und es konnten keine Abbauspuren gefunden werden. Demnach konnte proteolytische Restaktivität des Pepsins während der Assays ausgeschlossen werden.



**Abbildung 5 Chaperonaktivität der Pepsin C-Domäne.** Aggregationsassays mit thermisch denaturierter Citratsynthase: Citratsynthase (▲) ohne Pepsin C-Domäne, (●) mit gleicher Menge Pepsin C-Domäne, (■) 2,5-facher Überschuss an Pepsin C-Domäne, (◆) 5-facher Überschuss an Pepsin C-Domäne, (△) 5-facher BSA-Überschuss.

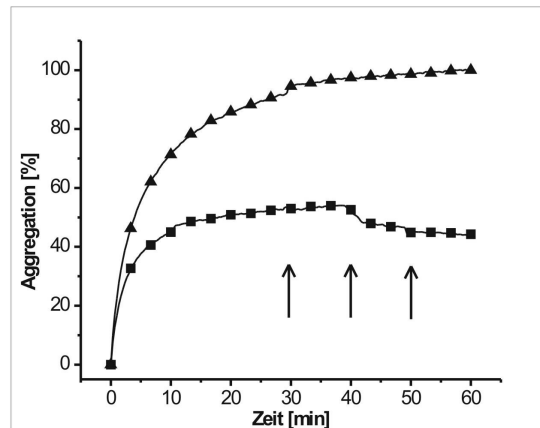
Pepsin besteht aus zwei Domänen mit unterschiedlicher Stabilität. Die N-Domäne ist generell weniger stabil als die C-Domäne und lag bei Assaybedingungen von 50° C wahrscheinlich entfaltet vor (Privalov et al., 1981). Deswegen könnte die Chaperonaktivität allein von der C-Domäne verursacht worden sein. Um dies zu überprüfen, wurden Chaperonassays mit der C-Domäne durchgeführt, die aus Pepsin durch limitierten proteolytischen Verdau durch Proteinase K hergestellt wurde. Bei pH 7,3 ist die N-Domäne entfaltet und daher Proteinase K sensitiv, wohingegen die C-Domäne gefaltet und Proteinase K resistent ist (Lin et al., 1993). Die C-Domäne besitzt als isolierte Domäne keine proteolytische Aktivität mehr. Die Chaperonaktivität wurde mit thermisch denaturierter

Citratsynthase getestet und es zeigte sich ein titrierbarer Rückgang der Aggregation auf 70%, 50% und 20% bei 1-, 2,5- und 5-fachen molaren Überschüssen von Pepsin C-Domäne (Abbildung 5). Das zeigt, dass die C-Domäne von Pepsin für Chaperonaktivität ausreichend war.

Das proteolytisch aktive Zentrum wird bei Pepsin von Resten gebildet, die zwischen den *psi-loops* liegen. Um zu überprüfen, ob das aktive Zentrum an der Bindung denaturierter Substratproteine und der Chaperonaktivität beteiligt ist, wurden Chaperonassays durchgeführt, in denen Pepsin mit dem kompetitiven Peptidinhibitor Pepstatin vorinkubiert war. Als Substratproteine wurden thermisch denaturierte Citratsynthase und chemisch denaturierte Rhodanese eingesetzt. Chaperonaktivität konnte durch Pepstatin weitestgehend verhindert werden. Dies wurde daran deutlich, dass die Aggregationskurven der Citratsynthase in Abwesenheit von Pepsin und mit 5-fachem Pepsinüberschuss mit Pepstatin fast identisch waren (Abbildung 4). Nur im Falle von Rhodanese mit 5-fachem Pepsinüberschuss konnte noch eine reduzierte Chaperonaktivität gemessen werden; in Anwesenheit von Pepstatin stieg die Aggregation von 20% auf 70% an.

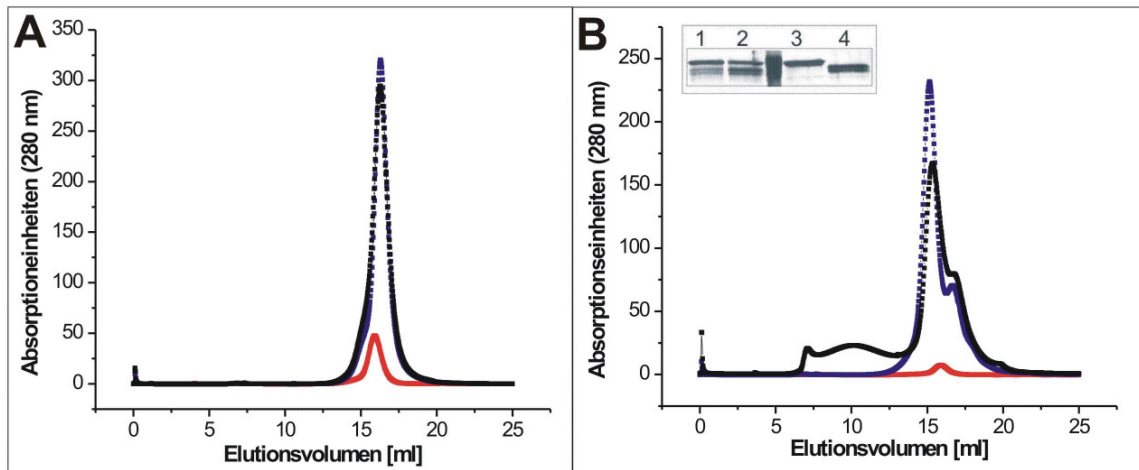
Auf den ersten Blick mag verwundern, dass Pepstatin die Chaperonaktivität in Assays mit thermisch induzierter Aggregation bei 50° C inhibiert. Es wurde nämlich angenommen, dass die N-Domäne bei solchen Temperaturen entfaltet vorliegt. Pepstatin sollte nicht an die isolierten C-Domäne binden. Jedoch stabilisiert Pepstatin beide Domänen und erhöht deren Schmelztemperatur. Dies ermöglichte die kompetitive Hemmung der Substratbindung im aktiven Zentrum. Mit der kompetitiven Hemmung der Chaperonaktivität konnte die Beteiligung des aktiven Zentrums zwischen den *psi-loops* gezeigt werden.

Um die Stärke der Wechselwirkung zwischen Pepsin und denaturierter Rhodanese abzuschätzen, wurde versucht, bereits gebildete Komplexe nach Erreichen der Plateauphase des Absorptionssignals durch Zugabe von Pepstatin zu dissoziieren. Falls Pepstatin Rhodanese aus geformten Komplexen mit Pepsin freisetzen könnte, so würde Aggregation freigesetzter Rhodanese nach Pepstatinzugabe und damit ein Anstieg des Absorptionssignals erwartet werden. Dies war jedoch nicht der Fall (Abbildung 6). Stattdessen blieb das Niveau des Aggregationssignals von der Pepstatinzugabe unbeeinflusst. Es ist davon auszugehen, dass keine Rhodanese aus den Komplexen mit Pepsin freigesetzt wurde, welches danach nämlich Aggregation verursacht hätte.



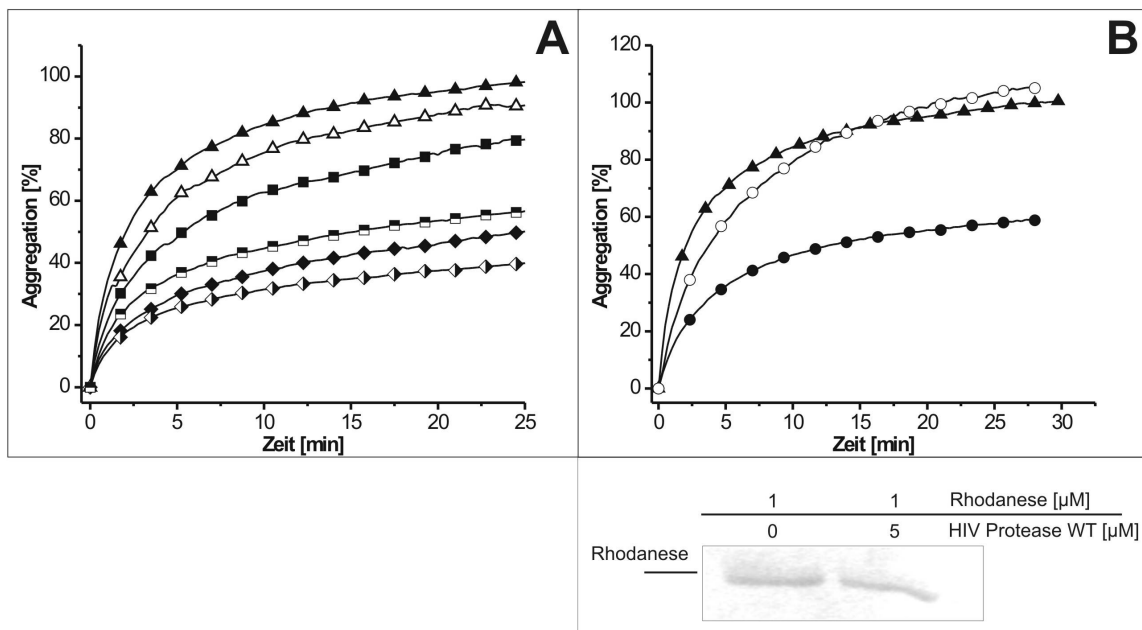
**Abbildung 6 Chaperonaktivität von Pepsin bei nachträglicher Zugabe von Pepstatin.** Aggregation von Rhodanese (▲) ohne Pepsin, (■) mit 2,5-fachen Pepsinüberschuss. Die Pfeile zeigen die Zeitpunkte der Pepstatinzugabe während der Messung an.

**Isolierung von Komplexen aus Citratsynthase und Pepsin** Um Informationen über die Stöchiometrie der Interaktion zwischen Pepsin und denaturierten Substratproteinen zu erhalten, wurde versucht, Komplexe aus Pepsin und Citratsynthase zu isolieren. Komplexbildung wurde durch 90-minütige Inkubation der Proteinmischung bei 50° C ausgelöst. Pepsin und Citratsynthase alleine, sowie die gleichen Proben bei 25° C, dienten als Kontrollen. Ohne Hitzeinkubation eluierten Pepsin und Citratsynthase als Mischung separater Proteine von einer Gelfiltrationssäule, was aus dem Vergleich mit dem Elutionsverhalten isolierter Proteine abgeleitet wurde (Abbildung 7). Dagegen resultierte die Inkubation bei 50° C in einem breiten Elutionspeak bei höheren Molekülmassen und einem schmalen Peak unmittelbar nach dem Ausschlussvolumen (8 ml). Weder Pepsin noch Citratsynthase alleine ließen nach Hitzeinkubation vergleichbare Peaks erkennen (Abbildung 7). Zusätzlich zeigte SDS-PAGE Analyse, dass der neu aufgetretene, breite Peak aus beiden Proteinen zusammengesetzt war: Pepsin und Citratsynthase. Weitergehende Analyse nach der Gelfiltration mit statischer Lichtstreuung zeigte, dass die Komplexe in keinem definierten Verhältnis vorlagen. Stattdessen wiesen sie eine sehr heterogene Größenzusammensetzung auf. Die heterogene Zusammensetzung ist konsistent mit der breiten Peakform. Bei Pepsin kann durch Hitzeinkubation eine geringe Verschiebung zu kleineren Elutionsvolumen beobachtet werden. Das unterstützt die bereits bestehende Vermutung, dass die N-Domäne bei 50° C denaturiert vorlag. Diese Resultate deuten darauf hin, dass Pepsin ausschließlich mit denaturierter, nicht mit nativer, Citratsynthase interagiert und die Bildung stabiler Komplexe die Löslichkeit denaturierter Citratsynthase erhält.



**Abbildung 7 Komplexbildung und Komplexisolierung von Pepsin mit denaturierter Citratsynthase.** Elutionsprofile von Gelfiltrationsläufen. (A) Inkubation bei 25° C: (rot) Citratsynthase, (blau) Pepsin, (schwarz) Mischung aus Citratsynthase und Pepsin. (B) Inkubation bei 50° C: (rot) Citratsynthase, (blau) Pepsin, (schwarz) Mischung aus Citratsynthase und Pepsin. Eingefügter Kasten: SDS-PAGE Analyse von Fraktionen des Gelfiltrationslauf der Mischung aus Pepsin und Citratsynthase (schwarz). Aufgetragene Fraktionen: bei 10 ml Elutionsvolumen (1), bei 15 ml Elutionsvolumen (2), zum Vergleich: Citratsynthase (3) und Pepsin (4).

**Chaperonaktivität der HIV-1 Protease** Für Assays mit HIV-1 Protease wurden die zu testenden Proteine in *E. coli* exprimiert, gereinigt und rückgefaltet. HIV-1 Protease wurde durch die Mutation D25N proteolytisch inaktiviert, da die katalytisch aktiven Aspartatreste des aktiven Zentrums entfernt wurden. Zusätzlich führten die Mutationen Q2C, C95A und L97C in der Mutante HPR-DS an bestimmten Stellen in den Dimerisierungssträngen Cysteinreste ein. John Louis und Kollegen fanden heraus, dass diese Cysteinreste Disulfidbrücken bilden können, und damit die Dimerisierungsstränge verdrehen und Dimerisierung verhindern (Louis et al., 2003). Um die Bildung der Disulfidbrücken zu überprüfen, wurde mit Ellmanns Reagenz die Menge an freien Thiolgruppen bestimmt und ins Verhältnis zur eingesetzten Proteinmenge gesetzt. Für HPR-DS ergab sich eine von drei möglichen freien Thiolgruppen und für HPR-M zwei von zwei möglichen. Demnach kann die Bildung einer Disulfidbrücke in HPR-DS angenommen werden. Beide HIV-1 Protease Mutanten wurden in Chaperonassays mit chemisch denaturierter Rhodanese getestet und zeigten Chaperonaktivität (Abbildung 8).



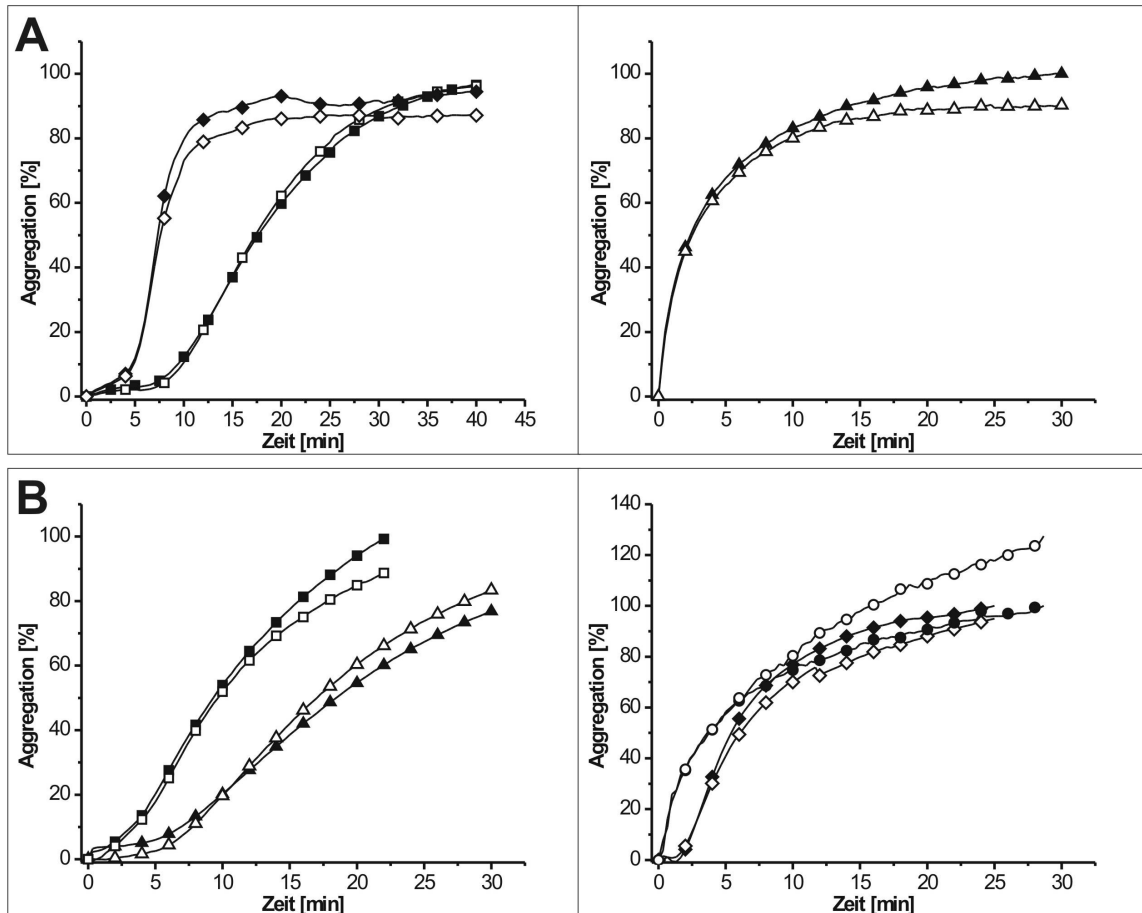
**Abbildung 8 Chaperonaktivität der HIV-1 Protease.** Aggregationsassays mit chemisch denaturierter Rhodanese: (A) Rhodanese (▲) ohne HIV-1 Protease, (■) 2,5-facher HPR-M Überschuss, (◆) 5-facher HPR-M Überschuss, (▣) 2,5-facher HPR-DS Überschuss, (◊) 5-facher HPR-DS Überschuss, (△) 5-facher BSA-Überschuss. (B, oben) Rhodanese (▲) ohne HPR-WT, (●) 5-facher HPR-WT Überschuss, (○) 5-facher HPR-WT Überschuss in Anwesenheit des HIV-1 Proteaseinhibitors. (Unten) SDS-PAGE Analyse der Testlösung nach Testende, um proteolytischen Abbau der Rhodanese durch HPR-WT auszuschließen.

Dimere von HPR-M reduzierten die Aggregation auf 80% und 50% bei 2,5-fachem und 5-fachem molarem Überschuss (Überschuss bezogen auf monomeres HPR-M). Monomere von HPR-DS reduzierten die Aggregation auf 60% und 40% bei 2,5-fachem und 5-fachem molarem Überschuss. Erkennbar war sowohl die Verlangsamung der Aggregation, als auch ihr vermindertes Ausmaß. Die Stärke der Aggregationsvermeidung nahm mit zunehmender Konzentration der eingesetzten HIV-1 Protease ebenfalls zu. Analog zu den Experimenten mit Pepsin wurde durch Einsatz eines kompetitiven Inhibitors die Beteiligung des aktiven Zentrums an der Chaperonaktivität überprüft. Es wurde der HIV-1 Proteaseinhibitor Ac-Leu-Val-Phe-Aldehyd eingesetzt. Jedoch konnte kein Effekt auf die Chaperonaktivität beobachtet werden. Eine mögliche Erklärung könnte sein, dass die Mutation D25N keine feste Bindung mehr ermöglichte, da die Beteiligung der negativen Ladung des Aspartatrestes fehlte. Deswegen wurde die wildtyp HIV-1 Protease HPR-WT in Chaperonassays getestet. Es zeigte sich, dass HPR-WT denaturierte Rhodanese proteolytisch nicht umsetzte. Vermutlich weil in Rhodanese keine HIV-1 Protease spezifische Schnitterkennungssequenzen vorkommen (Dunn et al., 1994). Aggregation von Rhodanese wurde bei 5-fachem molarem HPR-WT Überschuss auf 60% reduziert, was dem Wert bei Einsatz von 2,5-fachem Überschuss HPR-M entsprach. Anders als bei den D25N-Mutanten wurde bei der wildtyp Protease die Aggregationsvermeidung in Anwesenheit des Inhibitors vollständig verhindert (Abbildung 8).



Zusammengenommen wurde gezeigt, dass HIV-1 Protease, wie Pepsin auch, sowohl als Monomer wie auch als Dimer Chaperonaktivität zeigt und Substratbindung im aktiven Zentrum stattfindet.

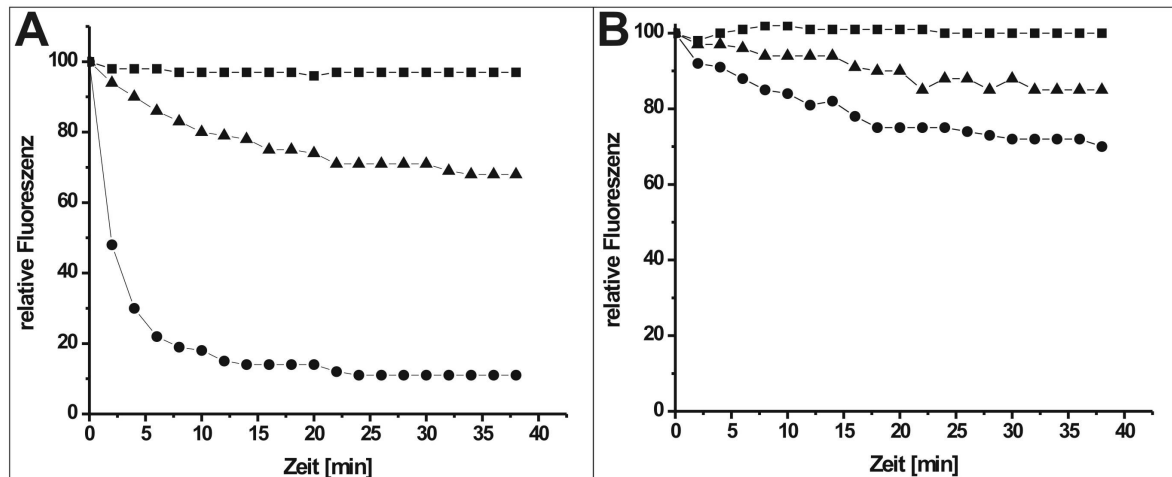
**Chaperonassays mit Thermolysin und Trypsin** Die Idee von der generellen Abstammung der Proteasen von Chaperonen wurde mit weiteren Proteasen aus der Proteinfamilie der Metallo- und Serinproteasen getestet. Falls Proteasen generell von Chaperonen abstammen, so sollte mit hoher Wahrscheinlichkeit bei den meisten Proteasen eine inhärente Chaperonaktivität messbar sein, sobald man die proteolytische Aktivität ausschaltet. Als Vertreter der Metalloproteasen wurde die Zinkprotease Thermolysin ausgewählt und durch den Zn-Chelator 1,10-Phenanthrolin inaktiviert. Durch Inkubation mit diesem Chelator blieb Calcium, das für die Stabilität des Thermolysins essentiell ist, am Protein gebunden. Thermolysin wurde mit den Substratproteinen Luciferase und Citratsynthase bei thermisch induzierter Aggregation getestet und mit Rhodanese bei chemisch induzierter Aggregation. Jeweils war das Absorptionssignal in An- und Abwesenheit von Thermolysin gleich (Abbildung 9 A). Chaperonaktivität konnte also in keinem der Fälle gemessen werden. Als Vertreter der Serinproteasen wurde Trypsin verwendet, das durch die Mutation S200A proteolytisch inaktiviert wurde. Dieses konnte in *E. coli* als Trypsinogen exprimiert, gereinigt und rückgefaltet werden. Da die Prosequenz das aktive Zentrum blockiert, musste Trypsinogen zu Trypsin prozessiert werden. Dies gelang durch Einsatz von Enterokinase, die hochspezifisch das Propeptid nach der Erkennungssequenz abspaltet. Trypsin wurde mit den Substraten Luciferase und Citratsynthase und thermisch induzierter Aggregation getestet. Hierbei musste die üblich Assaytemperatur von 50° C auf 43° C gesenkt werden, um die Aggregation von Trypsin selbst zu vermeiden. Rhodanese und Trypsin waren chemisch denaturierte Substrate zum Testen der Chaperonaktivität. Alle Signale waren jedoch in An- und Abwesenheit von Trypsin gleich. Es konnte keine Chaperonaktivität von Trypsin gemessen werden (Abbildung 9 B).



**Abbildung 9 Keine Chaperonaktivität von Thermolysin und Trypsin.** Aggregationsassays mit thermisch denaturierter Citratsynthase oder Luciferase und Assays mit chemisch denaturierter Rhodanese oder Luciferase. (A) Assays um Thermolysin zu testen: (Links) Luciferase<sub>thermisch denaturiert</sub> (Δ) ohne Thermolysin, (M) 10-facher Thermolysinüberschuss. Citratsynthase: (!) Ohne Thermolysin, (□) 10-facher Thermolysinüberschuss. (Rechts) Rhodanese (▲) ohne Thermolysin, (8) 10-facher Thermolysinüberschuss. (B) Assays um Trypsin zu testen: (Links) Luciferase<sub>thermisch denaturiert</sub> (!) ohne Trypsin, (□) 10-facher Trypsinüberschuss. Citratsynthase (▲) ohne Trypsin, (8) 5-facher Trypsinüberschuss. (Rechts) Rhodanese (.) ohne Trypsin, (-) 5-facher Trypsinüberschuss. Luciferase<sub>chemisch denaturiert</sub> (Δ) ohne Trypsin, (M) 5-facher Trypsinüberschuss.

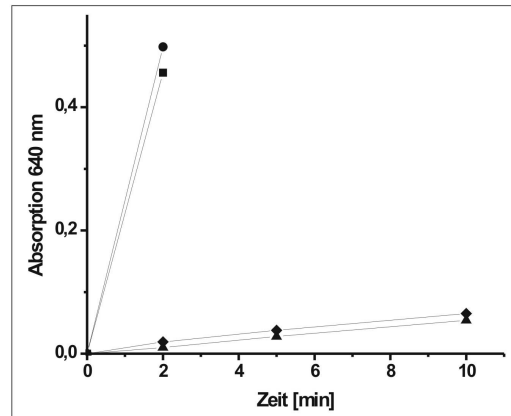
**GFP Entfaltungsassays mit ClpA Konstrukten** Die Proteinentfaltungsaktivität von Varianten des AAA+-Proteins ClpA wurde mit GFP getestet. Dazu wurde GFP mit ClpA Proteinen inkubiert und die Fluoreszenz des GFPs gemessen, die ein konzentrationsabhängiges Maß für die Faltung ist. Natives ClpA und ClpAΔN konnten in *E. coli* exprimiert und als lösliche Proteine gereinigt werden. Beide entfalteten GFP, sofern es mit einem Peptidanhang synthetisiert wurde (Abbildung 10). ClpAΔN hatte eine höhere Entfaltungsaktivität als natives ClpA. In beiden Fällen wurde GFP mit dem ssrA-tag schneller entfaltet als GFP mit der HIV-1 Protease Schnittstellenerkennungssequenz. Natives GFP wurde nicht entfaltet. HPR-ClpA und Saccharopepsin konnten nicht löslich exprimiert

werden und mussten nach Reinigung unter denaturierenden Bedingungen rückgefaltet werden. GFP-Entfaltung konnte mit diesen beiden Fusionsproteinen nicht beobachtet werden (ohne Abbildung). Während der Rückfaltung trat heftiger Niederschlag durch fehlgefaltetes Protein auf. Möglicherweise ist der lösliche Teil des rückgefalteten Proteins ebenfalls nicht nativ gefaltet.



**Abbildung 10 GFP Entfaltungsassays.** Die relative Fluoreszenz von GFP wurde gegen die Zeitdauer des Assays aufgetragen. (■) Natives GFP, (▲) GFP-HPRBS, (●) GFP-ssrA. (A) ClpAΔN, (B) ClpA.

**ATPase-Assays** Von allen ClpA-Varianten wurde die ATPase-Aktivität getestet. Es sollte überprüft werden, ob die höhere Entfaltungsaktivität von ClpAΔN gegenüber nativem ClpA mit einer höheren ATPase-Aktivität einhergeht und ob HPR-ClpA und Saccharopepsin-ClpA ATPase-Aktivität und damit nativ gefaltete AAA+-Domänen besitzen. ClpAΔN und ClpA zeigten nahezu gleiche ATPase-Aktivität, während die ATPase-Aktivität von HPR-ClpA und Saccharopepsin-ClpA wesentlich geringer war (Abbildung 11). Dies lässt darauf schließen, dass zwar lösliches Protein von HPR-ClpA und Saccharopepsin-ClpA erhalten wurde, welches aber nicht nativ gefaltet war und keine Aktivität zeigte.



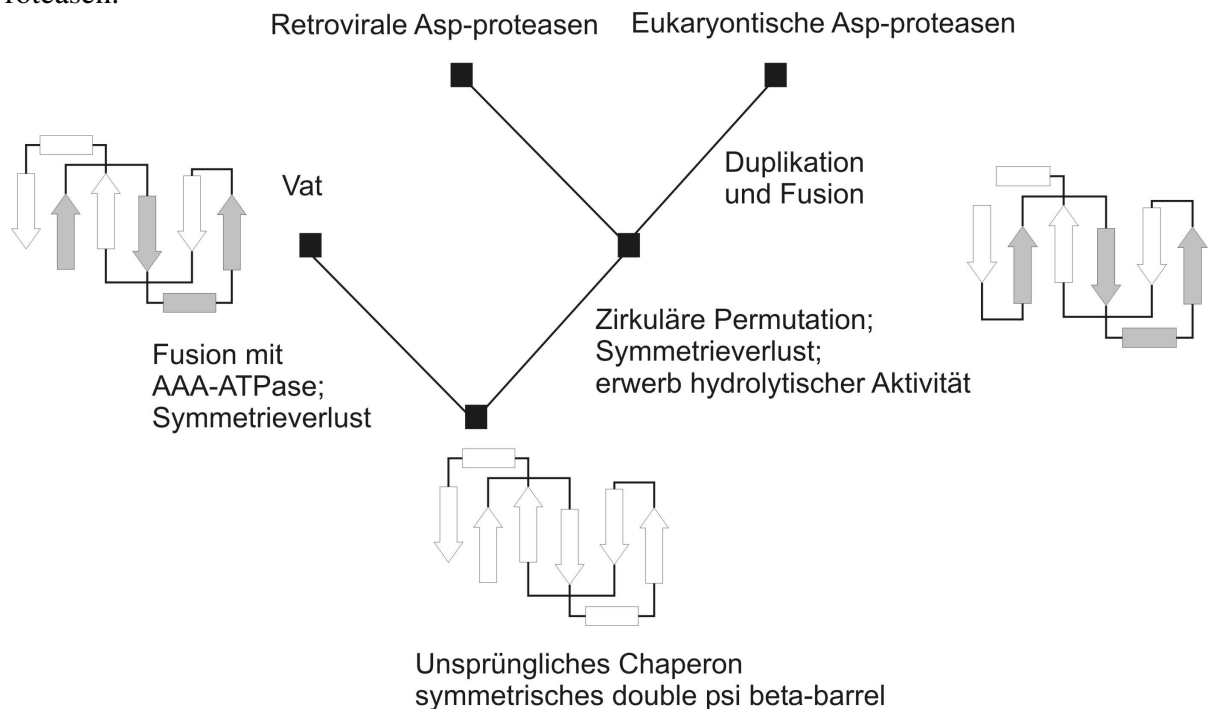
**Abbildung 11 ATPase-Assays von ClpA-Varianten.** Die ATPase-Aktivität wurde über die Freisetzung von Phosphat gemessen, welches einen Farbkomplex bildet, der Licht bei 640 nm absorbiert. Die Absorption ist gegen die Zeit aufgetragen (■) ClpA, (●) ClpA $\Delta$ N, (▲) HPR-ClpA, (◆) Saccharopepsin-ClpA.

#### 4.4 Diskussion

Eine inhärente Chaperonaktivität bei Aspartatproteasen wurde messbar, sobald ihre proteolytische Aktivität entfernt wurde. Pepsin, Pepsin C-Domäne und HIV-1 Protease Mutanten konnten die Aggregation von thermisch und chemisch denaturierten Substratproteinen konzentrationsabhängig verringern. Bei Trypsin und Thermolysin als Vertreter der Serin- und Metalloproteasen konnte keine Chaperonaktivität festgestellt werden. Daraus kann keine allgemeine inhärente Chaperonaktivität bei inaktivierten Proteasen abgeleitet werden. Die gemessene Chaperonaktivität der Aspartatproteasen wird in einem evolutionären Kontext interpretiert, wobei die gemessene Chaperonaktivität als vererbtes Überbleibsel eines ursprünglichen Vorfahren gedeutet wird.

Die Idee eines Vorfahren mit Chaperonaktivität bestärkt das Evolutionsszenario von Murray Coles und Kollegen (Coles et al., 1999), in dem von einer divergenten Evolution von Aspartatproteasen und Double psi beta-barrel Proteinen mit einem gemeinsamen symmetrischen Double psi beta-barrel Vorfahren mit Chaperonaktivität ausgegangen wird (Abbildung 12). Von diesem Vorfahren trennten sich die Cdc48 N-terminal domain-like Proteine ab, fusionierten mit AAA-ATPase Domänen und erhielten ihre Chaperonaktivität als Substratbindedomänen im Kontext der AAA-ATPasen. Parallel dazu, fand im Zweig der Aspartatproteasen eine zirkuläre Permutation und der Erwerb katalytisch aktiver Reste in der Substratbindetasche statt. Chaperonassays zeigten, dass die isolierte C-Domäne von Pepsin oder monomere HIV-1 Protease in Assays zur Aggregationsvermeidung ausreichend für die effektive Bindung entfalteter, aggregationsanfälliger Proteine waren. Dies deutet sehr stark darauf hin, dass der Chaperonvorgänger der Proteasen aus einer einzigen Domäne bestand. Im Gegensatz dazu verlangt proteolytische Aktivität die spezifische Wechselwirkung zweier

Domänen. Kompetitive Proteaseinhibitoren des aktiven Zentrums konnten die Chaperonaktivität verhindern. Dies deutet darauf hin, dass der Chaperonvorgänger der Proteasen die gleiche Bindungsstelle zwischen den *psi-loops* verwendete wie heutige Proteasen.



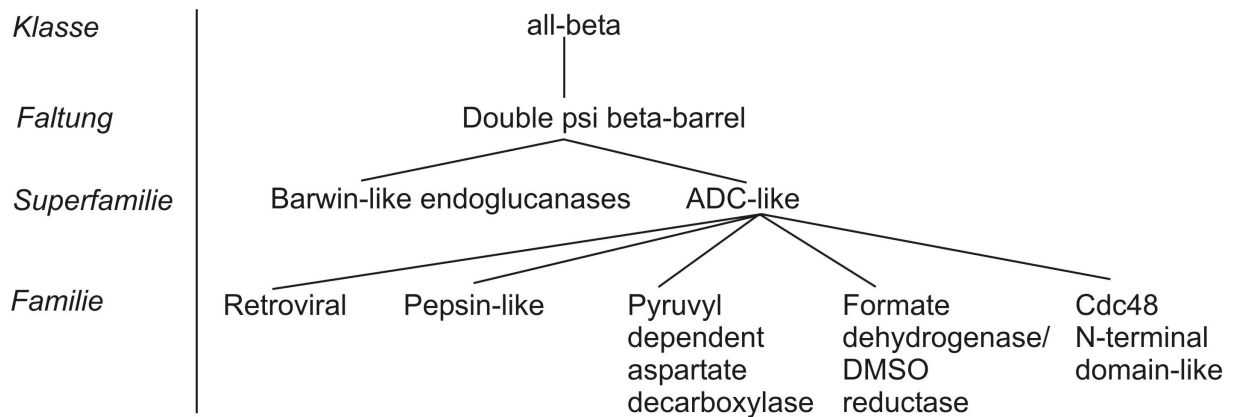
**Abbildung 12 Vorgeschlagene Evolutionshypothese der Cdc48 N-terminal domain-like Proteinen und Aspartatproteasen.** Das Szenario legt ein ursprüngliches Double psi beta-barrel Chaperon mit symmetrischen Hälften zu Grunde. Der linke Zweig führt über Symmetrieverlust und Fusion mit ATPase-Domänen zu AAA-ATPasen, z. B. zu Vat. Der rechte Zweig führt über eine zirkuläre Permutation, Symmetrieverlust und Erwerb hydrolytischer Aktivität zu einer potentiellen Vorgängerprotease für retrovirale und eukaryontische Proteasen, die daraus divergent entstanden sind. Die Entwicklung eukaryontischer Aspartatproteasen umfasst zusätzlich Genduplikation und Genfusion.

An der Bindungsstelle heutiger Double psi beta-barrel Domänen sind Aminosäurereste zwischen den *psi-loops* sowie andere Reste außerhalb der *psi-loops* beteiligt. Wenn Double psi beta-barrel Proteine und Aspartatproteasen einen gemeinsamen Vorfahren teilen, dann würde man für beide die gleiche Bindestelle für Substrate erwarten. Es konnte jedoch gezeigt werden, dass sich Bindestellen im Laufe der Evolution verschieben, besonders wenn die Domäne in einen anderen Domänenkontext eingebettet wird (Todd et al., 2002). Im Evolutionsszenario wird solch eine Spezialisierung des Domänenkontextes vorgeschlagen. Bei der Entfaltung proteolytischer Aktivität diente die ursprüngliche Chaperonaktivität nur noch der Substratbindung, die durch die proteolytische Spaltung des Substrats dann überdeckt wurde. Eine Verzweigung in virale und in eukaryontische Proteasen geschah von der ursprünglichen Aspartatprotease ausgehend. Virale Proteasen entwickelten unter dem Selektionsdruck der viralen Assemblierung eine hohe Substraterkennungsspezifität und

Dimerisierung. Eukaryontische Proteasen entwickelten hohe Effizienz bei geringer Spezifität als Teil des Verdauungssystems. Die Genfusion kann als Steigerung der Effizienz gedeutet werden, da das Zusammenfinden zweier Proteasedomänen zu einem aktiven Proteasekomplex eingespart werden kann. Das Fehlen einer prokaryontischen Aspartatprotease wurde als Argument gegen eine gemeinsame, ursprüngliche Vorgängerprotease von eukaryontischen und viralen Proteasen verwendet. Mit der gezeigten inhärenten Chaperonaktivität von Proteasen wird deutlich, dass der gemeinsame Vorfahre ein Chaperon gewesen sein könnte. Und Chaperone mit ähnlicher Faltung sind in heutigen Double psi beta-barrel Proteinen zu finden, die in allen Organismenreichen vorkommen.

Eine zirkuläre Permutation, welche topologisch gesehen die Faltungen Acid protease in Double psi beta-barrel umwandelt, ist ein häufiger genetischer Vorgang (Lindqvist and Schneider, 1997). Der Grund für diese zirkuläre Permutation konnte experimentell nicht untersucht werden. Dennoch soll hier eine mögliche Erklärung diskutiert werden. Im Falle der Aspartatproteasen könnte die Lage der Termini nach zirkulärer Permutation ein selektiver Vorteil gewesen sein. Im  $\beta\beta\alpha\beta$  Motiv wie in Vat-Nn liegen die Termini gegenüber den *psi-loops*. Dimerisierung geschieht bei der HIV-1 Protease durch zusätzliche Stränge an den Termini und auch bei Pepsin fand Fusion an den Termini statt. Dimerisierung ohne zirkuläre Permutation würde die aktiven *psi-loops* beider Domänen weit voneinander entfernt platzieren. Zirkuläre Permutation hingegen verschiebt die Termini in Nachbarschaft zu den *psi-loops* und Dimerisierung über die Termini erlaubt nun eine effiziente Wechselwirkung der *psi-loops* durch räumliche Nähe.

Ein zusätzliches Argument für die Verwandtschaft zwischen den Faltungen Acid protease und Double psi beta-barrel mag darin liegen, dass beide keine *superfolds* sind. Wenn konvergente Evolution angenommen wird, dann am ehesten für *superfolds*, da sie durch einen hohen Anteil von Supersekundärstrukturelementen und einer hohen Faltungseffizienz im Wettbewerb der Faltungen favorisiert werden. Faltungen hingegen, die keine *superfolds* sind, werden besser durch konvergente Evolution beschrieben, das die ökonomischen Eigenschaften der Natur und das Prinzip der Wiederverwendung einmal entstandener Strukturen besser widerspiegelt. Besonders für komplexe Faltungen wie die pseudoverknoteten Double psi beta-barrel ist eine parallele „Erfindung“ nur schwer vorstellbar. Wie unterschiedlich sich homologe Proteinfamilien entwickeln können wird eindrucksvoll von Alexey Murzin (Murzin, 1998) und Nick Grishin (Grishin, 2001) beschrieben. Insbesondere beschreibt Nick Grishin einige Beispiele von Faltungswechseln wie auch hier ein Faltungswechsel von Double psi beta-barrel zu Acid protease vorgeschlagen wird. Deswegen wird hier eine Alternative zur SCOP Klassifizierung vorgeschlagen, die Cdc48 N-terminal domain-like, die Retroviral und Pepsin-like Familien in einer Superfamilie zusammenfasst und die Idee einer gemeinsamen Abstammung gegenüber der Trennung aufgrund verschiedener Faltungen hervorhebt (Abbildung 13).



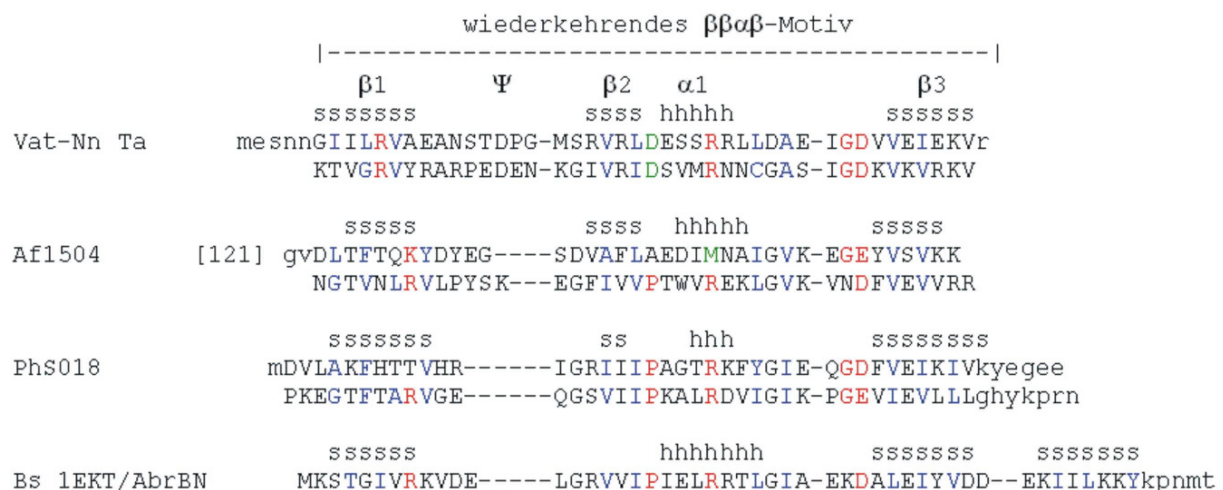
**Abbildung 13 Vorgeschlagene Klassifizierung nach dem SCOP Schema.** Links stehen die vier Hierarchieebenen der SCOP Klassifizierung. Rechts sind die Zweige der Double psi beta-barrel Proteine als Baumdiagramm abgebildet. Der Vorschlag streicht die Acid protease Faltung aus der bisherigen SCOP Klassifizierung und ordnet retrovirale und Pepsin-like Proteasen als Familien in die Superfamilie ADC-like Double psi beta-barrel ein, um die Verwandtschaft zwischen den Aspartatproteasen und Cdc48 N-terminal domain-like Proteinen zu berücksichtigen.

Die GFP-Entfaltungssassays zeigten, dass GFP-Varianten abhängig von ihrem Peptidanhang unterschiedlich schnell entfaltet werden. GFP mit dem *ssrA*-tag wird am schnellsten abgebaut, da es dem nativen Substrat von ClpA entspricht. GFP mit der zusätzlichen HIV-1 Protease Schnittstellenerkennungssequenz wurde ebenfalls entfaltet, aber deutlich langsamer als mit *ssrA*-tag. Natives GFP wurde nicht abgebaut. Ebenso konnte ein Einfluss der N-Domäne beobachtet werden. ClpAΔN entfaltete GFP viel schneller als natives ClpA. Die ATPase-Aktivität war in beiden Fällen jedoch gleich. Man kann daraus schließen, dass die N-Domäne einen regulierenden, selektiven Einfluss auf Kosten der Effizienz hat. Es konnte nicht erreicht werden, Proteasedomänen in einem AAA+-Kontext in GFP-Entfaltungssassays zu testen. Dies wurde durch die misslungene Rückfaltung der als *inclusion bodies* exprimierten Proteine HPR-ClpA und Saccharopepsin-ClpA verhindert, was durch die sehr geringe ATPase-Aktivität im Vergleich zu den nativen Proteinen ClpA und ClpAΔN deutlich wurde. Weitere Ansätze, vor allem mit dem Ziel, Entfaltungsaktivität von AAA+-Domänen zu modulieren, müssten erst eine lösliche Expression oder eine erfolgreiche Rückfaltung der chimären Proteine garantieren. Die meisten Proteasen, unter anderen HIV-1 Protease und Pepsin, werden in *E. coli* als *inclusion bodies* exprimiert. Es ist daher aus praktischen Gründen fraglich, ob diese beiden Proteine geeignete Kandidaten für N-Domänen von AAA+-ATPasen sind.

## 5 Evolution der Double Psi Beta-Barrel

### 5.1 Einleitung

In Kapitel 4 wurde vorgeschlagen, dass Double psi beta-barrel Proteine, wie Vat-Nn, und Aspartatproteasen von einem gemeinsamen Vorfahren abstammen. In Vat-Nn ist Symmetrie auf Sequenzebene erkennbar (Abbildung 14). Das bedeutet, dass man mit PSI-BLAST, einem Algorithmus zur Homologiedetektion, eine Hälfte des *barrels* mit der anderen Hälfte findet. Dies gibt einen Hinweis auf frühere Entwicklungsstufen und einen symmetrischen Vorfahren von Vat-Nn. Es ist damit denkbar, dass ein einfaches  $\beta\beta\alpha\beta$ -Element der Ursprung des Double psi beta-barrel Proteins Vat-Nn sein könnte, dass sich durch Genduplikation, Genfusion und wenigen symmetriebrechenden Mutationen entwickelte. Um dies zu überprüfen, suchte Andrei Lupas mit „Senser“, einem Algorithmus zur Suche von entfernt homologen Sequenzen (Koretke et al., 2001), nach homologen  $\beta\beta\alpha\beta$ -Elementen. Abbildung 14 zeigt das resultierende *alignment* einer begrenzten Auswahl gefundener homologer Proteine.



**Abbildung 14 Alignment von Vat-Nn Homologen.** Das *alignment* zeigt eine Auswahl der Ergebnisse der Suche mit Senser von Andrei Lupas. Zuoberst Vat-Nn, die N-terminale Domäne von Vat aus *Thermoplasma acidophilum*. Darunter der ORF Af1504 aus *A. fulgidus* ab Aminosäurerest 122, der ORF PhS018 aus *P. horikoshii* und der Transkriptionsfaktor AbrB aus *Bacillus subtilis*. Die jeweilige Sekundärstrukturvorhersage ist über den Sequenzen angegeben, wobei „h“ eine  $\alpha$ -helikale und „s“ eine  $\beta$ -Strang Umgebung ausdrückt. Konservierte hydrophobe Reste sind in blauer Farbe, konservierte hydrophile Reste sind in roter Farbe und auffällige Abweichungen konservierter Positionen sind in grüner Farbe hervorgehoben.

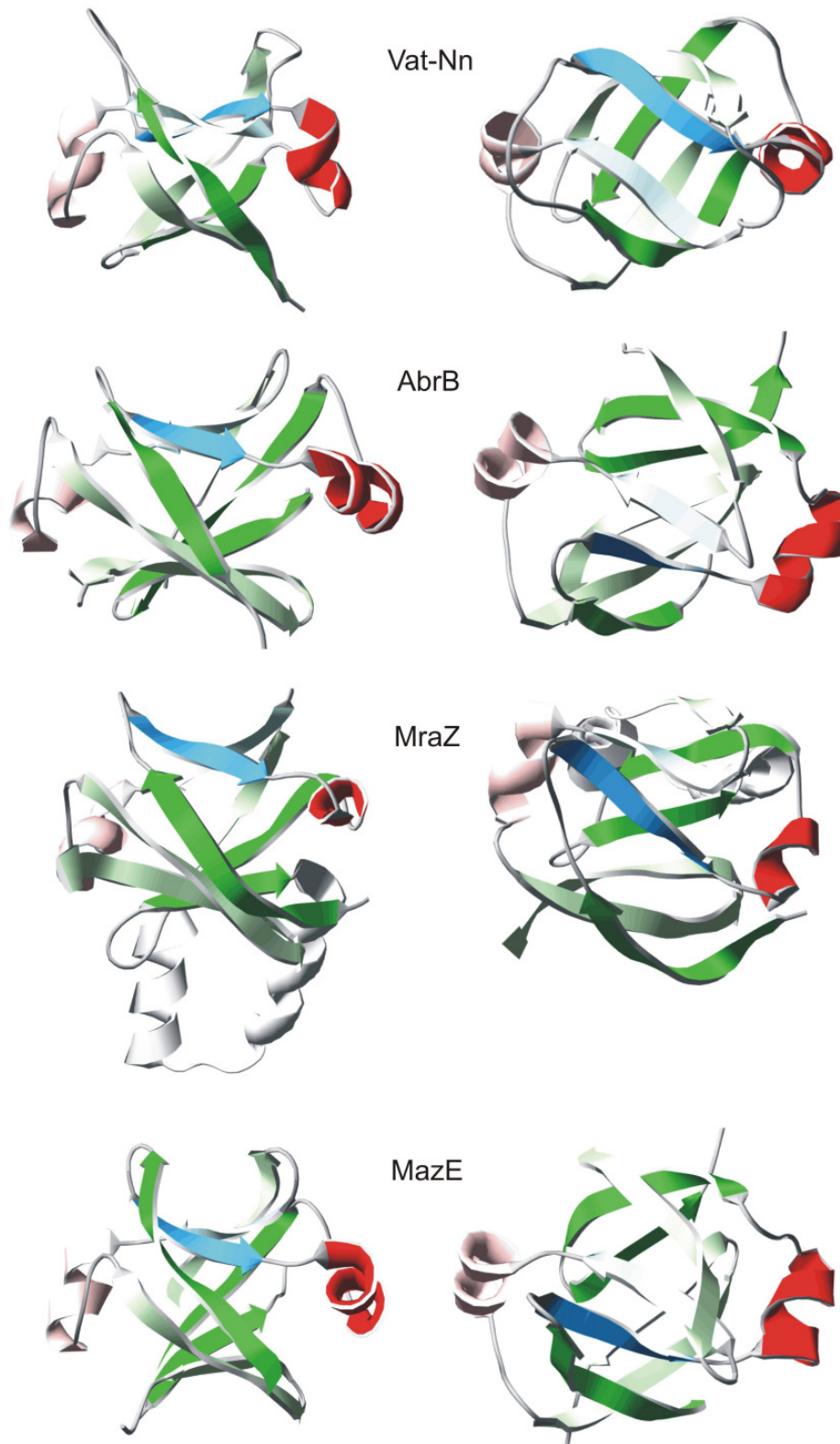
Der Transkriptionsfaktor AbrB von *Bacillus subtilis* ist der entfernteste gefundene Verwandte von Vat-Nn. Dazwischen liegen die ORFs Af1504 aus *A. fulgidus* und PhS018 aus *P. horikoshii*.



Zu diesem Zeitpunkt lag eine Struktur von AbrB vor, die ein  $\beta\alpha\beta$ -Motiv zeigte. Man nahm an, dass es über eine zirkuläre Permutation mit dem  $\beta\beta\alpha$ -Motiv von Vat-Nn verwandt war. Bioinformatische Studien und erneute NMR-Analyse von AbrB führten zur Revision der Struktur, die nun ein achtsträngiges *beta-barrel* mit zwei dimerisierenden  $\beta\beta\alpha\beta$ -Motiven zeigte (Coles et al., 2005) (Abbildung 15). Die Faltung von AbrB wurde als Swapped-hairpin barrel bezeichnet, da topologisches Vertauschen (englisch „swap“) der  $\beta$ 2-Stränge zu den *psi-loops* entsprechenden überkreuzten *loops* der Double *psi beta-barrel* Faltung führen würde. Strukturvergleich von Vat-Nn und AbrB zeigte die erwartete strukturelle Äquivalenz von Vat-Nn  $\beta$ 1 (aus  $\beta\beta\alpha\beta$ ) mit AbrB  $\beta$ 4 (aus  $\beta\beta\alpha\beta$ ). Aber der äquivalente  $\beta$ 4-Strang von AbrB lag, anders als erwartet, auf der symmetrisch gegenüberliegenden Peptidkettenseite des Dimers und konnte daher topologisch nicht durch zirkuläre Permutation mit  $\beta$ 1 von Vat-Nn korreliert werden. Deswegen wurde eine direkte Verwandtschaft dieser Proteine und dieser Faltungen über zirkuläre Permutation ausgeschlossen, obwohl eine Verwandtschaft zwischen beiden Faltungen aufgrund markanter Strukturähnlichkeiten erwogen wurde.

Es wurden Double *psi beta-barrel* und Swapped *hairpin-barrel* aufgrund der hohen Ähnlichkeit des  $\beta\alpha\beta$ -Motivs und der ähnlichen Erscheinung der *loops* zwischen  $\beta$ 1 und  $\beta$ 2 begrifflich zu den Cradle loop-barrel zusammengefasst. Zwei weitere Vertreter der Swapped *hairpin-barrel* mit bekannten Strukturen sind MazE und MraZ (Abbildung 15). MazE ist wie AbrB ein Dimer und ist Teil eines Toxin-Antitoxin Systems von *E. coli*. MraZ ist ein Monomer, dessen pseudosymmetrische Hälften durch ein helikalen *hairpin-turn* miteinander verbunden sind. MraZ ist ein Regulator der *E. coli* Zellwandbiosynthese. Aufgrund von Sequenz- und Strukturähnlichkeiten (Abbildung 14 und Abbildung 15) scheint die Verwandtschaft zwischen Double *psi beta-barrel* und Swapped *hairpin-barrel* plausibel. Es fehlt jedoch ein genetischer Mechanismus, der die evolutionäre Verwandtschaft erklären könnte. Zusätzliche Informationen waren nötig, um die Abstammung dieser Proteinfamilien zu verstehen.

Deshalb wurden die beiden Proteine PhS018 und Af1504 im Rahmen dieser Arbeit strukturell und funktionell untersucht. Die Struktur von PhS018 offenbarte eine neue Faltung, die RiSy-like barrel, mit der ein Evolutionsszenario entworfen werden konnte. Für Af1504 wurde eine Double *psi beta-barrel* Struktur vorhergesagt. Die Vorhersage konnte biochemisch unterstützt werden.



**Abbildung 15 Strukturen von Cradle loop-barrel Proteinen.** Von oben nach unten sind Vat-Nn (1cz4), AbrB (1yfb), MraZ (1n0G) und MazE (1ub4) dargestellt. Die rechte Abbildung geht aus der jeweils linken durch 90°-Drehung um die horizontale Achse hervor.  $\beta$ -Stränge sind grün-,  $\alpha$ -Helices rotfarbig markiert.  $\beta$ 2-Strang ist in blauer Farbe hervorgehoben. Die C-terminale Hälfte bzw. eine Hälfte des Dimers ist blasser abgebildet.

## 5.2 Materialien und Methoden

**Herstellung von PhS018** Die PhS018 Gensequenz wurde aus genomischer DNA von *P. horikoshii* OT3 (LGC Promochem) mit PCR amplifiziert. Mit Primern wurden terminale NdeI und HindIII Restriktionsschnittstellen eingeführt. Nach Restriktionsverdau wurde das Gen in den pET30b Expressionsvektor ligiert. Zur Expression wurde der Vektor in *E. coli* BL21 (DE3) Gold Zellen transformiert und in LB-Medium bei 37° C kultiviert. Die Zellen wurden bei OD ~ 0,6 mit 1 mM IPTG induziert und nach vier Stunden geerntet. Für die Isotopenmarkierung mit <sup>15</sup>N und <sup>15</sup>N/<sup>13</sup>C wurden die Zellen in M9 Minimalmedium (Anhang) kultiviert. <sup>15</sup>NH<sub>4</sub>Cl (0,7 g/l) und Glucose bzw. U-<sup>13</sup>C-Glucose waren dabei die einzigen Stickstoff- und Kohlenstoffquellen. Das lösliche Protein wurde mit Kationenaustauschchromatographie (SP Sepharose FastFlow, 10 ml, GE Healthcare) in 50 mM Natriumphosphatpuffer pH 7,4 und Elution mit einem linear ansteigenden NaCl-Gradienten von 50-525 mM NaCl über 10 Säulenvolumen gereinigt. Unmittelbar nach der Elution von der Säule präzipitierte das Protein. Das Präzipitat wurde zweimal durch Suspendieren in PBS (20 mM Natriumphosphat pH 7,2, 150 mM NaCl, 2 mM KCl), anschließender Zentrifugation und Dekantieren des Überstands gewaschen. Das Proteinpräzipitat wurde in 50 mM Natriumphosphatpuffer pH 3, 50 mM NaCl, 0,02% Natriumazid aufgelöst und bis 500 µM für NMR Analyse ankonzentriert. Die Proteinmenge wurde über Absorption bei 280 nm und die Reinheit wurde mit SDS-PAGE bestimmt. Es wurden noch folgende Additive zur Erhöhung der Löslichkeit von PhS018 bei neutralen pH-Werten getestet: 0,3 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,5 M KCl, 0,2 M MgCl<sub>2</sub>, 0,5 M Arginin, 10% Glycerol, 1 M Betain, 0,02% Triton X-100, 0,65% CHAPS, 0,73% Octylglucopyranosid, 0,01% SB14 Zwittergent, 100 mM Tris mit 300 mM NaCl, 1 mM ATP mit 5 mM MgCl<sub>2</sub>, 15% Ethanol, 3 M TMAO, 0,45 M Arginin mit 10% Glycerol (alles Endkonzentrationen).

**NMR Strukturaufklärung** Murray Coles (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) nahm am Bayrischen NMR Zentrum der TU München die NMR-Spektren auf und berechnete die dreidimensionale Struktur von PhS018 (pdbid 2GLW).

**DNA Bindung von PhS018** PhS018 wurde zusammen mit PCR-erzeugter DNA (~ 300 Basenpaare lang) 30 Minuten lang bei Raumtemperatur inkubiert. Anschließend wurde das Migrationsverhalten der DNA in einem 2% Agarosegel beobachtet. Der Testansatz enthielt 100 mM Tris, pH 8, 1 mM ATP, 10 mM MgCl<sub>2</sub>, 15% Glycerol, 0,9 pmol DNA, 0,55 mg/ml BSA, 412 mM KCl. Um verschiedene Proteinendkonzentrationen zu testen, ohne die Zusammensetzung des Testansatzes zu verändern, wurde stets 1 µl einer entsprechend vorverdünnten PhS018-lösung (50 mM Natriumphosphat pH 3, 50 mM NaCl) zu 19 µl Testlösung gegeben. Als Kontrollprotein für unspezifische DNA-Bindung wurde Lysozym (14,3 kD, pI 9,3 ) verwendet, was. PhS018 (10,3 kD, pI 9,0) in Größe und Ladung ähnelt. Um zu testen, ob PhS018 selektiv an verschiedene DNA-Sequenzen bindet, wurden zwei

unterschiedliche PCR-Produkte mit etwa gleicher Länge verglichen. Zum einen die Promoter-Sequenz von PhS018, auf der eine spezifische Bindestelle für PhS018 denkbar wäre. Zum anderen eine heterologe DNA-Sequenz als Kontrolle: Gen der HIV-1 Protease. Da PhS018 bei neutralen pH-Werten zur Aggregation neigt, wurde nur der Überstand des Testansatzes nach vorangehender Zentrifugation auf das Agarosegel aufgetragen.

**Strukturvorhersage von Af1504** Für die Vorhersage von *coiled coil* Strukturen wurde das Programm „Coils“ von Andrei Lupas mit der kompletten Sequenz von Af1504 als Eingabe verwendet (Lupas et al., 1991). Für die *ab initio* Strukturvorhersage mit „Rossetta“ (Rohl et al., 2004) wurde die Sequenz von Af1504 verwendet, für die keine *coiled coil* Struktur vorhergesagt wurde: G122-R201. Der gleiche Sequenzabschnitt wurde als Eingabe für „HHpred“ verwendet; ein Programm zur empfindlichen Homologiesuche (Soding et al., 2005). Gesucht wurde gegen die pdb70 (gefilterte RCSB Protein Data Bank mit Strukturen von Proteinen, die weniger als 70% Sequenzidentität aufweisen) vom 20.04.06, mit maximal 8 PSI-BLAST Iterationen, mit einem E-value threshold von 0,001 für PSI-BLAST und mit lokalen *alignments*. Als Templat für die Strukturvorhersage aufgrund von Homologie wurde das von HHpred als ähnlichstes eingestufte Protein, Vat-Nn, verwendet. Für die Modellierung wurde das Programm „Modeller“ verwendet. (Sali et al., 1995). Dem von HHpred dafür automatisch generierten *alignment* von Vat-Nn und Af1504 fehlte der erste  $\beta$ -Strang. Dieser wurde gemäß dem *alignment* in Abbildung 14 manuell ergänzt.

**Herstellung von Af1504** Das *af1504* Gensegment wurde mit PCR aus *A. fulgidus* DSM 4303 genomischer DNA (LGC Promochem) vervielfältigt und isoliert. Mit Primern wurden die Restriktionsschnittstellen NdeI und XhoI eingeführt, über die das verdaute PCR-Produkt in das Expressionsplasmid pET30b (Novagen) ligiert wurde. Für die Expression wurde das Plasmid in *E. coli* BI21 (DE3) Gold Zellen transformiert. Die Zellen wuchsen bei 37° C bis zu einer OD von ~ 0,6, wurden mit 1 mM IPTG induziert und nach vier Stunden geerntet. Nach Resuspendierung in 20 mM HEPES, pH 7,2, 150 mM NaCl wurden die Zellen lysiert. Lösliches Af1504 im Überstand wurde mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 20 mM HEPES, pH 7,2 mit einem linear ansteigendem NaCl-Gradienten über 10 Säulenvolumen von 150-575 mM NaCl gereinigt. In den Puffern für die Chromatographie war 1 mM DTT enthalten. Fraktionen mit Af1504 wurden vereinigt und auf Eis solange mit Ammoniumsulfat versetzt bis heftiger Niederschlag auftrat. Der Niederschlag wurde durch Zentrifugation vom Überstand getrennt, in 50 mM Natriumphosphat, pH 7,4, 50 mM NaCl gelöst und anschließend mit Gelfiltration in 50 mM Natriumphosphat, pH 7,4, 50 mM NaCl (Superdex S200 26/60, GE Healthcare) gereinigt. Die Proteinmenge wurde über Absorption bei 280 nm und die Reinheit wurde mit SDS-PAGE bestimmt.

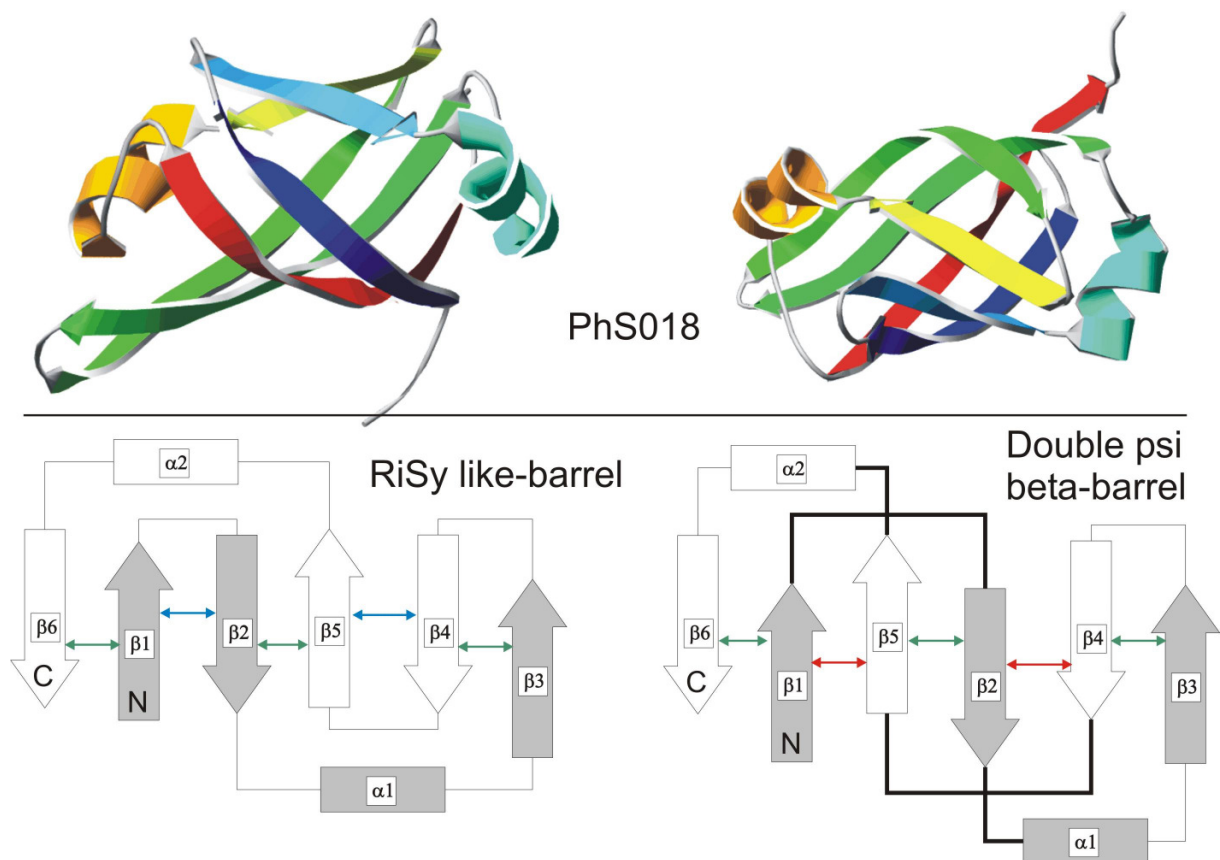
**Biochemische Analyse von Af1504** CD-Spektren wurden mit einem Jasco CD-Spektropolarimeter J-810 und Fluoreszenzspektren wurden mit einem Jasco FP-6500 Spektrofluorometer aufgenommen. Die Anregungswellenlänge für Tryptophanfluoreszenz betrug 293 nm. Analytische Gelfiltration (Superose 6 10/300 GL, GE Healthcare) wurde in PBS durchgeführt und das Elutionsprofil wurde über die Absorption bei 280 nm verfolgt. Molekülmassenabschätzung erfolgte durch Vergleich mit Elutionsvolumina von Proteinen bekannter Molekülgröße (Ferritin 440 kD, Aldolase 150 kD, BSA 66 kD, Carboanhydrase 29 kD, Aprotinin 6 kD). Für Native PAGE wurde ein 10% diskontinuierliches Polyacrylamidgel nach Laemmli ohne SDS (Anhang) verwendet. Für limitierte Proteolyse wurden 10 µl Tris-gepufferter Lösung mit 0,8 µg/µl Af1504 verwendet und die Proteasen Chymotrypsin, Trypsin und Proteinase K mit je 0,6 µg/ml Endkonzentration zugegeben. Inkubiert wurde 30 Minuten lang bei Raumtemperatur, danach wurde mit PMSF gestoppt und die Proteinzusammensetzung mit SDS-PAGE analysiert. Für Kreuzvernetzung der Proteine mit Glutaraldehyd wurde PBS-gepufferte Proteinlösung mit 1 µg/µl gereinigtes Af1504 verwendet und Glutaraldehyd auf eine Endkonzentration von 0,16% zugegeben. Inkubiert wurde 60 Minuten auf Eis und 10 µl des Ansatzes wurden mit 2µl 3M Tris-Stammlösung gestoppt und anschließend mit SDS-PAGE analysiert.

### 5.3 Ergebnisse

**Expression von PhS018** Das Protein PhS018 wurde in *E. coli* als lösliches Protein exprimiert und wurde aus dem Lyseüberstand gereinigt. Nach der Reinigung wurde das Protein in Phosphatpuffer mit pH 3 gelöst. Oberhalb von pH 3 war das Protein nahezu unlöslich. CD Analyse zeigte ein Spektrum mit hohem Anteil an  $\beta$ -Faltblatt Sekundärstrukturen, was an einem breiten negativem Signal bei 218 nm zu erkennen war (ohne Abbildung). Es wurden Additive getestet, welche die Löslichkeit von PhS018 bei neutralen pH-Werten erhöhen sollten. Von den verschiedenen Additiven zur Erhöhung der Löslichkeit bei pH 7 zeigte nur Arginin einen Einfluss. Mit 0,7 M Arginin bleiben bei pH 7 etwa 50 µM PhS018 löslich, bei 1 M Arginin sind es 100 µM. Die hohen Argininkonzentrationen erschwerten die weitere biophysikalische Analyse. Mit NMR-Spektroskopie konnte dennoch ein ähnliches Spektrum wie bei pH 3 aufgenommen werden. Die NMR-Spektren eigneten sich trotzdem nicht zur Strukturaufklärung, weil einige Signale durch Arginin verdeckt waren und weil die erreichbare Konzentration von PhS018 noch zu niedrig war. Die Löslichkeitsversuche haben ergeben, dass unter allen Bedingungen stets eine kleine Menge PhS018 löslich war. Dies ermöglichte Experimente, wie DNA-Bindungstests, die nur kleine Mengen PhS018 erforderten.

**NMR-Struktur von PhS018** Das Protein ist ein hochsymmetrisch sechssträngiges *barrel* Protein aus zwei  $\beta\beta\alpha\beta$ -Einheiten und ähnelt den Double psi beta-barrel Proteinen (Abbildung

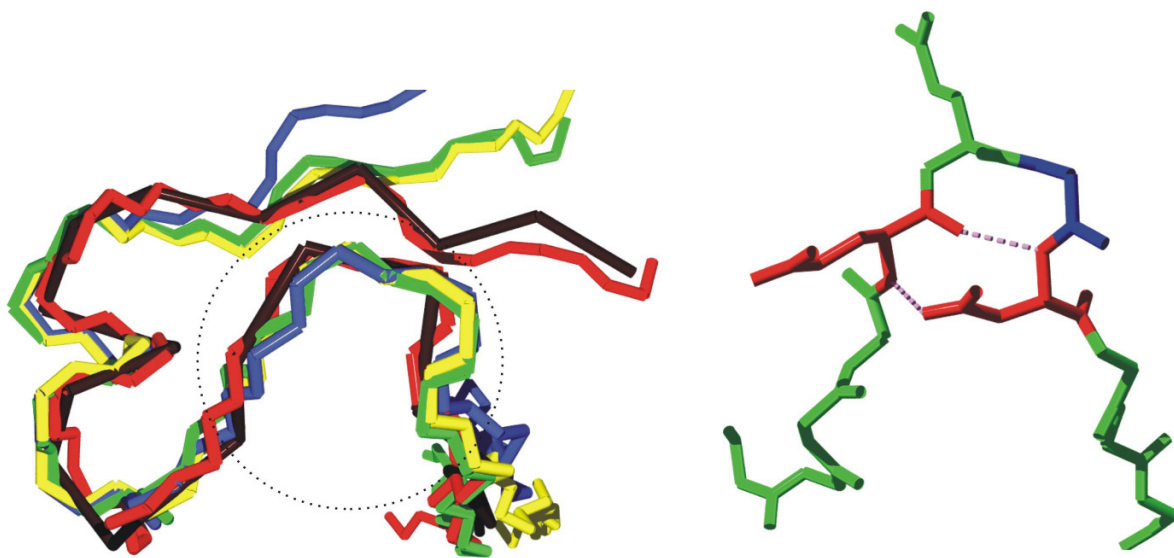
16). Unterschiede zwischen der Double psi beta-barrel - und der PhS018 Struktur bestehen im Muster der  $\beta$ -Strang-Interaktionen. In Double psi beta-barrel Proteinen sind  $\beta 1$  mit  $\beta 5$  und  $\beta 2$  mit  $\beta 4$  parallel verknüpft (Abbildung 16). Hierbei sind die zwei symmetrischen Hälften vollkommen miteinander verzahnt, da jeder Strang nur mit Strängen aus der symmetrieverwandten Hälfte verknüpft ist. In PhS018 findet man stattdessen Verknüpfung von  $\beta 1$  mit  $\beta 2$  und  $\beta 4$  mit  $\beta 5$ . Hier findet die Strangverknüpfung innerhalb der symmetrischen Hälften statt.  $\beta 2$  und  $\beta 5$  bilden jeweils nur eine einzige Wasserstoffbrücke mit  $\beta 1$  beziehungsweise mit  $\beta 4$ . Dies lässt das Protein als relativ offenes *barrel* erscheinen. Starke Wechselwirkung der antiparallelen Stränge  $\beta 2/\beta 5$  sind jedoch in beiden Strukturen zu finden.



**Abbildung 16 Struktur von PhS018.** (Oben) Struktur von PhS018 als Bändermodell. Die Abfolge der Sekundärstrukturelemente ist vom N- zum C-Terminus farbkodiert von Blau über Grün und Gelb zu Rot. Die rechte Struktur geht aus der linken durch 90°-Drehung um die horizontale Achse hervor. (Unten) Topologischer Vergleich zwischen PhS018 (Links) und Vat-Nn (Rechts). Doppelpfeile deuten auf Verknüpfung der  $\beta$ -Stränge. (Blau) Antiparallele Stränge in derselben symmetrischen Hälfte. (Grün) Antiparallele Stränge in verschiedenen symmetrischen Hälften. (Rot) Parallele Stränge in verschiedenen symmetrischen Hälften.

In PhS018 ist dies die einzige Interaktion der beiden Symmetriehälften. Im topologischen Vergleich sind in PhS018 die Positionen der Stränge  $\beta 2$  und  $\beta 5$  gegenüber einem Double psi beta-barrel vertauscht. Diese sechssträngige Greek-key Topologie (Greek-key bezeichnet eine  $\beta$ -Faltblatt Anordnung, in der der erste und der letzte Strang ohne überkreuzende Elemente miteinander verknüpft sind) ähnelt der Faltung der Riboflavinsynthasen (SCOP:

Reductase/Isomerase/Elongation factor common domain). Durch den Austausch der Positionen von  $\beta 2$  und  $\beta 5$  wurden *psi-loop* Strukturen, wie sie in den Double *psi beta-barrel* Proteinen vorkommen, umgangen. Die *loops* zwischen  $\beta 1$ - $\beta 2$  und  $\beta 4$ - $\beta 5$  sind bei PhS018 zudem sehr kurz, was die Erwartung aus der Sequenzanalyse erfüllt (Abbildung 14). Sie bestehen aus nur drei Resten: H12-L14 und G57-Q59. Ein äquivalenter topologischer Austausch der  $\beta$ -Stränge ist in Swapped hairpin-barrel Proteinen zu beobachten. PhS018 unterscheidet sich von den Swapped hairpin-barrel Proteinen aber dadurch, dass es ein sechssträngiges anstatt eines achtsträngigen *barrel* formt. Ein charakteristisches Merkmal, das beide symmetrischen Hälften von PhS018 aufweisen und mit den Cradle loop-barrel Proteinen teilen, ist die GD Box Sequenz zwischen der  $\alpha$ -Helix und dem nachfolgendem  $\beta$ -Strang: hydrophob-x-hydrophob-x-x-G-sauer-x-hydrophob (Abbildung 14).



**Abbildung 17 Struktur der GD-Box.** (Links) Überlagerung von  $\beta\alpha\beta$  Fragmenten: Vat-Nn *N-terminal* (schwarz), Vat-Nn *C-terminal* (rot), PhS018 *N-terminal* (grün), PhS018 *C-terminal* (blau), AbrB (gelb). (Rechts) Der mit einem Kreis markierte, grüne Strang des linken Bildes.

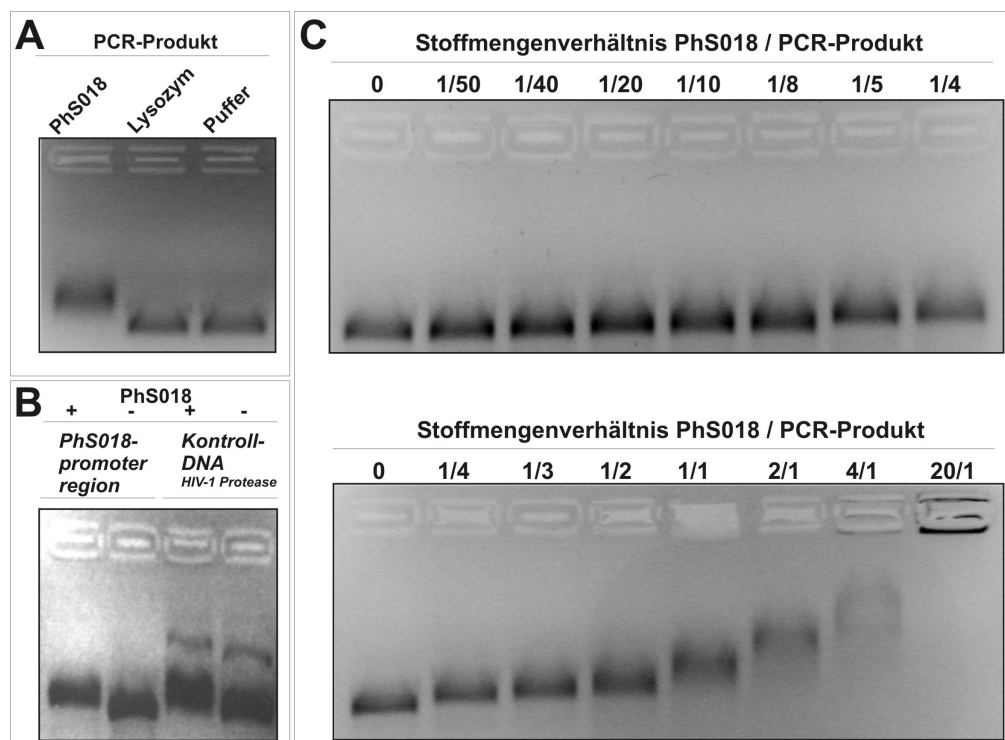
Der jeweils erste hydrophobe Rest (Y27 und I72 in PhS018) verankert die vorangehende  $\alpha$ -Helix im Kern des Proteins. Der folgende  $\beta$ -Strang verläuft quer über das *barrel* und wird vom zweiten hydrophoben Rest des Motivs (I29 und I74) im Kern verankert. Glycin besetzt die  $i+2$  Position eines Typ II  $\beta$ -turn und die folgende saure Seitenkette bildet eine Wasserstoffbrücke zum Amidstickstoff des Restes  $i$  (Abbildung 17). Ein Vergleich mehrerer GD-Box Strukturen von Cradle loop-barrel Proteinen und PhS018 zeigt, dass die GD-Box ein Teil eines konservierten  $\beta\alpha\beta$  Fragments ist (Abbildung 17). Die größte Strukturähnlichkeiten findet man bei beiden Hälften von PhS018 und AbrB mit etwa 0,5 Å RMSD (Tabelle 2). Die beiden Motive von Vat-Nn zeigen zueinander hohe Ähnlichkeit (0,78 Å RMSD), zeigen aber zu den Motiven von PhS018 und AbrB etwas höhere Abweichungen. RMSD-Werte liegen

hier zwischen 1,08 Å (Vat-Nn N-terminale Hälfte und PhS018 C-terminale Hälfte) und 1,46 Å (Vat-Nn C-terminale Hälfte und PhS018 N-terminale Hälfte).

**Tabelle 2 RMSD-Werte der paarweisen Überlagerung der C $\alpha$ -Ketten von  $\beta\alpha\beta$  Motiven mit GD-Box-turns.** Bei Vat-Nn und PhS018 wurden die N-terminale „N“ und die C-terminale Hälfte „C“ getrennt analysiert. Alle Werte in [Å].

	Vat-Nn N	Vat-Nn C	PhS018 N	PhS018 C	AbrB
Vat-Nn N					
Vat-Nn C	0,78				
PhS018 N	1,33	1,46			
PhS018 C	1,08	1,20	0,56		
AbrB	1,24	1,38	0,48	0,50	

**DNA-Bindung von PhS018** PhS018 verlangsamte die Migration von DNA-Fragmenten in Agarosegelen (Abbildung 18). Um einen unspezifischen Effekt aufgrund positiver Ladung des Proteins auszuschließen, wurde ein Kontrollversuch mit Lysozym durchgeführt, das in Größe und Ladung PhS018 ähnelt. Lysozym zeigte keinen Effekt auf die DNA-Migration (Abbildung 18 A).

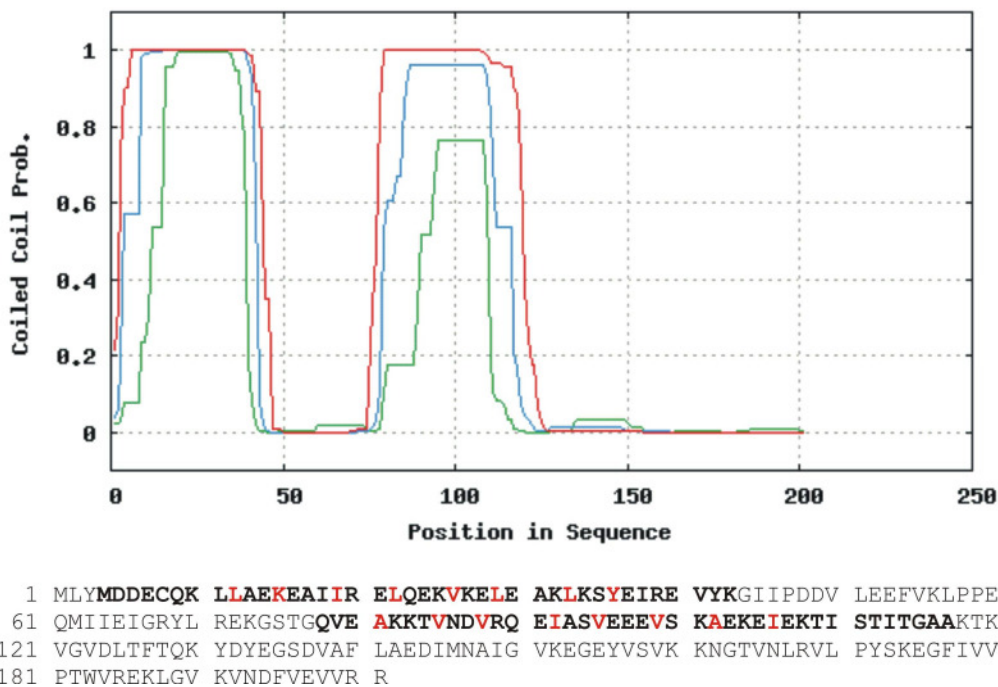


**Abbildung 18 PhS018 retardierte die Migration von DNA in Agarosegelen.** (A) Vergleich des Einflusses auf DNA-Migration von PhS018, Lysozym und Puffer ohne Protein. (B) Vergleich des Einflusses von PhS018 auf das Migrationsverhalten unterschiedlicher Nukleotidsequenzen. (C) DNA-Migrationsverhalten bei Anwesenheit steigender Mengen PhS018.



Um Spezifität der Bindung für verschiedene Nukleotidsequenzen zu testen, wurden zwei unterschiedliche DNA-Fragmente eingesetzt (Abbildung 18 B). Es konnte jedoch kein Unterschied des Retardierungseffektes bei Einsatz verschiedener DNA-Sequenzen festgestellt werden. Titrationsexperimente, bei denen eine definierte DNA-Menge mit unterschiedlichen Proteinmengen inkubiert wurde, zeigten, dass der Retardierungseffekt mit zunehmender Proteinmenge zunimmt (Abbildung 18 C). Die Retardierung nahm kontinuierlich und nicht sprunghaft zu, was auf mehr als eine mögliche Bindestelle auf der DNA-Sequenz schließen lässt.

**Struktur von Af1504** Das Programm „Coils“ detektierte in der Sequenz von Af1504 zwei N-terminal gelegene Abschnitte mit hohem Potential für *coiled coil* Strukturen. Die Abschnitte umfassen M4-K43 und Q78-A117 (Abbildung 19). Sie wurden so gewählt, dass in mindestens zwei der drei Berechnungsfenster über 90% Wahrscheinlichkeit für eine *coiled coil* Struktur angegeben wurde. Diese Bereiche könnten ein antiparalleles Vierhelixbündel und damit eine Dimerisierungsdomäne bilden.



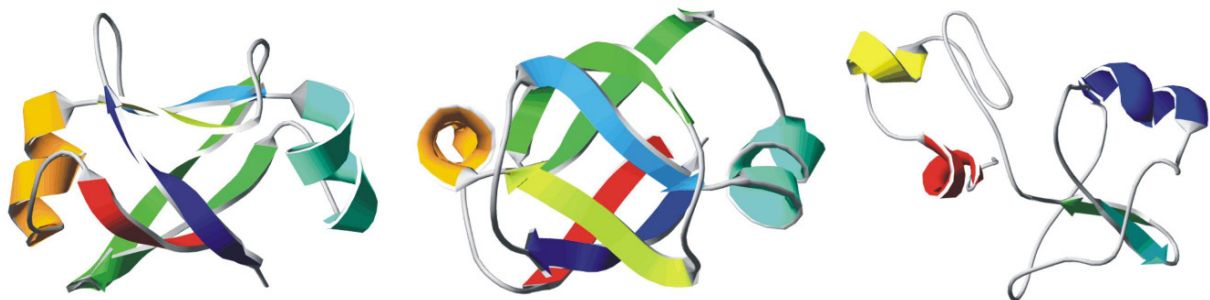
**Abbildung 19 Coiled coil Strukturvorhersage von Af1504.** Ausschnitt aus dem Resultatfenster des *coiled coil* Vorhersageprogramms „Coils“, in dem graphisch die Wahrscheinlichkeit für eine *coiled coil* Struktur gegen die Aminosäureposition der Af1504-Sequenz aufgetragen ist. Die drei Graphen entsprechen drei verschiedenen Berechnungsfenstern des Programms. In der darunter abgebildeten Sequenz von Af1504 ist der vorhergesagte *coiled coil* Bereich fett hervorgehoben und die hydrophoben, kernbildenden Reste sind rot markiert.

Der Sequenzabschnitt C-terminal zu diesen potentiellen *coiled coil* bildenden Sequenzen umfasst die Aminosäuren G122-R201. Mit Sensor wurde diese Sequenz bereits als homolog zu Vat-Nn erkannt und aligniert (Abbildung 14). In dieser Arbeit wurde der Bereich zusätzlich mit HHpred analysiert. HHpred ist ein Programm zur Suche von homologen Proteinsequenzen. Durch die Verwendung von Sequenzprofilen und Sekundärstrukturvorhersagen gewinnt es gegenüber PSI-BLAST und Sensor an Empfindlichkeit. Als Ergebnis liefert es eine Liste mit wahrscheinlich homologen Proteinen. Im Falle von Af1504<sub>122-201</sub> wurden vier Proteine gefunden, die mit über 20% Wahrscheinlichkeit homolog zu Af1504 sind – allesamt Double psi beta-barrel Proteine.

**Tabelle 3 Ergebnis der HHpred Analyse von Af1504 (G122-R201).**

Reihenfolge	Name	pdbid	Wahrscheinlichkeit [%]	E-value
1	Vat-Nn	1cz4_A	81,5	0,017
2	p97-N	1e32_A	63,4	0,4
3	Ufd1	1zc1_A	46,4	0,86
4	ADC	1pqh_A	23,5	7,1

Die niedrigen E-value Werte deuten an, dass diese Proteine mit weniger empfindlichen Algorithmen wahrscheinlich nicht gefunden worden wären. Zu den gefundenen Proteinen gehören Vat-Nn, die N-terminale Domäne von p97, das Ubiquitin fusion degradation protein Ufd1 und Aspartat-1-Decarboxylase ADC. Vat-Nn als bester Treffer wurde als Templat für Strukturmodellierung mit Hilfe von Homologie verwendet. Das Resultat ist eine Double psi beta-barrel Struktur mit den charakteristischen *psi-loops* (Abbildung 20).

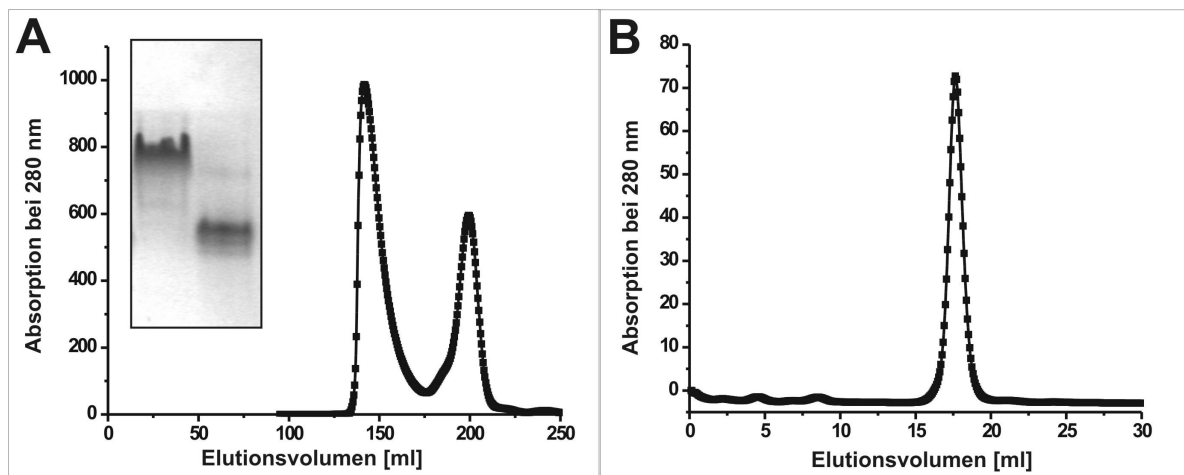


**Abbildung 20 Strukturmodelle von Af1504 (G122-R201) in Bänderdarstellung.** Die Abfolge der Sekundärstrukturelemente vom N- zum C-Terminus ist von Blau über Grün und Gelb nach Rot farbkodiert. (Links) Ergebnis der Modellierung aufgrund von Homologie, (Mitte) linke Struktur um 90° um die horizontale Achse gedreht, (Rechts) *ab initio* Modellierung von Rosetta.

Die *ab initio* Strukturvorhersage mit Rosetta ergab Strukturen ohne Ähnlichkeit zu bekannten Proteinen. Trotz der Alignierbarkeit von Af1504 mit Vat-Nn und AbrB aufgrund

der Ergebnisse von Senser (Abbildung 14), war selbst zu diesen beiden Proteinen keine Strukturähnlichkeit auszumachen.

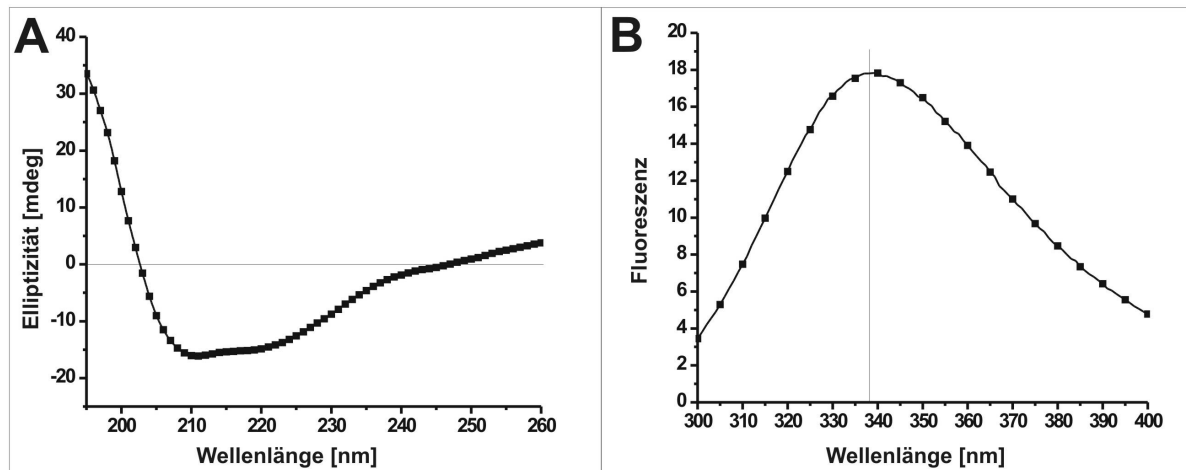
**Biochemische Analyse von Af1504** Das Protein Af1504 konnte rekombinant in *E. coli* hergestellt und gereinigt werden. Es eluierte bei der Reinigung von der Gelfiltrationssäule in zwei getrennten Fraktionen, was im UV-Absorptionsprofil sichtbar war (Abbildung 21). Über das Elutionsvolumen von der kalibrierten Gelfiltrationssäule wurden Molekülmassen von etwa 400 kD für die frühe und etwa 60 kD für die späte Fraktion abgeschätzt. 60 kD entsprechen im Rahmen der Messgenauigkeit der Größe eines dimeren Af1504 mit einer berechneten Molekülmasse von 52 kD. SDS-PAGE Analyse zeigte, dass beide Fraktionen Af1504 enthielten. Das Migrationsverhalten von Af1504 bei nativer PAGE zeigte einen Unterschied zwischen Protein aus der frühen und der späten Fraktion (Abbildung 21 A Kasten). Die Probe der frühen Fraktion migrierte bei nativer PAGE langsamer als eine Probe der späten Fraktion. Da es sich um das gleiche Protein mit stets gleicher Ladungsdichte handelt, deutet dies auf größere Proteinkomplexe in der früheren Fraktion. Die spätere Fraktion wurde erneut mit Gelfiltration untersucht und ergab einen einzigen Peak bei erwartetem Elutionsvolumen (Abbildung 21 B).



**Abbildung 21** Elutionsprofil des Gelfiltrationslaufs von Af1504. (A) Lauf während der Reinigung. Der eingefügte Kasten zeigt das Laufverhalten auf einem Nativgel. Die linke Gelspur enthält eine Probe des frühen Peaks, die rechte Gelspur enthält eine Probe des späten Peaks. (B) Elutionsprofil des Gelfiltrationslauf mit einer Probe des späten Peaks aus (A).

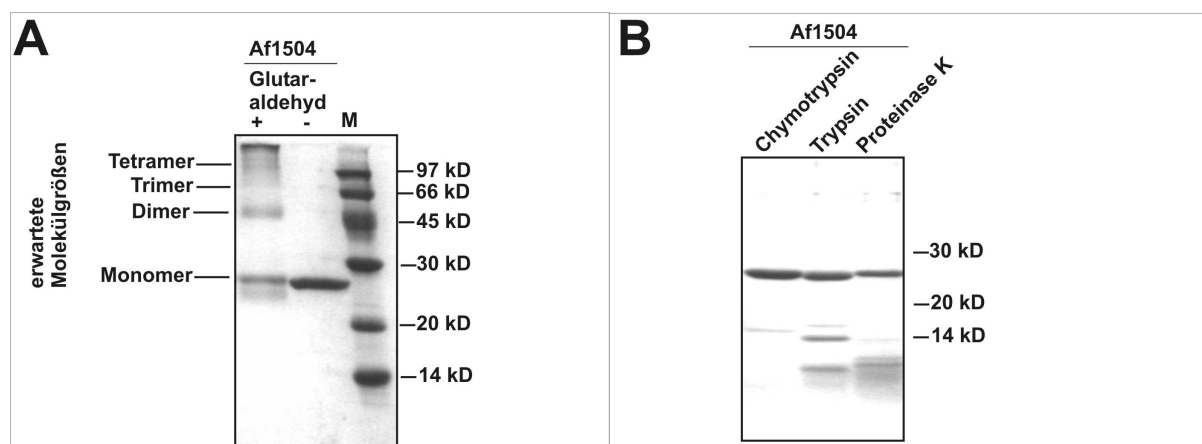
Zusammengenommen deutet dies auf zwei verschiedene Oligomerisierungsgrade von Af1504 hin, die nicht miteinander im Gleichgewicht stehen. Für weitere Analysen wurde nur das niedermolekulare Af1504 verwendet. Das CD-Spektrum zeigte zwei negative Peaks bei 209 nm und 221 nm wie sie typischerweise in  $\alpha$ -helikalen Proteinen erscheinen (Abbildung 22 A). Im CD-Spektrum dominierten die  $\alpha$ -helikalen Anteile, die vermutlich von der *coiled coil* Struktur stammten. Da  $\beta$ -Faltblätter geringere CD-Signale geben, können sie bei starken  $\alpha$ -helikalen Signalen überdeckt sein. Dennoch war ein Anteil von  $\beta$ -Faltblatt im CD-Spektrum

zu erkennen, dessen Signal ein negativer Peak bei 215 nm ist. Im gemessenen Spektrum wurde dies daran erkennbar, dass die  $\alpha$ -helikalen Peaks kaum voneinander getrennt waren. Das Spektrum der Tryptophanfluoreszenz zeigte maximale Intensität bei einer Wellenlänge von 338 nm (Abbildung 22 B). Dieser Wert zeigt eine hydrophobe Umgebung der angeregten Tryptophanreste an, wie sie im Inneren eines gefalteten Proteins vorkommt. Zum Vergleich liegt die Tryptophanfluoreszenz in wässriger Umgebung bei 358 nm.



**Abbildung 22 Spektroskopische Untersuchung von Af1504 (A)** CD-Spektrum von Af1504. **(B)** Spektrum der Tryptophanfluoreszenz von Af1504.

Kreuzvernetzung mit Glutaraldehyd ergab eine Proteinleiter zu höheren Molekülmassen. Die Bande, die einem Dimer oder einem Tetramer entsprechen waren sichtbar, die Trimerbande fehlte hingegen (Abbildung 23 A) Das Auftreten der Oligomere als vielfaches von zwei deutet in Übereinstimmung mit den Gelfiltrationsdaten auf eine Dimerstruktur.



**Abbildung 23 Biochemische Analyse von Af1504. (A)** SDS-PAGE Analyse der Kreuzvernetzung von Af1504 mit Glutaraldehyd. "M" steht für Markerproteine mit den rechts angegebenen Molekülmassen **(B)** Limitierte Proteolyse von Af1504.

Limitierte Proteolyse mit verschiedenen Proteasen ergab bei Trypsin zwei Abbaubanden auf einem SDS-PAGE Gel (Abbildung 23 B). Dies deutet auf die Zusammensetzung des Proteins

aus zwei kompakten, proteaseresistenten Domänen hin, die durch eine flexible, proteasesensitive Verbindung miteinander verbunden sind: Eine Bande bei 14 kD und eine bei etwa 9 kD (was durch den Marker nicht mehr abgedeckt ist und abgeschätzt werden musste). Das berechnete Gewicht von Af1504<sub>1-121</sub>, das den potentiellen *coiled coil* Bereich umfasst, beträgt 13,8 kD und Trypsinschnittstellen K118 und K120 sind am Ende dieses Bereichs vorhanden. Chymotrypsin baute Af1504 nicht ab. Proteinase K Verdau ergab nur eine Abbaubande bei etwa 9 kD. Bei dieser Bande fällt auf, dass sie als breites Band erscheint. Eine vergleichbare Bande bei 14 kD wie beim Trypsinverdau fehlte. Spaltung der Verbindungssequenz konnte die beiden getrennten Domänen destabilisieren, so dass durch die unspezifisch spaltende Proteinase K weiterer Abbau der einzelnen Domänen stattfand.

## 5.4 Diskussion

Auf der Suche nach evolutionären Verbindungen zwischen Double psi beta-barrel und Swapped hairpin-barrel Proteinen konnte ein Protein identifiziert und strukturell charakterisiert werden, welches intermediäre Eigenschaften aufweist: PhS018. Bei dieser Struktur handelt es sich um eine neue Faltung als weitere Variante innerhalb der Cradle loop-barrels. Die Struktur ähnelt dem Greek-key barrel der *E. coli* Riboflavinsynthasen und wurde hier als RiSy like-barrel bezeichnet. PhS018 ist wie die Double psi beta-barrel ein sechssträngiges pseudosymmetrisches *barrel*. Beide Hälften wechselwirken über nur ein Paar von  $\beta$ -Strängen, was das barrel offen erscheinen lässt. Wie in den Swapped hairpin-barrels fehlen überkreuzte psi-loops, die durch den topologischen Austausch von  $\beta 2$  mit  $\beta 5$  vermieden werden. Die loops zwischen  $\beta 1/\beta 2$  und  $\beta 4/\beta 5$  sind in PhS108 zwar mit drei Aminosäuren sehr kurz, ähneln aber strukturell und topologisch den cradle-loops. Zudem weist PhS018 das für Cradle loop-Proteine charakteristische antiparallele  $\beta 2/\beta 5$ -Strangpaar auf. Zusätzlich konnte in PhS018 ein  $\beta\alpha\beta$ -Fragment mit einem GD-Box-Motiv identifiziert werden. Dieses Fragment zeigt innerhalb der Cradle loop-barrel Proteine fast identische Strukturen. Zusammengenommen kann PhS108 als neue Faltung der Cradle loop-barrel Proteine betrachtet werden und steht intermediär zwischen den Double psi beta-barrel und den Swapped hairpin-barrel Proteinen, da es charakteristische Merkmale beider Gruppen aufweist.

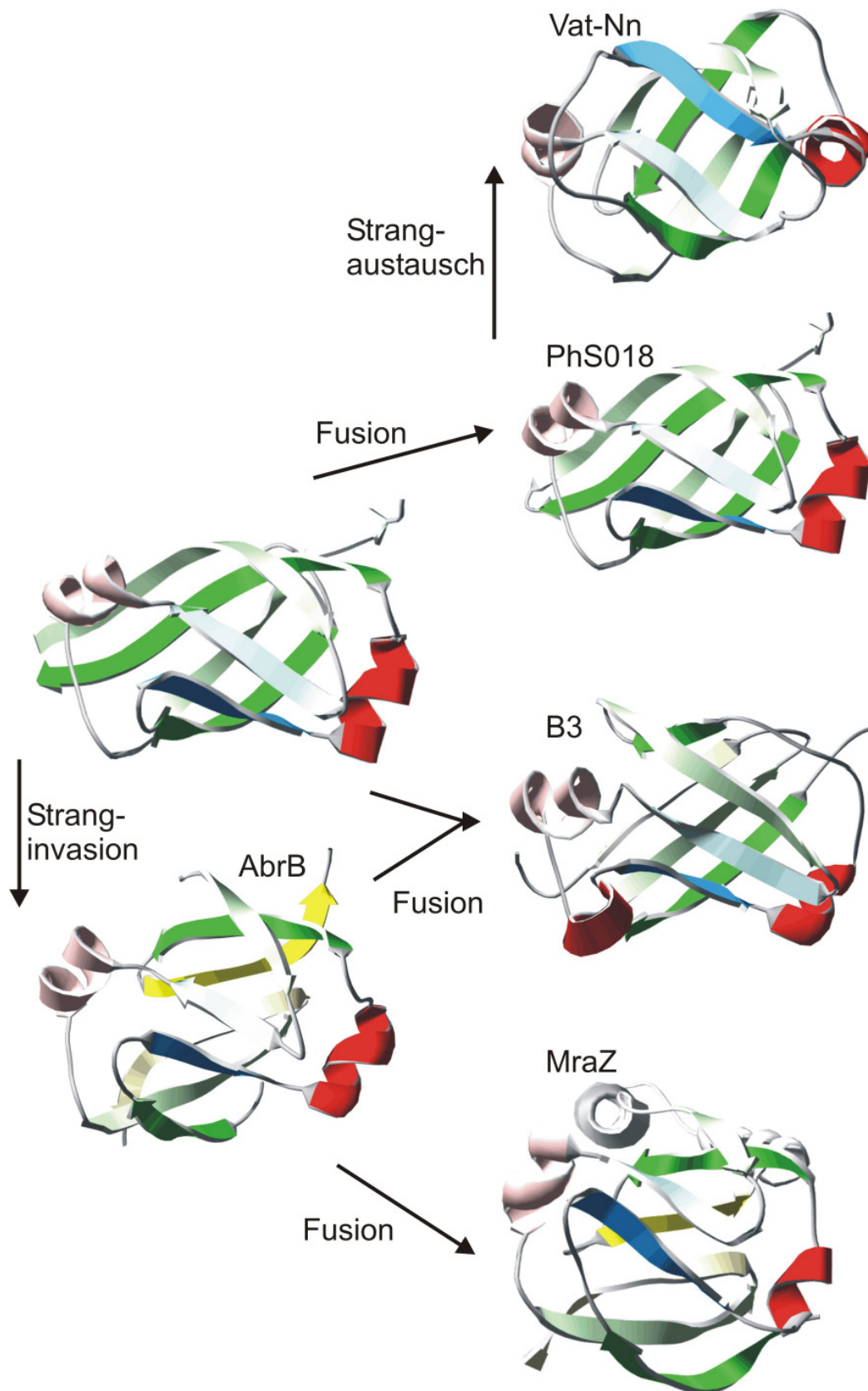
Die Eigenschaft von PhS018 DNA zu binden, deutet darauf, dass es wie AbrB ein Transkriptionsfaktor sein könnte. Eine Sequenz, an die PhS018 spezifisch bindet, konnte nicht gefunden werden.

Die Struktur von PhS018 erlaubt nun, einen evolutionären Zusammenhang zwischen Double psi beta-barrel und Swapped hairpin-barrel herzustellen (Abbildung 24). Aus der hohen Symmetrie von PhS018 kann ein dimerer Vorfahre von PhS018 plausibel abgeleitet werden, der durch Genduplikation und Genfusion zu PhS018 führte. Von PhS018 führt ein topologischer Strangaustausch (strand-swap) zu den Double psi beta-barrel Proteinen wie Vat-Nn. Einfügen eines  $\beta$ -Strangs (strand-invasion) in den PhS018 Vorfahren führt zu den

dimeren Swapped hairpin-barrel wie AbrB oder MazE. Dimerisierung führt zu den monomeren Swapped hairpin-barrel wie MraZ. Aus diesem Szenario ergibt sich eine potentielle Abstammung für ein weiteres ähnliches Protein: B3. B3 ist ein siebensträngiges *barrel*-Protein mit Cradle loop-barrel ähnlicher Struktur. Man kann es sich als Fusion aus dem PhS018 Vorgänger und einem monomeren Swapped hairpin-barrel vorstellen.

Das  $\beta\alpha\beta$ -Fragment könnte von einem ursprünglichen Peptid stammen (antecedent peptide, Lupas et al., 2001), welches als elementare Evolutionseinheit vorlag und dann durch Genfusion zu einer Domäne verschmolzen ist. Dieses Fragment findet sich nämlich mit hoher struktureller Ähnlichkeit in unterschiedlichen Umgebungen innerhalb der Cradle loop-barrel Proteinen wieder.

Af1504 wurde von HHpred mit hoher Wahrscheinlichkeit als homologes Protein der Double psi beta-barrel Proteine vorhergesagt und aufgrund dieser Homologie modelliert. Dies steht im Einklang mit dem *alignment* in Abbildung 14, das auf Ergebnissen des Programms Sensor beruht. Rosetta modellierte Strukturen ohne Ähnlichkeit zu bekannten Strukturen. Die Vorhersagen scheinen daher unrealistisch. Mögliche Schwierigkeiten bei dieser *ab initio* Modellierung könnten in der komplexen *psi-loop* Topologie oder in der Dimerisierung von Af1504 liegen. Es ist wahrscheinlich, dass Af1504 wie von HHpred vorhergesagt eine Double psi beta-barrel Domäne enthält, da die nächsten vier Sequenzverwandten Proteine Double psi beta-barrel Proteine sind. Die modellierte GD-Box in Af1504 lässt sich gut mit der GD-Box von PhS018 überlagern (RMSD 1,32 Å). Da die GD-Box innerhalb der Cradle loop-barrel stark konserviert ist, war eine solche auch in Af1504 zu erwarten. Die experimentellen CD- und Fluoreszenzdaten deuten insgesamt auf ein kompaktes, gefaltetes Protein mit  $\alpha$ -helikalen und  $\beta$ -Faltblatt Sekundärstrukturanteilen hin. Der  $\alpha$ -helikale Anteil ist wohl maßgeblich auf den *coiled coil* Bereich zurückzuführen und der  $\beta$ -Faltblattanteil auf die potentielle Double psi beta-barrel Domäne. Die vorhergesagte Domänenarchitektur, ein *coiled coil* Bereich vor einer Double psi beta-barrel Domäne, wird durch die Ergebnisse der limitierten Proteolyse unterstützt. Die erzeugten Banden von 14 und etwa 9 kD entsprechen den erwarteten Molekülmassen der vorhergesagten Domänen. Zwischen diesen Domänen befinden sich zwei Lysinreste (K118 und K120), die als Trypsinschnittstellen in Frage kommen. Die Abwesenheit von hydrophoben Resten zwischen den vorhergesagten Domänen erklärt die ausgebliebene Spaltung durch Chymotrypsin. Als wahrscheinliches Double psi beta-barrel Protein hat Af1504 für das entworfene Evolutionsszenario die gleiche Bedeutung wie Vat-Nn. Af1504 fusionierte mit einer *coiled coil* bildenden Struktur und zeigt damit eine neue Domänenkombination innerhalb der Double psi beta-barrel Proteinen



**Abbildung 24** Evolutionsszenario der Cradle loop-barrel. Links oben ist der potentielle dimere Vorfahre von PhS018 aus beiden Hälften der PhS018 Struktur abgebildet. Das zentrale  $\beta$ -Strangpaar  $\beta_2/\beta_5$  ist blau.  $\beta$ -Stränge, die exklusiv in Swapped hairpin-barrels vorkommen sind gelbfarbig.

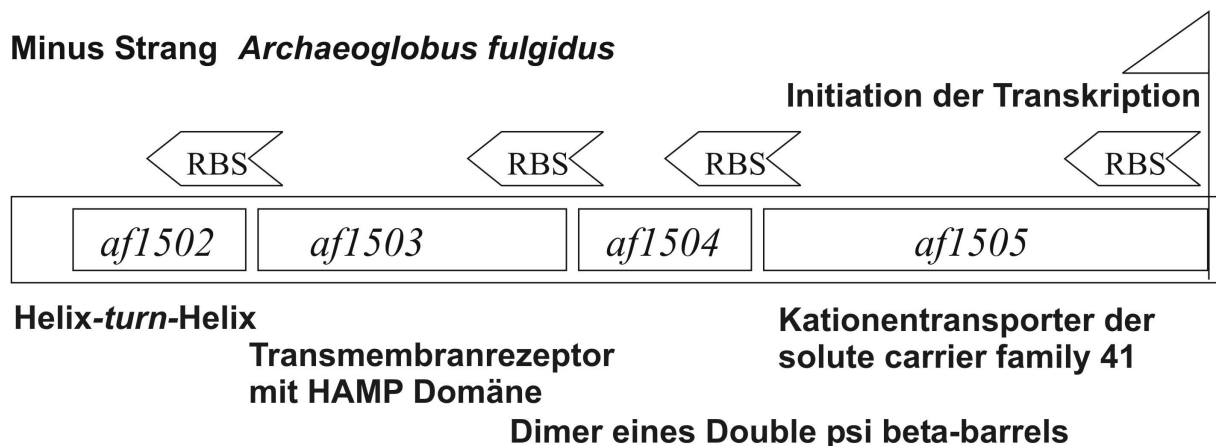
## 6 Prokaryontische Signaltransduktion

### 6.1 Einleitung

Durch Untersuchung des potentiellen Double psi beta-barrel Proteins Af1504 in Kapitel 5 wurde Aufmerksamkeit auf ein Operon in *Archaeoglobus fulgidus* gerichtet, welches neben dem *af1504*-Gen die Gene *af1502*, *af1503* und *af1505* enthält (Abbildung 25). Hans-Peter Klenk (e.gene Biotechnologie GmbH, Feldafing) identifizierte einen Promoter im idealen Abstand vor *af1505* mit einer Initiationssequenz der Transkription unmittelbar vor dem Startcodon der Translation. Zudem konnten Hinweise auf ribosomale Bindestellen vor den drei folgenden Genen gefunden werden. Somit sind die vier Gene *af1502*–*af1505* translational gekoppelt und eine funktionelle Interaktion der kodierten Proteine ist damit vorstellbar. Bis auf bioinformatische Vorhersagen war kein Protein dieser Gene charakterisiert worden.

Af1502 ist als kleines,  $\alpha$ -helikales Helix-turn-Helix Protein vorhergesagt. Eine Funktion ist nicht bekannt. Af1503 ist als Transmembranprotein mit einer extraplasmatische Domäne und einer cytoplasmatischen HAMP Domäne als einzige intrazelluläre Domäne vorhergesagt. Af1504 ist ein potentielles Double psi beta-barrel Protein (Kapitel 5) und Af1505 ist als Transmembranprotein der „solute carrier family 41“ vorhergesagt und wird in der Transport classification Database als  $Mg^{2+}$ -Transporter-E klassifiziert. HAMP Domänen kommen weit verbreitet in Prokaryonten als Teil von Signaltransduktionsproteinen vor. Damit liefert die HAMP Domäne eine erste wichtige Idee für die Funktion des Operons.

#### Minus Strang *Archaeoglobus fulgidus*



**Abbildung 25 Operonstruktur der *A. fulgidus* Gene *af1502*–*af1505*.** RBS steht für ribosomale Bindestellen. Unterhalb der Gensymbole stehen ihre vorhergesagten strukturellen oder funktionellen Eigenschaften.

Prokaryonten detektieren Stimuli in ihrer Umgebung durch modular aufgebaute Transmembranrezeptoren. Ligandenbindung findet meist extrazellulär statt und führt zur Verstärkung oder Abschwächung einer intrazellulären Phosphorylierungskaskade. Ein solches System erlaubt Bakterien einem chemischen Gradienten zu folgen. Dieses Verhalten wird als



Chemotaxis bezeichnet (Wadhams and Armitage, 2004). Dabei bindet der Ligand an die extrazelluläre Domäne eines Chemotaxisrezeptors. In *E. coli* gibt es Aspartat-, Serin-, Ribose- und Dipeptidrezeptoren. Zudem gibt es noch den Sauerstoffsensoren Aerotaxisrezeptor, der jedoch eine cytoplasmatische Rezeptordomäne mit noch unbekanntem Liganden besitzt. Die dimeren Rezeptoren wechselwirken über eine Phosphorylierungskaskade mit den Flagellenmotorproteinen, die das Bewegungsverhalten der Zellen durch ihre Drehrichtung beeinflussen. In Abwesenheit von chemischen Gradienten führen die Zellen einen „random walk“ aus, der aus alternierenden Schwimm- und Taumelbewegung besteht. In einem chemischen Gradienten kommt es zum „biased walk“, bei dem die Schwimm- gegenüber der Taumelbewegung erhöht wird, um schneller günstige Gebiete zu erreichen. Die Taumelbewegung dient stets der räumlichen Neuausrichtung der Zelle.

Neben chemotaktischen Bewegung in chemischen Gradienten gibt es viele weitere Stimuli mit ihren entsprechenden physiologischen Antworten. Beim Osmolaritätssensor EnvZ wird die Aktivität der Histidinkinasedomäne von EnvZ in Abhängigkeit der Osmolarität reguliert und das Signal führt über Aktivierung von Transkriptionsfaktoren zur differentiellen Expression von Membranporenproteinen.

Extrazelluläre und intrazelluläre Domänen sind oft durch HAMP Domänen unmittelbar nach der letzten Transmembranhelix miteinander verbunden. HAMP Domänen wurden ursprünglich von Tian Jin und Masayori Inouye als Linkersegment in Histidinkinasen und Chemotaxisrezeptoren entdeckt (Jin and Inouye, 1994) und später von L Aravind und Chris Ponting mit HAMP bezeichnet (Aravind and Ponting, 1999). Das Akronym HAMP stammt vom Erscheinen dieser Domäne in Histidinkinasen, Adenylatcyclasen, methylierten Chemotaxisproteinen und Phosphatasen. Derzeit sind über 4500 HAMP Domänen in der SMART Datenbank (Schultz et al., 1998) enthalten, wovon 4400 in Prokaryonten vorkommen und 100 in Archaeen. Über ein Fünftel aller Histidinkinasen und über zwei Drittel der Chemotaxisrezeptoren besitzen eine HAMP Domäne. Wenige HAMP Domänen werden auch in Pilzen, Pflanzen und Protisten gefunden. Ihr Ursprung ist wahrscheinlich durch lateralen Gentransfer aus endosymbiontischen Vorfahren zu erklären (Koretke et al., 2000).

Die Sequenzen der HAMP Domänen zeigen ein sich wiederholendes Muster aus sieben Aminosäuren, deren Positionen mit den Buchstaben *a-g* bezeichnet werden und Aminosäuren an den Positionen *a* und *d* hydrophobe Eigenschaften zeigen. Dieses Muster entspricht der „*heptad periodicity*“, die von Francis Crick als Markenzeichen von *coiled coil* Strukturen beschrieben wurde (Crick, 1953). Bioinformatische Analysen haben in den entsprechenden Regionen von Histidinkinasen potentielle *coiled coil* formende Sequenzen aufgezeigt (Singh et al., 1998). Reaktivitätsstudien mit Thiolgruppen von Cysteinresten, die gezielt in die HAMP Domäne des Aspartatrezeptors eingeführt wurden, konnten die amphipathische *coiled coil* Eigenschaften der HAMP Domänen experimentell bestätigen (Butler and Falke, 1998). Bis auf die Periodizität der hydrophoben Reste ist die Sequenz wenig konserviert und enthält keine invarianten Reste.

HAMP Domänen können zwischen verschiedenen Rezeptoren ausgetauscht werden, woraus funktionelle Hybride entstehen, die sich in ihrer Aktivität aber von den nativen Rezeptoren unterscheiden. Hybride wurden zwischen der Histidinkinase EnvZ und dem Chemotaxisrezeptor Tar konstruiert (Jin and Inouye, 1993), zwischen den Histidinkinasen NarX und CpxA (Appleman et al., 2003), sowie zwischen verschiedenen mycobakteriellen Adenylatcyclasen (Linder et al., 2004). Die Austauschbarkeit der HAMP Domänen lässt auf einen konservierten Signaltransduktionsmechanismus schließen. Meistens werden Signale *in cis* weitergegeben, d.h. innerhalb einer Polypeptidkette von der N-terminalen Rezeptordomäne zur C-terminalen Effektordomäne. Da es Rezeptoren gibt, die intrazellulär nur die HAMP Domäne besitzen, wird angenommen, dass ebenso ein *trans* Mechanismus möglich ist. Dieser beruht auf der regulierbaren Wechselwirkung der HAMP Domäne mit Adapterproteinen.

Die HAMP Domäne ist die letzte weit verbreitet vorkommende Domäne im Signaltransduktionssystem von Prokaryonten, deren Struktur noch nicht aufgeklärt wurde. Ihre Struktur ist aber besonders interessant, da sich die HAMP Domäne immer in einem konservierten Abstand direkt nach der letzten Transmembranhelix befindet und deshalb maßgeblich an der Übertragung von Signalen durch die Membran in die Zelle beteiligt sein muss. Für den molekularen Signalübertragungsmechanismus von der extrazellulären Rezeptordomäne durch die Zellmembran zur intrazellulären Effektordomäne und die molekulare Funktion der HAMP Domäne gibt es unterschiedliche Modellvorstellungen (Chervitz and Falke, 1996; Scott and Stoddard, 1994). Erstens gibt es ein Modell, welches die HAMP Domäne als starre, helikale Verlängerung der Transmembranhelix beschreibt, die translationale Bewegungen zwischen den Rezeptoruntereinheiten passiv weitergibt: „Pushrod-piston model“ (Chervitz and Falke, 1996). Zweitens stellt man sich vor, dass die HAMP Domäne je nach Signalzustand entweder eine *coiled coil* Struktur ausbildet oder als lose Helices an die Membran geheftet vorliegt: „Helix rearrangement model“ (Williams and Stewart, 1999). Das dritte Modell beschreibt die HAMP Domäne mit Ähnlichkeit zum Helix-Loop-Helix Transkriptionsfaktor „Max“, die zur Signalweiterleitung eine bisher noch unbekannt Konformationsänderung vollzieht (Koretke et al., 2003). Das vierte Modell beschreibt die Chemotaxisrezeptoren in einem Gleichgewicht zwischen einer dynamischen und einer eingefrorenen Konformation, die durch Rotation der beiden Rezeptoruntereinheiten um die gemeinsame Rezeptorachse ineinander übergehen: „Frozen dynamic dimer model“ (Kim, 1994).

In dieser Arbeit wurde die HAMP Domäne von Af1503 für die Strukturanalyse mit NMR-Spektroskopie ausgewählt. Sie schien aus praktischen Gründen besonders für die Strukturaufklärung geeignet, da sie von dem hyperthermophilen Organismus *A. fulgidus* stammt und hohe Stabilität versprach. Ferner ist sie als einzige intrazelluläre Domäne von Af1503, bis auf die N-terminale liegende Transmembranhelix, in keinen weiteren Domänenkontext eingebettet und sollte für sich allein stabil sein. Aus der Struktur der HAMP

Domäne wurde für sie ein molekularer Signaltransduktionsmechanismus abgeleitet. Mit dem Ziel, diesen Mechanismus zu belegen, wurden mehrere Mutationen und Fusionsproteine der HAMP Domäne hergestellt. Damit wurden dann weitergehende funktionelle und strukturelle Untersuchungen durchgeführt. Um eine Vorstellung von der Funktion des Transmembranproteins Af1503 zu erhalten, wurde nach einem Liganden von Af1503 gesucht. Die Bindung eines Liganden an die extraplasmatische Domäne schien wahrscheinlich, da Af1503 ein Rezeptor eines Signaltransduktionssystems sein könnte. Johannes Söding (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) analysierte die Sequenz von Af1503 mit dem Suchprogramm für homologe Sequenzen HHsenser und fand in der extraplasmatischen Domäne von Af1503 ein kurzes Motiv, das EF-Hand Motiven in Calciumbindenden Proteinen ähnelte. Deshalb wurde in dieser Arbeit die Calciumbindefähigkeit von Af1503 experimentell mit kalorimetrischer Titration untersucht.

## 6.2 Materialien und Methoden

**Herstellung von Af1502** Af1502 wurde als C-terminales Fusionsprotein zu GST hergestellt. Das Gensegment wurde mit PCR aus *A. fulgidus* DSM 4303 genomischer DNA (LGC Promochem) vervielfältigt und isoliert. Über Primer wurden die Restriktionsschnittstellen BamHI und XhoI eingeführt, über die das verdaute PCR-Produkt in das Expressionsplasmid pGex4T-1 (GE Healthcare) ligiert wurde. Für die Expression wurde das Plasmid in *E. coli* BI21 (DE3) Gold Zellen transformiert. Die Zellen wuchsen bei 37° C bis zu einer OD von ~ 0,6, wurden mit 1 mM IPTG induziert und nach vier Stunden geerntet. Nach Resuspendierung in PBS wurden die Zellen lysiert. Die lösliche Fraktion wurde auf eine GSH-Säule (GSH-Trap FF, 5 ml, GE Healthcare) aufgetragen, mit 20 Säulenvolumen PBS gewaschen und mit 10 mM Glutathion<sub>red</sub> in 50 mM Tris, pH 8 eluiert. Das Eluat wurde mit Thrombin behandelt (2 U/mg Fusionsprotein), um GST von Af1502 zu spalten. Thrombin wurde mit 0,1 mM PMSF inhibiert und die Proteinlösung wurde auf eine Gelfiltrationssäule (Superdex G75 26/60, GE Healthcare) aufgetragen und mit PBS eluiert.

**Herstellung von Af HAMP** Die HAMP Domäne von Af1503 (S278-E331) wurde als C-terminales Fusionsprotein zu GST hergestellt. Das Gensegment der HAMP Domäne wurde mit PCR aus *A. fulgidus* DSM 4303 genomischer DNA (LGC Promochem) vervielfältigt und isoliert. Über Primer wurden Restriktionsschnittstellen BamHI und XhoI eingeführt, über die das verdaute PCR-Produkt in das Expressionsplasmid pGex4T-1 (GE Healthcare) ligiert wurde. Für die Expression wurde das Plasmid in *E. coli* BI21 (DE3) Gold Zellen transformiert. Die Zellen wuchsen bei 37° C bis zu einer OD von ~ 0,6, wurden mit 1 mM IPTG induziert und nach vier Stunden geerntet. Nach Resuspendierung in PBS wurden die Zellen lysiert. Die lösliche Fraktion wurde auf eine GSH-Säule (GSH-Trap FF, 5 ml, GE Healthcare) aufgetragen, mit 20 Säulenvolumen PBS gewaschen und mit 10 mM Glutathion<sub>red</sub>

in 50 mM Tris, pH 8 eluiert. Das Eluat wurde mit Thrombin behandelt (2 U/mg Fusionsprotein), um GST von Af HAMP zu spalten. Die behandelte Lösung wurde nach Zugabe von 0,1 mM PMSF auf eine Gelfiltrationssäule (Superdex G75 26/60, GE Healthcare) aufgetragen und mit PBS eluiert. Gereinigtes Af HAMP wurde bis zu 6 mg/ml konzentriert und für NMR Analyse eingesetzt. Für  $^{15}\text{N}$ - und  $^{15}\text{N}/^{13}\text{C}$ -isotopenmarkierte Proben wurden Zellen in M9 Minimalmedium kultiviert.  $^{15}\text{NH}_4\text{Cl}$  und Glucose bzw.  $\text{U-}^{13}\text{C}$ -Glucose dienten als einzige Stickstoff bzw. Kohlenstoffquellen. Die Expression und Reinigung entsprach exakt dem Vorgehen bei der unmarkierten Probe. Um gemischte Proben aus  $^{15}\text{N}/^{13}\text{C}$ -markiertem Protein und unmarkiertem Protein herzustellen, wurden beide Proteine getrennt gereinigt, dann gleiche Proteinmengen gemischt, Guanidiniumchlorid (6 M Endkonzentration) zugegeben und zweifach gegen PBS dialysiert. Die Proteinmenge wurde mit einem Bicinchoninsäureassay und die Reinheit mit SDS-PAGE ermittelt.

**Herstellung der Af HAMP Mutanten** In die HAMP Domäne von Af1503 wurden Punktmutationen A291V, A291C, A291F und A291I eingeführt. Die Position 291 bezieht sich auf die Sequenz von Af1503. Af HAMP A291V, A291C, A291I und A291F wurden mit Primern, welche die entsprechende Mutation enthielten und mit Overlap-Extension PCR hergestellt. Die Reinigung der Proteine erfolgte wie die Reinigung der wildtyp HAMP Domäne.

**Strukturbestimmung der HAMP Domäne** Die Struktur von Af HAMP wurde mit NMR-Spektroskopie bestimmt. Aufnahme der Spektren, sowie deren Auswertung und die Strukturberechnung wurden von Murray Coles (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) durchgeführt (pdbid 2ASW). Ebenso bearbeitete Murray Coles das Protein Af HAMP A291V mit NMR-Spektroskopie.

**Drehung der Koordinaten der Af HAMP Domäne** Eine im Computer modellierte, gedrehte Struktur von Af HAMP wurde von Markus Gruber (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) erstellt.

**Herstellung von Af HAMP-DHP und Af HAMP A291V-DHP** Die Chimären bestehen aus der HAMP Domäne (S278-L326) von Af1503 von *A. fulgidus* und der DHP Domäne (K228-R289) des Osmolaritätssensors EnvZ von *E. coli*. Erst wurden von jedem Gensegment PCR-Produkte mit entsprechenden Überhängen für die Überlappung hergestellt. Mit Overlap-Extension PCR wurden beide Gensegmente verknüpft und amplifiziert. Überhängende terminale Primer führten NdeI und XhoI Restriktionsschnittstellen ein, mit denen die PCR Produkte in pET30b-Expressionsplasmid ligiert wurden. Für die Mutation diente Af HAMP A291V als Templat. Beide Proteine wurden auf die gleiche Weise gereinigt. Zellen wurden bei 37° C kultiviert und bei OD ~ 0,6 mit 1 mM IPTG induziert. Nach vier Stunden wurden

die Zellen geerntet, in 30 mM MOPS, pH 7, 50 mM NaCl resuspendiert und lysiert. Mit dem löslichen Protein im Überstand wurde mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 30 mM MOPS, pH 7 und einem linearen NaCl-Gradienten von 50-300 mM NaCl über 10 Säulenvolumen und von 300-525 mM NaCl über 5 Säulenvolumen durchgeführt. Fraktionen mit Zielprotein wurden vereinigt und mit 35% Ammoniumsulfat (Endkonzentration) versetzt. Nach Zentrifugation wurde der Niederschlag vom Überstand getrennt, in 20 mM MOPS, pH 7, 100 mM NaCl gelöst und auf eine Gelfiltrationssäule (Superdex G75 26/60, GE Healthcare) aufgetragen. Das Protein wurde mit 20 mM MOPS, pH 7, 100 mM NaCl eluiert und bis 60 mg/ml konzentriert. Die Proteinmenge wurde über Absorption bei 280 nm bestimmt und die Reinheit wurde mit SDS-PAGE nachgewiesen. Die Proben wurden für Kristallisationsversuche eingesetzt. Für die Markierung der Proteine mit Selenomethionin wurde das Plasmid in B834 (DE3)-Zellen transformiert, die in M9 Minimalmedium kultiviert wurden. Das Minimalmedium wurde mit jeder der biogenen Aminosäuren (0,1 mg/l) ergänzt. Statt Methionin wurde Selenomethionin verwendet. Die Expression und Reinigung wurden analog zu den unmarkierten Proben durchgeführt.

**Herstellung von Af HAMP- und Af HAMP A291V-EnvZ / cTsr / NarX** Die chimären Proteine bestehen aus der HAMP Domäne (S278-L326) von Af1503 und der Kinasedomäne von *E. coli* EnvZ- (K228-G450), der cytoplasmatischen Domäne des *E. coli* Serinrezeptors Tsr (M264-F554) oder der Kinasedomäne des Nitratrezeptors NarX (A224-E598) von *E. coli*. Erst wurden von jedem Gensegment PCR-Produkte mit entsprechenden Überhängen für die Überlappung hergestellt. Mit Overlap-Extension PCR wurden je zwei Gensegmente verknüpft und amplifiziert. Überhängende terminale Primer führten NdeI und XhoI Restriktionsschnittstellen ein, mit denen die PCR Produkte in pET30b-Expressionsplasmid ligiert wurden. Für die Mutation diente Af HAMP A291V als Templat. Die Reinigung aller Proteine verlief so wie die Reinigung von Af HAMP-DHP. Der einzige Unterschied bestand in der Verwendung einer anderen Gelfiltrationssäule: Superdex S200 26/60 statt Superdex G75 26/60.

**Strukturbestimmung von Af HAMP-DHP** Verschiedene Bedingungen zur Kristallisation des Proteins wurden von Ines Wanke (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) getestet. Röntgenbeugung von Kristallen, sowie anschließende Bestimmung der Proteinstruktur aus Elektronendichtekarten wurden von Kornelius Zeth (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) durchgeführt.

**Herstellung von Af1503 TM2CC** Af1503 besitzt zwei hydrophobe Transmembranbereiche. Für die Expression des Af1503-Rezeptors wurde der erste Transmembranbereich ausgespart und der zweite durch eine *coiled coil* bildende Sequenz ersetzt. Die genaue Sequenz ist im Anhang nachzulesen. Es wurden zwei PCR-Fragmente des *af1503*-Gens hergestellt. Das erste

Fragment umfasste den Bereich vor dem zweiten Transmembranbereich (T31-Q253) und das zweite Fragment den Bereich danach (T276-K338). Mit Primerüberhängen wurde eine Sequenz eingefügt, die für eine *coiled coil* bildende Aminosäuresequenz kodierte. Beide PCR-Fragmente wurden über ihre überlappenden Enden aneinandergesetzt, verlängert und durch PCR amplifiziert. Über terminale Primer wurden Restriktionsschnittstellen für NdeI und HindIII eingefügt. Die beiden Mutationen wurden mit Primern und Overlap-Extension PCR eingeführt. Das verdaute PCR-Produkt wurde in den pET30b-Expressionsvektor (Novagen) ligiert und für die Expression in *E. coli* BL21 (DE3) Gold Zellen transformiert. Die Expression wurde bei 37° C durchgeführt und induziert wurde mit 1 mM IPTG bei OD ~ 0,6. Die Zellen wurden vier Stunden nach Induktion geerntet, in 30 mM MOPS, pH 7, 50 mM NaCl suspendiert und lysiert. Lösliches Af1503 TM2CC wurde mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 30 mM MOPS, pH 7 mit linear ansteigendem NaCl-Gradienten über 10 Säulenvolumen von 50-525 mM NaCl gereinigt. Fraktionen mit Af1503 TM2CC wurden vereinigt und auf Eis mit 30% Ammoniumsulfat (Endkonzentration) versetzt. Der Niederschlag wurde durch Zentrifugation vom Überstand getrennt, in 20 mM MOPS, pH 7, 100 mM NaCl, 10 mM EDTA gelöst und anschließend mit Gelfiltration (Superdex S200 26/60, GE Healthcare) in 20 mM MOPS, pH 7, 100 mM NaCl gereinigt. Die Proteinmenge wurde über Absorption bei 280 nm bestimmt.

**Herstellung von Af1503 TM2CC D88A und D92A** Die Mutationen D88A und D92A wurden mit entsprechenden Primern und Overlap-Extension PCR in das Templat Af1503 TM2CC eingeführt. Das PCR Produkt wurde über terminale NdeI und HindIII Restriktionsschnittstellen in den pET30b-Expressionsvektor ligiert. Die Expressionsbedingungen waren die gleichen wie für Af1503 TM2CC und führten zu unlöslichen *inclusion bodies*. In diesem Fall wurden gröbere Verunreinigungen von den unlöslichen *inclusion bodies* durch mehrfaches suspendieren in 50 mM Tris pH 8, 200 mM NaCl und 1% Triton, zentrifugieren und anschließend dekantieren abgetrennt. Die *inclusion bodies* wurden in 50 mM Tris pH 7,5, 50 mM NaCl, 10 mM DTT und 8 M Harnstoff aufgelöst und mit Anionenaustauschchromatographie (QHP, 21 ml, GE Healthcare) in 50 mM Tris pH 7,5 und 8 M Harnstoff unter denaturierenden Bedingungen (6 M Harnstoff) und mit linear steigendem NaCl-Gradienten über 10 Säulenvolumen von 50-525 mM NaCl zur Elution gereinigt. Fraktionen mit Af1503 TM2CC D88A bzw. D92A wurden vereinigt (etwa 1 mg/ml) und durch zweimalige Dialyse gegen 20 mM MOPS, pH 7, 100 mM NaCl rückgefaltet. Endgültige Reinheit der Proteine wurde durch Ammoniumsulfatfällung und Gelfiltration analog zur Reinigung von Af1503 TM2CC erreicht. Die Proteinmenge wurde über Absorption bei 280 nm bestimmt und die Reinheit wurde mit SDS-PAGE nachgewiesen.

**Herstellung von Af1503 EP** Af1503 EP bezeichnet die extraplasmatische Domäne von Af1503 (T31-Q253). Das Gensegment wurde mit PCR amplifiziert und mit den

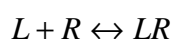
Restriktionsschnittstellen NdeI und HindIII versehen, über die das PCR Produkt in den pET30b-Expressionsvektor ligiert wurde. Die Reinigung verlief analog zur Reinigung von Af1503 TM2CC. Zur Gelfiltration wurde eine Superdex G75 26/60 statt einer Superdex S200 26/60 Gelfiltrationssäule eingesetzt. Die Proteinmenge wurde über Absorption bei 280 nm bestimmt und die Reinheit wurde mit SDS-PAGE nachgewiesen.

**CD und Fluoreszenz-Spektroskopie** CD-Spektren wurden mit einem Jasco CD-Spektropolarimeter J-810 gemessen. Das Gerät ist für Schmelzpunktbestimmungen mit einer temperierbaren Küvettenhalterung ausgestattet. Fluoreszenzspektren wurden mit einem Jasco FP-6500 Spektrofluorometer gemessen. Die Anregungswellenlänge für Tryptophanfluoreszenz betrug 293 nm.

**Isothermale kalorimetrische Titration** Für kalorimetrische Messungen wurde ein Microcal VP-ITC Mikrokalorimeter verwendet. Die Arbeitstemperatur betrug 30° C. Die Proteinkonzentration in der Zelle betrug 80 µM und die Ligandenkonzentration 2 mM. 3,8 µl Ligandenlösung wurden 30 mal zugegeben. In diesem Gerät war das Messzellvolumen 1,4295 ml. Nach der Datenaufnahme wurden die von der Software des Geräteherstellers automatisch angenäherten Integrationsgrenzen manuell für jede Injektion verfeinert. Eine Titration ohne Protein wurde als Leerwert verwendet und von den Messdaten mit Protein abgezogen. Mit der gleichen Software wurde eine sigmoidale Funktion an die Messdaten angelegt und daraus die Assoziationskonstante, die Bindungsstöchiometrie, die Reaktionsenthalpie und die Reaktionsentropie bestimmt. Im folgenden Absatz soll erläutert werden, wie sich die sigmoidale Bindungskurve erklären lässt und wie die Ergebnisparameter daraus bestimmt werden.

Durch Injektion eines Liganden zu einem Rezeptor entsteht bei exothermen Reaktionen Wärme oder es wird bei endothermen Reaktionen Wärme benötigt. Dies verursacht eine Temperaturänderung in der Messzelle gegenüber der mit Wasser gefüllten Referenzzelle des ITC-Gerätes. Die benötigte Energiemenge, um die Temperaturdifferenz durch Wärmeänderung nach Injektion des Liganden zu kompensieren, liefert das Messsignal. Eine adiabatische Isolierung der Zellen verhindert den Wärmeaustausch mit der Umgebung. Die Kompensationsenergie wird während des Experimentes aufgenommen, in der Auswertung für jede Injektion integriert und auf die Stoffmenge des Liganden normiert. Den theoretischen Hintergrund liefern das Massenwirkungsgesetz und die Gesetze der Thermodynamik:

Für die reversible Reaktion des Liganden L mit dem Rezeptor R zum Ligand-Rezeptor-Komplex LR gilt:



(Gleichung 1)

und aus dem Massenwirkungsgesetz folgt die Assoziationskonstante:

$$K_a = \frac{[LR]}{[L] \cdot [R]} \quad (\text{Gleichung 2})$$

mit den Gleichgewichtskonzentrationen [LR], [L] und [R].

Die Gesamtkonzentration des Liganden ist:

$$[L]_{tot} = [L] + [LR] \quad \text{und umgeformt } [L] = [L]_{tot} - [LR] \quad (\text{Gleichung 3})$$

und die Gesamtkonzentration des Rezeptors ist:

$$[R]_{tot} = [LR] + [R] \quad \text{mit Hilfe von Gleichung 2: } [R]_{tot} = [LR] + \frac{[LR]}{K_a \cdot [L]} \quad (\text{Gleichung 4})$$

Einsetzen von Gleichung 3 in Gleichung 4 ergibt:

$$[R]_{tot} = [LR] + \frac{[LR]}{K_a \cdot ([L]_{tot} - [LR])} \quad \text{und umgeformt}$$

$$[LR]^2 + [LR] \cdot \left( -[R]_{tot} - [L]_{tot} - \frac{1}{K_a} \right) + [R]_{tot} \cdot [L]_{tot} = 0 \quad (\text{Gleichung 5})$$

Gleichung 5 ist eine in [LR] quadratische Gleichung mit der Lösung:

$$[LR] = \frac{-b - \sqrt{b^2 - 4a}}{2} \quad (\text{Gleichung 6})$$

wobei sich a und b aus Gleichung 5 ergeben als

$$b = -[R]_{tot} - [L]_{tot} - \frac{1}{K_a} \quad (\text{Gleichung 7})$$

$$a = [R]_{tot} [L]_{tot} \quad (\text{Gleichung 8})$$

Nun wird die Ableitung von [LR] nach  $[L]_{tot}$  aus der Gleichung 6 gebildet, welche die Änderung der Liganden-Rezeptor-Konzentration bei steigender Liganden-Gesamtkonzentration angibt:



$$\frac{d[LR]}{d[L]_{tot}} = \frac{1}{2} + \frac{1 - \frac{1+u}{2} - \frac{w}{2}}{\sqrt{w^2 - 2 \cdot w \cdot (1-u) + (1+u)^2}} \quad (\text{Gleichung 9})$$

$$\text{mit } u = \frac{1}{K_a \cdot [R]_{tot}} \quad (\text{Gleichung 10})$$

$$\text{und } w = \frac{[L]_{tot}}{[R]_{tot}} \quad (\text{Gleichung 11})$$

Die gemessene Wärme ist proportional zur Bildung der Liganden-Rezeptor-Komplexes. Proportionalitätskonstanten sind dabei das konstante Zellvolumen  $V$  und die konstante molare Reaktionsenthalpie  $\Delta H^0$ :

$$dQ = V \cdot \Delta H^0 \cdot d[LR] \quad (\text{Gleichung 12})$$

Gleichung 12 eingesetzt in Gleichung 9 ergibt:

$$\frac{dQ}{V \cdot d[L]_{tot}} = \Delta H^0 \left( \frac{1}{2} + \frac{1 - \frac{1+u}{2} - \frac{w}{2}}{\sqrt{w^2 - 2 \cdot w \cdot (1-u) + (1+u)^2}} \right) \quad (\text{Gleichung 13})$$

die differenzielle Wärme  $\frac{dQ}{d[L]_{tot}}$  ist der experimentell bestimmte Parameter

Aus Gleichungen 11, 12 und 13 erkennt man die Abhängigkeiten der Titrationskurve hinsichtlich der Assoziationskonstante, Konzentrationen der Reaktionspartner und molarer Enthalpie.

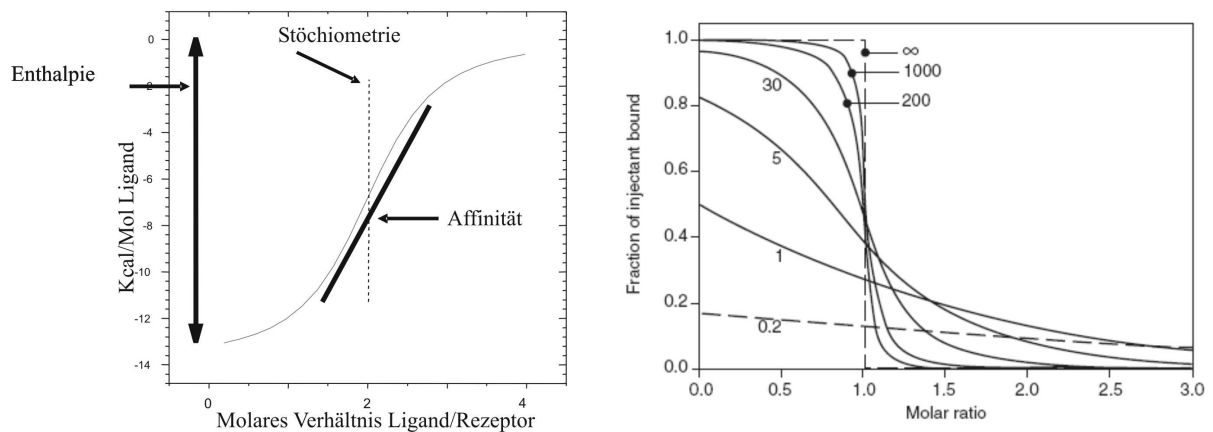
Wenn die Assoziationskonstante gegen  $\infty$  geht, dann geht  $u$  gegen Null und die gemessene Wärme ist nur noch abhängig von der injizierten Ligandenmenge (Gleichung 11). Jeder Ligand wird gebunden, solange noch genügend freier Rezeptor vorhanden ist. Ist der Äquivalenzpunkt der Titration erreicht, findet keine Bindung und damit keine Wärmeänderung mehr statt. Es resultiert eine stufenförmige Titrationskurve mit abruptem Übergang am Äquivalenzpunkt (Abbildung 26).

Geht die Assoziationskonstante gegen Null, dann ergibt sich eine fast lineare Funktion, in der die Ligandenkonzentration fast keinen Einfluss mehr hat. Es findet dann kaum Bindung und kaum Wärmeänderung statt (Abbildung 26).

Bei mittleren Assoziationskonstanten beschreibt die Funktion eine sigmoidale Kurve mit dem Übergang am Äquivalenzpunkt. Es hängt vom Verhältnis von der Assoziationskonstante zur Rezeptorkonzentration ab, wie stark die S-Form ausgeprägt ist. Um dieses Verhältnis zu beschreiben, wurde der Parameter  $c$  eingeführt, der reziprok zu  $u$  ist:

$$c = \frac{1}{u} = K_a \cdot [R]_{tot}$$

Gut auswertbare Kurven erhält man für  $10 < c < 100$ . (Abbildung 26) Dann hat die Kurve nämlich drei ausgeprägte Bereiche: Ein Plateau vor dem Äquivalenzpunkt, den Übergang durch den Äquivalenzpunkt und ein Plateau nach dem Äquivalenzpunkt. Durch Wahl der Ligandenkonzentration und der Injektionsvolumina erreicht man eine ausreichende Anzahl von Messpunkten in diesen drei Bereichen. Bei sehr hohen Assoziationskonstanten ist eine geringe Rezeptorkonzentration nötig. Sie muss aber hoch genug sein, um so viel Wärme zu erzeugen, dass die Empfindlichkeitsgrenze des Gerätes nicht unterschritten wird. Bei sehr niedrigen Assoziationskonstanten braucht man für ein ausgeprägtes Anfangsplateau hohe Rezeptorkonzentrationen. Bei Proteinen als Rezeptoren muss man Löslichkeitsgrenzen oder Grenzen von herstellbaren Proteinmengen beachten.



**Abbildung 26 Illustration von ITC-Kurven.** (Links) Parameter, die aus einem ITC-Experiment und der resultierenden Kurve abgelesen werden können. (rechts; aus Current Protocols in Protein Science) Einfluss des Verhältnisses zwischen Assoziationskonstante und Rezeptorkonzentration. Das Verhältnis ist durch den  $c$ -Wert ausgedrückt, der in der Abbildung an die jeweilige Kurve geschrieben wurde.

Durch die Messpunkte kann eine sigmoidale Kurve angepasst werden und man erhält daraus fünf Informationen. (Abbildung 26). Die Lage des Äquivalenzpunktes bestimmt die Stöchiometrie der Reaktion. Aus dem Höhenunterschied der Plateaus ergibt sich die

Reaktionsenthalpie  $\Delta H$ . Die Assoziationskonstante  $K_a$  ergibt sich aus der Kurvenanpassung und Gleichung 13. Aus der Assoziationskonstante und dem Massenwirkungsgesetz ergibt sich die freie Reaktionsenthalpie  $\Delta G$  und damit die Reaktionsentropie  $\Delta S$ .

**Analytische Gelfiltration** Zur Bestimmung der Oligomerisierung der Proteine Af1503 EP und Af1503 TM2CC sowie dessen Mutationen D88A und D92A wurde Gelfiltration (Superdex S200 Global 10/300, GE Healthcare) in 20 mM MOPS, pH 7, 100 mM NaCl durchgeführt. Zur Untersuchung des Einflusses von Calcium wurde in diesen Fällen der Puffer um 2 mM  $\text{CaCl}_2$  ergänzt. Das Elutionsprofil wurde über UV-Absorption bei 280 nm aufgezeichnet.

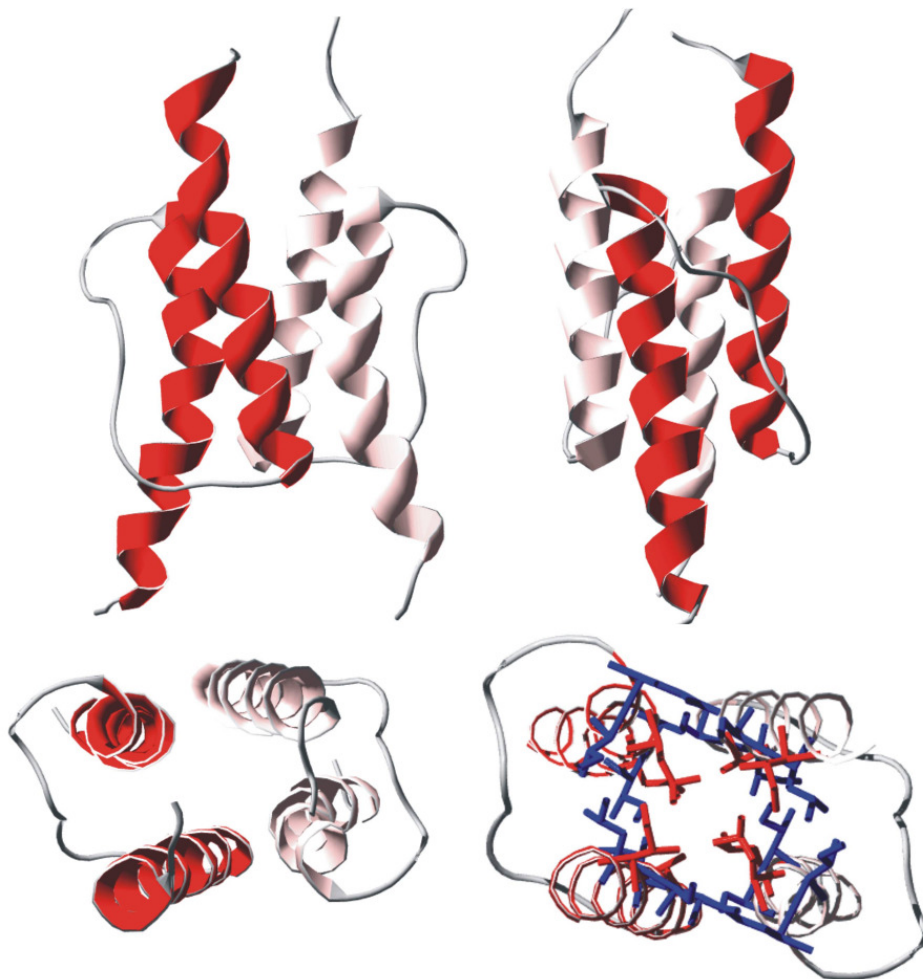
**E. coli Stämme für Chemotaxisassays** CP362 ( $\Delta(\text{tsr})$ -7028  $\Delta(\text{tar-tap})$  5201  $\Delta(\text{trg})$ -100 his leu rpsL ara lac thi tonA tsx xyl zbd:Tn5) sind *E. coli* Zellen, denen alle Chemotaxisrezeptoren fehlen. Das Plasmid pNT201 trägt eine Ampicillinresistenz und den *E. coli* Aspartatrezeptor Tar nach einem tac-Promoter. Beides wurde von Gerald Hazelbauer (University of Missouri, USA) zur Verfügung gestellt.

**Chemotaxisassays** Für Chemotaxisexperimente wurden chimäre Aspartatrezeptoren aus *E. coli* Tar und der HAMP Domäne von Af1503 hergestellt. Dafür wurden PCR-Fragmente des *tar*-Gens, welche die Proteinfragmente  $\text{tar}_{1-215}$  und  $\text{tar}_{262-630}$  kodieren, durch Overlap-Extension PCR mit einem PCR-Fragment verknüpft, das für A291C, A291V oder wildtyp Af1503<sub>280-326</sub> kodiert. Das fertige PCR-Produkt wurde mit XbaI und KpnI verdaut und in pNT201 ligiert. Das ergab drei Expressionsplasmide unter Kontrolle des tac-Promoters: Tar-AfHAMP (kodiert für  $\text{Tar}_{1-215}\text{Af1503}(\text{Wildtyp})_{280-326}\text{Tar}_{262-630}$ ), Tar-AfHAMP<sub>A291C</sub> (kodiert für  $\text{Tar}_{1-215}\text{Af1503}(\text{A291C})_{280-326}\text{Tar}_{262-630}$ ) und Tar-AfHAMP<sub>A291V</sub> (kodiert für  $\text{Tar}_{1-215}\text{Af1503}(\text{A291V})_{280-326}\text{Tar}_{262-630}$ ). Chemotaxisassays wurden als „swarm-plate Assays“ durchgeführt. Das Medium für die swarm-plates bestand aus 0,3 % Agar, 10g/l Trypton, 5 g/l NaCl und sofern nötig 100  $\mu\text{g/ml}$  Ampicillin und 20  $\mu\text{M}$  IPTG. In die Mitte jeder Platte wurden 5  $\mu\text{l}$  einer Flüssigkultur in LB von CP362 Zellen in logarithmischer Wachstumsphase pipettiert. Verschiedene Ansätze enthielten unterschiedliche Plasmidvarianten in den Zellen. Alle Platten wurden bei 30° C und 24 Stunden lang inkubiert und nach definierten Zeitabständen fotografiert, um die radiale Ausbreitungsgeschwindigkeit zu bestimmen. Um Bewegungsphänotypen und die Taumelfrequenz zu bestimmen, wurden die Zellen lichtmikroskopisch betrachtet. Dafür wurden dicht gewachsene Flüssigkulturen von CP362 Zellen 50-fach in LB-Medium verdünnt und nach 1,5 Stunden mit 0,1 mM IPTG induziert und drei weitere Stunden inkubiert. Um die Zellen in die logarithmische Wachstumsphase zu bringen, wurden die Zellen nochmals zehnfach in LB-Medium mit 0,1 mM IPTG verdünnt. Nach genau einer Stunde wurden sie für die mikroskopische Inspektion fünffach in PBS verdünnt und direkt mit 40-facher Vergrößerung und Phasenkontrast betrachtet.

**Adenylatcyclaseassays** Fusionsproteine, bei denen die HAMP Domäne N-terminal an die Adenylatcyclase Rv3645 von *Mycobacterium tuberculosis* fusioniert wurde und Adenylatcyclaseassays wurden von Franziska Berndt, Jürgen Linder und Professor Joachim Schultz am Pharmazeutischen Institut der Universität Tübingen durchgeführt. Weitere Einzelheiten können in (Linder et al., 2004) nachgelesen werden.

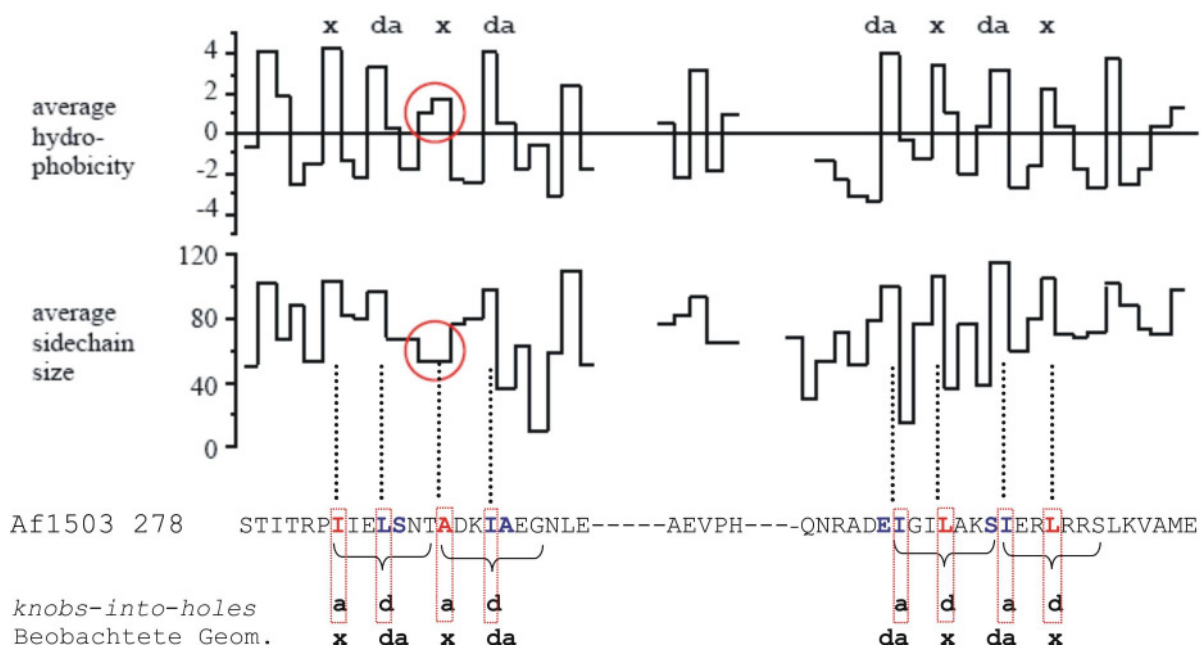
### 6.3 Ergebnisse

**Herstellung und Struktur der HAMP Domäne** Das Protein Af HAMP konnte als lösliches GST-Fusionsprotein exprimiert und gereinigt werden. Die HAMP Domäne konnte durch Thrombin von GST gespalten und isoliert werden. Die Struktur der HAMP Domäne wurde mit NMR-Spektroskopie ermittelt und ergab einen Satz von 20 Strukturen mit einem RMSD von 0,17 Å für alle Atome des Peptidrückgrats und 0,59 Å für alle nicht-Protonen Atome. Die Strukturen zeigen ein dimeres, paralleles Vierhelixbündel (Abbildung 27). Die Helices jedes Monomers ( $\alpha 1$  und  $\alpha 2$ ) sind gleich lang und um eine Umdrehung entlang der Helixachse versetzt. Beide Helices sind über einen rechtshändigen, 13 Aminosäure langen (G297-A309) *loop* miteinander verbunden.



**Abbildung 27 Struktur der HAMP Domäne.** Die oberen Bilder unterscheiden sich voneinander durch Drehung um die vertikale Achse. N-Termini befinden sich oben. Das Bild unten links geht aus dem Bild oben links durch Drehung um die horizontale Achse hervor. Unten rechts sind die Reste, die an der Kernverpackung teilnehmen farblich eingezeichnet. Rote Reste zeigen genau zur Achse des Helixbündels ( $x$ -Positionen), blaue Reste umrahmen den Kern ( $da$ -Positionen). N-Termini sind dem Betrachter zugewandt.

Die Topologie ähnelt der von basischen Helix-Loop-Helix Transkriptionsfaktoren, deren Helices jedoch durch einen linkshändigen *loop* miteinander verbunden sind. Die 13 Reste des *loops* sind etwas länger als die minimal nötige Anzahl von Resten, um die Entfernung der Helixenden von 26 Å zu überbrücken. Daher nimmt der *loop* eine fast gestreckte Konformation ein, von der nur Q306 abweicht und eine „ $\beta$ -bulge“ Konformation einnimmt. Die kürzesten *loops* in HAMP Domänen sind 12 Reste lang und dürften damit vollkommen gestreckt sein. Am Ende der ersten Helix befindet sich ein *turn*, der durch das fast invariante G297 möglich wird. Der *loop* mündet an seinem C-terminalen Ende in ein konserviertes D<sub>310</sub>-E-X-G<sub>313</sub>-Motiv, das als „capping“-Sequenz für  $\alpha$ 2 und damit als Markierung für den Beginn von  $\alpha$ 2 dient. Die Sequenz zeigt ein sich wiederholendes, sieben Aminosäuren langes, hydrophobes Muster (*heptad periodicity*, Abbildung 28).



**Abbildung 28** Sequenzeigenschaften der HAMP Domäne. Die abgebildete Sequenz stammt von der HAMP Domäne aus *A. fulgidus*. Die hydrophoben Reste sind mit roten Kästen gekennzeichnet. Einheiten der *heptad periodicity* sind durch waagrechte Klammern markiert. Darunter stehen Positionsbezeichnungen gemäß der regulären *knobs-into-holes* Geometrie wie es erwartet und gemäß dem „*complementary x-da*“ wie es beobachtet wurde. Die rot bzw. blau markierten Reste entsprechen den roten und blauen Seitenketten aus Abbildung 27 unten rechts. Das obere Histogramm stellt die durchschnittliche Hydrophobizität jeder Position nach Kyte-Doolittle dar und das untere Histogramm das durchschnittliche Seitenkettenvolumen in Å<sup>3</sup>. Für die Histogramme wurde ein *alignment* aller bekannten HAMP Domänen verwendet. Die roten Kreise kennzeichnen die Position A291, an der Mutationen eingeführt wurden.

Diese Periodizität, die das Markenzeichen von *coiled coil* Strukturen ist (Crick, 1953), ließ eine reguläre *knobs-into-holes* Geometrie der hydrophoben Reste erwarten. Die *knobs-into-holes* Geometrie beschreibt einen Rest als *knob*, der von Resten der symmetrie-verwandten

Helix (*hole*) umgeben ist. Daher sind symmetrie-verwandte Reste entlang der Helix nebeneinander angeordnet (Lupas and Gruber, 2005).

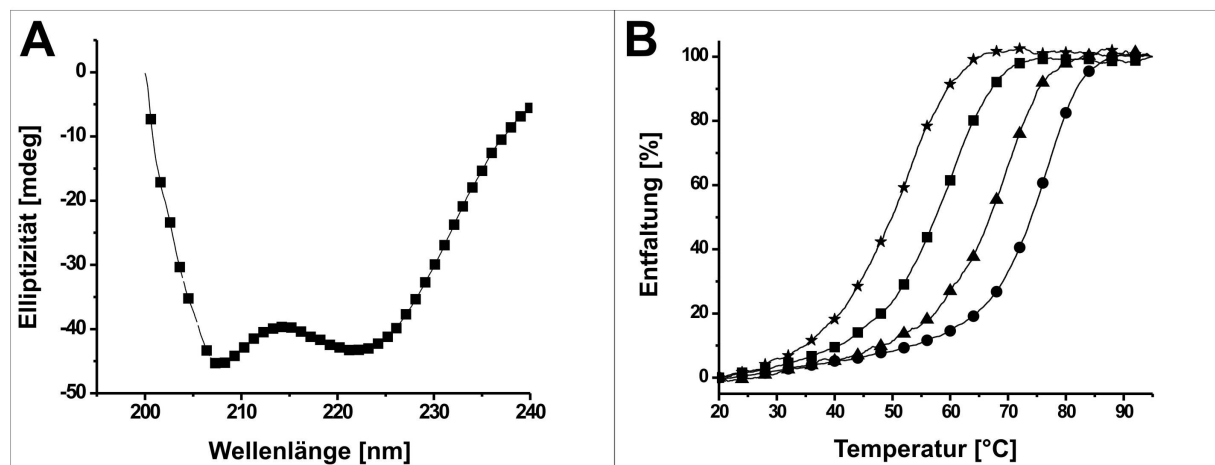
In der HAMP Domäne jedoch, weichen die Reste des Proteinkerns von dieser Geometrie ab und zeigen direkt auf die Helixachse und damit auf den symmetrisch liegenden Rest der gegenüberliegenden Helix (Abbildung 27 unten rechts). Diese Geometrie wurde in *coiled coil* Strukturen beobachtet, deren Sequenz vom kanonischen Wiederholungsmotiv abweicht (Gruber and Lupas, 2003). Diese abweichenden Geometrien werden durch zwei unterschiedliche geometrische Positionen für Reste des hydrophoben Kerns beschrieben. Zum einen gibt es die *x*-Position, in der die Reste direkt in Richtung Helixachse zeigen. Zum anderen gibt es *da*-Positionen, in denen die Reste seitwärts zeigen und eine zentrale Kavität bilden. In der HAMP Domäne gibt es vier Ebenen hydrophober Reste, die aus gemischten *x*- und *da*-Positionen bestehen. *x*-Positionen aus  $\alpha 1$  sind mit *da*-Positionen aus  $\alpha 2$  kombiniert und umgekehrt (Ebene 1: *x* I284 - *da* E311/I312; Ebene 2: *da* L287/S288 - *x* L315; Ebene 3: *x* A291- *da* S318/I319; Ebene 4: *da* I294/A295 - *x* L322; Reste in *coiled coils* werden gemäß ihrer geometrischen Position bezeichnet. Deshalb kann es vorkommen, dass ein Rest in zwei verschiedenen Konformationen zwei unterschiedliche Bezeichnungen trägt, sofern sich seine geometrische Lage durch eine Konformationsänderung verändert). Diese Geometrie wurde bisher nicht beschrieben und wurde hier erstmals als „*complementary x-da*“ benannt, da sich *x*-Positionen und *da*-Positionen gegenseitig ergänzen. Eine ähnliche Verpackungsgeometrie wurde bisher nur in der Tetramerisierungsdomäne von Lac-Repressor (1LBI, Reste 335-357) gefunden. In dieser Struktur sind die Helices jedoch antiparallel. Markus Gruber (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) versuchte, die experimentell ermittelte Struktur der HAMP Domäne mit der *complementary x-da* Geometrie rechnerisch in die kanonische *knobs-into-holes* Geometrie zu überführen. Es gelang, die gegebene Struktur durch 26° Drehung jeder Helix um ihre eigene Achse in die kanonische *knobs-into-holes* Geometrie zu drehen. Benachbarte Helices drehten dabei in entgegengesetzte Richtungen.

**Mutationen der Af HAMP Domäne** Die Af HAMP Domäne wurde rechnerisch durch Helixrotation von der *complementary x-da* in die *knobs-into-holes* Konformation überführt. Um sie tatsächlich ebenfalls in diese Konformation zu überführen wurden gezielte Mutationen in die HAMP Domäne von Af1503 eingeführt. Die Mutationen betrafen den Rest A291 und führten dort voluminösere Seitenketten ein. Diese Position im Inneren des *coiled coil* Kerns ist auffällig, da sie in HAMP Domänen von einem Rest mit kleiner Seitenkette eingenommen wird (Abbildung 28). Außerdem wurden andere HAMP Domänen an dieser Position bereits mutiert. Die HAMP Domänen der Histidinkinasen EnvZ und NarX haben an der A291 von Af1503 entsprechenden Position ebenfalls einen Alanin Rest. Mutationen von diesem Alanin führten zu phänotypischen Veränderungen (Appleman and Stewart, 2003; Tokishita et al., 1992).

**Tabelle 4 Schmelztemperaturen verschiedener HAMP Mutanten**

Af HAMP	Wildtyp	A291C	A291V	A291F
Schmelztemperatur [°C]	59	68	77	50

Zunächst wurden hier vier Varianten von Af HAMP hergestellt: Wildtyp, A291V, A291C, A291F. Die CD-Spektren waren jeweils identisch und widerspiegelten die  $\alpha$ -helikale Struktur der HAMP Domäne (Abbildung 29 A). Die Mutationen des Restes A291 zeigten einen großen Einfluss auf die Schmelztemperatur und damit auf die Stabilität der HAMP Domäne. (Abbildung 29 B und Tabelle 4). Die Mutation A291V erhöhte die Schmelztemperatur gegenüber dem Wildtyp um 18° C. Die Mutation A291F hingegen senkte die Schmelztemperatur um 9° C.



**Abbildung 29 CD-Spektrum und Schmelzkurven der HAMP Domäne.** (A) CD-Spektrum von Af HAMP (Wildtyp). (B) Mit einem CD-Spektropolarimeter aufgenommene Schmelzkurven der HAMP Domäne: Wildtyp (■), A291V (●), A291C (▲), A291F (★).

Ein Zusammenhang zwischen Seitenkettengröße und Schmelztemperatur konnte daher nicht abgeleitet werden, da Valin und Phenylalanin jeweils voluminösere Seitenketten besitzen als Alanin, aber die Schmelztemperatur entgegengesätzlich beeinflussen. Ebenso war es möglich, Af HAMP A291I herzustellen und zu reinigen, das ebenfalls gefaltet war, aber nicht weiter untersucht wurde. Es wurden die Mutanten Af HAMP I319D und I319C hergestellt, die zwar löslich exprimiert wurden, aber im CD-Spektrum keine Sekundärstrukturen erkennen ließen (ohne Abbildung). Für die folgenden Untersuchungen jeweils wurde nur ein ausgewählter Teil der Mutationen verwendet.

**Funktionelle Untersuchung von Af HAMP** Das nächste Ziel war es, funktionelle Effekte der HAMP Domäne und den Mutanten der HAMP Domäne auf Signaltransduktionssysteme zu untersuchen. Mit dem nativen Protein Af1503 und dem nativen Organismus *A. fulgidus*



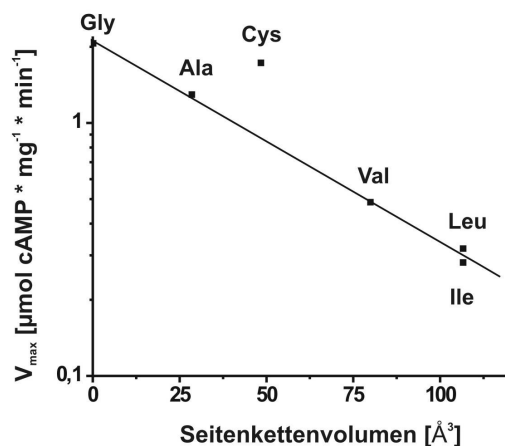
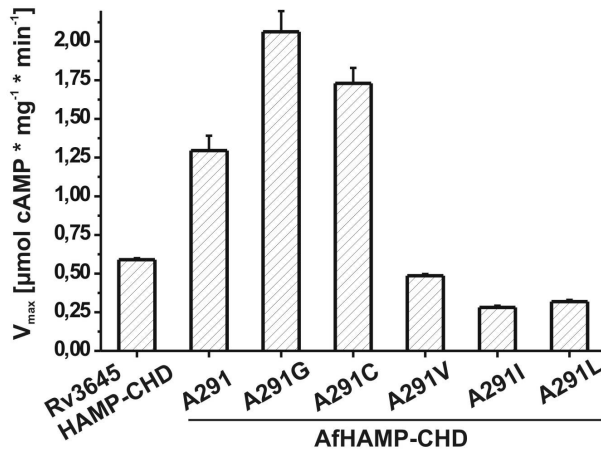
konnten keine Experimente durchgeführt werden. Für das Protein stand kein Assay zur Verfügung und der hyperthermophile Organismus *A. fulgidus* ließ sich unter gegebenen Laborbedingungen nicht kultivieren oder genetisch verändern. Um dennoch Effekte der *A. fulgidus* HAMP Domäne zu untersuchen, wurden HAMP Domänen aus experimentell zugänglichen Proteinen und Organismen durch Af HAMP substituiert und *in vitro* und *in vivo* Assays durchgeführt. Für *in vitro* Assays wurde die HAMP Domäne N-terminal an die Adenylatcyclase Rv3645 von *Mycobacterium tuberculosis* fusioniert und in Adenylatcyclase Assays von Franziska Berndt, Jürgen Linder und Joachim Schultz (Institut für Pharmazie, Universität Tübingen) getestet. Die Assays basierten auf dem Nachweis des radioaktiv markierten Reaktionsproduktes cAMP der Adenylatcyclase. Af HAMP beeinflusste die Aktivität der Adenylatcyclase von Rv3645. Af HAMP Wildtyp erhöhte die Aktivität der Adenylatcyclase gegenüber der nativen HAMP Domäne und Af HAMP A291V reduzierte die Aktivität (Abbildung 30 A und B). Dies zeigte zum einen die Austauschbarkeit von Af HAMP mit anderen HAMP Domänen, den aktivitätsregulierenden Effekt von Af HAMP und den Einfluss des Restes an Position 291 auf das Signaltransduktionsverhalten der HAMP Domäne. In weitergehenden Versuchen konnte eine Abhängigkeit zwischen der Seitenkettengröße an Position A291 und der Adenylatcyclaseaktivität gezeigt werden (Abbildung 30 A und B): Je größer die Seitenkette an der Position des A291 desto geringer die Adenylatcyclaseaktivität.

Ferner wurde der Einfluss von Af HAMP Wildtyp und von den Mutanten A291V und A291C auf ihre Signaltransduktionseigenschaften *in vivo* untersucht. Dafür wurde bakterielles, chemotaktisches Verhalten beobachtet. Unstimulierte Bakterien zeigen zwei alternierenden Bewegungsformen: Schwimm- und Taumelbewegung. In einem Gradienten von anziehenden oder abstoßenden Agenzien erhöht sich der Anteil der Schwimmbewegung gegenüber dem Taumeln, was durch eine Phosphorylierungskaskade ausgelöst, an dessen Anfang der modular aufgebaute Aspartatrezeptor Tar steht, der eine HAMP Domäne enthält. Diese HAMP Domäne (Tar<sub>214-266</sub>) wurde gegen die HAMP Domäne aus *A. fulgidus* ersetzt. Es wurden die wildtyp Sequenz von Af HAMP, sowie die Mutationen A291V und A291C getestet. *E. coli* Zellen, die diese chimären Rezeptoren exprimierten, verloren ihre Chemotaxisfähigkeit (Abbildung 31). In der Mitte der Platten sind die Zellen als helle kreisförmige Flächen zu sehen. Ohne Chemotaxisrezeptoren diffundierten die Zellen nur wenig vom Zentrum weg. Mit nativem Aspartatrezeptor folgten die Zellen dem Nährstoffgradienten, den sie selbst durch den Verbrauch der Nährstoffe in der Mitte der Platte erzeugten. Man sieht eine starke Ausbreitung zum Rand der Platte. Zellen mit chimären Rezeptoren zeigten minimale Ausbreitung, die etwas stärker war als bloße Diffusion. Sie war jedoch bei weitem nicht so ausgeprägt wie beim nativen Rezeptor. Dies war sowohl der Fall für die wildtyp Sequenz von Af HAMP, als auch für die Mutanten A291V und A291C. Allerdings ist die Unfähigkeit, einem chemischen Gradienten zu folgen, nur ein Teil des Phänotyps.

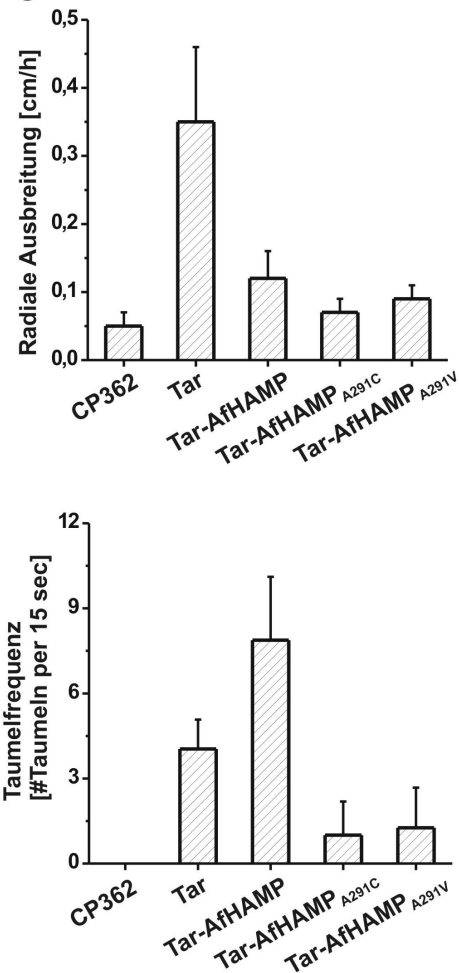
A

Af1503	275	FTT	STITRPIIELSNTADKIAEGNLEAEVPHQNR	EADEIGILAKSIERLRRSLKVAME	SLEEALK	338
Rv3645	275	LVA	MSIADPLRQLRWALSEVQRGNNAHQ-	IYDASELGLLQAGFNDMVRELSE	QRLRDLFGRY	338
Tar	211	GIR	MLLTPLAKIIAHIREIAGNLANTLT-	IDGRSEMGLAQSVSHMQRSLTDTVT	HVREGSDA	274
			*			
AfHAMP-CHD			STITRPIIELSNTADKIAEGNLEAEVPHQNR	EADEIGILAKSIERLRRSLKVAME	LRDLFGRY...	
Tar-AfHAMP	..	GIR	MITRPIIELSNTADKIAEGNLEAEVPHQNR	EADEIGILAKSIERLRRSLTDTVT	HVREGSDA...	

B



C



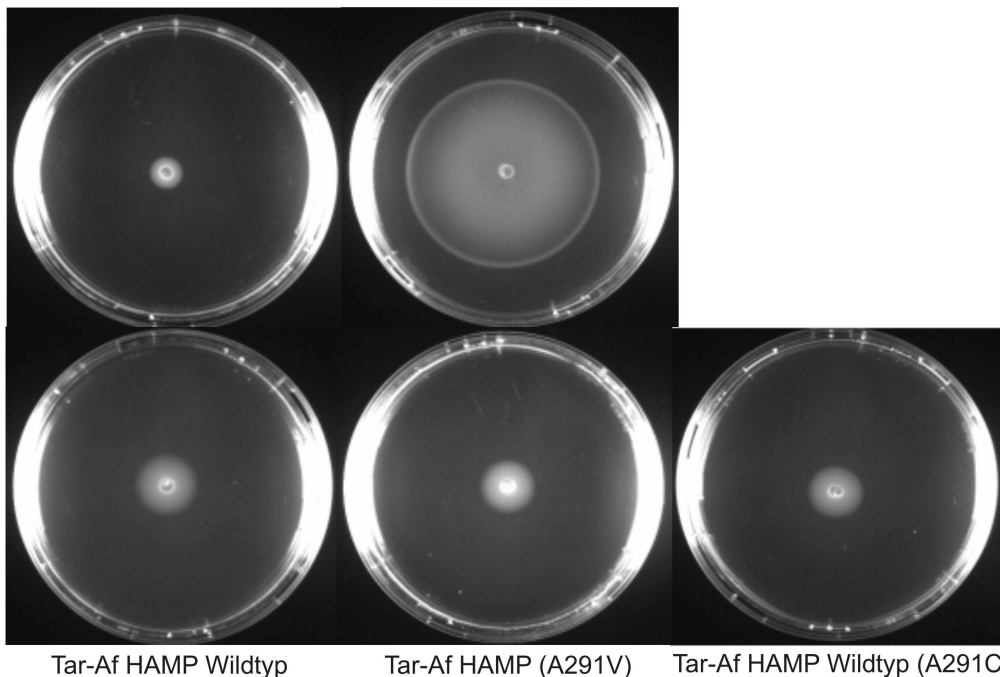
**Abbildung 30** Die HAMP Domäne von Af1503 beeinflusst die Aktivität der katalytischen Domäne der mycobakteriellen Adenylatcyclase Rv3645 *in vitro* und die Funktion des *E. coli* Aspartatrezeptors Tar *in vivo*. (A) Sequenz alignment, welches die Fusionsgrenzen der chimären Proteine AfHAMP-CHD und Tar-AfHAMP relativ zu den nativen Sequenzen von Af1503, Rv3645 und Tar wiedergibt. Die HAMP Domänen sind mit einem Kasten und die Position des mutierten Alanin Restes ist mit \* markiert. (B) Rv3645HAMP-CHD ist ein Fragment von Rv3645, das aus der mycobakteriellen HAMP Domäne und CHD Domäne besteht (CHD: cyclase homology domain). AfHAMP-CHD ist ein chimäres Protein, in dem die mycobakterielle CHD Domäne C-terminal an Af HAMP fusioniert wurde. Mutationen in den chimären Proteinen betreffen die Position, die A291 in Af1503 entspricht und werden folgendermaßen bezeichnet: A291G, A291C, A291V, A291I und A291L von AfHAMP-CHD. (Oben) Spezifische Aktivität der Adenylatcyclase. Fehlerbalken stellen die Standardabweichung dar ( $n > 3$ ). (Unten) Spezifische Aktivität der AfHAMP-CHD Chimären aufgetragen gegen die Seitenkettengröße des Restes in der Position "291". Jeder Datenpunkt ist mit der Abkürzung der entsprechenden Aminosäure markiert. Die Gerade stellt eine lineare Regression aller Datenpunkte bis auf A291C dar. (C) CP362 sind *E. coli* Zellen, die Defizient

für alle Chemotaxisrezeptoren sind. Tar und Tar-AfHAMP sind Rezeptoren, die in CP362 Zellen exprimiert wurden und den *E. coli* Aspartatrezeptor Tar oder Tar, in dem die native HAMP Domäne gegen Af HAMP ausgetauscht wurde, enthalten. Bei Tar-AfHAMP<sub>A291C</sub> und Tar-AfHAMP<sub>A291V</sub>, wurde Alanin, welches Rest 291 in Af1503 entspricht, gegen Cystein bzw. Valin ausgetauscht. (Oben) Radiale Schwimmgeschwindigkeit in swarm-plate Assays. (Unten) Taumelfrequenz der Zellen. Fehlerbalken stellen Standardabweichung dar (oben n = 6, unten n > 15).

Die Unfähigkeit für Chemotaxis kann nämlich sowohl durch erhöhte Taumel- als auch durch erhöhte Schwimmanteile in der Zellbewegung ausgelöst werden. Beide Störungen führen zum Verlust effektiver, gradientengetriebener Bewegung. Mikroskopische Betrachtung der Zellbewegung ließ erkennen, dass Af HAMP Wildtyp zu erhöhter Taumelbewegung führte und die Mutationen A291V und A291C zu erhöhter Schwimmbewegung (Abbildung 30 C).

In *in vivo* als auch *in vitro* wurde die A291C Mutante der HAMP Domäne getestet. *In vivo* verhielt sie sich wie die A291V Mutante, *in vitro* hingegen wie die wildtyp Domäne. Die Ursache für diese Diskrepanz bleibt unklar. Eine Disulfidbrücke zwischen den Helices scheint nicht ausgebildet gewesen zu sein, weil *in vitro* kein Unterschied zwischen reduzierenden und oxidierenden Reaktionsbedingungen zu beobachten war. Eine Eigenschaft der Thiolgruppe ist ihre hohe Polarisierbarkeit, die neben der Seitenkettengröße ebenfalls eine Rolle spielen könnte und sich je nach Bedingung und Domänenkontext gegensätzlich verhält.

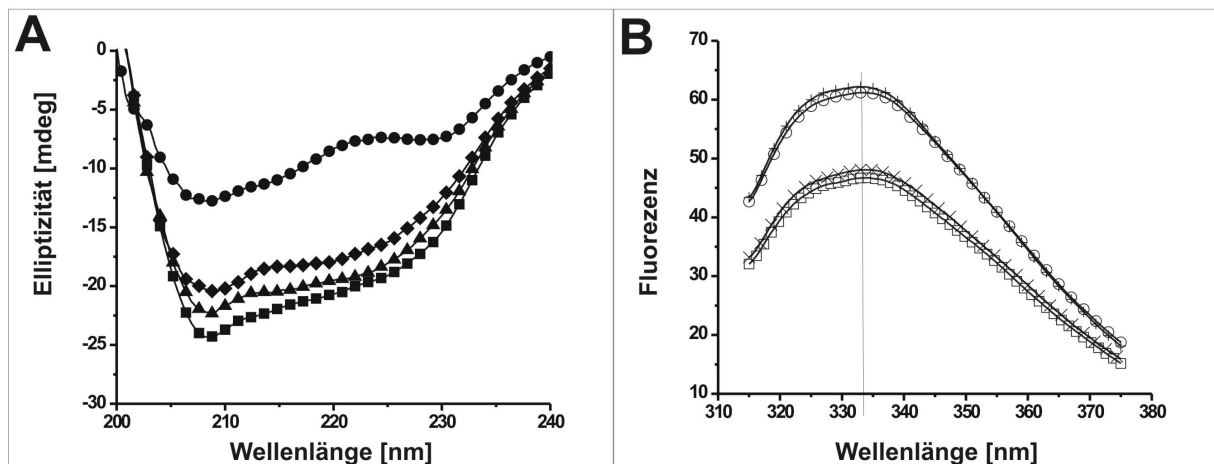
Ohne Chemotaxis-Rezeptoren    Mit nativem Aspartatrezeptor Tar



**Abbildung 31 Platten von swarm-plate Assays.** *E. coli* Zellen in LB-Medium wurden in die Mitte der Platte pipettiert. Die Zellen metabolisierten die Nährstoffe im Zentrum und erzeugten einen Nährstoffgradienten mit den höchsten Konzentrationen am Rand der Platte. Das Ausmaß, mit dem die Zellen den Gradienten entlang schwammen, zeigte sich im Radius der Kolonie (helle Kreise).

**NMR-Analyse von Af HAMP A291V** Mit NMR Analyse wurde versucht, die Struktur von Af HAMP A291V zu bestimmen. Dies gelang nicht, da ein schnelles Gleichgewicht zwischen zwei Konformationen vorlag. Dennoch konnten aus den Spektren bestimmte Eigenschaften für Af HAMP A291V abgeleitet werden. Af HAMP A291V zeigt eine kompakte Vierhelixbündelstruktur und ist der wildtyp HAMP Domäne sehr ähnlich. Die Verpackungsgeometrie ist entlang der Vierhelixbündelachse nicht verschoben. Die genaue Verpackungsgeometrie der kernbildenden Reste blieb jedoch unklar.

**Herstellung der Proteine Af1502, Af1503 TM2CC, Af1503 EP und Af1504** Die Proteine Af1502, Af1503 TM2CC, Af1503 EP (die extraplasmatische Domäne von Af1503) und Af1504 konnten als lösliche Proteine in *E. coli* exprimiert werden. Die Mutanten Af1503 TM2CC D88A und Af1503 TM2CC D92A wurden als unlösliche Proteine exprimiert. Der Transmembranbereich von Af1503 wurde durch eine hydrophilere *coiled coil* bildende Helix ersetzt, um die lösliche Expression zu begünstigen. Alle Proteine ergaben typische CD-Spektren für gefaltete Proteine (CD-Spektrum von Af1502 nicht abgebildet; von Af1503 Abbildung 32; von Af1504 Abbildung 22 und Kapitel 5). Aus dem negativen Peak bei 208 nm aller Af1503 Varianten sind  $\alpha$ -helikale Sekundärstrukturelemente zu erkennen.

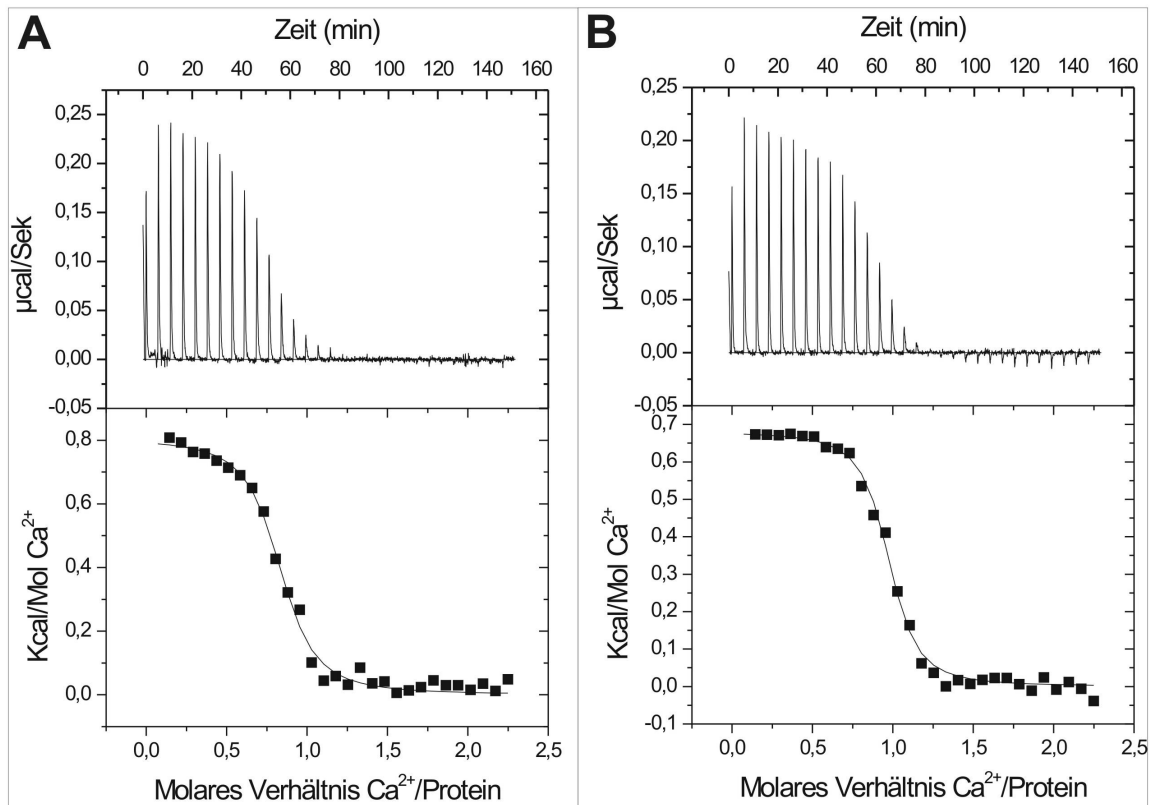


**Abbildung 32 CD- und Fluoreszenzspektren von Af1503 Proteinen.** (A) CD-Spektren von Af1503 TM2CC (■), Af1503 TM2CC D88A (▲), Af1503 TM2CC D92A (◆), Af1503 EP (●). (B) Fluoreszenzspektren von Af1503 TM2CC (○), Af1503 EP (+), Af1503 TM2CC D88A (□), Af1503 TM2CC D92A (X)

Allerdings ist der für  $\alpha$ -Helices erwartete negative Peak bei 222 nm nicht erkennbar. Dieser könnte durch Überlagerung von Signalen von  $\beta$ -Faltblattstrukturen oder ungeordneten Bereichen verdeckt sein. Weil der Peak bei 208 nm im Vergleich zu reinen  $\alpha$ -helikalen Spektren außergewöhnlich stark ausgeprägt ist, ist ein Anteil von ungeordneten Sekundärstrukturbereichen wahrscheinlich, die einen starken negativen Peak bei 200 nm erzeugen. Die extraplasmatische Domäne Af1503 EP zeigt einen schwachen, unerwarteten

negativen Peak bei 229 nm. Af1504 und Af1503 enthalten Tryptophanreste, die für Fluoreszenzspektroskopie genutzt wurden, um die Proteinumgebung des Tryptophans zu untersuchen. Bei Fluoreszenzanregung mit Licht der Wellenlänge 293 nm, waren maximale Fluoreszenzintensitäten bei folgenden Emissionswellenlängen zu beobachten: 338 nm bei Af1504 (Abbildung 22), 332,6 nm bei Af1503 EP und 333,6 nm bei Af1503 TM2CC und den mutierten Proteinen Af1503 TM2CC D88A und D92A (Abbildung 32). In allen Fällen deutet dies auf eine hydrophobe Umgebung des Tryptophanrestes hin, wie sie im Inneren eines gefalteten Proteins vorkommt. Af1503 TM2CC D88A und D92A zeigen identische CD- und Fluoreszenzspektren wie das unmutierte Protein Af1503 TM2CC. Daraus lässt sich schließen, dass ihre Rückfaltung in die native Struktur gelang. Mit kalibrierten Gelfiltrationssäulen wurde festgestellt, dass Af1503 TM2CC, sowie die Mutationen D88A und D92A als Tetramer vorliegen und Af1503 EP als Monomer (ohne Abbildung).

**Calciumbindung an Af1503** Mit dem Programm HHsenser wurde von Johannes Söding (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) ein potentielles EF-Hand Motiv in der extraplasmatischen Domäne von Af1503 (Af1503EP) gefunden:  $D_{88}LIQD_{92}PKFRSLGAQ$  und  $D_x[DN]_x[DNS]G_x(5)[ED]$  ist die EF-Hand Konsensussequenz. Daraufhin wurde die Bindung von Calcium an Af1503 experimentell mit kalorimetrischer Titration untersucht. Dazu wurde Proteinlösung von Af1503 TM2CC, Af1503EP, Af1503 TM2CC D88A oder Af1503 TM2CC D92A in eine adiabatisch isolierte Mikrokalorimeterzelle gegeben und mit Calcium titriert (Abbildung 33). Die entstehende Reaktionswärme wurde gemessen. Die molare Reaktionsenthalpie je Zugabe wurde gegen das molare Verhältnis aus Calcium/Protein aufgetragen. Daraus wurde die Bindungskonstante, die Stöchiometrie der Bindung und die freie Reaktionsenthalpie bestimmt. Die Dissoziationskonstante ergibt sich als reziproker Wert der Bindungskonstante. Die Ergebnisse der kalorimetrischen Messungen sind in Tabelle 5 zusammengefasst.



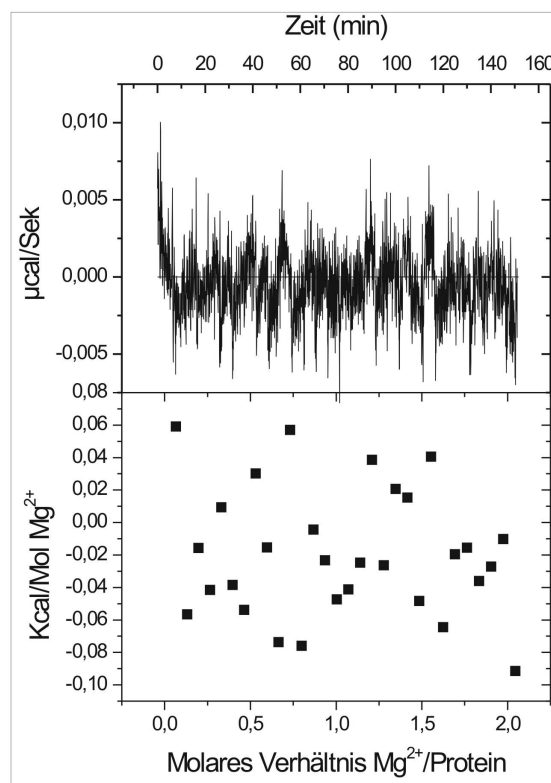
**Abbildung 33 ITC-Messung zur Calciumbindung an Af1503.** Im oberen Bereich ist jeweils die Reaktionswärme je Zugabe aufgetragen. In den unteren Bildteilen sind Kcal/Mol je Zugabe gegen das Stoffmengenverhältnis Ligand zu Rezeptor aufgetragen. (A) Calcium zu Af1503 TM2CC, (B) Calcium zu Af1503 EP.

Calciumbindung zu Af1503 TM2CC war eine endothermer Vorgang. Für eine exergone Reaktion müsste die Reaktion entropisch angetrieben sein. Dafür kommen Konformationsänderungen des Proteins in Betracht sowie der Verlust der Hydrathüllenstruktur der Calciumionen. Deswegen wurden CD- und Fluoreszenzspektren von Af1503 EP und Af1503 TM2CC mit und ohne Calcium aufgenommen. Unterschiede in den Spektren und damit Unterschiede im Sekundärstrukturgehalt mit und ohne Calcium konnten jedoch nicht beobachtet werden. Auch der Oligomerisierungsgrad wurde durch Calcium nicht beeinflusst, was durch analytische Gelfiltration gezeigt werden konnte (ohne Abbildung). Af1503 TM2CC war ein Tetramer, Af1503EP ein Monomer. Die gefundenen Bindungsstöchiometrien sind bei Af1503 EP 1:1 und bei Af1503 TM2CC 0,7 Calcium pro Proteinmonomer. Die Assoziationskonstante ist bei Af1503 EP mit  $1,4 \times 10^6 \text{ M}^{-1}$  gegenüber Af1503 TM2CC mit  $7,8 \times 10^5 \text{ M}^{-1}$  um Faktor 2 höher (Tabelle 5). Bei der Zugabe von Magnesium konnte keine Bindungswärme gemessen werden, was durch fehlende Bindung gedeutet werden kann (Abbildung 34). Um die Bindung des Calciums im Bereich der vorhergesagten EF-Hand zu lokalisieren, wurden im Protein Af1503 TM2CC die Aspartatreste D88 (Af1503 TM2CC D88A) und D92 (Af1503 TM2CC D92A) jeweils einzeln gegen Alanin ausgetauscht.

**Tabelle 5 Ergebnisse der Calciumbindungsstudien mit Af1503.**

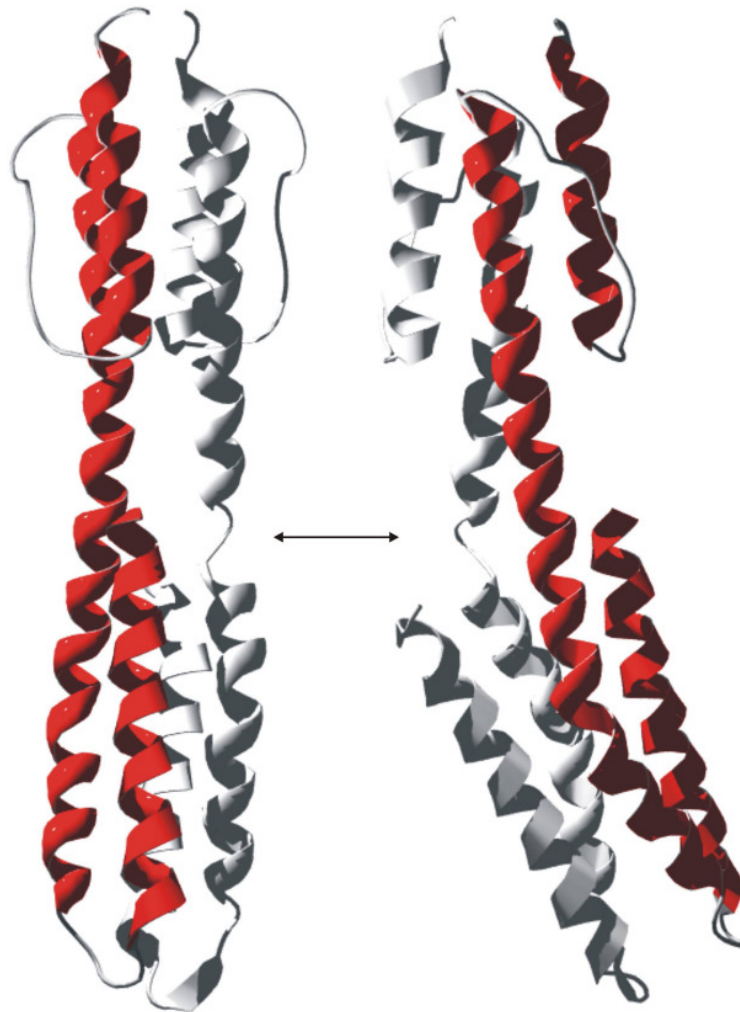
Protein	Af1503	Af1503	Af1503	Af1503
	TM2CC	EP	TM2CC	TM2CC
			D88A	D92A
Assoziationskonstante [ $M^{-1}$ ]	$7,8 \times 10^5$	$1,4 \times 10^6$	$5,3 \times 10^5$	$7,2 \times 10^5$
	$\pm 1,24 \times 10^5$	$\pm 2,17 \times 10^5$	$\pm 1,74 \times 10^5$	$\pm 4,0 \times 10^5$
Dissoziationskonstante [M]	$1,3 \times 10^{-6}$	$7 \times 10^{-7}$	$1,9 \times 10^{-6}$	$1,4 \times 10^{-6}$
Stöchiometrie Calcium/Protein	0,81	0,95	0,61	0,76
$\Delta H$ [Kcal/Mol]	$806 \pm 14$	$681 \pm 8$	$833 \pm 42$	$747 \pm 51$
$\Delta S$ [Kcal/(Mol K)]	30	30	29	29

Diese mutierten Proteine wurden als unlösliche *inclusion bodies* exprimiert und mussten denaturiert gereinigt und rückgefaltet werden. Die Rückfaltung gelang durch Entfernung des Harnstoffs durch Dialyse. Die rückgefalteten Proteine zeigten identische CD- und Fluoreszenzspektren wie Af1503 TM2CC. Die Calciumbindung an diese mutierten Proteine war ähnlich der Calciumbindung an Af1503 TM2CC (Tabelle 5).



**Abbildung 34 Keine Magnesiumbindung an Af1503.** Im oberen Bereich ist die Reaktionswärme je Zugabe aufgetragen. Im unteren Bildteilen ist Kcal/Mol je Zugabe gegen das Stoffmengenverhältnis: Magnesium zu Af1503 TM2CC aufgetragen.

**Struktur von Af HAMP-DHP** Die Struktur von Af HAMP-DHP wurde mit Röntgenkristallographie gelöst. Die HAMP Domäne ist wie in der NMR-Struktur ein Vierhelixbündel und ist zur NMR Struktur fast identisch (Peptidrückgrat RMSD 1,45 Å). Der Übergang zur DHP Domäne ist bei Arginin 56 asymmetrisch abgeknickt (Abbildung 35). Die Struktur der isolierten DHP-Domäne ist bereits bekannt (Tomomori et al., 1999) und ist identisch mit der Struktur im Kontext der HAMP-Domäne. Es handelt sich bei der DHP Domäne um ein antiparalleles Vierhelixbündel.



STITRPIIELSNTADKIAEGNLEAEVPHQNRAD EIGILAKSIER  
 LRRSLKQLADDRTLLMAGVSHDLRTP LTRIRLATEMMSEQD  
 GYLAESINKDIEECNAIIIEQFIDYLRLR

**Abbildung 35 Struktur von HAMP-DHP.** Strukturmodell der HAMP Domäne in Fusion mit der Dimerisierungsdomäne von EnvZ. Die N-Termini befinden sich oben. Die beiden Strukturen unterscheiden sich durch Rotation um die vertikale Achse. Die grau dargestellte Proteinkette ist keine durchgehende Helix, sondern ist an einer Stelle unterbrochen (markiert durch einen Pfeil), an der die Symmetrie des dimeren Moleküls bricht. Darunter die Sequenz des Fusionskonstruktes. Die Sequenz von Af HAMP ist hervorgehoben und der Argininrest im DRT Sequenzmotiv, an dem sich der Knick befindet ist rotfarbig.



## 6.4 Diskussion

Es gelang, die erste Struktur einer HAMP Domäne zu bestimmen. Die HAMP Domäne ist ein  $\alpha$ -helikales, paralleles, dimeres Vierhelixbündel. Die beobachtete Konformation der im Kern verpackten Reste ist in zweierlei Hinsicht ungewöhnlich und neuartig. Es ist das erste natürlich vorkommende, parallele *coiled coil* Vierhelixbündel mit *complementary x-da* Geometrie (alle anderen natürlichen oder durch Mutationen modifizierte Vierhelixbündel mit entsprechender *x-da* Geometrie sind antiparallel; parallele Bündel zeigen jeweils die kanonische *knobs-into-holes* Geometrie). Zudem ist es die einzige Struktur mit asymmetrischer Verpackung, nämlich einem gemischten *a-d-e* und *a-d-g* Kern. Das heißt  $\alpha 1$  verwendet die Reste *a*, *d* und *e*, um den hydrophoben Kern zu bilden und  $\alpha 2$  verwendet die Reste *a*, *d* und *g* (vgl. Abbildung 28) (die Reste einer Siebener-Wiederholung sind hierbei von *a* bis *g* durchnummeriert).

Die gefundene Struktur entsprach nicht den Erwartungen aus der Sequenzanalyse. Aufgrund des hydrophoben Musters und der *heptad periodicity* wurde eine kanonische *knobs-into-holes* Strukturkonformation erwartet. Diese kanonische *knobs-into-holes* Struktur konnte rechnerisch modelliert werden, indem jede Helix um  $26^\circ$  um ihre eigene Achse gedreht wurde. Da es sich um ein Signaltransduktionsprotein handelt, wurde die Hypothese aufgestellt, dass beide Konformationen möglich sind und jede Konformation einem Signalzustand entspricht. Um dies zu zeigen, wurden Mutationen in die HAMP Domäne eingeführt, welche die kanonische *knobs-into-holes* Konformation stabilisieren sollten. Eine geeignete Position schien A291 des Af1503 Proteins zu sein (Abbildung 28). An dieser *x*-Position sind Reste mit kleinen Seitenketten konserviert und es ist bekannt, dass kleine Seitenketten in *x*-Positionen bevorzugt sind (Gruber and Lupas, 2003).

Dies führte zur Idee, durch Vergrößerung der Seitenkette an dieser Position, die kanonische *knobs-into-holes* Konformation zu stabilisieren. Mehrere Mutationen wurden hergestellt und alle zeigten einen großen Einfluss auf die Thermostabilität der HAMP Domäne. Den größten Einfluss hatte die Mutante Af HAMP A291V. Hier zeigte sich eine Zunahme der Thermostabilität und der Schmelzpunkt lag  $18^\circ$  C über dem der wildtyp Domäne. Diese auffällige Stabilitätsveränderung gibt einen ersten Hinweis auf eine Konformationsänderung der HAMP Domäne, die durch die Mutation ausgelöst wurde.

*In vivo* Chemotaxisassays und *in vitro* Adenylatcyclaseassays zeigten, dass die wildtyp HAMP Domäne und die Mutation A291V Chemotaxisphänotypen und Adenylatcyclaseaktivitäten in entgegengesetztem Sinne beeinflussen. In den Adenylatcyclaseassays konnte sogar eine Abhängigkeit zwischen der Aktivität und des Seitenkettenvolumens an dieser Position festgestellt werden. Die Mutation A291C fügt sich nicht in diesen Zusammenhang ein. Eine mögliche Erklärung ist die hohe Polarisierbarkeit der Thiolgruppe. Diese Assays geben den zweiten Hinweis auf eine Konformationsänderung der HAMP Domäne, die durch Mutationen an Position A291 ausgelöst wurden. Zusätzlich zeigten diese Ergebnisse einen Zusammenhang zwischen dem Seitenkettenvolumen und der

Stärke der Veränderung der HAMP Domäne. Insbesondere zeigen die Ergebnisse, dass die Mutationen eine Konformationsänderung bewirkten, die ein zur wildtyp Domäne entgegengesetztes Signaltransduktionsverhalten erzeugten.

Es wurde versucht, die Struktur der A291V Mutante der HAMP Domäne mit NMR-Spektroskopie zu bestimmen. Eine genaue Struktur konnte aus den Messdaten nicht abgeleitet werden, da ein schnelles Gleichgewicht aus zwei Konformationen vorlag. Trotzdem konnte wichtige Informationen für Af HAMP A291V abgeleitet werden. Bei der Struktur handelt es sich um ein kompaktes Vierhelixbündel ähnlich zur wildtyp Struktur und die hydrophoben Reste der Untereinheiten sind, im Vergleich zur wildtyp Domäne, entlang der Bündelachse nicht verschoben. Diese Ergebnisse können einige der existierenden Modelle über den Signaltransduktionsmechanismus der HAMP Domäne ausschließen. Die Resultate widersprechen dem „Helix rearrangement“ Modell, weil die HAMP Domäne eine kompakte Struktur behält während sie unterschiedliche Zustände signalisiert. Die Resultate widersprechen dem „Pushrod-piston“ Modell, da die Untereinheiten in unterschiedlichen Signalzuständen nicht entlang der Bündelachse verschoben sind. Außerdem sind die Untereinheiten nicht durch Rotation um die Bündelachse verschoben, was gegen das „Frozen dynamic dimer“ Modell spricht.

Betrachtet man die NMR-Daten von Af HAMP und Af HAMP A291V zusammen mit den möglichen Konformationen einer *coiled coil* Struktur, so kommt für die HAMP Domäne nur die Drehung der Helices um  $26^\circ$  um ihre eigene Helixachse in die *knobs-into-holes* Konformation als molekularer Signaltransduktionsmechanismus in Frage.

Aus der Rotation jeder einzelnen Helix um ihre eigene Achse ergibt sich die Überlegung über die Symmetrie der kernbildenden Reste. Da die Reste in beiden Konformationen ineinander verzahnt sind, kann die Drehung nur zahnradartig erfolgen, bei der benachbarte Helices in entgegengesetzte Richtungen drehen. *Knobs-into-holes* ist eine symmetrische Verpackung. Deswegen muss die Drehung zur *complementary x-da* Geometrie eine Asymmetrie verursachen. Tatsächlich ist dies in Af HAMP zu beobachten, da  $\alpha 1$  die Reste *a*, *d* und *e* für den Kern verwendet und  $\alpha 2$  die Reste *a*, *d* und *g*.

Die angenommene Rotationsbewegung, um die Konformationen ineinander zu überführen, kann auch durch frühere Experimente unterstützt werden. Andrea Cochran und Peter Kim (Cochran and Kim, 1996) konnten bereits zeigen, dass GCN4 Leucin zipper als N-terminaler Fusionspartner des cytoplasmatischen Tar-Rezeptorfragments (C-terminal der Tar HAMP Domäne) zu funktionellen Chimären führt. Durch Einfügen von drei oder vier Resten, was einer Rotation gegen bzw. im Uhrzeigersinn entspricht, konnte der Phosphorylierungsgrad von assoziierten Kinasen erhöht bzw. erniedrigt werden. Kürzlich zeigten Martin Engelhard und Mitarbeiter, dass der Phototaxisrezeptor NpSR II von *Natronomonas pharaonis* bei Lichteinwirkung eine Drehung der letzten Transmembranhelix um  $15^\circ$  und einen translationalen Versatz entlang der Helixachse um  $0.9 \text{ \AA}$  vollzieht (Moukhametzianov et al., 2006; Wegener et al., 2001).

Es bleibt zu klären wie ein Signal die HAMP Domäne erreicht und welche Bedeutung die HAMP Domäne für die Signaltransduktion hat. Die HAMP Domäne kann als Konverter von undefinierten Eingangssignalen in ein definiertes Ausgangssignal dienen. Die Konformationsänderungen der extraplasmatischen Chemotaxisrezeptor Domänen scheint nämlich eine Mischung aus Rotation und Translation zu sein (Chervitz and Falke, 1996; Moukhametzianov et al., 2006). Für die HAMP Domäne sind jedoch nur die beiden definierten Konformationen vorstellbar. Man kann sich damit die HAMP Domäne als binären Schalter vorstellen, der umspringt, sobald ein Schwellenwert des Eingangssignals überschritten wurde. Die HAMP Domäne könnte somit eine Signalintegration ausführen. Ein breiter Aktivitätsbereich kann durch Steuerung des Gleichgewicht zwischen beiden Konformationen der HAMP Domäne erklärt werden.

Der vorgestellte Rotationsmechanismus in Einklang mit dem *cis* und dem *trans* Signaltransduktionsmechanismus. In *cis* wird die Rotation an die C-terminalen Domänen weitergegeben. Diese sind in EnvZ und Tar bekanntermaßen ebenfalls *coiled coil* Strukturen (Kim et al., 1999; Tomomori et al., 1999). Für den *trans* Mechanismus kann man sich vorstellen, dass die Rotation neue Interaktionsoberflächen bietet, über die Wechselwirkung mit Adapterproteinen ermöglicht oder entfernt wird.

Kristallographisch wurde die Struktur der HAMP Domäne in Fusion mit der DHP Domäne von EnvZ gelöst. Die Struktur dieser HAMP Domäne ist identisch mit der NMR-Struktur. Die DHP Domäne ist wie erwartet ein antiparalleles Vierhelixbündel. Der Bereich der die beiden Domänen verbindet, enthält in einer der beiden Ketten des dimeren Proteins einen Knick beim Arginin an Position 55. Hier bricht die Symmetrie im sonst symmetrischen Dimer. Dieser Knick könnte ein durch die Kristallpackung ausgelöstes Artefakt sein. Es ist aber auch möglich, dass er eine Rolle bei der Signaltransduktion spielt. Die Beobachtungen, dass er nur in einem Strang auftritt, die Helixstruktur lokal unterbrochen wird und der Knick in einer konservierten Verbindungssequenz zwischen HAMP Domänen und Effektoromänen (Glutaminsäure-Arginin-Threonin) (unveröffentlichte Ergebnisse von Andrei Lupas) vorkommt, sprechen gegen ein Kristallisationsartefakt. Die Schiefelage der DHP Domäne könnte eine Hebelwirkung verursachen, welche die Rotationsbewegung der HAMP Domäne bei der Signalweiterleitung verstärkt. Diese Vermutung muss durch Vergleiche mit der Struktur von Af HAMP A291V-DHP, was noch bearbeitet wird, bestätigt werden (Tabelle 6).

Um die Rotationsbewegung endgültig zu beweisen, ist es nötig, die Struktur HAMP Domäne in der postulierten *knobs-into-holes* Konformation strukturell zu lösen. Dafür sind Punktmutanten der HAMP Domäne hilfreich, die bereits der wildtyp Domäne entgegengesetzte Effekte auf Signaltransduktionsaktivitäten gezeigt haben. Das Protein Af HAMP A291V DHP wurde bereits gereinigt und kristallisiert. Die Strukturlösung steht noch aus. Ebenso wurde Af HAMP A291F gereinigt und NMR-Spektren wurden davon aufgenommen. Dieses Protein zeigt ein langsames Gleichgewicht mit getrennten Resonanzsignalen für beide Konformationen. Auch diese Strukturlösung dauert noch aus.

Tabelle 6 gibt einen Überblick über den Zwischenstand von Ansätzen zur strukturellen Charakterisierung der HAMP Domäne.

**Tabelle 6 Zwischenstand der Strukturaufklärung von Af HAMP und Fusionsproteine.**

Protein	Löslich exprimiert und gereinigt	Kristallisiert	Röntgenbeugung [Å]	NMR-Spektren aufgenommen	Struktur gelöst
HAMP	✓			✓	✓
HAMP A291V	✓	✓	2,1	✓	
HAMP A291C	✓	✓	10	✓	
HAMP A291I	✓	✓	1,3		
HAMP A291F	✓	✓	1,2	✓	
HAMP-DHP	✓	✓	1,2		✓
HAMP-DHP (Se-Met)	✓	✓			
HAMP A291V-DHP	✓	✓	3		
HAMP A291V-DHP (Se-Met)	✓	✓			
HAMP-EnvZ	✓	✓	3,5		
HAMP A291V-EnvZ	✓	✓			
HAMP-cTsr	✓	✓			
HAMP A291V-cTsr	✓	✓			
HAMP-NarX	Nach Reinigung ausgefallen				
HAMP A291V-NarX	Nach Reinigung ausgefallen				

Neben der Strukturaufklärung der HAMP Domäne wurde noch die Ligandenbindung an die extraplasmatische Domäne von Af1503 untersucht. Calciumbindung des Rezeptors Af1503 TM2CC konnte gezeigt werden. Die Calciumbindung findet in der extraplasmatischen Domäne statt, da auch diese isolierte extraplasmatische Domäne Af1503 EP Calcium bindet. Die Affinität ist mit einer Assoziationskonstante von  $1,4 \times 10^6 \text{ M}^{-1}$  bei Af1503 EP sehr hoch (Tabelle 5). Zudem konnte hohe Spezifität für Calcium gezeigt werden, da andere zweiwertige Kationen wie Magnesium keine Bindung zeigten (Abbildung 34). Für Af1503 EP konnte die Stöchiometrie von einem Calcium pro Af1503 EP sicher bestimmt werden. Die Stöchiometrie Calcium pro Af1503 TM2CC liegt deutlich unter eins bei etwa 0,7 (Tabelle 5) und wurde reproduzierbar gemessen. Mehrere Erklärungen scheinen für den unerwarteten Wert von 0,7 möglich. Zum einen könnte der Rezeptor während der Reinigung bereits Calcium gebunden haben, das durch Zugabe von EDTA während der Reinigung nicht

entfernt wurde. Dieses bereits gebundene Calcium würde Bindestellen blockieren. Weil das Protein als Tetramer vorliegt, könnte die Calciumbindestelle für zugesetztes EDTA während der Reinigung schlecht zugänglich gewesen sein. Beim monomeren Af1503 EP könnte diese Abschirmung ausgeblieben sein. Bei genauer Betrachtung der Bindungskurve von Af1503 TM2CC fällt auf, dass die ersten beiden Messwerte leicht höher liegen als die folgenden. Man könnte darin eine antikooperative Bindung vermuten. Das würde heißen, dass initial gebundenes Calcium die Affinitäten weiterer Calciumbindestellen erniedrigt und deswegen eine reduzierte Anzahl von Bindungsstellen vortäuscht. Antikooperative Bindung ist auch beim Aspartatrezeptor Tar beobachtet worden (Yeh et al., 1993). Bei Af1503 EP kommt aufgrund des monomeren Zustands keine antikooperative Bindung in Frage.

Es bleibt zu klären, welchen Einfluss Calciumbindung auf Af1503 hat. CD- und Fluoreszenzspektren zeigten nämlich keine Änderung der Sekundärstrukturgehalte an. Allerdings würden diese Techniken auch nicht ausreichen, um z.B. Rotationen zu erkennen, wie sie für die HAMP Domäne postuliert wurden. Auch der Oligomerisierungsgrad wurde durch Calcium nicht beeinflusst.

Durch Mutation der Aspartatreste des potentiellen EF-Hand Motivs konnte keine wesentliche Schwächung der Calciumbindung beobachtet werden. Mit den Mutationen wurde eine Carboxylatgruppe, von der Calciumbindung vermutet wurde, entfernt. Es wurde davon ausgegangen, dass sich der Verlust einer Ionenbindung in der Affinität sichtbar macht. Da dies nicht der Fall war, deuten die Ergebnisse darauf hin, dass dieses Sequenzmotiv nicht an der Calciumbindung beteiligt ist und keine EF-Hand Struktur ausbildet. Die Ursache der scheinbar falschen Vorhersage des Programms HHsenser könnte darin liegen, dass sehr geringe Sequenzgemeinsamkeiten zu falschen Annahmen über Verwandtschaften geführt haben.

Im Zusammenhang mit den Ergebnissen über die HAMP Domäne ist vorstellbar, dass Calciumbindung die beschriebene Rotation der HAMP Domäne in Af1503 auslöst. Da die Gene *af1502-af1505* translational gekoppelt sind, ist auch eine funktionelle Interaktion denkbar. Af1502 oder Af1504 könnten mit der HAMP Domäne Komplexe bilden und Adapter für weitere Signaltransduktionsproteine bilden könnten. In diesem Fall würde man die Beteiligung weiterer Enzyme wie Histidinkinasen erwarten. Man kann sich aber ebenso eine direkte Interaktion mit dem Kationentransporter Af1505 vorstellen. Interaktionen zwischen diesen Proteinen mit Kopräzipitationsversuchen konnten *in vitro* jedoch nicht nachgewiesen werden. Ebenso wurde nach Proteinen aus Lysat von *A. fulgidus* gesucht, die mit den gereinigten Proteinen Af1502, Af1503 und Af1504 kopräzipitieren. Auch hier konnte kein potentieller Interaktionspartner gefunden werden. Eine Schwierigkeit bei diesen Interaktionsstudien war, dass die Arbeitsbedingungen weit von den physiologischen Bedingungen des Archaeums *A. fulgidus* entfernt waren. Erschwerend für genetische oder *in vivo* Experimente ist, dass der Organismus *A. fulgidus* äußerst schwer zu kultivieren ist und keine genetischen Methoden zur Verfügung stehen.

Um Ideen über die Funktion eines regulierten Calciumstroms durch den Transporter Af1505 zu sammeln, kann man auf bereits durchgeführte Studien in *E. coli* und anderen Bakterien zurückgreifen. In prokaryontischen Signaltransduktionssystemen sind Proteine mit calciumbindenden, extraplasmatischen Cache Domänen bekannt (Anantharaman and Aravind, 2000). Es konnte außerdem gezeigt werden, dass sich intrazelluläre Calciumkonzentrationen bei Chemotaxis verändern (Tisa and Adler, 1995b) und dass Verarmung von Calcium im Cytosol zu fehlerhaftem Chemotaxisverhalten führt (Tisa and Adler, 1995a). Die spezifische Rolle des Calciums blieb bisher jedoch ungeklärt. Für die Quelle des Calciums gibt es zwei denkbare Möglichkeiten. Entweder ist es bereits im Cytosol in gebundener Form und wird bei Ligandenbindung freigesetzt oder es befindet sich in einem regulierten Austausch mit der Umgebung. Für letzteres spricht, dass Ionenkanalinhibitoren das chemotaktische Verhalten beeinflussen (Tisa and Adler, 1993). Ein regulierter Calciumstrom mit direktem oder indirektem Einfluss auf chemotaktisches Verhalten ist eine vorstellbare Funktion für die Proteine Af1502-Af1505 im Archaeum *A. fulgidus*.

## 7 Anhang

### 7.1 Exprimierte Proteinkonstrukte

In diesem Abschnitt sind die Nukleotidsequenzen der exprimierten Proteine aufgelistet und abgebildet. Die übersetzte Aminosäuresequenz steht in blauer Farbe darüber. Primer sind in grüner Schrift abgebildet und stehen an der Stelle, an der sie an die Nukleotidsequenz binden. Die Sequenz der „reverse Primer“ ist als Sequenz des Vorwärtsstrangs in kursiver Schrift geschrieben und sind entsprechend als komplementäre Sequenzen eingesetzt worden. Die Bezeichnung einer Mutation ist an entsprechender Stelle mit roter Farbe markiert.

#### HIV-1 Protease WT, HIV-1 Protease D25N

pET30b (NdeI, HindIII); 99 Aminosäuren; 10,8 kD; pI = 8,8;  $\epsilon_{(280\text{ nm})} = 12780\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus HIV-1 Protease in pbd2-Plasmid, bekommen von Dr. Felix Müller-Sarnowski  
 (Max-von-Pettenkofer-Institut, München)

```

ACTGACTGACTGCATATG
CCTCAGATCACTCTTTGGCAAC
1 P Q I T L W Q R P L V T I K I G G Q L K
1 CCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAATTAAG
GGAAGCTCTATTAATAACAGGAGCAGATGATACAG
GGAAGCTCTATTAATAACAGGAGCAGATGATACAG
      N      HPR-M
21 E A L L D T G A D D T V L E E M N L P G
61 GAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGA
41 R W K P K M I G G I G G F I K V R Q Y D
121 AGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGAT
61 Q I L I E I C G H K A I G T V L V G P T
181 CAGATACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACA
81 P V N I I G R N L L T Q I G C T L N F *
241 CCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTTAA
      GATTGGCTGTACTCTAAATTTTTAA
      GCTTAAGCTTAAGCTT
  
```

#### HIV-1 Protease Q2C, D25N, C95A, L97C

pET30b (NdeI, HindIII); 99 Aminosäuren; 10,8 kD; pI = 8,8;  $\epsilon_{(280\text{ nm})} = 12780\text{ M}^{-1}\text{ cm}^{-1}$   
 HIV-1 Protease D25N in pET30b

```

ACTGACTGACTGCATATG
CCTTGCATCACTCTTTGG
1 P C I T L W Q R P L V T I K I G G Q L K
1 CCTTGCATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAATTAAG
21 E A L L N T G A D D T V L E E M N L P G
61 GAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGA
41 R W K P K M I G G I G G F I K V R Q Y D
121 AGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGAT
  
```

```

61 Q I L I E I C G H K A I G T V L V G P T
181 CAGATACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACA
81 P V N I I G R N L L T Q I G A T C N F *
241 CCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTGCTACTTGC AATTTTTTAA
      GACTCAGATTGGCGCTACTTGC AATTTTTTAA
      GCTTAAGCTTAAGCTT

```

### Trypsinogen S200A

pET3a; 247 Aminosäuren; 26,6 kD; pI = 6,1;  $\epsilon_{(280 \text{ nm})} = 36160 \text{ M}^{-1} \text{ cm}^{-1}$   
 Erhalten von Diethard Mattanovich (Universität für Bodenkultur, Wien)  
 (Universität für Bodenkultur, Wien; (Hohenblum et al., 2004))

```

1 A P F D D D D K I V G G Y N C E E N S V
1  GCCCCCTTTGATGATGATGACAAGATCGTTGGGGGCTACAACGTGAGGAGAATTCTGTC
21 P Y Q V S L N S G Y H F C G G S L I N E
61 CCCTACCAGGTGTCCTGAATTCTGGCTACCACTTCTGTGGTGGCTCCCTCATCAACGAA
41 Q W V V S A G H C Y K S R I Q V R L G E
121 CAGTGGGTGGTATCAGCAGGCCACTGCTACAAGTCCCGCATCCAGGTGAGACTGGGAGAG
61 H N I E V L E G N E Q F I N A A K I I R
181 CACAACATCGAAGTCCTGGAGGGGAATGAGCAGTTCATCAATGCAGCCAAGATCATCCGC
81 H P Q Y D R K T L N N D I M L I K L S S
241 CACCCCAATACGACAGGAAGACTCTGAACAATGACATCATGTTAATCAAGCTCTCCTCA
101 R A V I N A R V S T I S L P T A P P A T
301 CGTGCAGTAATCAACGCCCGCGTGTCCACCATCTCTCTGCCACCGCCCTCCAGCCACT
121 G T K C L I S G W G N T A S S G A D Y P
361 GGCACGAAGTGCCTCATCTCTGGCTGGGGCAACACTGCGAGCTCTGGCGCCGACTACCCA
141 D E L Q C L D A P V L S Q A K C E A S Y
421 GACGAGCTGCAGTGCCTGGATGCTCCTGTGCTGAGCCAGGCTAAGTGTGAAGCCTCCTAC
161 P G K I T S N M F C V G F L E G G K D S
481 CCTGGAAAGATTACCAGCAACATGTTCTGTGTGGGCTTCCTTGAGGGAGGCAAGGATTCA
      CATGTCAGGGTGATGCTGGTGGCCCTGTGG
      CATGTCAGGGTGATGCTGGTGGCCCTGTGG
181 C Q G D A G G P V V C N G Q L Q G V V S
541 TGTCAGGGTGATGCTGGTGGCCCTGTGGTCTGCAATGGACAGCTCCAAGGAGTTGTCTCC
201 W G D G C A Q K N K P G V Y T K V Y N Y
601 TGGGGTGATGGCTGTGCCAGAAGAACAAGCCTGGAGTCTACACCAAGGTCTACAACCTAC
221 V K W I K N T I A A N S *
661 GTGAAATGGATTAAGAACACCATAGCTGCCAATAGCTAA

```

### ClpA

pET30b (NdeI, XhoI); 758 Aminosäuren; 84,2 kD; pI = 5,9;  $\epsilon_{(280 \text{ nm})} = 33015 \text{ M}^{-1} \text{ cm}^{-1}$   
 von *Escherichia coli* genomischer DNA

```

      GGAATTCCAT
      ATGCTCAATCAAGA AACTGGA AACTC
1 M L N Q E L E L S L N M A F A R A R E H
1  ATGCTCAATCAAGA AACTGGA AACTCAGTTTAAATATGGCTTTTCGCCAGAGCGCGGAGCAC
21 R H E F M T V E H L L L A L L S N P S A
61 CGTCATGAGTTTATGACCGTTCGAGCACTTGTACTGGCGCTGCTCAGTAACCCATCTGCC
41 R E A L E A C S V D L V A L R Q E L E A
121 CGGGAGGCGCTGGAAGCGTGTCTGTGGATTTGGTTGCGCTCCGTCAGGA AACTGGAAGCC
61 F I E Q T T P V L P A S E E E R D T Q P
181 TTTATTGAACAAACCACACCCGTTCTGCTGCCAGTGAAGAGGAGCGCGACACACAGCCC
81 T L S F Q R V L Q R A V F H V Q S S G R
241 ACGCTGAGTTTTTCAGCGTGTACTGCAACGTGCGGTCTTCCATGTCCAGTCCCTCCGGTTCG
101 N E V T G A N V L V A I F S E Q E S Q A

```



301 AATGAGGTTACCGGTGCAAACGTTCTGGTCGCTATCTTTAGCGAACAGGAGTCGCAGGCC  
121 A Y L L R K H E V S R L D V V N F I S H  
361 GCATATCTGTTGCGTAAACATGAAGTCAGCCGCTCTCGATGTGGTGAACCTTTATCTCTCAT  
141 G T R K D E P T Q S S D P G S Q P N S E  
421 GGCACGCGTAAAGACGAGCCGACACAGTCTTCTGATCCTGGCAGCCAGCCAAAACAGCGAA  
161 E Q A G G E E R M E N F T T N L N Q L A  
481 GAACAAGCTGGTGGGGAGGAACGTATGGAGAATTTACGACGAACCTGAATCAGCTTGCG  
181 R V G G G I D P L I G R E K E L E R A I Q  
541 CGCGTGGGCGGAATCGACCCACTGATTGGTCGTGAGAAGGAGCTGGAGCGTGCTATTGAG  
201 V L C R R R K N N P L L V G E S G V G K  
601 GTTCTCTGCCGTCGCCGTAAAAACAACCCGCTGCTGGTGGGGGAATCTGGTGTCCGTAAA  
221 T A I A E G L A W R I V Q G D V P E V M  
661 ACCGCGATTGCGGAAGGTCTTGCCTGGCGAATTGTTTCAGGGCGATGTGCCGGAAGTGATG  
241 A D C T I Y S L D I G S L L A G T K Y R  
721 GCTGACTGTACGATTTACTCTCTCGATATCGGTTCTCTGTTAGCGGGCACAAAATATCGC  
261 G D F E K R F K A L L K Q L E Q D T N S  
781 GGCGACTTTGAAAAACGTTTTAAAGCGTTGCTCAAGCAGCTGGAGCAGGACACTAACAGC  
281 I L F I D E I H T I I G A G A A S G G Q  
841 ATCCTGTTTATTGATGAGATCCACACCATTATCGGTCGGGTGCAGCGTCTGGTGGTTCAG  
301 V D A A N L I K P L L S S G K I R V I G  
901 GTCGATGCGGCTAACCTAATCAAACCGTTGCTCTCCAGCGGTAAAAATTCGTGTAATTGGT  
321 S T T Y Q E F S N I F E K D R A L A R R  
961 TCGACAACCTATCAGGAGTTCAGCAACATTTTCGAGAAAAGACCGTGCTCTGGCGCGTCGC  
341 F Q K I D I T E P S I E E T V Q I I N G  
1021 TTCCAGAAAATTGATATTACTGAACCCGTCGATCGAAGAAAATGTTCAAATCATCAATGGC  
361 L K P K Y E A H H D V R Y T A K A V R A  
1081 CTGAAACCGAAGTATGAAGCGCACCCAGCAGTGC GTTATACCGCAAAGCGGTGCGTGC  
381 A V E L A V K Y I N D R H L P D K A I D  
1141 GCGGTAGAGCTGGCGGTGAAATACATTAACGATCGTCATCTGCCGGATAAAGCCATTGAT  
401 V I D E A G A R A R L M P V S K R K K T  
1201 GTTATCGACGAAGCGGGCGCTCGCGCACGCTGATGCCGGTAAAGCAAACGCAAGAAAACC  
421 V N V A D I E S V V A R I A R I P E K S  
1261 GTTAATGTGGCGGATATTGAGTCCGTGGTGGCCGATTGCACGCATTCCAGAGAAGAGT  
441 V S Q S D R D T L K N L G D R L K M L V  
1321 GTTCTCAGAGTGATCGTGATACCCTGAAAAACCTCGGCGATCGCTTGAAAATGCTGGTC  
461 F G Q D K A I E A L T E A I K M A R A G  
1381 TTCGGTCAGGATAAAGCCATTGAGGCGCTGACTGAAGCCATTAAGATGGCGCGTGCAGGT  
481 L G H E H K P V G S F L F A G P T G V G  
1441 TTAGGTCACGAACATAAACCAGTTGGTTTCGTTTCTGTTTGGCCGCCCTACCGGGGTCGGG  
501 K T E V T V Q L S K A L G I E L L R F D  
1501 AAAACAGAGGTGACGGTACAGCTTTGAAAAGCTTTGGGCATTGAGCTTCTGCGCTTTGAT  
521 M S E Y M E R H T V S R L I G A P P G Y  
1561 ATGTCCGAGTATATGGAACGCCATACCGCTCAGCCGCTTATTGGTGCCTCCGGGATAC  
541 V G F D Q G F L L T D A V I K H P H A V  
1621 GTTGGTTTTGATCAGGGCGGTTTGTGACTGATGCGGTTCATCAAGCATCCACATGCGGTG  
561 L L L D E I E K A H P D V F N I L L Q V  
1681 CTGCTGCTGGACGAAATCGAGAAAAGCGCACCCGACGTGTTCAATATTCTGTTGCAGGTG  
581 M D N G T L T D N N G R K A D F R N V V  
1741 ATGGATAACGGTACGCTGACCGATAACAACGGACGCAAAGCAGACTTCCGTAACGTGGTG  
601 L V M T T N A G V R E T E R K S I G L I  
1801 CTGGTGTGACCACCAACGCCGGGTACGGGAAACTGAGCGCAAATCCATTGGTCTTATC  
621 H Q D N S T D A M E E I K K I F T P E F  
1861 CACCAGGATAACAGCACCGATGCGATGGAAGAGATCAAGAAGATCTTTACACCGGAATTC  
641 R N R L D N I I W F D H L S T D V I H Q  
1921 CGTAACCGTCTCGACAACATTATCTGGTTTGATCATCTGTCAACCGACGTGATCCATCAG  
661 V V D K F I V E L Q V Q L D Q K G V S L  
1981 GTGGTGGATAAATTCATCGTGCAGTTGAGGTTGAGCTGGATCAGAAAAGGTGTTTCTCTG  
681 E V S Q E A R N W L A E K G Y D R A M G  
2041 GAAGTGAGCCAGGAAGCGCGTAACTGGCTGGCCGAGAAAAGGTTACGACCGGGCAATGGGC  
701 A R P M A R V I Q D N L K K P L A N E L  
2101 GCTCGTCCGATGGCGCGTGTCCATCCAGGACAACCTGAAAAACCGCTCGCCAACGAACTG  
721 L F G S L V D G G Q V T V A L D K E K N  
2161 CTGTTTGGTTTCGCTGGTGGACGGCGGTACGGTCACCGTCGCGCTGGATAAAGAGAAAAAT

741 E L T Y G F Q S A Q K H K A E A A H \*  
 2221 GAGCTGACTTACGGATTCCAGAGTGCACAAAAGCACAAAGCGGAAGCAGCGCATTA  
GCGGAAGCAGCGCATTAACTCGAGCGG

### ClpA ΔN

pET30b (NdeI, XhoI); 605 Aminosäuren; 67,0 kD; pI = 6,37;  $\epsilon_{(280 \text{ nm})} = 31525 \text{ M}^{-1} \text{ cm}^{-1}$   
 von *Escherichia coli* genomischer DNA

GGAATTCATATG  
 GGCAGCCAGCCAAAACAGCGAAGAACAAG

1 G S Q P N S E E Q A G G E E R M E N F T  
 1 GGCAGCCAGCCAAAACAGCGAAGAACAAGCTGGTGGGGAGGAACGTATGGAGAATTCACG  
 21 T N L N Q L A R V G G I D P L I G R E K  
 61 ACGAACCTGAATCAGCTTGC GCGCGTGGGCGGAATCGACCCACTGATTGGTCGTGAGAAG  
 41 E L E R A I Q V L C R R R K N N P L L V  
 121 GAGCTGGAGCGTGCTATTCAGGTTCTCTGCCGTCGCCGTAACCAACCCGCTGCTGGT  
 61 G E S G V G K T A I A E G L A W R I V Q  
 181 GGGGAATCTGGTGTGCGTAAAACCGGATTGCGGAAGGTCTTGCTGGCGAATTGTTTACG  
 81 G D V P E V M A D C T I Y S L D I G S L  
 241 GGCGATGTGCCGGAAGTGATGGCTGACTGTACGATTTACTCTCTCGATATCGGTTCTCTG  
 101 L A G T K Y R G D F E K R F K A L L K Q  
 301 TTAGCGGGCACAAAATATCGCGGCGACTTTGAAAAACGTTTTAAAGCGTTGCTCAAGCAG  
 121 L E Q D T N S I L F I D E I H T I I G A  
 361 CTGGAGCAGGACACTAACAGCATCCTGTTTATTGATGAGATCCACACCATTAATCGGTGCG  
 141 G A A S G G Q V D A A N L I K P L L S S  
 421 GGTGCAGCGTCTGGTGGTTCAGGTCGATGCGGCTAACCTAATCAAACCGTTGCTCTCCAGC  
 161 G K I R V I G S T T Y Q E F S N I F E K  
 481 GGTAATAATTCGTGTAATTGGTTCGACAACCTATCAGGAGTTCAGCAACATTTTCGAGAAA  
 181 D R A L A R R F Q K I D I T E P S I I E E  
 541 GACCGTGCTCTGGCGCGTCGCTTCCAGAAAATTGATATTACTGAACCGTCGATCGAAGAA  
 201 T V Q I I N G L K P K Y E A H H D V R Y  
 601 ACTGTTCAAATCATCAATGGCCTGAAACCGAAGTATGAAGCGCACCACGACGTGCGTTAT  
 221 T A K A V R A A V E L A V K Y I N D R H  
 661 ACCGCAAAAGCGGTGCGTGC GCGCGGTAGAGCTGGCGGTGAAATACATTAACGATCGTCAT  
 241 L P D K A I D V I D E A G A R A R L M P  
 721 CTGCCGATAAAGCCATTGATGTTATCGACGAAGCGGGCGCTCGCGCACGCTGATGCC  
 261 V S K R K K T V N V A D I E S V V A R I  
 781 GTAAGCAAACGCAAGAAAACCGTTAATGTGGCGGATATTGAGTCCGTGGTGGCCCGTATT  
 281 A R I P E K S V S Q S D R D T L K N L G  
 841 GCACGCATTCCAGAGAAGAGTGTCTCAGAGTGATCGTGATACCCCTGAAAAACCTCGGC  
 301 D R L K M L V F G Q D K A I E A L T E A  
 901 GATCGCTTGAAAATGCTGGTCTTCGGTCAGGATAAAGCCATTGAGGCGCTGACTGAAGCC  
 321 I K M A R A G L G H E H K P V G S F L F  
 961 ATTAAGATGGCGCGTGCAGGTTTAGGTCACGAACATAAACCAGTTGGTTCGTTCCGTGTT  
 341 A G P T G V G K T E V T V Q L S K A L G  
 1021 GCCGGCCCTACCGGGGTCGGGAAAACAGAGGTGACGGTACAGCTTTCGAAAAGCTTTGGGC  
 361 I E L L R F D M S E Y M E R H T V S R L  
 1081 ATTGAGCTTCTGCGCTTTGATATGTCCGAGTATATGGAACGCCATAACCGTCAGCCGTCTT  
 381 I G A P P G Y V G F D Q G G L L T D A V  
 1141 ATTGGTGC GCGCTCCGGGATACGTTGGTTTTGATCAGGGCGGTTTGCTGACTGATGCGGT  
 401 I K H P H A V L L L D E I E K A H P D V  
 1201 ATCAAGCATCCACATGCGGTGCTGCTGCTGGACGAAATCGAGAAAAGCGCACCCGGACGTT  
 421 F N I L L Q V M D N G T L T D N N G R K  
 1261 TTCAATATTCTGTTGCAGGTGATGGATAACGGTACGCTGACCGATAACAACGGACGCAAA  
 441 A D F R N V V L V M T T N A G V R E T E  
 1321 GCAGACTTCCGTAACGTGGTGTGGTATGACCACCAACGCCGGGGTACGGGAAACTGAG  
 461 R K S I G L I H Q D N S T D A M E E I K  
 1381 CGCAAATCCATTGGTCTTATCCACCAGGATAACAGCACCGATGCGATGGAAGAGATCAAG  
 481 K I F T P E F R N R L D N I I W F D H L

```

1441 AAGATCTTTACACCGGAATTCCGTAACCGTCTCGACAACATTATCTGGTTTGATCATCTG
501 S T D V I H Q V V D K F I V E L Q V Q L
1501 TCAACCGACGTGATCCATCAGGTGGTGGATAAAATTCATCGTTCGAGTTGCAGGTTTCAGCTG
521 D Q K G V S L E V S Q E A R N W L A E K
1561 GATCAGAAAGGTGTTTCTCTGGAAGTGAGCCAGGAAGCGCGTAACTGGCTGGCCGAGAAA
541 G Y D R A M G A R P M A R V I Q D N L K
1621 GGTTACGACCGGGCAATGGGCGCTCGTCCGATGGCGCGTGCATCCAGGACAACCTGAAA
561 K P L A N E L L F G S L V D G G Q V T V
1681 AAACCGCTCGCCAACGAACACTGCTGTTTGGTTTCGCTGGTGGACGGCGGTCAGGTCACCGTC
581 A L D K E K N E L T Y G F Q S A Q K H K
1741 GCGCTGGATAAAGAGAAAAATGAGCTGACTTACGGATTCCAGAGTGCACAAAAGCACAAAG
601 A E A A H *
1801 GCGGAAGCAGCGCATTAA
      GCGGAAGCAGCGCATTAACTCGAGCGG

```

### HPR-ClpA Fusion

pET30b (NdeI, XhoI); 715 Aminosäuren; 79,0 kD; pI = 6,0;  $\epsilon_{(280\text{ nm})} = 44140\text{ M}^{-1}\text{ cm}^{-1}$   
 von HIV-1 Protease in pEt30 und ClpA in pET30

```

      ACTGACTGACTGCATATG
      CCTTGCATCACTCTTTGG
1   P C I T L W Q R P L V T I K I G G Q L K
1   CCTTGCATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAATTAAG
21  E A L L D T G A D D T V L E E M N L P G
61  GAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGA
41  R W K P K M I G G I G G F I K V R Q Y D
121 AGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTATCAAAGTAAGACAGTATGAT
61  Q I L I E I C G H K A I G T V L V G P T
181 CAGATACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACA
      CAGATTGGTGCCACTTGTAATTTTCGT
      CAGATTGGTGCCACTTGTAATTTTCGT
81  P V N I I G R N L L T Q I G A T C N F R
241 CCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTGCCACTTGTAATTTTCGT
      AAAGACGAGCCGACACAGTCTTC
      AAAGACGAGCCGACACAGTCTTC
101 K D E P T Q S S D P G S Q P N S E E Q A
301 AAAGACGAGCCGACACAGTCTTCTGATCCTGGCAGCCAGCCAAACAGCGAAGAACAAGT
121 G G E E R M E N F T T N L N Q L A R V G
361 GGTGGGGAGGAACGTATGGAGAATTTACGACGAACCTGAATCAGCTTGC GCGCGTGGGC
141 G I D P L I G R E K E L E R A I Q V L C
421 GGAATCGACCCACTGATTGGTTCGTGAGAAGGAGCTGGAGCGTGCTATTCAGGTTCTCTGC
161 R R R K N N P L L V G E S G V G K T A I
481 CGTCGCCGTA AAAACAACCCGCTGCTGGTGGGGGAATCTGGTGTCCGTA AAAACCGCGATT
181 A E G L A W R I V Q G D V P E V M A D C
541 GCGGAAGGTCTTGCCTGGCGAATTGTTGAGGGCGATGTGCCGGAAGTATGGCTGACTGT
201 T I Y S L D I G S L L A G T K Y R G D F
601 ACGATTTACTCTCTCGATATCGGTTCTCTGTTAGCGGGCACAAAATATCGCGGCGACTTT
221 E K R F K A L L K Q L E Q D T N S I L F
661 GAAAAACGTTTTAAAGCGTTGCTCAAGCAGCTGGAGCAGGACACTAACAGCATCTGTTT
241 I D E I H T I I G A G A A S G G Q V D A
721 ATTGATGAGATCCACACCATTATCGGTGCGGGTGCAGCGTCTGGTGGTCAGGTCGATGCG
261 A N L I K P L L S S G K I R V I G S T T
781 GCTAACCTAATCAAACCGTTGCTCTCCAGCGGTA AAAATTCGTGTAATTGGTTCGACAACC
281 Y Q E F S N I F E K D R A L A R R F Q K
841 TATCAGGAGTTCAGCAACATTTTCGAGAAAGACCGTGCTCTGGCGCGTCCGTTCCAGAAA
301 I D I T E P S I E E T V Q I I N G L K P
901 ATTGATATTACTGAACCGTCGATCGAAGAAACTGTTCAAATCATCAATGGCCTGAAAACCG
321 K Y E A H H D V R Y T A K A V R A A V E
961 AAGTATGAAGCGCACACGACGTGCGTTATACCGCAAAAAGCGGTGCGTGCGGCGGTAGAG

```

```

341 L A V K Y I N D R H L P D K A I D V I D
1021 CTGGCGGTGAAATACATTAACGATCGTCATCTGCCGGATAAAGCCATTGATGTTATCGAC
361 E A G A R A R L M P V S K R K K T V N V
1081 GAAGCGGGCGCTCGCGCACGCCTGATGCCGGTAAGCAAACGCAAGAAAACCGTTAATGTG
381 A D I E S V V A R I A R I P E K S V S Q
1141 GCGGATATTGAGTCCGTGGTGGCCCGTATTGCACGCATTCCAGAGAAGAGTGTTCCTCAG
401 S D R D T L K N L G D R L K M L V F G Q
1201 AGTGATCGTGATAACCTGAAAAACCTCGGCGATCGCTTAAAAATGCTGGTCTTCGGTTCAG
421 D K A I E A L T E A I K M A R A G L G H
1261 GATAAAGCCATTGAGGCGCTGACTGAAGCCATTAAGATGGCGCGTGCAGTTTAGGTAC
441 E H K P V G S F L F A G P T G V G K T E
1321 GAACATAAACCGGTTGGTTCGTTCTGTTTGCCGGCCCTACCGGGTTCGGGAAAACAGAG
461 V T V Q L S K A L G I E L L R F D M S E
1381 GTGACGGTACAGCTTTTCGAAAGCTTTGGGCATTGAGCTTCTGCGCTTTGATATGCCGAG
481 Y M E R H T V S R L I G A P P G Y V G F
1441 TATATGGAACGCCATAACCGTCAGCCGTCTTATTGGTGGCCTCCGGGATACGTTGGTTTT
501 D Q G G L L T D A V I K H P H A V L L L
1501 GATCAGGGCGGTTTTGCTGACTGATGCGGTCAATCAAGCATCCACATGCGGTGCTGCTGCTG
521 D E I E K A H P D V F N I L L Q V M D N
1561 GACGAAATCGAGAAAGCGCACCCGACGTGTTCAATATTCTGTTGCAGGTGATGGATAAC
541 G T L T D N N G R K A D F R N V V L V M
1621 GGTACGCTGACCGATAACAACGGACGCAAAGCAGACTTCCGTAACGTGGTGCTGGTGATG
561 T T N A G V R E T E R K S I G L I H Q D
1681 ACCACCAACGCCGGGGTACGGGAAACTGAGCGCAAATCCATTGGTCTTATCCACCAGGAT
581 N S T D A M E E I K K I F T P E F R N R
1741 AACAGCACCGATGCGATGGAAGAGATCAAGAAGATCTTTACACCGGAATTCCGTAACCGT
601 L D N I I W F D H L S T D V I H Q V V D
1801 CTCGACAACATTATCTGGTTTTGATCATCTGTCAACCGACGTGATCCATCAGGTGGTGGAT
621 K F I V E L Q V Q L D Q K G V S L E V S
1861 AAATTCATCGTCGAGTTGCAGTTTCAGCTGGATCAGAAAAGGTGTTTCTCTGGAAGTGAGC
641 Q E A R N W L A E K G Y D R A M G A R P
1921 CAGGAAGCGGTAACCTGGCTGGCCGAGAAAGGTTACGACCGGGCAATGGGCGCTCGTCCG
661 M A R V I Q D N L K K P L A N E L L F G
1981 ATGGCGCGTGTCTATCCAGGACAACCTGAAAAACCGCTCGCCAACGAACTGCTGTTTGGT
681 S L V D G G Q V T V A L D K E K N E L T
2041 TCGCTGGTGGACGGCGGTCAGGTCACCGCTCGCGCTGGATAAAGAGAAAAATGAGCTGACT
701 Y G F Q S A Q K H K A E A A H *
2101 TACGGATTCCAGAGTGCACAAAAGCACAAAGCGGAAGCAGCGCATTAA
GCGGAAGCAGCGCATTAACTCGAGCGG

```

### Saccharopepsin-ClpA Fusion

pET30b (NdeI, XhoI); 754 Aminosäuren; 83,1 kD; pI = 5,57;  $\epsilon_{(280\text{ nm})} = 560600\text{ M}^{-1}\text{ cm}^{-1}$   
 von Saccharopepsin in pTrxFus erhalten Rickey Yada (Department of Food Science,  
 University of Guelph, Guelph, Canada; (Tanaka and Yada, 1996) und ClpA in pET30

```

GGGGTACCCATATG
GCTTACTGGGAAGTCAAGTTTGAAGG
1 A Y W E V K F E G I G L G D E Y A E L E
1 GCTTACTGGGAAGTCAAGTTTGAAGGTATCGGTTTAGGCGACGAGTACGCCGAATTGGAG
21 S H G A A I D T G T S L I T L P S G L A
61 AGCCATGGTGCCGCCATCGATACTGGTACTTCTTTGATTACCTTGCCATCAGGATTAGCT
41 E M I N A E I G A K K G W T G Q Y T L D
121 GAAATGATTAATGCTGAAATTGGGGCCAAGAAGGGTTGGACCGGTCAATATACTCTAGAC
61 C N T R D N L P D L I F N F N G Y N F T
181 TGTAACACCAGAGACAATCTACCTGATCTAATTTTCAACTTCAATGGCTACAACCTTCACT
81 I G P Y D Y T L E V S G S C I S A I T P
241 ATGGGCCATACGATTACACGCTTGAAGTTTCAGGCTCCTGTATCTCTGCAATTACACCA
101 M D F P E P V G P L A I V G D A F L R K
301 ATGGATTTCCAGAACCTGTTGGCCCACTGGCCATCGTTGGTGGTATGCCTTCTTGGCTAAA

```

GCGGTTGGTTTGGCCAAAGCAATTCGTAAA  
GCGGTTGGTTTGGCCAAAGCAATTCGTAAA

121 Y Y S I Y D L G N N A V G L A K A I R K  
361 TACTATTCTATTTACGATTTGGGCAACAATGCGGTTGGTTTGGCCAAAGCAATTCGTAAA  
GACGAGCCGACACAGTCTTC  
GACGAGCCGACACAGTCTTC

141 D E P T Q S S D P G S Q P N S E E Q A G  
421 GACGAGCCGACACAGTCTTCTGATCCTGGCAGCCAGCCAAACAGCGAAGAACAAGCTGGT  
161 G E E R M E N F T T N L N Q L A R V G G  
481 GGGGAGAACGTATGGAGAATTTACGACGAACCTGAATCAGCTTGCGCGCTGGGCGGA  
181 I D P L I G R E K E L E R A I Q V L C R  
541 ATCGACCCACTGATTGGTCGTGAGAAGGAGCTGGAGCGTGCTATTCAGGTTCTCTGCCGT  
201 R R K N N P L L V G E S G V G K T A I A  
601 CGCCGTAAAAACAACCCGCTGCTGGTGGGGAAATCTGGTGTCCGTTAAAAACCGCGATTGCC  
221 E G L A W R I V Q G D V P E V M A D C T  
661 GAAGGTCTTGCCTGGCGAATTGTTTCAGGGCGATGTGCCGGAAGTGATGGCTGACTGTACG  
241 I Y S L D I G S L L A G T K Y R G D F E  
721 ATTTACTCTCTCGATATCGGTTCTCTGTTAGCGGGCACAAAATATCGCGCGACTTTGAA  
261 K R F K A L L K Q L E Q D T N S I L F I  
781 AAACGTTTTAAAGCGTTGCTCAAGCAGCTGGAGCAGGACACTAACAGCATCCTGTTTTATT  
281 D E I H T I I G A G A A S G G Q V D A A  
841 GATGAGATCCACACCATTATCGGTGCGGGTGCAGCGTCTGGTGGTCAGGTCGATGCGGCT  
301 N L I K P L L S S G K I R V I G S T T Y  
901 AACCTAATCAAACCGTTGCTCTCCAGCGGTAAAATTCGTGTAATTGGTTCGACAACCTAT  
321 Q E F S N I F E K D R A L A R R F Q K I  
961 CAGGATTCAGCAACATTTTCGAGAAAAGACCGTGTCTGGCGCGTCTCCAGAAAAT  
341 D I T E P S I E E T V Q I I N G L K P K  
1021 GATATTACTGAACCGTGCATCGAAGAAACTGTTCAAATCATCAATGGCCTGAAACCGAAG  
361 Y E A H H D V R Y T A K A V R A A V E L  
1081 TATGAAGCGCACACGACGTGCGTTATACCGCAAAAAGCGGTGCGTGCGGCGGTAGAGCTG  
381 A V K Y I N D R H L P D K A I D V I D E  
1141 GCGGTGAAATACATTAACGATCGTCATCTGCCGATAAAAGCCATTGATGTTATCGACGAA  
401 A G A R A R L M P V S K R K K T V N V A  
1201 GCGGGCGCTCGCGCACGCCTGATGCCGGTAAGCAAACGCAAGAAAACCGTTAATGTGGCG  
421 D I E S V V A R I A R I P E K S V S Q S  
1261 GATATTGAGTCCGTGGTGGCCCGTATTGCACGCATTCCAGAGAAGAGTGTTCCTCAGAGT  
441 D R D T L K N L G D R L K M L V F G Q D  
1321 GATCGTGATACCCTGAAAAACCTCGGCGATCGCTTGAAAATGCTGGTCTTCGGTCAGGAT  
461 K A I E A L T E A I K M A R A G L G H E  
1381 AAAGCCATTGAGGCGCTGACTGAAGCCATTAAGATGGCGCGTGCAGGTTTAGGTCACGAA  
481 H K P V G S F L F A G P T G V G K T E V  
1441 CATAAACCGGTTGGTTTCCTGTTTCCCGGCCCTACCGGGGTCGGGAAAACAGAGGTG  
501 T V Q A L S K A L G G I E L L R F D M S E Y  
1501 ACGGTACAGCTTTGAAAGCTTTGGGCATTGAGCTTCTGCGCTTTGATATGTCCGAGTAT  
521 M E R H T V S R L I G A P P G Y V G F D  
1561 ATGGAACGCCATAACCGTCAGCCGTCTTATTGGTGCGCCTCCGGGATACGTTGGTTTTGAT  
541 Q G G L L T D A V I K H P H A V L L L D  
1621 CAGGGCGGTTTGTGACTGATGCGGTCAATCAAGCATCCACATGCGGTGCTGCTGCTGGAC  
561 E I E K A H P D V F N I L L Q V M D N G  
1681 GAAATCGAGAAAGCGCACCCGACGTGTTCAATATTCTGTTGCAGGTGATGGATAACGGT  
581 T L T D N N G R K A D F R N V V L V M T  
1741 ACGCTGACCGATAACAACGGACGCAAAGCAGACTTCCGTAACGTGGTGTGTTGATGACC  
601 T N A G V R E T E R K S I G L I H Q D N  
1801 ACCAACGCCGGGTACGGGAAACTGAGCGCAAATCCATTGGTCTTATCCACCAGGATAAC  
621 S T D A M E E I K K I F T P E F R N R L  
1861 AGCACCGATGCGATGGAAGAGATCAAGAAGATCTTTACACCGGAATTCGTAACCGTCTC  
641 D N I I W F D H L S T D V I H Q V V D K  
1921 GACAACATTATCTGGTTTGATCATCTGTCAACCGACGTGATCCATCAGGTGGTGGATAAA  
661 F I V E L Q V Q L D Q K G V S L E V S Q  
1981 TTCATCGTCGAGTTGCAGGTTTCAGCTGGATCAGAAAAGGTGTTTCTCTGGAAGTGAGCCAG  
681 E L A R A E K G Y D R A M G A R P M  
2041 GAAGCGCGTAACCTGGCTGGCCGAGAAAAGTTACGACCGGGCAATGGGCGCTCCGATG  
701 A R V I Q D N L K K P L A N E L L F G S

```

2101 GCGCGTGTCCATCCAGGACAACCTGAAAAAACCGCTCGCCAACGAACTGCTGTTTGGTTTCG
721 L V D G G Q V T V A L D K E K N E L T Y
2161 CTGGTGGACGGCGGTCCAGGTCACCGTCGCGCTGGATAAAGAGAAAAATGAGCTGACTTAC
741 G F Q S A Q K H K A E A A H *
2221 GGATTCCAGAGTGCACAAAAGCACAAGGCGGAAGCAGCGCATTA
      GCGGAAGCAGCGCATTAACTCGAGCGG

```

## GFP

Clontech GFP-Vektor; 249 Aminosäuren; 26,8 kD; pI =5,8;  $\epsilon_{(280 \text{ nm})} = 22015 \text{ M}^{-1} \text{ cm}^{-1}$

```

1 S K G E E L F T G V V P I L V E L D G D
1 AGTAAAGGAGAAGAACTTTTCTACTGGAGTTGTCCCAATTCTTGTGAATTAGATGGTGAT
21 V N G H K F S V S G E G E G D A T Y G K
61 GTTAATGGGCACAAATTTTCTGTCTAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAA
41 L T L K F I C T T G K L P V P W P T L V
121 CTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTC
61 T T F S Y G V Q C F S R Y P D H M K R H
181 ACTACTTTCTCTTATGGTGTTCCAATGCTTTTCAAGATACCCAGATCATATGAAACGGCAT
81 D F F K S A M P E G Y V Q E R T I F F K
241 GACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTTCAAAA
101 D D G N Y K T R A E V K F E G D T L V N
301 GATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAAGGTGATACCCCTTGTTAAT
121 R I E L K G I D F K E D G N I L G H K L
361 AGAATCGAGTTAAAGGTATTGATTTTAAAGAAGATGAAACATTCTTGGACACAAATG
141 E Y N Y N S H N V Y I M A D K Q K N G I
421 GAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATC
161 K V N F K I R H N I E D G S V Q L A D H
481 AAAGTTAACTTCAAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCAT
181 Y Q Q N T P I G D G P V L L P D N H Y L
541 TATCAACAAAATACTCCAATTGGCGATGGCCCTGTCTTTTACCAGACAACCATTACCTG
201 S T Q S A L S K D P N E K R D H M V L L
601 TCCACACAATCTGCCCTTTGCAAAGATCCCAACGAAAAGAGAGACCACATGGTCTCTTCT
221 E F V T A A G I T H G M D E L Y K *
661 GAGTTTGTAAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAAatag

```

## GFP-HPRBS

pET30b (HindIII, BamHI); 249 Aminosäuren; 28,2 kD; pI =5,8;  $\epsilon_{(280 \text{ nm})} = 22015 \text{ M}^{-1} \text{ cm}^{-1}$   
 von Clontech GFP, eingesetzt in pET30b-Konstrukt mit

CATATG AAGCTT ... GGATCC GTAAGTTTCAATTTTCCACAAATAACATTA TAA  
 CTCGAG

```

      CCC AAGCTT
      AGTAAAGGAGAAGAACTTTTCTACTGGAGTTGTCCC
1 S K G E E L F T G V V P I L V E L D G D
1 AGTAAAGGAGAAGAACTTTTCTACTGGAGTTGTCCCAATTCTTGTGAATTAGATGGTGAT
21 V N G H K F S V S G E G E G D A T Y G K
61 GTTAATGGGCACAAATTTTCTGTCTAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAA
41 L T L K F I C T T G K L P V P W P T L V
121 CTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTC
61 T T F S Y G V Q C F S R Y P D H M K R H
181 ACTACTTTCTCTTATGGTGTTCCAATGCTTTTCAAGATACCCAGATCATATGAAACGGCAT
81 D F F K S A M P E G Y V Q E R T I F F K
241 GACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTTCAAAA
101 D D G N Y K T R A E V K F E G D T L V N
301 GATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAAGGTGATACCCCTTGTTAAT
121 R I E L K G I D F K E D G N I L G H K L

```

```

361 AGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTG
141 E Y N Y N S H N V Y I M A D K Q K N G I
421 GAATACAACATAACTCACACAATGTATACATCATGGCAGACAAAACAAAAGATGGAATC
161 K V N F K I R H N I E D G S V Q L A D H
481 AAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAAGCGTTCAACTAGCAGACCAT
181 Y Q Q N T P I G D G P V L L P D N H Y L
541 TATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTG
201 S T Q S A L S K D P N E K R D H M V L L
601 TCCACACAATCGCCCTTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCTTCTT
221 E F V T A A G I T H G M D E L Y K G S V
661 GAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAAAGGATCCGTA
      GCTGGGATTACACATGGCATGGATGAACTATACAAAAGGATCCGCG
241 S F N F P Q I T L *
721 AGTTTCAATTTTCCACAAATAACATTATAA

```

### GFP-ssrA

pTrc99-His6-GFP-ssrA, Erhalten von Peter Zwickl (Max-Planck-Institut für Biochemie, München; Referenz (Amann et al., 1988))

257 Aminosäuren, 28,9 kD; pI =6,0;  $\epsilon_{(280 \text{ nm})} = 23505 \text{ M}^{-1} \text{ cm}^{-1}$

```

1 M H H H H H H M A S K G E E L F T G V V
1 ATGCATCACCATCACCATCACATGGCTAGCAAAGGAGAAGAAGCTTTTCACTGGAGTTGTC
21 P I L V E L D G D V N G H K F S V S G E
61 CCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAG
41 G E G D A T Y G K L T L K F I C T T G K
121 GGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAATTTATTTGCACTACTGGAAAA
61 L P V P W P T L V T T F S Y G V Q C F S
181 CTACCTGTTCCATGGCCAACACTTGTCACTACTTTCTTATGGTGTTCATGCTTTTCC
81 R Y P D H M K R H D F F K S A M P E G Y
241 CGTTATCCGGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTAT
101 V Q E R T I S F K D D G N Y K T R A E V
301 GTACAGGAACGCACACTATATCTTTCAAAGATGACGGGAACACTACAAGACGCGTGCTGAAGTC
121 K F E G D T L V N R I E L K G I D F K E
361 AAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAAGGTATTGATTTTAAAGAA
141 D G N I L G H K L E Y N Y N S H N V Y I
421 GATGGAAACATTCTCGGACACAAACTCGAGTACAACATAACTCACACAATGTATACATC
161 T A D K Q K N G I K A N F K I R H N I E
481 ACGGCAGACAAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACATTGAA
181 D G S V Q L A D H Y Q Q N T P I G D G P
541 GATGGATCCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
201 V L L P D N H Y L S T Q S A L S K D P N
601 GTCCTTTTACCAGACAACCATTACCTGTGACACAATCTGCCCTTTCGAAAGATCCCAAC
221 E K R D H M V L L E F V T A A G I T H G
661 GAAAAGCGTGACCACATGGTCTTCTTGAGTTTGTAAGTCTGCTGGGATTACACATGGC
241 M D E L Y K A A N D E N Y A L A A *
721 ATGGATGAGCTCTACAAAGCTGCAAACGACGAAAACACTACGCTTTAGCAGCTTAA

```

### Trap-GroEL

Erhalten von Astrid Ursinus (Max-Planck-Institut für Entwicklungsbiologie, Tübingen) 546 Aminosäuren, 57,2 kD; pI =4,85;  $\epsilon_{(280 \text{ nm})} = 10555 \text{ M}^{-1} \text{ cm}^{-1}$

```

1 M A A K D V K F G N D A R V K M L R G V
1 ATGGCAGCTAAAGACGTAAAATTCGGTAACGACGCTCGTGTGAAAATGCTGCGCGCGCTA
21 N V L A D A V K V T L G P K G R N V V L
61 AACGTACTGGCAGATGCAGTGAAAGTTACCCTCGGTCCAAAAGGCCGTAACGTAGTTCTG
41 D K S F G A P T I T K D G V S V A R E I

```

```

121 GATAAATCTTTTCGGTGCACCGACCATCACCAAAGATGGTGTTCCTGCTCGTGAATC
61 E L E D K F E N M G A Q M V K E V A S K
181 GAACTGGAAGACAAGTTCGAAAATATGGGTGCGCAGATGGTGAAAGAAGTTGCCTCTAAA
81 A N D A A G K G T T T A T V L A Q A I I
241 GCAAACGACGCTGCAGGCAAGGGTACCACCACTGCAACCGTACTGGCTCAGGCTATCATC
101 T E G L K A V A A G M N P M D L K R G I
301 ACTGAAGGTCTGAAAGCTGTTGCTGCGGGCATGAACCCGATGGACCTGAAACGTGGTATC
121 D K A V T A A V E E L K A L S V P C S D
361 GACAAGCGGTTACCGCTGCAGTTGAAGAACTGAAAGCGCTGTCCGTACCATGCTCTGAC
141 S K A I A Q V G T I S A N S D E T V G K
421 TCTAAAGCGATTGCTCAGGTTGGTACCATCTCCGCTAACTCCGACGAAACCGTAGGTAAA
161 L I A E A M D K V G K E G V I T V E D G
481 CTGATCGCTGAAGCGATGGACAAAAGTCGGTAAAAGAAGGCGTTATCACCGTTGAAGACGGT
181 T G L Q D E L D V V E G M Q F D R G Y L
541 ACCGGTCTGCAGGACGAACTGGACGTGGTTGAAGGTATGCAGTTCGACCGTGGCTACCTG
201 S P Y F I N K P E T G A V E L E S P F I
601 TCTCCTTACTTCATCAACAAGCCGAAACTGGCGCAGTAGAACTGAAAGCCCGTTCATC
221 L L A D K K I S N I R E M L P V L E A V
661 CTGCTGGCTGACAAGAAAATCTCCAACATCCGCGAAATGCTGCCGTTCTGGAAGCTGTT
241 A K A G K P L L I I A E D V E G E A L A
721 GCCAAAGCAGGCAAACCGCTGCTGATCATCGCTGAAGATGTAGAAGGCGAAGCGCTGGCA
261 T L V V N T M R G I V K V A A V K A P G
781 ACTCTGGTTGTTAACACCATGCGTGGCATCGTGAAAGTCGCTGCGGTTAAAAGCACCGGGC
281 F G D R R K A M L Q D I A T L T G G T V
841 TTCGGCGATCGTCGTAAAGCTATGCTGCAGGATATCGCAACCCCTGACTGGCGGTACCGTG
301 I S E E I G M E L E K A T L E D L G Q A
901 ATCTCTGAAGAGATCGGTATGGAGCTGGA AAAAGCAACCCCTGGAAGACCTGGGTGAGGCT
321 K R V V I N K D T T T I I D G V G E E A
961 AAACGTGTTGTGATCAACAAAGACACCACCACTATCATCGATGGCGTGGGTGAAGAAGCT
341 A I Q G R V A Q I R Q Q I E E A T S D Y
1021 GCAATCCAGGGCCGTGTTGCTCAGATCCGTCAGCAGATTGAAGAAGCAACTTCTGACTAC
361 D R E K L Q E R V A K L A G G V A V I K
1081 GACCGTGAAAACTGCAGGAACGCGTAGCGAAACTGGCAGGCGGCGTTGCAGTTATCAAA
381 V G A A T E V E M K E K K A R V E D A L
1141 GTGGGTGCTGCTACCGAAGTTGAAATGAAAGAGAAAAAAGCACGCGTTGAAGATGCCCTG
401 H A T R A A V E E G V V A G G G V A L I
1201 CACGCGACCCGTGCTGCGGTAGAAGAAGGCGTGGTTGCTGGTGGTGGTGTTCGCTGATC
421 R V A S K L A D L R G Q N E D Q N V G I
1261 CGCGTAGCGTCTAAACTGGCTGACCTGCGTGGTCAGAACGAAGACCAGAACGTGGGTATC
441 K V A L R A M E A P L R Q I V L N C G E
1321 AAAGTTGCACTGCGTGCAATGGAAGCTCCGCTGCGTCAGATCGTATTGAACTGCGGCGAA
461 E P S V V A N T V K G G D G N Y G Y N A
1381 GAACCGTCTGTTGTTGCTAACCCGTTAAAGGCGGCGACGGCAACTACGGTTACAACGCA
481 A T E Y G N M I D M G I L D P T K V T
1441 GCAACCGAAGAATACGGCAACATGATCGACATGGGTATCCTGGATCCAACCAAAGTAACT
501 R S A L Q Y A A S V A G L M I T T E C M
1501 CGTTCTGCTCTGCAGTACGCAGCTTCTGTGGCTGGCCTGATGATCACCACCGAATGCATG
521 V T D L P K N D A A D L G A A G G M G G
1561 GTTACCGACCTGCCGAAAAACGATGCAGCTGACTTAGGCGCTGCTGGCGGTATGGGCGGC
541 M G G M G G M M *
1621 ATGGGTGGCATGGGCGGCATGATGTAA

```

## PhS018

pET30b (NdeI, HindIII); 92 Aminosäuren; 10,3 kD; pI = 9,0;  $\epsilon_{(280\text{ nm})} = 3840\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus genomischer DNA von *Pyrococcus horikoshii* (ATCC 700860)

```

CATATGGATGTTTTAGCAAAAATTCCATACCACCGTTCATAGAATAGGG
1 M D V L A K F H T T V H R I G R I I I P
1 ATGGATGTTTTAGCAAAAATTCCATACAACAGTTCATAGAATAGGGAGGATCATTTATCCAA

```



```

21 A G T R K F Y G I E Q G D F V E I K I V
61 GCAGGCACTCGAAAGTTCTATGGAATAGAGCAGGGTGATTTTGTGAAATTAATAAGTA
41 K Y E G E E P K E G T F T A R V G E Q G
121 AAGTATGAAGGTGAGGAACCGAAGGAAGGTACCTTCACAGCTCGAGTAGGTGAGCAGGGT
61 S V I I P K A L R D V I G I K P G E V I
181 TCCGTAATTATTCCAAAAGCCCTAAGGGATGTCATAGGAATAAAGCCCGGAGAGGGTTATA
81 E V L L L G H Y K P R N *
241 GAGTTCTACTACTGGGTCACTATAAACCCAGGAATTGA
      ATAGAGGTTCTGCTGCTGGGTCACTATAAAACCGCGTAATTAAGCTT

```

**Af1504**

pET30b (NdeI, XhoI); 201 Aminosäuren; 22,8 kD; pI = 5,0;  $\epsilon_{(280\text{ nm})} = 17420\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus genomischer DNA von *Archaeoglobus fulgidus* (ATCC 49558)

```

CATATGTTGTACATGGACGACGAATG
 1 M L Y M D D E C Q K L L A E K E A I I R
 1 ATGTTGTACATGGACGACGAATGTCAAAAATTGCTGGCAGAGAAAAGAGGCTATTATCAGA
21 E L Q E K V K E L E A K L K S Y E I R E
61 GAACCTCAAGAAAAAGTCAAGGAGCTTGAAGCAAAGTTAAAATCCTACGAGATTAGAGAA
41 V Y K G I I P D D V L E E F V K L P P E
121 GTTTATAAAGGAATAATACCCGATGATGTTCTGGAAGAGTTCGTGAAGCTTCCACCAGAA
61 Q M I I E I G R Y L R E K G S T G Q V E
181 CAGATGATCATAGAAATAGGGAGGTATCTGAGGGAGAAAAGGGTCAACAGGACAGGTTGAG
81 A K K T V N D V R Q E I A S V E E E V S
241 GCCAAGAAAACAGTTAATGACGTCAGGCAGGAAATAGCCTCGGTTCGAGGAAGAGGTTCTCA
101 K A E K E I E K T I S T I T G A A K T K
301 AAAGCCGAAAAAGAAATTGAAAAACAATATCAACCATAACGGGTGCGGCAAGACGAAAG
121 V G V D L T F T Q K Y D Y E G S D V A F
361 GTTGGTGTGACCTCACCTTCACACAAAAGTACGATTATGAGGGAAGCGACGTTGCTTTT
141 L A E D I M N A I G V K E G E Y V S V K
421 CTGGCGGAGGACATCATGAACGCCATTGGCGTCAAAGAGGGTGTGAGTACGTGAGCGTTAAG
161 K N G T V N L R V L P Y S K E G F I V V
481 AAAAATGGAACCGTGAACCTGAGGGTTTTACCCTACTCCAAGGAGGGATTTATTGTTGTC
181 P T W V R E K L G V K V N D F V E V V R
541 CCCACGTGGGTTAGGGAGAAGCTTGGTGTTAAGGTTAACGACTTTGTGGAGGTGGTTAGG
201 R *
601 AGATGA
      AGATGACTCGAG
      GTGGAGGTGGTTAGG

```

**Af1502**

pGex 4T-1 (BamHI, XhoI); 68 Aminosäuren; 8,0 kD; pI = 4,8;  $\epsilon_{(280\text{ nm})} = 5120\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus genomischer DNA von *Archaeoglobus fulgidus* (ATCC 49558)

```

GGATCCCATATGATAACCTACAAAAAAGCTTTTAGATGAGCTC
 1 M I T Y K K L L D E L K K E I G P I A K
 1 ATGATAACCTACAAAAAAGCTTTTAGATGAGCTCAAAAAGGAAATCGGGCCGATCGCAAAA
21 I F L N K A M E S L G Y D D V D D S N Y
61 ATCTTCCTGAACAAGGCAATGGAAAAGTTTGGGCTACGATGACGTGGACGACTCAAACCTAC
41 K E I L S V L K M N K E L R E Y V E I V
121 AAGGAAATTCTCAGCGTTCTGAAAATGAACAAGGAGCTTAGGGAGTATGTCGAAAATAGTC
61 E E R L E K E G *
181 GAGGAAAGGCTTGAAAAAGAGGGTTAA
      GAGGAAAGGCTTGAAAAAGAGGGTTAACTCGAGCGG

```

**Af HAMP**

pGex 4T-1 (BamHI, XhoI); 54 Aminosäuren; 6,0 kD; pI = 5,77;  
Aus genomischer DNA von *Archaeoglobus fulgidus* (ATCC 49558)

```

GGATCC
CATATGTCGACCATTACGAGACCCATAATC
  1 S T I T R P I I E L S N T A D K I A E G
  1 TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTGCCGACAAGATTGCCGAAGGA
 21 N L E A E V P H Q N R A D E I G I L A K
 61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
 41 S I E R L R R S L K V A M E *
 121 AGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATGGAGTGA
      CCTCAAGGTCGCCATGGAGTGACTCGAGCGG

```

**Af HAMP A291V**

pGex 4T-1 (BamHI, XhoI); 54 Aminosäuren; 6,0 kD; pI = 5,77;  
Aus synthetischen Primern

```

GGATCC
CATATGTCGACCATTACGAGACCCATAATC
      ACCATTACGAGACCCATAATCGAGCTGAGCAACACTTTGGACAAGATTGCCGAAGGA
  1 S T I T R P I I E L S N T V D K I A E G
  1 TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTGGTGACAAGATTGCCGAAGGA
      GAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
      AACCTTGAGGCTGAGGTTCCGCACCAGAAC
 21 N L E A E V P H Q N R A D E I G I L A K
 61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
      AGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATG
 41 S I E R L R R S L K V A M E *
 121 AGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATGGAGTGA
      CCTCAAGGTCGCCATGGAGTGACTCGAGCGG

```

**Af HAMP A291C**

pGex 4T-1 (BamHI, XhoI); 54 Aminosäuren; 6,0 kD; pI = 5,77;  
Aus synthetischen Primern

```

GGATCC
CATATGTCGACCATTACGAGACCCATAATC
      ACCATTACGAGACCCATAATCGAGCTGAGCAACACTTGGACAAGATTGCCGAAGGA
  1 S T I T R P I I E L S N T C D K I A E G
  1 TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTTGGACAAGATTGCCGAAGGA
      GAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
      AACCTTGAGGCTGAGGTTCCGCACCAGAAC
 21 N L E A E V P H Q N R A D E I G I L A K
 61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
      AGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATG
 41 S I E R L R R S L K V A M E *
 121 AGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATGGAGTGA
      CCTCAAGGTCGCCATGGAGTGACTCGAGCGG

```

**Af HAMP A291F**

pGex 4T-1 (BamHI, XhoI); 54 Aminosäuren; 6,0 kD; pI = 5,77;

Aus genomischer DNA von *Archaeoglobus fulgidus* (ATCC 49558)

GGATCC

CATATGTCGACCATTACGAGACCCATAATC

CATAATCGAGCTGAGCAACACTTTTGACAAGATTGCCGAAGGA

CATAATCGAGCTGAGCAACACTTTTGACAAGATTGCCGAAGGA

1 S T I T R P I I E L S N T A D K I A E G

1 TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTTTTGACAAGATTGCCGAAGGA

AACC

AACC

21 N L E A E V P H Q N R A D E I G I L A K

61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG

41 S I E R L R R S L K V A M E \*

121 AGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATGGAGTGA

CCTCAAGGTCGCCATGGAGTGAAGTACTCGAGCGG

**Af1503 TM2CC, Af1503 TM2CC D88A, Af1503 TM2CC D92A**pET30b (NdeI, HindIII); 308 Aminosäuren; 34,6 kD; pI = 5,2;  $\epsilon_{(280\text{ nm})} = 48485\text{ M}^{-1}\text{ cm}^{-1}$ Aus genomischer DNA von *Archaeoglobus fulgidus* (ATCC 49558)

GGAATTCCATATG

ACATCCTCTGCTGAAGAG

1 T S S A E E A K Q G V T T V S Q E Y L T

1 ACATCCTCTGCTGAAGAGGCGAAGCAGGGCGTTACAACAGTCAGCCAGGAATACCTGACC

21 K A G E E A V R M K A Q D L A L A V Q T

61 AAGGCCGGTGAGGAAGCTGTGAGGATGAAGGCACAGGATTTGGCCTTGGCAGTTCAGACC

CAACCTTTGACTTAATT

CAACCTTTGACTTAATT

CCATGCTCACAACCTTTGCCTTAATT

CCATGCTCACAACCTTTGCCTTAATT

AF1503 TM2CC D88A

A

41 Y I E A K M K L E N K T M L T T F D L I

121 TACATTGAGGCAAAGATGAAACTCGAGAACAAAACCATGCTCACAACCTTTGACTTAATT

CAAGCCCCGAAGTTCAGAAGTCTTG

CAAGCCCCGAAGTTCAGAAGTCTTG

A

AF1503 TM2CC D92A

CAAGACCCG

CAAGACCCG

61 Q D P K F R S L G A Q R W G A K E Y T W

181 CAAGACCCGAAGTTCAGAAGTCTTGAGCACAACGCTGGGGTGCGAAGGAGTACACCTGG

81 V G A G N K V A G R D V A V I L T H P A

241 GTTGGTGCTGGTAACAAAGTTGCGGGTAGGGATGTGGCAGTTATCCTAACGCATCCCCGC

101 F T G Q Y E K Y L G V D V A M L R W N E

301 TTCACCGGACAGTACGAAAAGTACCTTGGTGTGACGTTGCGATGCTGAGATGGAACGAA

121 T M P E L Y N L L L K I T E N P E A P K

361 ACAATGCCTGAGCTTTACAACCTGCTCCTCAAATAACCGAAAACCTGAAGCTCCAAAG

141 P V C G Y Y H W D D P E T P E K E E I P

421 CCGTTTTGTGGCTATTACCACTGGGATGACCCCGAAAACCTGAAAAGGAAGAAATTCCG

161 K Y L C H Y P T T I K V Y D P I S K G Q

481 AAATATCTCTGCCACTATCCCACGACCATAAAGGTCTATGACCCGATTTCCAAGGGTCAG

181 L W V V V G T S A Y I D G Y F Q Y L T Q

541 CTGTGGGTTGTTGTGGGAACCTCAGCCTACATCGACGGTTACTTCCAGTACCTCACCCAG

GCAAGAGTGTGAGGGTGCA

201 N P A N P A E N I A S E I S K S V E G A

601 AACCCCGCCAATCCTGCAGAGAACATCGCCAGTGAGATAAGCAAGAGTGTGAGGGTGCA

```

ATTTCAGCAGGTTAAAAACCTGCTGACCCCTGGCAGCGGATCGCGCTGAACAGATTGTC
GCAGCGGATCGCGCTGAACAGATTGTCAAC
221 I Q Q V K N L L T L A A D R A E Q I V N
661 ATTCAGCAGGTTAAAAACCTGCTGACCCCTGGCAGCGGATCGCGCTGAACAGATTGTCAAC
GATCTGGCGAGCACCACAACCTCGACCATTACGAGACCCATAATCGAG
241 D L A S T T T S T I T R P I I E L S N T
721 GATCTGGCGAGCACCACAACCTCGACCATTACGAGACCCATAATCGAGCTGAGCAACACT
261 A D K I A E G N L E A E V P H Q N R A D
781 GCCGACAAGATTGCCGAAGGAAACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGAT
281 E I G I L A K S I E R L R R S L K V A M
841 GAAATCGGTATTCTTGCAAAGAGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATG
301 E S L E E A L K *
901 GAGAGCCTTGAGGAGGCACTGAAATGA
GAGGAGGCACTGAAATGAAAGCTTGGG

```

### Af1503 EP

pET30b (NdeI, HindIII); 223 Aminosäuren; 25,2 kD; pI = 5,1;  $\epsilon_{(280\text{ nm})} = 48485\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus genomischer DNA von *Archaeoglobus fulgidus* (ATCC 49558)

```

GGAATTCATATG
ACATCCTCTGCTGAAGAG
1 T S S A E E A K Q G V T T V S Q E Y L T
1 ACATCCTCTGCTGAAGAGGCGAAGCAGGGCGTTACAACAGTCAGCCAGGAATACCTGACC
21 K A G E E A V R M K A Q D L A L A V Q T
61 AAGGCCGGTGAGGAAGCTGTCAGGATGAAGGCACAGGATTTGGCCTTGGCAGTTCAGACC
41 Y I E A K M K L E N K T M L T T F D L I
121 TACATTGAGGCAAAGATGAACTCGAGAACAAAACCATGCTCACAACCTTTGACTTAATT
61 Q D P K F R S L G A Q R W G A K E Y T W
181 CAAGACCCGAAGTTCAGAAGTCTTGGAGCACAACGCTGGGGTGCGAAGGAGTACACCTGG
81 V G A G N K V A G R D V A V I L T H P A
241 GTTGGTGCTGGTAACAAAGTTGCGGGTAGGGATGTGGCAGTTATCCTAACGCATCCCCGC
101 F T G Q Y E K Y L G V D V A M L R W N E
301 TTCACCGGACAGTACGAAAAGTACCTTGGTGTGACGTTGCGATGCTGAGATGGAACGAA
121 T M P E L Y N L L L K I T E N P E A P K
361 ACAATGCCTGAGCTTTACAACCTGCTCCTCAAATAACCGAAAACCTGAAAGCTCCAAAG
141 P V C G Y Y H W D D P E T P E K E E I P
421 CCGTTTTGTGGCTATTACCACTGGGATGACCCCGAAAACACCTGAAAAGGAAGAAATCCG
161 K Y L C H Y P T T I K V Y D P I S K G Q
481 AAATATCTCTGCCACTATCCCACGACCATAAAGGTCTATGACCCGATTTCCAAGGGTCAG
181 L W V V V G T S A Y I D G Y F Q Y L T Q
541 CTGTGGGTTGTTGTGGGAACCTCAGCCTACATCGACGGTTACTTCCAGTACCTCACCCAG
GCAAGAGTGTGAGGGTGCA
GCAAGAGTGTGAGGGTGCA
201 N P A N P A E N I A S E I S K S V E G A
601 AACCCCGCAATCCTGCAGAGAACATCGCCAGTGAGATAAGCAAGAGTGTGAGGGTGCA
ATTTCAGCAGTAAGTTTACTACGCTTTGGGTATTGCAGCGG
ATTTCAGCAGTAAGTTTACTACGCTTTGGGTATTGCAGCGG
221 I Q Q *
661 ATTCAGCAGTAAGTTTACTACGCTTTGGGTATTGCAGCGGGCGCTGAACAGATTGTCAAC
721 GATCTGGCGAGCACCACAACCTCGACCATTACGAGACCCATAATCGAGCTGAGCAACACT
781 GCCGACAAGATTGCCGAAGGAAACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGAT
841 GAAATCGGTATTCTTGCAAAGAGTATCGAAAGGCTGAGAAGGAGCCTCAAGGTCGCCATG
901 GAGAGCCTTGAGGAGGCACTGAAATGA
GAGGAGGCACTGAAATGAAAGCTTGGG

```

**Af HAMP-DHP, Af HAMP A291V -DHP**

pET30b (NdeI, XhoI); 111 Aminosäuren; 13,0 kD; pI = 5,0;  $\epsilon_{(280 \text{ nm})} = 2980 \text{ M}^{-1} \text{ cm}^{-1}$   
 Aus Af HAMP bzw. Af HAMP A291V und *Escherichia coli* genomischer DNA

```

CGGGATCCCATATG
TCGACCATTACGAGACCCATAATC

                                V      HAMP-DHP A291V
1  S T I T R P I I E L S N T A D K I A E G
1  TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTGCCGACAAGATTGCCGAAGGA
                                CAAAG
                                CAAAG
21 N L E A E V P H Q N R A D E I G I L A K
61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
   AGTATCGAAAGGCTGAGAAGGAGCCTCAAGCAACTGGCGGATGACCGCACGCTG
   AGTATCGAAAGGCTGAGAAGGAGCCTCAAGCAACTGGCGGATGACCGCACGCTG
41 S I E R L R R S L K Q L A D D R T L L M
121 AGTATCGAAAGGCTGAGAAGGAGCCTCAAGCAACTGGCGGATGACCGCACGCTGCTGATG
   61 A G V S H D L R T P L T R I R L A T E M
181 GCGGGGGTAAGCCATGATCTGCGCACGCCGCTGACGCGTATTTCGTCTGGCGACCGAGATG
   81 M S E Q D G Y L A E S I N K D I E E C N
241 ATGAGCGAGCAGGATGGCTACCTGGCAGAATCGATCAATAAAGATATCGAAGAGTGCAAC
101 A I I E Q F I D Y L R *
301 GCCATCATTGAGCAGTTTATCGACTACCTGCGCTAA
   GCCATCATTGAGCAGTTTATCGACTACCTGCGCTAACTCGAGCGG
  
```

**Af HAMP-EnvZ, Af HAMP A291V-EnvZ**

pET30b (NdeI, XhoI); 272 Aminosäuren; 30,4 kD; pI = 5,3;  $\epsilon_{(280 \text{ nm})} = 23950 \text{ M}^{-1} \text{ cm}^{-1}$   
 Aus Af HAMP bzw. Af HAMP A291V und *Escherichia coli* genomischer DNA

```

CGGGATCCCATATG
TCGACCATTACGAGACCCATAATC

                                V      HAMP-ENVZ A291V
1  S T I T R P I I E L S N T A D K I A E G
1  TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTGCCGACAAGATTGCCGAAGGA
                                CAAAG
                                CAAAG
21 N L E A E V P H Q N R A D E I G I L A K
61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
   AGTATCGAAAGGCTGAGAAGGAGCCTCAAGCAACTGGCGGATGACCGCACGCTG
   AGTATCGAAAGGCTGAGAAGGAGCCTCAAGCAACTGGCGGATGACCGCACGCTG
41 S I E R L R R S L K Q L A D D R T L L M
121 AGTATCGAAAGGCTGAGAAGGAGCCTCAAGCAACTGGCGGATGACCGCACGCTGCTGATG
   61 A G V S H D L R T P L T R I R L A T E M
181 GCGGGGGTAAGCCATGATCTGCGCACGCCGCTGACGCGTATTTCGTCTGGCGACCGAGATG
   81 M S E Q D G Y L A E S I N K D I E E C N
241 ATGAGCGAGCAGGATGGCTACCTGGCAGAATCGATCAATAAAGATATCGAAGAGTGCAAC
101 A I I E Q F I D Y L R T G Q E M P M E M
301 GCCATCATTGAGCAGTTTATCGACTACCTGCGCACCGGGCAGGAGATGCCGATGGAAATG
121 A D L N A V L G E V I A A E S G Y E R E
361 GCGGATCTCAACGCAGTACTCGGTGAGGTTATTGCTGCCGAAAGTGGCTATGAGCGGGAA
141 I E T A L Y P G S I E V K M H P L S I K
421 ATTGAAACTGCGCTTTACCCCGGTAGCATTGAGGTGAAAATGCACCCGCTGTGCATCAAA
161 R A A V A N V N A A R Y G N G W I K V
481 CGCGCGGTAGCGAATATGGTGGTGAACGCCCGCCGCTATGGCAATGGCTGGATCAAAATC
181 S S G T E P N R A W F Q V E D D G P G I
541 AGCAGCGGCACGGAGCCGAATCGCGCCTGGTTTCAGGTAGAAGATGACGGTCCGGGTATT
  
```

```

201 A P E Q R K H L F Q P F V R G D S A R T
601 GCGCCAGAACAGCGTAAGCACCTGTTCCAGCCGTTTGTTCGTGGCGACAGTGCGGCGCACT
221 I S G T G L G L A I V Q R I V D N H N G
661 ATTAGCGGCACGGGATTAGGGCTGGCGATTGTGCAGCGTATCGTGGATAACCATAACGGG
241 M L E L G T S E R G G L S I R A W L P V
721 ATGCTGGAGCTTGGCACCAGCGAACGGGGCGGGCTTTCATTTCGCGCCTGGCTGCCAGTG
261 P V T R A Q G T T K E G *
781 CCGGTAACGCGGGCGCAGGGCACGACAAAAGAAGGGTAA
      GGGCGCAGGGCACGACAAAAGAAGGGTAACTCGAGCGG

```

### Af HAMP-cTsr, HAMP A291V-cTsr

pET30b (NdeI, XhoI); 340 Aminosäuren; 36,5 kD; pI = 4,77;  $\epsilon_{(280\text{ nm})} = 6990\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus Af HAMP bzw. Af HAMP A291V und *Escherichia coli* genomischer DNA

```

      CGGGATCCCATATG
      TCGACCATTACGAGACCCATAATC
                                     V      HAMP-CTSR A291V
1 S T I T R P I I E L S N T A D K I A E G
1 TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTGCCGACAAGATTGCCGAAGGA
                                     CAAAG
                                     CAAAG
21 N L E A E V P H Q N R A D E I G I L A K
61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG
   AGTATCGAAAGGCTGAGAAGGAGCCTCATGCGTACCGTCCGGTGATGTGCGCAACG
   AGTATCGAAAGGCTGAGAAGGAGCCTCATGCGTACCGTCCGGTGATGTGCGCAACG
41 S I E R L R R S L M R T V G D V R N G A
121 AGTATCGAAAGGCTGAGAAGGAGCCTCATGCGTACCGTCCGGTGATGTGCGCAACGGGGCC
61 N A I Y S G A S E I A T G N N D L S S R
181 AATGCCATCTATAGCGGTGCCAGCGAAATCGCTACCGGCAATAACGATCTCTCTTCGCGC
81 T E Q Q A A S L E E T A A S M E Q L T A
241 ACCGAGCAACAGGCCGCTTCGCTGGAAGAGACGGCAGCCAGCATGGAGCAACTCACCGCA
101 T V K Q N A E N A R Q A S H L A L S A S
301 ACGGTGAAACAGAACGCCGAGAATGCGCGCCAGGCCAGCCACCTGGCGTTAAGTGCTTCT
121 E T A Q R G G K V V D N V V Q T M R D I
361 GAAACGGCGCAACGCGGGCGCAAAGTGGTAGATAACGTGGTGCAGACTATGCGTGATATC
141 S T S S Q K I A D I I S V I D G I A F Q
421 TCCACCAGTTCGCAGAAAATCGCCGATATTATCAGCGTAATTGACGGCATTGCCTTCCAG
161 T N I L A L N A A V E A R A G G E Q G R
481 ACCAATATTCTGGCTTTGAACGCGGGTGAAGCAGCGCGTGCGGGTGAGCAAGGACGT
181 G F A V V A G E V R N L A Q R S A A A
541 GGTTTTGCGGTGGTTGCGGGAGAAGTGCGTAATCTGGCCCAGCGTAGCGCCAGGCGGCT
201 R E I K S L I E D S V G K V D V G S T L
601 CGTGAAATTAAGCCTGATTGAAGACTCGGTGGGGAAAAGTGGATGTTGGCTCTACGCTG
221 V E S A G E T M A E I V S A V T R V T D
661 GTCGAAAGCGCCGGGGAAACAATGGCGGAGATTGTCAGTGCTGTGACCCGCGTGACGGAC
241 I M G E I A S A S D E Q S R G I D Q V G
721 ATTATGGGCGAAATAGCTTCTGCTTCTGATGAGCAGAGCCGTGGTATCGATCAGTTGGC
261 L A V A E M D R V T Q Q N A A L V E E S
781 TTAGCGGTTGCTGAGATGGACCGGGTAACTCAACAGAACGCTGCGCTGGTGGAAGAGTCT
281 A A A A A A L E E Q A S R L T E A V A V
841 GCCGCTGCCGCCGCGCTTGAAGAGCAGGCCAGTCGCTGACCGAAGCTGTGGCAGTG
301 F R I Q Q Q Q Q Q R E T S A V V K T V
901 TTCCGGATTTCAGCAACAGCAGCAACAGCAGCGTGAAACATCGGCTGTGGTAAAAACCGTG
321 T P A T P R K M A V A D S G E N W E T F
961 ACGCCAGCTACGCCGCTAAAATGGCAGTGGCAGATAGCGGGGAGAACTGGGAAACGTTT
      GGCAGATAGCGGGGAGAACTGGGAAACGTTT

```

341 \*  
 1021 TAA  
 TAACTCGAGCGG

### Af HAMP-NarX, HAMP A291V-NarX

pET30b (NdeI, XhoI); 424 Aminosäuren; 48,5 kD; pI = 5,75;  $\epsilon_{(280\text{ nm})} = 26065\text{ M}^{-1}\text{ cm}^{-1}$   
 Aus Af1503 HAMP bzw. Af1503 HAMP A291V und *Escherichia coli* genomischer DNA

CGGGATCCCATATG  
 TCGACCATTACGAGACCCATAATC

V      HAMP-NARX A291V

1 S T I T R P I I E L S N T A D K I A E G  
 1 TCGACCATTACGAGACCCATAATCGAGCTGAGCAACACTGCCGACAAGATTGCCGAAGGA

CAAAG  
 CAAAG

21 N L E A E V P H Q N R A D E I G I L A K  
 61 AACCTTGAGGCTGAGGTTCCGCACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAG  
 AGTATCGAAAGGCTGAGAAGGAGCCTCGCCGAAAAGTTATGCCGTACTIONGAGCAGCG  
 AGTATCGAAAGGCTGAGAAGGAGCCTCGCCGAAAAGTTATGCCGTACTIONGAGCAGCG  
 41 S I E R L R S L A E S Y A V L E Q R V  
 121 AGTATCGAAAGGCTGAGAAGGAGCCTCGCCGAAAAGTTATGCCGTACTIONGAGCAGCGGTT  
 61 Q E K T A G L E H K N Q I L S F L W Q A  
 181 CAGGAGAAAACCGCCGGGCTGGAGCATAAAAAATCAGATCCTCTCTTTTTTATGGCAGGCT  
 81 N R R L H S R A P L C E R L S P V L N G  
 241 AACCGCCGTTTGCATTCCCGCGCCCCGCTGTGTGAACGCTGTACCTGTACTCAACGGA  
 101 L Q N L T L L R D I E L R V Y D T D D E  
 301 TTACAGAATTTAACCTGCTACGTGATATCGAACTGCGGGTGTATGACACTGATGATGAA  
 121 E N H Q E F T C Q P D M T C D D K G C Q  
 361 GAGAATCATCAGGAGTTTACCTGCCAGCCAGATATGACTTGTGATGATAAAGGCTGCCAG  
 141 L C P R G V L P V G D R G T T L K W R L  
 421 CTCTGCCCGCGCGGCTATTACCCGTTGGCGATCGCGGCACAACCCCTGAAGTGGCGGCTG  
 161 A D T H T Q Y G I L L A T L P Q G R H L  
 481 GCTGACTCATACGCAGTACGGTATTTTTGCTGGCGACCCCTGCCGAGGGGGCGTCATCTT  
 181 S H D Q Q Q L V D T L V E Q L T A T L A  
 541 AGCCATGATCAACAACAACCTGGTGGATACCCTGGTTGAACAACCTACCGCCACGCTGGCG  
 201 L D R H Q E R Q Q Q L I V M E E R A T I  
 601 CTGGATCGCCATCAGGAACGTGAGCAACAGTTGATCGTGATGGAAGAGCGTGCCACCATT  
 221 A R E L H D S I A G Q S L S C M K M Q V S  
 661 GCGCGGAACTGCATGATTCTATTGCCAATCTCTCTCTTGCATGAAGTGCAGGTGAGT  
 241 C L Q M Q G D A L P E S S R E L L S Q I  
 721 TGCTTACAGATGCAGGGCGATGCGCTGCCAGAAAAGCAGCCGGAAGTGTAAAGTCAGATC  
 261 R N E L N A S W V Q L R E L L T T F R L  
 781 CGTAACGAACTGAATGCATCCTGGGTGAGTTGCGTGAATTGCTCACCACATTCCGTTTG  
 281 Q L T E P G L R P A L E A S C E E Y S A  
 841 CAGCTCACCGAGCCTGGATTACGTCCGGCGCTGGAGGCGAGTTGCGAAGAGTACAGCGCC  
 301 K F G F P V K L D Y Q L P P R L V P S H  
 901 AAATTTGGCTTCCCGGTGAAGCTGGATTATCAATTGCCGCTCGTCTGGTGCCTTCGCAT  
 321 Q A I H L L Q I A R E A L S N A L K H S  
 961 CAGGCAATCCACTTGTGCAAATTGCCCGTGAGGCATTAAGTAACGCCCTCAAACATTCCG  
 341 Q A S E V V V T V A Q N D N Q V K L T V  
 1021 CAAGCGAGTGAGGTGCTGGTGACGGTGGCGCAAAACGATAATCAGGTCAAACCTGACCGTC  
 361 Q D N G C G V P E N A I R S N H Y G M I  
 1081 CAGGATAACGGCTGCGGGCTGCCTGAAAATGCCATCCGCAGCAATCACTACGGCATGATA  
 381 I M R D R A Q S L R G D C R V R R R E S  
 1141 ATAATGCGCGACCGTGCGCAAAGTTTACGAGGCGATTGCCGCGTCCGCGCTCGTGAATCA  
 401 G G T E V V V T F I P E K T F T D V Q G

1201 GGTGGCACCGAAGTGGTGGTCACCTTTATTCCCGAAAAAACTTTACAGACGTCCAAGGA  
CACAGACGTCCAAGGA  
 421 D T H E \*  
 1261 GATACCCATGAGTAA  
GATACCCATGAGTAACTCGAGCGG

## Tar

### pNT201

Ursprungsorganismus *Escherichia coli*, erhalten von Gerald Hazelbauer (University of Missouri, USA), ursprünglich konstruiert von Mel Simon, Division of Biology, California Institute of Technology, Pasadena, USA (Borkovich et al., 1989)

1 M I N R I R V V T L L V M V L G V F A L  
 1 ATGATTAACCGTATCCGCGTAGTCACGCTGTTGGTAATGGTGCTGGGGGTATTTCGCACTG  
 21 L Q L I S G S L F F S S L H H S Q K S F  
 61 TTACAGCTTATTTCCGGCAGTCTGTTTTTTTCTTCCCTTCACCATAGCCAGAAGAGCTTT  
 41 V V S N Q L R E Q Q G E L T S T W D L M  
 121 GTGGTTTTCCAATCAATTACGGGAACAGCAGGGCGAGCTGACGTCAACCTGGGATTTAATG  
 61 L Q T R I N L S R S A V R M M M D S S N  
 181 CTGCAAACGCGCATTAACTGAGTCGTTTCAGCGGTACGGATGATGATGGATTCTCCAAT  
 81 Q Q S N A K V E L L D S A R K T L A Q A  
 241 CAACAAAGTAACGCCAAAGTTGAATTGCTCGATAGCGCCAGGAAAAACATTGGCGCAGGCA  
 101 A T H Y K K F K S M A P L P E M V A T S  
 301 GCGACGCATTATAAAAAATTCAAAGCATGGCACCGTTACCTGAAATGGTCGCTACCAGT  
 121 R N I D E K Y K N Y Y T A L T E L I D Y  
 361 CGTAATATTGATGAAAAATATAAAAACTATTACACAGCGTTAACTGAACTGATTGATTAT  
 141 L D Y G N T G A Y F A Q P T Q G M Q N A  
 421 CTAGATTATGGCAATACTGGAGCTTATTTTCGCTCAGCCAACCCAGGGAATGCAAAATGCA  
 161 M G E A F A Q Y A L S S E K L Y R D I V  
 481 ATGGGCGAAGCGTTTTGCTCAGTACGCCCTCAGCAGTGAAAAACTGTATCGCGATATCGTC  
 181 T D N A D D Y R F A Q W Q L A V I A L V  
 541 ACTGACAACGCAGATGATTACCGATTTGCCAGTGGCAACTGGCGGTTATCGCGCTGGTG  
 201 V V L I L V A W Y G I R R M L L T P L  
 601 GTGGTATTGATTCTGCTGTTGGTGGCGTGGTACGGCATTCCGCCGTATGTTGCTTCCGCTG  
 221 A K I I A H I R E I A G G N L A N T L T  
 661 GCAAAAATTATTGCTCACATTCGCGAAATCGCCGGTGGTAACCTGGCGAATACCTTGACC  
 241 I D G R S E M G D L A Q S V S H M Q R S  
 721 ATTGACGGGCGCAGTGAAATGGGCGACCTGGCGCAGAGCGTTTTACATATGCAACGCTCT  
 261 L T D T V T H V R E G S D A I Y A G T R  
 781 TTGACTGACACCGTCACTCATGTCCGCGAAGGTTTCAGATGCCATCTATGCCGGTACCCGT  
 281 E I A A G N T D L S S R T E Q Q A S A L  
 841 GAAATTGCGGCGGGCAACACCGATCTTTCTCCCGTACTGAACAGCAGGCATCCGCGCTG  
 301 E E T A A S M E Q L T A T V K Q N A D N  
 901 GAAGAAACTGCCGCCAGCATGGAGCAGCTCACCGCGACAGTGAAGCAAAACGCCGATAAC  
 321 A R Q A S Q L A Q S A S D T A Q H G G K  
 961 GCCCGCCAGGCCTCGCAACTGGCGCAAAGTGCTCCGACACCGCCAGCACGGCGGCAAAA  
 341 V V D G V V K T M H E I A D S S K K I A  
 1021 GTGGTGGATGGCGTAGTGAAAACGATGCATGAGATCGCCGATAGTTTCAAGAAAAATTGCC  
 361 D I I S V I D G I A F Q T N I L A L N A  
 1081 GACATTATCAGCGTTATCGACGGTATTGCCTTCCAGACTAATATCCTCGCGCTGAATGCC  
 381 A V E A R E A G E A Q G R G F A V A G E  
 1141 GCGGTTGAAGCCGCGCTGCGGGTGAACAGGGCCCGTGGTTTTGCCGTGGTGGCGGGTGAA  
 401 V R N L A S R S A Q A A K E I K A L I E  
 1201 GTGCGTAATCTTGCCAGTCGACGCGCCAGGCGGCAAAAAGAGATCAAAGCCCTCATTGAA  
 421 D S V S R V D T G S V L V E S A G E T M  
 1261 GACTCCGTCTCACGCGTTGATAACCGGTTCCGGTGCTGGTGCAAAAGCGCCGGGGAAACAATG  
 441 N N I V N A V T R V T D I M G E I A S A  
 1321 AACAAATATCGTCAATGCTGTCACTCGCGTACTGACATTATGGGCGAGATTGCATCGGCA  
 461 S D E Q S R G I D Q V A L A V S E M D R



1381 TCGGATGAACAGAGCCGTGGCATCGATCAAGTCGCATTGGCGGTTTCGGAAATGGATCGC  
 481 V T Q Q N A S L V Q E S A A A A A A L E  
 1441 GTCACGCAACAGAACGCATCGCTGGTGCAGGAATCAGCTGCCGCCGCCGCTGCGCTGGAA  
 501 E Q A S R L T Q A V S A F R L A A S P L  
 1501 GAACAGGCGAGTCTGTTTAAACGCAAGCAGTTTCCGCGTTCCGTCTGGCAGCCAGCCCCTC  
 521 T N K P Q T P S R P A S E Q P P A Q P R  
 1561 ACCAATAAACCGCAAACACCATCCCCTGCTGCCAGTGAGCAACCACCGGCTCAGCCACGA  
 541 L R I A E Q D P N W E T F \*  
 1621 CTGCGAATTGCTGAACAAGATCCAAACTGGGAAACATTTTGA

## Tar AfHAMP

Von pNT201 und Af HAMP bzw. Af HAMP A291V und Af HAMP A291C

ATGGTAGGTCTCAA  
 ATGATTAACCGTATCCGCGTAGTCACGC  
 1 M I N R I R V V T L L V M V L G V F A L  
 1 ATGATTAACCGTATCCGCGTAGTCACGCTGTTGGTAATGGTGTGGGGTATTTCGCACTG  
 21 L Q L I S G S L F F S S L H H S Q K S F  
 61 TTACAGCTTATTTCCGGCAGTCTGTTTTTTTTCTTCCCTTACCATAGCCAGAAGAGCTTT  
 41 V V S N Q L R E Q Q G E L T S T W D L M  
 121 GTGGTTTCCAATCAATTACGGGAACAGCAGGGCGAGCTGACGTCAACCTGGGATTTAATG  
 61 L Q T R I N L S R S A V R M M M D S S N  
 181 CTGCAAACGCGCATTAACTGAGTCGTTTACGCGGTACGGATGATGATGGATTCTCCAAT  
 81 Q Q S N A K V E L L D S A R K T L A Q A  
 241 CAACAAAGTAACGCCAAAGTTGAATTGCTCGATAGCGCCAGGAAAACATTGGCGCAGGCA  
 101 A T H Y K K F K S M A P L P E M V A T S  
 301 GCGACGCATTATAAAAAATTCAAAAGCATGGCACCGTTACCTGAAAATGGTCGCTACCAGT  
 121 R N I D E K Y K N Y Y T A L T E L I D Y  
 361 CGTAATATTGATGAAAAATATAAAAACTATTACACAGCGTAACTGAACTGATTGATTAT  
 141 L D Y G N T G A Y F A Q P T Q G M Q N A  
 421 CTAGATTATGGCAATACTGGAGCTTATTTGCTCAGCCAACCCAGGGAATGCAAAATGCA  
 161 M G E A F A Q Y A L S S E K L Y R D I V  
 481 ATGGGCGAAGCGTTTGTCTCAGTACGCCCTCAGCAGTGAAAACTGTATCGCGATATCGTC  
 181 T D N A D D Y R F A Q W Q L A V I A L V  
 541 ACTGACAACGCAGATGATTACCGATTTGCCAGTGGCAACTGGCGGTTATCGCGCTGGTG  
 CGTGGTACGGCATTTCGCCGTATGATTACAAGACCC  
 GCATTTCGCCGTATGATTACAAGACCCATA  
 201 V V L I L L V A W Y G I R R M I T R P I  
 601 GTGGTATTGATTCTGCTGGTGGCGTGGTACGGCATTTCGCCGTATGATTACAAGACCCATA  
 C TAR AFHAMP A291C  
 V TAR AFHAMP A291V  
 ATCGAGCTGAGC  
 221 I E L S N T A D K I A E G N L E A E V P  
 661 ATCGAGCTGAGCAACACTGCCGACAAGATTGCCGAAGGAAAACCTTGAGGCTGAGGTTCCG  
 GTATCGAAAAGGCTGAGAAGG  
 GAAGG  
 241 H Q N R A D E I G I L A K S I E R L R R  
 721 CACCAGAACAGGGCTGATGAAATCGGTATTCTTGCAAAGAGTATCGAAAAGGCTGAGAAG  
 AGCCTCACTGACACCGTC  
 AGCCTCACTGACACCGTCACTCATGTCCGC  
 261 S L T D T V T H V R E G S D A I Y A G T  
 781 AGCCTCACTGACACCGTCACTCATGTCCGCGAAGGTTTCAGATGCCATCTATGCCGGTACC  
 281 R E I A A G N T D L S S R T E Q Q A S A  
 841 CGTGAAATTGCGGCGGGCAACACCGATCTTCTCCCGTACTGAACAGCAGGCATCCGCG  
 301 L E E T A A S M E Q L T A T V K Q N A D  
 901 CTGGAAGAACTGCCGCCAGCATGGAGCAGCTCACC GCGACAGTGAAGCAAAACGCCGAT  
 321 N A R Q A S Q L A S D T A Q H G G  
 961 AACGCCCGCCAGGCTCGCAACTGGCGCAAAGTGCCTCCGACACCGCCAGCAGCGGCGGC  
 341 K V V D G V V K T M H E I A D S S K K I  
 1021 AAAGTGGTGGATGGCGTAGTGAAAACGATGCATGAGATCGCCGATAGTTCGAAGAAAATT

```

361 A D I I S V I D G I A F Q T N I L A L N
1081 GCCGACATTATCAGCGTTATCGACGGTATTGCCCTTCCAGACTAATATCCTCGCGCTGAAT
381 A A V E A A R A G E Q G R G F A V V A G
1141 GCCGCGGTTGAAGCCGCGCGTGC GG GTGAACAGGGCCGTGGT TTTGCCGTGGTGGCGGGT
401 E V R N L A S R S A Q A A K E I K A L I
1201 GAAGTGC GTAATCTTGCCAGTCGCAGCGCCCAGGCGGCAAAAAGAGATCAAAGCCCTCATT
421 E D S V S R V D T G S V L V E S A G E T
1261 GAAGACTCCGTCTCACGCGTTGATACCGGTTCCGGTGCTGGTCGAAAAGCGCCGGGGAAACA
441 M N N I V N A V T R V T D I M G E I A S
1321 ATGAACAATATCGTCAATGCTGTCACTCGCGTGACTGACATTATGGGCGAGATTGCATCG
461 A S D E Q S R G I D Q V A L A V S E M D
1381 GCATCGGATGAACAGAGCCGTGGCATCGATCAAGTCGCATTGGCGGTTTCGGAAAATGGAT
481 R V T Q Q N A S L V Q E S A A A A A A L
1441 CGCGTCACGCAACAGAACGCATCGCTGGTGCAGGAATCAGCTGCCGCCGCCGCTGCGCTG
501 E E Q A S R L T Q A V S A F R L A A S P
1501 GAAGAACAGGCGAGTCGTTTAACGCAAGCAGTTTCCGCGTTCGCTCTGGCAGCCAGCCCA
521 L T N K P Q T P S R P A S E Q P P A Q P
1561 CTCACCAATAAACCGCAAACACCATCCCGTCCTGCCAGTGAGCAACCACCGGCTCAGCCA
541 R L R I A E Q D P N W E T F *
1621 CGACTGCGAATTGCTGAACAAGATCCAAACTGGGAAAACATTTTGA
      GCTGAACAAGATCCAAACTGGGAAAACATTTTGAAGCGCTGAGACCTACCAT

```

## 7.2 Bemerkung zur Proteinreinigung

Hier werden Arbeitsschritte erläutert, die bei allen Proteinreinigungen gleich ausgeführt wurden.

Die Expressionskonstrukte wurden mit molekularbiologische Standardtechniken hergestellt (Sambrook and Russell, 2001). Die Korrektheit der Nukleotidsequenzen wurde durch DNA Sequenzierung (BigDye Terminator Sequencing, Applied Biosystems) überprüft. Nach der induzierten Überexpression wurden die Zellen durch Zentrifugation geerntet: Sorvall RC-6 Zentrifuge mit einem GS-3 Rotor und 5000 Umdrehungen/min für 20 Minuten bei 4° C. Die Zellen wurden bis zur Lyse bei -80° C aufbewahrt. Zur Lyse wurden sie in Lysepuffer resuspendiert und das Volumen des Lysepuffers in ml entsprach dem vierfachen des Nassgewichtes der Zellen in g (meist 3-5 g). Zur Inhibition von Serin- und Cysteinproteasen wurde 1 mM PMSF zugegeben. Um die Viskosität durch DNA herabzusetzen, wurde der Zellsuspension 1 mM MgCl<sub>2</sub> und 10 µg/ml DNase zugesetzt. Die Zellsuspension wurde durch viermaliges passieren einer French Press Zelle mit 1000 psi lysiert. Ab diesem Zeitpunkt der Reinigung wurde bei 4-8° C gearbeitet. Unlösliches Material wurde vom Überstand durch Zentrifugation getrennt: Sorvall RC-6 Zentrifuge mit einem SS34 Rotor und 15000 Umdrehungen/min für 30 Minuten. Dem Überstand wurde 5 mM EDTA zugegeben, um Metalloproteasen zu inhibieren. Für alle chromatographischen Arbeitsschritte wurden FPLC-Systeme verwendet: Pharmacia-FPLC oder Äkta FPLC von GE Healthcare. Alle FPLC-Puffer wurden vor Gebrauch filtriert (0,22 µm) und entgast. Alle Proben wurden vor dem Auftrag auf eine Chromatographiesäule filtriert (0,22 oder 0,45 µm). Fraktionen mit Zielprotein wurden über die entsprechende Proteinbande auf einem SDS-PAGE Gel identifiziert und mit der induzierbaren Überexpressionsbande aus der

Gesamtproteinmischung von *E. coli* verglichen. Für die Proteinmengenbestimmung über die Absorption bei 280 nm wurden die Extinktionskoeffizienten nach der Methode von Stanley Gill und Peter von Hippel berechnet (Gill and Hippel, 1989). Wenn die Proteinmenge mit dem Bicinchoninsäureassay bestimmt wurde, dann wurde 50 µl Probe mit 1 ml Assaylösung (Bicinchoninsäure-Lösung: CuSO<sub>4</sub>-Lösung im Verhältnis 50:1) gemischt, 30 Minuten bei 37°C inkubiert und die Lichtabsorption wurde photometrisch bei 562 nm gemessen. Als Eichproteine dienten der bovine pankreatische Trypsininhibitor und BSA.

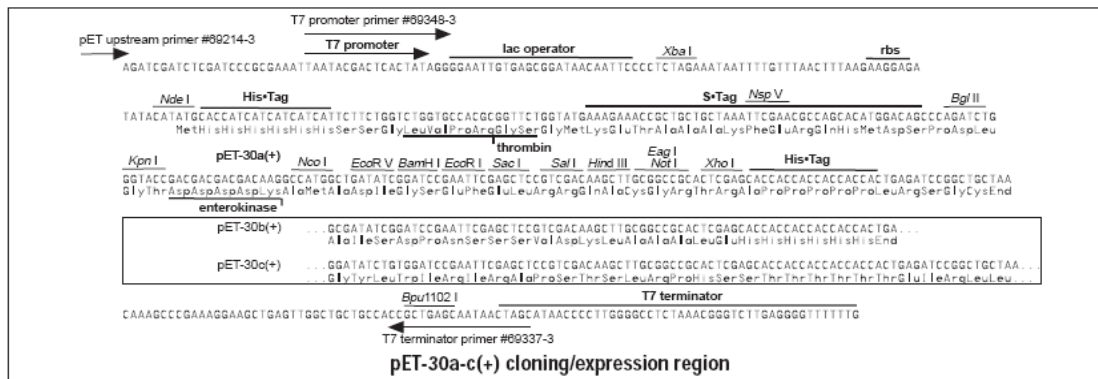
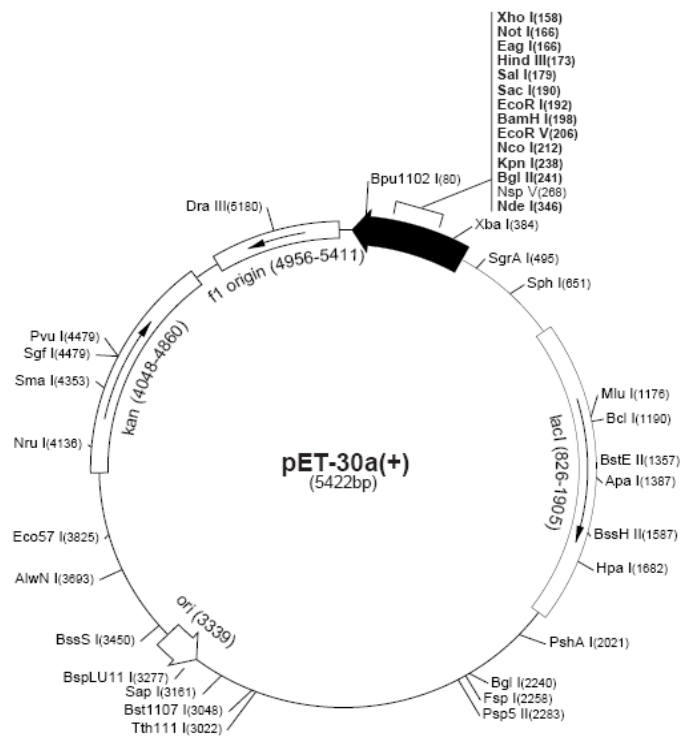
### 7.3 Vektorkarten

#### Vektorkarte pET30

**pET-30a(+) sequence landmarks**

T7 promoter	419-435
T7 transcription start	418
His* Tag coding sequence	327-344
S* Tag coding sequence	249-293
Multiple cloning sites	
( <i>Nco</i> I - <i>Xho</i> I)	158-217
His* Tag coding sequence	140-157
T7 terminator	26-72
<i>lac</i> I coding sequence	826-1905
pBR322 origin	3339
Kan coding sequence	4048-4860
f1 origin	4956-5411

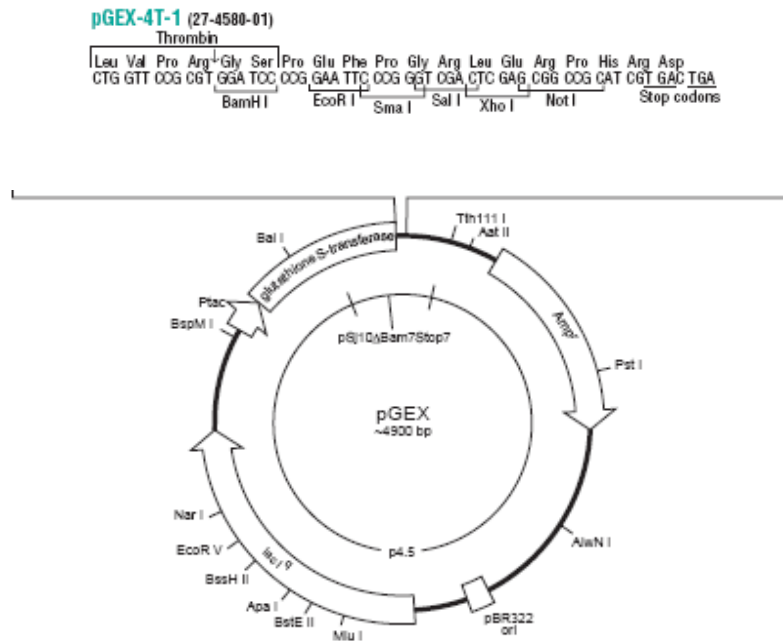
The maps for pET-30b(+) and pET-30c(+) are the same as pET-30a(+) (shown) with the following exceptions: pET-30b(+) is a 5421bp plasmid; subtract 1bp from each site beyond *Bam*HI at 198. pET-30c(+) is a 5423bp plasmid; add 1bp to each site beyond *Bam*HI at 198.



Quelle: Novagen "pETSystem Manual"

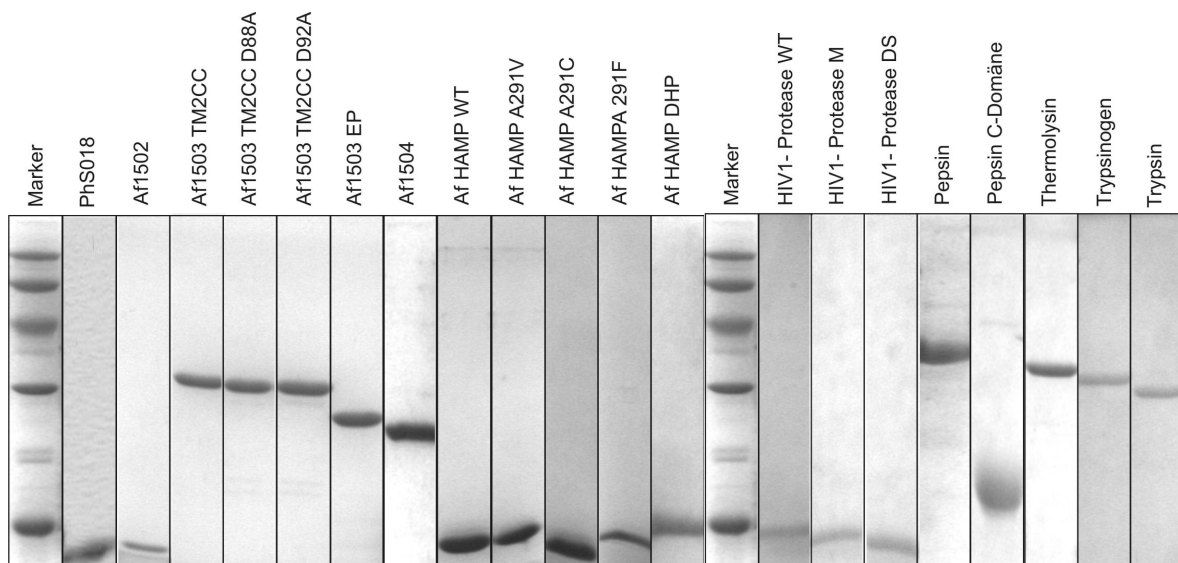
## Vektorkarte pGex 4T-1

Quelle: Amersham Bioscience "GST Gene Fusion System Handbook"



## 7.4 Reinheitskontrollen

Die Reinheit von exprimierten Proteinen, mit denen Ergebnisse in weiteren Experimenten erzielt wurden, ist mit SDS-PAGE geprüft worden. Abgebildet sind Gelspuren mit Molekülmassenmarker (97 kD, 66 kD, 45 kD, 30 kD, 20 kD, 14 kD).



## 7.5 Genotypen verwendeter E. coli Stämme

B121(DE3) [*E. coli* F<sup>-</sup> *lon ompT hsdS<sub>B</sub>* (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal dcm* λ (DE3)]

B121(DE3) Gold [*E. coli* Hte F<sup>-</sup> *lon ompT hsdS<sub>B</sub>* (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal dcm+* λ (DE3) *endA Tet<sup>r</sup>*]

C41(DE3) stammen von B121 (DE3) und tragen mindestens eine uncharakterisierte Mutation, welche die Überexpression von Membranproteinen oder toxischen löslichen Proteinen begünstigt.

CP362 *E. coli* (Δ(*tsr*)-7028 Δ(*tar-tap*) 5201 Δ(*trg*)-100 his leu rpsL ara lac thi tonA tsx xyl zbd:Tn5)

B834(DE3) [*E. coli* F<sup>-</sup> *lon ompT hsdS<sub>B</sub>* (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal dcm met* (DE3)]

KY2266 *E. coli* basiert auf MC4100 (F<sup>-</sup> *araD139* Δ(*argF-lac*)U169\* *rspL150 relA1 flbB5301 fruA25*‡ *deoC1 ptsF25 e14-*) mit zusätzlich Δ*lon-*, *clp-*, *hsIVU-* (Tc, Cm)

## 7.6 Rezepturen

### Polyacrylamid-Gele

Komponente	Sammelgel	10% Nativ	15% SDS
Wasser	4,9	9,5	4,2
Puffer	2,5	7,5	7,5
Acrylamid (37,5:1)	1,5	10	15
SDS (10%)	0,1 (nativ: 0)	0	0,3
Temed (2%)	0,5	1,5	1,5
Ammoniumpersulfat (1,4%)	0,5	1,5	1,5
Gesamt	10	30	30

Alle Angaben in [ml]

Puffer für Sammelgel: 0,5 M Tris-HCl, pH 6,8

Puffer für Trenngel: 1,5 M Tris, pH 8,8

### LB-Medium

10 g/l Trypton, 5 g/l Hefeextrakt, 5 g/l NaCl, pH 7,0

**M9 Minimalmedium**

Konzentration	Komponente
6 g/l	Na <sub>2</sub> HPO <sub>4</sub>
3 g/l	KH <sub>2</sub> PO <sub>4</sub>
0,5 g/l	NaCl
0,7 g/l	NH <sub>4</sub> Cl
2 mM	MgSO <sub>4</sub>
10 µM	CaCl <sub>2</sub>
10 µM	FeSO <sub>4</sub> (frisch)
0,2 %	Glucose
je 0,5 mg/l	Choline, Folsäure, Pantothensäure, Nicotinamid, Myo-Inositol, Pyridoxal, Thiamin, Riboflavin, Biotin
je 1 nM	Molybdat, Borsäure, CoCl <sub>2</sub> , CuSO <sub>4</sub> , MnCl <sub>2</sub> , ZnSO <sub>4</sub>

**7.7 Materialiennachweise**

Chemikalie	Hersteller
Standardchemikalien (NaCl, KCl, Na <sub>2</sub> HPO <sub>4</sub> , NaH <sub>2</sub> PO <sub>4</sub> , KH <sub>2</sub> PO <sub>4</sub> , NH <sub>4</sub> Cl, MgSO <sub>4</sub> , CaCl <sub>2</sub> , FeSO <sub>4</sub> , Glucose, Choline, Folsäure, Pantothensäure, Nicotinamid, Myo-Inositol, Pyridoxal, Thiamin, Riboflavin, Biotin, Molybdat, Borsäure, CoCl <sub>2</sub> , CuSO <sub>4</sub> , MnCl <sub>2</sub> , ZnSO <sub>4</sub> , Ammoniumpersulfat, SDS, EDTA)	Merck
MOPS, HEPES, DTT, Kanamycin, Ampicillin, Guanidinium chlorid, Harnstoff, Ammoniumsulfat, (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Applichem
TRIS, Acrylamid (37,5:1), TEMED, Glutaraldehyd, Glycerol	Roth
Lysozym (EC 3.2.1.17), Mitochondriale Citratsynthase aus Schweineherzen (EC 4.1.3.7), bovine adrenale Rhodanese (EC 2.8.1.1), Pepsin A aus Schweinen (EC 3.4.23.1), Thermolysin (EC 3.4.24.27), DNase, BSA, Eichproteine für Gelfiltration, ATP, Phosphocreatin, Phosphocreatinkinase, PMSF, Ampicillin, Kanamycin	Sigma
<sup>15</sup> NH <sub>4</sub> Cl, U- <sup>13</sup> C-Glucose	Eurisotop
IPTG	Gerbu
Bicinchoninsäure-Lösung und alkalische CuSO <sub>4</sub> -Lösung	Pierce
Low-molecular weight marker für SDS-PAGE	Amersham
	Bioscience
Genomische DNA von <i>A. fulgidus</i> und <i>P. horikoshii</i>	LGC Promochem
Luciferase (EC 1.13.12.7)	Promega
Proteinase K, Restriktionsenzyme, T4 DNA Ligase,	Fermentas

---

Deoxyribonukleotid-Mix für PCR	
DNA Phusion Polymerase	Finnzymes
DNA Taq Polymerase, T4 DNA Ligase	NEB
Pepstatin und Ac-Leu-Val-Phe-Aldehyd HIV-1 protease Inhibitor	Bachem
Thrombin	Calbiochem
Enterokinase	Roche
Vivaspin concentrator	Viviscience
Bottle-top Filter 0,22 µm	Millipore
Spritzenfilter 0,45 µm und 0,22 µm	Schleicher und Schüll
Dialysemembranen	Spetra/Por
Miniprep-Kit, PCR-Purification-Kit, Gel-Extraction-Kit	Qiagen
Primer	Operon, MWG Biotech, Sigma Genosys

---

## 7.8 Internetchweise

MPI-Toolkit	<a href="http://protevo.eb.tuebingen.mpg.de/toolkit/">http://protevo.eb.tuebingen.mpg.de/toolkit/</a>
HHpred	<a href="http://protevo.eb.tuebingen.mpg.de/toolkit/index.php?view=hhpred">http://protevo.eb.tuebingen.mpg.de/toolkit/index.php?view=hhpred</a>
HHsenser	<a href="http://protevo.eb.tuebingen.mpg.de/toolkit/index.php?view=hhsenser">http://protevo.eb.tuebingen.mpg.de/toolkit/index.php?view=hhsenser</a>
Modeller	<a href="http://protevo.eb.tuebingen.mpg.de/toolkit/index.php?view=modeller">http://protevo.eb.tuebingen.mpg.de/toolkit/index.php?view=modeller</a>
Coils	<a href="http://www.ch.embnet.org/software/COILS_form.html">http://www.ch.embnet.org/software/COILS_form.html</a>
Rosetta <i>ab initio</i> Internetch Server	Max-Planck-Institut für Entwicklungsbiologie, Tübingen
SMART	<a href="http://smart.embl-heidelberg.de/">http://smart.embl-heidelberg.de/</a>
Transport Classification Database	<a href="http://www.tcdb.org/">http://www.tcdb.org/</a>
RCSB Protein Data Bank	<a href="http://www.pdb.org/">http://www.pdb.org/</a>

## 8 Danksagung

Ich danke ...

Andrei Lupas für die Möglichkeit, meine Doktorarbeit in seiner Abteilung durchführen zu können. Ich danke ihm für die interessante Fragestellung und die exzellente wissenschaftliche Betreuung.

Jörg Martin für die Organisation des Laborbetriebs und die praktische Unterstützung.

Gabriele Dodt für die Vertretung meiner Arbeit in der Fakultät für Chemie und Pharmazie der Universität Tübingen

den Kooperationspartnern Murray Coles, Ines Wanke, Kornelius Zeth, Markus Gruber, Johannes Söding, Hans-Peter Klenk, Joachim Schultz, Jürgen Linder und Franziska Berndt

Felix-Müller Sarnowski, Diethard Mattanovich, Peter Zwickl, Astrid Ursinus, Rickey Yada, Gerald Hazelbauer und Angela Lilly für die Bereitstellung von Plasmid-DNA und *E. coli*-Stämmen

den Kollegen Indronil Chaudhuri, Birte Hernandez-Alvarez, Astrid Ursinus, Dirk Linke, Markus Gruber, Sergej Djuranovich, Manjunatha Karpenahalli, Vikram Alva, Sasha Shkumatov, Reinhard Albrecht, Moritz Ammelburg, Thomas Arnold, Kerstin Bär, Alexander Diemand, Silvia Deiss, Felipe Figuera, Julia Fischer, Tancred Frickey, Michael Gebert, Cedric Hobel, Birte Höcker, Nora Hornig, Karin Klein, Jürgen Lassak, Oliver Schmidt, Justyna Serek-Heuberger, Johannes Söding, Pawel Szczesny, Yakov Sergeev, Christoph Thomas, Vincent Truffault, Ines Wanke, Sonja Weinert, Yuen-Tsu Nicco Yu, Kornelius Zeth und Karin Lehmann



## 9 Literatur

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

Amann, E., Ochs, B., and Abel, K. J. (1988). Tightly Regulated Tac Promoter Vectors Useful for the Expression of Unfused and Fused Proteins in *Escherichia-Coli*. *Gene* 69, 301-315.

Anantharaman, V., and Aravind, L. (2000). Cache - a signaling domain common to animal Ca<sup>2+</sup> channel subunits and a class of prokaryotic chemotaxis receptors. *Trends in Biochemical Sciences* 25, 535-537.

Apic, G., Gough, J., and Teichmann, S. A. (2001). Domain combinations in archaeal, eubacterial and eukaryotic proteomes. *Journal of Molecular Biology* 310, 311-325.

Appleman, J. A., Chen, L. L., and Stewart, V. (2003). Probing conservation of HAMP linker structure and signal transduction mechanism through analysis of hybrid sensor kinases. *Journal of Bacteriology* 185, 4872-4882.

Appleman, J. A., and Stewart, V. (2003). Mutational analysis of a conserved signal-transducing element: the HAMP linker of the *Escherichia coli* nitrate sensor NarX. *Journal of Bacteriology* 185, 89-97.

Aravind, L., and Ponting, C. P. (1999). The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *Fems Microbiology Letters* 176, 111-116.

Bogard, L. D., and Deem, M. W. (1999). A hierarchical approach to protein molecular evolution. *Proceedings of the National Academy of Sciences of the United States of America* 96, 2591-2595.

Borkovich, K. A., Kaplan, N., Hess, J. F., and Simon, M. I. (1989). Transmembrane Signal Transduction in Bacterial Chemotaxis Involves Ligand-Dependent Activation of Phosphate Group Transfer. *Proceedings of the National Academy of Sciences of the United States of America* 86, 1208-1212.

Butler, S. L., and Falke, J. J. (1998). Cysteine and disulfide scanning reveals two amphiphilic helices in the linker region of the aspartate chemoreceptor. *Biochemistry* 37, 10746-10756.

Castillo, R. M., Mizuguchi, K., Dhanaraj, V., Albert, A., Blundell, T. L., and Murzin, A. G. (1999). A six-stranded double-psi beta barrel is shared by several protein superfamilies. *Structure* 7, 227-236.

Chervitz, S. A., and Falke, J. J. (1996). Molecular mechanism of transmembrane signaling by the aspartate receptor: A model. *Proceedings of the National Academy of Sciences of the United States of America* 93, 2545-2550.

Chothia, C., Gough, J., Vogel, C., and Teichmann, S. A. (2003). Evolution of the protein repertoire. *Science* 300, 1701-1703.

- Cochran, A. G., and Kim, P. S. (1996). Imitation of *Escherichia coli* aspartate receptor signaling in engineered dimers of the cytoplasmic domain. *Science* *271*, 1113-1116.
- Coles, M., Diercks, T., Liermann, J., Groger, A., Rockel, B., Baumeister, W., Koretke, K. K., Lupas, A., Peters, J., and Kessler, H. (1999). The solution structure of VAT-N reveals a 'missing link' in the evolution of complex enzymes from a simple beta alpha beta beta element. *Current Biology* *9*, 1158-1168.
- Coles, M., Djuranovic, S., Soding, J., Frickey, T., Koretke, K., Truffault, V., Martin, J., and Lupas, A. N. (2005). AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi beta barrels. *Structure* *13*, 919-928.
- Crick, F. H. C. (1953). The Packing of Alpha-Helices - Simple Coiled-Coils. *Acta Crystallographica* *6*, 689-697.
- Cui, Y., Wong, W. H., Bornberg-Bauer, E., and Chan, H. S. (2002). Recombinatoric exploration of novel folded structures: A heteropolymer-based model of protein evolutionary landscapes. *Proceedings of the National Academy of Sciences of the United States of America* *99*, 809-814.
- Dreveny, I., Kondo, H., Uchiyama, K., Shaw, A., Zhang, X. D., and Freemont, P. S. (2004). Structural basis of the interaction between the AAA ATPase p97/VCP and its adaptor protein p47. *Embo Journal* *23*, 1030-1039.
- Dunn, B. M., Goodenow, M. M., Gustchina, A., and Wlodawer, A. (2002). Retroviral proteases. *Genome Biol* *3*, REVIEWS3006.
- Dunn, B. M., Gustchina, A., Wlodawer, A., and Kay, J. (1994). Subsite Preferences of Retroviral Proteinases, In *Retroviral Proteinases*, pp. 254-&.
- Ellman, G. L. (1959). Tissue Sulfhydryl Groups. *Archives of Biochemistry and Biophysics* *82*, 70-77.
- Feng, D. F., and Doolittle, R. F. (1996). Progressive alignment of amino acid sequences and construction of phylogenetic trees from them, In *Computer Methods for Macromolecular Sequence Analysis*, pp. 368-382.
- Gerega, A., Rockel, B., Peters, J., Tamura, T., Baumeister, W., and Zwickl, P. (2005). VAT, the *Thermoplasma* homolog of mammalian p97/VCP, is an N domain-regulated protein unfoldase. *Journal of Biological Chemistry* *280*, 42856-42862.
- Gill, S. C., and Hippel, P. H. (1989). Calculation of Protein Extinction Coefficients from Amino acid Sequence Data. *Analytical Biochemistry* *182*, 319-326.
- Golbik, R., Lupas, A. N., Koretke, K. K., Baumeister, W., and Peters, J. (1999). The janus face of the archaeal Cdc48/p97 homologue VAT: Protein folding versus unfolding. *Biological Chemistry* *380*, 1049-1062.
- Grishin, N. V. (2001). Fold change in evolution of protein structures. *Journal of Structural Biology* *134*, 167-185.

- 
- Gruber, M., and Lupas, A. N. (2003). Historical review: Another 50th anniversary - new periodicities in coiled coils. *Trends in Biochemical Sciences* 28, 679-685.
- Grundy, W. N. (1998). Homology detection via family pairwise search. *Journal of Computational Biology* 5, 479-491.
- Hohenblum, H., Vorauer-Uhl, K., Katinger, H., and Mattanovich, D. (2004). Bacterial expression and refolding of human trypsinogen. *Journal of Biotechnology* 109, 3-11.
- Jin, T., and Inouye, M. (1993). Ligand-Binding to the Receptor Domain Regulates the Ratio of Kinase to Phosphatase-Activities of the Signaling Domain of the Hybrid Escherichia-Coli Transmembrane Receptor, Taz1. *Journal of Molecular Biology* 232, 484-492.
- Jin, T., and Inouye, M. (1994). Mutational Analysis of the Cytoplasmic Linker Region of Taz1-1, a Tar-Envz Chimeric Receptor in Escherichia-Coli. *Journal of Molecular Biology* 244, 477-481.
- Kageyama, T. (2002). Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. *Cellular and Molecular Life Sciences* 59, 288-306.
- Keller, M., Blochl, E., Wachtershauser, G., and Stetter, K. O. (1994). Formation of Amide Bonds without a Condensation Agent and Implications for Origin of Life. *Nature* 368, 836-838.
- Kendrew, J. C. (1950). The Crystal Structure of Horse Met-Myoglobin .1. General Features - the Arrangement of the Polypeptide Chains. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* 201, 62-&.
- Kendrew, J. C., and Parrish, R. G. (1957). The Crystal Structure of Myoglobin .3. Sperm-Whale Myoglobin. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* 238, 305-&.
- Kim, K. K., Yokota, H., and Kim, S. H. (1999). Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* 400, 787-792.
- Kim, S. H. (1994). Frozen Dynamic Dimer Model for Transmembrane Signaling in Bacterial Chemotaxis Receptors. *Protein Science* 3, 159-165.
- Koretke, K., Volker, C., Bower, M. L., and Lupas, A. (2003). Molecular evolution of histidine kinases, In *Histidine Kinases In Signal Transduction*, M. Inouye, and R. Dutta, eds. (New York: Academic Press), pp. 483-506.
- Koretke, K. K., Lupas, A. N., Warren, P. V., Rosenberg, M., and Brown, J. R. (2000). Evolution of two-component signal transduction. *Molecular Biology and Evolution* 17, 1956-1970.
- Koretke, K. K., Russell, R. B., and Lupas, A. N. (2001). Fold recognition from sequence comparisons. *Proteins-Structure Function and Genetics*, 68-75.
- Lee, D., Grant, A., Marsden, R. L., and Orengo, C. (2005). Identification and distribution of protein families in 120 completed genomes using Gene3D. *Proteins-Structure Function and Bioinformatics* 59, 603-615.

- Lin, X. L., Lin, Y. Z., and Tang, J. (1994). Relationships of Human-Immunodeficiency-Virus Protease with Eukaryotic Aspartic Proteases, In *Retroviral Proteases*, pp. 195-224.
- Lin, X. L., Loy, J. A., Sussman, F., and Tang, J. (1993). Conformational Instability of the N-Terminal and C-Terminal Lobes of Porcine Pepsin in Neutral and Alkaline-Solutions. *Protein Science* 2, 1383-1390.
- Lin, Y. Z., Fusek, M., Lin, X. L., Hartsuck, J. A., Kezdy, F. J., and Tang, J. (1992). Ph-Dependence of Kinetic-Parameters of Pepsin, Rhizopuspepsin, and Their Active-Site Hydrogen-Bond Mutants. *Journal of Biological Chemistry* 267, 18413-18418.
- Linder, J. U., Hammer, A., and Schultz, J. E. (2004). The effect of HAMP domains on class IIIb adenylyl cyclases from *Mycobacterium tuberculosis*. *European Journal of Biochemistry* 271, 2446-2451.
- Lindqvist, Y., and Schneider, G. (1997). Circular permutations of natural protein sequences: Structural evidence. *Current Opinion in Structural Biology* 7, 422-427.
- Louis, J. M., Ishima, R., Nesheiwat, I., Pannell, L. K., Lynch, S. M., Torchia, D. A., and Gronenborn, A. M. (2003). Revisiting monomeric HIV-1 protease - Characterization and redesign for improved properties. *Journal of Biological Chemistry* 278, 6085-6092.
- Lupas, A., Vandyke, M., and Stock, J. (1991). Predicting Coiled Coils from Protein Sequences. *Science* 252, 1162-1164.
- Lupas, A. N., and Gruber, M. (2005). The structure of alpha-helical coiled coils, In *Fibrous Proteins: Coiled-Coils, Collagen and Elastomers*, pp. 37-+.
- Lupas, A. N., Ponting, C. P., and Russell, R. B. (2001). On the evolution of protein folds: Are similar motifs in different protein folds the result of convergence, insertion, or relics of an ancient peptide world? *Journal of Structural Biology* 134, 191-203.
- Maxam, A. M., and Gilbert, W. (1977). New Method for Sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America* 74, 560-564.
- Moukhametzianov, R., Klare, J. P., Efremov, R., Baeken, C., Goppner, A., Labahn, J., Engelhard, M., Buldt, G., and Gordeliy, V. I. (2006). Development of the signal in sensory rhodopsin and its transfer to the cognate transducer. *Nature* 440, 115-119.
- Murzin, A. G. (1998). How far divergent evolution goes in proteins. *Current Opinion in Structural Biology* 8, 380-387.
- Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995). Scop - a Structural Classification of Proteins Database for the Investigation of Sequences and Structures. *Journal of Molecular Biology* 247, 536-540.
- Neurath, H. (1984). Evolution of Proteolytic-Enzymes. *Science* 224, 350-357.
- Nirenberg, M., and Matthaei, J. H. (1961). Dependence of Cell-Free Protein Synthesis in *E. coli* Upon Naturally Occurring or Synthetic Polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America* 47, 1588-&.

---

Orengo, C. A., Jones, D. T., and Thornton, J. M. (1994). Protein Superfamilies and Domain Superfolds. *Nature* 372, 631-634.

Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., and Thornton, J. M. (1997). CATH - a hierarchic classification of protein domain structures. *Structure* 5, 1093-1108.

Orengo, C. A., and Thornton, J. M. (2005). Protein families and their evolution - A structural perspective. *Annual Review of Biochemistry* 74, 867-900.

Park, S., Isaacson, R., Kim, H. T., Silver, P. A., and Wagner, G. (2005). Ufd1 exhibits the AAA-ATPase fold with two distinct ubiquitin interaction sites. *Structure* 13, 995-1005.

Pearl, F. M. G., Bennett, C. F., Bray, J. E., Harrison, A. P., Martin, N., Shepherd, A., Sillitoe, I., Thornton, J., and Orengo, C. A. (2003). The CATH database: an extended protein family resource for structural and functional genomics. *Nucleic Acids Research* 31, 452-455.

Privalov, P. L., Mateo, P. L., Khechinashvili, N. N., Stepanov, V. M., and Revina, L. P. (1981). Comparative Thermodynamic Study of Pepsinogen and Pepsin Structure. *Journal of Molecular Biology* 152, 445-464.

Quemeneur, E., Moutiez, M., Charbonnier, J. B., and Menez, A. (1998). Engineering cyclophilin into a proline-specific endopeptidase. *Nature* 391, 301-304.

Rao, J. K. M., Erickson, J. W., and Wlodawer, A. (1991). Structural and Evolutionary Relationships between Retroviral and Eukaryotic Aspartic Proteinases. *Biochemistry* 30, 4663-4671.

Rohl, C. A., Strauss, C. E. M., Misura, K. M. S., and Baker, D. (2004). Protein structure prediction using rosetta, In *Numerical Computer Methods, Pt D*, pp. 66-+.

Salem, G. M., Hutchinson, E. G., Orengo, C. A., and Thornton, J. M. (1999). Correlation of observed fold frequency with the occurrence of local structural motifs. *Journal of Molecular Biology* 287, 969-981.

Sali, A., Potterton, L., Yuan, F., Vanvlijmen, H., and Karplus, M. (1995). Evaluation of Comparative Protein Modeling by Modeler. *Proteins-Structure Function and Genetics* 23, 318-326.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edition edn: Cold Spring Harbor Laboratory).

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M., and Smith, M. (1977). Nucleotide-Sequence of Bacteriophage Phichi174 DNA. *Nature* 265, 687-695.

Sanger, F., and Tuppy, H. (1951). The Amino-Acid Sequence in the Phenylalanyl Chain of Insulin .1. the Identification of Lower Peptides from Partial Hydrolysates. *Biochemical Journal* 49, 463-481.

- Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., Levchenko, I., *et al.* (2004). Sculpting the proteome with AAA+ proteases and disassembly machines. *Cell* *119*, 9-18.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America* *95*, 5857-5864.
- Scott, W. G., and Stoddard, B. L. (1994). Transmembrane Signaling and the Aspartate Receptor. *Structure* *2*, 877-887.
- Singh, M., Berger, B., Kim, P. S., Berger, J. M., and Cochran, A. G. (1998). Computational learning reveals coiled coil-like motifs in histidine kinase linker domains. *Proceedings of the National Academy of Sciences of the United States of America* *95*, 2738-2743.
- Soding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* *33*, W244-W248.
- Soding, J., and Lupas, A. N. (2003). More than the sum of their parts: on the evolution of proteins from peptides. *Bioessays* *25*, 837-846.
- Tanaka, T., and Yada, R. Y. (1996). Expression of soluble cloned porcine pepsinogen a in *Escherichia coli*. *Biochemical Journal* *315*, 443-446.
- Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A., and Blundell, T. L. (1978). Structural Evidence for Gene Duplication in Evolution of Acid Proteases. *Nature* *271*, 618-621.
- Tisa, L. S., and Adler, J. (1993). Calcium-Ions Are Involved in *Escherichia-Coli* Chemotaxis (Vol 89, Pg 11804, 1992). *Proceedings of the National Academy of Sciences of the United States of America* *90*, 2097-2097.
- Tisa, L. S., and Adler, J. (1995a). Chemotactic Properties of *Escherichia-Coli* Mutants Having Abnormal Ca<sup>2+</sup> Content. *Journal of Bacteriology* *177*, 7112-7118.
- Tisa, L. S., and Adler, J. (1995b). Cytoplasmic Free-Ca<sup>2+</sup> Level Rises with Repellents and Falls with Attractants in *Escherichia-Coli* Chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* *92*, 10777-10781.
- Todd, A. E., Orengo, C. A., and Thornton, J. M. (2001). Evolution of function in protein superfamilies, from a structural perspective. *Journal of Molecular Biology* *307*, 1113-1143.
- Todd, A. E., Orengo, C. A., and Thornton, J. M. (2002). Plasticity of enzyme active sites. *Trends in Biochemical Sciences* *27*, 419-426.
- Tokishita, S., Kojima, A., and Mizuno, T. (1992). Transmembrane Signal Transduction and Osmoregulation in *Escherichia-Coli* - Functional Importance of the Transmembrane Regions of Membrane-Located Protein-Kinase, EnvZ. *Journal of Biochemistry* *111*, 707-713.
- Tomomori, C., Tanaka, T., Dutta, R., Park, H. Y., Saha, S. K., Zhu, Y., Ishima, R., Liu, D. J., Tong, K. I., Kurokawa, H., *et al.* (1999). Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. *Nature Structural Biology* *6*, 729-734.

---

Wadhams, G. H., and Armitage, J. P. (2004). Making sense of it all: Bacterial chemotaxis. *Nature Reviews Molecular Cell Biology* 5, 1024-1037.

Watson, J. D., and Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* 171, 737-738.

Wegener, A. A., Klare, J. P., Engelhard, M., and Steinhoff, H. J. (2001). Structural insights into the early steps of receptor-transducer signal transfer in archaeal phototaxis. *Embo Journal* 20, 5312-5319.

Williams, S. B., and Stewart, V. (1999). Functional similarities among two-component sensors and methyl-accepting chemotaxis proteins suggest a role for linker region amphipathic helices in transmembrane signal transduction. *Molecular Microbiology* 33, 1093-1102.

Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989). Conserved Folding in Retroviral Proteases - Crystal-Structure of a Synthetic Hiv-1 Protease. *Science* 245, 616-621.

Yakhnin, A. V., Vinokurov, L. M., Surin, A. K., and Alakhov, Y. B. (1998). Green fluorescent protein purification by organic extraction. *Protein Expression and Purification* 14, 382-386.

Yeh, J. I., Biemann, H. P., Pandit, J., Koshland, D. E., and Kim, S. H. (1993). The 3-Dimensional Structure of the Ligand-Binding Domain of a Wild-Type Bacterial Chemotaxis Receptor - Structural Comparison to the Cross-Linked Mutant Forms and Conformational-Changes Upon Ligand-Binding. *Journal of Biological Chemistry* 268, 9787-9792.

## Akademische Lehrer

Prof. Hans Bisswanger

Prof. Peter Bohley

PD Roland Brock

Prof. Gabriele Dodt

Prof. Ralf Dringen

Prof. Michael Duszenko

Prof. Karl Eisele

Prof. Kai-Uwe Fröhlich

Prof. Bernd Hamprecht

Prof. Afroditi Kapurniotu

Prof. Frank Madeo

Prof. Dieter Mecke

Prof. Helga Ninnemann

Prof. Hans-Georg Probst

PD Stanka Stoeva

Prof. Wolfgang Voelter

Prof. Ulrich Weber

Prof. Ulrich Weser

Prof. Heiner Wiesinger

Prof. Karl Walter Bock

Prof. Martin Brendle

Prof. Muhammad Iqbal Choudhary

Prof. Friedrich Gönnerwein

Prof. Wolfgang Pfeiffer

Prof. Helmut Pommer

Prof. Karl Poralla

Prof. Klaus Reutter

Prof. Michael Schwarz

Prof. Wolfgang Wohlleben

Prof. Klaus Albert

Prof. Günter Gauglitz

Prof. Günter Häfelinger

Prof. Hanspaul Hagenmaier

Prof. Volker Hoffmann

Prof. Günther Jung

Prof. Ekkehard Lindner

Prof. Martin E. Mayer

Prof. Heinz Oberhammer

Prof. Volker Schurig

Prof. Joachim Strähle

PD Karl-Heinz Wiesmüller

Prof. Klaus-Peter Zeller



---

## Lebenslauf

### Persönliche Daten

Michael Hulko

Geboren am 04.02.1977 in Loveland, USA

### Ausbildung

**1/2003 – 11/2006**

Promotion

„Evolution von Double psi beta-barrel Proteinen und Aspartatproteasen“

Max-Planck-Institut für Entwicklungsbiologie, Tübingen  
Abteilung Proteinevolution

Betreuer:

Prof. Andrei Lupas (Max-Planck-Institut, Tübingen)

Prof. Gabriele Dodt (Universität Tübingen)

**5/2002 – 12/2002**

Diplomarbeit

„Entwicklung einer temporären Immobilisierung von Peptiden auf Glasoberflächen zur langsamen Peptidfreisetzung in Zell-Chip-Anwendungen“

Institut für organische Chemie, Universität Tübingen

Betreuer:

Prof. Günther Jung

**10/1997 – 12/2002**

Studium der Biochemie: Eberhard-Karls-Universität, Tübingen

**10/1996 – 10/1997**

Zivildienst: Rotes Kreuz, Sindelfingen

**07/1987 – 07/1996**

Andreae-Gymnasium, Herrenberg

**07/1983 – 07/1987**

Peter-Rosegger Grundschule, Gärtringen