

Kapitel 1

Einleitung

Spezifische und fein aufeinander abgestimmte Protein-Protein-Wechselwirkungen spielen in beinahe allen zellulären Prozessen eine fundamentale Rolle.¹⁹⁰ Sie sind die Grundlage für die Komplexität lebender Zellen und Organismen. Ein nicht unwesentlicher Anteil dieser Protein-Protein-Wechselwirkungen wird durch spezialisierte Proteinuntereinheiten, den so genannten Protein-Interaktionsdomänen, vermittelt.²¹⁰ In dieser Arbeit werden auf der Basis experimenteller Daten theoretische Modelle der Wechselwirkung zwischen Protein-Interaktionsdomänen und ihren peptidischen Liganden erstellt. Diese Modelle geben nicht nur einen Einblick in den Wechselwirkungsmechanismus der untersuchten Domänen, es lassen sich aus ihnen auch generelle Erkenntnisse über die Mechanismen der Spezifität dieser Wechselwirkungen ableiten. Mit Hilfe dieser Erkenntnisse können neue wechselwirkende Peptide und Proteine ohne experimentelle Analyse vorhergesagt werden.

1.1 Protein-Protein-Wechselwirkungsnetzwerke und Protein-Interaktionsdomänen

Die Funktion eines biologischen Prozesses, einer Zelle bzw. eines Organismus wird maßgeblich durch ein hoch komplexes Protein-Protein-Wechselwirkungsnetzwerk bestimmt.¹⁹⁰ Diesem Netzwerk liegen eine Vielzahl spezifischer Protein-Protein-Wechselwirkungen zugrunde, die dabei ganz unterschiedliche Funktionen erfüllen. Dazu gehören strukturelle Aufgaben durch den Aufbau hochmolekularer Polymere,⁶⁸ die Bildung von Multiproteinkomplexen mit enzymatischer Funktion,^{41,211} sowie die Signalverarbeitung durch dynamische Protein-Protein-Wechselwirkungen.¹³⁴ Insbesondere die spezifischen und reversiblen Wechselwirkungen in Signaltransduktionswegen¹³⁴ werden zum Großteil durch so genannte Protein-Interaktionsdomänen vermittelt.

Protein-Protein-
Wechselwirkungen
und ihre Funktion

Modularer
Proteinaufbau und
Proteindomänen

Generell sind Proteindomänen genetisch mobile, strukturell und funktionell unabhängige Einheiten.¹⁴² Vor allem in höheren Organismen wird die Funktion eines Proteins wesentlich durch eine charakteristische Kombination verschiedener Domänen, der sogenannten Domänenarchitektur, bestimmt (Abbildung 1-3 und Abbildung 1-5).^{103,192} Dies ist vermutlich darauf zurückzuführen, dass die Evolution neuer Proteine mit komplexer Funktion durch die Kombination einmal erfolgreich entwickelter Domänen stark beschleunigt wird.¹⁴²

Sequenzbasierte
Identifizierung von
Domänenfamilien

Auf Grund ihrer entwicklungsgeschichtlichen Verwandtschaft können Familien von Proteindomänen an Hand der Homologie ihrer Aminosäuresequenzen identifiziert werden. In Datenbanken, wie PFAM (<http://www.sanger.ac.uk/Pfam>)¹⁶ oder SMART (<http://smart.embl-heidelberg.de>),¹⁶¹ werden dazu repräsentative Sequenzalignments für die einzelnen Familien gesammelt und ihre charakteristischen Sequenzmerkmale in Form von Sequenzprofilen und *Profile-Hidden-Markov-Modellen* (pHMM)⁴⁷ zusammengefasst. Mit Hilfe dieser Sequenzmodelle können dann neue Mitglieder einer Domänenfamilie in der vor allem durch Genomprojekte stetig wachsenden Zahl an Proteinsequenzen identifiziert werden.

Konservierte Struktur
der Domänen

Auf Grund des Sequenz-definiert-Struktur-Paradigmas⁹ besitzen alle Mitglieder einer Domänenfamilie dieselbe 3D-Raumstruktur.^{67,120} Dabei wird in dieser Arbeit die Tatsache ausgenutzt, dass Proteindomänen gewöhnlich unabhängig von ihrer Umgebung in den verschiedensten Proteinen und daher auch in isolierter Form die ihnen eigene Faltung annehmen. Meist besitzen sie eine kompakte und globuläre Struktur, in der N- und C-Termini räumlich nah zu einander positioniert sind.¹³⁴

Sequenz-Struktur-
Paradigma

Protein-Interaktions-
domänen

Protein-Interaktionsdomänen erkennen meist lineare und exponierte Sequenzmotive in ihren Zielproteinen über einen innerhalb einer Domänenfamilie konservierten Wechselwirkungsmechanismus.^{128,213} Auf Grund des konservierten Wechselwirkungsmechanismus sind die Liganden einer Domänenfamilie meist durch ein gemeinsames Kernmotiv gekennzeichnet. Verschiedene Spezifitäten innerhalb der Familie werden durch unterschiedliche Präferenzen für Reste, die das Kernmotiv flankieren, erreicht.^{1,13,176} An Hand der charakteristischen Konsensusmotive der spezifisch erkannten Liganden können die Vertreter einer Familie in Subfamilien ähnlicher Spezifität, die sogenannten Spezifitätsgruppen, eingeteilt werden (siehe Tabelle 1-3, Seite 11, und Tabelle 1-5, Seite 16).^{170,177} Trotz der vergleichsweise geringen Affinität im niedrigen mikromolaren bis hohen nanomolaren Bereich,¹³⁴ welche für die Reversibilität der Wechselwirkungen verantwortlich ist, zeichnen sich die meisten Protein-Interaktionsdomänen durch eine relativ hohe Spezifität für ihre Zielproteine aus.

Klassifizierung in
Spezifitätsgruppen

Reversibilität auf
Grund niedriger
Affinität

Einige Protein-Interaktionsdomänen, wie beispielsweise PDZ-Domänen,¹⁴¹ kommen auch schon in Prokaryoten vor, der Großteil ist aber parallel zur Entwicklung der komplexeren eukaryotischen Zellen, vor allem infolge des gesteigerten Bedarfs an Signalverarbeitung, entstanden (Tabelle 1-1).^{134,142} Für größere Domänenfamilien sind teilweise mehrere hundert paraloge Vertreter innerhalb eines Genoms zu finden. Die Domänen einiger Familien können zum Teil auch in mehreren Kopien innerhalb eines einzigen Proteins vorkommen (z.B. 13 PDZ-Domänen im *Multi-PDZ-domain-protein mMUPP1*).¹⁸⁸

Verbreitung der Protein-Interaktionsdomänen

Dies bedeutet, dass durch eine geringe Anzahl von Protein-Interaktionsdomänen-Familien eine große Anzahl an Wechselwirkungen vermittelt wird. Auf Grund der Konservierung der Sequenz, der Struktur und prinzipiell auch der Funktion in Form des Wechselwirkungsmechanismus sollten die Erkenntnisse, die für einige wenige Vertreter einer Familie gewonnen wurden, auf die übrigen, experimentell nicht untersuchten Vertreter übertragbar sein. Auf der Basis des Sequenz-bestimmt-Struktur⁹ und Struktur-bestimmt-Funktion-Paradigmas müssten damit für eine große Anzahl homologer Domänen, die nur auf Grund ihrer Sequenzhomologie identifiziert wurden, Wechselwirkungspartner vorhergesagt werden können.²⁵

Ziel der Analyse von Protein-Interaktionsdomänen

Sequenz-Struktur-Funktion-Paradigma

Tabelle 1-1 Anzahl paraloger Protein-Interaktionsdomänen innerhalb repräsentativer Spezies

Spezies	WW-Domänen		PDZ-Domänen		EVH1-Domänen		
	Domänen	Proteine	Domänen	Proteine	Domänen	Proteine	
Gesamt	761	417	3212	2022	71	70	
Eubacteria	0	0	750	607	0	0	
Archaea	0	0	23	21	0	0	
Eukaryota	761	417	2439	1394	71	70	
Pilze	<i>S. cerevisiae</i>	9	6	3	2	1	1
Pflanzen	<i>A. thaliana</i>	16	11	15	14	0	0
Würmer	<i>C. elegans</i>	45	27	131	94	4	3
Insekten	<i>D. melanogaster</i>	47	27	145	93	5	5
Säugetier	<i>M. musculus</i>	94	51	335	185	15	15
Mensch	<i>H. sapiens</i>	142	77	449	250	9	9

Die angegebenen Zahlen bezeichnen die Gesamtzahl der Domänen innerhalb eines Organismus oder Reichs bzw. die Anzahl der Proteine, welche Domänen der jeweiligen Familie enthalten. Die Zahlen sind der SMART-Datenbank Version 4.1 im *Genomic-Mode* (23.3.2005) entnommen.^{108,161} Diese Statistik wurde damit ausschließlich unter Verwendung komplett sequenzierter Genome erstellt.

1.2 Prinzipien der Protein/Liganden-Wechselwirkung

Spezifische Wechselwirkung infolge der Komplementarität

Die spezifische Wechselwirkung zwischen einer Protein-Interaktionsdomäne (allgemein einem Protein) und seinem Liganden (Protein oder Peptid) erfolgt über die Ausbildung einer Vielzahl nicht-kovalenter Interaktionen. Die Stärke der einzelnen Interaktionen hängt dabei sowohl vom Abstand, als auch von den Eigenschaften der wechselwirkenden Atomen ab. Daher müssen die Kontaktflächen (Epitope) der beiden Wechselwirkungspartner sowohl bezüglich der Passform, als auch der Eigenschaften (z.B. negative und positive Ladungen) zueinander komplementär sein.

Schlüssel-Schloss-Prinzip

Wechselwirkungs-Gleichgewicht

Nicht-kovalenten Interaktionen sind prinzipiell reversibel. Es besteht somit ein Gleichgewicht zwischen dem gebildeten Komplex (DL) und den beiden freien Wechselwirkungspartnern (Domäne D bzw. Ligand L , Gleichung 1:1).



Die Stärke der Wechselwirkung (Bindungsaffinität) wird auf der Grundlage des Massenwirkungsgesetz durch die Dissoziations- (K_d) bzw. Assoziationskonstante (K_a) beschrieben (Gleichung 1:2).

$$K_d = [D][L]/[DL] = K_a^{-1} \quad (1:2)$$

Über Gleichung 1:3 steht die Dissoziationskonstante gleichzeitig im direkten Zusammenhang zur freien Gibbs Energie der Bindung (ΔG , R allgemeine Gaskonstante, T Temperatur).

$$\Delta G = -RT \ln(K_a) = +RT \ln(K_d) \quad (1:3)$$

Eine starke bzw. affine Wechselwirkung wird damit durch eine niedrige K_d bzw. durch ein stark negatives ΔG angezeigt. Auf der Basis dieser biophysikalischen Grundlagen lassen sich die Wechselwirkungseigenschaften (Spezifität und Selektivität) einer Protein-Interaktionsdomäne quantitativ beschreiben.

Spezifität

Wie einführend erläutert wurde, hängt die selektive Bindung eines Liganden von der Form und den Eigenschaften der Bindungsstelle des Proteins ab. Auf Grund der Forderung der Komplementarität ergeben sich damit aus der Sicht des Proteins eine Reihe von Kriterien, welche ein potentieller Ligand erfüllen muss, um an das Protein zu binden. Die Gesamtheit dieser Kriterien definiert damit die Spezifität des Proteins. Gleichzeitig kann die Spezifität

Qualitative Definition der Spezifität

Quantitative Definition der Spezifität

auch quantitativ auf der Basis der Bindungsenergie/Bindungsaffinität definiert werden. Dabei wird davon ausgegangen, dass jedes der geforderten Kriterien einen definierten Beitrag zur gesamten Bindungsaffinität liefert. In Abhängigkeit von der Erfüllung der einzelnen ge-

forderten Kriterien ändern sich die Affinitätsbeiträge und damit die gesamte Bindungsaffinität. Die quantitative Spezifität beschreibt also den Zusammenhang zwischen der Ausprägung der einzelnen Kriterien und ihrem Affinitätsbeitrag.

Demzufolge unterscheiden sich „bindende“ von „nicht bindenden“ Liganden dadurch, wie diese die Kriterien erfüllen. Ein „nicht bindender Ligand“ bildet einerseits zu wenig vorteilhafte Kontakte aus, welche die Bindungsaffinität erhöhen, andererseits kommt es auf Grund der schlechten „Passform“ vermehrt zu unvorteilhaften Kontakten, welche die Bindungsaffinität weiter verringern.

„Nicht bindende Liganden“

Für die Wechselwirkung zwischen Protein-Interaktionsdomänen und ihren verschiedenen peptidischen Liganden hängen die Unterschiede in der Bindungsaffinität infolge des konservierten Bindungsmechanismus (siehe 1.3.2 und 1.4.1) fast ausschließlich von den unterschiedlichen Affinitätsbeiträgen der einzelnen Aminosäureseitenketten in den Liganden ab. Dies gilt natürlich nur unter der Voraussetzung, dass sich die Struktur der Domäne bei Bindung des Liganden (so gut wie) nicht ändert. Folglich kann die Spezifität ganz allgemein als von der Ligandensequenz abhängige Affinitätsfunktion definiert werden. Im Idealfall lässt sich daher die Spezifität einer Protein-Interaktionsdomäne durch die additiven Affinitätsbeiträge der Aminosäureseitenketten des Liganden, sowie deren Kooperativität beschreiben (siehe 3.5.2).

Quantitative Definition der Spezifität für Protein-Interaktionsdomänen

Spezifität als Funktion der Affinitätsbeiträge der einzelnen Ligandenreste

Mit Hilfe dieser Beschreibung der Spezifität lässt sich die Selektivität ebenfalls quantitativ definieren. Prinzipiell beschreibt die Selektivität eines Proteins die Trennschärfe zwischen „gewünschten“ und „ungewünschten“ Liganden aus der Menge aller möglichen Liganden. Für Protein-Interaktionsdomänen entspricht die Menge aller möglichen Liganden dem Sequenzraum, welcher durch die Anzahl der mit der Domäne in Kontakt stehenden Ligandenreste definiert wird. In Bezug auf diese Menge entspricht die Selektivität der Steilheit der Affinitätsfunktion, welche wiederum durch die Spezifität bestimmt wird. Dagegen bezeichnet die Selektivität bezogen auf konkurrierende Domänen die Größe der Überlappung der Spezifitäten, d.h. der Menge der Liganden, welche gleichzeitig von den konkurrierenden Domänen mit einer gegebenen Affinität gebunden werden können. Schließlich lässt sich die Selektivität einer Wechselwirkung auch auf funktionaler Ebene definieren. Selektivität bezeichnet in diesem Fall die Trennschärfe, mit der Zielproteine unterschiedlicher Funktion gebunden werden. Hierfür ist also nicht nur die Spezifität der Domäne, sondern auch die Verteilung der spezifisch gebundenen Erkennungsmotive in den Zielproteinen von entscheidender Bedeutung (siehe 4.5). Für die Definition der Selektivität ist also immer wichtig, auf

Selektivität

Qualitative Definition der Selektivität

Konzept des Ligandensequenzraums

Quantitative Definition der Selektivität für einzelne Domäne

Quantitative Definition der Selektivität für konkurrierende Domänen

Quantitative Definition der Selektivität auf funktionaler Ebene

was sie bezogen wird. Im Vergleich zum traditionellen Verständnis von Spezifität und Selektivität wird allerdings keine artifizielle Unterscheidung zwischen „Liganden“ und „Nicht-Liganden“ getroffen. Vielmehr wird das Unterscheidungskriterium in Form der Bindungsaffinität quantifiziert.

Mikroskopische und makroskopische Affinität

Tandemwiederholte Domänen und Bindungsmotive

Die vorangegangenen Gleichungen und Erläuterungen beschreiben den Spezialfall einer Wechselwirkung zwischen einer Domäne und einem Liganden mit je nur einer einzelnen Bindungsstelle (1:1 Komplexmodell). Insbesondere im Fall der hier untersuchten WW-Domänen kommen jedoch sowohl die Domänen als auch die prolinreichen Bindungsmotive häufig tandemwiederholt vor. Infolge der größeren Anzahl an Möglichkeiten der Komplexbildung ist damit die makroskopische Affinität größer als die intrinsische mikroskopische Affinität für die Wechselwirkung zwischen einer einzelnen Bindungsstelle und einer einzelnen Domäne. Hinzu kommen Kooperativitätseffekte zwischen benachbart gebundenen Domänen und Aviditätseffekte für die Fälle, in denen sowohl Domänen, als auch Liganden tandemwiederholt vorliegen. Durch diese übergeordneten Mechanismen wird die Spezifität und Selektivität der durch Protein-Interaktionsdomänen vermittelten Wechselwirkungen zusätzlich beeinflusst.

1.3 WW-Domänen

WW-Domänen sind kleine kompakte Protein-Interaktionsdomänen, die aus 35–40 Aminosäuren bestehen und als dreisträngiges antiparalleles β -Faltblatt vorliegen (Abbildung 1-2). Zum ersten Mal wurden WW-Domänen in *mYAP65* (65kD *Yes-kinase associated protein* aus der Maus)³⁰ als Protein-Interaktionsdomänen identifiziert, die prolinreiche Sequenzmotive in ihren Zielproteinen erkennen (siehe 1.3.2).³⁷ Die durch sie vermittelten Wechselwirkungen, mit Affinitäten im mikromolaren Bereich,¹¹³ sind für eine ganze Reihe unterschiedlicher zellulärer Prozesse von Bedeutung (siehe 1.3.3).¹⁷⁵

Domänen werden durch den Proteinnamen und einem Domänensuffix bezeichnet (siehe 2.1.1).

Tabelle 1-2 Sequenzalignment repräsentativer WW-Domänen

Domäne ^{a)}	UP-ID ^{b)}	Aminosäure-Sequenz ^{d)}			Domänengrenzen ^{e)}					
		β A	β B	β C	Beginn	Ende				
	Domänenposition ^{c)} ω	1	7	20	24	b	30	39	42	
<i>h</i> YAP65-WW	P46937	PLPAGWEMAKTSS	GORYFLNHIDQ	TTTWQDPRK	172	204				
<i>h</i> DYSTRO-WW	P11532	SVQGPWERAISP	NKVPYYINHET	QTTTCWDHPKM	3056	3088				
<i>y</i> RSP5-WW2	P39940	ELPSGWEQRFT	PEGRAYFVDHN	TRTTTWVDP	331	364				
<i>h</i> NEDD4-WW3	P46934	FLPKGWEVRHAP	NGRPFIDHNTK	TTTWEDPRL	449	481				
<i>h</i> PIN1-WW	Q13526	KLPPGWWEKMR	SRSSGRVYYFN	HITNASQWER	6	39				
<i>y</i> ESS1-WW	P22696	GLPTPWTVRYS	SKSKKREYFFN	PETKHSQWEE	30	63				
<i>h</i> FBP11-WW1	O75400	GAKSMWTEHK	SPDGRTYYYN	TETKQSTWEK	116	148				
<i>h</i> FBP11-WW2	O75400	LKSPWKEYKSD	SGKPYYYNSQ	TKESRWAKPKE	157	189				
<i>h</i> CA150-WW2	O14776	TAVSEWTEYKT	ADGKTYYYN	NRTLESTWEK	430	462				
<i>h</i> CA150-WW1	O14776	PTEEIWVENKTP	DGKVVYYN	ARTRESAWTK	132	164				
<i>h</i> GAS7-WW	O60861	ILPPGWQSYLS	PQGRRRYYV	NTTTNETTWER	14	46				
<i>y</i> PRP40-WW1	P33203	--MSIWKEAKD	ASGRIIYYN	TLTKKSTWEK	1	31				
<i>y</i> PRP40-WW2	P33203	LRENGWKAAKT	ADGKVVYYN	PPTRETSTW	40	72				
<i>h</i> FE65-WW	O00213	DLPAGWMRVQD	TSGTTYWHI	PTGTTQWEP	254	285				
<i>h</i> PQBP1-WW	O60828	GLPPSWYKVF	DPSSGLPYW	NADTDLVS	47	80				
<i>h</i> HYP109-WW1	Q9NT81	IEMGDWQEV	WDENTGSYY	YWNQTNEV	211	244				
<i>h</i> HYP109-WW2	Q9NT81	ATPKGWSSHW	DRDHRRYF	YVNEQSGE	592	625				
<i>y</i> YFB0-WW	P43582	QVPSGWKAV	FDDEYQTW	YYVDLSTN	10	43				
<i>h</i> FBP21-WW1	O75554	PSKGRWVEG	ITSEGYHYY	YDLISGAS	123	155				
<i>h</i> FBP21-WW2	O75554	AVKTVWVEG	LSGFTYY	YNTETGES	164	196				
<i>y</i> YPR152C-WW	Q06525	--MRGEWQEF	KTPAGKKYY	NKNTKQSR	1	32				
<i>y</i> YJQ8-WW	P46995	RLPPGWEI	IHENGRPL	YYNAEQK	476	507				
Prototype	PDB: 1E0M	GLPPGWDEY	KTHNGKTY	YYNHNTK	1	33				

Seitenkettenrichtung^{f)}

↑ dem Liganden zugewandt

↓ dem Liganden abgewandt

Strukturelle Reste

Funktionale Reste

^{a)} Domänennamen gebildet aus Organismus-Präfix, Proteinnamen und Domänen-Suffix (siehe 2.1.1). Die *Prototype* Domäne [PDB: 1E0M]¹¹² ist eine künstlich designte WW-Domäne.

^{b)} Primärer UniProt¹⁰ Zugriffscod des die WW-Domäne enthaltenden Proteins.

^{c)} Die konservierten strukturellen Domänenpositionen (ω) sind analog zu *hYAP65-WW* durchnummeriert (siehe 2.1.2.1).

^{d)} Sequenzalignment der Sequenzen der WW-Domänen. Die Einfärbung konservierter Aminosäureeigenschaften basiert auf der Konsensussequenz-Färbemethode von CLUSTAL-X.¹⁸¹ Oberhalb des Sequenzalignments ist der Sekundärstrukturkonsensus schematisch dargestellt.

^{e)} Position der ersten und letzten Aminosäure der Domänensequenz innerhalb des Gesamtproteins.

^{f)} Die für die Faltung der Domäne verantwortlichen strukturellen Reste sind orange unterlegt, die konservierten funktionalen Reste für die Ligandenbindung sind violett unterlegt. Die Pfeile zeigen die Richtung der Seitenkette an: (↑) in Richtung des Liganden oder (↓) nicht in Richtung des Liganden.

Die Domänenreste ω sind analog zu hYAP65-WW nummeriert (siehe 2.1.2.1).

Der Name WW-Domäne ist dabei auf zwei konservierte Tryptophanreste zurückzuführen ($W\omega 17$ und $W\omega 39$), die in einem Abstand von 20–22 Resten als charakteristisches Sequenzmerkmal auftreten (Tabelle 1-2).³⁰ Zwischen diesen beiden konservierten Tryptophanen besitzen WW-Domänen einen Cluster aus 2–4 aromatischen Aminosäuren (um $\omega 29$) sowie meist ein Glyzin an Position $\omega 25$. Daneben sind in den meisten WW-Domänen Proline sowohl am N-Terminus als auch in $\omega 42$ zu finden.

Das β -Faltblatt erlangt seine Stabilität durch einen Cluster konservierter hydrophober Aminosäuren ($W\omega 17$, $Y/F\omega 29$, $P\omega 42$, häufig auch $L/V\omega 13$ und $P\omega 14$),¹¹³ die auf der einen Seite des β -Faltblatts die N- und C-Termini fixieren (Tabelle 1-2 und Abbildung 1-2, orange Reste). Dadurch bleibt auf der anderen Seite des β -Faltblatts eine sattelförmige Oberfläche zur Bindung der prolinreichen Liganden frei, an der unter anderem die hoch konservierten Reste $Y/F\omega 28$, $T/S\omega 37$ und $W\omega 39$ beteiligt sind (Tabelle 1-2 und Abbildung 1-2, violette Reste).

1.3.1 Struktur und Funktion prolinreicher Sequenzmotive

Neben den WW-Domänen interagieren auch SH3-Domänen,²¹² EVH1-Domänen,¹³ UEV-Domänen¹⁴³ sowie GYF-Domänen¹²³ mit prolinreichen Sequenzmotiven in ihren Zielproteinen. Dabei gehören die prolinreichen Sequenzabschnitte auch zu den häufigsten Motiven in eukaryotischen Genomen.¹⁵³

Prolin ist eine ungewöhnliche Aminosäure, da die Seitenkette mit dem Stickstoffatom der eigenen Hauptkette zyklisiert ist. Dadurch wird die Flexibilität der Peptidhauptkette erheblich eingeschränkt. Folglich liegen prolinreiche Peptide in wässriger Lösung bevorzugt in der sogenannten Polyprolin-Typ-II(PPII)-Helix-Konformation vor (Abbildung 1-1).⁴⁰ Diese Konformation ist durch die Torsionswinkel des Peptidrückgrats von $\Phi \approx -78^\circ$ und $\Psi \approx 146^\circ$ charakterisiert.¹¹⁹ Dabei müssen aber nicht zwangsläufig alle Reste Prolin sein.^{3,166}

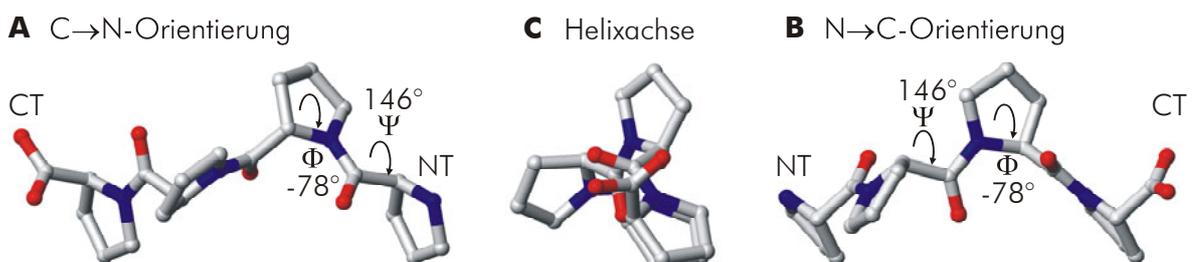


Abbildung 1-1 Die Polyprolin-Typ-II-Helix (PPII)

Polyprolin-Typ-II-Helix in (A) C→N-Orientierung, (B) mit Blick in Richtung der Helixachse und (C) in N→C-Orientierung. Exemplarisch ist je ein Φ - und Ψ -Hauptketten-Torsionswinkel für eine ideale PPII-Helix-Konformation angegeben. Die N- und C-Termini sind durch NT und CT bezeichnet.

Die PPII-Helix-Konformation ist eine relativ steife und langgestreckte linksgängige Helix mit jeweils 3 Resten pro Windung. Sie zeigt daher eine dreifache Rotationssymmetrie um die Helixachse³ und eine Pseudosymmetrie bezüglich ihrer N→C- bzw. C→N-terminalen Orientierung (Abbildung 1-1).²¹² Die Prolinringe interagieren als gut zugängliche hydrophobe Oberflächen bevorzugt mit aromatischen Ringen.²⁰³ Infolge der ebenfalls exponierten Carbonylgruppen der Hauptkette sind prolinreiche Peptidabschnitte trotz dieser exponierten hydrophoben Oberfläche gut wasserlöslich. Die Carbonylgruppen sind gleichzeitig hervorragende Wasserstoffbrücken-Akzeptoren, da sie in Folge der Seitenkettenzyklisierung relativ unflexibel und stark elektronegativ sind. Zusammen tragen diese Eigenschaften zur weiten Verbreitung der Wechselwirkung mit prolinreichen Sequenzmotiven bei.^{14,93,212}

1.3.2 WW-Domänen/Liganden-Bindungsmechanismus

Auf der Oberfläche der WW-Domänen bilden die beiden hochkonservierten aromatischen Aminosäuren Y/F ω 28 und W ω 39 eine hydrophobe Oberfläche (konservierte Prolin-Bindungsstelle, Abbildung 1-2), die komplementär zur Oberfläche der prolinreichen PPII-Helix ist.^{14,113,156,212} Die beiden aromatischen Ringe sind dabei so zueinander positioniert, dass zwei konsekutive Proline einer PPII-Helix koplanar zu den beiden aromatischen Ringen zu liegen kommen.^{14,76,87} Die konservierten virtuellen Ligandenpositionen (siehe 2.1.2.1), welche diese beiden Proline im Komplex mit einer WW-Domäne belegen, werden als $\lambda 0$ (koplanar zu W ω 39) und $\lambda 1$ (koplanar zu Y/F ω 28) bezeichnet. Zusätzlich werden in WW-Domänen/Liganden-Komplexen zwei charakteristische Wasserstoffbrücken zwischen den Hauptketten-Carbonylgruppen der Ligandenreste $\lambda - 2$ und $\lambda 1$ und der Indol-NH-Gruppe von W ω 39 bzw. der Hydroxylgruppe von S/T ω 37 ausgebildet (Abbildung 1-2).^{14,113,193}

Auf Grund der Konservierung des Wechselwirkungsmechanismus lassen sich virtuelle Ligandenpositionen λ in WW-Domänen/Liganden-Komplexen definieren. $\lambda 0$ und $\lambda 1$ kennzeichnen dabei die Reste, deren Seitenketten sich innerhalb der konservierten Prolin-Bindungsstelle an W ω 39 bzw. Y/F ω 28 anlagern (siehe 2.1.2.1).

Zusätzlich zu diesem zentralen Wechselwirkungsmechanismus wird die Spezifität einer WW-Domäne durch weitere Interaktionen bestimmt, die aber auf Untergruppen der WW-Domänenfamilie beschränkt sind (siehe Tabelle 1-3). So bindet beispielsweise eine Gruppe von WW-Domänen tyrosinhaltige Liganden (z.B. hDYSTRO-WW, Abbildung 1-2). Das Tyrosin in Ligandenposition $\lambda 3$ wird dabei von einer Tasche aufgenommen, die aus einem aliphatischen Rest (V/I/L) in Position $\omega 30$, einem Histidin in $\omega 32$ und einem basischen Rest (R/K/Q) in $\omega 35$ gebildet wird.^{76,113} Durch diese zusätzlichen Interaktionen mit flankierenden Ligandenresten entstehen sogenannte Spezifitätsgruppen. Dementsprechend war die Familie der WW-Domänen bislang in vier bis fünf mehr oder weniger gut definierte Gruppen eingeteilt worden (Tabelle 1-3, Seite 11).^{177,178}

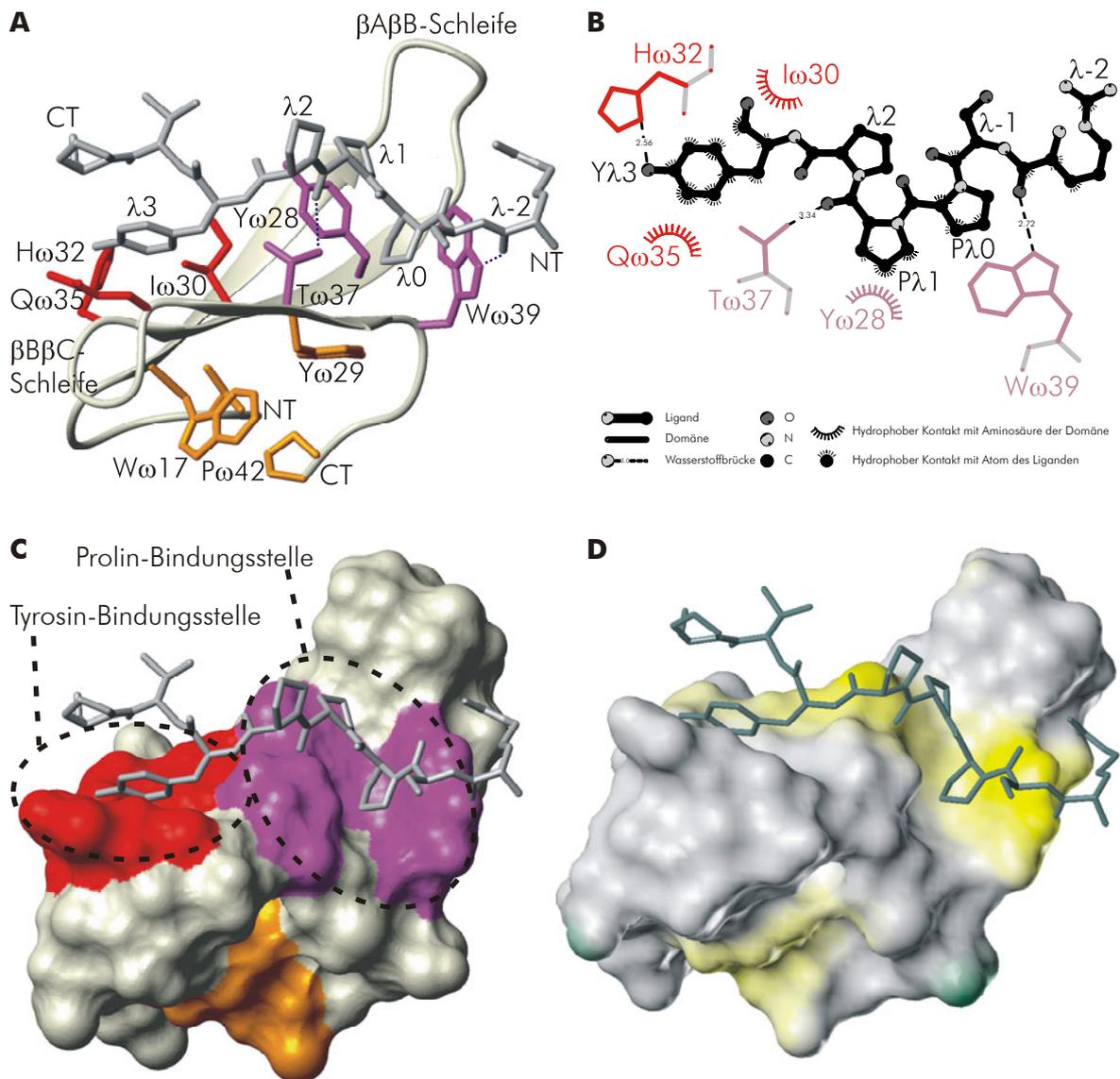


Abbildung 1-2 Struktur der hDYSTRO-WW Domäne aus Dystrophin im Komplex mit einem Peptid aus β -Dystroglycan

(**A**) Bändermodell Darstellung¹⁵¹ der hDYSTRO-WW Domäne [PDB: 1EG4].⁷⁶ Die Reste der konservierten Prolin-Bindungsstelle sind violett eingefärbt, die spezifitätsbestimmende Reste der Tyrosin-Bindungsstelle sind rot eingefärbt und die strukturellen Reste orange. Die Domänenreste ω sind analog zu hYAP65-WW durchnummeriert. Der Ligand (**RSPPPYVF**) aus β -Dystroglycan ist grau dargestellt. Er ist in C \rightarrow N-Orientierung gebunden. Die konservierten Ligandenpositionen sind durch $\lambda-1$, $\lambda 0$, $\lambda 1$ usw. gekennzeichnet, wobei $\lambda 0$ derjenige Rest ist, dessen Seitenkette gegen $W\omega 39$ packt und $\lambda 1$ derjenige Rest, dessen Seitenkette gegen $Y\omega 28$ packt. Die konservierten Wasserstoffbrücken ($\lambda-2$ zu $W\omega 39$ und $\lambda 1$ zu $S/T\omega 37$) sind gestrichelt dargestellt. Die N- und C-Termini der Domäne und des Liganden sind durch NT und CT gekennzeichnet. Die β -Stränge werden in N-C-terminaler Abfolge βA , βB und βC , die Schleifen entsprechend $\beta A\beta B$ - und $\beta B\beta C$ -Schleife genannt. (**B**) Schematische Darstellung der Interaktionen mittels LigPlot.¹⁹⁶ (**C**) Die Oberfläche der Domäne ist analog zu **A** eingefärbt. (**D**) Die Oberfläche³⁹ der Domäne ist an Hand der Hydrophobizität⁶⁰ eingefärbt: hydrophobe Bereiche gelb und hydrophile Bereiche grün.

Die Gruppen I und IV dieser Klassifizierung sind dabei sowohl strukturell (Abbildung 1-2),^{138,193} als auch funktionell relativ konsistent definiert. Dagegen sind für die anderen Gruppen weder die Konsensussequenzen der Liganden klar definiert, noch einzelne WW-Domänen eindeutig einer Gruppe zugeordnet.^{130,157,177,178} Auf Grund der unklaren Spezifität einer Vielzahl von WW-Domänen sollte in dieser Arbeit ein Verständnis der

Tabelle 1-3 Klassifizierung der WW-Domänen in Spezifitätsgruppen

Gruppe	Repräsentative Proteine	Konsensussequenz	Repräsentative Liganden
I	YAP65 DYSTROPHIN NEDD4	PPxY	PEBP2 Transkriptionsfaktor β-Dystroglycan ENaC Natriumkanal
II	FBP11 FE65	PPLP	Formin Mena
III	FE65 FBP21 HYP109 PQBP1	RPP, RxPP, PPR, PPxR, PGM	mRNA-Spleißfaktoren: SmB, U1C, NpwBP
IV	PIN1 ESS1	(poS/poT) P	RNA-Polymerase II Cdc25C Phosphatase

Diese Klassifizierung ist eine Konsolidierung verschiedener ähnlicher Klassifikationen.^{19,177,178}

Spezifitätsmechanismen entwickelt werden, welches unter anderem in einer verbesserten Klassifizierung der WW-Domänen an Hand ihrer Sequenz mündet.

Trotz großer struktureller Übereinstimmung binden die WW-Domänen der Gruppe I ihre Liganden in C→N-Orientierung, während WW-Domänen der Gruppe IV ihre Liganden in N→C-Orientierung binden (Orientierung des Liganden relativ zum ersten β-Strang βA der WW-Domäne).²¹³ Die Fähigkeit, prolinreiche Liganden in beiden Orientierungen binden zu können, beruht auf der Pseudosymmetrie der PPII-Helix (Abbildung 1-1).²¹² Im Gegensatz zu SH3-Domänen, die ebenfalls diese Fähigkeit besitzen,^{51,109} wurde für WW-Domänen einer Gruppe aber bislang nur entweder die eine oder die andere Orientierung beschrieben.

Die WW-Domänen der Gruppe IV unterscheiden sich von allen anderen dadurch, dass sie spezifisch phosphoserin(poS)- bzw. phosphothreonin(poT)-haltige Liganden mit einem **(poS/poT) P**-Motiv binden.^{130,193} Damit brechen diese WW-Domänen mit dem Paradigma der Bindung zweier konsekutiver Prolinreste durch die charakteristische hydrophobe Prolin-Bindungsstelle. Anstatt eines Prolins findet man in Komplexen der Gruppe IV WW-Domänen ein Phosphoserin oder Phosphothreonin an λ1.

Damit unterliegt die Wechselwirkung einer Gruppe IV WW-Domäne automatisch einer positiven Regulation: die Zielproteine werden nur dann gebunden, wenn ihre **SP**- oder **TP**-Motive spezifisch phosphoryliert wurden.¹⁷⁸ Im Gegensatz dazu gibt es nur wenige Hinweise auf eine möglicherweise negative Regulation der Wechselwirkung tyrosinhaltiger Liganden mit Domänen der Gruppe I durch Phosphorylierung des Tyrosins im Liganden.^{36,171}

1.3.3 Funktioneller Kontext der WW-Domänen

Die Domänenarchitektur für einige repräsentative WW-Domänen enthaltende Proteine ist schematisch in Abbildung 1-3 dargestellt. WW-Domänen können dabei sowohl als einzelne Domäne (*hYAP65*) als auch in mehreren Kopien vorkommen (*hNEDD4*).

Infolge der Verbreitung in Proteinen unterschiedlicher Funktion sind WW-Domänen an den verschiedensten zellulären Funktionen beteiligt: dazu gehören unter anderem der Ubiquitin-vermittelte Proteinabbau durch NEDD4-ähnliche Ubiquitinligasen (z.B. *hNEDD4*¹⁷² und *yRSP5*¹⁹⁷), die Vermittlung von Interaktionen mit dem Zytoskelett (z.B. *hDYSTRO*¹⁵⁰), die Transkription (z.B. *hPQBP1*⁹⁶ und *hYAP65*²⁰⁸), die Nachbearbeitung der mRNA durch Spleißen (z.B. *yPRP40*¹⁸) sowie die Regulation der Mitose (z.B. *hPIN1*^{164,207}).

Darüber hinaus besteht ein medizinisches Interesse an WW-Domänen, da sie direkt und indirekt in Zusammenhang mit einer Reihe menschlicher Krankheitsbilder gebracht werden.

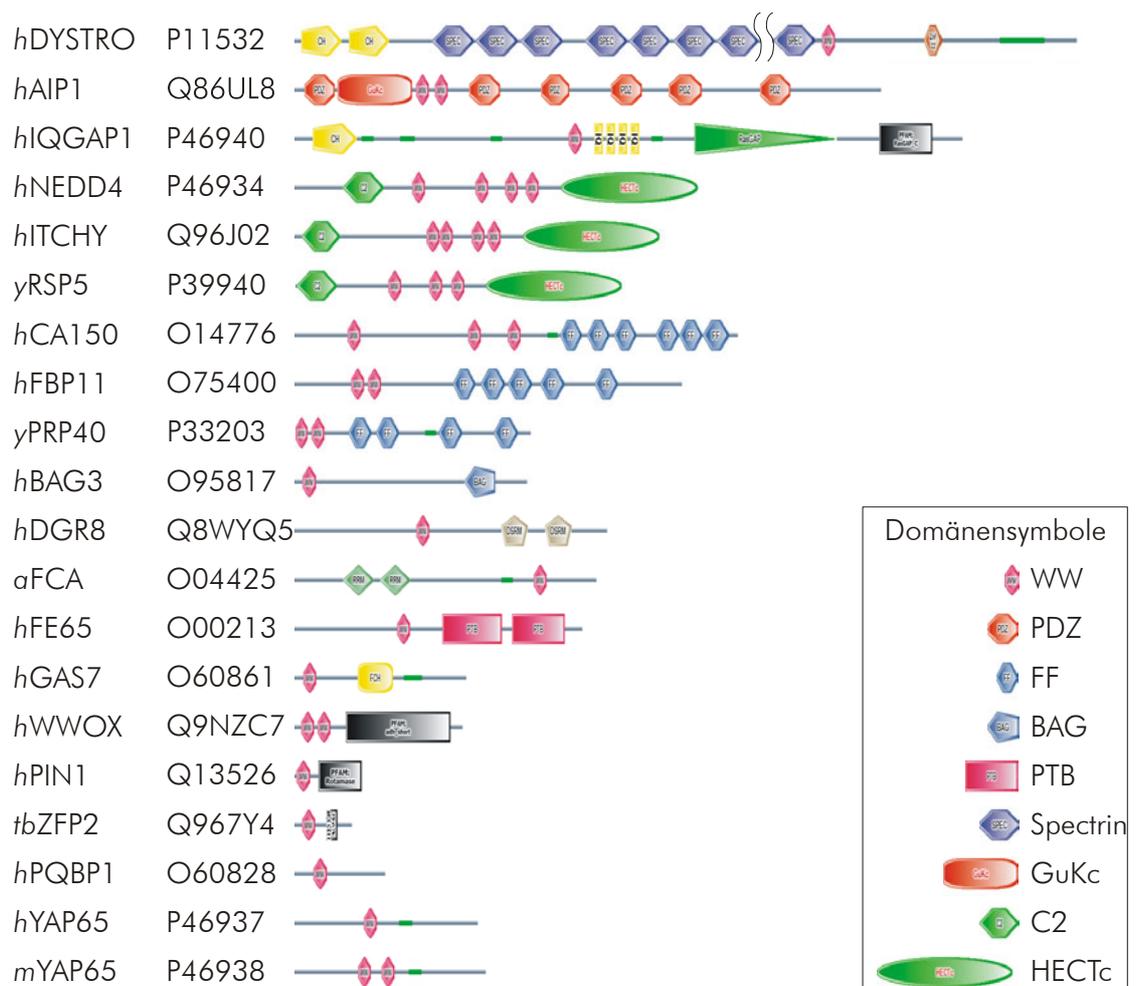


Abbildung 1-3 Domänenarchitektur repräsentativer WW-Domänen enthaltender Proteine

Domänenarchitektur nach SMART-Datenbank.¹⁶¹ Die häufigsten Domänensymbole sind in der Legende mit ihrer SMART-Bezeichnung aufgeführt. Für eine detaillierte Beschreibung dieser und weiterer Domänen, siehe <http://smart.embl-heidelberg.de> sowie im Text. Die Proteine sind durch ihre gängige Abkürzung und ihren primären UniProt¹⁰ Zugriffscode benannt.

So wird das Liddle-Syndrom, eine erbliche Bluthochdruckerkrankung, durch Störung des Ubiquitin-vermittelten Abbaus eines Natriumkanals ausgelöst. Auf molekularer Ebene basiert diese Störung auf einer Deletion oder Punktmutation des **xPPxY**-Zielmotivs in diesem Natriumkanal, mit dem normalerweise die WW-Domänen von NEDD4-Ubiquitin-Ligasen interagieren.¹⁷² Im Falle der Muskeldystrophie führen Deletionen in der C-terminalen Region des Dystrophins, in der sich unter anderem eine WW-Domäne befindet, zur schweren Form der Duchenne-Muskeldystrophie. Diese WW-Domäne ist für die Wechselwirkung mit einem **xPPxY**-Motiv des β -Dystroglycan verantwortlich und damit essentiell für die Verknüpfung des Zytoskeletts mit der extrazellulären Matrix.^{150,177} Auch im Zusammenhang mit der Alzheimer Krankheit scheinen mehrere Proteine mit WW-Domänen eine Rolle zu spielen: sowohl *hFE65*¹⁵⁴ als auch *hPIN1*¹¹⁰ wurden mit der Krankheit in Verbindung gebracht. Darüber hinaus wird ein Zusammenhang zwischen WW-Domänen enthaltenden Proteinen und Chorea Huntington,¹³² sowie der Entstehung maligner Tumore vermutet.^{206,207} Das in dieser Arbeit angestrebte Verständnis der Wechselwirkungs- und Spezifitätsmechanismen der WW-Domänen sollte nicht nur dabei helfen, die Rolle der WW-Domänen bei der Entstehung dieser Krankheiten zu verstehen, sondern es sollten auch neue Möglichkeiten für die therapeutische Intervention an diesen Wechselwirkungspunkten aufgezeigt werden.

1.4 PDZ-Domänen

PDZ-Domänen gehören zu den häufigsten Protein-Interaktionsdomänen (Tabelle 1-1).¹³⁴ Ihr Name leitet sich von den drei Proteinen ab, in denen sie entdeckt wurden: *post synaptic density protein 95* (PSD95),⁹⁴ *disc large* (Dlg)³⁸ und *zonula occludens 1* (ZO1).²⁰⁵

Strukturell sind PDZ-Domänen etwa 90 Aminosäuren lange globuläre Proteindomänen, die ein 6-strängiges β -Fass (β -Stränge β A bis β F) flankiert von 2 α -Helices (α A und α B) bilden (Abbildung 1-4).^{45,118} Auf der Sequenzebene wurden sie an Hand des charakteristischen **GLGF**-Motivs identifiziert, welches maßgeblich an der Bindung der extremen C-Termini vieler PDZ-Domänen-Liganden beteiligt ist (Tabelle 1-4 und Abbildung 1-4).⁹⁴ Funktionell spielen die von PDZ-Domänen vermittelten Wechselwirkungen mit Affinitäten im mikromolaren Bereich^{65,79} vor allem bei der Organisation der Komplexe in Signaltransduktionswegen,⁷⁹ dem Aufbau von Proteinnetzwerken, in submembranösen Kompartimenten^{94,98} und der Vermittlung von Wechselwirkungen mit dem Zytoskelett¹²⁴ eine Rolle (1.4.2).

Die konservierten Domänenpositionen werden an Hand ihrer Position relativ zum Beginn des vorausgehenden konservierten Sekundärstrukturelements bezeichnet, z.B. α B8 für den 8. Rest nach Beginn der zweiten α -Helix α B (siehe 2.1.2.2).

Tabelle 1-4 Sequenzalignment repräsentativer PDZ-Domänen

Domäne ^{a)}	UP-ID ^{b)}	Aminosäure-Sequenz ^{d)}		Domänengrenzen ^{e)}		
		Beginn	Ende	Beginn	Ende	
Domänenpositionen ^{c)}		3. 5. 1. 1. 3. 5. 1. 3. 5. 1. 3. 1. 3. 5. 1. 3. 1. 3. 5. 7. 9. 1. 3. 5.				
hPSD95-PDZ1	P78352	EITLERGN--SGLGFSIAGGTDNPHIGDD	---PSIFITKIIPG-GAAAQDGLRIVNDSILFVNE	---VDVREVTHSAAVEALREAGS---	IVRLVYM	108 - 192
hPSD95-PDZ2	P78352	EIKLIKGP--KGLGFSIAGGTVGNQHIPGD	---NSIYVTKIIEG-GAAHKDGLRLOIGDKILAVNS	---VGLIEDVMHEDAVAALNTYD---	VVYLKVA	203 - 287
hPSD95-PDZ3	P78352	RIVLHRGS--TGLGFNIIGGEDG	---EGEIFISFILAG-GPADLSEGLRKGQILSVNG	---VDLRNASHEQAALAKNAGQ---	TVTITIAQ	356 - 434
dDLG-PDZ1	P31007	DIQLERGN--SGLGFSIAGGTDNPHIGTD	---TSIYITKLSIG-GAAAADGRLSINDIIVSVND	---VSVVDVPHASAVDALRKAAG---	VVRLHVK	216 - 300
dDLG-PDZ2	P31007	EIDLKVGK--KGLGFSIAGGTVGNQHIPGD	---NGIYVTKITDG-GRAQVDGRISIGDKLIAVRTNGSEKNLENVTHELAVATLKSITD	---KVTLLIIG	330 - 418	
dDLG-PDZ3	P31007	TITITQGP--QGLGFNIIGGEDG	---QGIYVSFILAG-GPADLSEGLRKGQILSVNN	---VNLTHATHEEAQALATSGG---	VVTLIAQ	506 - 584
hZO1-PDZ1	Q07157	TVTILHRAPG-FGFGIATSGGRDNPFPQSGE	---TSIVISDVLKG-GPAEQG--LQENDRVAMVNG	---VSMDNVEHAFVQQLRKSQK---	NAKITIR	11 - 95
hZO1-PDZ2	Q07157	KVTILVSRKNEYGLRLAS	---HTFVKEISQD-SLAARDGNIQEGDVVLKING	---TVTENMSLTDKATLIERSKG---	KLRMIVQ	174 - 249
hZO1-PDZ3	Q07157	KLVEFRKG--DSVGLRLAGGND	---VGLFVAGVLED-SPAKEG--LEEGQILRVNN	---VDFNTIIREEAVLFLLDLPKGE--	EVTLIAQ	411 - 489
hNOS-PDZ	P29475	SVRLFRRK--GGLGFLVKERVSK	---PPVLIISDLIRG-GAAEQSGLIQAGDIIILAVNG	---RPLVDLSYDSALEVLVAGIAS---	ETHVLI	17 - 96
hLIMK1-PDZ	P53667	LVSIPASSH-GKRLSVSIDDPPHGGCGTEHSH	TVRVQGVDPGCMSPDVKNSIHVGDRIILEING	---TPIRNVPLDEIDLLIQETS---	LLQLTLE	165 - 255
dINAD-PDZ1	Q24008	MVTLDKTKG-KSFGICIVRGEVKDSPNKT	---TGLFIKGIIVPD-SPAHLICGRKLVGDRILSLNG	---KDVRSNTEQAVIDLLEKADF---	KILELEI	17 - 103
hPTPBL-PDZ2	Q64512	EVELAKTD--GSLGISVTTGVNNTSVRH	---GGIYVKAIIIPK-GAAESDGRHKGDRVLAVNG	---VSEGATEKQAVETLRNTGQ---	VVHLLIE	1357 - 1439
cLIN2-PDZ	P54936	LVOFKDTQ--EPMGITLKVND	---GRCFVARIHG-SPIHROATLHVGDIREING	---MSVANRSVESLQEMLRDARG---	QVTFKII	546 - 624
hDVL1-PDZ	O14640	TVTILNMRH--HFLGISIVGQSNDRGD	---GGIYIGSIMKG-GAVAADGRIEPGDMLLQVND	---VNFENMSNDDAVRVLREIVSQTGPISLTV	490 - 568	
hCASK-PDZ	O14936	LVOFKNTD--EPMGITLKMNEL	---NHCIVARIMHG-GMIHROGTLHVGDIREING	---ISVANQTVLQKMLREMRG---	SITFKIV	490 - 568
hEM55-PDZ	Q00013	LIOFEKVT--EPMGITLKLNEK	---QSCIVARILHG-GMIHROGSLHVGDILEING	---TNVINSVDLQKAMRETKG---	MISLKYI	71 - 149
hLIMK1-PDZ	P53667	LCCLEKGP--NGYGFHHLHGKKG	---LGOYIRLVEPG-SPAEKAG--LLAGDKLIQANG	---ENVEKETHQQVSRIRFAALN---	AVRLLIIV	14 - 91
hX11-PDZ1	Q02410	DVFTEKQK--EILGVVIVESGWGSIL	---PTVILIANMHRG-GPAEKSGKLNIGDQIMSING	---TSLVGLPLSTCQSLIKGLENQS--	RVKLNIV	656 - 740
hX11-PDZ2	Q02410	TVLIRRDRLRYQLGFSVON	---GIIICSLMRG-GIAERGG--VRVGHRIIEING	---QSVVATPEKIVHILLSNAVG---	EIHMKTM	747 - 820
rGRIP1-PDZ7	P97879	KVTILYKDSGMEDEFGRVADGLLE	---KGVYVKNIRPA-GPGDLGG--LKPYDRLLQVNH	---VRRDEDCDLVPLIAESGN---	KLDLVIS	988 - 1067
mSNA1-PDZ	Q61234	RVRVEKADA--GGLGISIKGGRENK	---MPLILSKIFRG-LAADQTEALFVGDALISVNG	---EDLSSATHDEAVQALAKTKTG---	EVVLEVK	81 - 161
hERBIN-PDZ	Q9NR18	RVRVEKAD--PELGFISGGVGGGRGNFRPDD	DGIFVTRVQPE-GPAEKL--LQPGDKLIQANG	---YSFINIEHQAVSLLKTFON---	TVLELIIV	1323 - 1407
hAF6-PDZ	P55196	TVTILKXQ--NMGSLSVAAKGAQDK	---LGIYVKSIVKRG-GAADVDGRLAAGDQLLSVDG	---RSLVGLSQRFAELMTRTSS---	VVTLIEVA	993 - 1073

Liganden-Bindungsstelle^{f)}

^{a)} Domänennamen gebildet aus Organismus-Präfix, Proteinamen und Domänen-Suffix (siehe 2.1.1).

^{b)} Primärer UniProt¹⁰ Zugriffscodes des die PDZ-Domäne enthaltenden Proteins.

^{c)} Die konservierten strukturellen Domänenpositionen sind nach der Standardnomenklatur durchnummeriert (siehe 2.1.2.2). Die Positionsbezeichnungen bestehen aus einem Sekundärstrukturpräfix und der Position, gezählt vom Anfang des jeweiligen Sekundärstrukturelements, z. B. β B1.

^{d)} Sequenzalignment der Sequenzen der PDZ-Domänen. Die Einfärbung konservierter Aminosäureeigenschaften basiert auf der Konsensussequenz-Färbemethode von CLUSTAL-X.¹⁸¹ Oberhalb des Sequenzalignments ist der Sekundärstrukturkonsensus schematisch dargestellt.

^{e)} Position der ersten und letzten Aminosäure der Domänensequenz innerhalb des Gesamtproteins.

^{f)} Aminosäurereste der PDZ-Domäne, welche die Oberfläche der Bindungsstelle formen, sind mit farbigen Streifen unter- und oberhalb des eigentlichen Alignments markiert. Diejenigen Reste, deren Seitenketten mit der konservierten Ligandenposition λ 0 interagieren, sind mit einem violetten Balken verbunden (IF0), mit $\lambda-1$ interagierende Reste mit einem blassgrünen Balken (IF-1), mit $\lambda-2$ interagierende mit einem orangenen (IF-2) und mit $\lambda-3$ interagierende mit einem dunkelgrünen (IF-3) (vergleiche Abbildung 1-4).

1.4.1 PDZ-Domänen/Liganden-Bindungsmechanismus

Charakteristisch für PDZ-Domänen ist ihre Fähigkeit spezifisch den C-Terminus ihrer Zielproteine zu binden.^{98,122} Dabei wird die C-terminale Carboxylgruppe des Liganden durch Wasserstoffbrücken mit den Hauptketten-Amidgruppen der GLGF-Schleife in der Bindungstasche zwischen β B und α B fixiert (Abbildung 1-4).⁹⁴ Gleichzeitig formen die vorausgehenden vier C-terminalen Ligandenreste einen zu β B antiparallelen β -Strang, der dadurch integraler Bestandteil des β -Faltblatts der PDZ-Domäne wird.⁷⁹

Als Folge dieses Bindungsmechanismus sind die Positionen der letzten vier C-terminalen Reste des Liganden relativ zur PDZ-Domäne stark konserviert.^{79,200} Dementsprechend können auch für die PDZ-Domänen/Liganden-Komplexe virtuelle Ligandenpositionen definiert werden. Die Position des C-terminalen Rests wird dabei als λ 0 bezeichnet. Die dazu N-terminal liegenden Reste besetzen entsprechend die Positionen $\lambda-1$, $\lambda-2$, $\lambda-3$ usw.

Auf Grund der Konservierung des Wechselwirkungsmechanismus lassen sich virtuelle Ligandenpositionen λ in PDZ-Domänen/Liganden-Komplexen definieren. λ 0 kennzeichnet dabei den C-terminalen Rest. Die vorausgehenden Reste werden folglich rückwärtszählend als $\lambda-1$, $\lambda-2$, etc. bezeichnet (siehe 2.1.2.2).

(Abbildung 1-4, siehe 2.1.2.2). Auf Grund dieser hochkonservierten Positionen interagieren die Seitenketten der letzten vier Ligandenreste mit vier definierten Bereichen der PDZ-Domäne, den sogenannten Interaktionsflächen IF0 bis IF-3.^{170,200} Die Form und Eigenschaften der Oberfläche dieser Interaktionsflächen definieren die Spezifität einer PDZ-Domäne für die letzten vier C-terminalen Reste ihrer Liganden.

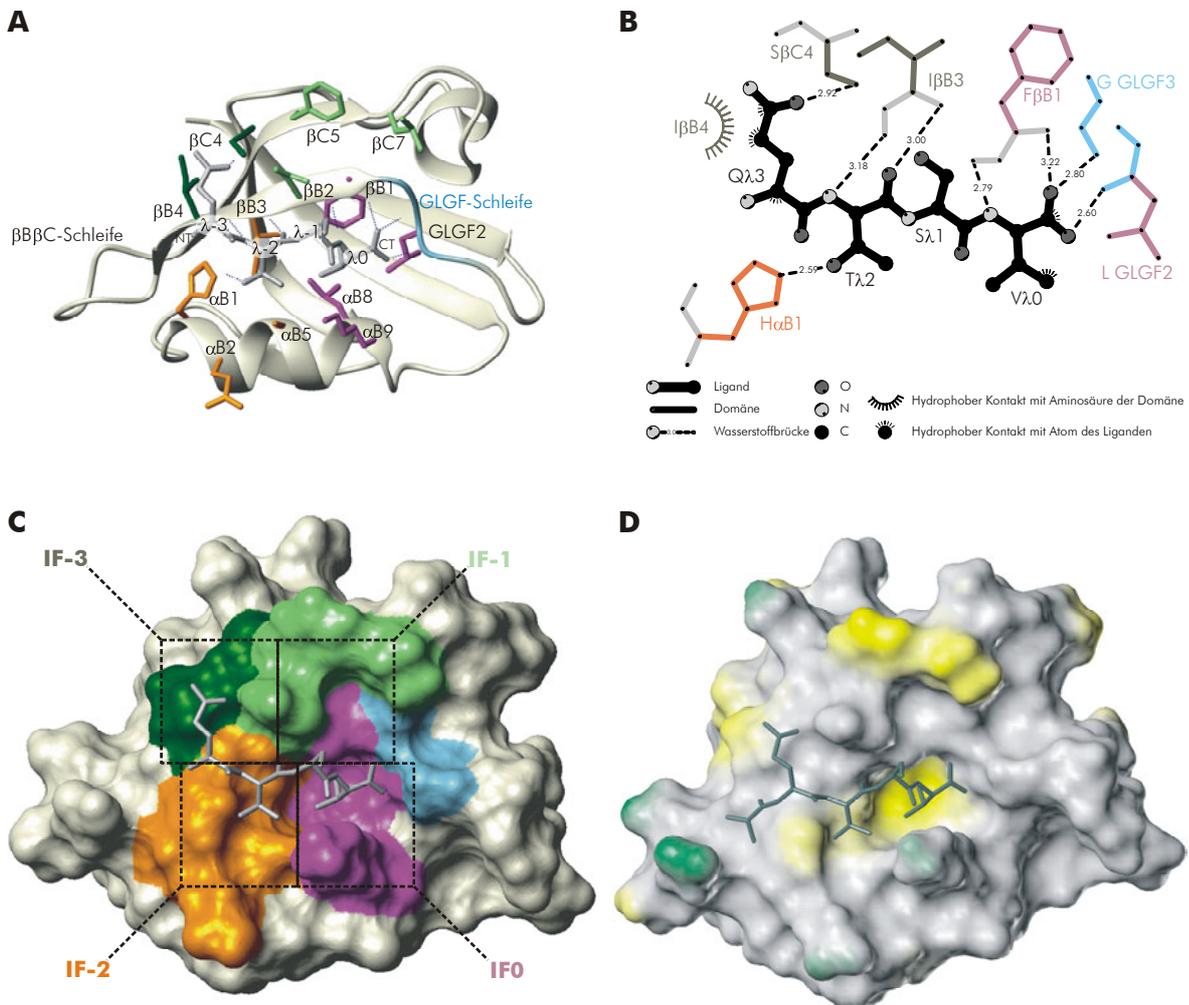


Abbildung 1-4 Struktur der rPSD95-PDZ3 Domäne im Komplex mit einem Peptid aus CRIPT

(A) Bändermodell¹⁵¹ der rPSD95-PDZ3 Domäne [PDB: 1BE9].⁴⁵ Die β-Stränge werden in N-C-terminaler Abfolge βA, βB usw., die α-Helices αA und αB, sowie die Schleifen entsprechend βAβB-, βBβC-Schleifen usw. genannt. Die konservierte GLGF-Schleife ist hellblau dargestellt. Die konservierten Domänenpositionen werden anhand ihrer Position relativ zum Beginn des vorausgehenden konservierten Sekundärstrukturelements bezeichnet, z.B. αB8 für den 8. Rest nach Beginn der zweiten α-Helix αB (siehe 2.1.2.2). Der Ligand (QTSV_{COOH}) aus β-Dystroglycan ist grau dargestellt. Die konservierten Ligandenpositionen werden ausgehend vom C-Terminus (CT) durch λ0, λ-1, λ-2 und λ-3 bezeichnet. Der Ligand wird in Form eines zusätzlichen antiparallelen β-Strangs gebunden. Die Hauptkette der Ligandenreste λ0 bis λ-3 bilden dazu typische Wasserstoffbrücken zur Hauptkette des βB-Strangs der Domäne aus (blau gestrichelte Linien). Die freie Hauptketten-Carboxylgruppe des λ0 Rests wird in der Bindungstasche durch Wasserstoffbrücken mit den Hauptketten-Amidgruppen der GLGF-Schleife fixiert. Die Seitenketten der Domänenreste der konservierten Bindungstasche bilden entsprechend für die Ligandenpositionen λ0 bis λ-3 die sogenannten Interaktionsflächen IF0 bis IF-3 aus. IF0 ist violett, IF-1 hellgrün, IF-2 orange und IF-3 dunkelgrün eingefärbt (siehe auch Tabelle 1-4). Die Wasserstoffbrücke zwischen Threonin/ Serin in λ-2 und Histidin in αB1 ist charakteristisch für die Erkennung von Liganden der Gruppe I (Tabelle 1-5). **(B)** Schematische Darstellung der Interaktionen mittels LigPlot.¹⁹⁶ **(C)** Die Interaktionsflächen IF0 bis IF-3 sind auf der Domänenoberfläche analog zu **A** eingefärbt. **(D)** Die Oberfläche³⁹ der Domäne ist an Hand der Hydrophobizität⁶⁰ eingefärbt: hydrophobe Bereiche gelb und hydrophile Bereiche grün.

Tabelle 1-5 Klassifizierung der PDZ-Domänen in Spezifitätsgruppen

Gruppe	Repräsentative Proteine	Konsensussequenz	Repräsentative Liganden
I	PSD95 SNA1 PICK1	x (S/T) xΦ_{COOH}	NMDA-Rezeptoreinheiten Spannungsabhängige Na ⁺ -Kanäle
II	GRIP CASK PICK1	xΦxΦ_{COOH}	Neurexin AMPA-Rezeptoreinheit GluR2
III	nNOS X11	x (D/E) xΦ_{COOH}	Melatoninrezeptor KIF17

Diese Klassifizierung ist eine Konsolidierung verschiedener ähnlicher Klassifikationen.^{79,88,170,189}

Basierend auf den Konsensussequenzen der selektiv gebundenen Liganden wurden auch die PDZ-Domänen in Spezifitätsgruppen eingeteilt (Tabelle 1-5). Erwartungsgemäß spielen für die Spezifität generell vor allem die letzten vier C-terminalen Reste eine maßgebliche Rolle.^{89,90,99} Dabei ist die hohe Selektivität für Aminosäuren in den Ligandenpositionen $\lambda 0$ und $\lambda - 2$ darauf zurückzuführen, dass die Seitenketten beider Reste in mehr ($\lambda 0$) oder weniger ($\lambda - 2$) ausgeprägte Taschen der PDZ-Domäne eintauchen (Abbildung 1-4). Dagegen schmiegen sich die Seitenketten der Reste in $\lambda - 3$ und $\lambda - 1$ eher an der Domänenoberfläche an.

In dieser Klassifizierung sind die Gruppen I und II allgemein anerkannt,¹⁷⁰ während weitere Gruppen in der Literatur uneinheitlich definiert und verwendet werden.^{21,65,79,88,189} Allerdings gibt es auch immer mehr Berichte über sogenannte dualspezifische PDZ-Domänen, d.h. PDZ-Domänen, die sowohl Liganden der Gruppe I als auch II binden.^{23,88,200} Darüber hinaus lassen sich eine ganze Reihe von PDZ-Domänen/Liganden-Wechselwirkungen keiner der konventionellen Gruppen zuordnen.⁸⁸ Versuche, die Spezifität mit einigen wenigen Sequenzmerkmalen der PDZ-Domänen zu korrelieren, sind bislang gescheitert.^{21,189} Auf Grund dieser unklaren Verhältnisse sollte in dieser Arbeit ein Verständnis der Spezifität von PDZ-Domänen entwickelt werden, um unter anderem die Dual-Spezifität zu erklären. Gleichzeitig sollte eine Methode zur quantitativen Analyse und Beschreibung der Spezifität von PDZ-Domänen entwickelt werden, welche nicht mehr den Einschränkungen der artifiziellen Einteilung in Spezifitätsgruppen unterliegt.

Zusätzlich zur Bindung C-terminaler Liganden können einige PDZ-Domänen, wie z.B. mSNA1-PDZ, auch als β -Finger exponierte interne Motive binden (siehe [PDB: 1QAV]).⁶⁹ Die β -Finger werden dabei über den gleichen Mechanismus wie C-terminale Peptide gebunden. Dabei verhindert die scharfe Biegung der β -Schleife in der zum C-Terminus ana-

Multi-PDZ Proteinen, die fast ausschließlich aus Tandem-Wiederholungen von PDZ-Domänen bestehen. Ihre Funktion ist wahrscheinlich die Vernetzung vieler Wechselwirkungspartner.

Prinzipiell spielen die meisten Proteine, die PDZ-Domänen enthalten, eine Rolle bei der Organisation von Proteinkomplexen sowie deren Lokalisation in definierte subzelluläre Kompartimente.⁶⁵ Dazu gehören unter anderem die Organisation von Signaltransduktionswegen (*dINAD*),¹¹⁷ die Aufrechterhaltung der Zellpolarität (*dPAR6*),¹²⁵ das *Protein-Sorting* (*mLIN10*, *mLIN7* und *mLIN2*),¹⁶³ die Ausbildung von Rezeptoren-Cluster (*hPSD95*)¹⁰⁶ sowie die Modulation der Aktivität des Bindepartners (*hNHERF* und *hCAP70*).²²

Auch für einige humane Krankheitsbilder konnte in den letzten Jahren ein Zusammenhang mit PDZ-Domänen-vermittelten Wechselwirkungen festgestellt werden. So wird das Usher-Syndrom (angeborene Taubheit und Blindheit) durch Mutationen in einem Gen ausgelöst, welches für das PDZ-Domänen enthaltende Protein Harmonin (*hHARM*) kodiert.^{24,194} Daneben konnten Verbindungen zwischen Mutationen in dem PDZ-Protein Pariaxin und der Dejerine-Sottas Neuropathie,^{26,165} der Deletion des Proteins Dishevelled (*hDVL1*) und dem DiGeorge-Syndrom¹⁴⁰ sowie der Deletion der LIM-Kinase-1 (*hLIMK1*) und dem Williams-Syndrom¹⁸⁰ nachgewiesen werden.

1.4.3 Untersuchte PDZ-Domänen

In dieser Arbeit wurden exemplarisch die PDZ-Domänen aus den Proteinen *hAF6*, *hERBIN* und *mSNA1* untersucht.

hAF6-PDZ

hAF6 [UP: P55196] ist das Protein des Gens *all-1 Fusion Partner on Chromosome 6*, welches häufig in akuten lymphoblastischen Leukämien auf Grund einer chromosomalen Translokation fusioniert mit dem Gen *all-1* vorliegt. Neben einer PDZ-Domäne enthält *hAF6* noch zwei *Ras-association* Domänen (RA), eine *Forkhead-association* Domäne (FHA) und ein *Dilute* Domäne (DIL), sowie mehrere prolinreiche Regionen (Abbildung 1-5). Damit kann *hAF6* sowohl als Adapterprotein fungieren als auch über die RA-Domänen in den MAP-Kinase-Signalweg eingreifen. Für *hAF6* wurde eine Beteiligung am Aufbau und der Aufrechterhaltung von Zell-Zell-Kontakten und der Zellpolarität, insbesondere in embryonalen Geweben, beschrieben.^{80,214} Daneben ist bekannt, dass *hAF6* über seine PDZ-Domäne mit verschiedenen Ephrin-Rezeptorkinase-Subtypen,⁷¹ dem *Breakpoint-Cluster-Region* (BCR) Protein,¹⁴⁷ dem *Junctional-Adhesion-Molecule* (JAM)⁴⁶ sowie dem Zell-Zell-Adhäsionsprotein Nectin¹⁷⁹ interagiert. Darüber hinaus bindet AF6 über die RA-Domänen nicht

nur die kleinen GTPasen der Ras-Familie¹⁴⁶ sondern auch ZO1,²⁰⁹ ein Protein der *Tight-Junction* Zell-Zell-Kontakte.

Das *ErbB2 interacting protein* (*hERBIN*, [UP: Q9NR18]) wurde ursprünglich als Wechselwirkungspartner der *Epidermal-Growth-Factor-Receptor related Kinase* (*ERB2*) identifiziert.²⁹ Es gehört zur Familie der *Leucine-Rich-Repeat-and-PDZ-containing* Proteine (*LAP*) und enthält neben einer PDZ-Domäne noch eine Region mit 16 *Leucine-Rich-Repeats* (*LRR*, siehe Abbildung 1-5).³³ *LAP*-Proteine scheinen als Gerüst-Proteine für die Aufrechterhaltung der spezifischen Zellmorphologie und der Apikal-Basal-Polarität von Epithelzellen verantwortlich zu sein. Über die PDZ-Domäne interagiert *hERBIN* unter anderem mit p120-ähnlichen Catenin-Proteinen,⁸⁵ der Integrin β 4-Untereinheit⁴⁹ und dem *ERB2*-Rezeptor,²⁹ für dessen basolaterale Lokalisation *hERBIN* verantwortlich zu sein scheint. Neben der Rolle als Adapterprotein, greift *hERBIN* auch regulierend in den MAP-Kinase-Signalweg ein, in dem es die Aktivierung von Raf-1 durch Ras über eine Wechselwirkung mit dem LRR-Motiv inhibiert.⁷⁷

hERBIN-PDZ

α 1-Syntrophin (*hSNA1*, [UP: Q13424]) ist ein Protein des Dystrophin-assoziierten Proteinkomplexes und verknüpft membranständige Rezeptoren mit dem Zytoskelett über die Wechselwirkung mit Dystrophin.⁵⁹ Die in dieser Arbeit untersuchte Domäne *mSNA1-PDZ* [UP: Q61234] des homologen murinen Proteins unterscheidet sich nur durch ein Alanin an Position α B7 anstelle eines Valins in *hSNA1-PDZ*. *hSNA1* besitzt neben einer PDZ-Domäne auch eine PH-Domäne (Abbildung 1-5). Über seine PDZ-Domäne lokalisiert *hSNA1* eine Reihe zytoplasmatischer Proteine in submembranöse Kompartimente im Bereich des Dystrophin-assoziierten Proteinkomplexes: dazu gehören die neuronale NO-Synthase (*nNOS*), deren PDZ-Domäne durch ein internes β -Fingermotiv gebunden wird,^{32,69} die Glutaminase L¹²⁶ und die stress-aktivierte Proteinkinase 3 (*SAPK3*).⁶⁶ Gleichzeitig rekrutiert *hSNA1* über seine PDZ-Domäne auch eine Reihe von Membranproteinen, unter anderem das Aquaporin4² und die spannungsabhängigen Natrium-Kanäle *CIN4* und *CIN5*.⁵⁹ Diese wiederum scheinen teilweise mit den zytoplasmatischen Effektorproteinen in funktionaler Beziehung zu stehen.

mSNA1-PDZ

1.5 Analyse und Vorhersage von Protein-Protein-Wechselwirkungen

Nach der Sequenzierung kompletter Genome, allen voran das des Menschen,^{103,192} ist die Aufklärung der komplexen Protein-Protein-Wechselwirkungsnetzwerke einer der nächsten Schritte auf dem Weg zum Verständnis der gesamten Organismen. Hierfür wurden eine Reihe experimenteller und theoretischer Methoden entwickelt, deren Spektrum von der Hochdurchsatz-Analyse kompletter Wechselwirkungsnetzwerke zur detaillierten Analyse einzelner Wechselwirkungen und von der rein experimentellen Analyse zur rein theoretischen Vorhersage reicht. Im Folgenden werden einige wichtige Entwicklungen der letzten Jahre vorgestellt, um die hier vorliegende Arbeit in einen Rahmen zu setzen. Diese befindet sich quasi im Zentrum des Spektrums, zwischen der Analyse einzelner Wechselwirkungen und ganzer Netzwerke. Vor allem liegt aber die Stärke der vorliegenden Arbeit in der effizienten Kombination der experimentellen Analysen und theoretischen Vorhersagen.

Die Wechselwirkungsmodelle der vorliegenden Arbeit basieren auf der Kombination experimenteller Analyse und theoretischer Vorhersage

1.5.1 Experimentelle Analyse von Protein-Protein-Wechselwirkungen

Zur systematischen experimentellen Analyse von Protein-Protein-Wechselwirkungsnetzwerken sind in den letzten Jahren eine ganze Reihe von Techniken entwickelt worden. Dazu gehören vor allem die *Yeast-Two-Hybrid* Technik,^{52,83,187} die Kombination von Tandem-Affinitäts-Reinigung und massenspektrometrischer Proteinidentifizierung^{57,70} und die Anwendung sogenannter Proteom-Chips.²¹⁵ Der Vergleich dieser Methoden an Hand der systematischen Analysen des Protein-Protein-Wechselwirkungsnetzwerks der Bäckerhefe (*Saccharomyces cerevisiae*) zeigt allerdings deutliche Schwächen auf: die identifizierten Netzwerke der verschiedenen Studien überlappen nur relativ gering.^{7,107,195} Scheinbar werden jeweils nur vergleichsweise kleine Ausschnitte des gesamten Wechselwirkungsnetzwerks identifiziert. Das liegt unter anderem daran, dass die einzelnen Methoden vorrangig verschiedene Arten von Wechselwirkungen detektieren. So lassen sich reversible Wechselwirkungen besser durch die *Yeast-Two-Hybrid* Technik identifizieren, während permanente Wechselwirkungen hauptsächlich durch Affinitäts-Reinigung detektiert werden können.⁷ Hinzu kommt das viel grundlegendere Problem, dass Protein-Protein-Wechselwirkungsnetzwerke hochdynamische Systeme sind, die sich nicht einfach durch einzelne Momentaufnahmen erfassen lassen. Erst die Kombination der verschiedenen experimentellen Ergebnisse

Systematische Analyse des Protein-Protein-Wechselwirkungsnetzwerks

Geringe Überlappung der globalen Analysen

Präferenzen der einzelnen Techniken

Dynamik der Wechselwirkungsnetzwerke

Kombination experimenteller und theoretischer Analysen

und zusätzlich die theoretische Modellierung von Wechselwirkungen wird daher letztendlich ein Gesamtbild ergeben.^{12,84,190}

Die reversiblen und damit hochdynamischen Wechselwirkungen der Protein-Interaktionsdomänen werden gerade aus diesen Gründen nur in geringem Maße von den globalen Ansätzen erfasst. Auf Grund der meist linearen Sequenzmotive, die von Protein-Interaktionsdomänen in ihren Zielproteinen erkannt werden, kann das Wechselwirkungsspektrum dieser Domänen effizient – allerdings nur *in vitro* – mit Hilfe von Phage-Display^{55,167,168} und Peptid- bzw. Protein-Bibliotheken^{28,74,170} untersucht werden. Diese Techniken wurden in letzter Zeit auch für genomweite Analysen aller Domänen einer Familie verwendet, unter anderem für WW-Domänen^{75,130} und SH3-Domänen.^{104,184}

Analyse der Wechselwirkungen von Protein-Interaktionsdomänen

Protein-Interaktionsdomänen erkennen meist lineare Sequenzmotive

1.5.2 Theoretische Vorhersage von Protein-Protein-Wechselwirkungen

Klassischerweise wird eine Familie von Protein-Interaktionsdomänen basierend auf der Domänen/Liganden-Spezifität in sogenannte Spezifitätsgruppen eingeteilt.^{19,79,170,177,178,189} Solche Klassifizierungssysteme auf Ebene der Ligandensequenzen wurden für die WW- und PDZ-Domänen schon vorgestellt (siehe 1.3.2 und 1.4.1 sowie Tabelle 1-3 und Tabelle 1-5). Ihre Aussagekraft ist aber limitiert, da einerseits die Zuordnung einzelner Domänen auf Grund teilweise widersprüchlicher experimenteller Ergebnisse unklar ist. Andererseits lassen sich Domänen nicht allein auf der Basis ihrer Sequenzen zuordnen, da für die meisten Spezifitätsgruppen die charakteristischen spezifitätsbestimmenden Reste der Domänen nicht definiert sind. Nur sie würden aber die Vorhersage der Spezifitätsgruppe für eine gegebene Domäne allein an Hand der Domänensequenz erlauben.¹⁸⁹ Für die WW-Domänen konnte solch ein gruppenspezifisches Sequenzmotiv nur für die tyrosinspezifischen Domänen identifiziert werden.^{76,113} Im Falle der PDZ-Domänen wurde bislang nur die Korrelation zwischen einem Histidin in der Position $\alpha B1$ (Abbildung 1-4) und der Bindung von Gruppe I Liganden (Tabelle 1-5) identifiziert. Allerdings schließt ein Histidin in dieser Position auch nicht die Bindung von Gruppe II Liganden aus.^{23,88} Weitere Versuche dieser Art waren ebenfalls erfolglos.^{21,189}

Klassifizierung von Protein-Interaktionsdomänen an Hand ihrer Spezifität

Limitierung der vorhandenen Klassifizierungssysteme

Unabhängig von der speziellen Anwendung für Protein-Interaktionsdomänen gibt es eine Reihe von Methoden zur Identifizierung solcher Sequenz-Aktivitäts-Beziehungen: dazu gehören die *Correlated Mutation Analysis*,⁶¹ die *in silico Two-Hybrid* Technik (i2H),¹³⁵ die *Prediction Of Unknown Sub-types* (PROUST),⁶⁴ *Support-Vector-Machine*-basierte Methoden²⁵ und

Sequenzbasierte Methoden: Sequenz-Aktivitäts-Beziehungen

die *Analysis of Complementary Residues*.²⁰¹ Die Anwendung der meisten dieser Methoden auf Protein-Interaktionsdomänen ist schwierig, vor allem auf Grund der relativ großen Anzahl an Liganden pro Domäne. Dagegen nutzt die iSPOT Methode³¹ gerade diesen Aspekt erfolgreich für die Vorhersage aus. iSPOT analysiert unter Einbeziehung struktureller Information die Korrelation der experimentell bestimmten Aminosäurepräferenzen an den einzelnen Ligandenpositionen mit der Aminosäuresequenz der Bindungsstelle in den Domänen.

Prinzipiell bieten Proteinstrukturen eine Fülle zusätzlicher Informationen, allen voran die räumliche Anordnung der Aminosäuren und damit die Form und Eigenschaft der Protein-oberfläche, die eine erweiterte Bewertung der Wechselwirkung zweier Proteine ermöglicht.⁷

Aus dem Bereich des *rational drug designs*¹⁰¹ sind auch einige erfolgreiche Methoden zur Vorhersage von Wechselwirkungen auf der Ebene der Proteinstruktur bekannt, insbesondere

das *Molecular docking*.^{63,169} Die klassischen Docking-Methoden lassen sich jedoch nicht einfach auf Protein-Interaktionsdomänen übertragen, da sowohl die peptidischen Liganden zu flexibel, als auch die Bindungsstellen auf den Domänen zu flach sind. Diesem Problem

begegnet eine Gruppe anderer Vorhersagemethoden mit der Annahme einer bekannten räumlichen Anordnung der Wechselwirkungspartner zueinander. Die Anordnung wird dabei der Struktur homologer Komplexe entnommen. Damit wird das Problem der Vorhersage vorrangig auf die Bewertung der Komplementarität der Eigenschaften bzw. Strukturkompatibilität der Sequenzen reduziert. Erfolgreich angewendet wurden unter anderem die InterPreTS-Methode auf eine Vielzahl von Proteinkomplexen der Hefezelle,^{5,6} die Protein-Designmethode Perla auf das Design von PDZ-Domänen mit veränderter Spezifität,¹⁴⁹ die COMBINE-Methode auf das Design von Neuramidase-Inhibitoren¹⁹⁸ sowie eine ganze Reihe verschiedener strukturbasierter Methoden auf die Vorhersage von Haupt-Histokompatibilitäts-Komplex (MHC)/Peptid- und T-Zell-Rezeptor/Peptid-Wechselwirkungen.^{152,159}

Ein weiterer Vorteil der strukturbasierten Methoden ist, dass die Erkenntnisse, welche für Protein-Protein-Wechselwirkungen gewonnen wurden, in das Design von kleinen organischen Substanzen einfließen können. Teilweise lassen sich die resultierenden Wechselwirkungsmodelle direkt für die Vorhersage der Wechselwirkungen mit Peptidomimetika verwenden.¹²⁷ Allerdings wird die Kenntnis der Struktur des Proteins und zum Teil des Liganden vorausgesetzt, was häufig eine schwerwiegende Anforderung darstellt.

Strukturbewusste
Methoden auf
Sequenzebene

Strukturbasierte
Methoden:
Struktur-Aktivitäts-
Beziehungen

Docking

Limitierungen des
Dockings für Protein-
Interaktionsdomänen

Homologiebasierte
Methoden

Transferleistung
strukturbasierter
Methoden

1.6 Zielsetzung und Strategie

Das Verständnis von Protein-Protein-Wechselwirkungen ist auf Grund ihrer Bedeutung für die molekulare Funktion zellulärer Prozesse eines der fundamentalen Ziele der biomedizinischen Forschung. Auch im Zusammenhang mit pathologischen Dysfunktionen, wie zum Beispiel maligner Tumore, Alzheimer oder Parkinson, spielen Protein-Protein-Wechselwirkungen eine entscheidende Rolle. Sie stellen damit interessante, wenn auch schwer beeinflussbare, molekulare Angriffspunkte für potentielle Therapien dar. Hier soll das Verständnis der molekularen Wechselwirkungsmechanismen zu Ideen führen, wie an diesen Angriffspunkten interveniert werden kann.

Relevanz

Das Ziel der vorliegenden Arbeit ist, Erkenntnisse über die Wechselwirkungs- und Spezifitätsmechanismen sowie die Selektivität von Protein-Interaktionsdomänen im Allgemeinen sowie von WW- und PDZ-Domänen im Speziellen zu gewinnen. Dazu sollten statistische Wechselwirkungsmodelle auf der Basis experimenteller Wechselwirkungsdaten in Kombination mit der theoretischen Analyse der Sequenzen und Strukturen für repräsentative Vertreter beider Domänenfamilien trainiert werden. Mit diesen Wechselwirkungsmodellen sollten nicht nur neue potentielle Wechselwirkungspartner für experimentell nicht untersuchte Domänen bzw. Liganden, sondern auch die entsprechenden Wechselwirkungsaffinitäten vorhergesagt werden.

Verständnis der Wechselwirkungs- und Spezifitätsmechanismen

Training statistischer Wechselwirkungsmodelle

Vorhersage von Wechselwirkungspartner und -affinitäten

Die praktische Anwendung der erstellten Wechselwirkungsmodelle sollte durch das rationale Design spezifischer Liganden mit gesteigerter Affinität demonstriert werden. Solche designte Liganden wären vielfältig einsetzbar, beispielsweise als Kompetitoren in zellbiologischen Experimenten oder als Leitstrukturen für das Design kleiner pharmakologischer Substanzen zur Intervention an definierten Protein-Protein-Wechselwirkungen.

Rationales Design peptidischer Liganden

Ein weiteres Ziel bestand darin, die Leistung und die Anforderungen verschiedener Methoden zur Erstellung von Wechselwirkungsmodelle auf der Sequenz-, der Tertiärstruktur- und der Quartärstrukturebene zu vergleichen. Dabei wurde insbesondere die Anwendbarkeit der *Comparative-Molecular-Field-Analyse* (CoMFA der Domänenstrukturen) und der *Comparative-Binding-Energy-Analyse* (COMBINE der Domänen/Liganden-Komplexstrukturen) auf Protein-Interaktionsdomänen untersucht (für Grundlagen dieser Methoden siehe Anhang D). Gleichzeitig wurde damit die Fragestellung bearbeitet, in wieweit sich die für wenige Vertreter einer Familie gewonnen Erkenntnisse durch die verschiedenen Wechselwirkungsmodelle auf die übrigen Mitglieder übertragen lassen.

Wechselwirkungsmodelle auf verschiedenen Informationsebenen

Quantifizierung der
Spezifität

Schließlich wurde mit dem Ziel, die Spezifität einer Protein-Interaktionsdomäne zu quantifizieren, die Methode der *Quantitativen Spezifitäts-Profile* (QSP) entwickelt. Die generierten QSP-Modelle beschreiben dabei die Spezifität als kontinuierliche Affinitätsfunktion über den gesamten potentiellen Liganden-Sequenzraum (siehe 1.2). Damit sollte die Vorhersage der Bindungsaffinität für alle potentiellen Ligandensequenzen einer Protein-Interaktionsdomäne und somit ein quantitativer Vergleich der Selektivitäten verschiedener Domänen ermöglicht werden. Letztendlich bilden QSP-Modelle eine wertvolle Grundlage für die Simulation von Domänen-vermittelten Wechselwirkungen auf dem Gebiet der Systembiologie.^{72,174,216}

Quantifizierung der
Selektivität