

# **Untersuchung von multivalenten Protein-Protein Interaktionen zwischen WW-Domänen und prolinreichen Sequenzen**

Dissertation zur Erlangung des akademischen Grades des Doktors der  
Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie  
der Freien Universität Berlin

vorgelegt von

**Anne Wartenberg**

aus Prenzlau

2012



---

Diese Arbeit wurde von April 2008 bis Oktober 2012 am Leibniz-Institut für Molekulare Pharmakologie (FMP) unter Anleitung von Prof. Dr. Hartmut Oschkinat angefertigt.

1. Gutachter: Prof. Dr. Hartmut Oschkinat

Leibniz-Institut für Molekulare Pharmakologie (FMP)

2. Gutachter: Prof. Dr. Christian Freund

Freie Universität Berlin

Disputation am: 25.04.2013

---

---

---



---

*„Ohne Spekulation gibt es keine neue Beobachtung“  
Charles Darwin*

---



## **DANKSAGUNG**

Die hier dargestellte Arbeit wurde am Leibniz-Institut für Molekulare Pharmakologie in der Arbeitsgruppe „NMR unterstützte Strukturbioogie“ unter Betreuung von Prof. Dr. Hartmut Oschkinat angefertigt. Hiermit möchte ich mich ganz herzlich bei allen Mitgliedern dieser Arbeitsgruppe für die erfolgreiche Zusammenarbeit und die fantastische Arbeitsatmosphäre bedanken. Einigen Personen möchte ich wegen ihres bewundernswerten Engagements und ihrer fortwährenden Unterstützung ganz besonders danken, da der Erfolg dieser Arbeit ohne sie undenkbar gewesen wäre.

Ich möchte mich zu allererst bei meinem Betreuer und Mentor Prof. Hartmut Oschkinat bedanken. Danke, dass Du mir die Möglichkeit gegeben hast, dieses spannende Thema zu bearbeiten und mich durch inspirierende, aber auch manchmal kritische Diskussionen immer wieder herausgefordert hast.

Herrn Prof Dr. Christian Freund möchte ich für die Übernahme des Koreferates und die kritischen und hilfreichen Diskussionen insbesondere am Anfang meiner Promotion danken.

Frau Dr. Anne Diehl danke ich für die sehr freundliche Aufnahme in ihr Laborteam, ihre andauernde Hilfs- und Diskussionsbereitschaft und ihre exzellente und so wichtige Unterstützung bei der Herstellung rekombinanter Proteine. Danke, dass Du sowohl bei beruflichen als auch bei privaten Problemen immer eine Lösung parat hast.

Dr. Peter Schmieder und Monika Beerbaum möchte ich für die Aufnahme der Triple-Resonanzspektren und für die kompetente Beratung bei allen „NMR-Fragen“ danken.

Bei Frau Dagmar Krause & Frau Angelika Ehrlich aus der Arbeitsgruppe Beyermann möchte ich mich für die Herstellung der beiden synthetischen WW-Domänen, der prolinreichen Peptidliganden sowie der SPOT-Membranen bedanken.

Dr. Michael Schümann danke ich für die massenspektrometrische Analyse der SILAC-Proben und dafür, dass er sich die Zeit genommen hat, alle meine Fragen bezüglich der wirklich sehr „umfangreichen SILAC-Excel-Tabellen“ zu beantworten.

Jun. Prof. Dr. Sandro Keller, Dir möchte ich besonders für die stets schnellen und ausführlichen Antworten auf alle meine „ITC-Fragen“ und für die mentale Unterstützung danken. Dein außergewöhnliches Engagement ist einfach bewundernswert!

---

Bei meinen lieben Laborkollegen, insbesondere Dr. Janet Zapke, Anja Voreck, Dr. Vivien Lange, Dr. Janina Hahn, Doc Hiller, Dr. Tobias Werther, Nils Cremer und meinem „Banknachbarn“ Paul Grzesik möchte ich mich für die wunderbare Laboratmosphäre und für so manche kleine Ablenkungen vom Laboralltag bedanken.

Marcel Jurk möchte ich für seine tollen Phyton-Skripte danken, die er mir so bereitwillig zur Verfügung gestellt hat. Außerdem danke ich Dir für Deine Hilfsbereitschaft bei der Durchführung von AUZ-Läufen und der CD-Spektroskopie und für die vielen hilfreichen Diskussionen.

Liebe Natalja, liebe Lilo und liebe Martina. Danke für die unterhaltsamen Mittagspausen, bei denen man so schön „abschalten“ konnte!

Bei Dr. Barth van Rossum möchte ich mich für die wunderschöne Abbildung der PPII-Helix bedanken.

Nicht zuletzt möchte ich den Sekretärinnen Katharina Tebel, Katrin Wittig, Andrea Steuer, Alex Kiesling und Heidi Petschick danken. Ihr macht einen ganz tollen Job und behaltet alles im Überblick.

Ganz herzlich möchte ich mich auch bei Dr. Britta Kunert bedanken. Du bist eine echte Freundin! Danke für Deine Ehrlichkeit, Deine Unterstützung und dafür, dass Du immer ein offenes Ohr für mich hast. Was würde ich nur ohne unsere „Telefonsessions“ machen!

Meinen lieben Eltern und meinem Opa Fredi möchte ich für ihre Unterstützung während meiner Ausbildungsjahre danken. Besonders meinem Papa und meinem Opa Fredi danke ich für die lieben aufmunternden Worte während der „Schreibphase“. Ich habe endlich fertig!

Zu guter Letzt möchte ich einem ganz besonderen Menschen danken, Matthias Dorn. Du bist zum wichtigsten Menschen in meinem Leben geworden. Dein grenzenloses Vertrauen in mich und das Gelingen meiner Arbeit hat mir unendlich geholfen, diese schwierige Phase zu meistern. Danke für Deine Unterstützung bei der Auswertung der Relaxations- und AUZ-Daten sowie für die Anleitung zur Aufnahme heteronuklearer NMR-Spektren. Du hast mir gezeigt, dass ich mich auch in schwierigen Situationen immer auf Dich verlassen kann! Danke für Deine Liebe und Deine Geduld!

---

# **INHALT**

Inhalt .....	3
Abbildungen.....	6
Tabellen.....	8
Abkürzungen .....	10
1 Einleitung.....	12
1.1 Die Bedeutung von Proteinen und Protein-Interaktionen .....	12
1.2 Proteindomänen I - modulare Funktionseinheiten.....	14
1.3 Die Multivalenz .....	15
1.1.1 Das Prinzip der Multivalenz.....	15
1.1.2 Thermodynamische Ansätze.....	16
1.3.1 Multivalenz in biologischen Systemen .....	19
1.4 PROTEINDOMÄNEN II – PRD, PRS & WW-Domänen .....	20
1.4.1 Prolinreiche Sequenzen-bindende Domänen (PRD).....	20
1.4.1.1 PPII-Helices und Prolinreiche Sequenzen (PRS) .....	21
1.4.2 Die WW-Domäne.....	22
1.4.2.1 Strukturelle Eigenschaften und Organisation .....	24
1.4.2.2 Bindeeigenschaften.....	25
1.4.2.3 Klassifizierung der WW-Domänen anhand ihrer Bindungsmotive.....	26
1.4.2.4 Klinische Relevanz.....	27
1.5 CA150, ein WW-Domänen-haltiger Transkriptionsfaktor.....	28
1.5.1 Strukturelle Merkmale .....	28
1.5.1.1 Strukturelle Aspekte der drei CA150-WW-Domänen.....	28
1.5.2 Modifikationen von CA150 .....	30
1.5.3 Pharmakologische Bedeutung von CA150 in Huntington .....	30
1.5.4 Physiologischer Kontext von CA150.....	31
2 Zielstellung .....	34
3 Material und Methoden.....	36
3.1 Chemikalien, Materialien und Geräte .....	36
3.2 Verwendete Software .....	38
3.3 Molekularbiologische Arbeiten .....	39
3.3.1 Oligonukleotide und Vektoren .....	39
3.3.2 Polymerase-Ketten-Reaktion (PCR).....	40
3.3.3 DNA-Gelelektrophorese.....	41
3.3.4 DNA-Gelextraktion.....	41
3.3.5 Klonierung .....	41
3.3.6 Herstellung chemisch kompetenter E. coli-Zellen .....	41
3.3.7 Transformation von E. coli .....	42
3.3.8 Plasmidpräparation.....	42

---

3.3.9	Verifizierung der Plasmide .....	42
3.3.10	Gerichtete Mutagenese .....	42
3.4	Biotechnologische Methoden .....	43
3.4.1	Escherichia coli Stämme .....	43
3.4.2	Anzuchtmedien für Bakterienkulturen .....	43
3.4.3	Glycerin-Dauerkulturen .....	44
3.4.4	Anzuchtbedingungen .....	44
3.4.5	Kontrolle der Proteinexpression .....	44
3.4.6	Ermittlung optimaler Expressionsbedingungen .....	44
3.4.7	Expression isotopenmarkierter WW-Domänen .....	45
3.5	Reinigung rekombinanter Proteine .....	45
3.5.1	Zellaufschluss .....	45
3.5.2	Proteinreinigung über Metall-Chelat-Chromatographie .....	46
3.5.3	Abspaltung des Reinigungs-Tags .....	46
3.5.4	Proteinreinigung durch Gelfiltration .....	46
3.5.5	Konzentration von Proteinen .....	46
3.6	Biochemische Analysemethoden .....	47
3.6.1	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese .....	47
3.6.2	Berechnung von Proteinkonzentrationen .....	47
3.6.3	Prolinreiche Peptidliganden für NMR-Titrationsen und ITC-Experimente .....	49
3.6.4	SPOT-Analysen .....	50
3.6.4.1	Herstellung der Zellulosemembran-gebundenen Peptidliganden .....	50
3.6.4.2	Durchführung von SPOT-Analysen .....	51
3.7	Lösungs-NMR Spektroskopie .....	51
3.7.1	Aufnahme mehrdimensionaler, heteronuklearer NMR-Spektren zur sequentiellen Zuordnung der Resonanzen des Proteinrückgrates .....	52
3.7.2	Analyse der Interaktion zwischen CA150-WW-Domänen und prolinreichen Peptidliganden .....	52
3.7.3	Messung <sup>15</sup> N-basierter Spinrelaxation ( <sup>15</sup> N-R1, <sup>15</sup> N-R2 & <sup>1</sup> H <sup>15</sup> N-HetNOE) .....	53
3.8	Massenspektrometrische Analyse (MS) .....	53
3.9	Isotherme Titrationskalorimetrie (ITC) .....	54
3.10	Analytische Ultrazentrifugation (AUZ) .....	54
3.11	CD-Spektroskopie .....	55
3.12	Triple-SILAC Pulldown .....	56
3.12.1	SILAC: Kultivierung der HeLa-Zellen .....	57
3.12.2	Pulldown .....	57
3.12.3	SILAC: MS-Analyse .....	57
4	Ergebnisse .....	59
4.1	Klonierung WW-Domänen-codierender-Vektoren .....	59
4.2	Expressionsanalyse der WW-Varianten .....	60
4.3	Präparation der WW-Varianten .....	62

---

---

4.4	Analyse der WW-Varianten.....	63
4.4.1	NMR-spektroskopische Untersuchung des Faltungszustandes verschiedener WW-Varianten .....	64
4.4.2	Massenspektrometrische Analyse von uniform $^{15}\text{N}^{13}\text{C}$ -markierter WW23-F7 .....	65
4.4.3	CD-spektroskopische Sekundärstrukturanalyse von WW23-F7 .....	66
4.5	NMR-Spektroskopische Untersuchungen.....	68
4.5.1	Sequentielle Zuordnung von WW23-F7.....	68
4.5.2	Strukturdynamische Eigenschaften von WW23-F7 .....	70
4.6	Untersuchung von monovalenten und multivalenten SF1 <sub>PRS</sub> /CA150 <sub>WW</sub> -Wechselwirkungen .....	74
4.6.1	SPOT-Analyse.....	74
4.6.1.1	Synthetische Einzel-WW-Varianten (WW2 und WW3) .....	77
4.6.2	$^1\text{H}^{15}\text{N}$ -HSQC-Titrationsexperimente mit WW23 .....	79
4.6.2.1	Erzeugung von WW23-Mutanten.....	83
4.6.3	NMR-Titrationsexperimente mit CA150 WW23-Mutanten .....	86
4.6.4	Isotherme Titrationskalorimetrie .....	90
4.7	Triple SILAC Pulldown-Experiment .....	95
4.8	Studien zur CA150 <sub>PRS</sub> /CA150 <sub>WW</sub> Interaktion.....	97
4.8.1	SPOT-Analyse.....	97
4.8.2	NMR-Titrationsexperimente .....	98
4.8.3	Isotherme Titrationskalorimetrie (ITC) .....	101
4.8.4	Analytische Ultrazentrifugation (AUZ).....	101
4.9	Interaktion von CA150-WW23 mit „Kurzen“ PP-Peptiden .....	104
5	Diskussion .....	106
5.1	Funktionelle Aktivität der verwendeten WW-Varianten .....	106
5.2	Die strukturelle Organisation von CA150 WW23 .....	108
5.3	Interaktionen zwischen CA150-WW-Domänen und der SF1 <sub>PRS</sub> .....	110
5.4	Interaktionen zwischen CA150-WW-Domänen und der N-terminalen CA150 <sub>PRS</sub> .....	112
5.4.1	Peptidinduzierte Strukturveränderung von WW23 .....	113
5.5	„xP“- und „xP2“- Bindetasche.....	114
5.5.1	Monovalente oder Bivalente Bindung.....	117
5.5.2	Bindemotive der Einzel-WWs (WW1, 2&3).....	119
5.5.3	Modelle multivalenter CA150 <sub>WW</sub> /CA150 <sub>PRS</sub> -Wechselwirkungen.....	123
	Zusammenfassung .....	126
	Summary .....	128
	Literatur .....	130
	Eidesstattliche Erklärung.....	141
6	Anhang .....	142
6.1	Sequenzinformationen.....	142
6.2	NMR-Daten .....	143

---

## **ABBILDUNGEN**

Abbildung 1: Beispiele multivalenter Systeme .....	15
Abbildung 2: Typen multivalenter Interaktionen .....	16
Abbildung 3: Thermodynamische Gleichgewichte einer mono- und bivalenten Interaktion .....	17
Abbildung 4: Der Einfluss des Linkers auf die Effektive Konzentration .....	18
Abbildung 5: Das Hämagglutinin des Influenzavirus bindet auf der Zelloberfläche an den trivalenten Rezeptor aus Sialinsäure .....	19
Abbildung 6: PPII-Helix aus 5 Prolin-Resten.....	21
Abbildung 7: Struktur und Bindemechanismus der Dystrophin WW-Domäne .....	26
Abbildung 8: Modulare Architektur des Transkriptionsfaktors CA150 .....	28
Abbildung 9: Sequenzvergleich der CA150 WW-Domänen.....	28
Abbildung 10: Backbone-Strukturen der CA150-WW-Domänen.....	29
Abbildung 11: Native und fibrilläre Struktur von CA150 WW2 .....	29
Abbildung 12: Netzwerk ausgewählter CA150 (TCERG1) Interaktionen.....	31
Abbildung 13: Schema der Expressionsuntersuchungen .....	45
Abbildung 14: Schema des Triple-SILAC-Pulldowns.....	56
Abbildung 15: Übersicht der WW-Domänen-Varianten .....	59
Abbildung 16: Vergleich der Löslichkeit von CA150 WW23-F6 und CA150 WW23-F7.....	62
Abbildung 17: IMAC und TEV-Spaltung von WW23.....	62
Abbildung 18: Gelfiltration von WW23 .....	63
Abbildung 19: $^1\text{H}^{15}\text{N}$ -HSQC-Spektren der WW23 Varianten F11 und F7.....	64
Abbildung 20: $^1\text{H}^{15}\text{N}$ HSQC-Spektren von WW1-F5 und WW3-F12 .....	65
Abbildung 21: Massenspektrum von uniform $^{15}\text{N}^{13}\text{C}$ -markiertem WW23-F7 .....	66
Abbildung 22: Überlagerung des realen und des rekonstruierten CD-Spektrums von WW23 .....	67
Abbildung 23: Bekannte Sekundärstrukturelemente in WW23-F7 .....	67
Abbildung 24: Zugeordnetes HSQC-Spektrum von WW23-F7 .....	69
Abbildung 25: Zugeordnete $^1\text{H}^{15}\text{N}$ -Korellationen von WW23-F7.....	69
Abbildung 26: $^{15}\text{N}$ -Spinrelaxationsdaten und Orderparameter von WW23-F7.....	70
Abbildung 27: $^{15}\text{N}$ Relaxation von WW23 in Interaktion mit CA_P1 .....	72
Abbildung 28: Vergleich der Dynamikeigenschaften von WW23 mit und ohne Peptidliganden .....	73
Abbildung 29: SPOT-Analyse von SF1-prolinreichen Peptiden und CA150-WWs .....	75
Abbildung 30: ITC-Interaktionsexperimente mit WW2 und WW3.....	77
Abbildung 31: $1\text{D}^{-1}\text{H}$ -Spektren der synthetischen und der rekombinanten WW3.....	78
Abbildung 32: $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von WW23 in An- und Abwesenheit verschiedener SF1 Peptide.....	80
Abbildung 33: Analyse der Änderung der $^1\text{H}^{15}\text{N}$ chemischen Verschiebungen von WW23 nach Peptidzugabe .....	81
Abbildung 34: Bindungstaschen WW2 und WW3 .....	83
Abbildung 35: Vergleich der Expressionsstärke von WW23 WT und WW23-Mutanten.....	84



---

Abbildung 36: Überlagerung der HSQC-Spektren von WW23 WT und WW23_W32A.....	84
Abbildung 37: Überlagerung der HSQC-Spektren von WW23 WT und WW23_W131A.....	85
Abbildung 38: Thermogramm der „klassischen“ und der reversen WW23+SF1_P4 Titration.....	93
Abbildung 39: ITC-Plots der Interaktionen von WW23(-M) mit SF1_P4.....	94
Abbildung 40: SPOT-Analyse mit prolinreichen Peptiden der CA150 <sub>PRS</sub> und CA150 <sub>WW</sub> -Domänen.....	98
Abbildung 41: CA_P1-induzierte <sup>1</sup> H <sup>15</sup> N chemische Verschiebungsdifferenzen in WW23.....	99
Abbildung 42: UV-Vis Spektrum von WW23.....	102
Abbildung 43: AUZ-Analyse von WW23 und WW23/CA_P1.....	103
Abbildung 44: Wichtige TCERG1-Interaktionen.....	106
Abbildung 45: Ergebnis der PsiPred-Vorhersage für WW23.....	108
Abbildung 46: CA150 <sub>WW2</sub> im Komplex mit PTPPLPP.....	114
Abbildung 47: WW2- und WW3-Sequenzabschnitt mit „xP“- und „xP2“-Bindungstasche.....	115
Abbildung 48: "xP" und "xP2"-Tasche von CA150 <sub>WW3</sub> .....	115
Abbildung 49: CA150 <sub>WW</sub> -Bindepotionen der SF1 <sub>PRS</sub> .....	122
Abbildung 50: CA150 <sub>WW</sub> -Bindepotionen der N-terminalen CA150 <sub>PRS</sub> .....	123
Abbildung 51: Modelle multivalenter CA150 <sub>WW</sub> /CA150 <sub>PRS</sub> -Wechselwirkungen.....	124
Abbildung 52: <sup>1</sup> H <sup>15</sup> N-HSQC-Spektren von WW23_W32A in An- und Abwesenheit verschiedener SF1 Peptide.....	146
Abbildung 53: <sup>1</sup> H <sup>15</sup> N-HSQC-Spektren von WW23_W131A in An- und Abwesenheit verschiedener SF1 Peptide.....	147
Abbildung 54: <sup>1</sup> H <sup>15</sup> N-HSQC-Spektren von WW23 in An- und Abwesenheit verschiedener monovalenter PP-Peptide.....	148
Abbildung 55: Analyse der Änderung der <sup>1</sup> H <sup>15</sup> N chemischen Verschiebungen von WW23, WW23_W32A und WW23_W131A nach SF1_P1-Zugabe.....	149
Abbildung 56: Analyse der Änderung der <sup>1</sup> H <sup>15</sup> N chemischen Verschiebungen von WW23, WW23_W32A und WW23_W131A nach SF1_P3.....	150
Abbildung 57: Analyse der Änderung der <sup>1</sup> H <sup>15</sup> N chemischen Verschiebungen von WW23, WW23_W32A und WW23_W131A nach SF1_P4-Zugabe.....	151
Abbildung 58: <sup>1</sup> H <sup>15</sup> N-HSQC-Spektren von WW23, WW23_W32A und WW23_W131A in An- und Abwesenheit verschiedener SF1 Peptide.....	152
Abbildung 59: Analyse der Änderung der <sup>1</sup> H <sup>15</sup> N chemischen Verschiebungen von WW23, WW23_W32A und WW23_W131A nach CA_P1-Zugabe.....	153
Abbildung 60: Ergebnis der Talos+ Strukturdynamik- und Sekundärstrukturvorhersage.....	153

---

## TABELLEN

Tabelle 1.2.I: Ausgewählte Proteindomänen in ihrem funktionellen Kontext .....	14
Tabelle 1.4.I: Anzahl der WW-Domänen und WW-Domänen-haltiger Proteine innerhalb repräsentativer Spezies .....	23
Tabelle 1.4.II: Publierte Strukturen von Tandem-WW-Domänen.....	24
Tabelle 1.4.III: Gegenüberstellung verschiedener WW-Klassifizierungen .....	27
Tabelle 1.5.I: Interaktionspartner von CA150.....	32
Tabelle 1.5.II: Ergebnisse der Mutagenesestudie von <i>Goldstrohm et al.</i> 2001 [112].....	33
Tabelle 3.1.I: Geräte .....	36
Tabelle 3.1.II: Spezielle Chemikalien.....	37
Tabelle 3.1.III: Molekularbiologische und biochemische Kits.....	37
Tabelle 3.1.IV: Nährmedien und Lösungen.....	37
Tabelle 3.2.I: Software .....	38
Tabelle 3.3.I: Verwendete Oligonukleotide .....	39
Tabelle 3.3.II: Verwendete Expressionsvektoren .....	40
Tabelle 3.3.III: In dieser Arbeit generierte Plasmide.....	40
Tabelle 3.4.I: Verwendete Bakterienstämme .....	43
Tabelle 3.6.I: Molekulargewichte (MW) und molare Extinktionskoeffizienten ( $\epsilon_{280}$ ) der rekombinant erzeugten WW-Domänen-Fusionsproteine.....	48
Tabelle 3.6.II: Molekulargewichte und molare Extinktionskoeffizienten ( $\epsilon_{280}$ ) rekombinant erzeugter WW-Domänen .....	48
Tabelle 3.6.III: Sequenzen rekombinant und synthetisch hergestellter CA150 <sub>WW</sub> -Varianten .....	48
Tabelle 3.6.IV: Prolinreiche Peptidliganden .....	49
Tabelle 3.6.V: Sequenzen der SF1-SPOT-Peptide .....	50
Tabelle 3.6.VI: Sequenzen der CA150 <sub>PRS</sub> -SPOT-Peptide .....	50
Tabelle 3.7.I: Übersicht der zur Zuordnung der Proteinerückgrat-Resonanzen verwendeten NMR-Spektren von CA150 WW23. ....	52
Tabelle 3.10.I: In SV-Experimenten eingesetzte WW23-Konzentrationen .....	55
Tabelle 3.10.II: In SV-Experimenten eingesetzte Konzentrationen von WW23 und CA_P1 .....	55
Tabelle 3.12.I: SILAC-Isotopen .....	57
Tabelle 4.1.I: Ergebnisse der PCR's.....	59
Tabelle 4.2.I: Expressionslevel der WW-Varianten in Vollmedium.....	60
Tabelle 4.4.I: Mit "CDSSTR" berechnete Sekundärstruktur-Anteile von WW23.....	67
Tabelle 4.5.I: $S_2$ und $T_c$ von WW23-Strukturelementen.....	71
Tabelle 4.5.II: $S_2$ und $T_c$ von WW23 in Interaktion mit CA_P1 .....	72
Tabelle 4.6.I: Ausgewählte SF1-PP-Peptide .....	76
Tabelle 4.6.II: Analyse der SPOT-Ergebnisse bezüglich bekannter WW-Bindemotive.....	76
Tabelle 4.6.III: Stärke der peptidinduzierten Signalverschiebungen .....	82
Tabelle 4.6.IV: Stärke der SF1_P1-induzierten Signalverschiebungen.....	87
Tabelle 4.6.V: Stärke der SF1_P3-induzierten Signalverschiebungen.....	88

---

Tabelle 4.6.VI: Stärke der SF1_P4-induzierten Signalverschiebungen .....	89
Tabelle 4.6.VII: ITC-Ergebnisse der WW-SF1_P1 Interaktion .....	91
Tabelle 4.6.VIII: ITC-Ergebnisse der WW-SF1_P3 Interaktion .....	92
Tabelle 4.6.IX: ITC-Ergebnisse der WW-SF1_P4 Interaktion .....	94
Tabelle 4.7.I: Anreicherungsverhältnis der WW-Interaktionspartner SILAC-Pulldowns .....	95
Tabelle 4.7.II: Mittels SILAC-Pulldown identifizierte bekannte und "neue" CA150-WW- Interaktionspartner .....	96
Tabelle 4.7.III: SILAC-Bindepräferenzen der CA150-Interaktionspartner für WW2, WW3 und WW23.....	97
Tabelle 4.8.I: Für NMR- und ITC-Studien ausgewähltes CA150-Peptid.....	98
Tabelle 4.8.II: Stärke der CA_P1-induzierten Signalverschiebungen.....	100
Tabelle 4.8.III: Ergebnisse WW/CA_P1-Interaktionen.....	101
Tabelle 4.8.IV: Mit "Sednterp" berechnete Parameter .....	103
Tabelle 4.8.V: Daten der AUZ-Experimente mit WW23.....	103
Tabelle 4.9.I: Kurze PP-Peptide .....	104
Tabelle 4.9.II: Stärke der SF1_KP1 und CA_KP1-induzierten Signalverschiebungen.....	105
Tabelle 5.1.I: Neue CA150 <sub>WW</sub> -Interaktionspartner .....	107
Tabelle 5.5.I: Übersicht zur Bindung der PP-Peptide an die „xP“- bzw. „xP2“-Tasche .....	116
Tabelle 5.5.II: Prolinreiche Peptide, PP-Motive, Bindekonstanten und Interaktionsstellen .....	117
Tabelle 5.5.III: Bindungsmotive der WW-Bindungsstudien.....	120
Tabelle 5.5.IV: Bindemotive der CA150-WW-Domänen.....	121
Tabelle 6.2.I: Zugeordnete <sup>1</sup> H-, <sup>15</sup> N- und <sup>13</sup> C chemische Verschiebungen des WW23-Tandems .	143

---

## **ABKÜRZUNGEN**

6xHis	6x-Polyhistidin-Anhang
®	eingetragene Marke ( <i>registered trademark</i> )
1D	eindimensional
2D	zweidimensional
3D	dreidimensional
AccNo	<i>Accession number</i>
APS	Ammoniumpersulfat
AS	Aminosäure
AUZ	Analytische Ultrazentrifugation
bp	Basenpaar
Cam <sup>R</sup>	Chloramphenicol-Resistenz
CA-P1	CA150 Peptid 1 HNAVMR <u>GP</u> PLMR <u>PP</u> PFGM <u>MR</u> GP <u>PP</u> PR <u>PP</u> FG
CD	Zirkulardichroismus
C-terminal	carboxyterminal
C-Terminus	Carboxyterminus
D <sub>2</sub> O	Deuteriumoxid
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
Ek	Enterokinase
c <sub>End</sub>	Endkonzentration
fw	vorwärts ( <i>forward</i> )
GST	Glutathion-S-Transferase
HeLa	Zervixkarzinom-Zelllinie (nach Spenderin Henrietta Lacks benannt)
HSQC	<i>heteronuclear single quantum coherence</i>
IMAC	immobilisierte Metall-Affinitätschromatographie ( <i>immobilized metal affinity chromatography</i> )
IPTG	Isopropyl-β-D-thiogalactopyranosid
ITC	isotherme Titrationskalorimetrie
Kan <sup>R</sup>	Kanamycin-Resistenz
kb	Kilobasen
K <sub>d</sub>	Dissoziationskonstante
kDa	Kilodalton
LB	<i>Lysogenybroth</i> (von Luria und Bertani beschriebenes Vollmedium)
LIC	ligationsunabhängige Klonierung ( <i>ligation-independent cloning</i> )
MS	Massenspektrometrie
MW	Molekulargewicht ( <i>molecular weight</i> )
MWCo	Molekulargewichts <i>Cut-Off</i> - Porengröße
n.H.	natürliche Häufigkeit / Isotopenverteilung
NMR	<i>nuclear magnetic resonance</i> (Kernspinresonanz)
N-Terminus	Aminoterminus
OD <sub>600nm</sub>	optische Dichte (bei 600 nm)
SF1_P4	SF1 Peptid 4 H <u>PP</u> PP <u>SH</u> Y <u>G</u> PL <u>PP</u> WHQK
PBS	phosphatgepufferte Salzlösung ( <i>phosphate buffered saline</i> )
PCR	Polymerase-Ketten-Reaktion

---

PDB	<i>Protein Data Bank</i> ( <a href="http://www.pdb.org">http://www.pdb.org</a> )
pET	Plasmid zur Expression, vermittelt durch die T7 RNA-Polymerase
ppm	<i>parts per million</i>
rpm	Umdrehungen pro Minute ( <i>rounds per minute</i> )
SDS-PAGE	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese
SF1_P1	SF01 Peptid 1 <u>QPPPPPMNQGP</u> PHPPGHHGPP
SF1_P3	SF01 Peptid 3 <u>MPPPPPPPSGQPPPPPSG</u> PL
SILAC	<i>Stable isotope labeling with amino acids in cell culture</i>
Sm <sup>R</sup>	Streptomycin-Resistenz
Tag	Protein-Tag
TEMED	N,N,N',N'-Tetramethylethylamin
TEV	Tabakäzvirus ( <i>tobacco etch virus</i> )
™	nichteingetragene Marke ( <i>unregisteredtrademark</i> )
Tris	Tris(-hydroxymethyl-)aminomethan
UV/VIS	ultraviolett / sichtbar ( <i>ultraviolet / visible</i> )
WW	WW-Domäne
PP-	Poly-Prolin
Seq	Sequenz

---

---

# 1 EINLEITUNG

## 1.1 ***DIE BEDEUTUNG VON PROTEINEN UND PROTEIN-INTERAKTIONEN***

Proteine sind die Stoffklasse lebender Materie mit der bei weitem höchsten Diversität. Sie bilden die Elemente von Stützstrukturen, dienen der Speicherung und dem Stofftransport, der Signalübermittlung, der Bewegung und der Abwehr von Fremdstoffen. Zusätzlich regulieren Proteine in Form von Enzymen den Stoffwechsel, indem sie chemische Reaktionen in der Zelle selektiv beschleunigen können. Eine Grundvoraussetzung für diese Funktionsvielfalt ist ihre Eigenschaft, andere Moleküle spezifisch, reversibel und modulierbar zu binden. Jedes Protein besitzt dazu ein unverwechselbares Oberflächenprofil, geprägt durch die Seitenketten seiner Aminosäurereste.

Ein sehr prominentes Beispiel sind die Antikörper der menschlichen Immunabwehr, deren Konturen so vielfältig sein können, dass sie praktisch jedes beliebige Molekül spezifisch binden und damit „erkennen“ können. Die Bindung entsprechender Liganden ([lateinisch] *ligare* = binden) an Teilbereiche der Proteinoberfläche ist oftmals auch mit einer Veränderung der physikochemischen Eigenschaften des Proteins verbunden, wodurch die Interaktion zu weiteren Zielmolekülen spezifisch verändert werden kann. Liganden sind beispielsweise Kohlenhydrate, Nukleotide, Lipide, Nucleinsäuren, Xenobiotika, kleine organische Moleküle, Ionen, Peptide und andere Proteine.

Proteomweite Interaktionsstudien zeigten, dass ein zellulärer Phänotyp nicht von einzelnen Proteinen, sondern vielmehr durch ein sehr komplexes Zusammenspiel von hunderten, manchmal tausenden unterschiedlichen Proteinen bestimmt wird. So umfasst beispielsweise das humane Interaktom ca. 650.000 Protein-Protein-Interaktionen [1]. Das Verständnis der diesen Wechselwirkungen zugrundeliegenden molekularen Mechanismen ist eine wichtige Herausforderung der aktuellen Forschung, da eine ganze Reihe von Krankheiten, wie z.B. verschiedene Krebserkrankungen, Alzheimer, Parkinson, Huntington, Rheuma und Diabetes unter anderem auf fehlregulierte Protein-Interaktionen zurückgeführt werden konnten [2].

Neuropathische Erkrankungen wie Alzheimer, Parkinson und Huntington sind durch den Verlust von Neuronen und das Auftreten synaptischer Anomalien charakterisiert, die zu schweren Beeinträchtigungen der kognitiven Fähigkeiten führen können. Durch histologische Untersuchungen von Gehirnschnitten betroffener Patienten, konnten pathologische Anreicherungen von Proteinen in Form unlöslicher Fibrillen und Plaques identifiziert werden. Wie sich herausstellte, beruht die Bildung solcher Fibrillen und Plaques auf einer Wechselwirkung zwischen fehlgefalteten Proteinen.

Pharmazeutika, welche fehlregulierte pathogene Interaktionen beeinflussen können, eröffnen die Möglichkeit, unerwünschte Wechselwirkungen zu unterdrücken und die ursprüngliche Funktion des Proteins wiederherzustellen.

---

---

Bevor allerdings derartige Wirkstoffmoleküle kreiert werden können, ist es zwingend erforderlich, die zu Grunde liegenden Wechselwirkungen zwischen den Interaktionspartnern besser zu verstehen.

Ein Charakteristikum molekularer Wechselwirkungen ist die Dissoziationskonstante ( $K_d$ ), welche ein Maß für die Affinität zwischen den beteiligten Molekülen ist. Wechselwirkungen zwischen Proteinen zellulärer Signalwege sind meist relativ schwach ( $K_d$   $10^{-3}$  -  $10^{-9}$  M), da diese ihre Liganden nicht dauerhaft, sondern nur vorübergehend binden sollen.

Bei einigen Proteinen führt die Ligandenbindung zu einer Konformationsänderung, welche die Affinität zu weiteren Liganden beeinflusst. Als Beispiel sei hier Calmodulin genannt, welches bis zu vier  $Ca^{2+}$ -Ionen bindet und dadurch eine Konformationsumwandlung erfährt, die an das „Aufklappen eines Regenschirms“ erinnert. Die dabei freigelegten neuen Interaktionsstellen können nun weitere Signalproteine wie die  $Ca^{2+}$ /Calmodulin-abhängige Proteinkinase (CaM-Kinase) binden und aktivieren [3,4].

Auch Interaktionen eines Liganden zu verschiedenen Ligandenbindungsstellen (Multivalenz, vgl. Abschn. 1.3) innerhalb eines Proteins sind möglich. Dieses Phänomen tritt oft im Zusammenhang mit allosterischen Effekten auf. So kann die Bindung eines Liganden an eine Bindungsstelle, die Bindung eines weiteren Liganden an eine weitere Bindungsstelle erleichtern (positive Allosterie) bzw. erschweren (negative Allosterie).

Der allosterische Effekt ist von großer Bedeutung für die Regulation von Proteinen. So kann beispielsweise Hämoglobin mittels allosterischer Effektoren wie 2,3-Bisphosphoglycerat die Sauerstoffaufnahme und -abgabe den physiologischen Bedürfnissen präzise anpassen [5,6].

Proteine können zudem kurzfristig oder dauerhaft chemisch verändert und damit etwa „an-“ oder „abgeschaltet“ werden. Die häufigste Form der kovalenten Modifikation ist die Phosphorylierung durch Kinasen bzw. Dephosphorylierung durch Phosphatasen. Die Proteinphosphorylierung besitzt große Bedeutung für zahlreiche zelluläre Signalwege.

Viele Regulations- und Signalproteine verändern ihre Eigenschaften je nach Phosphorylierungsgrad. So kann eine Phosphorylierung einerseits allosterische Übergänge bewirken, zum anderen kann aber auch durch das Einbringen der negativen Ladungen des Phosphats die Oberfläche des Proteins so verändert werden, dass lokal eine neue Bindungsstelle geschaffen oder blockiert wird. Eine regulatorisch wirksame Phosphorylierung spielt beispielsweise bei Transkriptionsfaktoren eine wichtige Rolle. Infolge dieser Modifikation können diese beispielsweise dimerisieren (z.B. CREB-1 [7]) und daraufhin in den Zellkern einwandern.

Voraussetzung für die Funktionsvielfältigkeit vieler Proteine ist deren modularer Aufbau aus distinkten Funktionseinheiten, den sogenannten Proteindomänen.

---

## 1.2 PROTEINDOMÄNEN I - MODULARE FUNKTIONSEINHEITEN

Die meisten humanen Proteine setzen sich aus mehreren modularen Funktionseinheiten (sogenannten Proteindomänen) zusammen [8], wobei durch die charakteristische Kombination verschiedener Module (Domänenarchitektur) maßgeblich die Funktion des Proteins bestimmt wird [9,10]. Proteindomänen sind in charakteristischerweise dreidimensional gefaltete und strukturell kohärente funktionelle Einheiten [11]. Als die kleinste (ca. 35-40 AS), kompakteste und zugleich stabilste Faltung einer Aminosäuresequenz kann eine Proteindomäne (z.B. WW-Domäne) einzeln oder gleich in mehrfacher Kopie im Protein vorliegen.

Auf genetischer Ebene werden humane Proteindomänen durch so genannte „Shuffled“ Exons codiert [12]. Diese können auf Grund ihrer Transponierbarkeit ähnlich eines „Baukastenprinzips“ zu den DNA-Matrizen verschiedenster Proteine kombiniert werden. Dementsprechend ist die Anzahl der codierenden Gene deutlich geringer, als die Zahl der exprimierbaren Proteinspezies.

In Tabelle 1.2.1 sind exemplarisch verschiedenste Proteindomänen und ihr entsprechender funktioneller Kontext aufgezeigt.

**Tabelle 1.2.1: Ausgewählte Proteindomänen in ihrem funktionellen Kontext**

Funktioneller Kontext	Domäne
Apoptose	BH1-4, BIR, CARD, DEATH, DED, DH, TRAF
Calcium Bindung	EF-Hand
Chromatin-Regulation	BROMO, BTB/POZ, Chromo, CSD, SWIRM, TUDOR
Zytoskelett-Modulation	CH, FH2, GEL
Methyl-Lysin-Bindung	LRR, MBT
Phospho-Serin/Threonin Bindung	14-3-3, BRCT, FF, FHA, MH2, POLO-BOX, <b>WW</b>
Phospho/Tyrosin-Bindung	PTB, SH2
Phospholipid-Bindung	BAR, BEACH, C1, C2, CALM, ENTH, FERM, FYVE, GLUE, GRAM, PH, PX
Protein-Abbau	F-Box, HECT, RING, SOCS, UBA, <b>WW</b>
Bindung prolinreicher Sequenzen	EVH1, GYF, SH3, <b>WW</b>
Ubiquitin Bindung	CUE, GAT, MIU, NZF, UEV, UIM
Vesikel Transport	EH, SNARE
Genregulation	<b>WW</b> , FF, ZnF_C4, THAP, NAM

Besonders bei der Ausbildung von Multiproteinkomplexen ist ein modularer Aufbau von Vorteil. Einige Proteindomänen fungieren dort als nicht katalytische Adaptoren, die durch die Ausbildung verschiedenster, oftmals multivalenter Protein-Protein-Wechselwirkungen die einzelnen Komplex-Untereinheiten zusammenhalten. Das Phänomen der Multivalenz (Prinzip, Thermodynamik) und deren Bedeutung in biologischen Systemen soll im Folgenden näher betrachtet werden.



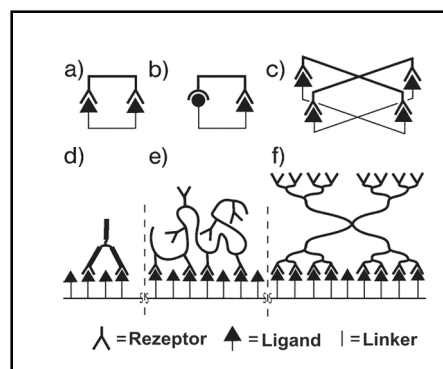
### 1.3 DIE MULTIVALENZ

Multivalenz ([lateinisch] Vielwertigkeit, Möglichkeit vieler Lösungen) beschreibt die Wechselwirkung zwischen Rezeptoren und Liganden, die jeweils multiple Bindungsstellen besitzen. Die Bezeichnungen „Ligand“ und „Rezeptor“ werden hier verwendet, um die individuellen Komponenten der multivalenten Spezies zu identifizieren. Der Rezeptor ist in diesem Fall die Komponente, welche den Liganden unter Verwendung einer oberflächlichen Vertiefung oder Tasche bindet. Einzelne Liganden können durch ein molekulares Gerüst, den so genannten „Linker“, miteinander verbunden sein, wodurch Kombinationen aus gleich- oder verschiedenartigen multivalenten Liganden entstehen (siehe Abbildung 1). In multivalenten Rezeptoren sind einzelne Bindungsstellen oftmals in distinkten Domänen lokalisiert, die ihrerseits ebenfalls durch eine Linkersequenz verknüpft sein können.

#### 1.1.1 DAS PRINZIP DER MULTIVALENZ

Eine multivalente Interaktion kann laut *Krishnamurthy et al.* [13] als Vorgang definiert werden, bei dem multiple gleichartige Bindungsereignisse zwischen zwei Komponenten (Ligand:Rezeptor) simultan stattfinden (siehe Abbildung 1). Man kann die Multivalenz auf der Grundlage der Anzahl ( $i$ ) der Interaktionen zwischen Ligand und Rezeptor in drei Kategorien unterteilen:

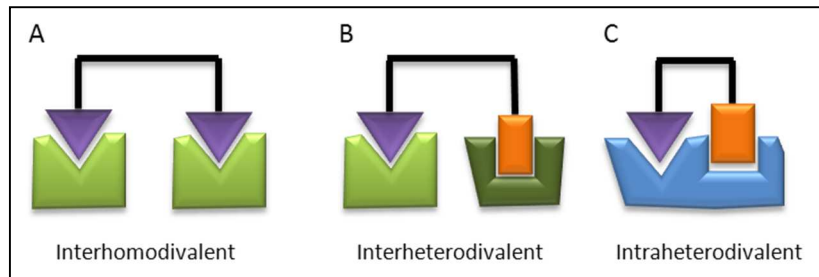
- Die Bivalenz (Abbildung 1 a, b, d):  
Hier finden zwei Interaktionen ( $i$ ) an zwei gleich- oder verschiedenartigen Bindungsstellen zweier Spezies statt ( $i=2$ ).
- Die Oligovalenz (Abbildung 1 c, e, f):  
Hier findet eine diskrete Anzahl Interaktionen ( $i$ ) an mehreren gleich- oder verschiedenartigen Bindungsstellen zweier Spezies statt ( $2 < i \leq 10$ ).
- Die Polyvalenz  
Hier findet eine große Zahl Interaktionen ( $i$ ) an mehreren gleich- oder verschiedenartigen Bindungsstellen zweier Spezies statt ( $i > 10$ ).



**Abbildung 1: Beispiele multivalenter Systeme**

(a) Ein bivalenter Ligand bindet einen bivalenten Rezeptor, (b) Ein heterobivalenten Ligand bindet an einen Rezeptor mit zwei verschiedenen Bindestellen, (c) Ein tetravalenter Ligand bindet an einen tetravalenten (oligovalenten) Rezeptor (d) Ein bivalenter Antikörper bindet an eine Oberfläche, (e) Ein Polymer bindet an eine Oberfläche, (f) Ein Dendrimer bindet an eine Oberfläche (entnommen aus [13])

Es können zwischen beiden Komponenten also zwei oder mehr gleichartige (Homo-Multivalenz) oder verschiedenartige (Hetero-Multivalenz) molekulare Bindungsereignisse auftreten (vgl. Abbildung 2).



**Abbildung 2: Typen multivalenter Interaktionen**

Gezeigt sind verschiedene Typen multivalenter Interaktionen zwischen Ligand- und Rezeptormolekülen. (A) interhomovalente, (B) interheterovalente, (C) intraheterovalente Interaktionen. (nach [14])

Die Mehrzahl der bis heute synthetisierten multivalenten Moleküle binden interhomovalent (Abbildung 2A). Das bedeutet, dass zwei verbundene identische Liganden die gleichen Bindestellen in zwei identischen Rezeptoren bedienen [14]. *Choi* zeigte, dass derartige Moleküle multivalent auf der Oberfläche von Viren, Bakterien oder Zellen gebunden werden [15]. Auf diesem Prinzip beruht beispielsweise die Wechselwirkung zwischen Influenza Virus und Zelloberfläche (vgl. Abschnitt 1.3.1).

Die in biologischen Systemen am häufigsten auftretende Form der Multivalenz ist vermutlich die Hetero-Multivalenz. So wirkt z.B. Rapamycin, ein immunsuppressiver Wirkstoff aus *Streptomyces hygroscopicus*, interheterodivalent (vgl. Abbildung 2B). Dieser Wirkstoff kann gleichzeitig zwei cytoplasmatische Proteine, FKBP 12 (*FK506 binding protein*) und FRB (*FKBP-rapamycin binding domain*) binden, und so den ternären FKBP-Rapamycin-FRB Komplex ausbilden [16]. Es konnte gezeigt werden, dass synthetische Rapamycinderivate in der Lage sind, verschiedene intrazelluläre Signalwege zu modifizieren. Zu diesen pharmakologisch sehr interessanten Prozessen zählen die Relokalisation von Proteinen [17,18], die bedingte Induktion der Apoptose [19], der Proteinabbau [20] und das konditionelle Spleißen [21].

Eine weitere Form der multivalenten Interaktion ist die Intraheterodivalenz. Dieser Mechanismus beschreibt die simultane Bindung eines multivalenten Liganden an mehrere verschiedene Bindestellen eines Rezeptormoleküls (vgl. Abbildung 2C).

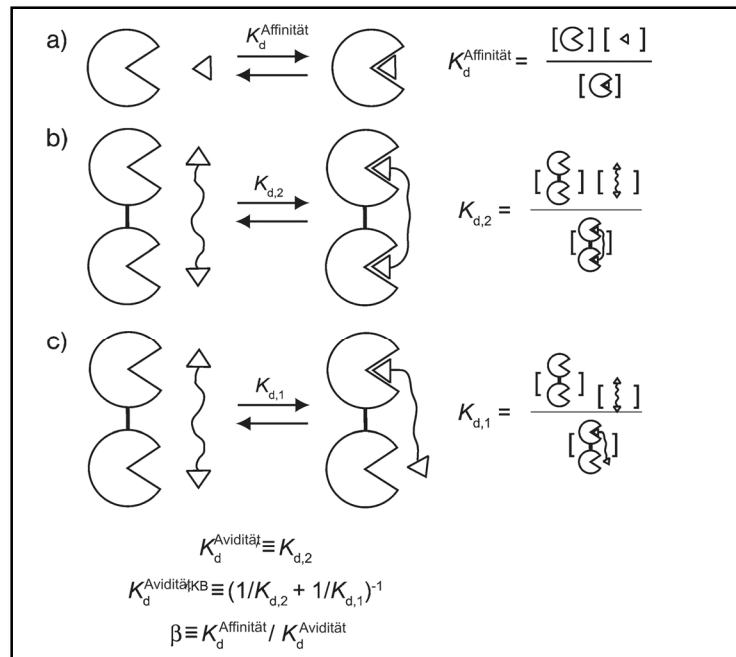
### 1.1.2 THERMODYNAMISCHE ANSÄTZE

Eine konkrete thermodynamische Beschreibung der vielfältigen multivalenten Interaktionen stellt auf Grund der enormen Komplexität eine große Herausforderung dar. Da Homo-Multivalenz einfacher zu definieren ist als Hetero-Multivalenz, die grundlegenden Prinzipien aber die gleichen sind, wird bei der Beschreibung der Multivalenz häufig nur die Homo-Multivalenz betrachtet. Während eine monovalente Bindung hauptsächlich durch die Bindungsaffinität ( $K_d$ ) des Liganden bestimmt wird, muss

bei multivalenten Bindungen auch der Einfluss des Linkers (Länge, Flexibilität / Rigidität) berücksichtigt werden [13,22–24]. Nach *Kitov und Bundle* kann die Affinität multivalenter Liganden nicht einfach als die Summe der Bindungskräfte zwischen den distinkten Interaktionsstellen beschrieben werden [23]. Die Stärke einer multivalenten Interaktion wird vielmehr durch die Avidität ( $K_d^{\text{Avidität}}$ ) ausgedrückt. Diese berechnet sich aus dem Verhältnis der Konzentrationen der verschiedenen Rezeptor:Ligand-Komplexe mit  $n$  Rezeptor:Ligand-Interaktionen ( $i=n$ ) zum vollständig dissoziiert vorliegenden Rezeptor ( $i=0$ ) (Abbildung 3 b) [13]. *Kitov und Bundle* [23] definieren die Avidität ( $K_d^{\text{Avidität, KB}}$ ) wie folgt:

$$K_d^{\text{Avidität, KB}} \equiv (1/K_{d,n} + 1/K_{d,n-1} + \dots + 1/K_{d,1})^{-1}$$

$K_{d,n}$ ,  $K_{d,n-1}$  und  $K_{d,1}$ : Dissoziationskonstanten von Rezeptor:Ligand-Komplexen mit  $n$  Rezeptor:Ligand Interaktionen ( $i=n$ ),  $n-1$  Interaktionen und einer Interaktion ( $i=1$ ), jeweils im Verhältnis zum vollständig dissoziierten Rezeptor:Ligand-Komplex ( $i=0$ ) (Abbildung 3 b, c).

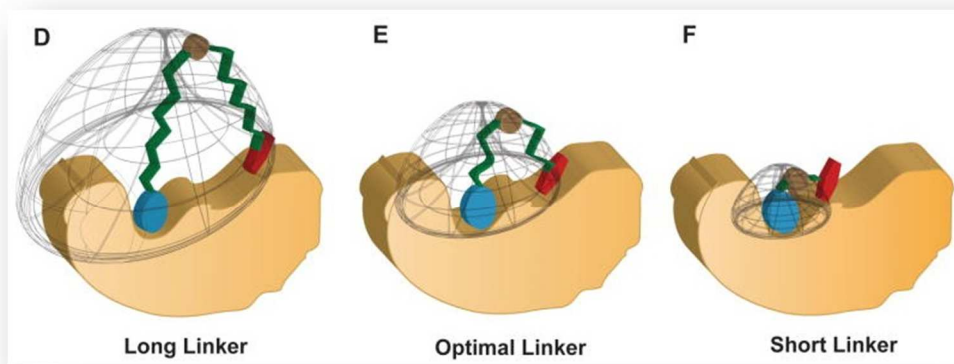


**Abbildung 3: Thermodynamische Gleichgewichte einer mono- und bivalenten Interaktion**

Thermodynamische Gleichgewichte zur Definition von Affinität, Avidität und Verstärkung (a) Ein Monovalenter Ligand bindet einen monovalenten Rezeptor mit einer Dissoziationskonstante ( $K_d^{\text{Affinität}}$ ) (b) Der oligovalente Ligand (hier bivalent) bindet einen Rezeptor mit der gleichen Valenz mit einer Dissoziationskonstante  $K_{d,2}$  entsprechend dem Gleichgewicht zwischen vollständig gebundenen Rezeptor und freien Rezeptor- und Ligand-Molekülen (c) Der bivalente Rezeptor kann auch den bivalenten Liganden mit nur einer Rezeptor-Ligand Interaktion binden, der Komplex hat eine Dissoziationskonstante  $K_{d,1}$ . *Krishnamurthy et al.* [13] definierten die Avidität (die Gesamtheit der Affinitäten,  $K_d^{\text{Avidität}}$ ) als  $K_{d,2}$ . Um alle Rezeptor::Ligand Komplexe einzubeziehen, definierten *Kitov und Bundle* [23] die Avidität ( $K_d^{\text{Avidität, KB}}$ ) dagegen durch die Beziehung:  $(1/K_{d,2} + 1/K_{d,1})^{-1}$ . Die Verstärkung ( $\beta$ ) ist das Verhältnis zwischen Affinität und Avidität (Entnommen aus [13]).

Eine umfassende thermodynamische Betrachtung multivalenter Wechselwirkungen ist im Rahmen dieser Arbeit nicht möglich. An dieser Stelle sei auf die Arbeiten von *Kitov und Bundle* [23] und *Krishnamurthy et al.* [13] verwiesen.

Ein intuitiveres Verständnis von Multivalenzeffekten bietet das Konzept der so genannten „effektiven“ Konzentration ( $C_{\text{eff}}$ ) [25]. In diesem Parameter sind alle bindungsrelevanten thermodynamischen Einflussgrößen zu einer fiktiven Ligandenkonzentration zusammengefasst. Wenn beispielsweise in einem bivalenten System bereits ein Ligand an einen Rezeptor bzw. eine Bindestelle gebunden hat, erhöht sich durch die resultierende räumliche Nähe oder günstige Orientierung die Wahrscheinlichkeit, dass auch der zweite Ligand an den zweiten Rezeptor bindet. Der zweite Rezeptor scheint also effektiv von einer viel höheren Ligandenkonzentration umgeben zu sein, als tatsächlich vorliegt. Experimente mit bivalenten Liganden zeigten, dass es möglich ist, die effektive Konzentration durch eine Veränderung der Linkerlänge zwischen den Liganden zu kontrollieren [26] (siehe Abbildung 4).



**Abbildung 4: Der Einfluss des Linkers auf die Effektive Konzentration**

Gezeigt sind schematische Proteinmoleküle mit gebundenen bivalenten Liganden. Der Linker zwischen den Liganden ist länger (D), gleichlang (E) bzw. kürzer (F), als der Abstand zwischen den zwei Bindestellen im Protein. Abhängig von seiner Länge, kann der Linker die Wahrscheinlichkeit einer bivalenten Interaktion erhöhen oder verringern. Das ist äquivalent zu einer Erhöhung oder Verringerung der effektiven Konzentration ( $C_{\text{eff}}$ ). (Abbildung wurde entnommen aus [14].)

Zusammenfassend lässt sich festhalten, dass die Avidität eines multivalenten Liganden nicht ausschließlich durch die Affinitäten der monovalenten Liganden, sondern auch durch die Länge und konformationelle Dynamik des Linkers bestimmt wird. Des Weiteren können auch enthalpische oder entropische Kompensationseffekte den Gewinn der Multivalenz verringern [13]. Interessant ist, dass in multivalenten Systemen die Dissoziation eines multivalenten Komplexes schrittweise geschieht, wobei die Dissoziationsraten durch Zugabe eines kompetitierenden monovalenten Liganden erhöht werden können [27,28]. Aufgrund des komplexen Zusammenspiels sich gegenseitig beeinflussender Wechselwirkungen, ist es bis heute noch nicht gelungen, die molekularen Mechanismen multivalenter Interaktionen zufriedenstellend aufzuklären.

### 1.3.1 MULTIVALENZ IN BIOLOGISCHEN SYSTEMEN

Multivalenzeffekte können vielfach in biologischen Systemen beobachtet werden. Zu den bekanntesten multivalenten Prozessen zählen u. A. Interaktionen zwischen Proteinen, zwischen Proteinen und ihren Liganden, Antigen-Antikörper-, Virus-Zell-Interaktionen, sowie die Wechselwirkungen zwischen bakteriellen Toxinen und ihren Ziel-Zellen [15,24,29–31]. Ebenfalls zu den multivalenten Prozessen zählt das Prinzip der Selbstorganisation, also der spontanen Assoziation von Molekülen zu wohldefinierten Strukturen (durch nichtkovalente Wechselwirkungen zusammengehalten) [32–35]. In der Natur wird dieses Prinzip elegant am Beispiel des Tabakmosaikvirus genutzt, welches sich spontan aus 2000 identischen Aminosäure-Untereinheiten zu einer helikalen Struktur um einen einzelnen RNA-Strang zusammenbaut [36]. Auf DNA-Ebene sorgen polyvalente Kontakte mit polymeren Histonen dafür, dass die DNA spezifisch in der Coiled-Coil-Konformation eingelagert wird [22].

Multivalente Interaktionen führen häufig zu starken Affinitätserhöhungen zwischen den Interaktionspartnern. So werden beispielsweise Interaktionen zwischen Liganden mit geringen Oberflächen durch das mehrfache repräsentieren von Liganden- und Akzeptoreinheiten verstärkt (Bsp. Verkleben von Zellen). Ein äußerst prominentes multivalentes Molekül ist der trimere Hämagglutinkomplex des Influenza Virus. Multivalente Sialinsäuremoleküle auf der Zelloberfläche binden den Hämagglutinin-Rezeptor auf der Virusoberfläche [37] und sorgen so dafür, dass die Selektivität und Avidität der Interaktion dramatisch ansteigt (Faktor  $10^7$ ) [38] (siehe Abbildung 5)



**Abbildung 5: Das Hämagglutinin des Influenzavirus bindet auf der Zelloberfläche an den trivalenten Rezeptor aus Sialinsäure**

Das Influenza A Glycoprotein Hämagglutinin erkennt verschiedene Konformationen der Sialinsäure-Moleküle. Die Bindung dieser Moleküle erleichtert das Eintreten des Virus in die Zelle. Eine einzelne Zelle kann durch mehrere Viren infiziert werden. (Entnommen aus <http://www.fluinfectomics.ubc.ca/home/spotlight-1/>)

Affinitätsanstiege durch Multivalente Interaktionen können auch auf makroskopischer Ebene beobachtet werden. Ein Paradebeispiel liefert der Gecko, welcher durch Millionen extrem feiner Haarstrukturen („Spatulae“) unter seinem Fuß in der Lage ist sich kopfüber mit nur einem Zeh an einer frisch polierten Glasplatte zu halten. Die Spatulae passen sich

---

genau an die Oberfläche an, auf der sich der Gecko bewegt, wobei die Summe der einzelnen van der Waals-Wechselwirkungen den Gecko an der Decke hält [39]

Durch die Multivalenz wird evolutionäre Effizienz dadurch erzielt, dass existierende Interaktionen vervielfältigt, statt neu kreiert werden. In Multiproteinkomplexen kann durch Heteromultivalenz nicht nur die Stärke, sondern auch die Spezifität einer Bindung erhöht werden. So wird in zellulären Signalwegen durch den modulartigen Proteinaufbau in Kombination mit multivalenten Effekten die spezifische Erkennung einer Vielzahl von Liganden durch ein einziges Protein ermöglicht [22]. Zelluläre Antworten können nicht nur an- und abgeschaltet werden, durch multivalente Wechselwirkungen werden auch abgestufte Zellantworten ermöglicht. RXR (retinoid X receptor) fungiert beispielsweise bei Anwesenheit seines Liganden als Transkriptionsfaktor RXR-L [40], welcher an spezielle CRBP-II Elemente (cellular retinol binding protein II element) der DNA bindet. Die Affinität von RXR-L für ein CRBP-II Element allein ist schwach. Erhöht sich jedoch die Konzentration des Liganden in der Zelle, entstehen multimere Spezies von RXR-L, die multivalent mit den CRBP-II Elementen interagieren, wodurch die Affinität zur DNA erhöht und damit einhergehend die Expressionsrate des Zielproteins gesteigert wird.

Andere Transkriptionsfaktoren können Expressionsraten dadurch regulieren, dass ihre verschiedenen Proteindomänen multivalent und dadurch spezifisch mit verschiedenen Proteinen, wie z.B. Untereinheiten der RNA-Polymerase II, Spleißfaktoren, Histon-modifizierenden Enzymen und anderen Transkriptionsfaktoren wechselwirken können.

## **1.4 PROTEINDOMÄNEN II – PRD, PRS & WW-DOMÄNEN**

### **1.4.1 PROLINREICHE SEQUENZEN-BINDEnde DOMÄNEN (PRD)**

Viele Proteindomänen, die in Multiproteinkomplex-vermittelten Signalwegen vorkommen, binden multivalent an oberflächenexponierte prolinreiche Sequenzen (kurz: PRS, siehe Abschnitt 1.4.1.1) ihrer Interaktionspartner. Bis heute können sechs Familien der Prolinreichen Sequenzen-bindenden Domänen (PRD) unterschieden werden: Die SH3 (*Src-homology-3*)-Domänen [41,42], die EVH1-Domänen [43–45], die GYF- oder CD2-Domänen [46,47], die UEV-Domänen [48,49], das aus einer Domäne bestehende Profilin [50,51] und die in dieser Arbeit untersuchten WW-Domänen [52,53]. Ein charakteristisches Merkmal aller PRDs sind hoch konservierte, räumlich nahe beieinander liegende und an der Oberfläche lokalisierte Ansammlungen aromatischer Aminosäuren, die auch als „aromatische Wiege“ [54] bezeichnet werden und die Erkennung spezifischer PRS ermöglichen [55].

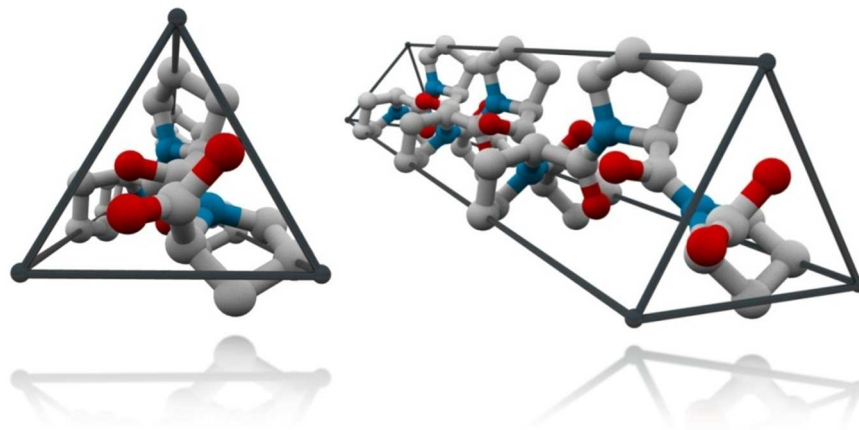
Die PRDs erkennen ein i.d.R. 3-6 Reste umfassendes, Prolin enthaltendes „Kernmotiv“ eines 5-10 Aminosäuren langen Sequenzabschnittes auf der Oberfläche des Bindepartners [55]. Das Kernmotiv wird häufig von mehreren Mitgliedern einer Domänenfamilie erkannt [55]. PRDs interagieren mit den PRS im niedrigaffinen Bereich ( $K_d$  zwischen 1-500  $\mu$ M) und benötigen zusätzliche flankierende Epitope, um die

---

notwendige Spezifität zu erreichen [56]. Durch die in den PRD räumlich sehr nahe beieinander liegenden N- und C-Termini, kann eine PRD leicht in ein Trägerprotein eingepasst werden, ohne dessen Raumstruktur wesentlich zu verändern [55]. Diese Eigenschaft stellt den Zugang zur PRS-Bindestelle sicher und schafft dadurch eine gute Voraussetzung zur Rekrutierung von Zielproteinen [55].

#### **1.4.1.1 PPII-Helices und Prolinreiche Sequenzen (PRS)**

Prolinreiche Sequenzen (PRS) sind in prokaryotischen und eukaryotischen Proteinen weit verbreitet. Sie enthalten mehrere aufeinanderfolgende Proline, von denen aber nur einige essentiell für die Bindung an PRDs zu sein scheinen. Allen PRS ist gemein, dass sie in Lösung bevorzugt eine linksgängige Poly-Prolin Type II (PPII)-Helix Konformation (mit  $\Phi=-75^\circ$ ,  $\Psi=145^\circ$  [57]) einnehmen, ähnlich der von poly-L-Prolin und Kollagen [58,59]. Eine Helix-Drehung wird dabei aus drei Prolin-Resten gebildet, man spricht hierbei von einer Pseudo  $C_3$  Rotationssymmetrie. Zusätzlich besitzt die PPII-Helix eine  $C_2$  pseudosymmetrische Rotationsachse senkrecht zur langen Helixachse, entsprechend erinnert die Grundform der PPII-Helix stark an ein trigonales Prisma (vgl. Abbildung 6 ).



**Abbildung 6: PPII-Helix aus 5 Prolin-Resten**

Gezeigt ist eine typische PPII-Helix bestehend aus 5 Prolin-Resten, von der Seite (rechts) und in der Draufsicht (links).  $\Phi=-75^\circ$ ,  $\Psi=145^\circ$  Eine Sequenz von vier oder mehr Prolinen in einer Reihe nimmt in Lösung eine bevorzugte PPII-Helix-Konformation an (Autodesk MAYA 2012, in Zusammenarbeit mit *Barth van Rossum*).

Aus dieser Architektur resultiert eine sehr stabile Grundstruktur. So konnte die PPII-Helix nicht nur in pankreatischen Polypeptid-Hormonen und Neuropeptiden, sondern auch als wichtiges Strukturelement in Kollagen gefunden werden [60,61]. Aber auch globuläre Proteine besitzen PPII-Helices. Entsprechend wurden auf den Oberflächen verschiedenster Proteine zahlreiche kurze Sequenzen lokalisiert, welche eine PPII-Konformation einnehmen [62]. Interessanterweise enthielten nur sehr wenige dieser Sequenzen gänzlich Prolin-Reste, vielmehr wird in derartigen Polypeptidketten die PPII-Helix-Konformation dadurch stabilisiert, dass jeder dritte Aminosäurerest durch ein Prolin repräsentiert wird  $(Pxx)_n$  [59].

---

Die PRS besitzen nicht nur als strukturbildendes Element Bedeutung. Wie oben bereits angesprochen, bilden sie auch die Grundlage für eine Vielzahl von Protein-Protein-Interaktionen. Tandemartig wiederholte PRS können darüber hinaus Polyphenole, Peptidoglycane und Zytoskelettproteine binden [59,63,64]. Nicht periodisch wiederholte PRS vermitteln dagegen hauptsächlich Protein-Protein-Interaktionen [59].

Strukturbiologische Untersuchungen haben gezeigt, dass PRDs die PRS nur erkennen können, wenn diese in der PPII-Helix-Konformation vorliegen [56]. Durch die Ausbildung einer Wasserstoffbrücke zwischen einem konservierten Tryptophan auf der Oberfläche der PRD und einem Rückgrat-Carbonylsauerstoffatom der zentralen PRS wird die Wechselwirkung maßgeblich verstärkt [55,65,66]. Dazu ermöglicht die C<sub>2</sub> Pseudosymmetrie der PPII-Helix die PRS-Bindung in zwei Orientierungen (vorwärts (N-C-terminal) und rückwärts (C-N-terminal)). Interessanterweise werden dabei sowohl die gleichen Erkennungselemente des Peptids und Wasserstoffbrückendonoren der Protein-Seitenketten, als auch die gleichen hydrophoben Bindestellen der Domäne verwendet [55]. Bisher konnten beide Bindungsvarianten für SH3- und WW-Domänen sowie für Profilin nachgewiesen werden. Dabei ist anzumerken, dass die Domänen eine Präferenz für jeweils nur eine Orientierung zeigten [55].

#### **1.4.2 DIE WW-DOMÄNE**

Im Jahre 1994 studierten *Bork und Sudol* [67] mittels computerbasierter Analysemethoden die Sequenzen verschiedenster Proteine und entdeckten dabei eine sich häufig wiederholende Sequenz, welche aufgrund zweier typischer Tryptophan-Reste (WW) in ihrer Consensus-Sequenz fortan als WW-Domäne (Synonyme: WWP-Domäne [68], Rsp5-Domäne [69]) bezeichnet wurde. Die WW-Domäne besitzt zusätzlich zu den zwei konservierten Tryptophanen einen in seiner Positionierung unveränderlichen Prolin-Rest, weshalb sie auch unter den Namen WWP-Domäne bekannt ist. Bislang konnten über 4700 dieser kleinen (38-40 AS) globulären Domänen [67] in über 3200 Proteinen lokalisiert werden (vgl. Tabelle 1.4.I). Erstaunlicherweise findet man die WW-Domäne nicht in Bakterien. Es scheint, als stünde das Auftreten der WW-Domäne im Zusammenhang mit der evolutionären Entstehung des Zellkerns, da die Anzahl WW-Domänen-haltiger Proteine, ausgehend von einzelligen eukaryotischen Organismen (*Saccharomyces*) über Insekten (*Drosophila*) bis hin zum Menschen, stark zunimmt (siehe Tabelle 1.4.I).

---



**Tabelle 1.4.I: Anzahl der WW-Domänen und WW-Domänen-haltiger Proteine innerhalb repräsentativer Spezies**

Familie	Proteine	Spezies	WW-	
			Domänen	Proteine
Gesamt		272	4771	3214
Pilze	665	<i>S. cerevisiae</i>	8	5
Pflanzen	349	<i>A. thaliana</i>	28	19
Fadenwürmer	142	<i>C. elegans</i>	42	27
Insekten	473	<i>D. melanogaster</i>	71	45
Amphibien	103	<i>X. laevis</i>	31	53
Säugetiere	589	<i>M. musculus</i>	218	128
		<i>H. sapiens</i>	261	171

Die angegebenen Zahlen repräsentieren die Gesamtzahl der WW-Domänen bzw. WW-Domänen enthaltender Proteine innerhalb einer Familie und einer repräsentativen Spezies. Die Daten sind der Pfam Datenbank (pfam.sanger.ac.uk Version 26.0 2011/06) entnommen.

Interessant ist, dass die 171 humanen WW-Domänen-haltigen Proteine keiner bestimmten Proteinfamilie zugeordnet sind, vielmehr wird ein breites Funktionsspektrum abgedeckt. Darunter sind Zytoskelett-assoziierte Proteine (Dystrophin) ebenso vertreten, wie regulatorische Proteine (Yap65 *Yes-associated protein*) und Transkriptionsfaktoren (z.B. TCERG1 *Transcription elongation regulator 1*). Aber auch im Ubiquitin-vermittelten Proteinabbau spielen WW-Domänen-haltige Proteine ((Nedd4) eine wichtige Rolle. Sogar im späten Stadium einiger Virusinfektionen (Knospung) haben WW-Domänen ihre Bedeutung. So sind beispielsweise Interaktionen zwischen den Nedd4-WW-Domänen und einigen viralen Proteinen Voraussetzung für die Entwicklung einiger Viren [70]. Dazu konnte gezeigt werden, dass das Adenovirus *Adenovirus dodecahedron* Proteine in humane Zellen transduziert, indem es die Nedd4-WW-Domänen 2 und 4 als universelle Adaptoren nutzt [71].

WW-Domänen kommen häufig mehrfach im gleichen Protein vor, wobei sich die WW-Domänensequenzen, abgesehen von der Consensus-Sequenz, z.T. deutlich unterscheiden können [52]. Einzelne WW-Domänen verschiedener Proteine besitzen aber häufig eine hohe Sequenzhomologie [52]. Durch alternatives Spleißen ist zudem die Möglichkeit gegeben, die Anzahl und die Position der WW-Domänen im Protein zu verändern [72]. In diesem Zusammenhang werden oft auch die Linkersequenzen zwischen den Domänen variiert. Daher ist anzunehmen, dass auch Länge und Struktur des Interdomänen-Linkers für die physiologische Funktion des Proteins von Bedeutung sind. Der funktionelle Kontext der Linkersequenz ist bisher jedoch noch wenig erforscht.

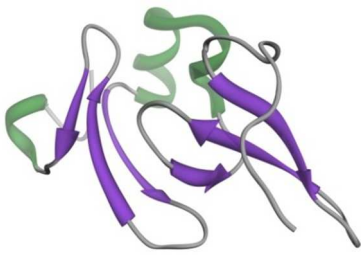

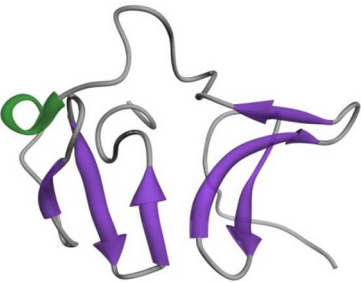

Im humanen Nedd4-2 Protein verringert eine Phosphorylierung an definierten Stellen der Linkerregionen (S<sup>338</sup> im WW1-2-Linker bzw. S<sup>444</sup> im WW2-3-Linker) beispielsweise die Interaktion von Nedd4-2 zum epithelialen Na<sup>+</sup>-Kanal (ENaC) und führt so zu einer erhöhten ENaC-Expression an der Zelloberfläche, wodurch die Na<sup>+</sup>-Ionen-Balance und damit einhergehend der Blutdruck reguliert wird [73].

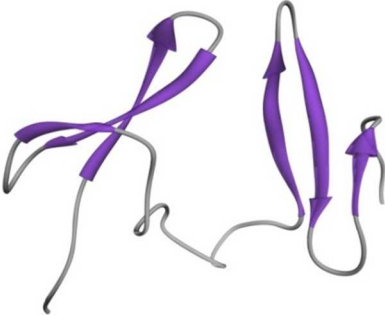
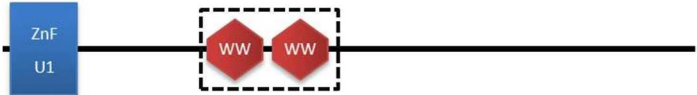


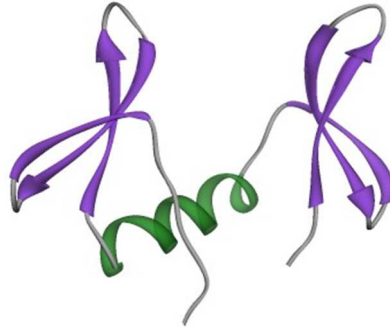

### 1.4.2.1 Strukturelle Eigenschaften und Organisation

Der strukturelle Kernbereich der WW-Domäne besteht aus einem leicht gekrümmten, drei-strängigen antiparallelen  $\beta$ -Faltblatt. Das N-terminale Tryptophan der Consensus-Sequenz ist essentiell für die Faltung [53]. Es wird von zwei Prolin-Resten bedeckt, wodurch sich eine hydrophobe Aufwölbung auf der konvexen Seite des Faltblattes ausbildet [53]. Das zweite Tryptophan der Consensus-Sequenz ist auf der Proteinoberfläche exponiert. Es ist Teil eines hydrophoben Clusters, in welchem die aromatischen Seitenketten eines Tyrosins und eines Tryptophans im Winkel von etwa  $90^\circ$  angeordnet sind und so die Form der Prolin-Bindungstasche bestimmen [53,74].

In der Protein-Datenbank (PDB, <http://www.rcsb.org>) sind bislang 125 Strukturen von isolierten WW-Domänen und 45 Strukturen von WW-Domänen im Komplex mit einem prolinreichen Liganden veröffentlicht. Von tandemartig angeordneten WW-Domänen wurden seit 2002 jedoch nur 5 Strukturen gelöst (siehe Tabelle 1.4.II). Da die WW-Domänen in der Mehrzahl der Proteine nicht einzeln, sondern tandemartig angeordnet sind, wurden in den letzten Jahren große Anstrengungen unternommen mehr über deren Struktur und Funktion zu erfahren.

**Tabelle 1.4.II: Publierte Strukturen von Tandem-WW-Domänen**

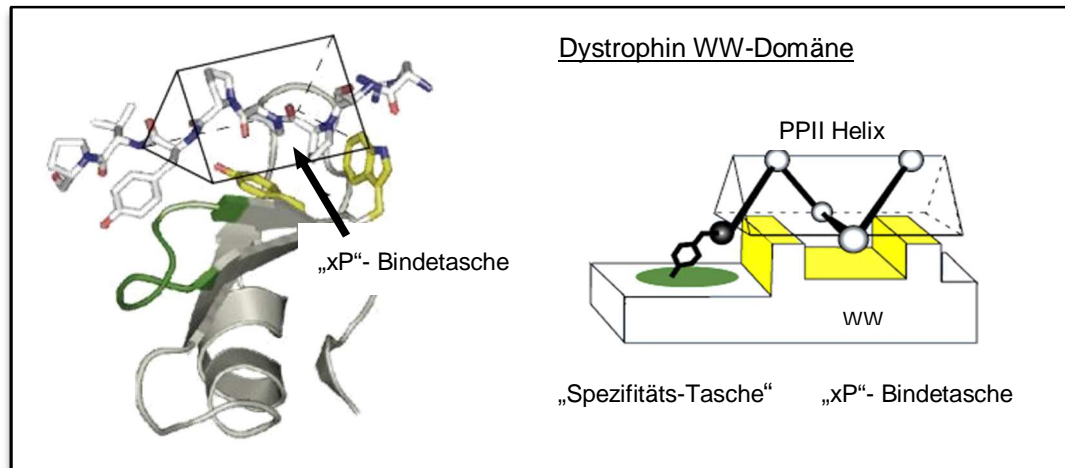
Tandem WW-Domäne	Informationen
<b>WW1-2 FBP11/HYPA</b> 	<i>Prä-mRNA-processing factor 40 homolog A</i> Autoren: Y. Jiang; H. Hu Methode: Lösungs-NMR PDB Code: 2L5F Veröffentlicht: 2/2011 Spezies: <i>H. sapiens</i> 
<b>WW2-3 SMURF2</b> 	<i>E3 Ubiquitin-Protein Ligase SMURF2</i> Autoren: J. D. Forman-Kay; H. Lin; A. Chong; J. Wrana Methode: Lösungs-NMR PDB-Code: 2KXQ Veröffentlicht: 10/2010 Spezies: <i>H. sapiens</i> 

Tandem WW-Domäne	Informationen
FBP21 WW1-2	<p data-bbox="695 259 975 286"><i>Formin-binding protein 21</i></p> <p data-bbox="695 315 1158 342">Autoren: X. Huang; J. Zhang; J. Wu; Y. Shi</p> <p data-bbox="695 349 959 376">Methode: Lösungs-NMR</p> <p data-bbox="695 383 890 409">PDB-Code: 2JXW</p> <p data-bbox="695 439 940 465">Veröffentlicht: 12/2008</p> <p data-bbox="695 472 911 499">Spezies: <i>H. sapiens</i></p>
	
WW3-4 von Suppressor of Deltex	<p data-bbox="695 674 1018 701"><i>Suppressor of Deltex (Su(dx))</i></p> <p data-bbox="695 707 991 734">E3 Ubiquitin-Protein Ligase</p> <p data-bbox="695 763 1294 831">Autoren: O. Y. Fedoroff; J. M. Avis; A. P. Golovanov, M. Baron; S.A. Townson</p> <p data-bbox="695 837 959 864">Methode: Lösungs-NMR</p> <p data-bbox="695 871 884 898">PDB-Code: 1TK7</p> <p data-bbox="695 927 927 954">Veröffentlicht: 7/2004</p> <p data-bbox="695 960 975 987">Spezies: <i>D. melanogaster</i></p>
	
WW1-2 Prp40	<p data-bbox="695 1144 1358 1171"><i>Yeast splicing factor pre-mRNA processing protein 40 (Prp40)</i></p> <p data-bbox="695 1189 1262 1216">Autoren: S. Wiesner, G. Stier, M. Sattler, M.J. Macias</p> <p data-bbox="695 1223 959 1249">Methode: Lösungs-NMR</p> <p data-bbox="695 1256 890 1283">PDB-Code: 106W</p> <p data-bbox="695 1290 940 1317">Veröffentlicht: 10/2002</p> <p data-bbox="695 1323 940 1350">Spezies: <i>S. cerevisiae</i></p>
	

Die Informationen zur modularen Architektur der Proteine wurden mit Hilfe von SMART [75,76] zusammengetragen.

### 1.4.2.2 Bindeeigenschaften

Die WW-Domänen binden bevorzugt PRS mit zwei direkt benachbarten Prolin-Resten, interagieren aber auch mit PRS, in denen das erste der beiden Proline durch eine andere Aminosäure, wie beispielsweise Leuzin substituiert ist. Man spricht daher vom „xP“-Motiv, das mit der „xP“-Bindetasche der Domäne wechselwirkt [56] (siehe Abbildung 7).



**Abbildung 7: Struktur und Bindemechanismus der Dystrophin WW-Domäne**

Links ist die Struktur der Dystrophin WW-Domäne im Komplex mit einem prolinreichen Liganden (PDB-Code: 1EG4) gezeigt. Rechts ist schematisch der generelle Mechanismus der Liganden-Erkennung dargestellt. Der Kernbereich der Interaktionsfläche („Core-Recognition-Surface“) besteht aus einer „xP“-Bindetasche, welche durch aromatische Aminosäuren (gelb) gebildet wird. Die angrenzende Spezifitäts-Tasche (grün) ist ein weniger konserviertes Motiv in den WW-Domänen (entnommen aus [56])

Einige WW-Domänen besitzen zusätzlich eine zweite „xP“-Bindetasche („xP2“). *Kato et al.* bestätigten nicht nur das Vorkommen der „xP2“-Bindetasche in WW-Domänen der Klasse II und III (Einteilung nach *Sudol* siehe Abschnitt 1.4.2.3, Bsp. CA150-WW2), sondern konnten durch Mutagenese-Studien zusätzlich die Bedeutung der „xP2“ für das Bindungsereignis untermauern [77].

Interessanterweise können einige WW-Domänen sogar an phosphorylierte Serin- (*poS*) oder Threonin-Reste (*poT*) binden [78]. Beispiele hierfür sind die WW-Domänen der essentiellen mitotischen Prolyl-Isomerase Pin1 und die der Ubiquitin-Ligase Nedd-4.

### 1.4.2.3 Klassifizierung der WW-Domänen anhand ihrer Bindungsmotive

WW-Domänen können anhand ihrer bevorzugten Bindungsmotive in verschiedene WW-Klassen eingeordnet werden. Es existieren verschiedene mehr oder weniger präzise Klassifizierungsansätze, von denen hier zwei vorgestellt werden sollen:

*Marius Sudol* (2000, [79]) unterteilte die WW-Domänen, entsprechend ihren Bindepräferenzen in vier Klassen. WW-Domänen der Klasse I binden PPxY (**PY**)-Motive. WW-Domänen der Klasse II binden PPLP (**PL**)-Motive, denen oft lange Prolin-Sequenzen vorausgehen. Die WW-Domänen der Klasse III erkennen Motive mit Arginin- oder Lysin-Resten (**PR**-Motiv) und die WW-Domänen der Klasse IV erkennen Motive mit phosphorylierten Serin- oder Threonin-Resten (*poS*/*poTP*-Motiv) (vgl. Tabelle 1.4.III).

Die zweite neuere Klassifikation nach *Ball et al.* 2005 [55], die u.a. auf umfangreichen Substitutionsanalysen von *Otte et al.* [80] und *Macias et al.* [81] basiert, unterteilt die WW-Domänen in fünf Klassen. *Ball et al.* ergänzte die Klassifizierung von *Sudol* um WW-

Domänen, die bevorzugt reine Prolin-Motive erkennen (Klasse V). Darüber hinaus wurde die Klasse III in zwei Subklassen unterteilt. Es gilt zu beachten, dass WW-Domänen deren Sequenzspezifitäten sich überschneiden, nicht eindeutig einer Klasse zugeordnet werden können.

**Tabelle 1.4.III: Gegenüberstellung verschiedener WW-Klassifizierungen**

WW-Klasse <i>Sudol</i> (Ball [55])	Beispiele	Motiv <i>Sudol</i> [79]	Motiv <i>Otte et al.</i> [80]
I (I)	YAP65, Nedd-4, Dystrophin, CA150-1,3	PPxY (PY)	(L/P)Pp(Y/poY)
II (II)	FBP11, CA150-2	PPLP (PL)	PPLPp
III (IIIa)	FE65	PR	(p/Φ)P(p/g)PPpR
III (IIIb)	FBP21	PR	(p/Φ)PP(R/K)gpPp
IV (IV)	PIN1, Nedd-4, CA150-1	poS/poTP	(poS/poT)P
(V)	PRP40-2, PIN1, CA150-2		(p/ Φ)PPPPP

Φ: hydrophobe Aminosäure, poY: Phosphotyrosin, poS: Phosphoserin, poT: Phosphothreonin. Kleine Buchstaben repräsentieren bevorzugte, aber nicht konservierte Aminosäurereste. Die Nomenklatur der Peptidmotive erfolgte nach *Aasland et al.* 2002 [82]

#### 1.4.2.4 Klinische Relevanz

Die WW-Domäne ist ursprünglich als ein sich wiederholendes Modul in Onkogen-assoziierten Proteinen identifiziert worden und daher für die Krebsforschung von großer Bedeutung. So tragen Patienten mit Brust- oder Ovarialkrebs häufig Mutationen in der WWOX Chromosomenregion (16p23), welche einen potentiellen Tumorsuppressor mit zwei WW-Domänen codiert [83]. Der regelmäßige Ausfall der WWOX-Expression korreliert bei Brustkrebs-Patienten mit dem Status des Östrogen-Rezeptors [84].

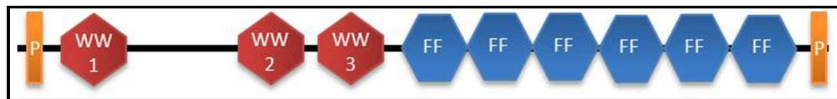
In den letzten Jahren wurden einige PRDs als mögliche therapeutische Zielstrukturen identifiziert. So ist eine weitere mit WW-Domänen assoziierte Krankheit das „Liddle-Syndrom“. Dabei handelt es sich um eine Hypertensionskrankheit (erheblicher Bluthochdruck). Das „Liddle-Syndrom“ resultiert aus einer Mutation im PPxY-Motiv der zytoplasmatischen Domäne des epithelialen Natrium-Kanals (ENaC), das durch die Nedd4-WW-Domäne(n) erkannt wird [85,86]. Auch Huntington (HD, *Huntington Disease*), eine neurodegenerative Erbkrankheit, konnte mit WW-Domänen in Zusammenhang gebracht werden. So wird die prolinreiche Region im Huntingtin-Protein (HTT) durch die CA150-WW-Domänen erkannt, was eine wichtige Rolle im Krankheitsverlauf zu spielen scheint [87] (siehe auch Abschnitt 1.5). Für die Ausbildung des „*Golabi-Ito-Hal* Syndroms“, einer bestimmten Form der X-Chromosom-vermittelten mentalen Retardierung, wird eine „Missense Mutation“ in der WW-Domäne des Poly-Glutamin-Bindeproteins PQBP-1 verantwortlich gemacht [88]. Fehlregulierte WW-Domänen scheinen darüber hinaus auch mit Alzheimer [89] und der „*Dychehenne-Muskeldystrophie*“ [54] assoziiert zu sein.

## 1.5 CA150, EIN WW-DOMÄNEN-HALTIGER TRANSKRIPTIONSFAKTOR

CA150 (*co-activator of 150 kDa; Cancer Antigen 150*) ist ein Transkriptionsfaktor (Synonym: TCERG1; *transcription elongation regulator 1*), der sich bei TATA-Box-abhängigen Promotoren auf die Elongation der RNA-Polymerase II auswirkt [90]. Er ist in den so genannten „Speckle“-Regionen, Substrukturen des Zellkerns, in denen sich Prä-mRNA Spleißfaktoren anreichern, lokalisiert [91] und wurde ursprünglich aus Zellkernextrakten von HeLa-Zellen isoliert. Weitere Untersuchungen zeigten, dass CA150 als transkriptioneller Co-Faktor wirkt, der u.a. die Tat- (*Transactivator of Transcription*) Gen-Aktivierung des Humanen Immundefizienz-Virus (HIV-1, *human immunodeficiency virus type 1*) regulieren kann [92]. Suné und Garcia-Blanco zeigten später, dass eine transiente Überexpression von CA150 die basale und die Tat-aktivierte Transkription am HIV-1 Promotor durch eine verstärkte Inhibition der Elongation reduzierte [90].

### 1.5.1 STRUKTURELLE MERKMALE

Je nach Isoform besteht das humane CA150 aus 1077-1098 Aminosäuren entsprechend einem Molekulargewicht von ca. 124 kDa. Das Protein besitzt eine modulare Architektur und enthält drei N-terminale WW-Domänen, sechs C-terminal aufeinanderfolgende, tandemartig angeordnete FF-Domänen sowie je eine N- und C-terminale PRS [92–95] (vgl. Abbildung 8).



**Abbildung 8: Modulare Architektur des Transkriptionsfaktors CA150**

P: prolinreiche Region, WW: WW-Domäne; FF: FF-Domäne

Die Struktur des vollständigen Proteins konnte bis heute nicht aufgeklärt werden. Es sind jedoch hochaufgelöste Strukturen der isolierten WW-Domänen und FF-Domänen verfügbar.

#### 1.5.1.1 Strukturelle Aspekte der drei CA150-WW-Domänen

Die Sequenzanalyse der drei CA150-WW-Domänen zeigt, dass sich die WW1 und WW2 sehr ähneln, die WW3-Sequenz jedoch am divergentesten ist. Die Consensus-Sequenzen sind jedoch in allen drei Domänen absolut identisch (vgl. Abbildung 9).

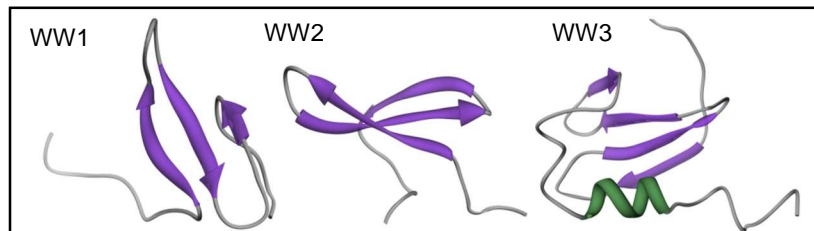
WW1	P	P	T	E	E	I	W	V	E	N	K	T	P	D	G	K	V	Y	Y	Y	N	A	R	T	R	E	S	A	W	T	K	P	D	G
WW2	A	T	A	V	S	E	W	T	E	Y	K	T	A	D	G	K	T	Y	Y	Y	N	N	R	T	L	E	S	T	W	E	K	P	Q	E
WW3	P	I	P	G	T	P	W	C	V	V	W	T	G	D	E	R	V	F	F	Y	N	P	T	T	R	L	S	M	W	D	R	P	D	D

**Abbildung 9: Sequenzvergleich der CA150 WW-Domänen**

Die Rechtecke markieren die Positionen der drei beta-Stränge. Reste der Consensus-Sequenz sind orange hinterlegt. Reste die in zwei Domänen identisch sind, wurden blau hinterlegt.

Die drei WW-Domänen besitzen sehr unterschiedliche Bindeeigenschaften. So zeigt die WW1 Bindungscharakteristika der Klassen I und IV [80] (vgl. Abschnitt 1.4.2.3), die WW2 die der Klassen II und V [80,96–98] und die WW3 könnte der Klasse I zugeordnet werden, verhält sich aber sehr promiskuitiv [98]. Zudem ist anzumerken, dass nur die WW2 über eine zweite „xP“-Bindetasche („xP2“) verfügt (siehe Abschnitt 5.5).

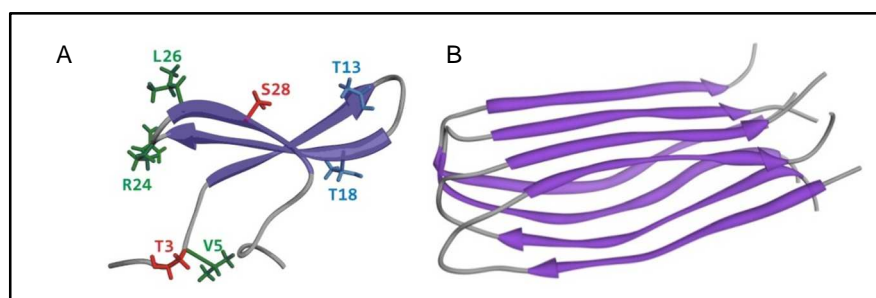
Mit Hilfe der hochauflösenden NMR-Spektroskopie konnten die Strukturen der drei WW-Domänen gelöst werden. Diese sind in Abbildung 10 schematisch dargestellt.



**Abbildung 10: Backbone-Strukturen der CA150-WW-Domänen**

Cartoon-Darstellung der Backbone-Strukturen der ersten (1, PDB-Code: 2YSI), zweiten (2, PDB-Code: 1E0L) und dritten (3, PDB-Code: 2DK7) WW-Domäne von CA150.

Die CA150-WW2 besitzt die gleiche Sequenz, wie FBP28, ein kleines dreisträngiges antiparalleles  $\beta$ -Faltblatt-Protein aus Mäusen [99]. FBP28 ist in der Lage *in vitro* innerhalb weniger Mikrosekunden mit einer offenbar zweistufigen Kinetik das drei-strängige  $\beta$ -Faltblatt auszubilden [100,101]. Aufgrund der Sequenzidentität kann dieser Faltungsmechanismus auch auf die CA150-WW2 übertragen werden. Interessant ist, dass die WW2 unter physiologischen Bedingungen (37°C, pH 7.0) hochgradig geordnete amyloide Fibrillen ausbilden kann, wobei der Fibrillenbildung eine denaturierte Form der WW2 vorausgeht [102]. Mittels Festkörper-NMR-Spektroskopie konnte die Struktur dieser Fibrillen gelöst werden (Abbildung 11 B).



**Abbildung 11: Native und fibrilläre Struktur von CA150 WW2**

Cartoonstruktur der **A** nativen WW2 von CA150 (PDB-Code: 1E0L), **A**) Seitenketten, die im Fibrillenstadium in Interaktion stehen, sind in identischen Farben dargestellt. (entnommen aus [102]), **B** Cartoon der fibrillären WW2 (PDB: 2NNT).

Amyloide Fibrillen werden häufig in Krankheitsstadien aber auch *in vitro* bei vielen Proteinen und Peptiden ausgebildet [103]. Da pathogene Fibrillenbildungen bereits mit einer Reihe schwerer humaner Krankheiten, einschließlich Alzheimer, assoziiert werden



---

konnten [104], besteht der Verdacht, dass auch CA150-WW2-Fibrillen pathogen wirken könnten. In vivo wurden bis heute allerdings noch keine CA150-WW2-Fibrillen nachgewiesen.

### **1.5.2 MODIFIKATIONEN VON CA150**

Durch alternatives Spleißen können zwei Transkriptvarianten (Isoformen) des humanen CA150 Proteins generiert werden, wobei in Isoform2 eine 21 Aminosäuren lange Sequenz (AS 379-399) im Linkerbereich zwischen WW1 und WW2 fehlt. Das CA150 der Maus existiert ebenfalls in zwei Isoformen. Diese werden im Gegensatz zu den humanen Versionen in einer Variante mit sechs und einer mit nur fünf FF-Domänen realisiert [105]. Im parasitären Wurm *Schistosoma mansoni* konnten sogar vier verschiedene, geschlechterspezifisch exprimierte, CA150-Transkriptvarianten identifiziert werden [106].

Eine andere Modifikation von CA150 ist dessen Phosphorylierung an S<sup>460</sup> und T<sup>468</sup> (Linkerbereich zwischen WW2 und WW3) [107], wobei über die strukturellen oder physiologischen Auswirkungen derzeit noch keine Informationen vorliegen.

### **1.5.3 PHARMAKOLOGISCHE BEDEUTUNG VON CA150 IN HUNTINGTON**

Ein sehr prominenter Interaktionspartner von CA150 ist das Huntingtin-Protein (HTT) [87]. Zwar ist es bislang noch nicht gelungen die genaue Rolle dieser Interaktion im HD-Krankheitsverlauf aufzuklären, fest steht jedoch, dass eine Überexpression von CA150 in Ratten, welche das mutierte HTT exprimierten, eine Vermehrung der neuritischen Aggregate zur Folge hat [108].

*Holbert et al.* haben gezeigt, dass eine Veränderung der Anzahl der Glutamin-Alanin-Wiederholungen im Linkerbereich von CA150-WW1-2 (normalerweise 38 Q-A Wiederholungen) das Ausbruchsalter von Huntington beeinflusst [87]. *Arango et al.* zeigten, dass CA150 aktiv bleibt, aber vermehrt im Zytoplasma angereichert wird, wenn die Q-A gänzlich fehlen [108].

In humanen Populationen gibt es drei verschiedene CA150 Allele (A, B, C). Erstaunlicherweise wird Allel A ausschließlich in nicht an HD-erkrankten Menschen und Allel C ausschließlich in HD-Patienten gefunden. Allel B trat hingegen gleichermaßen in nicht HD- und HD-Patienten auf. Das Allel C steht mit einem früheren Ausbruch der Krankheit (~25 Jahre gegenüber 40 Jahre bei Allel B) und einer höher expandierten CAG Wiederholung im CA150 in Verbindung [109]. Letztere führt zu einer Verlängerung der (Q-A)<sub>38</sub> Wiederholungen im WW1-2-Linkerbereich.

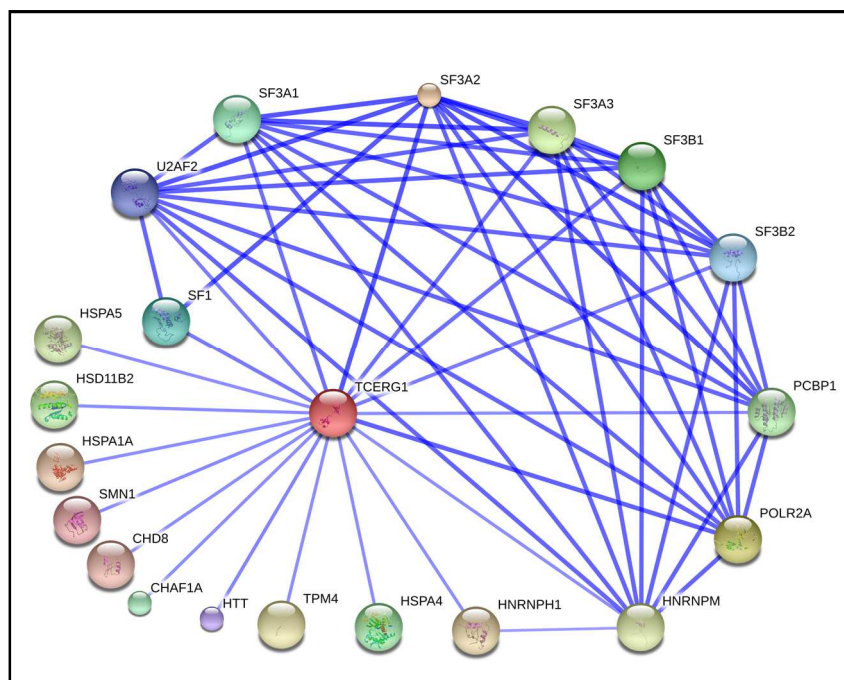
*Arango et al.* [108] nehmen an, dass eine Fehlfunktion von CA150 und eine gleichzeitige Mutation in Huntingtin maßgeblich am Ausbruch von HD beteiligt sind, da CA150 zum einen mit HTT wechselwirkt und zum anderen gezeigt werden konnte, dass es in Abhängigkeit vom HD-Stadium zunehmend im Gehirn der HD-Patienten akkumuliert [87].

---



### 1.5.4 PHYSIOLOGISCHER KONTEXT VON CA150

Mit Hilfe der Massenspektrometrie konnte CA150 als ein Bestandteil des aktiven Spleißosoms identifiziert werden [110,111]. Weitere Untersuchungen zeigten, dass CA150 mit einer Vielzahl von Transkriptions- und Spleißkomponenten in Interaktion steht (siehe Tabelle 1.5.I). Ein wichtiger Interaktionspartner ist der Spleißfaktor SF1, der mit den CA150-WW-Domänen 1 und 2 wechselwirkt [112]. Beide Proteine interagieren ihrerseits mit dem Spleißfaktor U2AF<sup>65</sup> [91,113]. Es wird vermutet, dass CA150 als eine Art „Brücke“ fungiert, die SF1 mit der elongierenden RNAPII verbindet, und dadurch einerseits die Erkennung der mRNA und andererseits das Spleißen erleichtert [112,114]. Ferner konnten auch Wechselwirkungen zwischen CA150 und Komponenten des SF3-Spleißkomplexes gezeigt werden [112,115]. Über die FF-Domänen tritt CA150 auch in direkte Wechselwirkung mit der phosphorylierten CTD der RNA-Polymerase II (RNAPII) [116]. Letztere verbindet in eukaryotischen Organismen die Transkription mit der prä-mRNA-Prozessierung (Spleißen). Zusätzlich konnte gezeigt werden, dass die FF-Domänen für die Lokalisation von CA150 in den „Speckles“ erforderlich sind [91]. Wie *Sánchez-Alvarez et al.* herausfanden, interagiert CA150 darüber hinaus mit Tat-Co-Faktoren. So konnten Wechselwirkungen zu P-TEFb, zum Tat-spezifischen Spleißfaktor Tat-SF1 sowie zur RAP30-Untereinheit des Transkriptionsfaktors TFIIF nachgewiesen werden [91]. Neben den aufgeführten Beispielen interagiert CA150 noch mit vielen weiteren Proteinen bzw. Proteinkomplexen. Eine Auswahl dieser vielfältigen direkten und indirekten Wechselbeziehungen ist in Abbildung 12 und Tabelle 1.5.I dargestellt.



**Abbildung 12: Netzwerk ausgewählter CA150 (TCERG1) Interaktionen**

CA150 (TCERG1) und dessen Interaktionspartner sind als verschiedenfarbige Kugeln dargestellt, siehe auch Tabelle 1.5.I. Die Stärke der Verbindungslinien repräsentiert die Stärke der Wechselbeziehung. Die Abbildung wurde mit STING 9.0 erzeugt [117]

**Tabelle 1.5.I: Interaktionspartner von CA150**

Abk.	Name	Funktion
CHAF1A	<i>Chromatin assembly factor 1, subunit A</i>	Chromatin Assemblierung, Kernkomponente des CAD-1 Komplexes
CHD8	<i>Chromodomain helicase DNA binding protein 8</i>	DNA-Helikase/ Transkriptionsrepressor
HNRNPH1	<i>Heterogeneous nuclear ribonucleoprotein H1</i>	Zellkernprotein
HNRNPM	<i>Heterogeneous nuclear ribonucleoprotein M</i>	Prä-mRNA-bindendes Protein, Spleißen
HSD11B2	<i>Hydroxysteroid (11-beta) dehydrogenase 2</i>	Enzym, moduliert intramolekulare Glukokortikoid-Level
HSPA1A	<i>Heat shock 70 kDa protein 1A</i>	Chaperon
HSPA4	<i>Heat shock 70 kDa protein 4</i>	Chaperon
HSPA5	<i>Heat shock 70 kDa protein 5</i>	Chaperon
HTT	<i>Huntingtin</i>	Mikrotubuli-vermittelter Transport, oder Visikelfunktion
PCBP1	<i>Poly(rC) binding protein 1</i>	ss-DNA-bindendes Protein
POLR2A	<i>DNA-directed RNA polymerase II polypeptide A</i>	Polymerase
SF1	<i>Splicing factor 1</i>	Spleißen
SF3A1	<i>Splicing factor 3a, subunit 1</i>	Spleißen
SF3A2	<i>Splicing factor 3a, subunit 2</i>	Spleißen
SF3A3	<i>Splicing factor 3a, subunit 3</i>	Spleißen
SF3B1	<i>Splicing factor 3b, subunit 1</i>	Spleißen
SF3B2	<i>Splicing factor 3b, subunit 2</i>	Spleißen
SF3B4	<i>Splicing factor 3b, subunit 4 [115]</i>	Spleißen
SMN1	<i>Survival of motor neuron 1</i>	Telomer
TPM4	<i>Tropomyosin 4</i>	Aktinfilamentbindung
U2AF2	<i>U2 small nuclear RNA auxiliary factor 2</i>	Prä-mRNA Spleißen

Mittels Mutagenesestudien konnten *Goldstrohm et al.* zeigen, dass die Repressionswirkung von CA150 durch die WW-Domänen WW1 und WW2 vermittelt wird [112].

Durch Mutationen der drei zentralen Tyrosine in der Sequenz der ersten bzw. der zweiten WW-Domäne zu Alanin konnte die Repressionswirkung von CA150 jeweils um 35% verringert werden. Die gleichzeitige Mutation beider WW-Domänen reduzierte diese sogar um 63%. Dazu wirkte sich die Mutation in der WW2 auch negativ auf die Bindung an SF1 aus. Die Mutation des FFY-Motives in WW3 zu drei Alaninen, zeigte hingegen keinerlei Auswirkung auf die Repressionswirkung. Tabelle 1.5.II fasst die Ergebnisse von *Goldstrohm et al.* zusammen.

**Tabelle 1.5.II: Ergebnisse der Mutagenesestudie von Goldstrohm et al. 2001 [112]**

<b>Mutierte Reste</b>	<b>Mutation</b>	<b>Effekt</b>
148-150 (in WW1)	YYY → AAA	Um 35% Verringerte Repression der Transkription
446-448 (in WW2)	YYY → AAA	Verlust der Interaktion zu SF1 Um 35% Verringerte Repression der Transkription
545-547 (in WW3)	FFY → AAA	Kein Effekt zu beobachten
148-150 (in WW1)	YYY → AAA	Um 63% Verringerte Repression der Transkription
446-448 (in WW2)	YYY → AAA	

Die Ergebnisse lassen vermuten, dass CA150 über die WW-Domänen multivalent mit seinen Interaktionspartnern wechselwirkt. Für diese These spricht auch die Tatsache, dass die WW-Domänen WW1 und WW2 mit der PRS von SF1 interagieren. Dabei blieb jedoch unklar, ob beide WW-Domänen gleichzeitig verschiedene Polyprolin-Motive der PRS binden, oder mit ein und demselben Motiv interagieren. Die Beantwortung dieser Frage war ein wichtiges Ziel der vorliegenden Arbeit.

## **2 ZIELSTELLUNG**

Viele humane Krankheiten, wie Alzheimer [89], Krebs [83]), Muskeldystrophie [54], „Liddle-Syndrom“ [118] und Huntington [119] können auf fehlregulierte WW-Domänen-vermittelte Signalwege zurückgeführt werden. WW-Domänen sind daher pharmakologisch sehr bedeutend. Erstmals beschrieben wurden sie 1994 von *Bork und Sudol* [67], die diesen Domänen auf Grund zweier hochkonservierter Tryptophane (W) innerhalb ihrer Consensus-Sequenz auch ihren Namen verliehen. Mit nur ca. 40 Aminosäuren gehören sie zu den kleinsten Proteindomänen. Sie binden prolinreiche Sequenzmotive (PRS) und sind, mit Ausnahme von Bakterien, hoch konserviert in allen Organismengruppen präsent. Bisher wurden über 4000 WW-Domänen in über 3000 Proteinen unterschiedlichster Funktionsklassen identifiziert. Eines dieser Proteine ist der humane Transkriptionsfaktor CA150 (*co-activator of 150 kDa, Synonym: TCERG1 transcription elongation regulator 1*), der neben sechs FF-Domänen und zwei N- und C-terminal lokalisierten PRS, gleich drei WW-Domänen besitzt, die durch außergewöhnlich lange Linkersequenzen (WW1-2: 165 AS; WW2-3 65 AS) miteinander verknüpft sind. CA150 ist an der Regulation der Elongation der RNA-Polymerase II, der Koordinierung der Transkription und der prä-mRNA-Prozessierung beteiligt (vgl. Abschnitt 1.5). Unter anderem konnten Interaktionen der WW- und FF-Domänen zur RNA-Polymerase II [112,116] sowie Wechselwirkungen der WW-Domänen (1 und 2) mit den PRS des Spleißfaktors SF1 [112] nachgewiesen werden. Durch CA150 werden die fundamentalen Prozesse Transkription und Spleißen miteinander verknüpft. Wie gezeigt werden konnte, ist eine bestimmte Isoform von CA150 maßgeblich in den Krankheitsverlauf von Huntington (HD, *Huntington Disease*) involviert (vgl. Abschnitt 1.4). So beeinflusst die Anzahl von (Q-A)-Wiederholungen innerhalb der WW1-WW2-Linkerregion das HD-Ausbruchalter [87]. Zusätzlich konnte gezeigt werden, dass CA150 verstärkt an eine krankheitsspezifisch mutierte Form des Huntingtin-Proteins (HTT) bindet [87], bisher ist der Einfluss dieser Interaktion auf den Krankheitsverlauf jedoch noch unklar. Fest steht, dass eine Überexpression von CA150 in Ratten, die auch das mutierte HTT exprimieren, zur Vermehrung der HD-typischen neuritischen Aggregate führt [108]. Es besteht daher ein großes Interesse, die molekularen Mechanismen CA150-vermittelter Wechselwirkungen besser zu verstehen. Die Vielzahl der bisher identifizierten CA150-Interaktionspartner und das mehrfache Vorhandensein gleicher Funktionseinheiten legen nahe, dass CA150 mit vielen Zielproteinen multivalent wechselwirken könnte. Die Multivalenz könnte zum einen über das Tandem der drei aufeinanderfolgenden WW-Domänen und zum anderen über die beiden PRS realisiert sein. Derartige Interaktionen sind bis zu Beginn dieser Arbeit nur in zwei Fällen [120,121] untersucht worden.

Der Fokus dieser Arbeit liegt daher auf der systematischen Analyse von Interaktionen der isolierten WW-Domänen und Tandem-WW-Konstrukten aus CA150 mit Teilbereichen der PRS aus SF1 sowie Teilbereichen der N-terminalen PRS aus CA150 selbst. Zusätzlich

---

sollten Erkenntnisse zur strukturellen Organisation der drei WW-Domänen und der Linkerregionen zwischen WW1 und WW2 sowie WW2 und WW3 gewonnen werden. Dadurch sollte u.a. geklärt werden, ob die CA150 WW-Domänen selbst miteinander interagieren und dadurch eine Art Autoinhibition stattfindet, wie sie bereits für Su(dx) beschrieben wurde [122]. In diesem Zusammenhang sollte auch geprüft werden, ob die isolierten WW-Domänen bzw. Tandem-WW-Konstrukte artifizielle Oligomere bilden. Zusätzlich sollte untersucht werden, ob die N-terminale CA150<sub>PRS</sub> an die CA150-eigenen WW-Domänen binden kann und dadurch möglicherweise autoinhibitorisch wirkt. Bezüglich der strukturellen Organisation und der Bindungseigenschaften der dritten WW-Domäne von CA150 (WW3) finden sich in der Literatur widersprüchliche Angaben. So berichteten *Otte et al. 2003* [80] von einer ungefalteten WW3, die mit den dort untersuchten Peptiden keinerlei Wechselwirkungen zeigte. Im Gegensatz dazu fanden *Hu et al. 2004* [98] Interaktionen zwischen WW3 und mehreren prolinreichen Motiven, klassifizierten diese jedoch nicht eindeutig. Ein weiteres Ziel dieser Arbeit bestand deshalb darin, über eine Interaktionsanalyse die Bindungspräferenzen der drei WW-Domänen anhand verschiedener prolinreicher Peptide der SF1<sub>PRS</sub> präziser zu definieren.

Um diese Fragenstellungen beantworten zu können, sollten verschiedene WW-Domänen-Konstrukte kloniert und anschließend rekombinant in *E. coli* exprimiert werden. Zur Aufreinigung der Proteine sollten entsprechende Reinigungsstrategien entwickelt werden. Durch Triple-SILAC-basierte Co-Präzipitationsexperimente sollten die WW-Varianten hinsichtlich ihrer Funktionalität getestet werden, indem bekannte Bindepartner aus HeLa-Zelllysaten gefischt und deren Präferenz für Einzel- oder Tandem-WW-Domänen analysiert werden sollten. Durch SPOT-Experimente sollten die Bindepräferenzen von Einzel- und Tandem-WW-Domänen für Peptide der SF1<sub>PRS</sub> mit einem bzw. zwei prolinreichen Motiven untersucht werden. In einer weiteren SPOT-Analyse sollte die Bindung der CA150<sub>WWS</sub> an prolinreiche Peptide der N-terminalen CA150<sub>PRS</sub> getestet werden. Basierend auf den SPOT-Analysen sollten mono- und bivalente Peptide für weitere Interaktionsstudien ausgewählt werden. Mittels Lösungs-NMR-Spektroskopie sollten die Resonanzen des Proteinerückgrates (<sup>1</sup>H<sub>N</sub>, <sup>15</sup>N, C<sub>α</sub>, C<sub>β</sub>, C<sub>o</sub>) einer Tandem-WW-Domäne sequentiell zugeordnet werden. Auf Grundlage dieser Zuordnung sollten über <sup>15</sup>N-Spinrelaxationsexperimente strukturdynamische Eigenschaften der freien und mit einem multivalenten PRS-Peptid assoziierten Tandem-WW ermittelt werden. Durch HSQC-basierte NMR-Bindungsstudien sollten Charakteristika verschiedener PRS-WW-Interaktionen auf atomarer Ebene erfasst werden. Um dabei mono- von bivalenten Interaktionen unterscheiden zu können, sollten Tandem-WW-Konstrukte mit zwei aktiven und Konstrukte mit entweder einer inaktiven WW2 oder WW3-Domäne untersucht werden. Mittels Isothermer Titrationskalorimetrie sollten zusätzlich die thermodynamischen Aspekte mono- und bivalenter PRS-Peptid/WW-Interaktionen erfasst werden. Über die Analytische Ultrazentrifugation sollte der Oligomerisierungszustand eines Tandem-WW-Konstruktes bei Ab- und Anwesenheit eines multivalenten PRS-Peptides untersucht werden.

---

## 3 MATERIAL UND METHODEN

### 3.1 CHEMIKALIEN, MATERIALIEN UND GERÄTE

Soweit nicht anders vermerkt, wurden alle in dieser Arbeit verwendeten Chemikalien von den Firmen Merck (Darmstadt), Sigma-Aldrich (München), Roth (Karlsruhe), Biomol (Hamburg) und VWR International (Darmstadt) in höchst möglicher Reinheit bezogen. Das Säulenmaterial Superdex 75 für die Gelfiltration von Amersham Biosciences (Little Chalfont, GB), das TALON™-Material für die Metallchelate-Affinitätschromatographie von BD Biosciences Clontech (Palo Alto, USA), die gepackten 1ml His GraviTrap™ Nickel-Sepharosesäulen von GE-Healthcare Biosciences AB (Uppsala, Schweden) bezogen. Die <sup>13</sup>C-Glukose, das <sup>15</sup>N-Ammoniumchlorid und das Deuteriumoxid wurden von Euroisotop (Saint-Aubin Cedex, Frankreich), Cambridge Isotopes (Andover, MA, USA), Sigma-Aldrich/ISOTEC (München) geliefert. Für die Amplifikation von DNA wurden Enzyme und Chemikalien der Firmen New England Biolabs (Schwalbach), Novagen®/Merck (Darmstadt) und Invitrogen (Darmstadt) verwendet.

Im Rahmen dieser Arbeit wurden folgende Geräte für die Proteinreinigung und diverse biochemische und biophysikalische Analysen verwendet:

**Tabelle 3.1.1: Geräte**

<b>Inkubatoren:</b>	Multriton INFORS(Leipzig) Innova™ Incubator (New Brunswick Scientific)	
<b>Zentrifugen:</b>	Avanti™ J-25 Centrifuge (Beckman)	Rotor: F10BCI 6x500y(FIBERLite®) J25.50 JLA10.500
	Beckman Coulter Allegra® X-22R Allegra® 21R Eppendorf Centrifuge 5417R	Rotor: SX4250 S4180 Rotor: F45-30-11
<b>Calorimeter:</b>	VP-ITC ITC200	
<b>FPLC:</b>	FPLC-System (Pharmacia Biotech)	
<b>Schwenktisch:</b>	Kisker BioShaker 3D (Steinfurt)	
<b>Spektrometer</b>	BrukerAvance™ III 600 MHz BrukerAvance™ II 750 MHz	

**Tabelle 3.1.II: Spezielle Chemikalien**

Substanz	Hersteller
<sup>15</sup> N-Ammoniumchlorid	Cambridge Isotope Laboratories
Autoinduktionsmedium (Overnight Express™)	Novagen®
Bovines Serumalbumin (BSA)	Sigma-Aldrich®
Bromphenolblau	Roth®
Chloramphenicol	Roth®
Complete® Protease Inhibitor Mix (EDTA-frei)	Roche-Diagnostics®
Coomassie Brilliant Blau G250 & R250	Roth®
D-(+)- <sup>13</sup> C-Glucose	Cambridge Isotope Laboratories
D-(+)-Biotin	Roth®
D-(+)-Glucose, wasserfrei	Roth®
Deuteriumoxid (D <sub>2</sub> O) 99,8%	Eurisotop®
Dithiothreitol (DTT)	Roth®
Ethidiumbromid	Roth®
Isopropyl-β-D-thiogalactopyranosid (IPTG)	Roth®
Kanamycin	Sigma-Aldrich®
Streptomycin	Sigma-Aldrich®
PBS DUBLECCO w/o Ca <sup>2+</sup> , Mg <sup>2+</sup>	BIOCHROM AG
<i>rabbit polyclonal anti-GST Z-5</i>	Santa Cruz, California, USA
<i>HRP-coupled anti-rabbit IgG antibody</i>	Rockland, Gilbertsville, Pennsylvania, USA
„enhanced chemiluminescence substrate“ (SuperSignal West Pico)	Pierce®

**Tabelle 3.1.III: Molekularbiologische und biochemische Kits**

Kit	Verwendung	Hersteller
BCA Protein Assay Kit	Bestimmung von Proteinkonzentrationen	Pierce®
FastDigest® Restriction System	Restriktionsverdau von DNA	Fermentas®
KOD HotStart PCR Kit	Polymerase-Kettenreaktion	Novagen®
Ligation Independent Cloning Kit (LIC)	Ligationsfreie Klonierung von DNA-Sequenzen in LIC-kompatible Vektoren	Novagen®
Montage DNA Gel Extraction Kit	DNA-Extraktion aus Agarosegelen	Millipore®
Nucleospin® Plasmid Preparation Kit	Isolierung von Plasmiden aus <i>E. coli</i>	Macherey-Nagel®

**Tabelle 3.1.IV: Nährmedien und Lösungen**

Medium	Zusammensetzung (für 1 Liter)
LB-Medium	5 g Hefeextrakt 10 g Pepton 10 g NaCl
LB-Agar	5 g Hefeextrakt 10 g Pepton 10 g NaCl 15 g Agar-Agar

Medium	Zusammensetzung (für 1 Liter)
2xM9-Medium	200 ml 10xM9-Salzlösung 20 ml Spurenelemente-Lösung (100x) 2 mM MgSO <sub>4</sub> 300 µM CaCl <sub>2</sub> 3 mg Thiamin-HCl 3 mg Biotin 8 g Glucose 2 g NH <sub>4</sub> Cl
10xM9-Salzlösung	72,8 g Na <sub>2</sub> HPO <sub>4</sub> 20 g KH <sub>2</sub> PO <sub>4</sub> 5 g NaCl
100x Spurenelemente-Lösung	5 g EDTA 500 mg FeSO <sub>4</sub> 50 mg ZnCl <sub>2</sub> 10 mg CuSO <sub>4</sub>

### 3.2 VERWENDETE SOFTWARE

Für die Bearbeitung, Auswertung und graphische Darstellung experimenteller Daten wurden die folgenden Programme, Softwarepakete und webbasierten Anwendungen benutzt.

**Tabelle 3.2.1: Software**

Anwendung	Hersteller / Referenz
CCPN 2.1.5	[123]
Corel Draw 11	Corel Corporation®
CurveFit 1.4	<a href="http://www.palmer.hs.columbia.edu/software/curvefit.html">http://www.palmer.hs.columbia.edu/software/curvefit.html</a> [124] (© Arthur G. Palmer)
DichroWeb	<a href="http://dichroweb.cryst.bbk.ac.uk/html/home.shtml">http://dichroweb.cryst.bbk.ac.uk/html/home.shtml</a> [125,126]
MendeleyDesktop	<a href="http://www.mendeley.com">http://www.mendeley.com</a> , 2009 Mendeley Ltd©
MS Office 2003 / 2007 / 2010	Microsoft® Corporation
Notepad++ 4.9.1	<a href="http://notepad-plus-plus.org">http://notepad-plus-plus.org</a>
Origin® 7.0/8.5	OriginLab®
PrimerX	<a href="http://www.bioinformatics.org/primerx/">http://www.bioinformatics.org/primerx/</a> [127]
ProtParam	<a href="http://web.expasy.org/protparam">http://web.expasy.org/protparam</a> [128]
Pymol	<a href="http://pymol.org">http://pymol.org</a> [129,130]
SEDPHIT	<a href="http://analyticalultracentrifugation.com">http://analyticalultracentrifugation.com</a> [131,132]
Sednterp 1.09	<a href="http://www.jphilo.mailway.com/download.htm">http://www.jphilo.mailway.com/download.htm</a> [133] (University of New Hampshire)
SEDPHAT 9.04	<a href="http://analyticalultracentrifugation.com">http://analyticalultracentrifugation.com</a> [134–140]
SMART	<a href="http://smart.embl-heidelberg.de/">http://smart.embl-heidelberg.de/</a> [75,76]
STRING 9.0	<a href="http://string-db.org/">http://string-db.org/</a> [117]
Talos+	<a href="http://spin.niddk.nih.gov/NMRPipe/talos">http://spin.niddk.nih.gov/NMRPipe/talos</a> [141]
Tensor2	<a href="http://rmni.iqfr.csic.es/HTML-manuals/TENSORV2_DOC/theory.html">http://rmni.iqfr.csic.es/HTML-manuals/TENSORV2_DOC/theory.html</a> [142]
Topspin 2.1	Bruker BioSpin®
UCSF-Chimera	<a href="http://www.cgl.ucsf.edu/chimera">http://www.cgl.ucsf.edu/chimera</a> [143]
Vector-NTI® Advance 11	Invitrogen®



### 3.3 MOLEKULARBIOLOGISCHE ARBEITEN

#### 3.3.1 OLIGONUKLEOTIDE UND VEKTOREN

Die in dieser Arbeit verwendeten Oligonukleotide (siehe Tabelle 3.3.I) wurden von der Firma Biotex (Berlin) bezogen.

**Tabelle 3.3.I: Verwendete Oligonukleotide**

Name	Sequenz
CA150_fw_16	5'-AAC-CTG-TAT-TTC-CAG-GGA-GGG-GAG-CTC-AGG-ATG-GCC-CAA-C-3'
CA150_fw_26	5'-AAC-CTG-TAT-TTC-CAG-GGA-GCC-TTG-AGG-TTC-CGA-GGT-CC-3'
CA150_fw_31	5'-AAC-CTG-TAT-TTC-CAG-GGA-GGT-CCG-GCT-CCC-CCA-CCA-AAT-3'
CA150_fw_40	5'-AAC-CTG-TAT-TTC-CAG-GGA-ATG-CGA-GGC-CCA-CCA-CCT-CTG-3'
CA150_fw_67	5'-AAC-CTG-TAT-TTC-CAG-GGA-GGA-CGT-CCT-CCT-TTT-GAT-CC-3'
CA150_fw_128	5'-AAC-CTG-TAT-TTC-CAG-GGA-GCA-CTA-CCT-CCT-ACG-GAG-G-3'
CA150_fw_331	5'-AAC-CTG-TAT-TTC-CAG-GGA-ACC-GTT-CCC-CAG-CCG-CAC-CC-3'
CA150_fw_408	5'-AAC-CTG-TAT-TTC-CAG-GGA-GCT-GGA-GCA-ACA-GCA-GTT-TCT-G-3'
CA150_fw_506	5'-AAC-CTG-TAT-TTC-CAG-GGA-GCT-CCT-ATT-CCT-GGT-ACT-CC-3'
CA150_rev_167	5'-GAG-GAG-AAG-CCC-GGT-TTA-CTT-AAC-TCC-ATC-TGG-CTT-GGT-CC-3'
CA150_rev_450	5'-GAG-GAG-AAG-CCC-GGT-TTA-CTC-TTC-TAA-CTT-TTC-TTT-TTC-C-3'
CA150_rev_548	5'-GAG-GAG-AAG-CCC-GGT-TTA-GCC-AAT-CAG-ATC-ATC-AGG-TCG-GTC-3'
CA150_rev_559	5'-GAG-GAG-AAG-CCC-GGT-TTA-TTT-TTT-ATG-AGG-GGG-CTC-CTG-3'
sense EK_LIC_TEV	5'-GAC-GAC-GAC-AAG-ATG-GAA-AAC-CTG-TAT-TTC-CAG-3'
WW23_W32A_fw	5'-GAA-CAT-TAG-AAT-CAA-CCG-CGG-AAA-AAC-CCC-AAG-AAC-3'
WW23_W32A_rev	5'-GTT-CTT-GGG-GTT-TTT-CCG-CGG-TTG-ATT-CTA-ATG-TTC-3'
WW23_W131A_fw	5'-CTC-GTC-TTT-CTA-TGG-CGG-ACC-GAC-CTG-ATG-3'
WW23_W131A_rev	5'-CAT-CAG-GTC-GGT-CCG-CCA-TAG-AAA-GAC-GAG-3'
S-Tag-Primer	5'-CGA-ACG-CCA-GCA-CAT-GGA-CA-3'

Die Oligonukleotide WW23\_W32A und WW23\_W131A wurden für die Mutagenese von CA150 WW23 verwendet. Der S-Tag-Primer wurde zur Sequenzierung und der Primer „sense EK\_LIC\_TEV“ zur Einführung der Schnittstelle für die TEV-Protease verwendet (3.3.5). Primer mit der Bezeichnung „fw“ dienen zur Synthese codierender Halbstränge und mit „rev“ bezeichnete Primer zur Synthese komplementärer Halbstränge.

Das Plasmid pOT7 (cDNA Klon: IMAG p958C08335Q/AccNo.: BC014151; BE410347; Gen: TCERG1(Isoform2)) der Firma imaGenes GmbH diente als DNA-Matrize zur Klonierung der verschiedenen WW-Domänen (WW1, WW2, WW3) aus CA150/TCERG1 in die Vektoren pET30 Ek/LIC, pET41 Ek/LIC und pCDF2 EK/LIC (Merck, Novagen®).

Darmstadt). Die verwendeten Vektoren sind in Tabelle 3.3.II und die generierten Plasmide in Tabelle 3.3.III aufgeführt.

**Tabelle 3.3.II: Verwendete Expressionsvektoren**

Vektor	Beschreibung	Referenz
pET30 EK/LIC	Klonierungs- und Expressionsvektor, T7 Promoter, Kan <sup>R</sup> ; N-terminal 6xHis Tag + S-Tag; C-terminal 6xHis Tag	Merck, Novagen®
pET41 EK/LIC	Klonierungs- und Expressionsvektor, T7 Promoter, Kan <sup>R</sup> ; N-terminal GST-Tag; 6xHis Tag + S-Tag; C-terminal 6xHis Tag	Merck, Novagen®
pCDF2 EK/LIC	Klonierungs- und Expressionsvektor, T7 Promoter, Sm <sup>R</sup> ; N-terminal 6xHis Tag; C-terminal S-Tag	Merk, Novagen®

**Tabelle 3.3.III: In dieser Arbeit generierte Plasmide**

Bezeichnung	Insert AS	Domäne	Ausgangs-Vektor	Resistenz
pET30EK/LIC CA150 F1	CA150 16-166	WW1	pET30 EK/LIC	Kan <sup>R</sup>
pET30EK/LIC CA150 F3	CA150 26-166	WW1	pET30 EK/LIC	Kan <sup>R</sup>
pET30EK/LIC CA150 F4	CA150 31-166	WW1	pET30 EK/LIC	Kan <sup>R</sup>
pET30EK/LIC CA150 F5	CA150 68-166	WW1	pET30 EK/LIC	Kan <sup>R</sup>
pET30EK/LIC CA150 F6	CA150 406-546	WW23	pET30 EK/LIC	Kan <sup>R</sup>
pET30EK/LIC CA150 F7	CA150 406-558	WW23	pET30 EK/LIC	Kan <sup>R</sup>
pET30EK/LIC CA150 F10	CA150 16-558	WW123	pET30 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F5	CA150 67-166	WW1	pET41 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F7	CA150 406-558	WW23	pET41 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F8	CA150 16-450	WW12	pET41 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F10	CA150 16-558	WW123	pET41 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F11	CA150 331-558	WW23	pET41 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F12	CA150 506-558	WW3	pET41 EK/LIC	Kan <sup>R</sup>
pET41EK/LIC CA150 F13	CA150 406-450	WW2	pET41 EK/LIC	Kan <sup>R</sup>
pCDF2 EK/LIC CA150 F13	CA150 406-450	WW2	pCDF2 EK/LIC	Sm <sup>R</sup>
pCDF2 EK/LIC CA150 F12	CA150 506-558	WW3	pCDF2 EK/LIC	Sm <sup>R</sup>
pCDF2 EK/LIC CA150 F10	CA150 16-558	WW123	pCDF2 EK/LIC	Sm <sup>R</sup>

### 3.3.2 POLYMERASE-KETTEN-REAKTION (PCR)

Verschiedene DNA-Fragmente wurden mittels Polymerase-Kettenreaktion (PCR, *polymerase chain reaction*) amplifiziert. Zur Anlagerung von Oligonukleotiden wurde ein Temperaturgradient (59-66°C) benutzt. Die Reaktionen wurden in 200 µL Reaktionsgefäßen in einem *Mastercycler® egradient S* (Eppendorf, Hamburg) durchgeführt. PCR-Reaktionen erfolgten entsprechend der dargestellten Vorschrift.

Reaktionsansatz:		PCR-Profil:		
1 µL	DNA Matrize	94°C	5 min	} 10x
5 µL	10xPuffer	94°C	30 sec	
5 µL	2 mM dNTPs	42°C	1 min	
2 µL	25 mM MgSO <sub>4</sub>	72°C	1 min	} 35x
3 µL	Rev-Primer	94°C	30 sec	
3 µL	Fw-Primer	59-66°C	1 min	
1 µL	KOD-Polymerase (1U/µL)	72°C	1 min	
ad 50 µL	H <sub>2</sub> O	72°C	7 min	
		4°C	4	

Die verwendete KOD-*HotSart*-DNA-Polymerase (Merck, Novagen®) aus *Thermococcus kodakaraensis* ist thermostabil und verfügt über eine hohe Lesegenauigkeit.

### 3.3.3 DNA-GELELEKTROPHORESE

Die durch PCR gewonnenen DNA-Fragmente und die geschnittene Plasmid-DNA wurden mittels Gelelektrophorese analysiert. 1,2% bzw. 2% Agarosegele wurden mit 1xTAE-Puffer (*Modified Tris-Acetat EDTA TAE-Buffer*, Millipore®) und 0,5 µg/mL Ethidiumbromid hergestellt. Die DNA-Proben wurden mit 1/5 Vol. 6x-Probenpuffer (Fermentas, St. Leon-Rot) versetzt und bei 110V in *Hoefler HE33 Mini HorizontalSubmarine Units* (GE Healthcare, AmershamBiosciences, Uppsala, Schweden) aufgetrennt. Als Molekulargewichtsstandards wurden GeneRuler™ 1 kb DNA Ladder (#SM0311) und GeneRuler™ 100 bp DNA Ladder (#SM0241) mitgeführt.

### 3.3.4 DNA-GELEXTRAKTION

Die Isolierung der PCR-Fragmente aus dem Agarose-Gel erfolgte mittels Ultrafree®-DA der Firma Amicon® nach Angaben des Herstellers.

### 3.3.5 KLONIERUNG

Klonierungen erfolgten nach dem LIC (*ligation-independent cloning*, (Merck, Novagen®)-System nach Herstellerangaben. Die kodierenden Sequenzen der verschiedenen WW-Domänen des CA150/TCERG1 wurden in die Vektoren pET30 EK/LIC, pET41 EK/LIC und pCDF2 EK/LIC kloniert. Über eine zweistufige PCR wurde unmittelbar vor die zu klonierende Zielsequenz die „ENLYFQG“-Erkennungssequenz für die TEV-Protease (TEV, *tobaccoetchvirus*) eingefügt. Die PCR-Produkte wurden mit dem Reagenziensatz *SureClean Plus* (Bioline, London, UK) gereinigt.

### 3.3.6 HERSTELLUNG CHEMISCH KOMPETENTER E. COLI-ZELLEN

Die Herstellung chemisch kompetenter Zellen erfolgte nach der CaCl<sub>2</sub>-Methode von Sambrook (1989) [144]. Dazu wurden die jeweiligen *E. coli*-Stämme auf einer LB<sub>Agar</sub>-Platte ausgestrichen. Nach 24h Inkubation (37°C) wurde von dieser Agarplatte eine Start-Kultur beimpft, welche bei 37 °C und 180 rpm über Nacht inkubiert wurde. Anschließend wurde die Hauptkultur auf eine OD<sub>600nm</sub> 0,1 inokuliert und bei 37°C und 160 rpm bis zu einer OD<sub>600nm</sub> 0,3 inkubiert. Danach wurde die Kultur für 30 min auf Eis abgekühlt und im

nächsten Schritt bei 4°C, 2500xg für 10 min zentrifugiert. Das Pellet wurde in 0,1 M CaCl<sub>2</sub> (1/10 des Ausgangsvolumens) resuspendiert und 30 min auf Eis inkubiert. Der Ansatz wurde dann erneut zentrifugiert (4°C, 2500xg, 10 min) und die pelletierten Zellen in 0,1 M CaCl<sub>2</sub> + 10% Glycerin (1/20 des Ausgangsvolumens) aufgenommen. Die kompetenten Zellen wurden entweder direkt transformiert oder bei -80°C gelagert.

### 3.3.7 TRANSFORMATION VON *E. COLI*

Zur Transformation von *E. coli* wurden 100 µL kompetente Zellen mit 100-200 ng Plasmid-DNA gemischt und 10 min auf Eis inkubiert. Anschließend erfolgte ein Hitzeschock bei 42°C für 90 Sekunden. Der Ansatz wurde für weitere 5 min auf Eis gestellt, anschließend mit 1 mL LB-Medium versetzt und 45-60 min bei 37°C inkubiert. Etwa 100 µL des Ansatzes wurden auf einer LB-Selektions-Platte (mit entsprechendem Antibiotikum) ausgestrichen und bei 37°C über Nacht inkubiert.

### 3.3.8 PLASMIDPRÄPARATION

Die Präparation von Plasmid-DNA erfolgte mit dem *NucleoSpin<sup>®</sup>Plasmid*-Kit der Firma Macherey & Nagel (Düren, Deutschland) nach Angaben des Herstellers. Zur Kontrolle wurde die Plasmid-DNA auf einem Agarosegel elektrophoretisch aufgetrennt (siehe Abschnitt 3.3.3).

### 3.3.9 VERIFIZIERUNG DER PLASMIDE

Die isolierten Plasmide wurden mittels Restriktionsverdau analysiert. Die Behandlung mit verschiedenen Restriktionsenzymkombinationen (*NdeI*, *EcoRI*, *PstI*, *KpnI*, *BamHI* (FastDigest<sup>™</sup>, Fermentas GmbH)) lieferte erste Hinweise auf den korrekten Einbau der PCR-Produkte. Die fehlerfreie Insertion der Kodierungssequenzen in die pET-Vektoren wurde durch DNA-Sequenzierung kontrolliert. Diese wurde mit den vektorspezifischen Oligonukleotiden T7-Promoter, T7-Terminator (Merck, Novagen<sup>®</sup>) bzw. S-Tag (Biotex) von der Firma Invitex (Berlin) durchgeführt. Die Auswertung der Sequenzierungsergebnisse erfolgte mit der Software VectorNTIAdvance<sup>®</sup> 11 (Invitrogen, Paisley, UK).

### 3.3.10 GERICHTETE MUTAGENESE

Zur Einführung von Mutationen in CA150 WW23 wurde die Zirkuläre PCR-Mutagenese nach *Weiner et al.* 1994 [145] angewendet. Die entsprechenden Oligonukleotide (Tabelle 3.3.I) wurden unter Verwendung der Software PrimerX [127] generiert. Die PCR erfolgte nach der dargestellten Vorschrift.

Reaktionsansatz:	PCR-Profil:				
10x Puffer	5 µL	Sequenz	Zyklen	Temp.	Zeit
dsDNA (50 ng)	x µL	I	1	95°C	30 sec
Rev-Primer (125 ng)	1 µL	II	16	95°C	30 sec
Fw-Primer (125 ng)	1 µL			55°C	1 min
25 mM MgSO <sub>4</sub>	4 µL			68°C	25s/ kb Plasmid
KOD Hot StartPolymerase (1U/µL)	1 µL	III		4°C	4
Ad H <sub>2</sub> O	50 µL				

Die Reaktionsansätze wurden im Anschluss mit dem Restriktionsenzym *DpnI* (4h bei 37°C) verdaut. Anschließend wurden je 10 µL restriktionsverdautes PCR-Produkt in 100 µL CaCl<sub>2</sub>-kompetente DH5α-Zellen transformiert.

### 3.4 BIOTECHNOLOGISCHE METHODEN

#### 3.4.1 ESCHERICHIA COLI STÄMME

Die in dieser Arbeit genutzten *E.coli*-Stämme sind mit ihren jeweiligen genetischen Charakteristika in der nachfolgenden Tabelle gelistet.

**Tabelle 3.4.1: Verwendete Bakterienstämme**

<i>E.coli</i> Stamm	Genotyp	Referenz
GigaCells® (NovaBlueGigaSingles)	endA1 hsdR17(r <sub>K12</sub> <sup>-</sup> m <sub>K12</sub> <sup>+</sup> )supE44 thi-1 recA gyrA96 relA1 lacF' [proA <sup>+</sup> B <sup>+</sup> lacI <sup>q</sup> ZΔM15::Tn10] (Tet <sup>R</sup> )	Novagen®
BL21(DE3)	F <sup>-</sup> omp ThsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) gal dcm (DE3)	Studier&Moffatt 1986 [146]; Novagen®
Rosetta™ 2(DE3)	F <sup>-</sup> ompThsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) gal dcm (DE3) pRARE2(Cam <sup>R</sup> )	Novagen®
DH5α	F <sup>-</sup> φ80lac15 D(lacZY A-argF) U169 recA1 hsdR17(rK-mK+) phoA supE44 λ- thi-1 gyrA96relA1	Invitrogen

#### 3.4.2 ANZUCHTMEDIEN FÜR BAKTERIENKULTUREN

Als Vollmedium wurde LB-Medium bestehend aus 0,5% Hefeextrakt, 1% Pepton und 1% NaCl (pH 7,2-7,5) verwendet. Für die Herstellung von LB-Agar-Platten wurde diesem Ansatz 1,5% Agar-Agar zugesetzt. Zur rekombinanten Herstellung Isotopen-markierter Proteine wurde M9-Minimalmedium [147] mit 2-facher Konzentration M9-Salz verwendet. Eine 10-fache M9-Salz-Stammlösung entsprach 450 mM Na<sub>2</sub>HPO<sub>4</sub>, 150 mM KH<sub>2</sub>PO<sub>4</sub>, 85 mM NaCl (pH 7,2-7,3). Eine 100-fache Spurenelementlösung enthielt 17 mM EDTA, 3 mM FeSO<sub>4</sub>, 300 µM ZnCl<sub>2</sub>, 30 µM CuSO<sub>4</sub> (pH 7,5-7,7). Das M9-Minimalmedium setzte sich wie folgt zusammen.

Zusammensetzung für 1 L Minimalmedium	
100x Spurenelemente	10 mL
1 M MgSO <sub>4</sub>	1 mL
1 M CaCl <sub>2</sub>	0,3 mL
10x M9-Salz	2x100 mL
Glukose bzw. <sup>13</sup> C Glukose)	4 g
Thiamin/HCl 1 mg/mL	1,5 mL
Biotin 0,1 mg/mL	15 mL
NH <sub>4</sub> Cl bzw. <sup>15</sup> N NH <sub>4</sub> Cl	0,5 g
Ad H <sub>2</sub> O	1 L

Zur Selektion wurden den Nährmedien, entsprechend der Resistenzmarker der verwendeten *E. coli* Zellen, Antibiotika in folgenden Endkonzentrationen ( $c_{\text{End}}$ ) zugesetzt: 40  $\mu\text{g/mL}$  Kanamycin, 34  $\mu\text{g/mL}$  Chloramphenicol und 50  $\mu\text{g/mL}$  Streptomycin.

### **3.4.3 GLYCERIN-DAUERKULTUREN**

Dauerkulturen von *E.coli*-Stämmen wurden wie folgt erzeugt. Nachdem die Kulturen in LB-Medium eine Optische Dichte $_{600\text{nm}}$  von 0,7 erreichten. Anschließend wurden 500  $\mu\text{L}$  Kultur mit 500  $\mu\text{L}$  LB-Medium (mit 30% Glycerin) versetzt ( $c_{\text{End}}$ : LB + 15% Glycerin) und bis zur weiteren Verwendung bei  $-80^{\circ}\text{C}$  gelagert.

### **3.4.4 ANZUCHTBEDINGUNGEN**

Übernacht-Vorkulturen wurden von Glycerin-Dauerkulturen oder von frischen Transformationsplatten beimpft. Anschließend wurde ein entsprechendes Kultivierungs-Medium mit der frischen Übernacht-Kultur auf eine Optische Dichte ( $\text{OD}_{600\text{nm}}$ ) von 0,1 inokuliert. Die Zellen wuchsen, soweit nicht anders vermerkt, schüttelnd bei  $37^{\circ}\text{C}$ . Das Wachstum wurde durch die Messung der Optischen Dichte ( $\text{OD}_{600\text{nm}}$ ) an einem Spektralphotometer (Utraspec 2000, Pharmacia) bei 600 nm verfolgt.

### **3.4.5 KONTROLLE DER PROTEINEXPRESSION**

Für die Überexpression von Zielproteinen wurden die Zellen in LB- oder M9-Medium bei  $37^{\circ}\text{C}$  unter Schütteln bei 110 rpm kultiviert. Bei Erreichen einer  $\text{OD}_{600\text{nm}}$  von 0,8-1 wurde die Temperatur auf 20 oder  $25^{\circ}\text{C}$  gesenkt und anschließend mit 1 mM IPTG induziert. Die Zellernte erfolgte ca. 17-20 h nach dem Induktionszeitpunkt durch Zentrifugation.

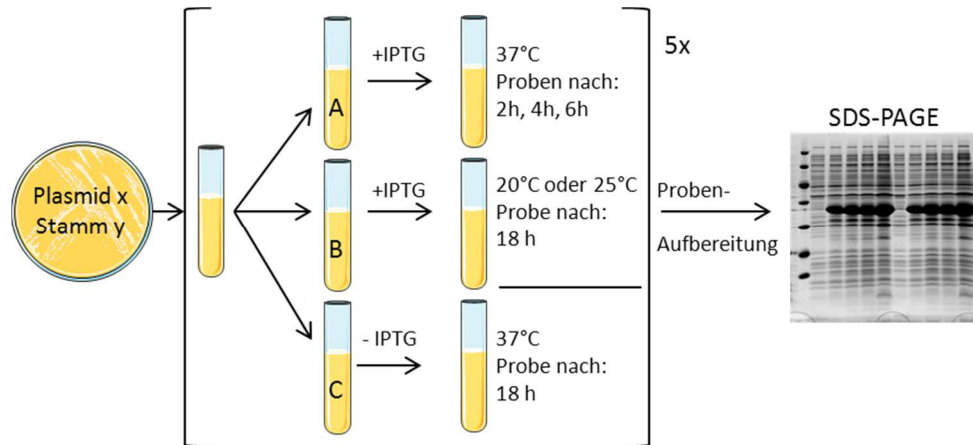
Die Kontrolle der Proteinexpression erfolgte über SDS-PAGE (3.6.1). Dazu wurden entsprechend 1  $\text{OD}_{600\text{nm}}$  Zellen entnommen und 1 min bei 14.000 rpm pelletiert. Das Pellet wurde anschließend mit 25  $\mu\text{L}$  Sarkosyllyse-Puffer (1% N-Lauroylsarcosin, 1 mM NaOH, 2x Probenpuffer (3.6.1)) versetzt und 15 min bei  $96^{\circ}\text{C}$  inkubiert. Nach Zugabe von 75  $\mu\text{L}$  1x Probenpuffer wurden 10  $\mu\text{L}$  dieses Ansatzes auf eine SDS-PAGE aufgetragen.

Zur Abschätzung der Löslichkeit exprimierter Zielproteine wurden die Proteinpellets alternativ mit BugBuster® (MERCK) aufgeschlossen. Hierbei wurden sowohl Protein-Proben vor dem Zentrifugationsschritt (5 min, 20.800xg,  $4^{\circ}\text{C}$ ) als nach diesem (vom löslichen Proteinüberstand) genommen. Die Proben wurden anschließend erhitzt (je 10  $\mu\text{L}$  Probe + 10  $\mu\text{L}$  2x Probenpuffer, Inkubation: 6 min,  $96^{\circ}\text{C}$ ) und 10  $\mu\text{L}$  zur Analyse auf ein SDS-Gel aufgetragen.

### **3.4.6 ERMITTLUNG OPTIMALER EXPRESSIONSBEDINGUNGEN**

Zur Ermittlung optimaler Expressionsbedingungen wurden, ausgehend von einer frischen Transformationsplatte, je 5 Klone pro Plasmid und Bakterienstamm in 4 mL LB-Medium angezogen. Von diesen Vorkulturen wurden am nächsten Tag je 3x 4 mL LB-Medium auf 0,1  $\text{OD}_{600}$  inokuliert (Kulturen A, B und C in Abbildung 13). Bei Erreichen von 0,7-1  $\text{OD}_{600}$  wurden die Kulturen (A und B) mit 1 mM IPTG induziert. Kultur C wurde nicht induziert

und diente als Referenz. Kultur A wurde nach der Induktion bei 37°C und Kultur B bei 20°C (1. Testreihe) oder 25°C (2. Testreihe) inkubiert. Nach 2h, 4h, 6h und 18h Expression erfolgte jeweils eine Probenentnahme. Von Kultur C wurde nur nach 18h Proben entnommen (nicht induzierte Referenz).



**Abbildung 13: Schema der Expressionsuntersuchungen**

Alle Proben wurden per SDS-PAGE analysiert (siehe Abschnitt 3.4.5). In einem weiteren Test wurde die IPTG-Konzentration variiert (0.5-1 mM). Als Expressionswirte wurden die Stämme BL21(DE3) und Rosetta™ 2(DE3) getestet.

### 3.4.7 EXPRESSION ISOTOPENMARKIERTER WW-DOMÄNEN

Die Herstellung uniform  $^{15}\text{N}$  und uniform  $^{13}\text{C}$ ,  $^{15}\text{N}$ -markierter WW-Domänen erfolgte in 2 L-Kulturflaschen á 400 mL Medium. Zunächst wurde eine 50 mL LB-Vorkultur mit 1 mL einer LB-Übertag-Kultur beimpft und über Nacht bei 25°C inkubiert. Die Kultur wurde anschließend 10 min bei 1600xg und 20°C zentrifugiert und das erhaltene Zellpellet in 400 mL isotonenmarkiertem M9-Medium (3.4.2) resuspendiert. Bis zum Erreichen einer  $\text{OD}_{600\text{nm}}$  0,7-1 wurde die Kultur bei 37°C und 110 rpm inkubiert. Danach wurde die Kultur auf 25°C abgekühlt und durch Zugabe von IPTG ( $c_{\text{End}} = 1 \text{ mM}$ ) die Expression des Zielproteins eingeleitet. Die Kulturen wurden 16-20h nach Induktion durch Zentrifugation geerntet (3.5.1.).

## 3.5 REINIGUNG REKOMBINANTER PROTEINE

### 3.5.1 ZELLAUFSCHLUSS

Zum Aufschluss der Zellen wurde das Zellpellet in Puffer A (20 mM  $\text{NaPO}_4$  (pH 7.5), 150 mM NaCl, 1 mM  $\text{MgCl}_2$ , 0,2  $\mu\text{L/mL}$  Benzonase (Novagen), Protease-Inhibitor-Mix Complete® (ohne EDTA)) resuspendiert (10 ml Puffer pro Gramm Biofeuchtmasse). Die Zellsuspension wurde am Zellaufschlussgerät (EmulsiFlex-C3, Avestin Europe GmbH, Mannheim) bei 1500 bar für 5 min aufgeschlossen. In einem anschließenden Zentrifugationsschritt (30 min, 58.000xg, 4°C) wurden die unlöslichen Bestandteile

pelletiert. Der lösliche Proteinüberstand wurde direkt für die weitere Proteinreinigung verwendet oder in flüssigem Stickstoff schockgefroren und bei -80 °C gelagert.

### **3.5.2 PROTEINREINIGUNG ÜBER METALL-CHELAT-CHROMATOGRAPHIE**

Für die Reinigung der CA150-WW-Domänen wurde das lösliche Zelllysate nach Filtration (0,45 µM) einer Metallchelate-Chromatographie unterzogen. Hierzu wurde entweder die Chromatographieranlage Profinia der Firma Bio-Rad oder Gravitationsssäulen (MoBiTec GmbH, Göttingen) verwendet. Es wurde Nickel(II)-Ionen oder Kobalt(II)-Ionen (TALON™) haltiges Säulenmaterial genutzt, welches vor der Chromatographie nach Angaben des Herstellers gereinigt und mit den entsprechenden Metall-Ionen beladen wurde. Vor der Reinigung der WW-Domänen wurde das Säulenmaterial zunächst mit PufferB (20 mM Na-PO<sub>4</sub> (pH 7.5), 150-300 mM NaCl) equilibriert. Bei der Verwendung von Gravitationsssäulen wurde die Proteinlösung zusammen mit dem Säulenmaterial im Rotationsschüttler inkubiert (1 h bei 10°C). Anschließend wurde das Säulenmaterial in die Gravitationsssäule überführt. Zur Entfernung unspezifisch gebundener Proteine erfolgte ein Waschschriff mit 15 Säulenvolumen (SV) PufferB. In zwei weiteren Waschschriffen á 15xSV wurden dem PufferB 5 mM bzw. 10 mM Imidazol zugesetzt. Die Elution erfolgte mit 250 mM Imidazol in PufferB. Das Eluat wurde in 2 mL Fraktionen gesammelt und mittels SDS-PAGE analysiert.

### **3.5.3 ABSPALTUNG DES REINIGUNGS-TAGS**

Die proteolytische Abspaltung des Reinigungs-Tags erfolgte während einer Dialyse gegen ein mindestens 100-faches Volumen Puffer B (+1 mM β-Mercaptoethanol) bei 22°C über Nacht (Dialyseschlauch: Spectra/Por®, MWCO: 3,5 kDa, 6-8 kDa, 12-14 kDa). Dazu wurden die ausgewählten Elutionsfraktionen der 1. Reinigung (siehe Abschnitt 3.5.2) mit TEV-Proteaselösung (1 OD<sub>280</sub> TEV-Protease auf 100 OD<sub>280</sub> Fusionsprotein) versetzt. Im Anschluss erfolgte eine weitere Metallchelate-Chromatographie (3.5.2) wobei hier der 1. Durchlauf aufgefangen und zur weiteren Reinigung mittels Gelfiltration aufbewahrt wurde. Die verwendete TEV-Protease wurde in der AG Oschkinat rekombinant hergestellt.

### **3.5.4 PROTEINREINIGUNG DURCH GELFILTRATION**

Für die Gelfiltration wurde eine 120 mL Säule verwendet. Als Säulenmaterial diente Superdex75. Der Lauf erfolgte mit einer Geschwindigkeit von 1 mL/min. Bei Bedarf wurde bei diesem Schritt das Puffersystem (ITC-Puffer (3.9), NMR-Puffer (3.7)).

### **3.5.5 KONZENTRATION VON PROTEINEN**

Die Konzentration der gereinigten Proteine erfolgte in einer Tisch-Zentrifuge mittels Amicon® Ultra Centrifugal Filter Devices (MILLIPORE, MWCO: 3,5 kDa, 10 kDa, 30 kDa) nach Angaben des Herstellers.



### 3.6 BIOCHEMISCHE ANALYSEMETHODEN

#### 3.6.1 NATRIUMDODECYLSULFAT-POLYACRYLAMIDGELELEKTROPHORESE

Proteine wurden mittels Natriumdodecylsulfat-Polyacrylamidgelelektrophorese (SDS-PAGE) nach *Laemmli* [148] oder nach *Schägger & von Jagow* [149] aufgetrennt. Für die Elektrophorese wurden Hoefer SE250 *Mini-Vertical Gelelectrophoresis Units* (GE Healthcare, Amersham Biosciences) oder das *Dual Triple Wide Mini-Vertical Electrophoresis System* (C.B.S. Scientific Company) verwendet.

Die 15% bzw. 10%igen SDS-Polyacrylamidgele setzten sich aus Trenngel (15% bzw. 10% Acrylamid/Bisacrylamid; 370 mM Tris-HCl (pH 8,8); 0,1% SDS; 0,07% APS; 0,05% TEMED) und Sammelgel (5% Acrylamid/Bisacrylamid; 80 mM Tris-HCl (pH 6,8); 0,1% SDS; 0,05% APS; 0,1% TEMED) zusammen. Die Elektrophorese nach *Laemmli* [148] erfolgte in 25 mmol/L Tris; 190 mM Glycin; 0,1% SDS. Proteinproben wurden 1:1 mit 2x-Probenpuffer (125 mM Tris-HCl (pH 6,8); 4% SDS; 20% Glycerin; 0,01% Bromphenolblau; 5% 2-Mercaptoethanol) versetzt und 6-10 min bei 96 °C inkubiert. Folgende Molekulargewichtsstandards wurden entsprechend der Proteinprobe mitgeführt: *Spectra<sup>TM</sup> Multicolor Low Range Protein Ladder* #SM1861 (Fermentas), *PageRuler<sup>TM</sup> Plus Prestained Protein Ladder* #SM1811 (Fermentas) und *Unstained Protein Molekular Weight Marker* #SM0431 (Fermentas). Die Färbung der SDS-Polyacrylamidgele erfolgte nach *Schägger & von Jagow* [149]. Dazu wurden die Gele 10 min in 10% Essigsäure und 50% Ethanol fixiert, 30 min mit 0,025% *Coomassie Brilliant Blue G250* und 0,025% *Coomassie Brilliant Blue R250* (Carl Roth) in 10% Essigsäure gefärbt und abschließend in 10% Essigsäure entfärbt.

#### 3.6.2 BERECHNUNG VON PROTEINKONZENTRATIONEN

Die Konzentration der gereinigten Proteine wurde durch Absorptionsmessung bei einer Wellenlänge von 280 nm am UV/VIS-Spektralphotometer *NanoDrop2000c* (ThermoFisher Scientific, Waltham, MA, USA) bestimmt. Hierbei kam das Lambert-Beer'sche Gesetz zur Anwendung.

$$A = c * d * \varepsilon$$

Dabei beschreibt  $A$  die Absorption bei 280 nm,  $c$  die zu bestimmende Proteinkonzentration [M],  $d$  die Schichtdicke der Küvette [cm] und  $\varepsilon$  den molaren Extinktionskoeffizienten bei 280 nm. Letzterer wurde entsprechend der Primärsequenz der Proteine, mit dem Programm *ProtParam* (<http://expasy.org/tools/protparam.html>) ermittelt.

**Tabelle 3.6.I: Molekulargewichte (MW) und molare Extinktionskoeffizienten ( $\epsilon_{280}$ ) der rekombinant erzeugten WW-Domänen-Fusionsproteine**

WW-Domäne	Bezeichnung	codiert auf	kDa	$\epsilon_{280}$ Abs 0.1%
WW1 (16-166)	WW1 S His	pET30EK/LIC CA150 F1	21,1	0,733
WW1 (26-166)	WW1 S His	pET30EK/LIC CA150 F3	19,4	0,797
WW1 (31-166)	WW1 S His	pET30EK/LIC CA150 F4	18,6	0,831
WW1 (67-166)	WW1 S His	pET30EK/LIC CA150 F5	16,5	1,028
	WW1 GST	pET41EK/LIC CA150 F5	43,2	1,391
WW2 (406-450)	WW2 GST	pET41EK/LIC CA150 F13	37,8	1,630
	WW2 His	pCDF2 F13	8,0	2,303
WW3 (506-558)	WW3 GST	pET41EK/LIC CA150 F12	38,6	1,621
	WW3 His	pCDF2 F12	8,9	2,198
WW123 (16-558)	WW123 S His	pET30EK/LIC CA150 F10	62,6	0,892
	WW123 GST	pET41EK/LIC CA150 F10	88,7	1,071
	WW123 His	pCDF2 F10	58,9	0,881
WW12 (16-450)	WW12 GST	pET41EK/LIC CA150 F8	66,0	1,168
	WW12 S His	pET30EK/LIC CA150 F8	39,3	0,863
WW23 (406-548)	WW23 S His	pET30EK/LIC CA150 F6	20,9	1,676
WW23 (406-558)	WW23 S His	pET30EK/LIC CA150 F7	23,6	1,545
	WW23 GST	pET41EK/LIC CA150 F7	50,3	1,583
WW23 (406-558) W32A	WW23_W32A S His	pET30EK/LIC CA150 F7 W32A	23,5	1,318
WW23 (406-558) W131A	WW23_W131A S His	pET30EK/LIC CA150 F7 W131A	23,5	1,318

GST = Glutathion-S-Transferase-Tag, S His = S-Tag und Hexahistidin-Tag, His = Hexahistidin-Tag

**Tabelle 3.6.II: Molekulargewichte und molare Extinktionskoeffizienten ( $\epsilon_{280}$ ) rekombinant erzeugter WW-Domänen**

WW-Domäne	AS in CA150	kDa	$\epsilon_{280}$
WW1	68-166	10,8	1,434
WW2	406-450	5,4	3,156
WW3	506-558	6,2	2,891
WW23	406-558	17,9	1,955
WW23_W32A	406-558, W436A	17,8	1,658
WW23_W131A	406-558, W535A	17,8	1,658

**Tabelle 3.6.III: Sequenzen rekombinant und synthetisch hergestellter CA150<sub>WW</sub>-Varianten**

Konstrukt	Sequenz
WW1	GGRPPFDPNMPPMPPPGGIPPPMGPPLHQRPPFMPPPMSSMPPPPGMMFPPGMPPVTAPGT PALPPTTEIIVVENKTPDGKVVYYNARTRESAWTKPDGVK
WW2	GAGATAVSEWTEYKTADGKTYYYNNRTLESTWEKPQELKEKEKLEE
WW3	GAPIPGTPWCVVWTGDERVFFYNPTTRLSMWDRPDDLIGRADVDKIIQEPHKK
WW23	GAGATAVSEWTEYKTADGKTYYYNNRTLESTWEKPQELKEKEKLEEKIKEPIKEPSEEPLMET EEEDPKKEEPIKEEKEEEMTEEEKAAQKAKPVATAPIPGTPWCVVWTGDERVFFYNPTTRL SMWDRPDDLIGRADVDKIIQEPHKK
WW23_W32A	GAGATAVSEWTEYKTADGKTYYYNNRTLESTAEKPQELKEKEKLEEKIKEPIKEPSEEPLMET EEEDPKKEEPIKEEKEEEMTEEEKAAQKAKPVATAPIPGTPWCVVWTGDERVFFYNPTTRL SMWDRPDDLIGRADVDKIIQEPHKK

Konstrukt	Sequenz
WW23_W131A	GAGATAVSEWTEYKTADGKTYYYNNRTLESTWEKPQELKEKEKLEEKIKEPIKEPSEEPLMET EEEDPKEEPIKEIKEEPKEEEMTEEEKAAQKAKPVATAPIPGTPWCVVWTGDERVFFYNPTTRL SMADRPDDLIGRADVDKIIQEPHKK
WW2 <sub>synthetisch</sub>	AC-GATAVSEWTEYKTADGKTYYYNNRTLESTWEKPQELKEKEKLEE-NH <sub>2</sub>
WW3 <sub>synthetisch</sub>	AC-APIPGTPWCVVWTGDERVFFYNPTTRLSMWDPRDDLIGRADVDKIIQEPHKK-NH <sub>2</sub>

AC: Acetyl-Gruppe; Die Bereiche der WW-Domänen sind Fett dargestellt. Rot hervorgehobene Glycine sind nicht Teil der WW-Sequenzen, sondern resultieren aus dem Einbau der TEV-Schnittstelle (siehe Abschnitt 3.3.5)

Die synthetischen WW-Domänen WW2<sub>synthetisch</sub> und WW3<sub>synthetisch</sub> wurden mittels Festphasensynthese in der Arbeitsgruppe Beyermann (FMP) hergestellt.

### 3.6.3 PROLINREICHE PEPTIDLIGANDEN FÜR NMR-TITRATIONEN UND ITC-EXPERIMENTE

Die in verschiedenen Bindungsstudien (siehe Abschnitt 3.7.2 und 3.9) verwendeten prolinreichen Peptidliganden (siehe Tabelle 3.6.IV) wurden mittels Festphasensynthese in der Arbeitsgruppe Beyermann (FMP) hergestellt. Alle Peptide enthielten N-Terminal ein freies Amin und C-Terminal eine Amidgruppe. Bei der Synthese wurde die Fmoc-Chemie angewendet. Die Peptide wurden anschließend mittels Hochleistungsflüssigkeitschromatographie (HPLC *high pressure liquid chromatography*) gereinigt. Die Endprodukte besaßen einen Reinheitsgrad von 95-98%. Anschließend wurden die Produkte in der Arbeitsgruppe von Dr. Eberhard Krause (FMP) mittels Massenspektrometrie auf ihr korrektes Molekulargewicht hin untersucht.

**Tabelle 3.6.IV: Prolinreiche Peptidliganden**

Bezeichnung	Sequenz	AS	kDa
SF1_P1	QPPPPPMNQGPHPGHHGPP-NH <sub>2</sub>	426-445	2,070
SF1_P2	SGVYRLHQGKGMPPPPMGM-NH <sub>2</sub>	458-477	2,169
SF1_P3	MPPPPPPPSGQPPPPSGPL-NH <sub>2</sub>	478-497	1,939
SF1_P4	HPPPPPSHYGPLPPWHQK-NH <sub>2</sub>	n.d.	2,067
CA_P1	HNAVMRGPPLMRPPPPFGMMRGPppppRPPFG-NH <sub>2</sub>	37-68	3,694
SF1_KP1	VQPPLPPGA-NH <sub>2</sub>	573-581	0,875
CA_KP2	RGPPPPPR-NH <sub>2</sub>	57-64	0,873

Peptide deren Sequenz aus SF1 stammt beginnen mit SF1\_ (ausgenommen SF1\_P4), die aus CA150 mit CA\_ (die entsprechenden Sequenzpositionen sind unter „AS“ aufgeführt. Das Histidin (Rot) im CA\_P1 ist nicht in der nativen CA150 Sequenz enthalten.

### 3.6.4 SPOT-ANALYSEN

#### 3.6.4.1 Herstellung der Zellulosemembran-gebundenen Peptidliganden

Die SPOT-Membranen wurden von *Angelika Ehrlich* (Arbeitsgruppe Beyermann (FMP)) hergestellt. Die Peptidsequenzen (Tabelle 3.6.V) wurden dabei mittels *Auto-Spot Robot* (ASP 222: Invartis AG, Köln, Germany) unter Verwendung von Standardprotokollen der Fmoc-Chemie [150] auf eine Whatman 50 Zellulose-Membran synthetisiert.

**Tabelle 3.6.V: Sequenzen der SF1-SPOT-Peptide**

	Sequenz	SF1_Position
1	TTQSRPPWMNSGPSESRPYH	361-380
2	NGPPPPWMQPPPPPMNQGPH	418-437
3	QPPPPPMNQGPHPPGHGPP	426-445
4	QGPHPGHHGPPMDQYLGS	434-453
5	SGVYRLHQGKGMPPPPMGM	458-477
6	MMPPPPMGMMPPPPPPSGQ	469-488
7	MPPPPPPSGQPPPPPSGPL	478-497
8	QPPPPPSGPLPPWQQQQQP	488-507
9	LPPWQQQQQPPPPPPSSS	497-516
10	WQQQQQPPPPPPSSSMAS	500-519
11	PPPPPPSSSMASSTPLPWQ	507-526
12	IPPWQQQAAAAASPGAPQM	541-560
13	QQQAAAAASPGAPQMGGNP	545-564
14	MVPLPPGVQPPLPPGAPPPP	566-585
15	QPPLPPGAPPPPPPPGSA	574-593
16	APPPPPPPPGSAGMMYAPP	581-600
17	APPPPPPPMDPSNFVTMMG	598-617
18	VAGMPPFGMPPAPPPPPQN	620-639

**Tabelle 3.6.VI: Sequenzen der CA150<sub>PRS</sub>-SPOT-Peptide**

	Sequenz	CA150_Position
1	FRGPAPPNAVMRGPPPLMR	29-48
2	VMRGPPPLMRPPPPFGMMRG	39-58
3	LMRPPPPFGMMRGPPPPRP	46-65
4	GPPPPRPPFGRPPFDPNMP	58-77
5	PFGRPPFDPNMPPPPGGI	66-85
6	PPMPPGGIPPMGPPHLQR	77-96
7	PPPMGPPHLQRPPFMPPMS	68-105
8	RPPFMPPMSMPPPPGMMF	96-115
9	PPMSSMPPPPGMMFPPGMPP	102-121

### 3.6.4.2 Durchführung von SPOT-Analysen

Spot-Analysen wurden ausschließlich mit GST-fusionierten WW-Domänen durchgeführt. Die SPOT-Membran wurde zunächst in mehrere Streifen mit analogen Peptidmustern geteilt und jeweils mit einer der zu untersuchenden WW-Varianten inkubiert. Alle Waschschriffe, sowie die Inkubationsschritte mit Protein- und Antikörperlösung wurden auf einem Schwenktisch in *Western-Blot-Immuno-Schalen* (ROTH) durchgeführt. Zur Entfernung eventueller Synthesereste wurden die Membranstreifen mit 1 mL 100% Ethanol gewaschen. Es folgten drei Waschschriffe á 10 min mit 1 mL PBS-Puffer (PBS DUBLECCO w/o  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$  (BIOCHROM AG) (pH 7,5)). Im Anschluss wurde die Membran über Nacht bei 10°C mit PBS-Puffer + 5% Magermilch inkubiert. Es folgte ein weiterer Waschschriff mit PBS (10 min). Die Membranstreifen wurden mit jeweils 1  $\mu\text{M}$  GST-WW-Fusionsprotein gelöst in PBS und 2% Magermilch für 1h bei 25°C inkubiert. Anschließend wurde die Membran 3x mit 1 mL PBS gewaschen und danach mit dem primären Antikörper (*rabbit polyclonal anti-GST* (Z-5, Santa Cruz, California, USA)) in einer Verdünnung von 1:1.000 für 1h bei 25°C inkubiert. Nach drei weiteren PBS-Waschschriffen folgte ein Inkubationsschriff (1h, 25°C) mit einem 1:10.000 verdünnten sekundären, HRP-gekoppelten (HRP: *horseradishperoxidase*) Antikörper (*HRP-coupled anti-rabbit IgG antibody* (Rockland, Gilbertsville, Pennsylvania, USA)). Nach zwei abschließenden Waschschriffen konnte durch Zugabe von jeweils 500  $\mu\text{L}$  „*enhanced chemiluminescence substrate*“ (SuperSignal West Pico, Pierce Illinois, USA) die Detektion und Auswertung am Lumilmager™ (Diagnostics, Mannheim, Germany) erfolgen. Alle Versuche zur SF1<sub>PRS</sub>/CA150<sub>WW</sub>- bzw. zur CA150<sub>PRS</sub>/CA150<sub>WW</sub>-Interaktion wurden 3-mal wiederholt. Als Kontrolle wurde jeweils ein Streifen einer SPOT-Membran mit reinem GST inkubiert.

## 3.7 LÖSUNGS-NMR SPEKTROSKOPIE

Die Lösungs-NMR Experimente wurden an einem Bruker AVANCE® 600 MHz Spektrometer (Magnetfeldstärke: 14,1 Tesla) oder an einem Bruker AVANCE® 750 MHz Spektrometer (Magnetfeldstärke: 17,6 Tesla) unter Verwendung von „5 mm Triple-Resonanz PFG (Z-Achse) Kryoprobenköpfen“ durchgeführt. Für die Messungen wurden 5 mm NMR-Röhrchen (NORELL®, Landisville, NJ, USA), befüllt mit 550-600  $\mu\text{L}$  Probe, verwendet. Triple-Resonanz-Experimente für die Zuordnung des Proteinrückgrads wurden bei 14°C, Titrations- und  $^{15}\text{N}$ -Spinrelaxationsexperimente bei 7-10°C durchgeführt. In allen NMR-Experimenten wurde ein NMR-Puffer bestehend aus 20 mM  $\text{NaPO}_4$  (pH 7.0), 50 mM NaCl und 10%  $\text{D}_2\text{O}$  eingesetzt (Ausnahme: Probe für die Sequenzielle Zuordnung pH 6.5). Experimente für die. Das  $\text{H}_2\text{O}$ -Signal wurde mittels „Watergate-Pulssequenz“ [151] runterdrückt. Die Prozessierung von NMR-Spektren, und die Erstellung einiger Konturplots erfolgten mit „TopSpin® 2.1“ (Bruker®). Die Ermittlung von Signalintensitäten, die Erfassung chemischer Verschiebungen, die Erstellung vieler Konturplots sowie die Zuordnung von Resonanzen in 2D- & 3D-Spektren, wurden mit dem Softwarepaket

„CCPNMR 2.1“ [123] durchgeführt. Die Messungen wurden von *Dr. Peter Schmieder, Monika Beerbaum* und *Matthias Dorn* begleitet.

### **3.7.1 AUFNAHME MEHRDIMENSIONALER, HETERONUKLEARER NMR-SPEKTREN ZUR SEQUENTIELLEN ZUORDNUNG DER RESONANZEN DES PROTEINRÜCKGRATES**

Für die Zuordnung der chemischen Verschiebungen der Atome  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}$ ,  $^{13}\text{CO}$ ,  $^{13}\text{C}\alpha$  und  $^{13}\text{C}\beta$  innerhalb des Proteinrückgrates von CA150-WW23 wurden 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC-Spektren (HSQC = Heteronuclear Single Quantum Correlation) [152] sowie HSQC-basierte Versionen der dreidimensionalen, heteronuklearen Triple-Resonanz-Spektren HNCA, HN(CO)CA, HNCACB, CBCA(CO)NNH, HN(CA)CO und HNCO [153–155] aufgenommen. Dazu wurde uniform  $^{13}\text{C}^{15}\text{N}$ -markiertes CA150-WW23 verwendet. Zur sequentiellen Zuordnung von Tryptophan-Seitenketten ( $^1\text{H}\epsilon$ -,  $^{15}\text{N}\epsilon$ -Resonanzen) wurden ein  $^{15}\text{N}$ -editiertes NOE-Spektrum (HSQC-NOESY) von uniform  $^{15}\text{N}$ -markiertem CA150-WW23 aufgenommen. In Tabelle 3.7.I sind die durchgeführten NMR-Experimente zusammengefasst.

**Tabelle 3.7.I: Übersicht der zur Zuordnung der Proteinrückgrat-Resonanzen verwendeten NMR-Spektren von CA150 WW23.**

Spektrum	Art	Ausgewertete Resonanzen
HSQC	2D	$\text{NH}_{(i)}$
HSQC-NOESY	2D	$\text{N}\epsilon$ , $\text{H}\epsilon$ von Tryptophan-Seitenketten
HNCA	3D	$\text{C}\alpha_{(i-1)}$ und $\text{C}\alpha_{(i)}$
HN(CO)CA	3D	$\text{C}\alpha_{(i-1)}$
HNCACB	3D	$\text{C}\alpha_{(i-1)}$ und $\text{C}\beta_{(i-1)}$ , $\text{C}\alpha_{(i)}$ und $\text{C}\beta_{(i)}$
CBCA(CO)NNH	3D	$\text{C}\alpha_{(i-1)}$ und $\text{C}\beta_{(i-1)}$
HNCO	3D	$\text{CO}_{(i-1)}$
HN(CA)CO	3D	$\text{CO}_{(i-1)}$ und $\text{CO}_{(i)}$

*i* bzw. *i-1* stehen hierbei für die Resonanzen innerhalb einer beliebigen Aminosäure (*i*) und der dieser sequentiell vorangehenden Aminosäure (*i-1*).

### **3.7.2 ANALYSE DER INTERAKTION ZWISCHEN CA150-WW-DOMÄNEN UND PROLINREICHEN PEPTIDLIGANDEN**

Zur Identifizierung der Aminosäuren in CA150-WW-Domänen, welche die Interaktion mit verschiedenen prolinreichen Peptidliganden vermitteln, wurden  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von uniform  $^{15}\text{N}$ -markiertem CA150-WW23, CA150-WW23\_W32A sowie CA150-WW23\_W131A in An- und Abwesenheit entsprechender Peptidliganden (siehe Tabelle 3.6.IV, außer SF1\_P2) aufgenommen. Es wurden Peptid-Protein-Konzentrationsverhältnisse von 0:1, 2:1, 5:1 und 10:1 untersucht, wobei die Proteinkonzentration innerhalb einer Messreihe konstant gehalten wurde. Vor jeder Messreihe wurden die jeweilige WW-Variante und der zu testende Peptidligand separat gegen NMR-Puffer

dialysiert und die Proben anschließend mit 10% D<sub>2</sub>O versetzt. Alle Spektren wurden in TopSpin2.1 mit identischen Parametern prozessiert und mittels CCPN 2.1.5 auf Veränderungen der chemischen Verschiebung sequentiell zugeordneter <sup>1</sup>H<sub>N</sub><sup>15</sup>N-Kreuzsignale analysiert. Die konzentrationsabhängige Veränderung der <sup>1</sup>H<sub>N</sub>- und <sup>15</sup>N-chemischen Verschiebung eines Kreuzsignals wurde nach folgender Gleichung in einen gewichteten NH-Vektorbetrag ( $\Delta\sigma$ ) umgerechnet:

$$\Delta\sigma_{HN} = \sqrt{(\Delta\sigma^{1H_N})^2 + (0,15 * \Delta\sigma^{15N})^2}$$

$\Delta\sigma^{1H_N}$  und  $\Delta\sigma^{15N}$  sind die peptidinduzierten Unterschiede der <sup>1</sup>H<sub>N</sub> und <sup>15</sup>N chemischen Verschiebungen eines entsprechenden HSQC-Kreuzsignals des untersuchten Proteins.

### 3.7.3 MESSUNG <sup>15</sup>N-BASIERTER SPINRELAXATION (<sup>15</sup>N-R1, <sup>15</sup>N-R2 & <sup>1</sup>H<sup>15</sup>N-HETNOE)

Zur Erfassung moleküldynamischer Parameter wurden die longitudinale (R1) und transversale (R2) Relaxationsrate sowie die Verstärkung der heteronuklearen <sup>1</sup>H<sup>15</sup>N-NOE's (<sup>1</sup>H<sup>15</sup>N-HetNOE; NOE = *Nuclear Overhauser Effect*) der Amidstickstoffkerne von uniform <sup>15</sup>N-markiertem CA150-WW23 gemessen [156,157]. Die dazu durchgeführten Experimente basierten auf <sup>1</sup>H<sup>15</sup>N-korrelierten pseudo-3D HSQC-Spektren [152]. <sup>15</sup>N-R1-Experimente wurden mit Relaxationsperioden von 13,5; 53,5; 103,5; 153,5; 203,5; 303,5; 403,5; 603,5; 903,5; 2003,5 und 5003,5 ms und <sup>15</sup>N-R2-Experimente mit Relaxationsperioden von 4; 8; 16; 24; 32; 40; 80; 120; 160; 200 und 240 ms aufgenommen. Die Intensitätswerte korrespondierender <sup>1</sup>H<sup>15</sup>N-Kreuzsignale wurden mit „CCPNMR 2.1.5“ ermittelt, gegen die Dauer der Relaxationsperiode aufgetragen und mittels „CurveFit“ [124] über eine monoexponentielle Fitfunktion daraus die Relaxationsraten berechnet. Die HetNOE's wurden mit „CCPNMR 2.1.5“ quantifiziert. Aus der Kombination der Relaxationsdaten jeder eindeutig zugeordneten Amidgruppe des Proteins wurde mit dem Programm „TENSOR2“ die globale Rotationskorrelationszeit ( $\tau_c$ ) des gesamten Proteins, die  $\tau_c$  der einzelnen WW-Domänen sowie die „Orderparameter“ ( $S^2$ ) der einzelnen Amidgruppen berechnet [158–163]. Die Konvertierung der experimentell ermittelten Relaxationsdaten in ein „Tensor2-kompatibles“ Dateiformat wurde mit Hilfe eines von *Marcel Jurk* (AG Schmieder, FMP-Berlin) geschriebenen Python-Skriptes bewerkstelligt.

## 3.8 MASSENSPEKTROMETRISCHE ANALYSE (MS)

Die Massenspektrometrische Analyse von <sup>15</sup>N<sup>13</sup>C-markierter WW23 (siehe Abschnitt 4.4.2) wurde von *Lilo Handel* (AG Oschkinat) an einem MALDI-Voyager System

---

4041 von Applied Biosystems durchgeführt. Es wurde im „Linear Mode“ im Bereich 5.000-50.000 Da gemessen. Als Matrix wurde Sinapinsäure verwendet.

### **3.9 ISOTHERME TITRATIONS-KALORIMETRIE (ITC)**

Die ITC-Messungen wurden bei 7°C an der VP-ITC oder an der ITC<sub>200</sub> (Microcal<sub>Inc</sub>, Northampton, MA) durchgeführt. Alle Titrationsexperimente erfolgten in ITC-Puffer bestehend aus 20 mM TRIS (pH 7.7 @ 7°C) und 50 mM NaCl. Protein und Peptid-Ligand wurden vor der Messung 18h gegen ITC-Puffer dialysiert und anschließend entgast. Die eingesetzten Proteinkonzentrationen in der Messzelle variierten zwischen 40 und 380 µM, die Peptidkonzentrationen in der Spritze zwischen 0.5 und 4 mM. Die Titrationsschritte variierten zwischen 5-10 µL an der VP-ITC und zwischen 0.5-2 µL an der ITC<sub>200</sub>. Es wurde ein Konzentrationsverhältnis zwischen WW-Domäne und prolinreichem Peptid von 1:10 angestrebt. Zur Quantifizierung möglicher Verdünnungseffekte wurden Kontrollexperimente durchgeführt, indem entweder ITC-Puffer in die Protein- bzw. Peptidlösung oder Protein- bzw. Peptidlösung in den ITC-Puffer titriert wurde.

Die Integration der Reaktionswärmen und die Basislinienkorrektur wurden mittels der Gerätesoftware Microcal-Origin® durchgeführt und ggf. manuell nachbearbeitet. Die weitere Auswertung der so prozessierten Thermogramme erfolgte mit SEDPHAT v.9.4 [134–140]. Alle Experimente einer Interaktionsreihe wurden zunächst separat und anschließend global gefittet. Die Konfidenzintervalle der Parameter  $K_d$  und  $\Delta H$  wurden unter Verwendung der in SEDPHAT integrierten MonteCarlo-Analyse ermittelt.

### **3.10 ANALYTISCHE ULTRAZENTRIFUGATION (AUZ)**

Die AUZ wurde mit einer *Beckman Optima XL-I* analytischen Ultrazentrifuge, bestückt mit einer Absorptions- und Interferenzoptik (675 nm Laser), durchgeführt. Für die Sedimentationsgeschwindigkeitsexperimente (SV) wurden Zweikammer-Titanzellen verwendet, wobei die Probenkammer mit 400 µL Proteinlösung und die Referenzzelle mit 400 µL Dialysepuffer gefüllt wurden. Alle Messungen erfolgten in einem 4-Loch-Roter bei einer Rotorgeschwindigkeit von 50.000 rpm. Proteinlösungen und Referenzpuffer wurden vor jedem Lauf ins Dialysegleichgewicht gebracht (24h Dialyse gegen 1000-fachen Überschuss Puffer bei 8°C). Die Sedimentation von CA150 WW23 wurde bei 10°C, die von CA150 WW23 in Anwesenheit des Peptidliganden CA\_P1 bei 12°C verfolgt. Die Absorption sedimentierender Spezies wurde bei einer Wellenlänge von 307 nm (für WW23) bzw. bei 305 nm (für WW23+Peptid CA\_P1) gemessen, wobei jeweils alle 8 min ein Scan mit einer radialen Auflösung von 0,01 cm durchgeführt wurde.

Die Auswertung der Messergebnisse erfolgte mit der Software SEDFIT [131,132]. Die dazu benötigten Parameter wie die Pufferdichte ( $\rho$ ), die Pufferviskosität ( $\eta$ ) und das partialspezifische Volumen ( $\bar{v}$ ) des Proteins wurden mit dem Programm

---



SEDNTERP [133] berechnet. Die eingesetzten Proteinmengen sind den nachfolgenden Tabellen zu entnehmen.

**Tabelle 3.10.I: In SV-Experimenten eingesetzte WW23-Konzentrationen**

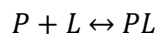
Zelle 1	Zelle 2	Zelle 3
8 µM WW23	163 µM WW23	318 µM WW23

**Tabelle 3.10.II: In SV-Experimenten eingesetzte Konzentrationen von WW23 und CA\_P1**

	Zelle 1	Zelle 2	Zelle 3
CA150 WW23	250 µM	250 µM	250 µM
CA_P1	32 µM	65 µM	765 µM
Sättigung* von WW23	~9%	~17%	~84%

\*Die Sättigung von WW23 wurde auf Grundlage der ITC - Bindekonstante von 109 µM errechnet.

Der prozentuale Anteil von mit CA\_P1 assoziierter WW23 wurde nach folgender Gleichung berechnet:



$P = \text{Protein}; L = \text{Ligand}$

$$K_d = \frac{[P] \cdot [L]}{[PL]}$$

$$[P_0] = [P] + [PL], [L_0] = [L] + [PL]$$

$$K_d = \frac{([PL] - [P_0]) \cdot ([PL] - [L_0])}{[PL]}$$

$$K_d = \frac{[PL]^2 - ([PL] \cdot [P_0] - [L_0] \cdot [PL] + [P_0] \cdot [L_0])}{[PL]}$$

$$K_d = [PL] - [P_0] - [L_0] + \frac{[P_0] \cdot [L_0]}{[PL]}$$

$$[L_0] = \frac{x \cdot ([P_0] \cdot (x - 1) - K_d)}{(x - 1)}$$

Bsp.: für  $K_d = 109 \mu\text{M}$  und  $P_0 = 250 \mu\text{M}$

$$[L_0] = \frac{x \cdot (250 \mu\text{M} \cdot (x - 1) - 109 \mu\text{M})}{(x - 1)}$$

$x = \text{Sättigung } 0,1 = 10\%$

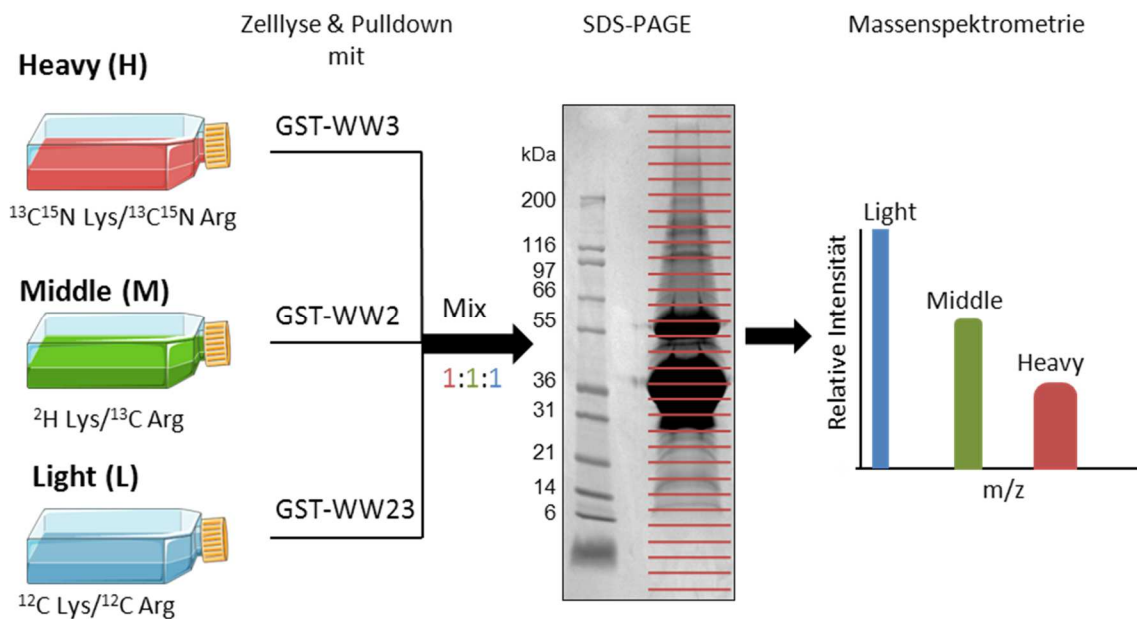
### 3.11 CD-SPEKTROSKOPIE

Der Anteil der Sekundärstrukturelemente von CA150 WW23 wurde mittels CD-Spektroskopie an einem *Jasco J-720 CD Spectrophotometer* (Jasco, Inc., Easton, Maryland) bestimmt. Hierzu wurde CA150 WW23 zunächst gegen CD-Puffer (20 mM

NaPO<sub>4</sub> (pH 7.5), 50 NaF) dialysiert. Für die Messung wurden 7 µM Protein (WW23) in einer 1 mm Quarzküvette eingesetzt. Das CD-Spektrum wurde im Bereich von 260-190 nm (0.2 nm Schrittweite) aufgenommen. Die Messung erfolgte in CD-Puffer bei einer Temperatur von 8°C. Als Referenz wurde ein Spektrum vom CD-Puffer aufgenommen, welches anschließend von den Messdaten subtrahiert wurde. Für jeden spektralen Datenpunkt wurden jeweils 20 Akkumulationen gemittelt. Die Auswertung des Spektrums erfolgte mittels Dichroweb [125,126] unter Verwendung des Referenzdatensatzes 7 [125,164].

### 3.12 TRIPLE-SILAC PULLDOWN

Im Rahmen einer Proteomanalyse sollte die Frage beantwortet werden, ob WW23 im Kontext des zellulären Proteoms auf Grundlage multivalenter Effekte andere Bindungseigenschaften als die Einzel-WWs (WW2&WW3) besitzt. Zusätzlich sollte dadurch auch die Bindungsaktivität der WW-Varianten bestätigt werden. Die durchgeführte Proteomanalyse beruhte auf einer MS-basierten Quantifizierung von WW-spezifischen Proteinen, die dazu nach dem SILAC-Prinzip (*Stable isotope labeling with amino acids in cell culture*) [165] metabolisch markiert wurden (siehe Abbildung 14).



**Abbildung 14: Schema des Triple-SILAC-Pulldowns**

Während der Kultivierungsphase werden drei HeLa-Zellpopulationen unterschiedlich metabolisch bzw. isotopenspezifisch (Heavy, Middle, Light) markiert (Triple-SILAC). In Co-Präzipitationsexperimenten (sogenannten Pulldowns) werden die entsprechenden Zelllysate anschließend mit GST-WW Fusionsproteinen inkubiert. Danach werden die mit den WW2-, WW3- bzw. WW23 Interaktionspartnern angereicherten Co-Präzipitate im Verhältnis 1:1:1 gemischt und die vorhandenen Proteine mittels SDS-PAGE molekulargewichtsabhängig separiert. Die Proteine innerhalb einzelner Gelbanden (bzw. Molekulargewichtsbereiche, siehe Streifenmaske) werden dann tryptisch verdaut und mittels Massenspektrometrie analysiert. Anhand der spezifischen Massendifferenzen zwischen identischen, aber unterschiedlich isotope markierten Peptiden wird dann deren relatives Verhältnis quantifiziert.

### 3.12.1 SILAC: KULTIVIERUNG DER HELA-ZELLEN

Die Kultivierung der HeLa-Zellpopulationen wurde von *Daniela Kosslick* und die Gewinnung der Zelllysate von *Stefan Klippel* (beide AG Freund) durchgeführt. Die Kultivierung erfolgte in DMEM-basierten SILAC-Medien (*Dulbecco's ModiWed Eagle Medium*) [165,166]. Das DMEM wurde dazu mit entsprechend isotoopenmarkiertem Arginin (21 mg/mL) und Lysin (48 mg/mL) in versetzt (siehe Tabelle 3.12.1).

**Tabelle 3.12.1: SILAC-Isotopen**

SILAC-Medium	Arginin	Lysin	Massedifferenz*
<i>Light</i>	$^{12}\text{C}$	$^{12}\text{C}$	Arg0 + Lys0
<i>Medium</i>	$^{13}\text{C}_6$	$^2\text{H}_4$	Arg6 + Lys4
<i>Heavy</i>	$^{13}\text{C}_6\ ^{15}\text{N}_4$	$^{13}\text{C}_6\ ^{15}\text{N}_2$	Arg10 + Lys8

\* gegenüber der natürlichen Isotopenzusammensetzung

### 3.12.2 PULLDOWN

Je 25  $\mu\text{L}$  Glutathion-Sepharose-Beads (GE-Healthcare) wurden mit GST-WW2, GST-WW3 bzw. GST-WW23 (jeweils 14  $\mu\text{M}$  bzw. 3-fache Matrixsättigung) beladen und für 2h bei 4°C inkubiert. Anschließend wurden die Beads 3x mit kaltem PBS-Puffer gewaschen und für 5 min bei 500xg und 4°C zentrifugiert. Nach Entfernung des PBS-Überstandes wurden die beladenen Beads mit je 4,5 mg HeLa-Zelllysate versetzt und über Nacht rotierend bei 4°C inkubiert.

GST-WW23 + *Light* Lysat

GST-WW2 + *Medium* Lysat

GST-WW3 + *Heavy* Lysat

Anschließend wurden alle Ansätze 3x mit 1 mL eiskaltem PBS gewaschen und 5 min bei 500xg und 4°C zentrifugiert. Der Überstand wurde verworfen und die Ansätze im Verhältnis 1:1:1 vereint. Nach einem weiteren Zentrifugationsschritt wurden die Beads mit 60  $\mu\text{L}$  5x Probenpuffer versetzt und 5 min bei 96°C inkubiert. Abschließend wurden die Beads abfiltriert (Ultrafree® Centrifugal Unit 0.45  $\mu\text{M}$  pore size, MILLIPORE) und das Filtrat einer SDS-PAGE (Novex® 4-20% Tris-Glycine Gel 1.0 mm Gradienten-Gel) unterzogen.

### 3.12.3 SILAC: MS-ANALYSE

#### > MS-Probenvorbereitung

Nach Coomassiefärbung wurden die Proteinbanden ausgeschnitten und zunächst mit  $\text{dH}_2\text{O}$ , dann mit 50% (v/v) ACN in 50 mM Ammoniumhydrogencarbonat und abschließend mit 50 mM Ammoniumhydrogencarbonat gewaschen. Danach wurden die Banden durch Zugabe von ACN dehydriert und in einer Vakuumzentrifuge getrocknet.

---

Durch Zugabe von je 0,05 µg Trypsin in 50 mM Ammoniumhydrogencarbonat wurden die Proteinbanden 17h bei 37°C enzymatisch verdaut. Danach wurde die Reaktion durch Zugabe von 20 µL 0,5% (v/v) TFA in ACN gestoppt und die Proben einer 2 minütigen Ultraschallbehandlung unterzogen. Der flüssige Überstand wurde in einer Vakuumzentrifuge eingeeengt. Der Rückstand wurde abschließend durch Ultraschall (3 min) in 6 µL 5% (v/v) ACN, 0,1% TFA gelöst und bis zur MS-Analyse bei -20°C gelagert.

### **> LC-MS/MS und Datenprozessierung**

Die LC-MS/MS Analyse und die Datenprozessierung wurden von *Dr. Michael Schümann* (AG-Krause) durchgeführt.

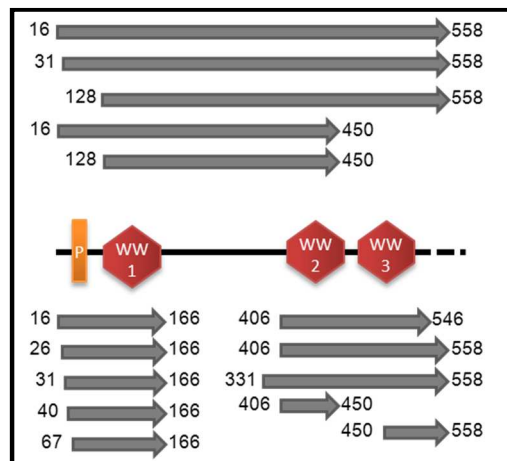
Die tryptisch-verdauten Proteinfragmente wurden an einem LTQ-Orbitrap Hybrid-Massenspektrometer (*Thermo Fischer*), ausgestattet mit einem *reversed-phase capillary liquid chromatography system* (Eksigent 2D Nanoflow LC System, *Axel Semarau GmbH*), vermessen. Die Identifikation und Quantifizierung der Proteine erfolgte mit dem MaxQuant Software Packet (Version 1.0.12.31) [167]. Die generierten Peaklisten (msm files) wurden dabei über einen MASCOT Server (Version 2.2, Matrix Science Ltd.) gegen eine humane IPI Proteindatenbank (Version 3.52) abgeglichen. Bei der Proteinidentifikation wurden Methionin-oxidationen und Acrylamidmodifikationen von Cysteinen berücksichtigt.

---

## 4 ERGEBNISSE

### 4.1 KLONIERUNG WW-DOMÄNEN-CODIERENDER-VEKTOREN

Ein wichtiger Schritt für das Studium der CA150 WW-Domänen und deren Wechselwirkungen war die Herstellung T7-Promotor-basierter WW-codierender pET-Vektoren durch biotechnologische Klonierungsarbeit. Durch den Einsatz entsprechender Oligonukleotid-Kombinationen (vgl. Tabelle 4.1.I) sollten mittels PCR die benötigten DNA-Fragmente, welche WW-Varianten unterschiedlichster Länge und WW-Kombination (siehe Abbildung 15) codieren, gewonnen werden.



**Abbildung 15: Übersicht der WW-Domänen-Varianten**

Schematischer Ausschnitt aus der CA150-Sequenz (Mitte). Die Pfeile repräsentieren die zur Klonierung ausgewählten CA150-Sequenzbereiche. Die Zahlen repräsentieren die Position der Aminosäuren in CA150. P: prolinreiche Region; WW: WW-Domäne.

Zwecks Optimierung der PCR-Bedingungen wurden Parameter, wie Template-, dNTP-Oligonukleotid- und Magnesiumkonzentration variiert. Unter Verwendung von Temperaturgradienten wurde der „Annealing“-Schritt der PCR hinsichtlich maximaler Produktausbeute optimiert. Die für die Synthese der einzelnen WW-Sequenzen verwendeten Oligonukleotide und die Ausbeute der entsprechenden PCR sind in der nachfolgenden Tabelle zusammengefasst.

**Tabelle 4.1.I: Ergebnisse der PCR's**

Fragment	Oligonukleotid CA150_	Position in CA150	PCR 1	PCR 2	Länge [bp]	Domäne
F1	fw_16, rev_167	16-166	+	+	504	WW1
F2	fw_26, rev_167	26-166	-	-	459	WW1
F3	fw_31, rev_167	31-166	+	+	462	WW1
F4	fw_40, rev_167	40-166	+	+	435	WW1
F5	fw_67, rev_167	67-166	+	+	351	WW1
F6	fw_408, rev_548	406-546	+	+	468	WW23

Fragment	Oligonukleotid CA150_	Position in CA150	PCR 1	PCR 2	Länge [bp]	Domäne
F7	fw_408, rev_559	406-558	+	+	510	WW23
F8a	fw_16, rev_450	16-450	-	-	1302	WW12
F8	fw_128, rev_450	128-450	+	+	987	WW12
F9	fw_16, rev_559	16-558	-	-	1671	WW123
F10	fw_31, rev_559	31-558	+	+	1500	WW123
F11	fw_331, rev_559	331-558	+	+	702	WW23
F12	fw_506, rev_559	506-558	+	+	177	WW3
F13	fw_408, rev_450	406-558	+	+	177	WW2
F14	fw_67, rev_559	67-558	+	+	1491	WW123

Gelistet ist das Ergebnis der PCRs, +: korrektes PCR-Produkt, - kein PCR-Produkt

Wie Tabelle 4.1.I zeigt, war es trotz intensiver Bemühungen nicht möglich die DNA-Fragmente für die Varianten F2 (26-166; WW1), F8a (16-450; WW12) und F9 (16-558; WW123) herzustellen. Die erzeugten DNA-Fragmente der übrigen Konstrukte, wurden unter Verwendung der Ligationsunabhängigen Klonierungsmethode (LIC) in verschiedene *E. coli* Vektoren (pET30EK/LIC; pET41EK/LIC; pCDF2EK/LIC) eingebaut. So entstandene Plasmide wurden durch DNA-Sequenzierung (vorgenommen durch die Firma *Invitac*) analysiert und unter Zuhilfenahme der Software VectorNTI® auf ihre korrekte Basensequenz hin überprüft.

## 4.2 EXPRESSIONSANALYSE DER WW-VARIANTEN

Diverse Untersuchungen zur Optimierung der Expressionsausbeuten an WW-Varianten, in denen sowohl Parameter, wie Temperatur, *E. coli* Expressionswirt und Dauer der Expressionsphase, als auch die Menge des Induktionsmittels IPTG variierten, wurden wie in Abschnitt 3.4.5 beschrieben durchgeführt. In Tabelle 4.2.I sind die Ergebnisse dieser Untersuchungen hinsichtlich der relativen Ausbeuten an löslich überexprimierten Zielproteinen zusammengestellt. Optimaler Expressionsstamm war *E. coli* BL21 (DE3) Rosetta2. Die höchsten Proteinausbeuten wurden bei einer IPTG-Konzentration von 1 mM, einer Expressionstemperatur von 25°C und einer Expressionsdauer von 18 Stunden erzielt.

**Tabelle 4.2.I: Expressionslevel der WW-Varianten in Vollmedium**

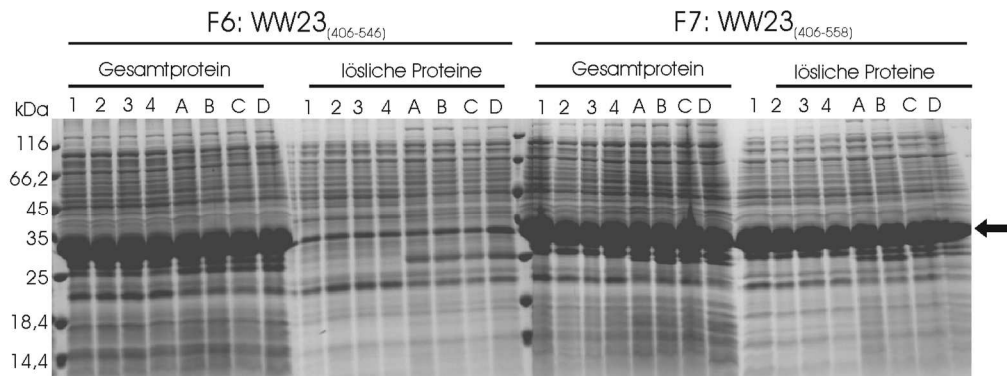
WW-Variante	AS	Tag	relative Expressionsstärke	WW-Variante in Lösung
WW1-F1	16-166	S His	+	-
WW1-F3	31-166	S His	+	-
WW1-F4	40-166	S His	+	-
WW1-F5	67-166	S His	++	+++
		GST S His	++	+++
WW23-F6	406-546	S His	+++	-
WW23-F7	406-558	S His	+++	+++
		GST S His	+++	+++
WW12-F8	128-450	S His	+	+++
		GST S His	+	-
WW123-F10	31-558	S His	-	-
		GST S His	-	-

WW-Variante	AS	Tag	relative Expressionsstärke	WW-Variante in Lösung
WW23-F11	331-558	S His	+	+
		GST S His	++	++
WW3-F12	506-558	His	+	+
		S His	+	+++
WW2-F13	406-558	GST S His	++	+++
		His	+	+
		S His	+	+++
WW123-F14	67-558	GST S His	++	+++
		GST S His	-	-

Die relative Expressionsstärke wurde auf Basis der Bandenintensität der Zielproteine abgeschätzt ((-) keine; (+) geringe; (++) mittlere; (+++) starke Expression). Die Menge löslich exprimierter WW-Varianten wurde über den Vergleich der Bandenintensität des entsprechenden Zielproteins im Totallysat und in der löslichen Proteinfraction abgeschätzt ((-) keine; (+) ~10%; (++) ~50%; (+++) >90% löslich exprimiertes Zielprotein). S: S-Tag; GST: Glutathion-S-Transferase-Tag; His: (Histidin)<sub>6</sub>-Tag. AS: Aminosäureposition in CA150

Trotz intensiver Bemühungen war es nicht möglich die 3-WW-Domänen-Varianten F10 und F14 rekombinant in *E.coli* zu exprimieren. Die 2-WW-Domänen-Variante (WW12; F8) konnten zwar erfolgreich, jedoch nur in geringen Ausbeuten rekombinant hergestellt werden. WW12 konnte ebenfalls erfolgreich exprimiert werden, unterlag aber einem sehr schnellen proteolytischen Abbau, welchem trotz Durchführung der weiteren Arbeitsschritte bei 4-12°C und durch die Zugabe von Proteaseinhibitoren (Complete® Protease Inhibitor Mix (EDTA-frei)) nicht entgegengewirkt werden konnte. Dementsprechend wurde auf die weitere Arbeit mit diesem Konstrukt verzichtet. Alle Einzel-WW-Varianten (WW1-F5, WW2-F13 und WW3-F12) konnten rekombinant exprimiert werden, wobei bei den Einzel-WW-Domänen, welche neben dem His-Tag noch über einen Glutathion-S-Sepharose-Tag (GST-Tag) verfügten die besten Ergebnisse erzielt werden konnten. Lediglich die Einzel-WW-Domänen mit S- und Hexahistidin-Tag, sowie die mit zusätzlichem Glutathion-S-Sepharose-Tag (GST-Tag) waren zu 100% löslich und konnten erfolgreich aufgereinigt werden (vgl. Abschnitt 4.3).

Ein vergleichsweise sehr unterschiedliches Expressions- bzw. Löslichkeitsverhalten zeigten die beiden WW23-Varianten F6 und F7 (beide mit S His-Tag) (Abbildung 16). Die mit einer C-terminalen Helix terminierte, um 12 Aminosäuren längere WW23-Variante F7 konnte in *E. coli* erfolgreich in löslicher Form hergestellt werden. Die kürzere Variante F6 wurde zwar ebenso stark exprimiert wie F7, war aber unlöslich. Die C-terminale Helix scheint demzufolge essentiell für die lösliche Expression eines WW23-Konstruktes zu sein.



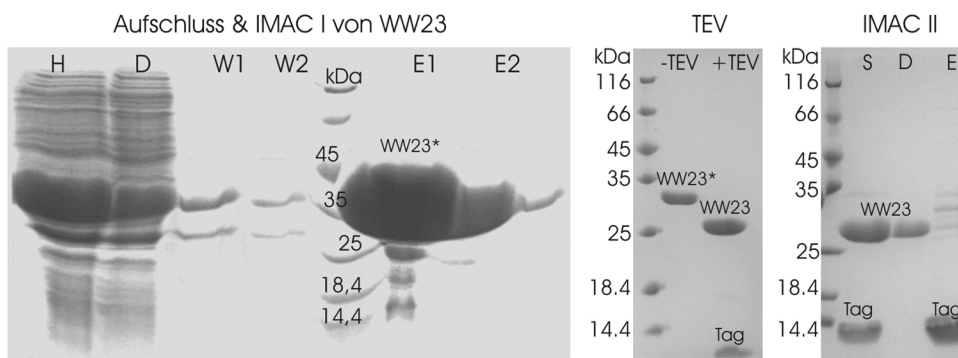
**Abbildung 16: Vergleich der Löslichkeit von CA150 WW23-F6 und CA150 WW23-F7**

SDS-Gele aus der Expressionsanalyse von WW23-F6 und WW23-F7. Die Spuren 1-4 zeigen die Expression in *E.coli* BL21 (DE3) und die Spuren A-D zeigen die Expression in *E.coli* BL21(DE3) Rosetta2. Aufgetragen wurden der Gesamtproteinextrakt und der lösliche Proteinüberstand nach Zentrifugation. Beide WW23-Varianten zeigen in der SDS-PAGE ein deutlich verzögertes Laufverhalten, Entsprechend sind die Proteinbanden nicht bei erwartungsgemäß 21 kDa (F6) bzw. 18 kDa (F7) lokalisiert, sondern oberhalb der 25 kDa Bande des Markerproteins (ca. 30 kDa, siehe Pfeil).

### 4.3 PRÄPARATION DER WW-VARIANTEN

Alle WW-Varianten wurden, wie in Abschnitt 3.5.2 beschrieben, mittels IMAC über den Hexahistidin-Tag aufgereinigt. In Abbildung 17 sind die Reinigungsschritte exemplarisch für WW23-F7 gezeigt. Über eine Gelfiltration wurde der Reinheitsgrad der isolierten Proteine weiter erhöht und der Erfolg der Methode per SDS-PAGE (Abbildung 18) dokumentiert.

#### ➤ Reinigung von $^{15}\text{N}^{13}\text{C}$ -markierter WW23-F7



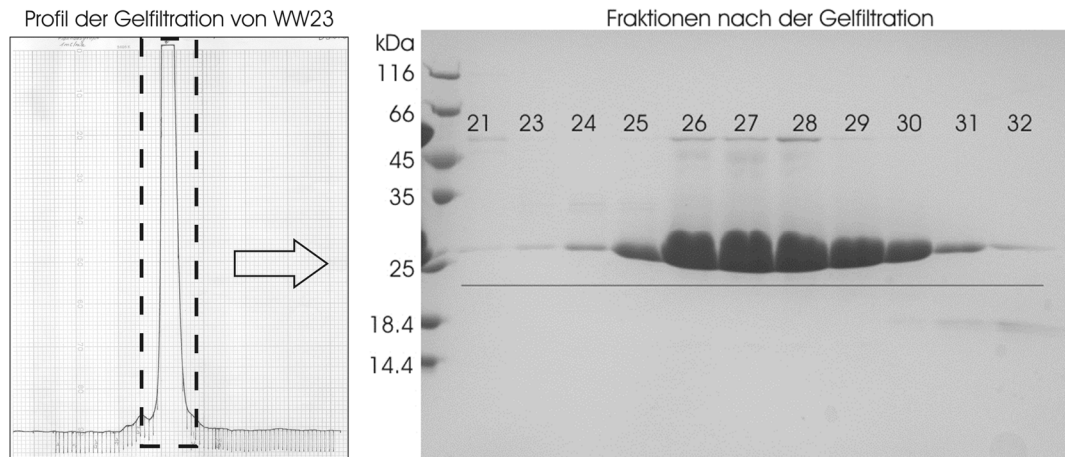
**Abbildung 17: IMAC und TEV-Spaltung von WW23**

Das SDS-Gel zeigt verschiedene Schritte der Reinigung von  $^{15}\text{N}^{13}\text{C}$ -markierter WW23. Neben dem Proteinmarker wurden folgende Proben aufgetragen: H: Zellhomogenisat, D: Durchflussfraktionen der IMAC I & II, Waschfraktionen, des 1. Waschschrtes (W1) und 2. Waschschrtes (W2), Eluatfraktionen der IMAC I (E1, E2) und der IMAC II (E), WW23 vor (-TEV) und nach TEV-Spaltung (+TEV, S). WW23\*: Fusionsprotein

Mit Ausnahme der Einzel-WW-GST- (WW1-F5-, WW2-F13- und WW3-F12-GST) und der Einzel-WW-S-His-Fusionsproteine (WW2-F13-S-His und WW3-F12-S-His) konnte der Reinigungs-Tag bei allen WW-Varianten vollständig abgespalten werden. Diverse Versuche zur Abspaltung des GST- und GST-S-His-Tags der Einzel-WWs mit TEV-



Protease, Enterokinase sowie Thrombin (über 1-3 Tage) schlugen fehl, da die Proteine präzipitierten oder abgebaut wurden. Als Fusionsproteine konnten die Einzel-WW-GST-Varianten aber erfolgreich in den SPOT-Analysen (Abschnitt 4.6.1) und im SILAC-Pulldown (Abschnitt 4.7) eingesetzt werden.



**Abbildung 18: Gelfiltration von WW23**

Links: Ausschnitt aus dem Chromatogramm der Gelfiltration von WW23. Es wurde eine 320 mL Superdex75 Säule mit 3 mL Probenvolumen beladen. Der Lauf erfolgte mit einer Geschwindigkeit von 1 mL/min. Das Fraktionsvolumen betrug 4 mL. Die Fraktionen (21-32) innerhalb des gestrichelten Rechtecks wurden mittels SDS-PAGE (Rechts) analysiert.

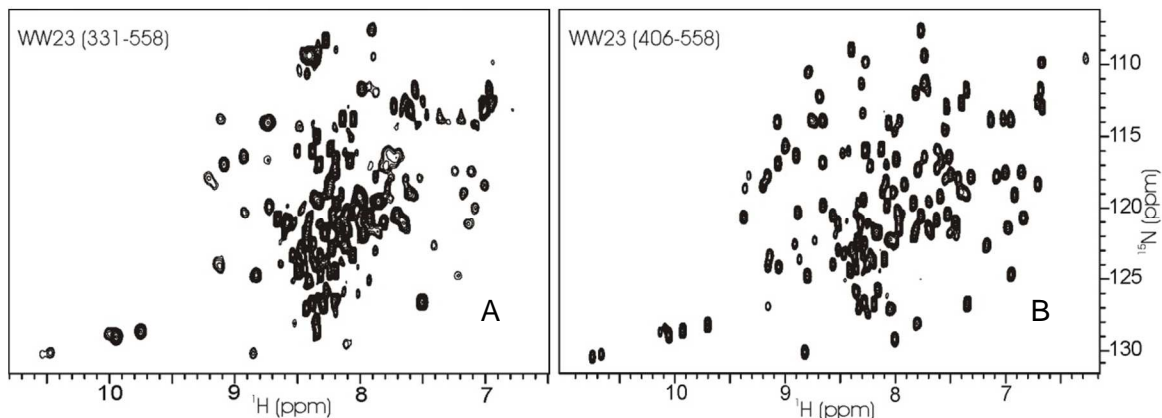
Alle WW-Varianten konnten stets als monomere Spezies separiert werden. Abgesehen von den GST-Fusionsproteinen waren alle WW-Varianten relativ stabil und problemlos bei  $-80^{\circ}\text{C}$  lagerfähig. Die Handhabung der GST-Fusionsproteine gestaltete sich etwas schwieriger, da diese relativ schnell präzipitierten und nur nach Zugabe von 50% Glycerol zum Puffer eingefroren werden konnten.

#### 4.4 ANALYSE DER WW-VARIANTEN

Um beurteilen zu können, welche WW23-Variante (WW23-F7 und WW23-F11) sich für die geplanten Bindungsstudien (NMR und ITC) zur Untersuchung multivalenter Interaktionen am besten eignen könnte, wurden diese zunächst mittels heteronuklearer 2D Lösungs-NMR-Spektroskopie bezüglich ihres Faltungszustandes untersucht. Ausschlaggebendes Kriterium war dabei die Anzahl und Dispersion von  $^1\text{H}^{15}\text{N}$ -Kreuzsignalen innerhalb der 2D  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren entsprechend  $^{15}\text{N}$ -markierter Konstrukte. Die in Betracht kommende WW-Variante wurde zwecks Aufnahme verschiedener 3D NMR-Spektren für die sequentielle Zuordnung (siehe Abschnitt 3.7.1& 4.5.1) dann uniform  $^{15}\text{N}^{13}\text{C}$ -markiert hergestellt. Der erzielte Isotopen-Anreicherungsgrad ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ) wurde massenspektrometrisch untersucht und mittels CD-Spektroskopie die Sekundärstrukturanteile erfasst. Letzteres sollte insbesondere Erkenntnisse über die strukturelle Beschaffenheit des Interdomänen-Linkers in WW23 liefern.

#### 4.4.1 NMR-SPEKTROSKOPISCHE UNTERSUCHUNG DES FALTUNGSZUSTANDES VERSCHIEDENER WW-VARIANTEN

Jedes Protein besitzt ein einzigartiges und charakteristisches  $^1\text{H}^{15}\text{N}$ -HSQC-Spektrum, welches im übertragenen Sinne auch als „Fingerabdruck“ eines Proteins bezeichnet wird. Anhand der Anzahl, Frequenz und Dispersion der darin abgebildeten  $^1\text{H}^{15}\text{N}$ -Kreuzsignale können auch ohne eine Zuordnung Rückschlüsse über den Faltungszustand und die Erfolgsaussichten für eine sequentielle Resonanzzuordnung gezogen werden. Die  $^1\text{H}_\text{N}$  und  $^{15}\text{N}$  chemischen Verschiebungen der Amidgruppen von in unstrukturierten bzw. ungefalteten Sequenzbereichen lokalisierten Aminosäuren unterscheiden sich, bezogen auf einen Aminosäuretyp, nur sehr wenig. Im  $^1\text{H}^{15}\text{N}$ -HSQC-Spektrum sind solche Bereiche an sehr intensiven und stark überlagerten Signalen im mittleren Spektralbereich zu erkennen. Gut aufgelöste und über einen großen Spektralbereich verteilte Signale sind dagegen ein Zeichen für das Vorhandensein von definierten Sekundär- und Tertiärstrukturen. Für eine erfolgreiche sequentielle Zuordnung sollte die Anzahl der sichtbaren Signale (abzüglich der Signale von Seitenketten) möglichst mit der Zahl der im Protein vorhandenen Aminosäuren korrelieren. Dabei ist zu beachten, dass Proline und die N-terminale Aminogruppe grundsätzlich nicht im  $^1\text{H}^{15}\text{N}$ -Spektrum detektiert werden. In Abbildung 19 sind die  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von 2 unterschiedlich langen WW23-Varianten dargestellt.

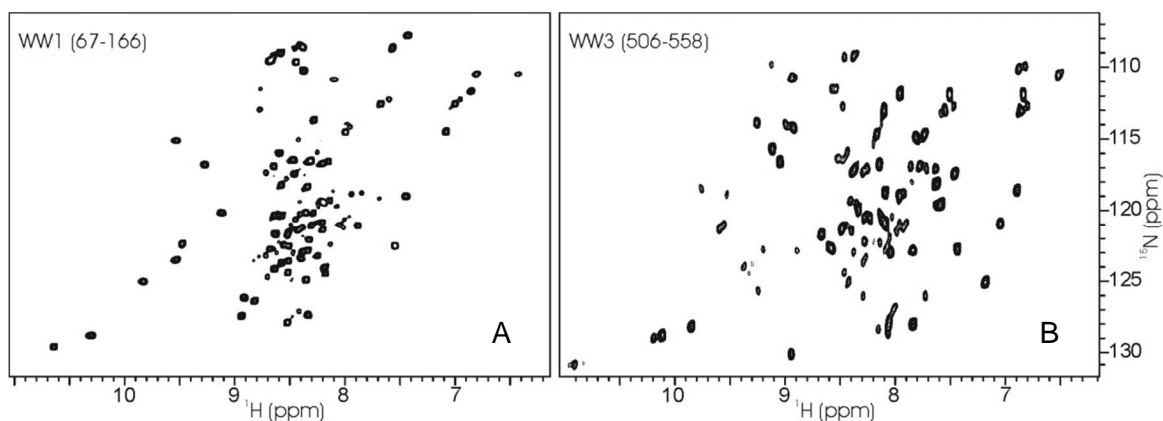


**Abbildung 19:**  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren der WW23 Varianten F11 und F7

$^1\text{H}^{15}\text{N}$ -HSQC Spektrum von uniform  $^{15}\text{N}$  markiertem WW23-F11 (WW23<sub>331-558</sub>) (A) und WW23-F7 (WW23<sub>406-558</sub>) (B). Aufgenommen an einem 600 MHz Spektrometer bei 280K.

Obwohl WW23-F11 75 Aminosäuren länger ist, als WW23-F7, sind in dessen Spektrum bedeutend weniger Signale vorhanden, als im Spektrum von WW23-F7. Die Zahl der in WW23-F7 detektierten Kreuzsignale entspricht dagegen in etwa dem Erwartungswert. Das  $^1\text{H}^{15}\text{N}$ -HSQC von WW23-F7 zeigt darüber hinaus eine erheblich bessere Signaldispersion und eine differenziertere spektrale Signalverteilung. Alle weiteren Untersuchungen wurden deshalb mit WW23-F7 durchgeführt.

Da im Rahmen dieser Arbeit auch Interaktionsstudien mit einzelnen WW-Domänen durchgeführt werden sollten, wurde deren Faltungszustand ebenfalls NMR-spektroskopisch untersucht. In Abbildung 20 sind die  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren der beiden erfolgreich rekombinant herstellbaren Einzeldomänen WW1(-F5) und WW3(-F12) dargestellt.



**Abbildung 20:  $^1\text{H}^{15}\text{N}$  HSQC-Spektren von WW1-F5 und WW3-F12**

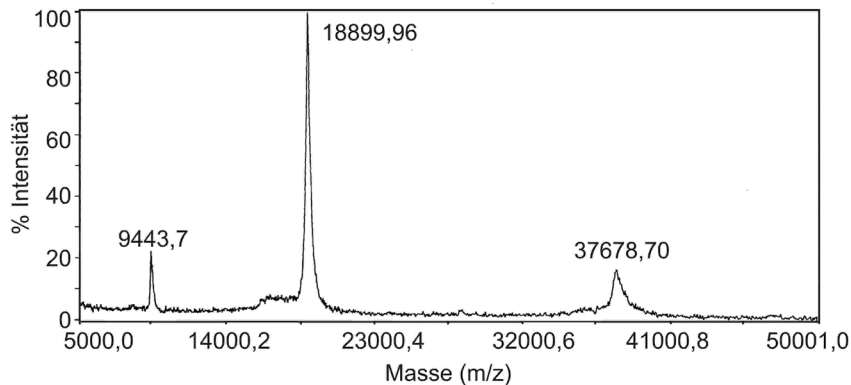
$^1\text{H}^{15}\text{N}$ -HSQC Spektren von uniform  $^{15}\text{N}$  markiertem WW1 (WW1<sub>67-166</sub>, F5) (A) und WW3 (WW3<sub>506-558</sub>, F12) (B). Aufgenommen an einem 600 MHz Spektrometer bei 280K.

Die Spektren beider Proteine sind sehr gut aufgelöst und die Anzahl der detektierten Kreuzsignale entspricht jeweils dem Erwartungswert. Im HSQC von WW1 liegen auffällig viele Signale im Bereich zwischen 8-9 ppm, was für einen hohen Anteil  $\alpha$ -helikaler Strukturanteile sowie möglicherweise auch unstrukturierter Bereiche spricht. Denkbar wäre, dass die in diesem Protein vorhandenen Polyprolin-Helices für die „helix-lastige“ Signalverteilung verantwortlich sind. Das HSQC von WW3 zeigt eine spektrale Signalverteilung, wie für ein aus gleichermaßen  $\alpha$ -helikalen und  $\beta$ -strängigen Strukturanteilen bestehendes Protein zu erwarten ist. Das spricht gegen die Annahme von *Otte et al 2003* [80], nach der WW3 eine relativ unstrukturierte Domäne wäre. Da in dem von *Otte et al.* untersuchten WW3-Konstrukt die C-terminale  $\alpha$ -Helix fehlt ist anzunehmen, dass deren Anwesenheit für die Faltung von WW3 essentiell ist.

#### **4.4.2 MASSENSPEKTROMETRISCHE ANALYSE VON UNIFORM $^{15}\text{N}^{13}\text{C}$ -MARKIERTER WW23-F7**

Um den Anreicherungsgrad der Isotopen  $^{15}\text{N}$  und  $^{13}\text{C}$  in rekombinantem, uniform  $^{15}\text{N}^{13}\text{C}$ -markiertem WW23-F7 zu bestimmen wurde die für die sequentielle Zuordnung (vgl. Abschnitt 4.5.1) hergestellte Proteinprobe zunächst massenspektrometrisch analysiert. Das Massenspektrum in Abbildung 21 zeigt 3 distinkte Spezies mit Massen von 9.444 Da, 18.900 Da und 37.679 Da. Da das berechnete Molekulargewicht von vollständig  $^{15}\text{N}^{13}\text{C}$ -markierter WW23-F7 18.875 Da beträgt und ein methodenbedingter Messfehler von  $\pm 50$  Da berücksichtigt werden muss, ist die Spezies mit  $m = 18.900$  Da dem einfach geladenen ( $m/z$ ) und die Spezies mit  $m = 9.444$  Da dem zweifach geladenen ( $m/2z$ )

Molekülion von WW23-F7 zuzuordnen. Die dritte Spezies ( $m = 37.678$  Da) könnte einem WW23-Dimer entsprechen. Da sich derartige artifizelle Dimere häufig während des Ionisierungs-Prozesses bilden, kann hier keine Aussage darüber getroffen werden, ob die Probe tatsächlich dimeres WW23-F7 enthält. Mittels Analytischer Ultrazentrifugation (AUZ) wurde diese Fragestellung deshalb noch genauer untersucht (siehe Abschnitt 4.8.4). Auf Basis des unteren Grenzwertes für die Masse des einfach geladenen Molekülions (18.850 Da) ergibt sich ein Anreicherungsgrad von mindestens 97,5%.

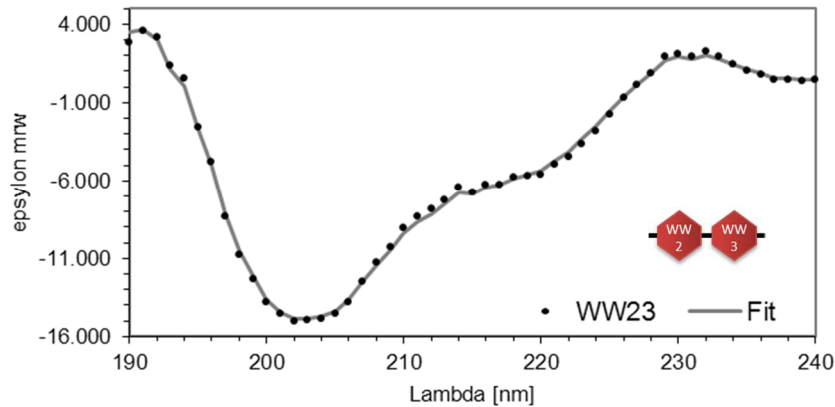


**Abbildung 21: Massenspektrum von uniform  $^{15}\text{N}^{13}\text{C}$ -markiertem WW23-F7**

Das Signal höchster Intensität wurde auf 100% normiert. Die Zahlen an den Signalmaxima geben die korrespondierenden Molekulargewichte in Dalton an.

#### **4.4.3 CD-SPEKTROSKOPISCHE SEKUNDÄRSTRUKTURANALYSE VON WW23-F7**

Um Informationen über die Sekundärstrukturanteile von WW23-F7 zu gewinnen, wurde ein CD-Spektrum des Proteins aufgenommen (siehe Abschnitt 3.11) und mit dem internetbasierten Programm „Dicroweb“ [125,126] daraus die Anteile der erfassbaren Sekundärstrukturen ( $\alpha$ -Helix,  $\beta$ -Stränge und Schleifen) ermittelt. Die Quantifizierung der Sekundärstrukturanteile erfolgte mit dem „CDSSTR-Algorithmus“ [168] auf Basis der CD-Spektren von Referenzproteinen mit bekannten Sekundärstrukturanteilen (Referenzdatenbank 7). Durch eine iterative Kombination von Teilbereichen verschiedener Referenzspektren wird das gemessene CD-Spektrum möglichst genau rekonstruiert und daraus die entsprechenden Sekundärstrukturanteile des untersuchten Proteins abgeleitet. Abbildung 22 zeigt das experimentell gemessene und das rekonstruierte CD-Spektrum.



**Abbildung 22: Überlagerung des realen und des rekonstruierten CD-Spektrums von WW23**

Der Verlauf des gemessenen CD-Spektrums ist durch Punkte, das mittels CDSSTR-Algorithmus“ rekonstruierte CD-Spektrums als durchgehende Linie dargestellt.

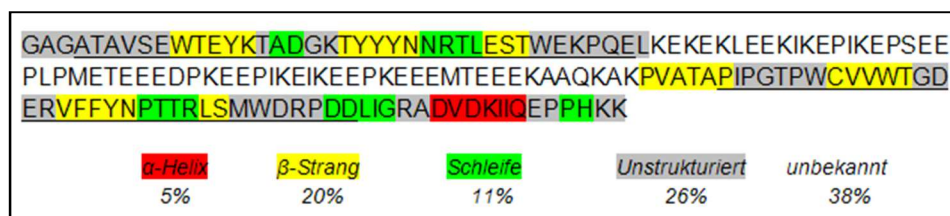
Das experimentell gemessene CD-Spektrum konnte nahezu vollständig durch ein rekonstruiertes CD-Spektrum abgebildet werden. Daher ist anzunehmen, dass die im Protein tatsächlich vorhandenen Sekundärstrukturelemente relativ genau bestimmt wurden. Die Resultate der Analyse sind in Tabelle 4.4.I dargestellt.

**Tabelle 4.4.I: Mit "CDSSTR" berechnete Sekundärstruktur-Anteile von WW23**

	WW23
$\alpha$ -Helix	4%
$\beta$ -Strang	33%
Schleife	22%
unstrukturiert	40%
nrmsd*	0,028

\* Normalized root-mean-square-deviation [169]

Ein Vergleich der berechneten Sekundärstrukturelemente in WW23-F7 mit den PDB-Strukturinformationen von WW2 und WW3 (WW2-PDB-Code 1E0L; WW3-PDB-Code 2DK7) (siehe Abbildung 23) zeigt vergleichbare Anteile an  $\alpha$ -Helix.



**Abbildung 23: Bekannte Sekundärstrukturelemente in WW23-F7**

Unstrukturierte Anteile sind grau,  $\alpha$ -helikale rot,  $\beta$ -Strang und Schleifen Bereiche sind gelb bzw. grün hinterlegt. Die Regionen der WW-Domänen (WW2 und WW3) wurden unterstrichen.

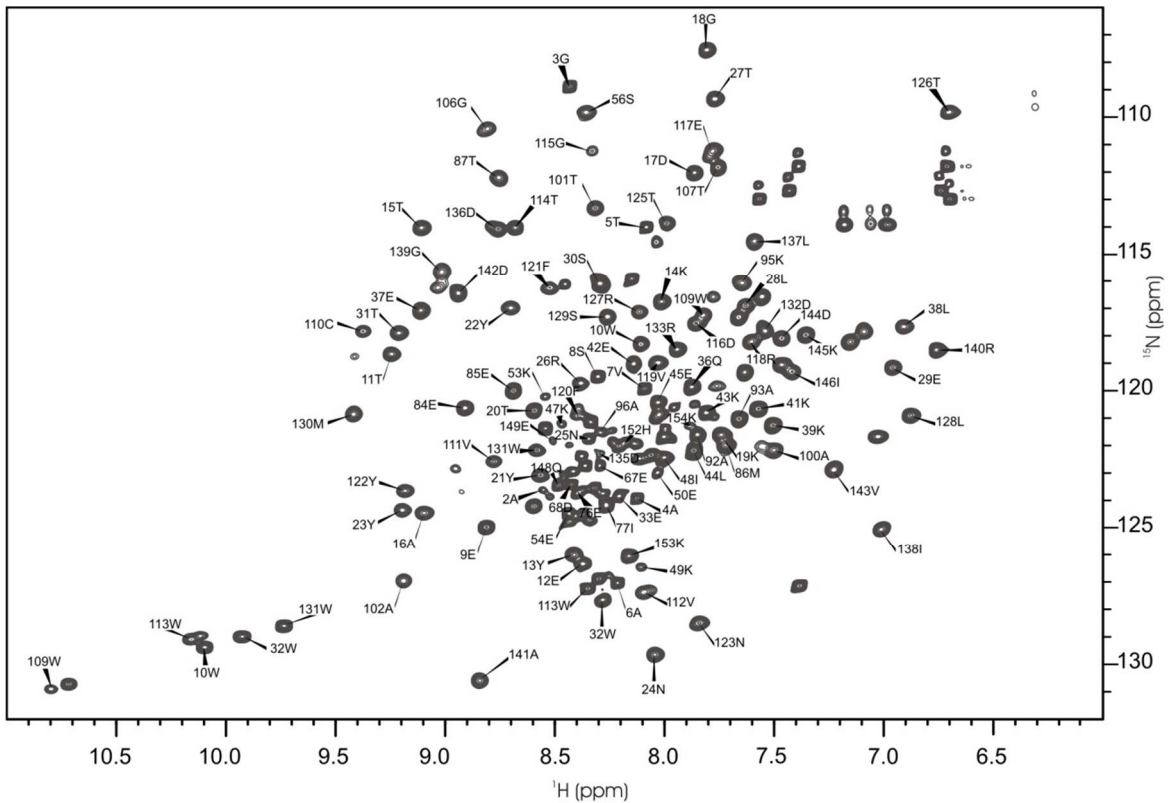
Somit kann davon ausgegangen werden, dass sich innerhalb des Interdomänenlinkers keine  $\alpha$ -helikalen Strukturelemente befinden. Vergleicht man den  $\beta$ -Strang-Anteil von WW23-F7 mit WW2 und WW3 wird deutlich, dass dieser in WW23-F7 größer ist, als die Summe der  $\beta$ -Strang-Anteile von WW2 und WW3. Es erscheint daher wahrscheinlich, dass in WW23 zusätzliche  $\beta$ -Strang-Elemente vorliegen. Diese könnten an den Domänengrenzen, im Interdomänen-Linker oder in beiden Bereichen lokalisiert sein. Da die unstrukturierten Anteile und Schleifen in WW23 zusammen ca. 62% ausmachen, muss der Interdomänen-Linker aber auch unstrukturierte Bereiche enthalten.

#### **4.5 NMR-SPEKTROSKOPISCHE UNTERSUCHUNGEN**

Für die Lösungs-NMR-spektroskopische Untersuchung von Proteinen muss die zu untersuchende Probe einige Voraussetzungen erfüllen. Da die Methode relativ unempfindlich ist, werden, abhängig vom Experimenttyp, recht hohe Proteinkonzentrationen benötigt. Die erforderlichen Proteinkonzentrationen bewegen sich in einem Bereich von 0,05 mM für die Aufnahme von  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren bis  $>0,5$  mM für die Aufnahme von Triple-Resonanz-Spektren. Für viele NMR-Experimente müssen die Proteine zudem uniform mit den Isotopen  $^{15}\text{N}$  ( $^1\text{H}^{15}\text{N}$ -HSQC) bzw.  $^{15}\text{N}$  und  $^{13}\text{C}$  (Triple-Resonanz-Spektren) markiert sein. Zusätzlich müssen die Proteine sehr rein vorliegen ( $>90\%$  Reinheitsgrad) und während der oft mehrere Tage dauernden Messungen stabil bleiben. Bei der Untersuchung der WW-Domänen hat sich zudem herausgestellt, dass hier tiefere Temperaturen (7-14°C) zu qualitativ besseren Spektren führen.

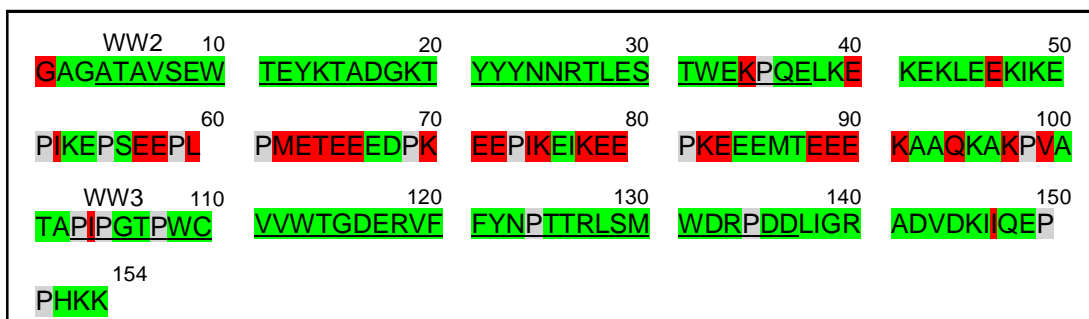
##### **4.5.1 SEQUENTIELLE ZUORDNUNG VON WW23-F7**

Um moleküldynamische Eigenschaften erfassen und die mit dem jeweiligen prolinreichen Peptidliganden interagierenden Aminosäuren der WW-Domänen identifizieren zu können, war es zwingend erforderlich, die  $^1\text{H}_\text{N}^{15}\text{N}$ -Resonanzen von WW23-F7 den entsprechenden Aminosäuren sequentiell zuzuordnen. Dazu wurden eine Reihe verschiedener 3D-Spektren für die Sequentielle Zuordnung der  $^1\text{H}_\text{N}$ -,  $^{15}\text{N}$ -,  $^{13}\text{C}_\alpha$ -,  $^{13}\text{C}_\beta$ - und  $^{13}\text{C}_\text{O}$ -Resonanzen aufgenommen (siehe Abschnitt 3.7.1). Auf Basis dieser Spektren konnten ca. 75% der  $^1\text{H}^{15}\text{N}$ -Korellationen sequentiell eindeutig zugeordnet werden (siehe Abbildung 24 & Abbildung 25, vollständige Resonanzliste im Anhang, Tabelle 6.2.I).



**Abbildung 24: Zugeordnetes HSQC-Spektrum von WW23-F7**

Gezeigt ist ein  $^1\text{H}^{15}\text{N}$ -HSQC-Spektrum von WW23-F7 aufgenommen an einem 750 MHz-Spektrometer bei 287K.



**Abbildung 25: Zugeordnete  $^1\text{H}^{15}\text{N}$ -Korellationen von WW23-F7**

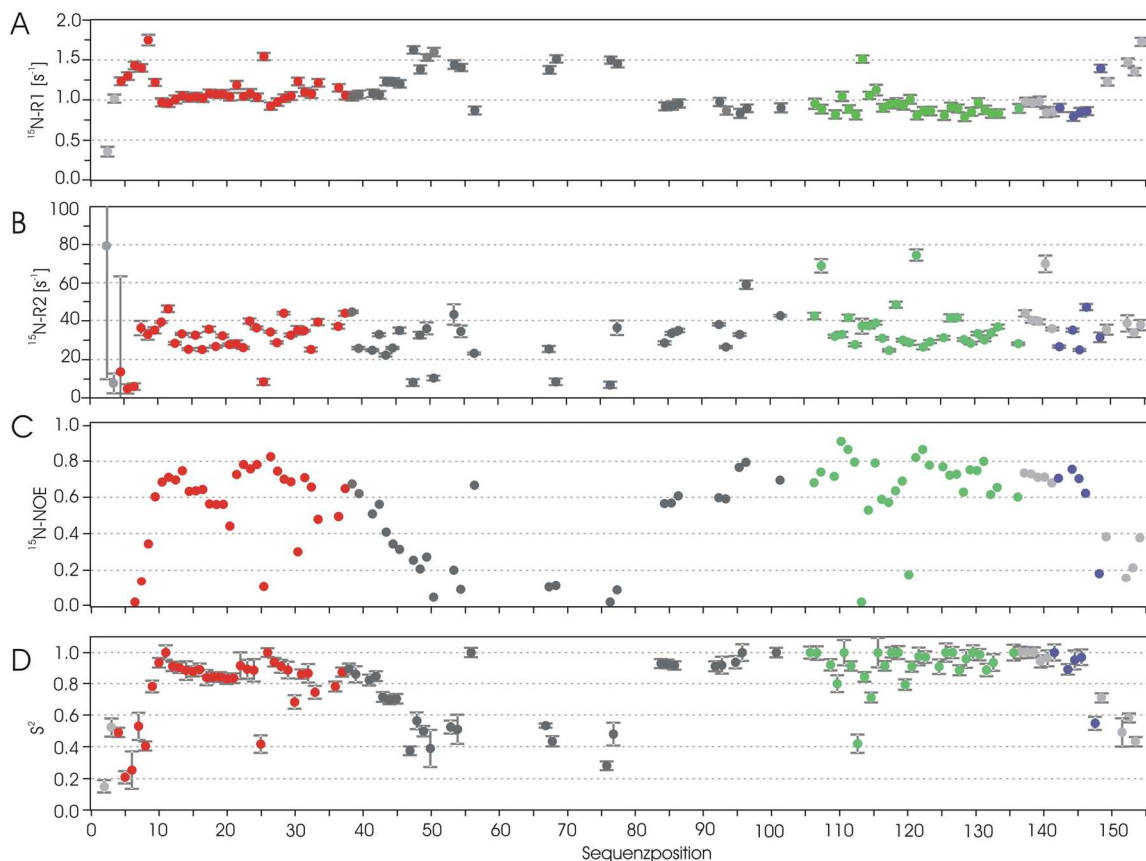
Aminosäuren mit eindeutig zugeordneten  $^1\text{H}^{15}\text{N}$ -Korellationen sind grün, nicht zugeordnete AS rot und Proline grau hinterlegt. Das in der Wildtypsequenz N-terminal vorhandene Leuzin wurde durch Einführung der Erkennungssequenz der TEV-Protease durch Glycin ersetzt. Die laut UniProt-Datenbank (Eintrag: TCRG1\_HUMAN) zu den Domänen gehörenden Aminosäuren sind unterstrichen.

Die sequentielle Zuordnung der Resonanzen im Bereich des Interdomänen-Linkers erwies sich auf Grund der Häufigkeit sehr ähnlicher Sequenzabschnitte und der vielen Proline als schwierig. Zwar konnten viele Resonanzen zu einem Spinsystem verknüpft, der korrespondierende Aminosäuretyp aber nicht eindeutig identifiziert werden. Da die Resonanzen der einzelnen WW-Domänen fast vollständig identifiziert werden konnten, wurde aus zeitlichen Gründen auf weitere Zuordnungsstrategien bzw. NMR-Experimente zur Vervollständigung der Zuordnung des Linkerbereiches verzichtet.



### 4.5.2 STRUKTURDYNAMISCHE EIGENSCHAFTEN VON WW23-F7

Durch die Messung der Longitudinalen (R1) und transversalen (R2) Relaxationsraten der  $^{15}\text{N}$ -Amidstickstoffkerne, sowie durch die Messung der heteronuklearen  $^1\text{H}^{15}\text{N}$ -NOE-Verstärkung ( $^{15}\text{N}$ -HetNOE) sollte die Dynamik des Proteinrückgrates von WW23-F7 in Ab- und Anwesenheit eines interagierenden PP-Peptides näher charakterisiert werden (3.7.3). Auf Basis einer „modellfreien“ Herangehensweise („Model-free-Approach“ [156,159–163]) konnten aus der Kombination dieser Daten strukturdynamische Eigenschaften abgeleitet werden, ohne die Art und Amplitude der der Dynamik zu Grunde liegenden Bewegungsmodi genau zu kennen. Zur Beschreibung der lokalen Dynamik einer Aminosäure innerhalb des Peptidrückgrates wurde der sogenannte „Orderparameter“ ( $S^2$ ) berechnet. Sein Wert stellt ein Maß für die strukturelle Beweglichkeit der betrachteten Aminosäure dar und umfasst einen Wertebereich von 0 (sehr dynamischer Rest) bis 1 (sehr rigider Rest). Die Berechnung von  $S^2$  erfolgte mit dem Programm „Tensor2“ [158]. Die dieser Analyse zu Grunde liegenden, experimentell ermittelten,  $^{15}\text{N}$ -basierten Relaxationsraten bzw.  $^{15}\text{N}$ -HetNOE's und die daraus berechneten Orderparameter sind in Abbildung 26 (WW23) bzw. Abbildung 27 (WW23 mit Peptid) zusammengefasst.



**Abbildung 26:**  $^{15}\text{N}$ -Spinrelaxationsdaten und Orderparameter von WW23-F7

Gezeigt sind Plots der experimentell ermittelten Werte, der longitudinalen (A) und der transversalen (B)  $^{15}\text{N}$ -Spinrelaxationsraten, der  $^1\text{H}^{15}\text{N}$ -NOE's (C) und der berechneten Orderparameter (D). Alle Werte sind gegen die entsprechende Sequenzposition aufgetragen. Die Fehlerbalken entsprechen absoluten Fehlerwerten. Die Werte sind für WW2 rot, für den Interdomänen-Linker dunkelgrau, für WW3 grün und für die C-terminale Helix blau dargestellt. Alle  $^{15}\text{N}$ -basierten Spinrelaxationsdaten wurden an einem 600 MHz Spektrometer bei 280K aufgenommen.

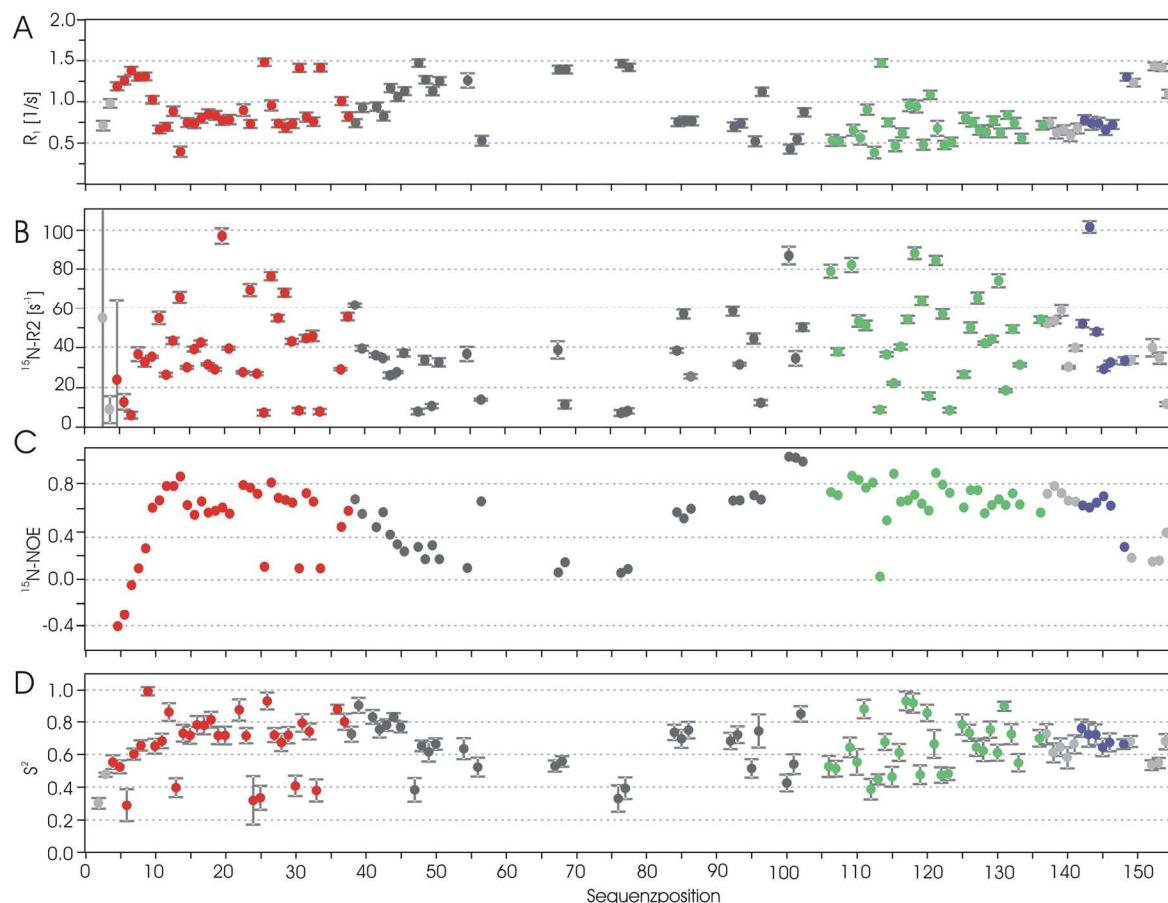


Die Relaxationsdaten ( $^{15}\text{N-R1}$ ,  $^{15}\text{N-R2}$ ,  $^{15}\text{NHetNOE}$ ,  $S^2$ ) von WW23-F7 deuten darauf hin, dass die funktionell zur WW2-Domäne gehörende Aminosäuresequenz C-terminal länger ist, als in der UniProt-Datenbank angegeben (Eintrag: TCRG1\_HUMAN). Die  $S^2$ -Werte von L<sup>38</sup>, K<sup>39</sup>, K<sup>41</sup> und E<sup>42</sup> sind  $>0,8$  und somit einem rigiden Strukturelement zuzuordnen, welches durchaus noch Teil der WW2-Domäne sein könnte. Ab K<sup>43</sup> nimmt die Mobilität deutlich zu ( $S^2 = 0,72$ ; HetNOE = 0,51) und steigt bei den nachfolgenden Aminosäuren weiter an. Das deutet darauf hin, dass sich die C-terminale WW2-Domänengrenze zwischen E<sup>42</sup> und K<sup>43</sup> befindet. Wie der abrupte Anstieg der Rigidität zwischen E<sup>9</sup> ( $S^2 = 0,40$ ) und W<sup>10</sup> ( $S^2 = 0,78$ ) zeigt, liegt die N-terminale Domänengrenze von WW2 genau vor W<sup>10</sup>.

Die N-terminale Domänengrenze von WW3 kann aufgrund der unvollständigen Zuordnung im Bereich von K<sup>81</sup> bis V<sup>99</sup> nicht exakt lokalisiert werden. Wie in Abbildung 26D erkennbar ist, liegen die  $S^2$ -Werte aller in diesem Sequenzabschnitt zugeordneten Aminosäuren (E<sup>84</sup>, E<sup>85</sup>, M<sup>86</sup>, A<sup>92</sup>, A<sup>93</sup>, K<sup>95</sup>, A<sup>96</sup>) jedoch oberhalb von 0,9. Für ein isoliertes, nur 10-15 Aminosäuren langes Sequenzelement ist eine so geringe Mobilität eher unwahrscheinlich, sodass davon auszugehen ist, dass dieser Sequenzbereich fest mit WW3 assoziiert sein muss. Der mittlere  $S^2$ -Wert der definitiv zu WW3 gehörenden Reste (siehe Tabelle 4.5.I), welcher ebenfalls größer als 0,9 ist, bestätigt diese Annahme zusätzlich. Diese Argumentation gilt auch für die C-terminale  $\alpha$ -Helix (L<sup>137</sup>-I<sup>146</sup>), deren mittlerer  $S^2$ -Wert 0,96 beträgt. Demnach scheint der gesamte Sequenzbereich von E<sup>84</sup> - I<sup>146</sup> funktionell der CA150-WW3-Domäne zugeordnet zu sein. Mit Tensor2 wurden für WW23-F7 sowie für die 2 distinkten, auf Basis der Orderparameter eingegrenzten Einzeldomänen zusätzlich die Rotationskorrelationszeiten ( $T_c$ ) berechnet (siehe Tabelle 4.5.I). Da die  $T_c$  die Gesamtdynamik einer jeweils betrachteten Struktureinheit abbildet, sollte deren Vergleich Informationen darüber liefern, ob sich die Domänen unabhängig voneinander bewegen können oder nicht. Für diese Berechnung wurden lt. Tensor2-Anleitung nur Reste mit  $^{15}\text{N-HetNOE}$ 's größer oder gleich 0,7 herangezogen. Da die Struktur von WW23 unbekannt ist, wurde ein isotroper Diffusionstensor angenommen (Kugelmodell). Dabei stellte sich heraus, dass WW3 eine scheinbar längere  $T_c$  besitzt (19,4 ns), als das Gesamtprotein (18,7 ns). Das ist ein physikalisch unsinniges Resultat und bedeutet, dass WW23 keinen isotropen Diffusionstensor besitzen kann und demnach nicht durch ein Kugelmodell zu beschreiben ist.

**Tabelle 4.5.I:  $S^2$  und  $T_c$  von WW23-Strukturelementen**

Sequenzabschnitt	MW	$S^2$	$\sigma$	$T_c$ [ns]	$\sigma$
1-154	Gesamt	0,80	$\pm 0,22$	18,7	$\pm 0,15$
10-42	WW2	0,86	$\pm 0,11$	17,7	$\pm 0,22$
43-83	Linker	0,55	$\pm 0,18$	-	-
84-146	WW3	0,93	$\pm 0,10$	19,4	$\pm 0,20$
137-146	C-terminale Helix	0,97	$\pm 0,04$	-	-



**Abbildung 27:**  $^{15}\text{N}$  Relaxation von WW23 in Interaktion mit CA\_P1

Gezeigt sind Plots der experimentell ermittelten Werte, der longitudinalen (A) und der transversalen (B)  $^{15}\text{N}$ -Spinrelaxationsraten, der  $^1\text{H}^{15}\text{N}$ -NOE's (C) und der berechneten Orderparameter (D). Alle Werte sind gegen die entsprechende Sequenzposition aufgetragen. Die Fehlerbalken entsprechen absoluten Fehlerwerten. Die Werte sind für WW2 rot, für den Interdomänen-Linker dunkelgrau, für WW3 grün und für die C-terminale Helix blau dargestellt. Alle  $^{15}\text{N}$ -basierten Spinrelaxationsdaten wurden an einem 600 MHz Spektrometer bei 283K aufgenommen.

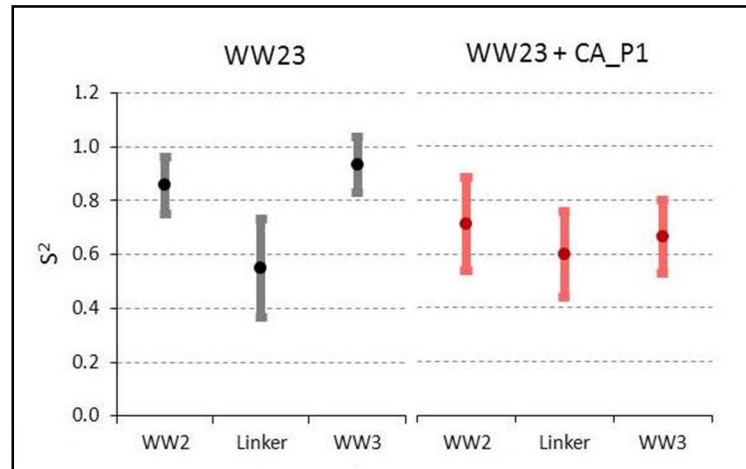
Durch die Zugabe des potentiell multivalenten Peptidliganden CA\_P1 verändert sich die lokale Dynamik innerhalb beider WW-Domänen z.T. erheblich (vgl. Abbildung 27 & Tabelle 4.5.II) und wird im Mittel deutlich größer.

**Tabelle 4.5.II:**  $S^2$  und  $T_c$  von WW23 in Interaktion mit CA\_P1

Sequenzabschnitt	MW	$S^2$	$\sigma$	$\tau_c$	$\sigma$
				[ns]	
1-154	Gesamt	0,66	$\pm 0,16$	17,3	$\pm 0,20$
10-42	WW2	0,71	$\pm 0,17$	14,7	$\pm 0,29$
43-83	Linker	0,60	$\pm 0,16$	-	-
84-146	WW3	0,67	$\pm 0,14$	23,8	$\pm 0,33$

Da sich die Messtemperatur und Zusammensetzung der Proben in den Experimenten mit und ohne Peptid unterscheiden, können die gemessenen Daten jedoch nicht direkt miteinander verglichen werden. Zulässig ist hier nur der Vergleich der relativen

Dynamikunterschiede zwischen WW2, Linker und WW3 der Probe ohne Peptid mit denen der Probe mit Peptid.



**Abbildung 28: Vergleich der Dynamikeigenschaften von WW23 mit und ohne Peptidliganden**

Gezeigt sind die für einzelne WW23-Sequenzbereiche (WW2<sub>(10-42)</sub>, Linker<sub>(43-83)</sub> und WW3<sub>(84-146)</sub>) ermittelten mittleren Orderparameter  $S^2$  (Punkte) inklusive des Bereichs der einfachen Standardabweichung (Balken) bei Abwesenheit (links) und Anwesenheit (rechts) von CA\_P1.

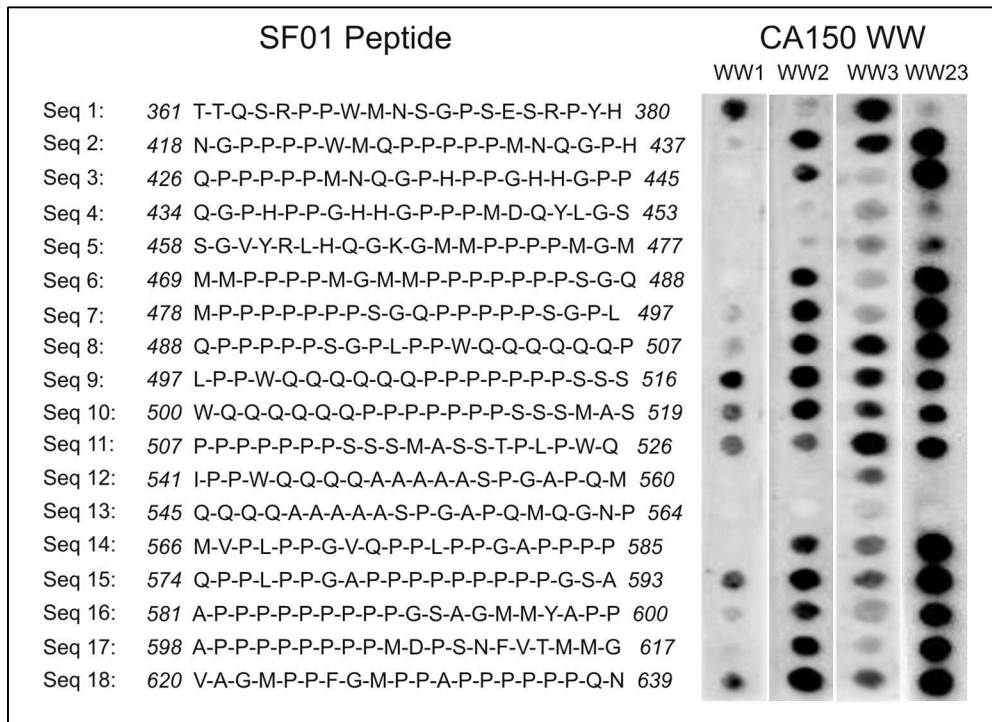
Wie der in Abbildung 28 gezeigte Vergleich der mittleren  $S^2$ -Werte der einzelnen Sequenzelemente deutlich macht, führte die Anwesenheit des Peptides zu einer Angleichung der mittleren lokalen Dynamiken von WW2, Linker und WW3. Das könnte ein Indiz dafür sein, dass die beiden WW-Domänen gleichzeitig mit dem multivalenten CA\_P1 interagieren und dadurch strukturell und dynamisch gekoppelt sind. Dafür spricht auch der etwas höhere mittlere  $S^2$ -Wert des Linkers ( $S^2 = 0,6$  mit Peptid;  $S^2 = 0,55$  ohne Peptid). Leider lässt sich diese These nicht über einen Vergleich der Rotationskorrelationszeiten belegen (siehe Tabelle 4.5.II), da auch der WW23-CA\_P1-Komplex keinen isotropen Diffusionstensor besitzt ( $T_c$ -WW23 <  $T_c$ -WW3).

## **4.6 UNTERSUCHUNG VON MONOVALENTEN UND MULTIVALENTEN SF1<sub>PRS</sub>/CA150<sub>WW</sub>-WECHSELWIRKUNGEN**

Der Spleißfaktor 1 (SF1) ist ein bedeutender Interaktionspartner von CA150 [112]. Die Wechselwirkung beider Proteine stellt eine wichtige Verbindung zwischen der Transkriptionsmaschinerie und dem Spleißapparat der Zelle her [115]. *Goldstrom et al.* haben bereits 2001 gezeigt, dass die Interaktion durch die Bindung prolinreicher Sequenzen aus SF1 an die WW1- und WW2-Domänen aus CA150 vermittelt wird [112], konnten die interagierenden Sequenzmotive aber nicht eindeutig lokalisieren. Zur Eingrenzung der entsprechenden Sequenzmotive wurde in einer systematischen Peptid-SPOT-Analyse daher als erstes geprüft, welche Teilbereiche der SF1<sub>PRS</sub>-Sequenz an welche WW-Domäne binden können. Auf Basis dieser Analyse wurden dann 3 in SF1<sub>PRS</sub> vorkommende Peptidsequenzen und ein künstlich aus 2 prolinreichen Sequenzmotiven zusammengesetztes Peptid ausgewählt und mittels NMR-Titration sowie ITC-Experimenten deren Affinität und Spezifität zu verschiedenen CA150<sub>WW</sub>-Varianten untersucht.

### **4.6.1 SPOT-ANALYSE**

Um alle potentiell in der SF1<sub>PRS</sub>-Sequenz vorkommenden und an mindestens eine der 3 CA150<sub>WW</sub>-Domänen bindenden prolinreichen Sequenzmotive zu erfassen, wurde die SF1<sub>PRS</sub>-Sequenz durch 18, jeweils 20 Aminosäuren lange, sich teilweise N- und C-terminal überlappende Peptidsequenzen auf einer SPOT-Membran abgebildet und mit verschiedenen CA150<sub>WW</sub>-Varianten inkubiert (siehe Abschnitt 3.6.4). Ziel dieser Analyse war es darüber hinaus, mono- und bivalente Peptide zu finden, welche die Tandem-WW-Domäne (WW23) stärker oder schwächer als die Einzeldomänen (WW2 u. WW3) binden und sich daher besonders für das Studium multivalenter Effekte eignen könnten. Da die gebundenen WW-Domänen über eine Kombination aus primärem AntiGST (Z5) Antikörper (Kaninchen) und sekundärem HRP-gekoppelten Anti-Kaninchen IgG Antikörper mittels Chemilumineszenz (*Peroxydase*-Substrat) detektiert wurden, korreliert die Intensität der Spots mit der Bindungsaffinität zwischen Peptid und WW-Variante. In Abbildung 29 ist das Ergebnis der SPOT-Analyse zusammengefasst.



**Abbildung 29: SPOT-Analyse von SF1-prolinreichen Peptiden und CA150-WWs**

Gezeigt ist das Ergebnis der SPOT-Analyse mit prolinreichen Peptiden aus SF1 (links: AS-Sequenzen, Zahlen stehen für die Sequenzposition in SF1). Relative Intensitätsunterschiede zwischen den Spots spiegeln relative Affinitätsunterschiede zwischen den WW-Varianten und den fixierten prolinreichen Peptiden wieder.

Die SPOT-Analyse zeigt, dass entgegen der bisherigen Annahme von *Goldstrohm et al.* [112] alle drei WW-Domänen an die SF1<sub>PRS</sub> binden können. Darüber hinaus wird deutlich, dass die drei Domänen unterschiedliche Bindungspräferenzen besitzen. Drei stärker wechselwirkende Bereiche der SF1<sub>PRS</sub> lassen sich deutlich abgrenzen: SF1<sub>361-445</sub>; SF1<sub>469-526</sub> und SF1<sub>566-639</sub>. WW1 bindet Motive in diesen drei Bereichen nur schwach, wohingegen WW2 und WW3 eine deutlich höhere Affinität, aber auch eine geringere Spezifität aufweisen. Interessanterweise bindet Seq1 gut an WW1 und WW3, aber kaum an WW2 und WW23. Diese Sequenz enthält zwar ein putatives WW-Bindemotiv (RPPW), da die geringe Affinität zu WW23 jedoch ein artifizieller Effekt sein könnte, wurde Seq 1 nicht in weiteren Experimenten eingesetzt. Ein ähnliches Verhalten zeigte auch Seq 12. Dieses Peptid wird nur von WW3, nicht aber von WW23 gebunden und enthält ein ähnliches putatives Bindemotiv (IPPW). Die Sequenzen 2, 3, 5-7, 14, 15 und 18 zeigten putative Multivalenzeffekte. Das heißt, sie binden an WW23 stärker als an die Einzeldomänen. Besonders deutliche Effekte zeigten Seq 3, Seq 5 und Seq 7, sodass diese für weitere Studien ausgewählt wurden (siehe Tabelle 4.6.I). Seq 8 enthält ein bekanntes (PPPPP) und ein putatives (PLPP) Bindemotiv für WW-Domänen, bindet an WW2, WW3 und WW23, zeigt aber keinen Multivalenzeffekt. Um in weiteren Experimenten zu prüfen ob PLPP tatsächlich ein WW-Bindemotiv darstellt, wurde ein analoges, aber leichter zu synthetisierendes (keine polyQ-Sequenz) Peptid kreiert (siehe Tabelle 4.6.I).

**Tabelle 4.6.I: Ausgewählte SF1-PP-Peptide**

Bezeichnung	Sequenz	AS	kDa	GST-		
				WW2	WW3	WW23
SF1_P1	QPPPPPMNQGHPPGHHGPP-NH <sub>2</sub>	426-445	2,070			
SF1_P2	SGVYRLHQGKGMPPPPMGM-NH <sub>2</sub>	458-477	2,169			
SF1_P3	MPPPPPPPSGQPPPPPSGPL-NH <sub>2</sub>	478-497	1,939			
SF1_P4	HPPPPPSHYG <del>L</del> PPWHQK-NH <sub>2</sub> *	n.d.	2,067			

\* Sequenz wurde in Anlehnung an SF1-PP<sub>(488-507)</sub> synthetisiert; \*\* SPOT-Analyse zu SF1<sub>(488-507)</sub>

Die Verwendung von SF1\_P2 gestaltete sich äußerst schwierig, da das Peptid stark hygroskopisch war und bereits während der Einwaage verklumpte. Deshalb wurde auf die weitere Arbeit mit SF1\_P2 verzichtet. In Tabelle 4.6.II sind die Ergebnisse der SPOT-Analyse in Beziehung zu den bekannten WW-Bindemotiven (nach *Otte et al.* [80]) gezeigt.

**Tabelle 4.6.II: Analyse der SPOT-Ergebnisse bezüglich bekannter WW-Bindemotive**

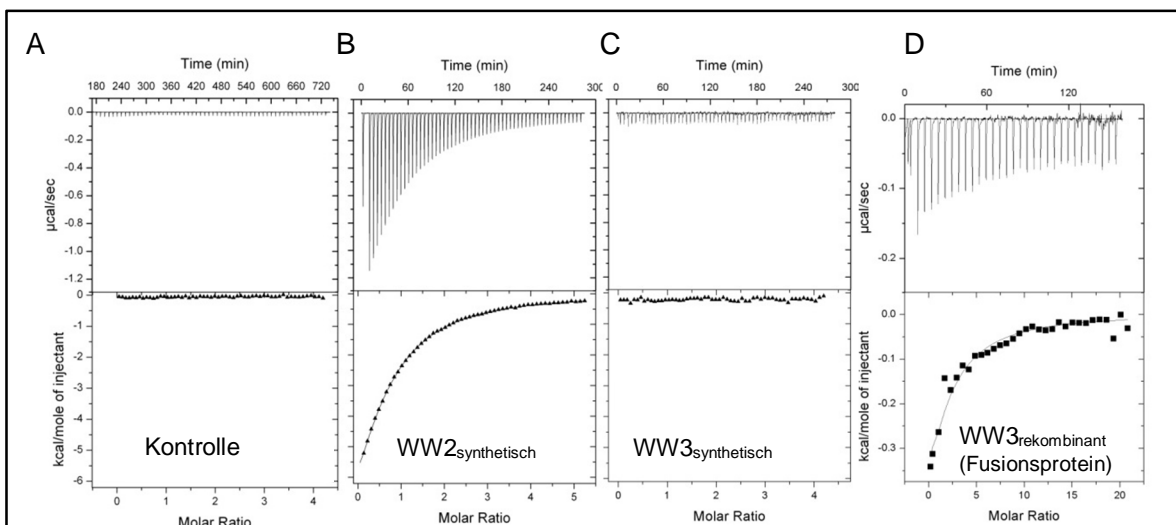
SF1 Peptid	(L/P)Pp(Y/poY)	PPLPp	(p/Φ)P(p/g)PPpR	(p/Φ)PP(R/K)gpPp	(p/Φ)PPPPP	WW1	WW2	WW3	WW23
361-380	-	-	-	-	-	++	+	+++	+
418-437	-	-	-	-	X	-	++	++	+++
426-445	-	-	-	-	X	-	++	+	+++
434-453	-	-	-	-	-	-	-	+	+
458-477	-	-	-	-	-	-	-	+	+
469-488	-	-	-	-	X	-	++	+	+++
478-497	-	-	-	-	X	+	++	+	+++
488-507	-	X*	-	-	X	+	++	++	+++
497-516	-	-	-	-	X	++	++	++	++
500-519	-	-	-	-	X	+	++	+	++
507-526	-	-	-	-	X	+	+	+++	++
541-560	-	-	-	-	-	-	-	+	-
545-564	-	-	-	-	-	-	-	-	-
566-585	-	X	-	-	-	-	++	+	+++
574-593	-	X*	-	-	X	++	++	+	+++
581-600	-	-	-	-	X	-	++	+	++
598-617	-	-	-	-	X	-	++	+	++
620-639	-	-	-	-	X	++	+++	+	++

Analyse der jeweils in den untersuchten Peptiden enthaltenen bekannten WW-Bindemotive (nach *Otte et al.* [80]). Die linke Spalte repräsentiert die Position der Peptidsequenz in SF1. X: Motiv vorhanden, X\*: Motiv reverse vorhanden, (-) Motiv nicht vorhanden, Rechts: (-) keine Reaktion; (+) schwache, (++) mittlere, (+++) stärkere Spotintensität (vgl. Abbildung 29)

Wie die Übersicht zeigt, enthalten die ausgewählten SF1<sub>PRS</sub>-Peptide zwei bekannte Bindemotive für WW-Domänen (PPLPp & (p/Φ)PPPPP). Erstaunlicherweise zeigt sich hier, dass die CA150<sub>WWs</sub> (WW2&WW3) auch das reverse Motiv von PPLPp: PLPP binden. Bekanntermaßen bevorzugt WW2 5er Prolin-Sequenzen ((p/Φ)PPPPP) bindet aber auch an PPR- ((p/Φ)P(p/g)PPpR; (p/Φ)PP(R/K)gpPp) und PPLP(p)-Motive [98]. Diese SPOT-Analyse bestätigt, dass WW2 eine Präferenz für (p/Φ)PPPPP- und PPLP(p)-Motive hat. Die WW3 konnte bisher nicht eindeutig in eine der klassifizierten Gruppen eingeordnet werden, da sie Bindungseigenschaften verschiedener Gruppen aufzeigt (z.B. PPxY) [98]. Es wird daher vermutet, dass sie eher promiskuitiv mit einem breiten Repertoire an prolinreichen Motiven wechselwirken kann. Wie oben bereits angeführt, scheint WW3 aber mit den putativen Bindemotiven IPPW und RPPW zu interagieren.

#### 4.6.1.1 Synthetische Einzel-WW-Varianten (WW2 und WW3)

Da WW2 nur als His-Fusionsprotein und WW3 lediglich in kleinsten Mengen rekombinant hergestellt werden konnte, wurde die Strategie verfolgt, die Einzeldomänen synthetisch herzustellen. Erste ITC-Titrationsexperimente (siehe Abbildung 30) zeigten jedoch nur für WW2<sub>synthetisch</sub>, nicht jedoch für WW3<sub>synthetisch</sub> die erwartete Wechselwirkung zum Peptid SF1\_P3 (vgl. Tabelle 4.6.I).

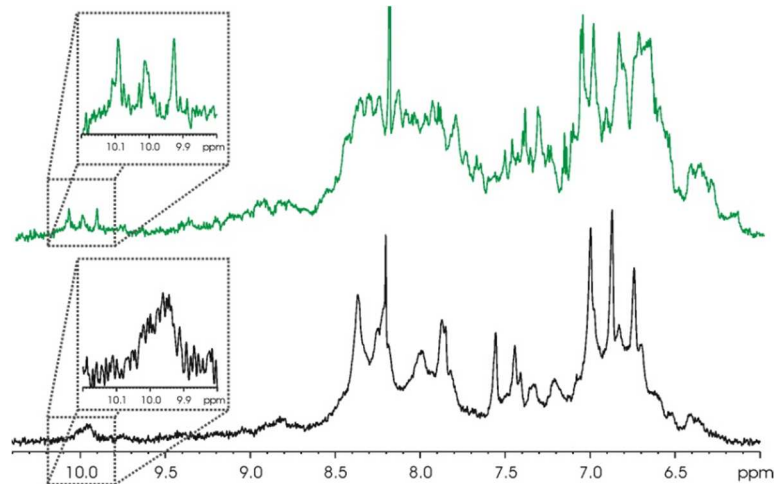


**Abbildung 30: ITC-Interaktionsexperimente mit WW2 und WW3**

Gezeigt sind die Thermogramme und Bindungsisotherme der ITC-Experimente mit SF1\_P3 und Einzel-WW-Varianten. Titriert wurden **A:** 2 mM SF1\_P3 (MPPPPPPPSGQPPPPPSGPL-NH<sub>2</sub>) in Puffer (Kontrollexperiment), **B:** 2 mM SF1\_P3 in 0,2 mM WW2<sub>synthetisch</sub>, **C:** 2 mM SF1\_P3 in 0,2 mM WW3<sub>synthetisch</sub> und **D:** 0,80 mM SF1\_P3 in 0,085 mM WW3<sub>rekombinant</sub>. Die Messungen erfolgten bei 7°C. Die Auswertung erfolgte mittels Origin® Software.

Ein Kontrollexperiment mit einem rekombinant erzeugten WW3-S-His-Fusionsprotein bestätigte, wie zuvor die SPOT-Analyse, dass eine Wechselwirkung zwischen SF1\_P3 und WW3 stattfindet (vgl. Abbildung 30D). Auf Grund der geringen WW3-Konzentration (85 µM) waren die thermodynamischen Effekte klein und das Signal zu Rausch-Verhältnis entsprechend gering.

Da die synthetische WW3 im Gegensatz zur rekombinant hergestellten Domäne nicht aktiv war, wurde der Faltungszustand beider Proteine NMR-spektroskopisch überprüft. Abbildung 31 zeigt die zu diesem Zweck aufgenommenen  $^1\text{H}$ -Spektren.



**Abbildung 31: 1D- $^1\text{H}$ -Spektren der synthetischen und der rekombinanten WW3**

$^1\text{H}$ -Spektrum von WW3<sub>rekombinant</sub> (grün) und WW3<sub>synthetisch</sub> (schwarz). Aufgenommen an einem 600 MHz Spektrometer bei 280K.

Die Spektren beider Proteine zeigen deutliche Unterschiede. Die scharfen Signale im Bereich von 8,5-6 ppm im Spektrum von WW3<sub>synthetisch</sub> deuten auf vorwiegend ungefaltetes Protein hin. Auch die drei gut aufgelösten Signale der Tryptophanseitenketten (um 10 ppm) im Spektrum der rekombinanten WW3 fehlen im Spektrum von WW3<sub>synthetisch</sub>. Lediglich ein schwaches breites Signal ist vorhanden. Das bestätigt die Vermutung, dass die synthetische WW3 einen anderen Faltungszustand einnimmt, als das rekombinante Protein. Damit erklärt sich auch die Inaktivität von WW3<sub>synthetisch</sub> im zuvor durchgeführten ITC-Experiment.

Zum Studium der monovalenten Einzel-WW-Interaktionen wurde daher ein anderer Ansatz verfolgt (siehe Abschnitt 4.6.2.1.)

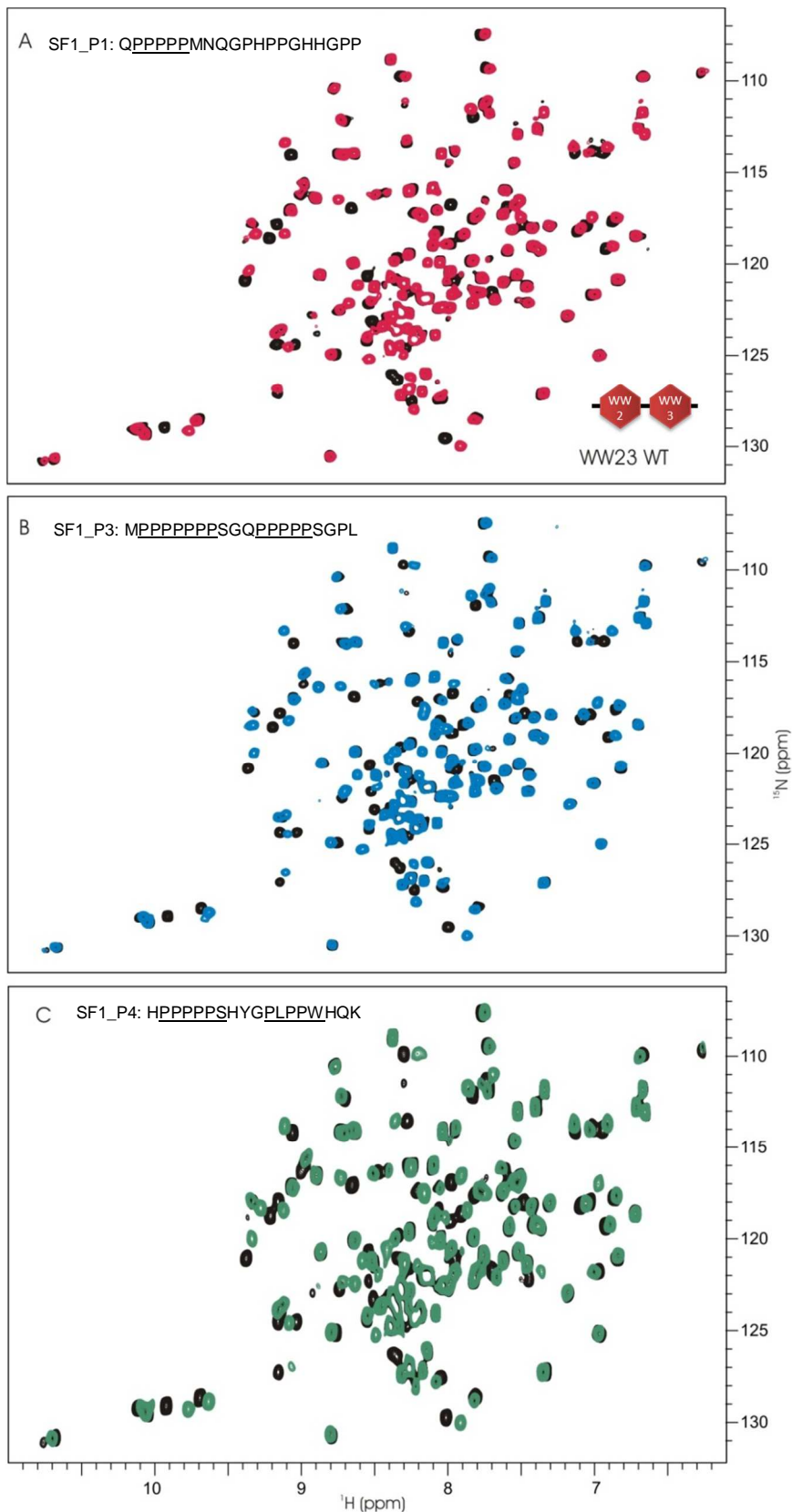


### **4.6.2 $^1\text{H}^{15}\text{N}$ -HSQC-TITRATIONSEXPERIMENTE MIT WW23**

Die Lösungs-NMR-Spektroskopie ist eine etablierte Methode zur Untersuchung von Interaktionen zwischen Proteinen und kleinen Peptid-Liganden, insbesondere, wenn diese Interaktionen eher schwacher Natur sind. Durch den Einsatz uniform  $^{15}\text{N}$ -markierten Proteins, können dessen Signale, sowie eine durch eine Protein-Peptid-Interaktion auftretende Signalveränderung detektiert werden. Da die meisten  $^1\text{H}^{15}\text{N}$ -Korellationen von WW23 zugeordnet werden konnten (siehe Abschnitt 4.5.1), war es möglich die an der Interaktion beteiligten Aminosäuren zu identifizieren. Das Ziel dieser Experimente bestand weniger in der Bestimmung einer Bindekonstante für die einzelnen WW-PRS Interaktionen, sondern vielmehr sollten diese Experimente Aufschluss darüber geben, ob eine oder beide Domänen in WW23 mit dem Peptid wechselwirken. Dazu stellte sich die Frage, ob sich die Peptidzugabe auch auf die  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen von Aminosäuren im Linker-Bereich auswirkt. Veränderungen der chemischen Umgebung des Linkers könnten ein Hinweis für eine strukturelle Veränderung in diesem Bereich sein.

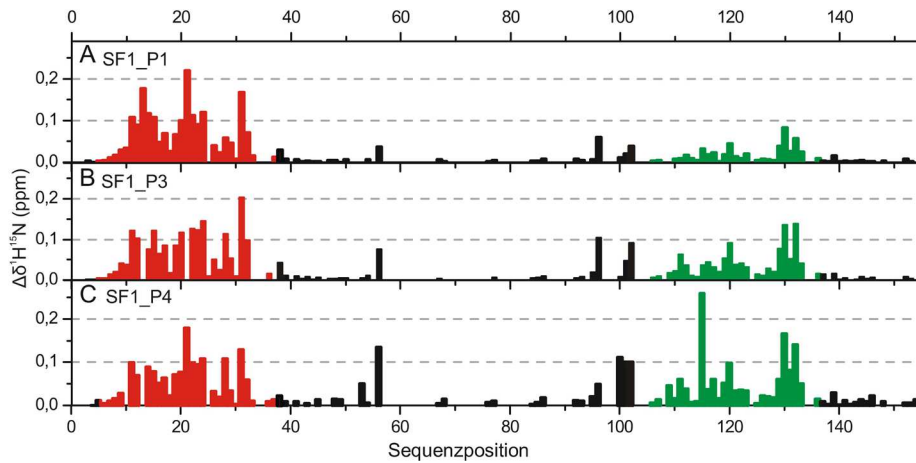
Abbildung 32 zeigt Überlagerungen der  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von  $^{15}\text{N}$ -markierter WW23 ( $c = 50 \mu\text{M}$ ) bei An- und Abwesenheit eines 10-fachen Überschusses an SF1\_P1 (A), SF1\_P3 (B) und SF1\_P4 (C). Hier wird deutlich, dass alle drei Peptide mit WW23 interagieren, da sich die  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen vieler Kreuzsignale signifikant verändern.

Die Größe der peptidinduzierten Signalveränderungen wurde systematisch analysiert (siehe Abschnitt 3.7.2) und auf die Sequenz abgebildet (siehe Abbildung 33).



**Abbildung 32:  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von WW23 in An- und Abwesenheit verschiedener SF1 Peptide**

Überlagerung von  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren aus NMR-Titrationsexperimenten. WW23 ( $c = 50 \mu\text{M}$ ) in Abwesenheit vom Peptidliganden (schwarz), WW23 in Anwesenheit eines 10-fachen Überschusses von SF1\_P1 (rot), SF1\_P3 (blau), SF1\_P4 (grün). Aufgenommen an einem 600 MHz Spektrometer bei 280K.



**Abbildung 33: Analyse der Änderung der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen von WW23 nach Peptidzugabe**

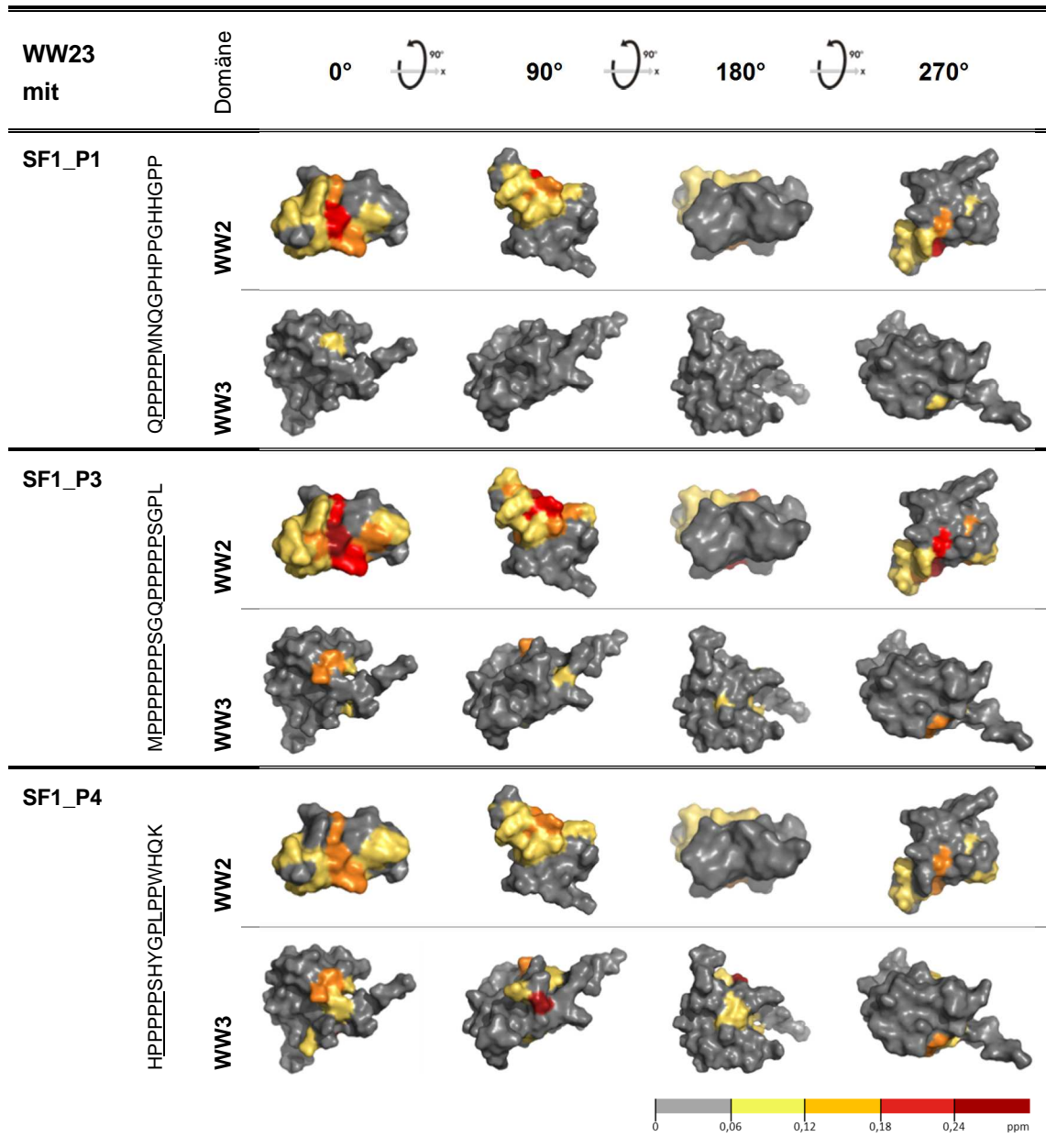
Die peptidinduzierten Differenzen der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen ( $\Delta\sigma^{15}\text{N}$ ) von WW23 Kreuzsignalen wurden gegen die entsprechende Sequenzposition aufgetragen. Die Differenzen der zur WW2<sub>(4-37)</sub>-Kerndomäne gehörenden Resonanzen sind rot, die zur WW3<sub>(103-136)</sub>-Kerndomäne gehörenden Resonanzen grün und alle anderen Resonanzen schwarz dargestellt.

Abbildung 33 ist klar zu entnehmen, dass jeweils beide WW-Domänen, mit den getesteten Peptiden (SF1\_P1, SF1\_P3 & SF1\_P4) interagieren. Darüber hinaus wird deutlich, dass SF1\_P1 (A) und SF1\_P3 (B) die chemische Umgebung der WW2-Domäne deutlich stärker beeinflussen als die der WW3-Domäne. SF\_P1 scheint mit WW3 nur sehr schwach zu interagieren. Diese Beobachtungen sind konform mit den Ergebnissen der SPOT-Analyse (siehe Tabelle 4.6.I).

Interessanterweise zeigt auch das S<sup>56</sup> des Linker-Bereichs bei allen drei Titrationsexperimenten eine starke Änderung seiner Chemischen Verschiebung. Außerdem ist in allen 3 Peptid/WW-Interaktionen eine Änderung der Chemischen Verschiebung N-terminal vor WW3 gelegener Resonanzen (A<sup>100</sup>, T<sup>101</sup>, A<sup>102</sup>) erkennbar. Das deutet, neben den Ergebnissen der Untersuchung der Struktur-dynamischen Eigenschaften von WW23 (Abschnitt 4.5.2) zusätzlich darauf hin, dass die Reste A<sup>100</sup>, T<sup>101</sup> und A<sup>102</sup> Teil der WW3-Domäne sind.

Da die dreidimensionalen Strukturen von WW2 und WW3 bekannt sind (WW2: PDB: 1E0L, WW3: PDB: 2DK7), konnten die peptidinduzierten  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungsdifferenzen direkt auf die entsprechenden Strukturen abgebildet werden.

Die Absolutwerte der Chemischen Verschiebungsdifferenzen wurden dazu in verschiedene Wertebereiche eingeteilt und anhand einer Farbskala dargestellt. Dadurch konnten die Interaktionsstellen zwischen WW-Domäne und Peptid auf der Proteinoberfläche visualisiert werden (siehe Tabelle 4.6.III).

**Tabelle 4.6.III: Stärke der peptidinduzierten Signalverschiebungen**

Gezeigt sind die Strukturen von WW2 (PDB: 1E0L) und WW3 (PDB: 2DK7). Die gewichteten Vektorbeträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}$ - $^{15}\text{N}$  chemischen Verschiebungen der WW23-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala).

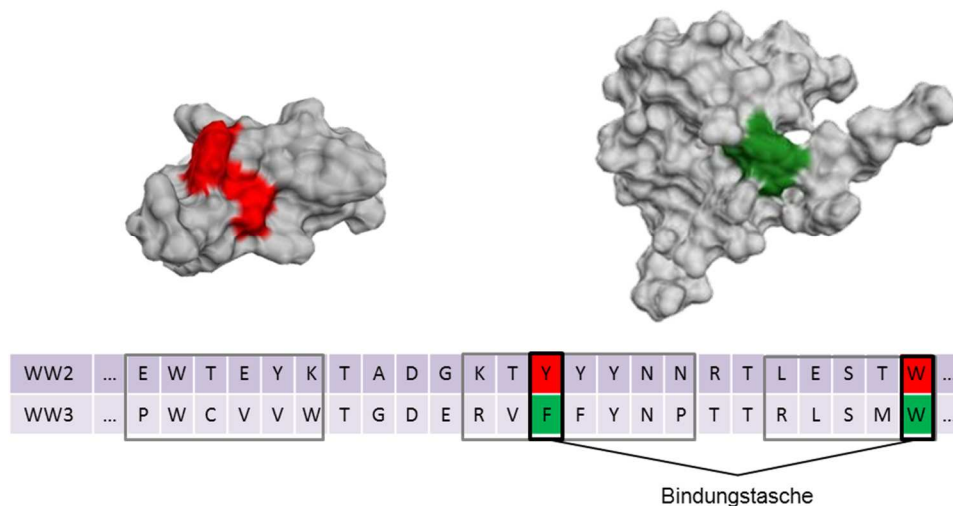
Die Strukturplots in der Tabelle zeigen deutlich, dass alle drei SF1-Peptide (SF1\_P1, SF1\_P3 und SF1\_P4) mit nahezu dem gleichen Oberflächenbereich von WW2 bzw. WW3 wechselwirken. Da die Beträge der peptidinduzierten Signalverschiebungen (siehe Farbskala in Tabelle 4.6.III) in Kombination mit der Zahl der betroffenen Reste (vom Peptid beeinflusste Proteinoberfläche) qualitativ die Stärke der Interaktion widerspiegeln, erlauben die Strukturplots zusätzlich eine Aussage über die Affinität der interagierenden Spezies.

Demzufolge interagieren die Peptide SF1\_P1 und SF1\_P3 deutlich stärker mit WW2 als mit WW3. Die Wechselwirkung beider Peptide zu WW3 scheint zudem relativ unspezifisch zu sein, da jeweils nur ein sehr kleiner Bereich der Domäne an der Bindung beteiligt ist. SF1\_P4 interagiert weniger stark mit WW2, hat aber eine höhere Affinität zu WW3 und scheint mit dieser Domäne auch eine spezifischere Wechselwirkung einzugehen, als die anderen Peptide.

Da diese Untersuchung keine quantitativen Aussagen zur Bindungsstärke ( $K_d$ ) erlauben, können jedoch keine Aussagen über Multivalenzeffekte getroffen werden. Die Affinitäten der Peptidliganden wurden daher in ITC-Experimenten bestimmt (siehe Abschnitt 4.6.4). Um dabei mono- von bivalenten Effekten differenzieren zu können war es zwingend erforderlich, WW23-Mutanten zu erzeugen, in denen jeweils eine der beiden WW-Domänen inaktiv ist.

#### 4.6.2.1 Erzeugung von WW23-Mutanten

WW23-„knockout-Mutanten“ haben gegenüber der Untersuchung der Einzel-WW-Domänen den Vorteil, dass die Affinitäten von WW2 und WW3 separat erfasst werden können, ohne die globale Struktur bzw. die Domänenorganisation von CA150-WW23 und damit die strukturelle Basis potentieller Multivalenzeffekte zu zerstören. Zur Inaktivierung einer WW-Domäne wurde jeweils das charakteristische Tryptophan der „xP“-Bindungstasche (vgl. Abbildung 34) durch Alanin ersetzt (vgl. Abschnitt 3.3.10).



**Abbildung 34: Bindungstaschen WW2 und WW3**

Gezeigt sind die Strukturen von WW2 (PDB: 1E0L, links) und WW3 (PDB: 2DK7, rechts). Die „xP“-Bindungstasche der WW-Domäne wurde in den Strukturmodellen und im Sequenzausschnitt (unten) Rot (WW2) bzw. Grün (WW3) hervorgehoben. Die umrahmten Sequenzabschnitte bilden die drei WW-typischen  $\beta$ -Strang-Bereiche aus.

Die erzeugten WW23-Mutanten (WW23\_W32A und WW23\_W131A) wurden unter vergleichbaren Kultivierungsbedingungen ebenso stark wie der Wildtyp (WW23 WT) exprimiert und die Proteine zeigten in der SDS-PAGE die gleiche Laufverzögerung (vgl.







Die HSQC-Spektren von WW23\_W131A und WW23 WT sind sich ebenfalls sehr ähnlich (siehe Abbildung 39). Im Spektrum der W131A-Mutante fehlen erwartungsgemäß die Kreuzsignale des Tryptophans ( $W^{131}$ ). Die Kreuzsignale verschiedener WW3-Resonanzen sind auf Grund der mutationsbedingt veränderten chemischen Umgebung verschoben (vgl. Abbildung 37, blaue Kreise). Gleichzeitig sind auch Kreuzsignale von Resonanzen verschoben, welche sich N-terminal ( $K^{95}$ ,  $A^{100}$ ,  $T^{101}$ ,  $A^{102}$ ) bzw. C-terminal ( $L^{137}$ ,  $D^{144}$ ,  $I^{146}$ ) vom WW3<sub>(103-136)</sub>-Kernbereich befinden (vgl. Abbildung 37, rote Kreise). Diese Reste sind demnach strukturell eng mit dem WW3-Kernbereich gekoppelt. Das ist, zusätzlich zu dem Ergebnis aus der Untersuchung der strukturdynamischen Eigenschaften von WW23 (siehe Abschnitt 4.5.2), ein weiterer Hinweis dafür, dass diese Reste struktureller Bestandteil von WW3 zu sein scheinen. Analog zur W32A-Mutation, beeinflusst auch die W131A-Mutation das NH-Signal von  $S^{56}$ . Interessanterweise und im Gegensatz zum W32A-Austausch verändert die W131A-Mutation jedoch nicht die Signalintensität, sondern die chemische Verschiebung von  $S^{56}$ . Die W131A-Mutation verändert auch die chemische Verschiebung von 11 weiteren, sequentiell nicht zugeordneten Kreuzsignalen des Linkerbereiches (vgl. Abbildung 37, grüne Kreise) und wirkt sich auch geringfügig auf die chemischen Verschiebungen von  $A^{16}$ ,  $Y^{22}$  und  $W^{32}$  innerhalb der WW2-Domäne aus. Zusammen deuten die betrachteten spektralen Unterschiede zwischen dem Wildtyp und der W131A-Mutante stark darauf hin, dass WW2 und WW3 in Interaktion stehen und dass diese Interaktion durch den Austausch von W131 gegen Alanin verändert wird.

Die NMR-spektroskopische Untersuchung der beiden punktmutierten Proteine zeigte übereinstimmend, dass die Mutationen die globale Struktur zwar wenig beeinflussen, im gesamten Protein aber eine Reihe lokaler Strukturveränderungen bewirken. Da sich die Punktmutation einer WW-Domäne jeweils auch auf den Linker und die andere WW-Domäne auswirken, ist davon auszugehen, dass alle drei Strukturelemente miteinander in Interaktion stehen müssen.

#### **4.6.3 NMR-TITRATIONSEXPERIMENTE MIT CA150 WW23-MUTANTEN**

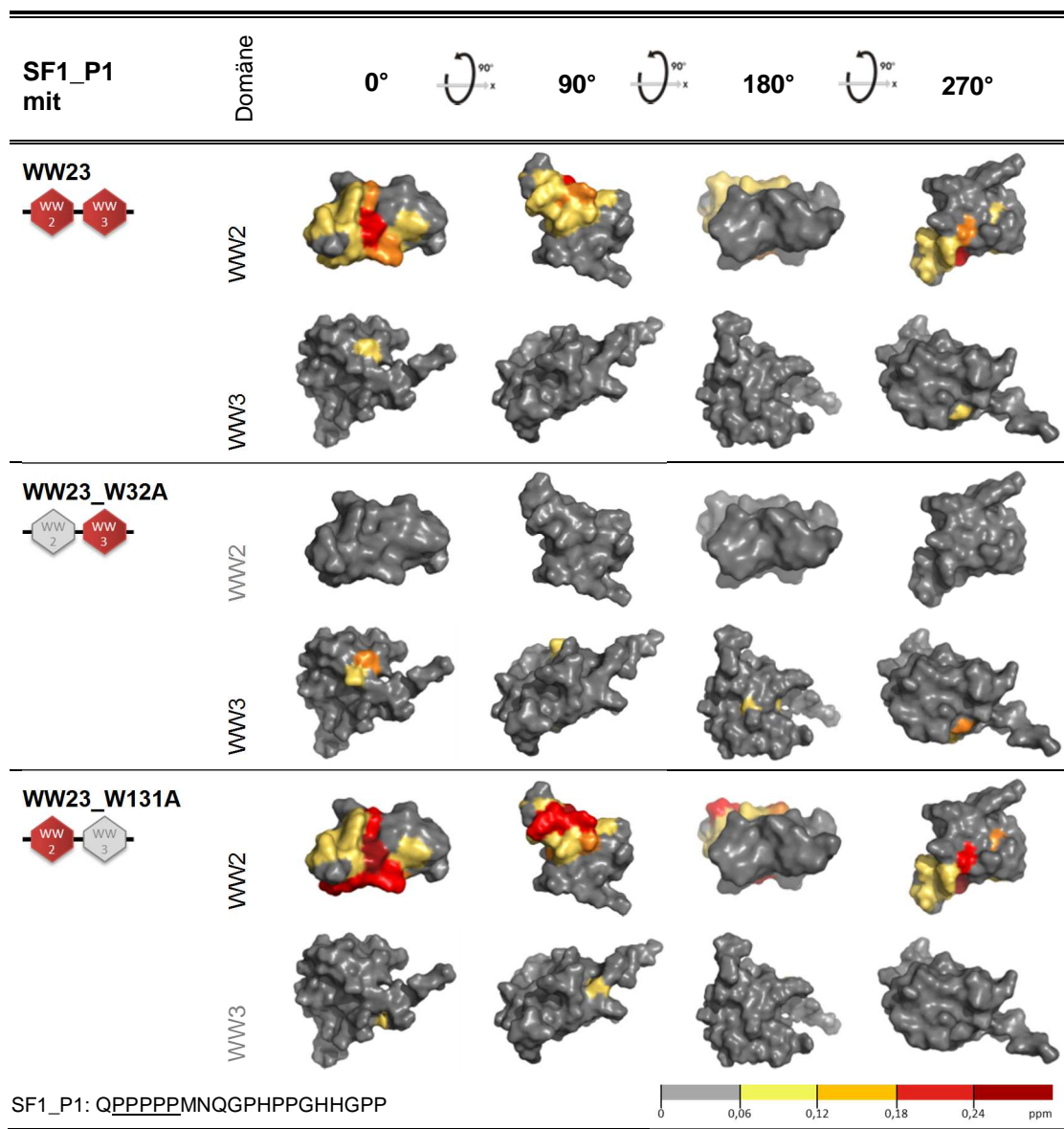
Um die Bindungsfähigkeit bzw. die Inaktivität der mutierten WW-Domänen zu testen, wurden NMR-Titrationsexperimente durchgeführt. Dazu wurden jeweils 2D- $^1H^{15}N$ -Korellationsspektren (HSQCs) von  $^{15}N$ -markierter WW23-M (WW23\_W32A oder WW23\_W131A, je  $c = 50 \mu M$ ) bei An- und Abwesenheit eines 10-fachen Überschusses an SF1\_P1, SF1\_P3 bzw. SF1\_P4 aufgenommen (siehe Anhang Abbildung 52 & Abbildung 53).

Die Größe der peptidinduzierten Veränderungen der  $^1H_N^{15}N$  chemischen Verschiebungen wurde, wie in Abschnitt 3.7.2 beschrieben, systematisch analysiert, Intensitätsveränderungen von  $^1H^{15}N$ -Kreuzsignalen wurden jedoch nicht betrachtet. Da die dreidimensionalen Strukturen von WW2 und WW3 bekannt sind (WW2: PDB:1E0L, WW3: PDB: 2DK7), konnten die peptidinduzierten  $^1H_N^{15}N$  chemischen Verschiebungsdifferenzen direkt auf die entsprechenden Strukturen abgebildet werden. Dazu wurden die



Differenzen der gewichteten NH-Vektoren (siehe Abschnitt 3.7.2) in verschiedene Wertebereiche eingeteilt und anhand einer Farbskala dargestellt. Dadurch konnten die Interaktionsstellen zwischen WW-Domäne und Peptid auf der Proteinoberfläche visualisiert und entsprechend die Aktivität der nicht mutierten und die Inaktivität der mutierten WW-Domäne veranschaulicht werden (siehe Tabelle 4.6.IV-Tabelle 4.6.VI).

**Tabelle 4.6.IV: Stärke der SF1\_P1-induzierten Signalverschiebungen**

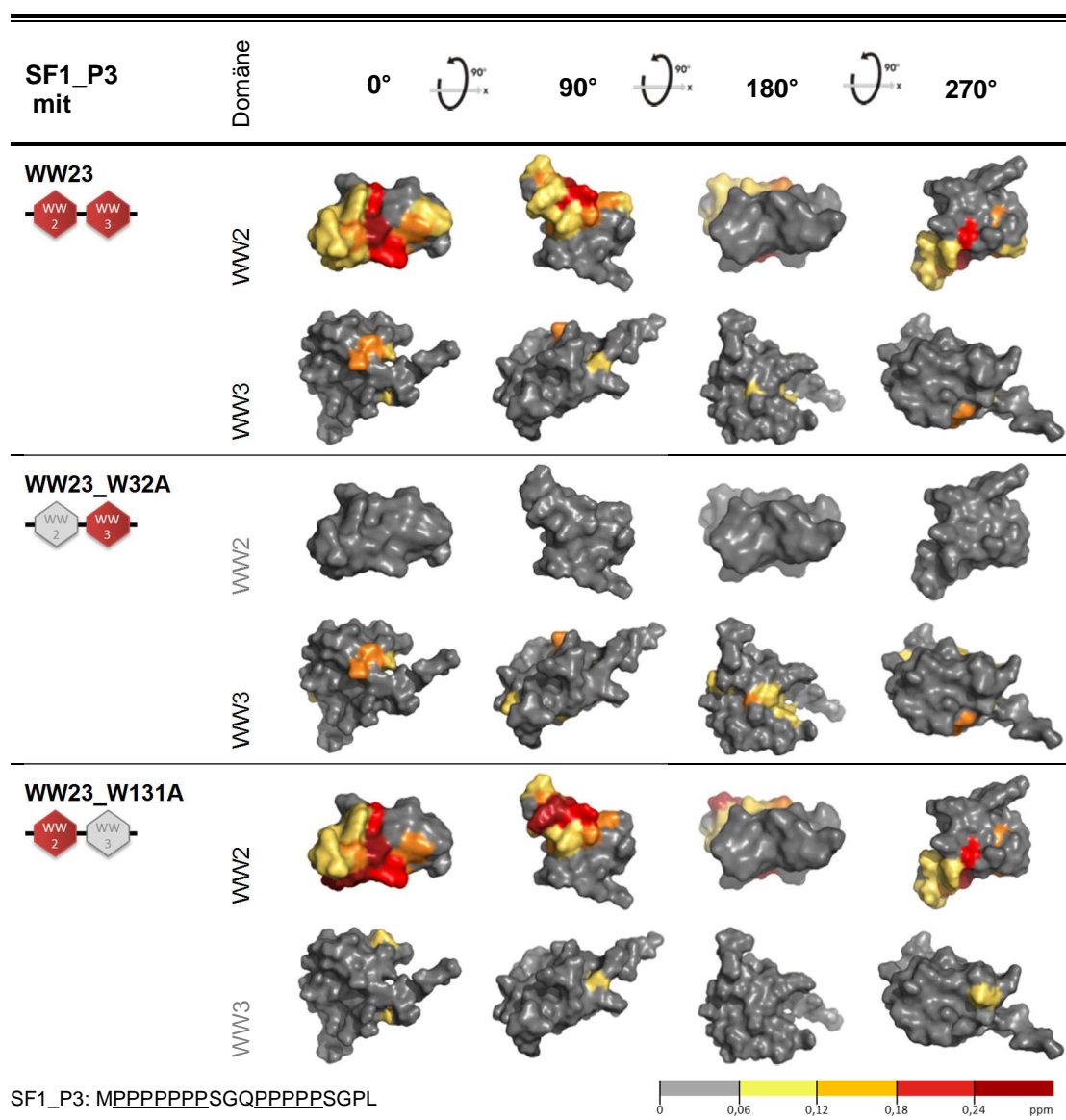


Gezeigt sind die Strukturen von WW2 (PDB: 1E0L) und WW3 (PDB: 2DK7). Die gewichteten Vektorbeträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen der WW23-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala).

Da die verwendeten Farben im Prinzip die Stärke der Interaktion widerspiegeln (siehe Farbskala in Tabelle 4.6.IV), erlauben die Strukturplots zusätzlich eine qualitative Aussage über die Affinität der interagierenden Spezies. Die Strukturplots in der Tabelle zeigen deutlich, dass SF1\_P1 mit der gleichen Oberfläche der funktionellen WW2 in

WW23 und WW23\_W131A wechselwirkt, die Affinität zur WW2 in WW23\_W131A jedoch stärker ist. Die Inaktivierung von WW3 in WW23\_W131A scheint sich positiv auf die Affinität zwischen WW2 und SF1\_P1 auszuwirken. Ähnliches gilt für WW3. Durch die Mutation W32A, wird die Bindung von WW2 an SF1\_P1 verhindert (vgl. Tabelle 4.6.IV, WW23\_W32A), gleichzeitig verstärkt sich, im Vergleich zur Wildtypsituation, die Interaktion zur WW3 leicht. Die Strukturplots der WW3 der WW23\_W131A-Mutante zeigen lediglich einen kleinen Interaktionsbereich zum Peptid. Da dieser nicht mit der charakteristischen SF1\_P1-Interaktionsoberfläche (vgl. Tabelle 4.6.III) übereinstimmt, kann davon ausgegangen werden, dass die Bindungsfähigkeit von WW3 praktisch vollständig ausgeschaltet werden konnte (siehe auch Anhang, Abbildung 55).

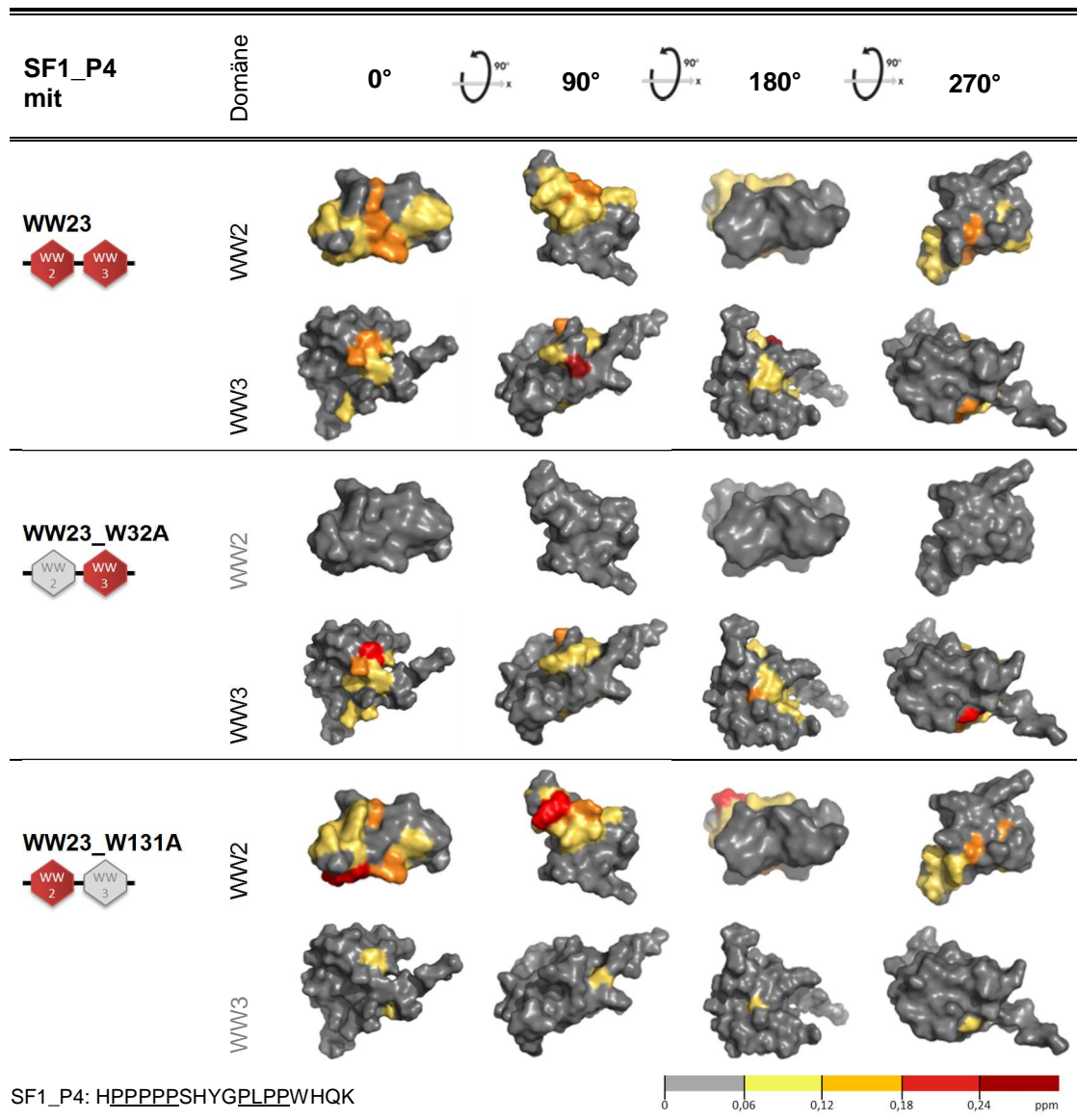
**Tabelle 4.6.V: Stärke der SF1\_P3-induzierten Signalverschiebungen**



Gezeigt sind die Strukturen von WW2 (PDB: 1E0L und WW3 (PDB: 2DK7). Die gewichteten Vektorbeträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}$ / $^{15}\text{N}$  chemischen Verschiebungen der WW23-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala).

Ähnliches gilt für die Wechselwirkung der WW3 in WW23\_W131A mit SF1\_P3. Dabei ist anzumerken, dass eine kleine Interaktionsfläche zum Peptid bestehen bleibt, diese aber wesentlich kleiner ist als in WW23 (vgl. Tabelle 4.6.V). Die WW2 in WW23\_W32A ist dagegen nicht mehr in der Lage an SF1\_P3 zu binden, dafür verstärkt sich die Affinität der WW3 von WW23\_W32A zu SF1\_P3 gegenüber der Wildtypsituation. Die WW2 in WW23\_W131A bindet ebenfalls etwas besser an SF1\_P3. So führen die Mutationen auch bei der Wechselwirkung mit SF1\_P3 zu leicht erhöhten Affinitäten der bindungsaktiven WW-Domänen (siehe auch Anhang, Abbildung 56).

**Tabelle 4.6.VI: Stärke der SF1\_P4-induzierten Signalverschiebungen**



Gezeigt sind die Strukturen von WW2 (PDB: 1E0L) und WW3 (PDB: 2DK7). Die gewichteten Vektorbeträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen der WW23-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala).

In der Interaktion mit SF1\_P4 zeigen beide Mutanten im Vergleich zum Wildtyp (WW23) leicht veränderte Bindungsmuster. Durch die Mutation und Inaktivierung einer WW-

Domäne wird die Affinität der funktionellen WW-Domäne zu SF1\_P4 erhöht, aber das Peptid bindet anders an die WW-Domänen-Oberfläche.

Die Bindung von SF1\_P4 an WW2 in WW23\_W32A konnte vollständig unterdrückt werden. Demgegenüber besteht noch eine sehr schwache Wechselwirkung zwischen SF1\_P4 und der WW3 von WW23\_W131A. Vergleicht man diese aber mit der zur WW3 in WW23, so kann von einem nahezu vollständigen Funktionsverlust der WW3 ausgegangen werden (siehe auch Abbildung 57 im Anhang).

Zusammenfassend lässt sich festhalten, dass durch die jeweilige Mutation (WW2: W32A, WW3: W131A) die entsprechende WW-Domäne praktisch vollständig inaktiviert werden konnte. WW23\_W32A und WW23\_W131A sind daher uneingeschränkt für die kalorimetrische Bestimmung der Bindungsstärken ( $K_d$ ) zwischen der jeweils noch intakten WW-Domäne (WW2 bzw. WW3) und den SF1-Peptiden nutzbar.

#### **4.6.4 ISOTHERME TITRATIONS-KALORIMETRIE**

In allen chemischen Reaktionen wird Energie in Form von Wärme erzeugt oder verbraucht, man spricht dabei von exothermen bzw. endothermen Reaktionen. Die Isotherme Titrationskalorimetrie (ITC) bietet durch die direkte Detektion dieser Reaktionswärmen Zugang zu thermodynamischen Parametern wie der Reaktionsenthalpie ( $\Delta H$ ), der Reaktionsentropie ( $\Delta S$ ) und der freien Gibbsenergie ( $\Delta G^\circ$ ) sowie der Stöchiometrie und der Dissoziationskonstante interagierender Spezies. Vorteil der Methode ist, dass die zu untersuchenden Spezies chemisch nicht modifiziert oder isotonenmarkiert werden müssen und frei in Lösung diffundieren können.

Dass die SF1-Peptide SF1\_P1 und SF1\_P3 sowie das künstliche, aus der SF1-PRS-Sequenz abgeleitete, potentiell bivalente Peptid SF1\_P4 mit beiden WW-Domänen in CA150-WW23 interagieren, konnte bereits durch die NMR-Titrations in Abschnitt 4.6.2 gezeigt werden. Unklar blieb dabei jedoch, ob die Peptide separat bzw. monovalent an die WW-Domänen binden oder ein Peptidmolekül mit beiden Domänen gleichzeitig und damit bivalent wechselwirkt. Bei bekannter Aktivität der eingesetzten Spezies (Protein und Peptid) und bekanntem Reaktionsmechanismus, wäre diese Frage mittels ITC-Experimenten prinzipiell zu beantworten. Da jedoch die Aktivität der eingesetzten Proteine und Peptide unbekannt war und sich deren aktive Konzentration auf Grund der schwachen Wechselwirkungen nicht zufriedenstellend aus den Bindungsisothermen ableiten ließen, erlaubten die ITC-Experimente in diesem Fall jedoch keine verlässlichen Aussagen über die tatsächliche Reaktionsstöchiometrie. Erschwerend kommt hinzu, dass bei der Betrachtung der stöchiometrischen Verhältnisse auch berücksichtigt werden müsste, dass monovalente und bivalente Interaktionen gleichzeitig auftreten und somit zu einem Gemisch verschiedener Protein-Peptidkomplexe führen können. Denkbar wäre die Bildung der Komplexe:  $(WW2_{\text{Peptid}})\text{-WW3}$ ,  $(WW2_{\text{Peptid}})\text{-(WW3}_{\text{Peptid}})$ ,  $WW2\text{-(WW3}_{\text{Peptid}})$ ,  $(WW23_{\text{Peptid}})$ . Aus diesen Gründen konnten für die Unterscheidung von mono- und bivalenten Interaktionen nur die Dissoziationskonstanten herangezogen werden. Ausschlaggebendes Kriterium für die Annahme einer bivalenten Wechselwirkung, war




eine signifikant höhere Affinität zwischen Peptid und Wildtypprotein, als zwischen Peptid und den WW23-Mutanten.

Da der genaue Reaktionsmechanismus unbekannt war, und auch kein Modell zur Verfügung stand, mit welchem eine Kombination aus mono- und bivalenten Interaktionen beschrieben werden konnte, wurden alle Bindungsisotherme mit einem 1 zu 1 Modell gefittet. Die kalorimetrischen Messungen wurden entsprechend der in Abschnitt 3.9. dargestellten Vorschrift durchgeführt und unter Verwendung der Programme Origin 7.0 (Origin Lab®) und SEDPHAT 9.4 [134–140] ausgewertet.

Alle WW-Peptid Interaktionen sind sehr schwach mit Bindekonstanten im mikromolaren Bereich. Da die Reaktionswärmen sehr gering waren, wurden höhere Proteinmengen (100-400  $\mu\text{M}$ ) und Peptidkonzentrationen (bis 3 mM) als für die ITC üblich, eingesetzt. So konnte ein  $K_d$ -Bereich bis zu einem oberen Limit von ca. 1 mM abgedeckt und das Signal zu Rausch-Verhältnis optimiert werden. Auf Grund der schwachen Interaktionen konnte oftmals jedoch keine vollständige Proteinsättigung erreicht werden. Darüber hinaus erzeugten die hohen Peptidkonzentrationen z.T. große Verdünnungswärmen. Entsprechend war für die Auswertung der Messdaten die SEDPHAT-Software vorteilhafter, da die Verdünnungswärme/-kälte dort als zusätzlicher Fit-Parameter einbezogen werden konnte. Zudem konnten mit SEDPHAT durch Monte-Carlo-Analyse die Konfidenzintervallgrenzen für  $K_d$  und  $\Delta H$  bestimmt werden. Da die Bindungsisotherme wegen des unvorteilhaften c-Wertes nur unvollständig durch Messwerte abgebildet werden konnten, sind die  $\Delta H$ - und  $\Delta S$ -Werte mit einer großen Unsicherheit behaftet und nur der Vollständigkeit halber angegeben.

Die Resultate der Titrationsexperimente sind in den Tabelle 4.6.VII bis Tabelle 4.6.IX dargestellt.

**Tabelle 4.6.VII: ITC-Ergebnisse der WW-SF1\_P1 Interaktion**

SF1_P1: QPPPPPMNQGPHPPGHHGPP	$\Delta G$ [kcal/mol]	$\Delta H$ [kcal/mol]	$-T\Delta S$ [kcal/mol]	$K_d$ [ $\mu\text{M}$ ]
 WW23	-4	-16 -(21-13)	12	452 (352-582)
 WW23_W32A	n.B.	n.B.	n.B.	>> 1.000
 WW23_W131A	-4	-13 -(18-10)	9	807 (620-1.052)

Die Best-Fit Werte und die 95% Konfidenzintervalle (in Klammern) der Titrationen von WW23; WW23\_W32A und WW23\_W131A mit SF1\_P1. Alle Bindungsisotherme einer Titrationsreihe wurden Global gefittet und per Monte Carlo Analyse anschließend die Konfidenzintervalle ermittelt.




Die Ergebnisse der WW/SF1\_P1-Titration in Tabelle 4.6.VII zeigen deutlich, dass SF1\_P1 an WW2 bindet ( $K_d = 807 \mu\text{M}$ ), aber nur sehr schwach mit WW3 wechselwirkt ( $K_d \gg 1 \text{ mM}$ ). Das ist konsistent mit den Resultaten aus der SPOT-Analyse (Tabelle 4.6.I) und der NMR-Titration (Tabelle 4.6.IV). Ebenfalls konform mit der SPOT-Analyse bindet SF1\_P1 deutlich besser an den Wildtyp ( $K_d$ :  $452 \mu\text{M}$ ) als an WW23\_W131A (WW2).



Die NMR-Titration ließ zwar darauf schließen, dass das Peptid eher schwächer an den Wildtyp bindet als an WW23\_W131A, indizierte aber auch, dass SF1\_P1 möglicherweise unspezifisch bzw. sehr schwach mit WW3 wechselwirkt. Die hier ebenfalls detektierte, sehr schwache WW3/SF1\_P1-Interaktion könnte daher Teil einer bivalenten Wechselwirkung und damit die Ursache für die beobachtete Affinitätssteigerung sein.

Denkbar wäre, dass dieser schwache Beitrag die Wahrscheinlichkeit einer korrekten Orientierung zwischen WW2 und SF1\_P1 erhöht und dadurch deren Assoziation begünstigt.

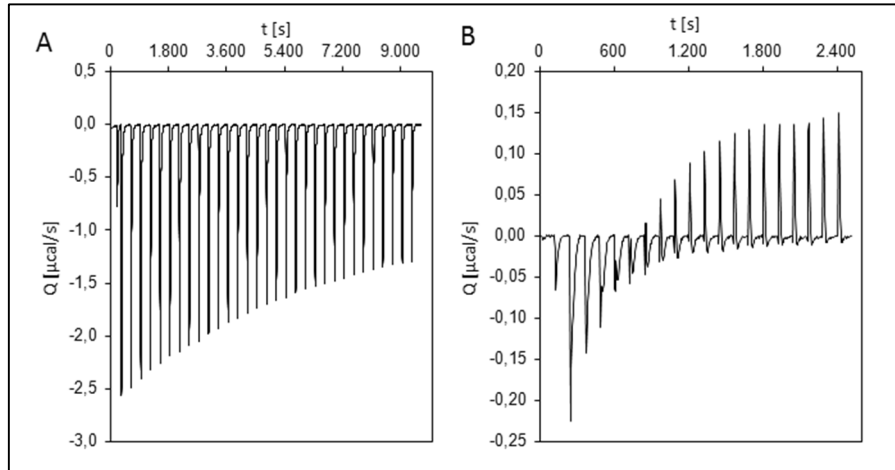
**Tabelle 4.6.VIII: ITC-Ergebnisse der WW-SF1\_P3 Interaktion**

SF1_P3: MPPPPPPPSGQPPPPPSGPL	$\Delta G$ [kcal/mol]	$\Delta H$ [kcal/mol]	$-T\Delta S$ [kcal/mol]	$K_d$ [ $\mu M$ ]
 WW23	-5	-7 -(8-7)	2	55 (49-60)
 WW23_W32A	-4	-13 -(16-11)	9	634 (506-808)
 WW23_W131A	-5	-11 -(13-10)	6	123 (104-148)

Die Best-Fit Werte und die 95% Konfidenzintervalle (in Klammern) der Titrations von WW23; WW23\_W32A und WW23\_W131A mit SF1\_P3. Alle Bindungsisotherme einer Titrationsreihe wurden Global gefittet und per Monte Carlo Analyse anschließend die Konfidenzintervalle ermittelt.

Die Ergebnisse der WW/SF1\_P3-Titration in Tabelle 4.6.VIII zeigen deutlich, dass SF1\_P3 den Wildtyp deutlich affiner bindet ( $K_d = 55 \mu M$ ) als die einzelnen WW-Domänen (WW2:  $K_d = 123 \mu M$ ; WW3:  $K_d = 634 \mu M$ ). Das ist konsistent mit den Resultaten aus der SPOT-Analyse (Tabelle 4.6.I) und der NMR-Titration (Tabelle 4.6.V). Bei dieser Wechselwirkung handelt es sich mit großer Wahrscheinlichkeit um eine bivalente Interaktion.

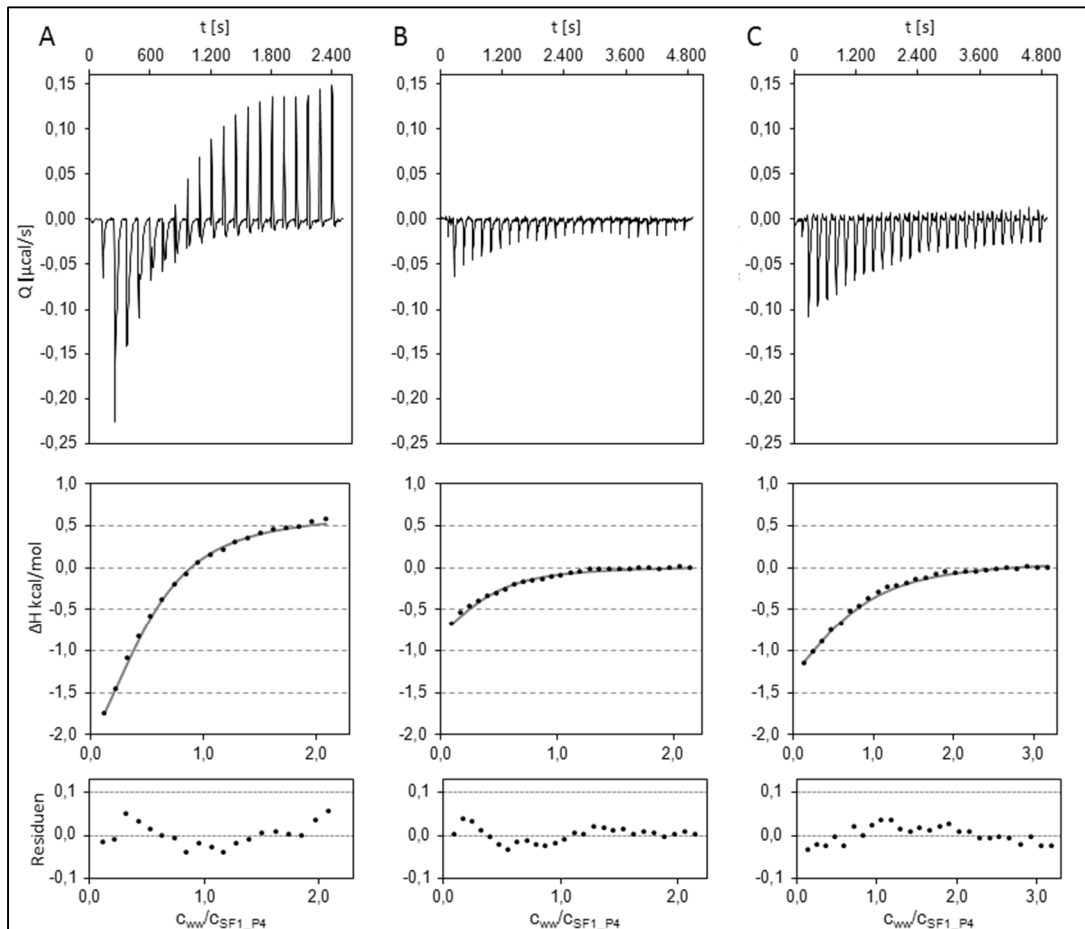
Im Gegensatz zu den beiden zuvor betrachteten Peptiden gestaltete sich die kalorimetrische Untersuchung der WW/SF1\_P4-Interaktion bedeutend schwieriger. Erste Titrationsexperimente mit SF1\_P4 (HPPPPPSHYGPLPPWHQK) zeigten zum einen sehr starke Verdünnungseffekte, zum anderen konnte das Protein nicht vollständig gesättigt werden (siehe Abbildung 38A).



**Abbildung 38: Thermogramm der „klassischen“ und der reversen WW23+SF1\_P4 Titration**

Thermogramm der Titration von **A** 2 mM SF1\_P4 in 0,2 mM WW23 und **B** 1,5 mM WW23 in 0,15 mM SF1-P4 (reverse Titration). Gemessen wurde an der VP-ITC (A) und an der ITC<sub>200</sub> (B) bei 7°C.

Das Thermogramm dieser Titration ließ sich nicht zufriedenstellend auswerten, und ergab eine unverhältnismäßig schwache Bindekonstante von 393  $\mu\text{M}$ . Diese fiel deutlich schwächer aus als die NMR-Titrationsexperimente (vgl. Tabelle 4.6.VI) vermuten ließen. Offensichtlich wurden die Reaktionswärmen durch die starken Verdünnungseffekte stark verfälscht. Um dieses Problem zu umgehen wurden die Titrationsexperimente dann revers durchgeführt, wodurch die Verdünnungseffekte stark reduziert werden konnten. Die Titration des Wildtyps in SF1\_P4 zeigte dennoch reproduzierbar ein ungewöhnliches Thermogramm. Nachdem die Reaktion zunächst exotherm verlief, wurde ab einer WW/Peptid-Stöchiometrie von 1 dem System plötzlich Energie entzogen und die Reaktion schien endotherm zu verlaufen (vgl. Abbildung 38B und Abbildung 39). Interessanterweise konnte dies nicht bei der Bindung an die Mutanten beobachtet werden (vgl. Abbildung 39).






**Abbildung 39: ITC-Plots der Interaktionen von WW23(-M) mit SF1\_P4**

Gezeigt sind die Thermogramme (oben), Bindungsisotherme (Mitte) und die Residuen (unten) aus den Titrations von (A) 1,5 mM WW23, (B) 1,5 mM WW23\_W32A und (C) 1,78 mM WW23\_W131A in 0,15  $\mu$ M SF1\_P4. Gemessen wurde an der ITC<sub>200</sub> bei 7°C

Die Ergebnisse der globalen Auswertung der entsprechenden Messreihen sind in Tabelle 4.6.IX zusammengefasst.

**Tabelle 4.6.IX: ITC-Ergebnisse der WW-SF1\_P4 Interaktion**

SF1_P4: HPPPPPSHYG <del>LP</del> PWHQK	$\Delta G$ [kcal/mol]	$\Delta H$ [kcal/mol]	$-T^*\Delta S$ [kcal/mol]	$K_d$ [ $\mu$ M]
 WW23_rev	-6	-4 -(4,7-3,5)	-2	39 (30-50)
 WW23_W32A_rev	-6	-1 -(1,6-1,2)	-4	39 (32-48)
 WW23_W131A_rev	-5	-2 -(2,6-2,2)	-3	59 (51-69)

Die Best-Fit Werte und die 95% Konfidenzintervalle (in Klammern) der Titrations von WW23; WW23\_W32A und WW23\_W131A mit SF1\_P4. Alle Bindungsisotherme einer Titrationsreihe wurden Global gefittet und per Monte Carlo Analyse anschließend die Konfidenzintervalle ermittelt.

Wie die Daten in Tabelle 4.6.IX zeigen, bindet SF1\_P4 mit vergleichbarer Affinität an WW2 (WW23\_W131A:  $K_d = 59 \mu$ M), WW3 (WW23\_W32A:  $K_d = 39 \mu$ M) und an das Wildtypprotein (WW23:  $K_d = 39 \mu$ M). Eine bivalente Wechselwirkung zwischen Peptid und Wildtypprotein scheint somit ausgeschlossen.



### 4.7 TRIPLE SILAC PULLDOWN-EXPERIMENT

Um die Frage zu beantworten, ob WW23 im Kontext des zellulären Proteoms auf Grundlage multivalenter Effekte veränderte Bindungseigenschaften besitzt, als die Einzel-WW-Domänen (WW2 & WW3), wurden Triple-SILAC-basierte Pulldown Experimente durchgeführt.

Diese Proteomanalyse beruht auf einer massenspektrometrischen Quantifizierung von spezifisch an WW-Domänen bindenden Proteinen. Dazu wurden HeLa-Zelllysate nach dem Triple-SILAC-Prinzip (*Stable isotope labeling with amino acids in cell culture*) [165,170] metabolisch markiert und deren zelluläres Proteom im Pulldown spezifisch gegen die WW2, WW3, und WW23 Domäne von CA150 angereichert.

Die co-präzipitierten WW-Ligand-Komplexe wurden in einer SDS-PAGE aufgetrennt, die Proteinbanden anschließend ausgeschnitten, tryptisch verdaut und einer MS-Analyse unterzogen (siehe Abschnitt 3.12). Die Anreicherungsverhältnisse wurden aus der massenspektrometrischen Quantifizierung der identifizierten Proteine abgeleitet. Als Grundlage der Anreicherung eines Markierungszustandes (Light/Medium/Heavy), sprich der verstärkten Bindung der entsprechenden WW-Variante galt ein Verhältnis von 3:1.

**Tabelle 4.7.I: Anreicherungsverhältnis der WW-Interaktionspartner SILAC-Pulldowns**

	WW2>WW3	WW2=WW3	WW2<WW3
WW23>Einzel-WW	0	2	0
WW23=Einzel-WW	3	>800	2
WW23<Einzel-WW	26	150	108

Das Ergebnis dieser Untersuchung zeigte deutlich, dass entgegen den Erwartungen lediglich zwei (NAV1; CA150) von über 1.000 identifizierten und quantifizierten Proteinen eine höhere Affinität zur Doppel-WW-Variante WW23 besitzen (siehe Tabelle 4.7.I). Interessanterweise ist eines dieser Proteine CA150 selbst. Ein Vergleich der gefundenen Peptid-Sequenzen aus CA150 mit der Sequenz der „Köder“-WW-Domänen, ergab keine Übereinstimmungen. Demnach befand sich CA150 während der Co-Präzipitation entweder im Komplex mit einem der CA150<sub>WW</sub>-Interaktionspartner oder in direkter Interaktion mit einer „Köder“-WW-Domäne. Letzteres würde bedeuten, dass die CA150-WW-Domänen an die CA150<sub>PRS</sub> binden.

Die SILAC-Analyse zeigte weiterhin, dass 110 quantifizierte Proteine eine höhere Affinität zur WW3 haben und nur 29 eine Präferenz für WW2 zeigen. Beim Großteil der quantifizierten Proteine (>800) ließ sich jedoch keine Präferenz für WW2 oder WW3 beobachten.

**Tabelle 4.7.II: Mittels SILAC-Pulldown identifizierte bekannte und "neue" CA150-WW-Interaktionspartner**

ID	Name	Verhältnis		
		WW2/WW23	WW3/WW23	WW3/WW2
CHD1	<i>Chromodomain-helicase-DNA-binding protein 1</i>	2,8	0,5	0,4
CHD3	<i>Chromodomain-helicase-DNA-binding protein 3</i>	10,9	1,0	0,2
CHD4	<i>Chromodomain-helicase-DNA-binding protein 4</i>	4,3	3,3	0,9
CHD8	<i>Chromodomain-helicase-DNA-binding protein 8</i>	2,8	0,5	0,4
HNRNPL	<i>Heterogeneous nuclear ribonucleoprotein L</i>	9,9	7,3	1,1
HNRNPM	<i>Heterogeneous nuclear ribonucleoprotein M</i>	10,3	8,0	0,9
HTT	<i>Huntingtin</i>	45,1	4,1	0,2
POLR2A	<i>DNA-directed RNA polymerase II polypeptide A</i>	2,6	2,5	1,1
POLR2B	<i>DNA-directed RNA polymerase II subunit RPB2</i>	1,6	1,6	1,1
RPB1	<i>RNA-polymerase Rpb1</i>	2,6	2,5	1,1
SF1	<i>Splicing factor 1</i>	38,8	10,6	0,9
SF1 (IF5)	<i>Splicing factor 1 isoform 5</i>	18,0	15,4	1,0
SF3A1	<i>Splicing factor 3a, subunit 1</i>	8,8	11,0	1,0
SF3A2	<i>Splicing factor 3a, subunit 2</i>	15,4	15,0	0,9
SF3A3	<i>Splicing factor 3a, subunit 3</i>	38,0	30,0	0,8
SF3B1	<i>Splicing factor 3b, subunit 1</i>	7,3	4,9	1,0
SF3B2	<i>Splicing factor 3b, subunit 2</i>	10,4	7,0	0,9
SF3B3	<i>Splicing factor 3b, subunit 3</i>	7,1	4,1	1,0
SF3B4	<i>Splicing factor 3b, subunit 4</i> [115]	14,6	14,9	1,1
SF3B5	<i>Splicing factor 3b, subunit 5</i>	42,8	27,4	1,0
SF4	<i>Splicing factor 4</i>	1,6	1,5	1,0
CA150 (TCERG)	<i>Transcription elongation regulator 1</i>	0,02	0,02	1,0

Durch die Proteomanalyse konnten 11 (siehe Tabelle 4.7.II, blau hinterlegt) der 20 bekannten (vgl. Abschnitt 1.5 Tabelle 1.5.I) CA150-Interaktionspartner quantifiziert werden. Zusätzlich konnten Interaktionen zu 12 funktionell ähnlichen Proteinen nachgewiesen werden (siehe Tabelle 4.7.II, grau hinterlegt). Dazu gehören die drei „*Chromodomain-helicase-DNA-binding*“-Proteine (CHD1, 3 und 4), ein weiteres „*Heterogeneous nuclear ribonucleo*“-Protein (HNRNPL), eine RNA-Polymerase (RPB1), sowie eine Untereinheit der RNA-Polymerase II (POLR2B) und vier weitere Spleißfaktoren (SF3B3, SF3B5, SF3B4 und SF4).

**Tabelle 4.7.III: SILAC-Bindepräferenzen der CA150-Interaktionspartner für WW2, WW3 und WW23**

	WW2	WW3	WW23
HTT	+++	++	+
CHD4, HNRNPL, HNRNPM, SF1, SF1 (IF5), SF3A1, SF3A2, SF3A3, SF3B1, SF3B2, SF3B3, SF3B4, SF3B5	++	++	+
CHD3	++	+	+
CHD1, CHD8, POLR2A, POLR2B, RPB1, SF4	+	+	+
CA150	+	+	++

Der Vergleich der Bindepräferenzen dieser Proteine in Tabelle 4.7.III zeigt deutlich, dass abgesehen von CA150 keines der Proteine besser an die Tandem-WW (WW23) als an die Einzel-WWs bindet. Auch hinsichtlich der Präferenz für WW2 oder WW3 zeigen die meisten Proteine (darunter viele der Spleißfaktoren, SF's) keine Unterschiede. Interessanterweise zeigen die Polymerasen (POLR2A, POLR2B, RBP1) für jede der drei WW-Domänen die gleiche Präferenz.

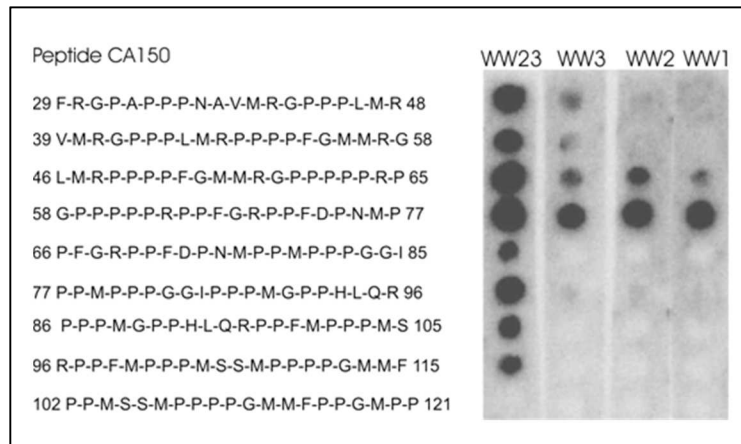
#### **4.8 STUDIEN ZUR CA150<sub>PRS</sub>/CA150<sub>WW</sub> INTERAKTION**

Mittels SILAC-MS (Abschnitt 4.7) wurde CA150 als möglicher Bindepartner seiner eigenen WW-Domänen identifiziert. Um die Frage zu klären, ob CA150 mit Hilfe seiner N-terminalen PRS (CA150<sub>PRS</sub>) in der Lage ist seine eigenen WW-Domänen zu binden, wurden verschiedene Peptide der CA150<sub>PRS</sub> ausgewählt und auf eine Cellulosemembran synthetisiert. SPOT-Analysen sollten den wechselwirkenden Bereich der CA150<sub>PRS</sub> näher definieren, um anschließend eine multivalente Peptidsequenz zur Synthese auszuwählen. Mit Hilfe der NMR-Spektroskopie sollte die Interaktionsstellen und mittels ITC die Affinität der mono- und potentiell bivalenter Wechselwirkung bestimmt werden.

##### **4.8.1 SPOT-ANALYSE**

Durch die SPOT-Analyse sollte geklärt werden, ob und an welchen Bereich der CA150<sub>PRS</sub> die WW-Domänen binden. Hierzu wurden neun teilweise überlappende 20 Aminosäuren lange Peptidsequenzen der CA150<sub>PRS</sub> auf eine Zellulosemembran synthetisiert (AG Beyermann).

Es wurden WW1-, WW2-, WW3- und WW23- GST-Fusionsproteine eingesetzt, welche zu den membran-fixierten CA150<sub>PRS</sub>-Sequenzen gegeben wurden (siehe Abschnitt 3.6.4). Indem die gebundenen WW-Domänen über eine Kombination aus primärem AntiGST (Z5) Antikörper und sekundärem HRP-gekoppelten Antikörper mittels Chemilumineszenz (*Peroxydase*-Substrat) detektiert wurden, konnten die Bindungsaffinitäten der WW-Liganden semi-quantitativ bestimmt werden.



**Abbildung 40: SPOT-Analyse mit prolinreichen Peptiden der CA150<sub>PRS</sub> und CA150<sub>WW</sub>-Domänen**

Gezeigt ist das Ergebnis der SPOT-Analyse, mit PP-Peptiden der N-terminalen CA150<sub>PRS</sub> (links: AS-Sequenzen, Zahlen stellen die Sequenzposition in CA150 dar). Die Intensität der Spots (rechts) ist ein relatives Maß für die Bindungsstärke der Interaktion zwischen der WW-Domäne und dem fixierten PP-Peptid.

Die SPOT-Analyse (siehe Abbildung 40) bestätigt eindeutig die Bindung der drei CA150-WW-Domänen (WW1, WW2, und WW3) und des WW23-Tandems an die N-terminale CA150<sub>PRS</sub>. Der Sequenzbereich zwischen 29-58 scheint multivalent mit dem WW23-Tandem zu interagieren, da dieser lediglich schwach mit WW3 aber sehr stark mit WW23 interagiert. Ein Teil der PRS-Sequenz (46-77) bindet auch an alle drei Einzel-Domänen, zeigt aber eine bedeutend höhere Affinität zum WW23-Tandem, was ebenfalls für einen multivalenten Interaktionsmechanismus spricht. An den Sequenzbereich 66-115 kann erstaunlicher Weise nur das WW23-Tandem binden. Auch das ist ein starker Hinweis darauf, dass dieser Sequenzbereich bzw. die darin enthaltenen Sequenzabschnitte ausschließlich auf Basis einer multivalenten Wechselwirkung gebunden werden. Um den Mechanismus der beobachteten multivalenten Wechselwirkung qualitativ (Interaktionsstellen) und thermodynamisch (Affinitäten) besser verstehen zu können, wurde für weitere Untersuchungen ein PRS-Bereich ausgewählt, welcher mono- als auch multivalente Bindungsmuster zeigte (siehe Tabelle 4.8.I).

**Tabelle 4.8.I: Für NMR- und ITC-Studien ausgewähltes CA150-Peptid**

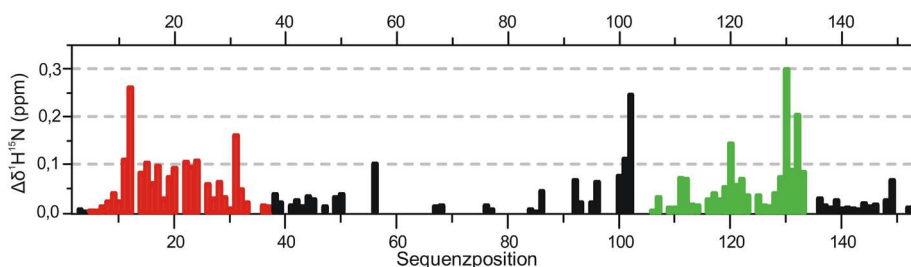
Bezeichnung	Sequenz	AS	kDa
CA_P1	HNAVMRGPPLMRPPPPFEGMMRGPPPPRPPFG-NH <sub>2</sub>	37-68	3,694

#### 4.8.2 NMR-TITRATIONSEXPERIMENTE

Mittels NMR-Spektroskopie wurde untersucht, welche Aminosäuren des WW23-Tandems mit CA\_P1 in Interaktion stehen bzw. von der Interaktion betroffen sind. Mit Hilfe der bereits in Abschnitt 4.6.2.1 vorgestellten WW23-Mutanten sollten zusätzlich Informationen darüber gewonnen werden, über welche Interaktionsstellen die per Peptid-SPOT nachgewiesenen monovalenten Interaktionen realisiert werden. Hierzu wurden <sup>1</sup>H<sup>15</sup>N-HSQC von <sup>15</sup>N-markierter WW23 bzw. den beiden WW23-Mutanten (c = 50 µM) bei An-

und Abwesenheit von CA\_P1 ( $c = 500 \mu\text{M}$ ) aufgenommen (siehe Anhang Abbildung 58 & Abbildung 59).

Die  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren aller untersuchten Proteinspezies zeigten bei einer Reihe von Kreuzsignalen signifikante Veränderungen der  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen. Bei einigen Kreuzsignalen verringerte sich zudem die Intensität. Letzteres könnte ein Hinweis auf konformativen Austausch sein. Im Folgenden wird jedoch nur auf die peptidinduzierten Veränderungen der  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen eingegangen. Zur Visualisierung wurden diese zum einen gegen die Sequenzposition aufgetragen, zum anderen auf die bekannten Strukturen von WW2 und WW3 abgebildet. Abbildung 41 und Tabelle 4.8.II fassen die entsprechenden Daten aus der Untersuchung des Wildtypproteins sowie der beiden WW23-Mutanten zusammen.

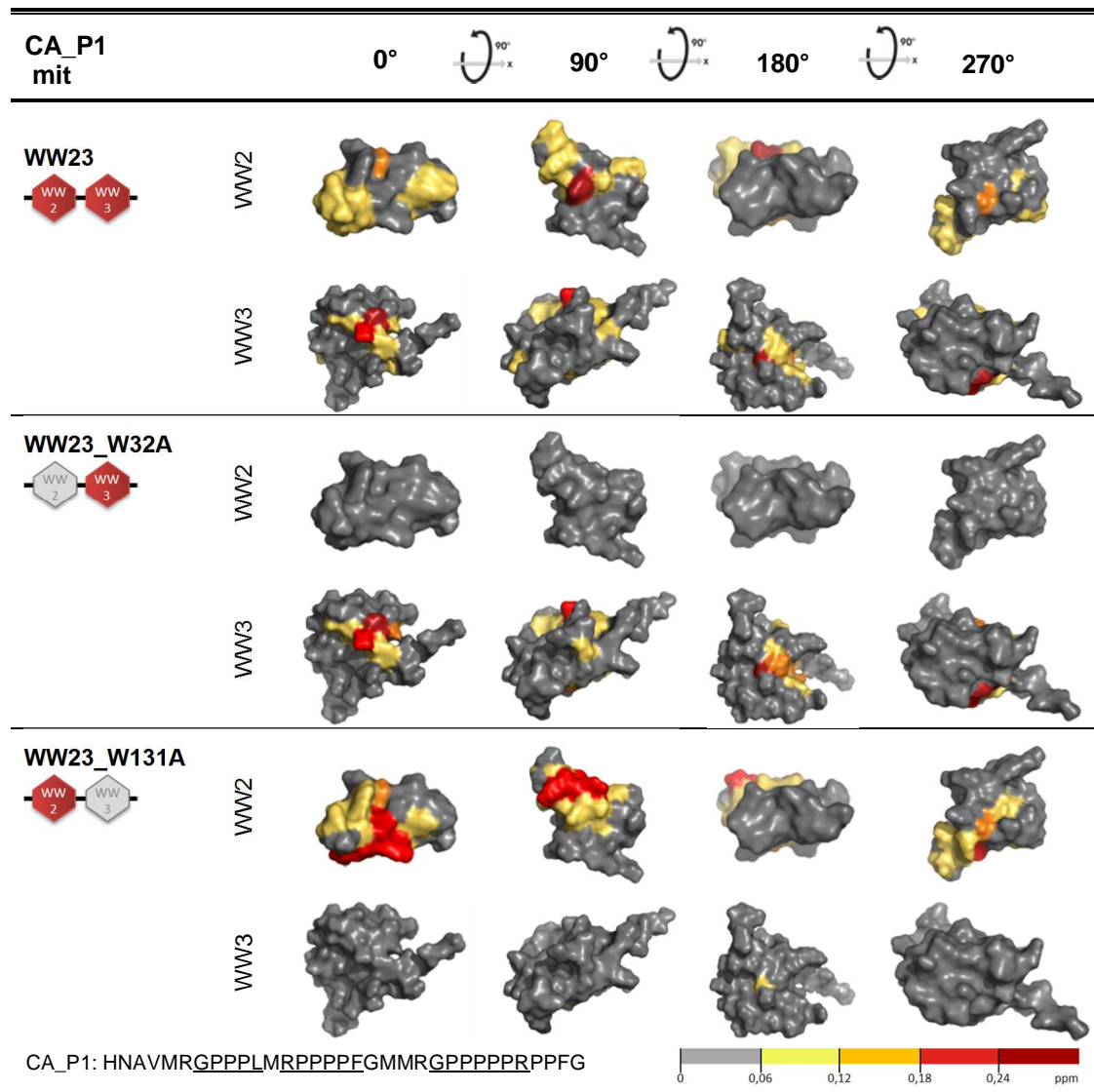


**Abbildung 41: CA\_P1-induzierte  $^1\text{H}^{15}\text{N}$  chemische Verschiebungsdifferenzen in WW23**

Die peptidinduzierten Differenzen der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen von WW23 Kreuzsignalen wurden in gewichtete NH-Vektoren ( $\Delta\sigma^1\text{H}^{15}\text{N}$ ) umgerechnet und gegen die entsprechende Sequenzposition aufgetragen. Die Werte der WW2-Kerndomäne (WW2<sub>(4-37)</sub>) sind rot, die der WW3-Kerndomäne (WW3<sub>(103-136)</sub>) grün und die aller anderen Resonanzen schwarz dargestellt.

Abbildung 41 ist klar zu entnehmen, dass CA\_P1 mit beiden WW-Domänen interagiert. Diese Beobachtungen sind konform mit den Ergebnissen der SPOT-Analyse (siehe Tabelle 4.6.I). Nahezu alle Reste innerhalb der beiden Kerndomänen zeigen eine signifikante Signalverschiebung ( $\Delta\sigma^1\text{H}^{15}\text{N} > 0,06 \text{ ppm}$ ). Auch die Kreuzsignale der N-terminal vor der WW3-Kerndomäne lokalisierten Aminosäuren (A<sup>92</sup>, A<sup>96</sup>, A<sup>100</sup>, T<sup>101</sup>, A<sup>102</sup>) verschieben sich signifikant. Wie in Abschnitt 4.5.2 beschrieben, scheinen diese Aminosäuren mit großer Wahrscheinlichkeit noch zur WW3-Domäne zu gehören. Interessanterweise verändert sich auch die chemische Umgebung von S<sup>56</sup> im Linker-Bereich. Da die dreidimensionalen Strukturen von WW2 und WW3 bekannt sind (WW2: PDB:1E0L, WW3: PDB: 2DK7), konnten die peptidinduzierten  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungsdifferenzen direkt auf die entsprechenden Strukturen abgebildet werden. Die Absolutwerte der Chemischen Verschiebungsdifferenzen wurden dazu in verschiedene Wertebereiche eingeteilt und diese anhand einer Farbskala dargestellt (siehe Tabelle 4.8.II).

Tabelle 4.8.II: Stärke der CA\_P1-induzierten Signalverschiebungen



Gezeigt sind die Strukturen von WW2 (PDB: 1E0L) und WW3 (PDB: 2DK7). Die gewichteten Vektorbeträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}$   $^{15}\text{N}$  chemischen Verschiebungen der WW23-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala)

Die Plots in Tabelle 4.8.II verdeutlichen zum einen, dass die mutierten Domänen keine Bindungsaktivität mehr besitzen und zeigen zum anderen, dass CA\_P1 mit WW2 und WW3 monovalent und mit dem WW23-Tandem monovalent oder bivalent interagieren kann. Deutlich wird auch, dass CA\_P1 mit der WW2-Domäne über andere Oberflächenareale wechselwirkt, wenn WW3 inaktiviert wurde. D.h. bei inaktiver WW3 bindet das Peptid über die klassischen Bindetaschen „xP“ und „xP2“ (siehe Abschnitt 5.5) an WW2. Bei der Interaktion mit dem Wildtyp werden diese Bindetaschen jedoch interessanterweise nicht genutzt. Dieser Unterschied könnte ein ernst zu nehmendes Indiz für eine bivalente Wechselwirkung zwischen WW23 und CA\_P1 sein. Die Interaktion zwischen CA\_P1 und WW3 richtet das Peptid so zur WW2-Domäne aus, dass es an diese vermutlich nur schwach binden kann, die Summe beider Wechselwirkungen scheint




aber stärker oder energetisch günstiger zu sein als die monovalente Interaktion zwischen Peptid und WW2.

### 4.8.3 ISOTHERME TITRATIONS-KALORIMETRIE (ITC)

Um die vermutlich bivalente Wechselwirkung zwischen WW23 und CA\_P1 zu verifizieren, wurde die Stärke der Wechselwirkung zwischen Peptid und WW23, WW23\_W32A sowie WW23\_W131A kalorimetrisch bestimmt. Die Annahme bestand darin, dass die bivalente Wechselwirkung signifikant stärker wäre, als eine monovalente Interaktion.

Alle untersuchten WW-Peptid Interaktionen sind schwach mit Bindekonstanten im mikromolaren Bereich. Da die Reaktionswärmen sehr gering waren, wurden höhere Proteinmengen (100-380  $\mu\text{M}$ ) und Peptidkonzentrationen (bis 4 mM) als für die ITC üblich, eingesetzt. So konnte ein  $K_d$ -Bereich bis zu einem oberen Limit von ca. 1 mM abgedeckt und das Signal zu Rausch-Verhältnis optimiert werden. Tabelle 4.8.III fasst die Ergebnisse der Messreihen zusammen.

**Tabelle 4.8.III: Ergebnisse WW/CA\_P1-Interaktionen**

CA_P1: HNAVMRGPPPLMRPPPPEGMRRGPPPPRRPPFG	$\Delta G$ [kcal/mol]	$\Delta H$ [kcal/mol]	$-T \cdot \Delta S$ [kcal/mol]	$K_d$ [ $\mu\text{M}$ ]
 WW23	-5	-11 -(12-10)	6	109 (99-121)
 WW23_W32A	-4	-41 -(42-40)	36	469 (439-503)
 WW23_W131A	-5	-11 -(14-9)	6	114 (94-139)

Die Best-Fit Werte und die 95% Konfidenzintervalle (in Klammern) der Titrations von WW23; WW23\_W32A und WW23\_W131A mit CA\_P1. Alle Bindungsisotherme einer Titrationsreihe wurden Global gefittet und per Monte Carlo Analyse anschließend die Konfidenzintervalle ermittelt.

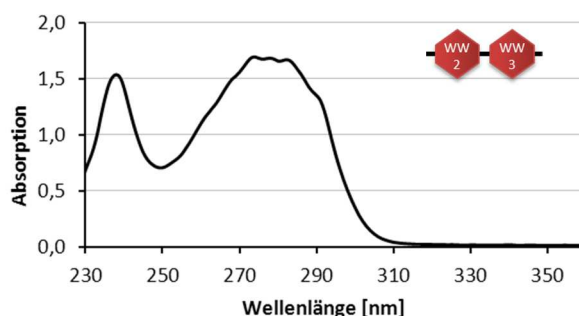
Da die Bindungsisotherme wegen des unvorteilhaften c-Wertes nur unvollständig durch Messwerte abgebildet werden konnten, sind die  $\Delta H$ - und  $\Delta S$ -Werte mit einer großen Unsicherheit behaftet und nur der Vollständigkeit halber angegeben. Wie die Ergebnisse in Tabelle 4.8.III zeigen, bindet CA\_P1 sehr schwach an WW3 ( $K_d$ : 469  $\mu\text{M}$ ) und schwach an WW2 ( $K_d$ : 114  $\mu\text{M}$ ). Entgegen den Erwartungen bindet CA\_P1 das WW23-Tandem ebenfalls nur schwach ( $K_d$ : 109  $\mu\text{M}$ ), vergleichbar mit der monovalenten CA\_P1/WW2-Interaktion. Es ist daher nicht möglich festzustellen, ob CA\_P1 mono- oder bivalent mit WW23 wechselwirkt.

### 4.8.4 ANALYTISCHE ULTRAZENTRIFUGATION (AUZ)

Die vorläufige Auswertung der Thermogramme aus den ITC-Experimenten mit WW23 und CA\_P1 ergab ein stöchiometrisches Verhältnis (N) zwischen Peptid und WW23 im Bereich von  $N = 0.5$ . Zum einen wurde daher eine auf bivalenter Wechselwirkung basierende peptidinduzierte Dimerisierung von WW23 in Betracht gezogen. Andererseits bestand die Möglichkeit einer Cystein-basierten Monomerverbrückung, da WW23 innerhalb der WW3-Sequenz ein Cystein enthält. Das WW-Domänen, welche an

geeigneter Stelle ein Cystein exponieren, Dimere ausbilden können, zeigte schon *Christoph Parthier* [171]. Zwar wurden in den während der Präparation aller WW23-Varianten routinemäßig durchgeführten Gelfiltrationen keine dimeren Spezies detektiert, aufgrund der geringen Auflösung dieser Methode ließ sich jedoch die Präsenz dimerer Spezies auch nicht zweifelsfrei ausschließen.

Mittels Analytischer Ultrazentrifugation wurde in Sedimentationsgeschwindigkeits-experimenten (SV) deshalb geprüft, ob WW23 generell oder in Anwesenheit von CA\_P1 Dimere ausbildet (s. Abschnitt 3.10). Dazu wurde einerseits das Sedimentationsverhalten von WW23 bei Konzentrationen von 8  $\mu\text{M}$  und 318  $\mu\text{M}$  verfolgt und andererseits die Sedimentation von WW23 ( $c = 250 \mu\text{M}$ ) in Gegenwart von drei verschiedenen Peptidkonzentrationen (entsprechend einer Proteinsättigung von: 9%, 17% und 84%) beobachtet. Aufgrund des hohen molaren Extinktionskoeffizienten von WW23 ( $\epsilon_{280}$  Abs. bei 0.1%: 1,96) konnte die Messung nicht, wie sonst üblich, im Absorptionsmaximum bei 280 nm erfolgen. Um eine geeignete Wellenlänge zu ermitteln, wurde vor den SV-Experimenten jeweils ein UV-Vis Spektrum von der Probe mit der höchsten Konzentration aufgenommen (siehe Abbildung 42).



**Abbildung 42: UV-Vis Spektrum von WW23**

Bei 307 nm (WW23) bzw. 305 nm (WW23 + Peptid) waren die Absorptionen gerade klein genug, um den gesamten Konzentrationsbereich zu erfassen. Insgesamt wurden jeweils 134 Absorptionsscans (0,003 cm radiale Auflösung, 8 min Scanintervall) bei 50.000 rpm aufgenommen. Die Auswertung der Daten und die Berechnung der dafür erforderlichen Parameter (siehe Tabelle 4.8.IV), wurde wie in Abschnitt 3.10 beschrieben durchgeführt.

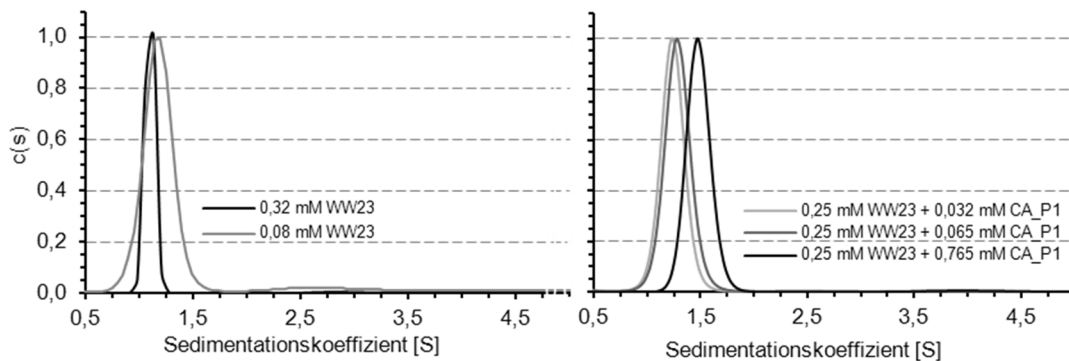
Die von „SEDPHIT“ [131,132] zur Berechnung der Friktionskoeffizienten-Verhältnisse ( $f/f_0$ ) sowie der davon abhängigen mittleren apparenten Molekulargewichte ( $MW_{app}$ ) benötigten Parameter (Pufferdichte ( $\rho$ ), Pufferviskosität ( $\eta$ ) und partial-spezifisches Volumen ( $\bar{v}$ ) von WW23), wurden mit „Sednterp“ [133] kalkuliert. Unter Berücksichtigung von Pufferzusammensetzung, pH-Wert, Temperatur und Aminosäuresequenz von WW23 ergaben sich die in Tabelle 4.8.IV aufgeführten Werte.



**Tabelle 4.8.IV: Mit "Sednterp" berechnete Parameter**

Lauf	$\rho$	$\eta$	$\bar{v}$	MW
WW23	1,00649 g cm <sup>-3</sup>	0,01333 N*s*m <sup>-2</sup>	0,7255	17,9 kDa
WW23+CA_P1	1,00629 g cm <sup>-3</sup>	0,012593 N*s*m <sup>-2</sup>	0,7255	17,9 + 3,7 kDa

Wie aus den in Abbildung 43 gezeigten  $c(S)$ -Verteilungen eindeutig hervorgeht, sedimentiert WW23 bei Abwesenheit (links) und Anwesenheit von CA\_P1 (rechts) im untersuchten Konzentrationsbereich in Form einer einzelnen Spezies. Die aus den Sedimentationsverläufen abgeleiteten Daten sind in Tabelle 4.8.V dargestellt.

**Abbildung 43: AUZ-Analyse von WW23 und WW23/CA\_P1**

Plots der aus SV-Experimenten ermittelten  $c(S)$ -Verteilung von WW23 ohne (links) und in Anwesenheit verschiedener CA\_P1-Konzentrationen (rechts). Die Sedimentation der Proteine wurde durch Absorptionsmessung bei 307 nm (WW23, rechts) bzw. 305 nm (WW23 +CA\_P1, links) verfolgt. Die Werte der Kurvenmaxima wurden jeweils auf 1 normiert.

**Tabelle 4.8.V: Daten der AUZ-Experimente mit WW23**

	Konzentration [ $\mu$ M]	Fitgüte (rmsd)	$f/f_0$	$S_{app}$ [S]	$MW_{app}$ [kDa]
WW23	8	0,0062	1,635	1,174	18,9
WW23	318	0,0067	1,720	1,098	18,5
WW23 CA_P1	250 32	0,0185	1,649	1,238	18,8
WW23 CA_P1	250 65	0,0185	1,571	1,285	18,7
WW23 CA_P1	250 765	0,0189	1,347	1,480	18,4

Die mittleren, apparenten Sedimentationskoeffizienten ( $S_{app}$ ) bzw. die daraus abgeleiteten, mittleren apparenten Molmassen ( $MW_{app}$ ) von WW23 liegen mit 18,9 kDa und 18,5 kDa nahe dem theoretischen Molekulargewicht von monomerem WW23 (17,9 kDa). Das bestätigt zweifelsfrei, dass WW23 unter den gewählten Bedingungen als Monomer vorliegt. Das Friktionskoeffizientenverhältnis ( $f/f_0$ ) von  $f/f_0 = 1,72$  (bei 318  $\mu$ M WW23) spricht dafür, dass die globale Molekülform des Proteins deutlich von der

Kugelform ( $f/f_0 = 1$ ) abweicht. WW23 könnte somit als stark elongiertes Monomer vorliegen. Vorstellbar wäre auch, dass beide WW-Domänen miteinander assoziiert sind, der Linkerbereich aber schleifenförmig herausragt oder dass die Konformation von WW23 dynamisch zwischen einem elongierten und einem kompakten Zustand wechselt.

Die Anwesenheit von CA\_P1 führte zu einer konzentrationsabhängigen Verringerung des  $f/f_0$ . Die Interaktion mit dem Peptid bewirkt demnach eine Konformationsänderung in Richtung einer kompakteren Molekülform. Das ist ein sehr starkes Indiz dafür, dass CA\_P1 gleichzeitig und bivalent mit beiden WW-Domänen interagiert. Dass sich offensichtlich ein WW23-CA\_P1-Komplex bildet, lässt sich mit der peptidkonzentrationsabhängigen Zunahme des  $S_{app}$ -Wertes belegen. Obwohl der Komplex ca. 3,5 kDa schwerer ist, als WW23 allein, ändert sich das  $MW_{app}$  der sedimentierenden Spezies nach Zugabe von CA\_P1 kaum. Da in die Berechnung von  $MW_{app}$  neben  $S_{app}$  auch  $f/f_0$  eingeht (siehe unten stehende Gleichung) und sich  $f/f_0$  peptidinduziert verringert, wird  $MW_{app}$  zu klein bestimmt. Die peptidinduzierte Bildung eines  $(WW23)_2$ -CA\_P1-Komplexes kann jedoch sicher ausgeschlossen werden. Die Bildung eines WW23-(CA\_P1)<sub>2</sub>-Komplexes scheint ebenfalls ausgeschlossen, da eine Massenzunahme von 39% (entsprechend zwei gebundenen Peptidmolekülen mit je 3,5 kDa) mit großer Wahrscheinlichkeit einen signifikanten Anstieg des  $MW_{app}$  bewirkt hätte.

$$MW_{app} = \frac{S_{app} * N * k * 6\pi\eta r_0}{(1 - \bar{v}\rho)}$$

$MW_{app}$  = mittleres Apparentes Molekulargewicht

$S_{app}$  = mittlerer apparenter Sedimentationskoeffizient

$N$  = Avogadro Zahl

$k$  = Friktionskoeffizientenverhältnis ( $f/f_0$ )

$\eta$  = Pufferviskosität

$r_0$  = hydrodynamischer Radius

$\bar{v}$  = partialspezifisches Volumen

$\rho$  = Pufferdichte

#### 4.9 INTERAKTION VON CA150-WW23 MIT „KURZEN“ PP-PEPTIDEN

Um die Frage zu klären, mit welcher Präferenz die beiden WW-Domänen in WW23 die prolinreichen Motive PPLPP und PPPPP binden, wurde die Interaktion der in Tabelle 4.9.I dargestellten „kurzen“ Peptidliganden mit WW23 NMR-spektroskopisch untersucht.

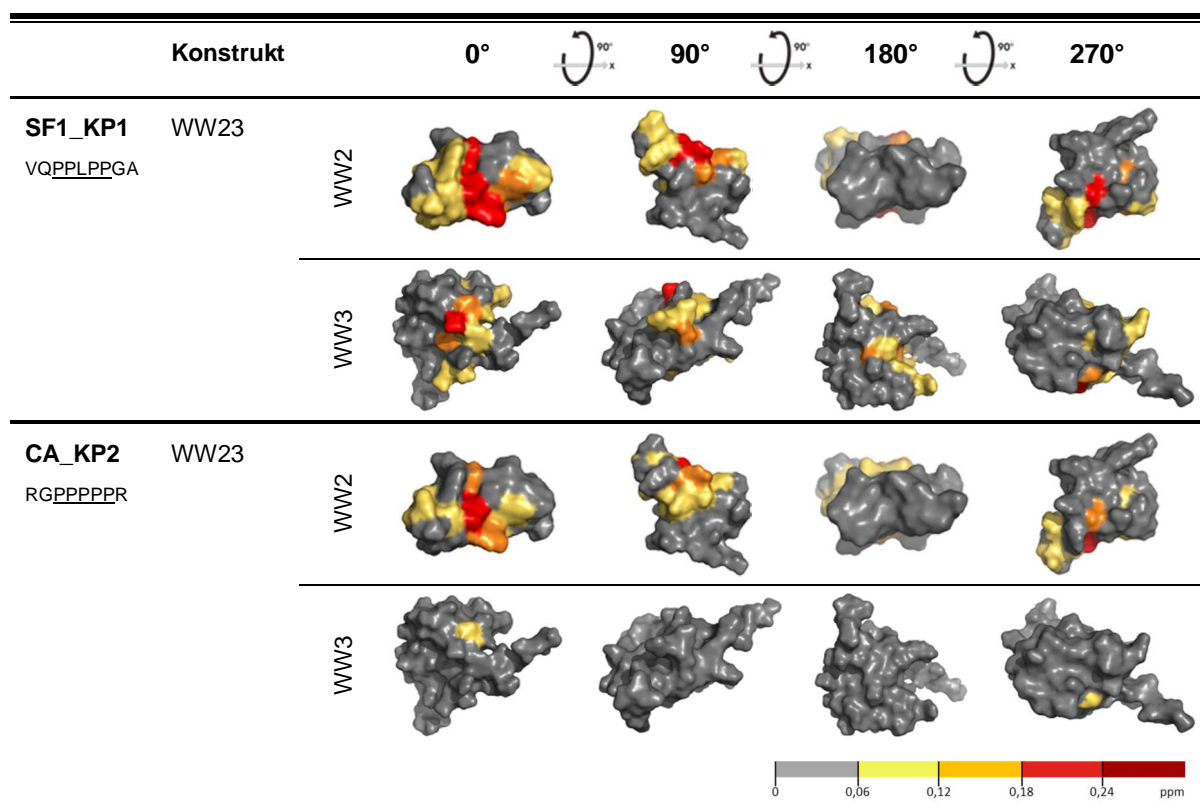
**Tabelle 4.9.I: Kurze PP-Peptide**

Bezeichnung	Sequenz	AS <sub>CA150</sub>	kDa
SF1_KP1	VQPPLPPGA	573-581	0,875
CA_KP2	RGPPPPPR	57-64	0,873

Dazu wurden  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von uniform  $^{15}\text{N}$ -markierter WW23 ( $c = 50 \mu\text{M}$ ) in Ab- und Anwesenheit eines 10fachen Überschusses ( $c = 500 \mu\text{M}$ ) an SF1\_KP1 bzw. CA\_KP2 aufgenommen und bezüglich Signalpositionsveränderungen untersucht (siehe auch Anhang, Abbildung 54).

Da die dreidimensionalen Strukturen von WW2 und WW3 bekannt sind (WW2: PDB:1E0L, WW3: PDB: 2DK7), konnten die peptidinduzierten  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungsdifferenzen direkt auf die entsprechenden Strukturen abgebildet werden. Die Absolutwerte der Chemischen Verschiebungsdifferenzen wurden dazu in verschiedene Wertebereiche eingeteilt und diese anhand einer Farbskala dargestellt (siehe Tabelle 4.9.II).

**Tabelle 4.9.II: Stärke der SF1\_KP1 und CA\_KP1-induzierten Signalverschiebungen**



Gezeigt sind die Strukturen von WW2 (PDB: 1E0L) und WW3 (PDB: 2DK7). Die gewichteten Vektorbeträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}^{15}\text{N}$  chemischen Verschiebungen der WW23-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala)

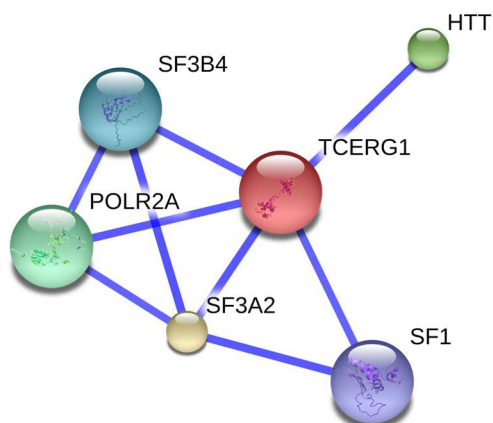
Aus den Ergebnissen der Titrationsen geht eindeutig hervor, dass WW2 erwartungsgemäß beide Motive (PPLPP & PPPPP) binden kann. Beide Peptide besetzten nahezu die gleichen Oberflächenareale und interagieren mit den beiden bekannten WW2-Bindetaschen „xP“ und „xP2“. Die WW3-Domäne bindet nur das PPLPP-Motiv.

## 5 DISKUSSION

Um ein besseres Verständnis grundlegender Mechanismen der multivalenten Wechselwirkung zwischen den CA150<sub>WW</sub> und der PRS von SF1 sowie der N-terminalen PRS von CA150 selbst zu erlangen, aber auch um Erkenntnisse bezüglich der Organisation der drei WW-Domänen in CA150 zu gewinnen, wurden in dieser Arbeit verschiedenste molekularbiologische, biochemische und biophysikalische Methoden angewendet. Die wichtigsten Forschungsergebnisse werden nachfolgend kurz zusammengefasst und unter Einbeziehung von Ergebnissen anderer Forschungsgruppen diskutiert.

### 5.1 FUNKTIONELLE AKTIVITÄT DER VERWENDETEN WW-VARIANTEN

Der humane Transkriptionsfaktor CA150 ist im Zellkern an der Regulation der Elongation der RNA-Polymerase II, der Koordinierung der Transkription und der prä-mRNA-Prozessierung beteiligt. Zudem spielt CA150 auch eine wichtige Rolle im Krankheitsverlauf von Huntington (vgl. Abschnitt 1.5). Einige wichtige Interaktionen von CA150 finden zu Spleißfaktoren, wie SF3B4 [115], SF3A2 [172,173] und SF1 [112], zur RNA-Polymerase POLR2A [112] und zum Huntingtin (HTT) [87] statt (vgl. Abbildung 44).



**Abbildung 44: Wichtige TCERG1-Interaktionen**

CA150 (TCERG1) und dessen Interaktionspartner sind als verschiedenfarbige Kugeln dargestellt. Die Stärke der Verbindungslinien repräsentiert die Stärke der Wechselbeziehung. Die Abbildung wurde mit STING 9.0 erzeugt [117]

Um die Funktionalität bzw. Bindungsaktivität der in dieser Arbeit eingesetzten WW-Varianten (WW2, WW3 & WW23) zu überprüfen, wurden Triple-SILAC-Pulldown Experimente durchgeführt. Wie in Abschnitt 4.7 dargestellt, konnten mit allen drei WW-Varianten Bindepartner aus dem Zelllysate gefischt werden. Dadurch konnte zweifelsfrei

bestätigt werden, dass die isolierten Konstrukte funktionell aktiv waren und in weiteren Interaktionsstudien eingesetzt werden konnten.

Neben den bereits bekannten Interaktionspartnern (siehe Tabelle 4.7.II) konnten zusätzlich 9 weitere CA150<sub>WW</sub>-Interaktionspartner identifiziert werden (siehe Tabelle 5.1.I). Nach einer Datenbankanalyse stellte sich heraus, dass diese entweder eine Funktion im Spleiß- oder Transkriptionsprozess innehaben. Es konnte auch eine Wechselwirkung aller drei WW-Varianten zu RBP1, einer RNA-Polymerase gezeigt werden. Anders als bei der RNA-Polymerase II (POLR2A), die mit den FF-Domänen von CA150 wechselwirkt [116], könnte RBP1 mit den WW-Domänen von CA150 interagieren. Da es sich bei den gefundenen Interaktionspartnern prinzipiell auch um Sekundärinteraktionen handeln kann, müsste das in weiteren Untersuchungen verifiziert werden.

**Tabelle 5.1.I: Neue CA150<sub>WW</sub>-Interaktionspartner**

ID	Name	Funktioneller Kontext
<b>CHD1</b>	<i>Chromodomain-helicase-DNA-binding protein 1</i>	Phosphoprotein, <b>Transkription</b>
<b>CHD3</b>	<i>Chromodomain-helicase-DNA-binding protein 3</i>	Phosphoprotein, <b>Transkription</b>
<b>CHD4</b>	<i>Chromodomain-helicase-DNA-binding protein 4</i>	Phosphoprotein, Acetylierung, <b>Transkription</b>
<b>HNRNPL</b>	<i>Heterogeneous nuclear ribonucleoprotein L</i>	Phosphoprotein, Acetylierung, Ribonukleoprotein
<b>POLR2B</b>	<i>DNA-directed RNA polymerase II subunit RPB2</i>	<b>Transkription</b>
<b>RPB1</b>	<i>RNA-polymerase Rpb1</i>	Phosphoprotein, Acetylierung, Methylierung, Ubl Konjugation, <b>Transkription</b>
<b>SF3B3</b>	<i>Splicing factor 3b, subunit 3</i>	<b>Spleißen</b>
<b>SF3B5</b>	<i>Splicing factor 3b, subunit 5</i>	<b>Spleißen</b>
<b>SF4</b>	<i>Splicing factor 4</i>	<b>Spleißen</b>
<b>CA150</b>	<b>Transcription elongation regulator 1</b>	Phosphoprotein, <b>Transkription</b>

Eine Wechselwirkung von CA150 mit SF3B3 ist zwar nicht in den Datenbanken gelistet, wurde aber erstmalig schon 2005 von *Ingham et al.* [172] beschrieben. Das überraschende Ergebnis des SILAC-Pulldown-Experimentes war jedoch, dass CA150 selbst als Interaktionspartner aus dem HeLa-Zelllysat gefischt werden konnte.

Um zu klären, ob CA150 mit seinen eigenen WW-Domänen wechselwirkt oder es sich dabei um eine Sekundärinteraktion handelt, wurden SPOT-Analysen, NMR-Titrationsexperimente und ITC-Untersuchungen mit verschiedenen CA150<sub>WW</sub>-Varianten und der N-terminalen CA150<sub>PRS</sub> durchgeführt (siehe Abschnitt 4.8 und Abschnitt 5.4).



den angrenzenden WW-Domänen zu gehören. Aus der Kombination der berechneten Orderparameter und der gemessenen  $^{15}\text{N}$ -Het-NOEs ließen sich vier klare Domänengrenzen ableiten. Daraufhin wurden die Sequenzbereiche zwischen den AS 10-42, 43-83 und 84-146 strukturell der WW2-Domäne, dem Linker bzw. der WW3-Domäne zugeordnet. Da die C-terminal an WW3 grenzende  $\alpha$ -Helix fest mit der Kerndomäne assoziiert zu sein scheint, ist davon auszugehen, dass beide Elemente eine funktionelle Struktureinheit bilden. Dafür sprechen auch die Resultate der Expressionstests (siehe Abbildung 16), wonach es nicht möglich war, ein lösliches WW23-Konstrukt ohne die C-terminale Helix zu exprimieren. Höchstwahrscheinlich ist die Helix für eine korrekte WW23-Faltung essentiell.

Aus der Strukturdynamischen Analyse ließ sich darüber hinaus ableiten, dass WW23 keinen isotropen Diffusionstensor besitzt und demzufolge keine globuläre Molekülform aufweist. Unklar blieb jedoch, ob WW2 und WW3 in Interaktion miteinander stehen oder sich relativ unabhängig voneinander bewegen können. Eine potentielle WW2-WW3-Interaktion könnte dabei sowohl intramolekular als auch intermolekular stattfinden. Da letzteres zur Bildung oligomerer Spezies führen würde, wurde der Oligomerisierungszustand von WW23 mittels AUZ untersucht. Wie sich dabei herausstellte, lag WW23 ausschließlich als monomere Spezies vor, sodass eine intermolekulare WW23-Interaktion ausgeschlossen werden konnte (siehe Abschnitt 4.8.4). Darüber hinaus konnte festgestellt werden, dass die globale Molekülform von WW23 deutlich von der Kugelform abweicht. WW23 könnte somit als stark elongiertes Monomer vorliegen. Vorstellbar wäre auch, dass beide WW-Domänen miteinander assoziiert sind, der Linkerbereich aber schleifenförmig herausragt oder dass die Konformation von WW23 dynamisch zwischen einem elongierten und einem kompakten Zustand wechselt.

Die NMR-spektroskopische Untersuchung von zwei WW23-Mutanten (WW23\_W23A und WW23\_W131A) zeigte, dass sich die W $\rightarrow$ A Punktmutation in der Bindetasche einer WW-Domäne nicht nur auf die chemische Umgebung innerhalb der mutierten WW-Domäne auswirkte, sondern auch die chemische Umgebung des Linkerbereiches und der anderen WW-Domäne verändert wurde. Da die Mutation einer Domäne nicht die Bindungsaktivität der jeweils anderen Domäne beeinträchtigte, und die HSQC-Spektren aller drei Proteine (WW23, WW23\_W32A & WW23\_W131A) in hohem Maße miteinander übereinstimmten, kann davon ausgegangen werden, dass die Punktmutation keine globalen Strukturveränderungen zur Folge hatte.

Mit hoher Wahrscheinlichkeit ist daher anzunehmen, dass die beiden WW-Domänen direkt miteinander wechselwirken.

### 5.3 INTERAKTIONEN ZWISCHEN CA150-WW-DOMÄNEN UND DER SF1<sub>PRS</sub>

Dass CA150 über seine WW-Domänen mit der SF1<sub>PRS</sub> interagiert, wurde bereits von *Goldstrohm et al.* [112] gezeigt. Das Hauptziel der hier durchgeführten Interaktionsstudien bestand darin, mehr über den Mechanismus dieser Wechselwirkung zu erfahren. Der Fokus lag dabei insbesondere auf der Beantwortung der Frage, ob die WW-Domänen multivalent mit der PRS interagieren. Dazu wurde in einer systematischen SPOT-Analyse geprüft, ob die einzelnen Domänen bestimmte Bereiche der SF1<sub>PRS</sub> bevorzugen. Ein starker Hinweis für eine multivalente Wechselwirkung wäre, wenn jeder WW-Domäne eine eindeutige und separate Bindestelle innerhalb der PRS zugeordnet ist. Wie die SPOT-Analyse zeigte, ließen sich den einzelnen WW-Domänen jedoch keine individuellen, spezifisch lokalisierten PRS-Sequenzbereiche zuordnen.

Vielmehr konnte gezeigt werden, dass entgegen der bisherigen Annahme von *Goldstrohm et al.* [112] alle drei WW-Domänen an die SF1<sub>PRS</sub> binden können. Der von *Goldstrohm* postulierte CA150<sub>WW</sub> SF1<sub>PRS</sub>-Interaktionsbereich (SF1<sub>400-500</sub>) konnte darüber hinaus erweitert (SF1<sub>418-639</sub>) und in drei Regionen unterteilt werden: SF1<sub>418-445</sub>; SF1<sub>469-526</sub> und SF1<sub>566-639</sub> (siehe Abschnitt 4.6.1). Zusätzlich wurde deutlich, dass die drei Domänen unterschiedliche Bindungspräferenzen besitzen. WW1 bindet Motive in diesen drei Bereichen nur schwach, wohingegen WW2 und WW3 eine deutlich höhere Affinität, aber auch eine geringere Spezifität aufweisen. Für eine genauere Analyse der Präferenzen bezüglich verschiedener Bindemotive siehe Abschnitt 5.5.2).

Die Untersuchung von Multivalenzeffekten beschränkte sich in dieser Arbeit auf das Studium von bivalenten Interaktionen zwischen WW23 und multivalenten Teilbereichen der SF1<sub>PRS</sub>, da keine WW12 bzw. WW123 Konstrukte erzeugt werden konnten. Verschiedene SF1-Sequenzen (2, 3, 5-7, 14, 15 und 18) zeigten putative Multivalenzeffekte in der Art, dass diese an WW23 stärker gebunden haben als an die Einzeldomänen. Besonders deutliche Effekte zeigten dabei Seq 3 (SF1\_P1), Seq 5 (SF1\_P2) und Seq 7 (SF1\_P3), sodass diese für weitere Studien ausgewählt wurden (s. Tabelle 4.6.I). Zusätzlich wurde eine synthetische Peptidsequenz, welche ein bekanntes (PPPPP) und ein putatives (PLPP) Bindemotiv für WW-Domänen enthält (SF1\_P4), kreiert (siehe Abschnitt 4.6.1).

Um die putativen Multivalenzeffekte zu verifizieren und die an der Interaktion beteiligten Aminosäuren in WW23 zu identifizieren, wurden NMR-Bindungsstudien durchgeführt. Diese sollten Aufschluss darüber geben, ob eine oder beide Domänen in WW23 mit dem Peptid wechselwirken. Die Experimente zeigten, dass jeweils beide WW-Domänen mit allen getesteten Peptiden (SF1\_P1, SF1\_P3 & SF1\_P4) interagieren. Darüber hinaus wurde deutlich, dass SF1\_P1 und SF1\_P3 die chemische Umgebung der WW2-Domäne deutlich stärker beeinflussen, als die der WW3-Domäne. SF\_P1 scheint mit WW3 nur sehr schwach zu interagieren. Diese Beobachtungen sind konform mit den Ergebnissen der SPOT-Analyse (siehe Tabelle 4.6.I). SF1\_P4 interagiert weniger stark mit WW2, hat aber eine höhere Affinität zu WW3, als die anderen Peptide.



Da die NMR-Bindungsstudien keine quantitativen Aussagen zur Bindungsstärke ( $K_d$ ) und zur Bindungsstöchiometrie erlaubten, wurden ITC-Experimente durchgeführt (siehe Abschnitt 4.6.4). Um dabei mono- von bivalenten Effekten differenzieren zu können wurden WW23-Mutanten erzeugt, in denen jeweils eine der beiden WW-Domänen inaktiv war (siehe Abschnitt 4.6.2.1 und 4.6.3). Ausschlaggebendes Kriterium für die Annahme einer bivalenten Wechselwirkung, war eine signifikant höhere Affinität zwischen Peptid und Wildtypprotein, als zwischen Peptid und den WW23-Mutanten.

Alle untersuchten Interaktionen waren überraschenderweise sehr schwach mit Bindekonstanten im mikromolaren Bereich, was die Aussagekraft der durchgeführten ITC-Experimente erheblich reduzierte. Da die Aktivität der eingesetzten Proteine und Peptide unbekannt war und sich deren aktive Konzentration auf Grund der schwachen Wechselwirkungen nicht zufriedenstellend aus den Bindungsisothermen ableiten ließen, erlaubten die ITC-Experimente in diesem Fall jedoch keine verlässlichen Aussagen über die tatsächliche Reaktionsstöchiometrie. Erschwerend kommt hinzu, dass bei der Betrachtung der stöchiometrischen Verhältnisse auch berücksichtigt werden müsste, dass monovalente und bivalente Interaktionen gleichzeitig auftreten und somit zu einem Gemisch verschiedener Protein-Peptidkomplexe führen können. Denkbar wäre die Bildung der Komplexe: (WW2<sub>Peptid</sub>)-WW3, (WW2<sub>Peptid</sub>)-(WW3<sub>Peptid</sub>), WW2-(WW3<sub>Peptid</sub>), (WW23<sub>Peptid</sub>). Aus diesen Gründen konnten für die Unterscheidung von mono- und bivalenten Interaktionen nur die Dissoziationskonstanten herangezogen werden.

Die ITC-Untersuchungen mit SF1\_P1 zeigten deutlich, dass SF1\_P1 schwach an WW2 bindet ( $K_d = 807 \mu\text{M}$ ), aber nur sehr schwach mit WW3 ( $K_d \gg 1 \text{ mM}$ ) wechselwirkt (siehe Tabelle 4.6.VII). Das ist konsistent mit den Resultaten aus der SPOT-Analyse (Tabelle 4.6.I) und der NMR-Titration (Tabelle 4.6.IV). Ebenfalls konform mit der SPOT-Analyse ist die Tatsache, dass SF1\_P1 wesentlich besser an den Wildtyp ( $K_d = 452 \mu\text{M}$ ) als an WW23\_W131A (WW2,  $K_d = 807 \mu\text{M}$ ) bindet. Diese Verdopplung der Affinität spricht deutlich für eine bivalente Wechselwirkung zwischen SF1\_P1 und WW23.

Die NMR-Titration ließ zwar darauf schließen, dass das Peptid eher schwächer an den Wildtyp bindet als an WW23\_W131A, indizierte aber auch, dass SF1\_P1 möglicherweise unspezifisch bzw. sehr schwach mit WW3 wechselwirkt. Die hier ebenfalls detektierte, sehr schwache WW3/SF1\_P1-Interaktion könnte daher Teil einer bivalenten Wechselwirkung und damit die Ursache für die beobachtete Affinitätserhöhung sein.

Denkbar wäre, dass dieser schwache Beitrag die Wahrscheinlichkeit einer korrekten Orientierung zwischen WW2 und SF1\_P1 und damit einhergehend die effektive Konzentration ( $C_{\text{eff}}$ ) erhöht, wodurch die Assoziation begünstigt wird.

Die Ergebnisse der WW/SF1\_P3-Titration zeigten, dass SF1\_P3 den Wildtyp deutlich affiner bindet ( $K_d = 55 \mu\text{M}$ ) als die einzelnen WW-Domänen (WW2:  $K_d = 123 \mu\text{M}$ ; WW3:  $K_d = 634 \mu\text{M}$ ) (siehe Tabelle 4.6.VIII). Das ist konsistent mit den Resultaten aus der SPOT-Analyse (siehe Tabelle 4.6.I) und der NMR-Titration (siehe Tabelle 4.6.V). Es kann

davon ausgegangen werden, dass es sich bei dieser Wechselwirkung mit großer Wahrscheinlichkeit um eine bivalente Interaktion handelt.

Zu SF1\_P4 (Tabelle 4.6.IX) konnten mit Bindekonstanten von 59  $\mu\text{M}$  (WW2), 39  $\mu\text{M}$  (WW3) und 39  $\mu\text{M}$  (WW23) zwar die stärksten Affinitäten, jedoch keine Affinitätszunahme beobachtet werden. Eine bivalente Wechselwirkung schien deshalb zunächst ausgeschlossen. Die NMR-Spektroskopische Untersuchung zeigte jedoch, dass sich die peptidinduzierten Signalverschiebungen in den HSQC-Spektren von WW23 und WW23\_W131A deutlich unterscheiden. Das Peptid interagiert mit der WW2-Domäne des Wildtyps über andere Oberflächenbereiche als mit der WW2 von WW23\_W131A (Tabelle 4.6.VI). Die Interaktion mit der aktiven WW3-Domäne scheint das Peptid sterisch anders zu WW2 auszurichten, sodass hier ebenfalls von einer bivalenten Wechselwirkung auszugehen ist. Aufgrund der geringen Distanz zwischen den beiden prolinreichen Sequenzen können sich wahrscheinlich nicht beide Motive gleichzeitig optimal zu beiden WW-Domänen orientieren (siehe Abbildung 4). Die bivalente Interaktion scheint zwar bevorzugt zu sein, führt laut den ITC-Daten aber zu keiner messbaren Affinitätserhöhung.

Zusammenfassend ist festzuhalten, dass das WW23-Tandem aus CA150 bivalent mit verschiedenen Teilbereichen der SF1<sub>PRS</sub> wechselwirkt und alle drei WW-Domänen verschiedene prolinreiche Motive der SF1<sub>PRS</sub> binden können. Mit großer Wahrscheinlichkeit ist deshalb davon auszugehen, dass CA150 *in vivo* multivalent mit SF1 interagiert.

Welche WW-Bindetaschen von den getesteten Peptiden genutzt werden und welche Präferenzen die WW-Domänen für verschiedene prolinreiche Motive zeigen, wird, auch unter Berücksichtigung der Ergebnisse aus CA150<sub>WW</sub> CA150<sub>PRS</sub> Interaktionsstudie, in den Abschnitten 5.5.1 & 5.5.2 diskutiert.

#### **5.4 INTERAKTIONEN ZWISCHEN CA150-WW-DOMÄNEN UND DER N-TERMINALEN CA150<sub>PRS</sub>**

Mittels SPOT-Analyse, NMR-Bindungsstudien und ITC-Experimenten konnte in dieser Arbeit erstmalig gezeigt werden, dass die WW-Domänen von CA150 mit einem Teilbereich der CA150-eigenen N-terminalen PRS (CA150<sub>37-68</sub>) wechselwirken können (siehe Abschnitt 4.8). Erste Anhaltspunkte für eine CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Wechselwirkung lieferte der SILAC-Pulldown (siehe Abschnitt 4.7 & 5.1). Diese Methode erlaubte jedoch keine Unterscheidung dahingehend, ob es sich bei den gefundenen Interaktionen um direkte Wechselwirkungen oder Sekundärinteraktionen handelte.

In einer SPOT-Analyse wurden deshalb systematisch verschiedene Sequenzabschnitte der N-terminalen CA150<sub>PRS</sub> auf ihre Interaktion mit den CA150-WW-Varianten (WW1, WW2, WW3 & WW23) getestet. Es konnte gezeigt werden, dass alle drei WW-Domänen an den PRS-Abschnitt 58-66 binden. Die WW3-Domäne konnte zusätzlich auch prolinreiche Motive im PRS-Bereich 29-58 erkennen. Interessanterweise konnte das

WW23-Tandem prolinreiche Motive innerhalb der gesamten PRS erkennen, selbst, wenn keine Interaktionen zu den Einzeldomänen nachweisbar waren. Zu allen untersuchten Teilbereichen hatte das WW23-Tandem ausnahmslos eine bedeutend höhere Affinität als die Einzel-Domänen, was ein starker Hinweis für das Vorhandensein multivalenter Wechselwirkungen war.

Um den Mechanismus der beobachteten multivalenten Wechselwirkung qualitativ (Interaktionsstellen) und thermodynamisch (Affinitäten) besser verstehen zu können, wurde für weitere Untersuchungen ein PRS-Bereich (CA150<sub>37-68</sub>) ausgewählt, welcher mono- als auch multivalente Bindungsmuster zeigte (CA\_P1, siehe Tabelle 4.8.I).

Die mit WW23 und den beiden WW23-Mutanten (WW23\_W32A & WW23\_W131A) durchgeführten NMR-Bindungsstudien bestätigten, dass CA\_P1 an beide WW-Domänen binden kann (vgl. Abschnitt 4.8.2). Deutlich wurde auch, dass CA\_P1 mit der WW2-Domäne über andere Oberflächenareale wechselwirkt, wenn WW3 inaktiviert wurde. D.h. bei inaktiver WW3 bindet das Peptid über die klassischen Bindetaschen „xP“ und „xP2“ (siehe Abschnitt 5.5) an WW2. Bei der Interaktion mit dem Wildtyp wurden diese Bindetaschen jedoch interessanterweise nicht genutzt. Dieser Unterschied war ein ernstzunehmendes Indiz für eine bivalente Wechselwirkung zwischen WW23 und CA\_P1. Die Interaktion zwischen CA\_P1 und WW3 orientiert das Peptid so zur WW2-Domäne, das es an diese vermutlich nur schwach binden kann. Die Summe beider Wechselwirkungen scheint aber energetisch günstiger zu sein als die monovalente Interaktion zwischen Peptid und WW2. Die ITC-Bindungsstudien bestätigten diese Vermutung. Es konnte gegenüber den monovalenten Interaktionen (WW2:  $K_d = 114 \mu\text{M}$ , WW3:  $K_d = 469 \mu\text{M}$ ) keine signifikante Affinitätserhöhung durch die bivalente Wechselwirkung zwischen CA\_P1 & WW23 ( $K_d = 109 \mu\text{M}$ ) beobachtet werden. Die mittels SPOT-Analyse detektierte Affinitätserhöhung scheint demnach ein Artefakt zu sein. Eine plausible Erklärung dafür wäre, dass die eingesetzten Einzel-WW-Varianten eine geringere Bindeaktivität besaßen, als die entsprechenden Domänen der Tandem-WW-Konstrukte.

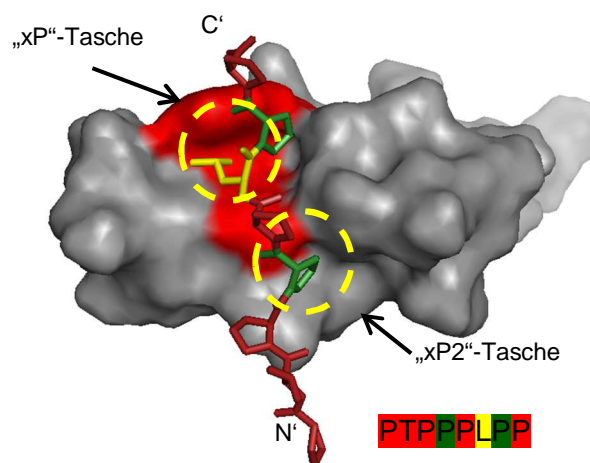
#### **5.4.1 PEPTIDINDUZIERTE STRUKTURVERÄNDERUNG VON WW23**

Die Auswertung der Thermogramme aus den ITC-Experimenten mit WW23 und CA\_P1 ergab ein stöchiometrisches Verhältnis (N) von Peptid zu WW23 im Bereich von  $N \sim 0.5$ . Neben einer monomolekularen bivalenten Wechselwirkung musste daher auch eine auf bivalenter Wechselwirkung basierende peptidinduzierte Dimerisierung von WW23 in Betracht gezogen werden. Das wurde mittels Sedimentationsgeschwindigkeitsexperimenten überprüft. Die Bildung peptidassoziierter Dimere konnte dadurch zweifelsfrei ausgeschlossen werden (siehe Abschnitt 4.8.4). Es konnte ebenfalls ausgeschlossen werden, dass WW23 zwei Peptidmoleküle bindet. Die AUZ-Daten zeigten aber eindeutig, dass CA\_P1 mit beiden WW-Domänen gleichzeitig und somit bivalent interagiert. Die Interaktion mit dem Peptid führte auch zu einer räumlichen Annäherung der WW-Domänen und damit einhergehend zu einer kompakteren Molekülform. Das

impliziert, dass sich die WW-Domänen auch in Abwesenheit eines Peptidliganden zeitweilig in räumlicher Nähe befinden müssen oder sogar miteinander interagieren. Letzteres scheint auch deshalb sehr plausibel, da sich die Punktmutation einer Domäne auch auf die chemische Umgebung der anderen Domäne auswirkte (siehe Abschnitt 5.2).

### 5.5 „xP“- UND „xP2“- BINDETASCHE

Die  $C_2$  Pseudosymmetrie der PPII Helix ermöglicht eine Bindung der PRS in zwei Orientierungen, vorwärts (+) und rückwärts (-). In beiden Bindungssituationen werden die gleichen Erkennungselemente des Peptids und Wasserstoffbrückendonoren der Protein-Seitenketten sowie die gleichen hydrophoben Bindestellen der WW-Domäne verwendet [55]. Die WW-Domänen zeigen oftmals Präferenzen für eine N-C oder eine C-N terminale Ausrichtung der Peptide [55]. Die CA150<sub>WW2</sub> besitzt, wie alle Mitglieder der WW-Domänen-Klasse II, neben der „xP“-Tasche eine weitere Bindungstasche, die sogenannte „xP2“-Tasche, durch welche wahrscheinlich die Bindung in zwei Orientierungen erst ermöglicht wird [96] (vgl. Abbildung 46).

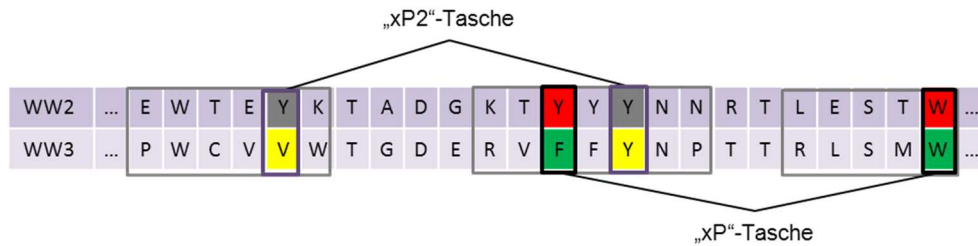


**Abbildung 46: CA150<sub>WW2</sub> im Komplex mit PTPPPLPP**

PDB-ID: 2RLY [96], die „xP“-Bindungstasche der CA150<sub>WW2</sub> wurde rot markiert

In dieser Arbeit konnte gezeigt werden, dass sich durch eine Punktmutation in der „xP“-Tasche von WW2 auch die chemischen Verschiebungen der Resonanzen ändern, welche die „xP2“-Tasche bilden ( $Y_{13}$  &  $Y_{23}$ ) (vgl. Abbildung 36). Da die Mutation die Bindeaktivität der Domäne nahezu vollständig ausschaltet, muss die „xP2“-Tasche funktionell mit der „xP“-Tasche gekoppelt bzw. funktionell von dieser abhängig sein.

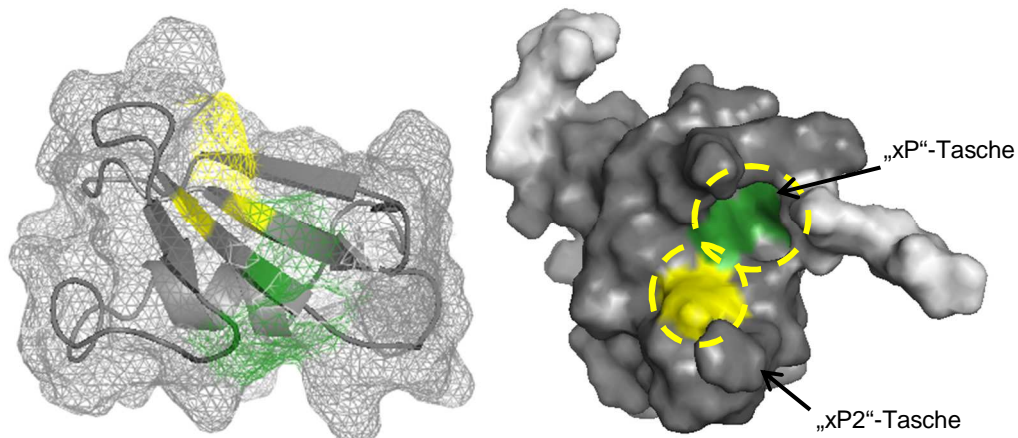
In den NMR-Bindungsstudien mit WW23 konnte gezeigt werden, dass abgesehen von CA\_P1, alle untersuchten prolinreichen Peptide mit beiden Bindetaschen wechselwirken. Nach Inaktivierung der WW3 von WW23 lagerte sich CA\_P1 ebenfalls an beide WW2-Bindetaschen an (vgl. Tabelle 4.8.II).



**Abbildung 47: WW2- und WW3-Sequenzausschnitt mit „xP“- und „xP2“-Bindungstasche**

Die Rechtecke markieren die Positionen der drei beta-Stränge. Die Aminosäuren der „xP“- und „xP2“-Tasche der WW2 wurden Rot bzw. Grün gekennzeichnet. Die Aminosäuren der „xP“- und potentiellen „xP2“-Tasche der WW3 sind Grün bzw. Gelb gekennzeichnet.

Basierend auf der Positionierung der „xP2“-Tasche in WW2 kann vermutet werden, durch welche Sequenzpositionen ( $V^{112}$  &  $Y^{122}$ ) eine solche „xP2“-Tasche in WW3 ausgebildet werden könnte (s. Abbildung 47 & Abbildung 48). Interessanterweise zeigten die  $^1H_N^{15}N$ -Kreuzsignale von  $V^{112}$  &  $Y^{122}$  nach der Punktmutation in der WW3-„xP“-Tasche ebenfalls eine starke Änderung ihrer chemischen Verschiebung (vgl. Abbildung 37).

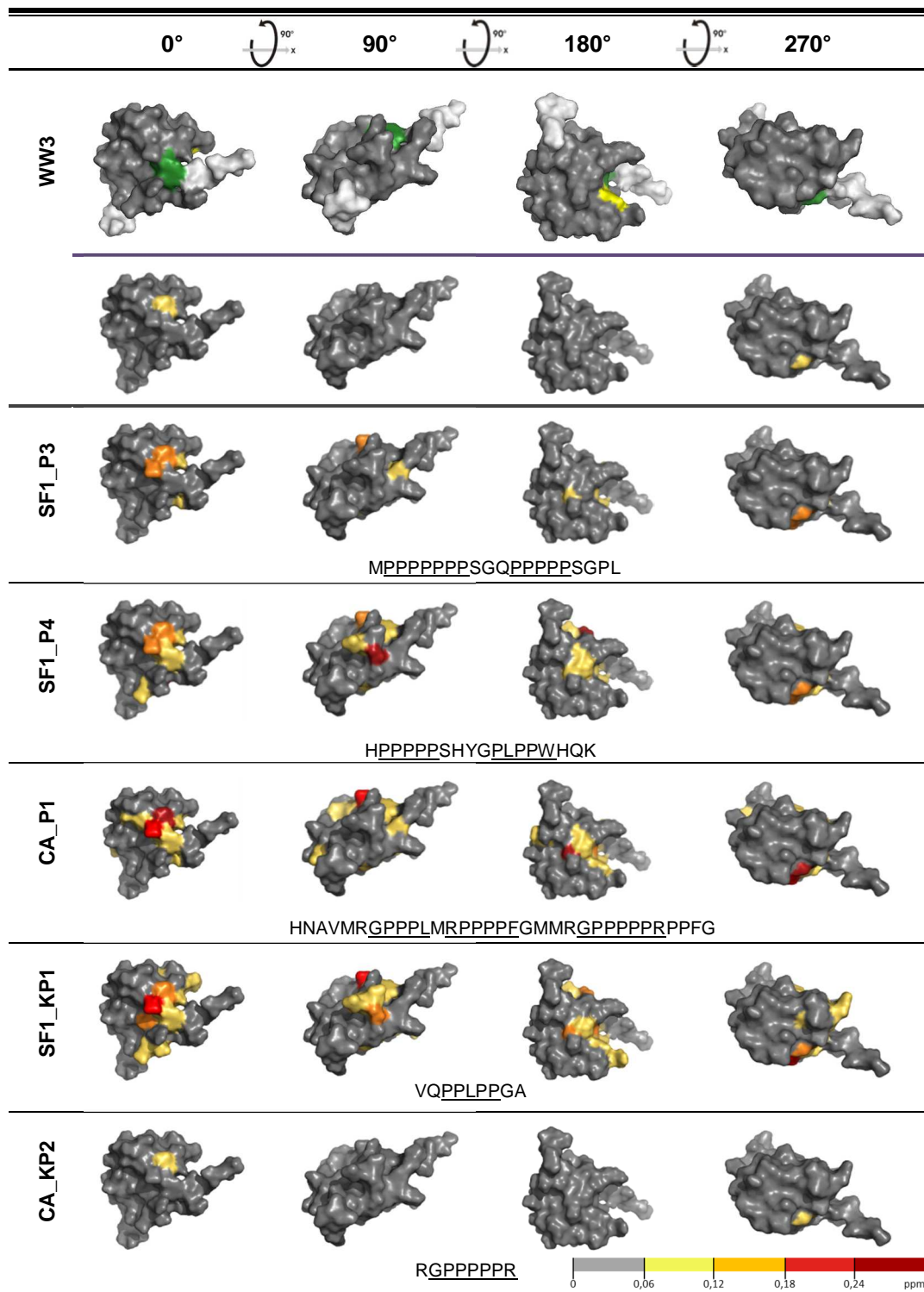


**Abbildung 48: „xP“ und „xP2“-Tasche von CA150<sub>WW3</sub>**

Gezeigt ist die Cartoon- (links) und die Oberflächenstruktur (rechts) von CA150 WW3 (PBD: 2DK7). Die Hellgrau gefärbten Bereiche (rechts) markieren die Positionen der (in der NMR-Struktur abgebildeten) N- und C-terminalen Protein-Tags. Die Position der „xP“-Tasche ist grün, die der potentiellen „xP2“-Tasche gelb eingefärbt.

Auf Grund dieser Beobachtung stellt sich die Frage, ob WW3 mit dieser potentiellen „xP2“-Tasche auch prolinreiche Peptidliganden binden könnte. Um dieser Frage nachzugehen, wurden in Tabelle 5.5.I zunächst die Positionen der „xP“- (grün) und „xP2“-Bindetasche und darunter die peptidinduzierten  $^1H_N^{15}N$  chemischen Verschiebungsdifferenzen direkt auf der WW3-Struktur abgebildet. Die Differenzen der gewichteten NH-Vektoren (siehe Abschnitt 3.7.2) wurden dazu in verschiedene Wertebereiche eingeteilt und anhand einer Farbskala dargestellt.

Tabelle 5.5.I: Übersicht zur Bindung der PP-Peptide an die „xP“- bzw. „xP2“-Tasche



Gezeigt sind die Strukturen von WW3 (PDB: 2DK7). Oben: Die Position der „xP“- und putativen „xP2“-Bindetasche sind grün bzw. gelb markiert. Darunter: Die gewichteten Vektorbeiträge der peptidinduzierten Änderungen der  $^1\text{H}_\text{N}$ / $^{15}\text{N}$  chemischen Verschiebungen der WW3-Resonanzen wurden, abhängig vom Betrag, verschiedenfarbig auf die Strukturen geplottet. Die verwendeten Farben repräsentieren verschiedene Wertebereiche der peptidinduzierten Signalverschiebungen (siehe Farbskala)

Aus der Übersicht wird deutlich, dass WW3 nur unspezifisch mit SF1\_P1, SF1\_P3 und CA\_KP2 wechselwirkt, denn zu keiner der „xP(2)“-Taschen ist eine Interaktion zu beobachten. Es konnte festgestellt werden, dass SF1\_P4 nur unter Verwendung der „xP“-

Tasche bindet. Interessanterweise benutzen CA\_P1 und SF1\_KP1 neben der „xP“- auch die postulierte „xP2“-Bindetasche von WW3. Es besteht daher eine große Wahrscheinlichkeit, dass die WW3 tatsächlich eine funktionelle „xP2“-Bindetasche besitzt. Wie die NMR-Bindungsstudien mit SF1\_P4 zeigten, ist diese jedoch nicht zwingend für die Bindung prolinreicher Peptidliganden erforderlich. Möglicherweise besitzt die „xP2“-Tasche der WW3 eher die Funktion einer „Spezifitäts-Tasche“. Eine solche Spezifitäts-Tasche besitzt beispielsweise die Dystrophin WW-Domäne, die ebenso wie die CA150-WW3 C-terminal von einer  $\alpha$ -Helix flankiert wird [56] (siehe auch Abbildung 7).

### 5.5.1 MONOVALENTE ODER BIVALENTE BINDUNG

Im Folgenden werden die in dieser Arbeit untersuchten monovalenten und bivalenten Bindungen zwischen prolinreichen Peptidliganden und Tandem-WW23-Varianten vergleichend analysiert. In Tabelle 5.5.II sind die Ergebnisse der NMR- und ITC-Bindungsstudien zusammengestellt. Gezeigt sind die in den jeweiligen Peptidsequenzen enthaltenen bekannten und putativen WW-Bindemotive, die vom Peptid genutzten WW-Bindetaschen („xP“, „xP2“) (aus NMR) und die ermittelten Dissoziationskonstanten (aus ITC).

**Tabelle 5.5.II: Prolinreiche Peptide, PP-Motive, Bindekonstanten und Interaktionsstellen**

Peptid Sequenz	Motive	WW2* K <sub>d</sub> [ $\mu$ M]	WW3** K <sub>d</sub> [ $\mu$ M]	WW23 K <sub>d</sub> [ $\mu$ M]
<b>SF1_P1</b> QPPPPPMNQGP <del>HP</del> PHGPP	xPPPPPx xPHPPx	807 (620-1052) xP; xP2	>> 1 mM - -	452 (352-582) 2: xP, xP2 3: -
<b>SF1_P3</b> MPPPPPPPSGQPPPPPSGPL	xPPPPPPPx xPPPPPx	123 (104-148) xP, xP2	634 (506-808) - -	55 (49-60) 2: xP; xP2 3: -
<b>SF1_P4</b> HPPPPPSHYG <del>PL</del> PPWHQK	xPPPPPx xPLPPx xPPW	59 (51-69) xP2	39 (32-48) xP, xP2	39 (30-50) 2: xP, xP2 3: xP
<b>CA_P1</b> HNAVMRGPPPLMRPPPPF <del>GMMR</del> GPPPPPRPPFG	xPPPPPx xPPPPx (xPPPL)	114 (94-139) xP, xP2	469 (439-503) xP, xP2	109 (99-121) 2: xP2 3: xP, xP2
<b>SF1_KP1</b> VQ <del>PL</del> PPGA	xPLPP PPLPx	+	+	+ 2: xP, xP2 3: xP, xP2
<b>CA_KP2</b> RGPPPPPR	xPPPPPx	+	-	+ 2: xP, xP2 3: -

\* WW23\_W131A; \*\* WW23\_W131A, (+) Bindung mittels NMR nachgewiesen, (-) NMR zeigt keine Bindung  
Zur WW/SF1\_KP1- bzw. WW/CA\_KP2-Wechselwirkung sind keine ITC-Daten vorhanden. x: beliebige Aminosäure, ausgenommen Prolin

Die WW/SF1\_P1-Interaktionsstudien zeigten, dass SF1\_P1 bivalent deutlich stärker an WW23 ( $K_d = 452 \mu\text{M}$ ), als monovalent an WW2 ( $K_d = 807 \mu\text{M}$ ) bindet. Trotz der geringen Affinitäten ist die Interaktion mit WW2 spezifisch, da beide WW2-Bindetaschen („xP“ und „xP2“) vom Peptid genutzt werden. Die Interaktion mit WW3 ist dagegen eher unspezifisch, da die WW3-„xP“-Tasche weder in der mono- noch in der bivalenten Bindungssituation genutzt wird. Der kleine Beitrag dieser unspezifischen Wechselwirkung ermöglicht jedoch eine bivalente Wechselwirkung die sich positiv auf die effektive Konzentration des Liganden auswirkt und dadurch die Affinität nahezu verdoppelt. SF1\_P3 zeigte einen ähnlichen Bindemechanismus. Auch dieses Peptid bindet spezifisch an beide WW2-Bindetaschen und interagiert unspezifisch mit WW3. Das führt, analog zu SF1\_P1, zu einer Verdopplung der Affinität in der bivalenten Bindungssituation, wobei die in SF1\_P3 enthaltenen prolinreichen Motive deutlich affiner an die WW-Domänen binden. Wie bereits in Abschnitt 5.3 diskutiert wurde, interagiert auch SF1\_P4 bivalent mit WW23. Gegenüber der monovalenten Interaktion führte das aber nicht zu einer Affinitäts-erhöhung. Im Unterschied zu SF1\_P1 und SF1\_P3 bindet das Peptid durch Nutzung der „xP“ und/oder „xP2“-Bindetaschen jedoch spezifisch und deutlich stärker an WW3. Sind beide WW-Domänen aktiv, bindet SF1\_P4 außerdem sterisch anders an die Domänen, als in den jeweiligen monovalenten Bindungssituationen. Im letztgenannten Fall bindet es an die „xP2“-Tasche der WW2 bzw. an die „xP“- und „xP2“-Tasche der WW3. Bivalent werden dagegen beide „xP“-Taschen der WW2, jedoch lediglich die „xP“-Tasche der WW3 benutzt. Warum SF1\_P4 monovalent nicht an die WW2-„xP“-Tasche bindet, ist indes unklar. In der bivalenten Bindungssituation können die Interaktionsstellen beider WW-Domänen möglicherweise nicht optimal bedient werden. Die bivalente Interaktion ist den NMR-Daten zufolge zwar bevorzugt, aber nicht wesentlich stärker, als die rein monovalente Wechselwirkung. SF1\_P4 enthält neben dem bekannten WW2-Bindemotiv xPPPPPx ein putatives WW-Bindemotiv xPLPP. Bislang konnte nur die Bindung von WW2 an die reverse Motivvariante PPLPx gezeigt werden [96,98]. Wie die in dieser Arbeit durchgeführte SF1<sub>PRS</sub>-CA150<sub>WW</sub>-SPOT-Analyse (siehe Abschnitt 3.6.4, Tabelle 4.6.II) indizierte, erkennen WW2 und WW3 mit hoher Wahrscheinlichkeit aber auch das xPLPP-Motiv. Der Mechanismus der bivalenten CA\_P1/WW23-Wechselwirkung zeigt Analogie zur bivalenten SF1\_P4/WW23-Interaktion in der Art, dass auch hier keine signifikante Affinitätserhöhung gegenüber der monovalenten Bindung detektiert wurde. Wie SF1\_P4, bindet auch CA\_P1 spezifisch an beide WW-Domänen, nutzt jedoch, abgesehen von der monovalenten Interaktion mit WW3, die vorhandenen Bindetaschen nach einem anderen Muster. Bei monovalenter Wechselwirkung bindet es stets an beide Bindetaschen der jeweils aktiven WW-Domäne, interagiert in der bivalenten Bindungssituation aber nicht mit der „xP“-Tasche in WW2. Interessanterweise bleibt die sehr schwache Wechselwirkung zwischen CA\_P1 und WW3 ( $K_d = 469 \mu\text{M}$ ) bei der bivalenten Interaktion erhalten, während die monovalent deutlich stärkere CA\_P1/WW2-Wechselwirkung ( $K_d = 114 \mu\text{M}$ ) durch Kontaktverlust zur WW2-„xP“-Tasche geschwächt wird. In ihrer Summe sind die intrinsisch schwachen Assoziationskräfte zwischen den prolinreichen Motiven und den beiden WW-Domänen zwar kaum größer, als die Kräfte der monovalenten CA\_P1/WW2-Assoziation, dennoch spiegelt sich hier deutlich das Multivalenzprinzip wieder, die



Vervielfachung der Gesamtaffinität durch die Kombination intrinsisch schwacher Einzelaffinitäten. Bei Betrachtung der CA\_P1-Sequenz fällt auf, dass darin neben dem bekannten WW2-Motiv GPPPPR zwei weitere putative prolinreiche Motive (xPPPPx und xPPPL) vorhanden sind, von denen zumindest eines von der WW3-Domäne spezifisch erkannt werden muss. Das bekannte WW2-Motiv scheidet von vornherein aus, da mittels NMR-Bindungsstudie gezeigt werden konnte, dass das monovalente Peptid CA\_KP2 (enthält das WW2-Motiv GPPPPR) nicht an WW3 bindet. Den Ergebnissen aus den Interaktionsstudien mit SF1\_P1 und SF1\_P3 folgend, kann jedoch nicht ausgeschlossen werden, dass WW3 generell keine xPPPPx-Motive bindet. Offensichtlich ist nur, dass falls es mit diesen Motiven wechselwirkt, diese Interaktion unspezifisch sein muss. Wie die NMR-Bindungsstudien mit SF1\_KP1 zeigten, binden beide WW-Domänen zudem das prolinreiche Motiv „PPLPP“. Um die Präferenzen aller drei CA150-WW-Domänen für verschiedene prolinreiche Motive genauer einzugrenzen, werden im folgenden Abschnitt die Bindungsmuster aller untersuchten prolinreichen Peptidsequenzen und WW-Domänen analysiert.

### **5.5.2 BINDEMOTIVE DER EINZEL-WWs (WW1, 2&3)**

Bezüglich der Bindeeigenschaften der WW-Domänen von CA150 gibt es nur für WW1 und WW2 genauere Informationen. Demnach interagiert WW1 bevorzugt mit Y, poY, poS, poT/R-Peptiden ((L/P)Pp(Y/poY)) und wurde dementsprechend den Klassen I und IV zugeordnet [80]. Auf Grund der Tatsache, dass diese Motive in den untersuchten PRS nicht vorkommen, war es überraschend, dass über die in dieser Arbeit durchgeführten SPOT-Analysen (Abschnitt 4.6.1 Abbildung 29) auch Interaktionen von WW1 zu mehreren Sequenzbereichen der SF1<sub>PRS</sub> gezeigt werden konnten. Die WW2-Domäne von CA150 besitzt Bindeeigenschaften der Klassen II und V, bindet also bevorzugt PPLPp-Motive und reine Prolin-Motive (p/ Φ)PPPPP [80,96–98]. Besonders die reinen Prolin-Motive, bestehend aus mindestens 5 aufeinanderfolgenden Prolinen, sind relativ häufig in der SF1<sub>PRS</sub> vertreten, was eine Vielzahl der beobachteten WW2-Interaktionen zu verschiedenen SF1<sub>PRS</sub>-Bereichen erklärt. Die CA150-WW3 konnte bisher nicht eindeutig in eine der WW-Klassen eingeordnet werden, da sie laut *Hu et al.* eher promiskuitiv bindet. *Hu et al.* beobachteten dennoch, dass die WW3 viele Eigenschaften von WW-Domänen der Klasse I (binden an (L/P)Pp(Y/poY)-Motive) aufweist [98]. Diese Motive sind in den untersuchten PRS-Bereichen (SF1 & CA150) nicht enthalten, dennoch konnten Interaktionen der WW3 mit mehreren Sequenzbereichen beobachtet werden (vgl. Abschnitt 4.6.1 & 4.8.1). In der SF1-SPOT-Analyse konnte gezeigt werden, dass WW2 und WW3 neben dem für WW2 bekannten PPLPp Motiv [96,98] ebenfalls in der Lage sind, das reverse Motiv xPLPP zu binden (vgl. Tabelle 5.5.III). Starke Interaktionen von WW1 und WW3 konnten auch zu diversen Sequenzbereichen (z.B. SF1<sub>361-380</sub>) beobachtet werden, die kein bisher bekanntes WW-Motiv enthalten. Folglich müssen WW1 und WW3 in der Lage sein, ein weiteres PP-Motiv zu binden. Zur Identifizierung potentiell neuer WW-Bindemotive wurden in Tabelle 5.5.III alle Ergebnisse der WW-Bindungsstudien zusammengetragen und hinsichtlich bekannter und putativer WW-Bindemotive ausgewertet.

Tabelle 5.5.III: Bindungsmotive der WW-Bindungsstudien

		xPPPx	xPPx	xPPx	PHPP	PPW	xPLPx	xLPP	PPLPx	PPLPP	(p/Φ)PPPPP	WW1	WW2	WW3	WW23
SF1 <sub>361-380</sub>	TTQSRPPW <u>M</u> NSGPSESRYH														
SF1 <sub>418-437</sub>	NGPPPPW <u>M</u> QPPPPMNQGP														
SF1 <sub>426-445</sub>	QPPPPMNQGP <u>H</u> PPGHHGPP														
SF1 <sub>434-453</sub>	QGPHPPGHHGPP <u>M</u> DQYLGS														
SF1 <sub>458-477</sub>	SGVYRLHQKGM <u>M</u> PPPPMGM														
SF1 <sub>469-488</sub>	MMPPPPMGMMPPPPPPSGQ														
SF1 <sub>478-497</sub>	MPPPPPPSGQPPPPPSGPL														
SF1 <sub>488-507</sub>	QPPPPSGPLPPWQQQQQP														
SF1 <sub>497-516</sub>	LPPWQQQQQPPPPPPSSS														
SF1 <sub>500-519</sub>	WQQQQQPPPPPPSSSMAS														
SF1 <sub>507-526</sub>	PPPPPPSSSMASSTLPWQ														
SF1 <sub>541-560</sub>	IPPWQQQAAAAASPGAPQM														
SF1 <sub>545-564</sub>	QQQAAAAASPGAPQMGNP														
SF1 <sub>566-585</sub>	MVPLPPGVQPPLPPGAPPP														
SF1 <sub>574-593</sub>	QPPLPPGAPPPPPPPGSA														
SF1 <sub>581-600</sub>	APPPPPPPPSAGMMYAPP														
SF1 <sub>598-617</sub>	APPPPPPPMDPSNFVMMG														
SF1 <sub>620-639</sub>	VAGMPPEGMPPAPPPPPQN														
SF1_P4	HPPPPSHYGPLPPWHQK														
SF1_KP1	VQPPLPPGA														
CA <sub>29-48</sub>	FRGPAPPNAVMRGPPPLMR														
CA <sub>39-58</sub>	VMRGPPPLMRPPPPFGMMRG														
CA <sub>46-65</sub>	LMRPPPPFGMMRGPPPPRP														
CA <sub>58-77</sub>	GPPPPRPPFGRPPFDPNMP														
CA <sub>66-85</sub>	PFGRPPFDPNMPPMPPGGI														
CA <sub>77-96</sub>	PPMPPGGI <sup>x</sup> PPPMGPPHLQR														
CA <sub>68-105</sub>	PPPMGPPHLQRPPFMPPMS														
CA <sub>96-115</sub>	RPPFMPPMSMPPPPGMMF														
CA <sub>102-121</sub>	PPMSSMPPPPGMMFPPGMPP														
CA_P1	HNAVMRGPPPLMRPPPPFGMM RGPPPPRPPFG														
CA_KP2	RGPPPPR														

Gezeigt sind mögliche Bindemotive der WW-Domänen (oben). Links: Die untersuchten Prolinreichen Sequenzen. Die Zahlen repräsentieren die Sequenzposition in SF1 bzw. CA150 (CA). Die bekannten bzw. putativen PP-Motive sind in der Peptidsequenz unterstrichen. Mitte: Die vorhandenen Motive (oben) sind farblich markiert (blau: putatives WW-Motiv, orange: bekanntes WW-Motiv). x: beliebige Aminosäure, ausgenommen Prolin; Φ: hydrophobe Aminosäure. Kleine Buchstaben repräsentieren bevorzugte, aber nicht konservierte Aminosäurereste. Die Nomenklatur der Peptidmotive erfolgte nach *Aasland et al.* 2002 [82]. Rechts: Die Ergebnisse der Bindungsstudien für WW1, WW2, WW3 und WW23. Domäne zeigte keine Bindung (Rot), schwache Bindung (Hellgrün) bzw. mittlere bis starke Bindung (Grün).

Nach manueller Auswertung der in Tabelle 5.5.III aufgeführten Bindungsmuster konnten den WW-Domänen eindeutige Bindepräferenzen zugewiesen werden. Alle prolinreichen Motive, für die zweifelsfrei eine bzw. keine Interaktion mit den entsprechenden WW-Domänen gefunden wurde, sind in Tabelle 5.5.IV zusammengestellt.

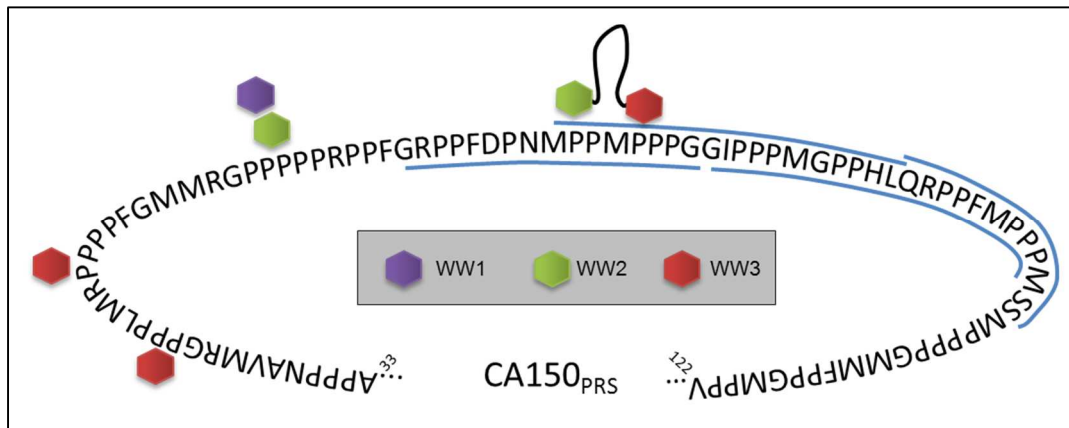
**Tabelle 5.5.IV: Bindemotive der CA150-WW-Domänen**

CA150-WW1	CA150-WW2	CA150-WW3
R P P W	R P P W	I P P W
M P P F	R P P F	L P P W
I P P W	G P P H	G P P H
L P P W	I P P W	R P P F
R P P F	L P P W	G P P P L
G P P H	F P P G	M P P P M
F P P G	G P P P P W	M P P P P M
G P P P P P R	M P P P P M	M P P P P G
Q P P P P P S	M P P P P G	R P P P P F
Q P P P P P M	R P P P P F	Q P P P P P M
A P P P P P P Q	X P P P P P X	G P P P P P R
Q P P P P P P S	P L P P	m P P P P P P S
M P P P P P P S	P H P P	A P P P P P P P M
A P P P P P P P M	P P L P p	A P P P P P P P P G
X P P P P P P P P		P H P P
		P P L P P
		P L P P

Gezeigt sind die für die CA150 WW1, WW2 bzw. WW3 gefundenen Bindepräferenzen. PP-Motive (grau) mit grün markierten Aminosäuren wurden eindeutig gebunden, PP-Motive mit rot markierten Aminosäuren wurden eindeutig nicht gebunden.

Allen Domänen ist gemein, dass sie Motive mit zwei oder mindestens fünf aufeinanderfolgenden Prolinen binden. Es lässt sich jedoch kein generalisiertes Motiv (xPPx, bzw. xPPPPx) ableiten, da die N- und C-terminal flankierenden Aminosäuren bzw. deren Kombinationen eine essentielle Rolle dabei spielen, ob ein PP-Motiv gebunden wird oder nicht. So wird beispielsweise das Motiv RPPW von WW1 und WW2 gebunden, von WW3 jedoch nicht. Wird das Tryptophan an Position 4 durch Phenylalanin ersetzt, kann WW2 das Motiv nicht mehr erkennen. Befindet sich an Position 1 anstelle des Arginins ein Isoleuzin, binden weder WW1 noch WW2, dagegen interagiert jetzt aber WW3 mit dem Motiv. Auf Basis der vorliegenden Ergebnisse können allerdings keine Aussagen darüber getroffen werden, welche Kombinationen von flankierenden Aminosäuretypen (positiv/negativ geladen, hydrophob/polar, aromatisch) vorhanden sein müssen, damit die PP-Motive von den einzelnen WW-Domänen erkannt werden. Um diesbezüglich Aussagen treffen zu können, wäre eine systematische Substitutions-Analyse erforderlich. Motive mit nur drei aufeinanderfolgenden Prolinen, sowie das PHPP-





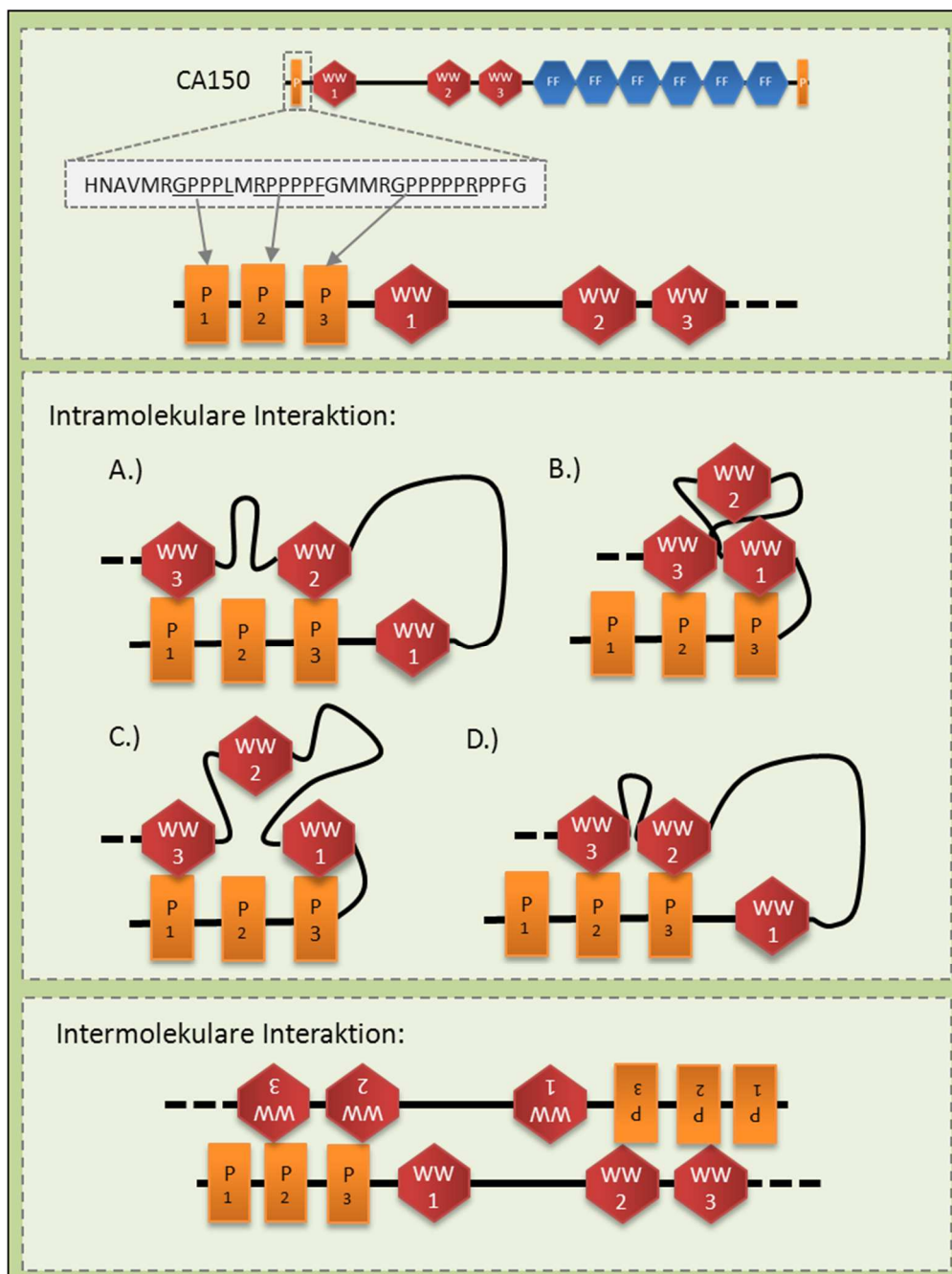
**Abbildung 50: CA150<sub>WW</sub>-Bindepotionen der N-terminalen CA150<sub>PRS</sub>**

Visualisierung der Bindepositionen der jeweiligen WW-Domäne innerhalb der N-terminalen CA150<sub>PRS</sub>. Die WW-Domänen sind durch Sechsecke repräsentiert; WW1: violett, WW2: grün, WW3 rot. Durch eine schwarzen Linie verbundene Domänen repräsentieren das WW23 Tandem. Untersuchte Sequenzbereiche, die ausschließlich vom WW23-Tandem erkannt werden, sind Blau hervorgehoben. Die Zahlen repräsentieren die Sequenzpositionen.

Das WW23-Tandem kann mit prolinreichen Motiven innerhalb fast des gesamten Sequenzbereiches (von 29 bis 107) wechselwirken (vgl. Abschnitt 4.8.1). Besonders interessant ist jedoch die Tatsache, dass die prolinreichen Motive von 66 bis 107 (blau markiert) ausschließlich vom WW23-Tandem erkannt werden, ohne dass hier eine Wechselwirkung mit einer der isolierten Domänen detektiert werden konnte. Die monovalenten Interaktionen scheinen demnach äußerst schwach zu sein, bilden aber die Basis einer deutlich nachweisbaren bivalenten Wechselwirkung. Anhand der identifizierten Bindepräferenzen und der lokalisierten Interaktionsstellen sind mehrere Modelle der CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Wechselwirkung vorstellbar, welche im Folgenden näher betrachtet werden.

### 5.5.3 MODELLE MULTIVALENTER CA150<sub>WW</sub>/CA150<sub>PRS</sub>-WECHSELWIRKUNGEN

Auf Basis der in dieser Arbeit untersuchten Wechselwirkungen zwischen den CA150 WW-Domänen und verschiedenen Teilbereichen SF1<sub>PRS</sub> sowie der N-terminalen CA150<sub>PRS</sub> sind vier Modelle eines intramolekularen und ein Modell eines intermolekularen multivalenten CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Interaktionsmechanismus denkbar. Diese sind in Abbildung 51 schematisch dargestellt.



**Abbildung 51: Modelle multivalenter CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Wechselwirkungen**

Oben: Domänenorganisation von CA150 (P: prolinreiche Region, WW: WW-Domäne; FF: FF-Domäne). Die CA\_P1-Sequenz mit den drei verifizierten Bindemotiven (P1: GPPPL, P2: RPPPPF, P3: GPPPPR) für die Einzel-WW-Domänen (WW1, WW2 bzw. WW3). Darunter: Modelle für eine intramolekulare Interaktion: (A) WW3 bindet an P1, WW2 bindet an P3, WW1 ist frei für andere Interaktionspartner. (B) WW3 bindet an P2, WW1 bindet an P3, WW2 ist frei für andere Interaktionspartner. (C) WW3 bindet an P1, WW1 bindet an P3, WW2 ist frei für andere Interaktionspartner. (D) WW3 bindet an P2, WW2 bindet an P3, WW1 ist frei für andere Interaktionspartner. Unten: Modell einer intermolekularen Wechselwirkung zwischen zwei CA150-Molekülen. WW2 und WW3 binden jeweils an P3 und P2 des anderen Moleküls.

Die WW3-Domäne kann entweder mit dem prolinreichen Motiv P1 (GPPPL) oder P2 (RPPPPF) interagieren, während WW1 und WW2 nur an P3 (GPPPPR) binden können. Eine bivalente Interaktion vorausgesetzt, ergeben sich daraus jeweils zwei Konformationen in denen entweder die WW1-Domäne (Abbildung 51A & B) oder die WW2-Domäne (Abbildung 51C & D) potentiell für weitere Interaktionspartner zur

---

Verfügung stehen. Die jeweils „freie“ WW-Domäne könnte demnach eine regulatorische Funktion besitzen oder durch die bivalente Wechselwirkung der beiden übrigen Domänen in eine günstige Position für die Interaktion mit einem Zielprotein gebracht werden. Weitere Spekulationen wären hier jedoch wenig sinnvoll, da bereits die gezeigten Modelle rein hypothetisch sind. Auf Basis der gefundenen CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Interaktionen wäre auch das in Abbildung 51 dargestellte Modell einer bivalenten intermolekularen Wechselwirkung zwischen zwei CA150-Proteinen denkbar. Da die Zahl der wechselwirkenden Interaktionsstellen hier auf Grund der Spiegelsymmetrie doppelt so hoch ist, wie im Falle einer intramolekularen CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Assoziation, könnte CA150 möglicherweise fest bindende Dimere bilden. Auch hier könnten die „freien“ WW1-Domänen regulatorische Einheiten sein, die durch Interaktion mit prolinreichen Liganden möglicherweise die Dissoziation von dimerem CA150 forcieren. Auch das ist rein spekulativ und wäre nur durch weitere intensive Forschungsbemühungen aufzuklären.

Neuste Forschungsergebnisse zeigten, dass die Serin/Threonin-Proteinkinase PLK1 mit Hilfe ihrer PBD-Domäne (PBD: Polo-Box-Domäne) in der Lage ist, das phosphorylierte FMPPPMSpoSM-Motiv (CA150<sub>99-107</sub>) innerhalb der N-terminalen CA150<sub>PRS</sub> zu erkennen (PDB-ID: 3Q1I) [174]. Entsprechend wäre eine Bedeutung von PLK1 für den hier postulierten intramolekularen Interaktionsmechanismus zu berücksichtigen. Man muss darüber hinaus auch in Betracht ziehen, dass einige Aminosäuren der CA150<sub>PRS</sub> *in vivo* potentiell phosphoryliert sein können und daraus andere Interaktionsmuster oder Bindepräferenzen resultieren können, als in dieser Arbeit erfasst werden konnten.

Für die tandemartig angeordneten WW-Domänen des Nedd4-ähnlichen *Suppressor of Deltex* (Su(dx)) konnte gezeigt werden, dass erst durch die Bindung eines PY-Liganden an WW3, eine autoinhibitorische Domänen-Domänen-Wechselwirkung zwischen WW3 und WW4 freigeschaltet wird. Erst durch die Trennung von WW3 konnte WW4 den vollständigen Faltungszustand einnehmen und war in der Lage, einen PP-Liganden zu binden [122]. Diese Konkurrenz zwischen WW-Domänen-Assoziation und PY-Motiv-Bindung beweist, dass eine tandemartige Organisation der WW-Domänen direkten Einfluss auf ihre Fähigkeit zur Bindung von Liganden nehmen kann. Da auch die WW-Domänen in CA150 tandemartig angeordnet sind, könnten ähnliche Mechanismen auch bei der Regulation von CA150 eine Rolle spielen.

---

## **ZUSAMMENFASSUNG**

Viele humane Krankheiten, wie Alzheimer [89], Krebs [83], Muskeldystrophie [54], „Liddle-Syndrom“ [118] und Huntington [119] können auf fehlregulierte WW-Domänen-vermittelte Signalwege zurückgeführt werden. Erstmals beschrieben wurden WW-Domänen 1994 von *Bork und Sudol* [67], die diesen Domänen auf Grund zweier hochkonservierter Tryptophane (W) innerhalb ihrer Consensus-Sequenz auch ihren Namen verliehen. Mit nur ca. 40 Aminosäuren gehören sie zu den kleinsten Proteindomänen. Sie binden bevorzugt prolinreiche Sequenzmotive (PRS) und sind, mit Ausnahme von Bakterien, hoch konserviert in allen Organismengruppen präsent. Der humane Transkriptionsfaktor CA150 (*Synonym: TCERG1 transcription elongation regulator 1*) ist eines dieser Proteine. Es besitzt neben sechs FF-Domänen und zwei N- und C-terminal lokalisierten PRS, gleich drei WW-Domänen, die durch außergewöhnlich lange Linkersequenzen miteinander verknüpft sind. CA150 interagiert mit verschiedensten Proteinen und ist so an der Regulation der Elongation der RNA-Polymerase II, der Koordinierung der Transkription und der prä-mRNA-Prozessierung beteiligt. Dementsprechend konnten unter anderem Interaktionen der WW- und FF-Domänen zur RNA-Polymerase II [112,116] sowie Wechselwirkungen der WW-Domänen (1&2) mit den PRS des Spleißfaktors SF1 [112] nachgewiesen werden. Darüber hinaus kann CA150 auch mit der mutierten Form des Huntingtin-Proteins interagieren, wodurch dem Protein eine wichtige Rolle im Krankheitsverlauf von Huntington zukommt [87]. Es besteht daher ein großes Interesse, die molekularen Mechanismen CA150-vermittelter Wechselwirkungen besser zu verstehen. Die Vielzahl der bisher identifizierten CA150-Interaktionspartner und das mehrfache Vorhandensein gleicher Funktionseinheiten legen nahe, dass CA150 mit vielen Zielproteinen multivalent wechselwirken könnte. Die Multivalenz könnte zum einen über das Tandem der drei aufeinanderfolgenden WW-Domänen und zum anderen über die beiden PRS realisiert sein. Der Fokus dieser Arbeit lag insbesondere auf der systematischen Analyse von Interaktionen der isolierten WW-Domänen und Tandem-WW-Konstrukten aus CA150 mit Teilbereichen der PRS aus SF1 (SF1<sub>PRS</sub>) sowie Teilbereichen der N-terminalen PRS aus CA150 selbst (CA150<sub>PRS</sub>). Um ein besseres Verständnis grundlegender Mechanismen der multivalenten Wechselwirkungen zu erlangen, aber auch um Erkenntnisse bezüglich der Organisation der drei WW-Domänen in CA150 zu gewinnen, wurde eine Vielzahl molekular-biologischer, biochemischer und biophysikalischer Methoden angewendet.

Die codierenden Sequenzen verschiedener CA150 WW-Varianten wurden in verschiedene Vektoren kloniert. Zur Maximierung der Ausbeuten unmarkierter bzw. isotoopenmarkierter Proteine wurden die Expressionsbedingungen optimal an die zu exprimierenden WW-Varianten angepasst. Für diverse GST-WW und His-WW-Konstrukte wurde eine IMAC-basierte Reinigungsstrategie entwickelt. Durch SILAC-Pulldown-Experimente konnte die Aktivität der rekombinanten WW-Domänen verifiziert werden,



indem bekannte sowie neue, bisher unbekannte, Interaktionspartner gefunden wurden. Der SILAC-Pulldown lieferte auch erste Hinweise dafür, dass CA150 mit sich selbst interagieren könnte. In einer SPOT-Analyse wurde die Interaktion zwischen Teilbereichen der N-terminalen CA150<sub>PRS</sub> und den drei CA150-WW-Domänen verifiziert. Darüber hinaus konnte mittels AUZ gezeigt werden, dass die bivalente Bindung einer ausgewählten CA150<sub>PRS</sub>-Sequenz dazu führt, dass ein WW23-Tandem eine Konformationsänderung erfährt. Eine artifizielle Oligomerisierung des WW23-Tandems konnte durch die AUZ ausgeschlossen werden. SPOT-Analysen mit SF1<sub>PRS</sub>-Teilbereichen zeigten neben mehreren Interaktionen mit WW1 und WW2, auch verschiedene Wechselwirkungen mit WW3. Durch die Aufnahme von HSQC-basierten Triple-Resonanz-Spektren von <sup>13</sup>C<sup>15</sup>N-markiertem WW23-Protein konnten 75% der <sup>1</sup>H<sup>15</sup>N-Korrelationen sequentiell eindeutig zugeordnet werden. Die Resonanzen der einzelnen WW-Domänen wurden fast vollständig identifiziert. Durch die Sequentielle Zuordnung war es möglich, die dynamischen Eigenschaften des Proteinerückgrades des WW23-Tandems NMR-spektroskopisch zu ermitteln. Dabei wurden neue Domänengrenzen für WW2 und WW3 aufgedeckt. Die strukturdynamischen Untersuchungen in Kombination mit sequenzbasierten (Pspred) und „*chemical shift*“-basierten (Talos+) Sekundärstrukturvorhersagen zeigten, dass der Linker zwischen WW2 und WW3 aus einem zentralen unstrukturierten Abschnitt besteht, der von zwei  $\alpha$ -helikalen Bereichen flankiert wird. Die sequentielle Zuordnung erlaubte zusammen mit den NMR-Bindungsstudien auch die Identifizierung der Interaktionsstellen für bestimmte bivalente Peptide. Dadurch konnte gezeigt werden, welche WW-Bindetasche („xP“, „xP2“) in der jeweiligen Bindungssituation vom Peptid genutzt wird und zusätzlich die Existenz einer postulierten „xP2“-Tasche innerhalb der WW3 bestätigt werden. Um in Bindungsstudien mono- und bivalente Wechselwirkungen eindeutig identifizieren zu können, war es notwendig, Mutanten vom WW23-Tandem zu erzeugen, in denen die Bindungsaktivität einer Domäne durch eine Punktmutation in der „xP“-Tasche ausgelöscht wurde. Der Einsatz der Mutanten in ITC-Bindungsstudien ermöglichte es, die Affinitäten von mono- und bivalenten Interaktionen zu ermitteln und vergleichend zu analysieren. In den ITC-Bindungsstudien wurden durchweg schwache Affinitäten detektiert. Aus den Ergebnissen der Interaktionsanalysen konnten zwei Mechanismen bivalenter Wechselwirkungen abgeleitet werden. Ersterer führte durch Kombination einer spezifischen und einer unspezifischen Wechselwirkung zu einer deutlichen Affinitätserhöhung. Der Zweite basiert auf zwei spezifischen Wechselwirkungen. Diese führten jedoch nicht zu einer verstärkten, sondern zu einer veränderten Interaktion in der Art, dass in der bivalenten Bindungssituation mindestens eine WW-Domäne über andere Oberflächenbereiche wechselwirkt, als bei der monovalenten Bindung. Durch Kombination der Daten der Interaktionsanalysen konnten auch die Bindepräferenzen der WW-Domänen ermittelt und bisher unbekannte prolinreiche Motive identifiziert werden. So konnte gezeigt werden, an welchen Positionen der CA150<sub>PRS</sub> die WW-Domänen binden können. Darauf aufbauend konnten hypothetische Modelle bivalent intramolekularer bzw. tetravalent intermolekularer CA150<sub>WW</sub>/CA150<sub>PRS</sub>-Interaktionsmechanismen abgeleitet werden.

## **SUMMARY**

Diverse human diseases, such as Alzheimer's [89], cancer [83], muscular dystrophy [54], „Liddle-syndrome“[118] and Huntington [119] have been implicated with miss regulated WW domain mediated pathways. WW domains were described first by *Bork und Sudol* [67] in 1994. Based on the presence of two highly conserved tryptophan residues (W) within the consensus sequence, they were named WW domains. They belong to the smallest protein domains containing on average 40 amino acids. They preferentially bind proline-rich sequence motifs (PRS) and are highly conserved in all organisms, except bacteria. One of these WW domain containing proteins is the human transcription factor CA150 (*synonym: TCERG1 transcription elongation regulator 1*). It consists additionally six FF domains, a N- and a C-terminal PRS as well as three WW domains, which are linked by exceptionally long linker sequences. CA150 is involved in the regulation of the elongation of RNA polymerase II, the coordination of transcription and pre-mRNA processing by interacting with many different proteins. Among others, interactions of the WW and FF domains with the RNA polymerase II [112,116] and interactions of the WW domains (1 & 2) with the PRS of the splicing factor SF1 [112] were detected. Moreover, CA150 also interacts with the mutated form of the huntingtin protein (HTT), whereby the protein plays an important role in huntington's disease [87].

It is of high medical interest to elucidate the molecular mechanisms of CA150 mediated interactions. Both, many previously identified CA150-interacting proteins and the presence of multiple equal functional units suggests that CA150 might interact multivalent with many target proteins. The multivalency could be realized on the one hand through the tandem of the three successive WW domains and on the other hand on the two PRS.

This work focuses mainly on the systematic analysis of interactions between isolated or tandem CA150 WW domain constructs and PRS motifs of SF1 (SF1<sub>PRS</sub>) as well as motifs of the N-terminal CA150<sub>PRS</sub>. To gain a better understanding of basic mechanisms of multivalent interactions, but also to get insights into the organization of the three WW domains in CA150, a variety of molecular biological, biochemical and biophysical methods have been used.

The coding sequences of different CA150 WW variants were cloned into different *E. coli* expression vectors. In order to yield efficient amounts of unlabeled and isotopic labeled WW variants, the expression conditions have been optimized. For various GST-WW and His-WW constructs, an IMAC-based purification strategy was developed. The activity of recombinant WW domains could be verified by SILAC pull-down experiments, by identifying known and new, previously unknown, interaction partners. The SILAC pull-down also provided the first evidence that CA150 might interact with itself. Furthermore, the interaction between sub-regions of the N-terminal CA150<sub>PRS</sub> and the three CA150 WW domains was verified by SPOT analysis experiments. Moreover, it was shown by means of analytical ultracentrifugation (AUC) that the bivalent binding of a selected CA150<sub>PRS</sub>

---

sequence causes a conformational change of a WW23 tandem. An artificial oligomerization of the WW23 tandems could be screened out by the AUC measurements. SPOT analysis with SF1<sub>PRS</sub> sub-regions showed besides several interactions with WW1 and WW2 also various interactions with WW3. By recording of HSQC-based triple-resonance spectra of <sup>13</sup>C<sup>15</sup>N-labeled WW23 protein, 75% of the <sup>1</sup>H<sup>15</sup>N-correlations were sequentially assigned unambiguously. The resonances of the individual WW domains were identified almost completely. The sequential assignment enabled the determination of dynamic characteristics of the protein-backbone of a WW23 tandem via solution-state NMR spectroscopy. Thereby new domain boundaries for WW2 and WW3 were revealed. The structural dynamic studies in combination with sequence-based (Ppsipred) and chemical shift-based (Talos+) secondary structure prediction showed that the linker between WW2 and WW3 consists of a central unstructured section flanked by two  $\alpha$ -helical regions. Both, the sequential assignment and the NMR binding studies, allowed the identification of interaction sites for specific bivalent peptides. It could be shown that the WW domain binding grooves ("xP", "xP2") are occupied by the bound proline-rich ligand in particular situations. Thereby the existence of a postulated "xP2"-groove in WW3 could be confirmed.

For the exact determination of mono- and bivalent interactions in different binding studies, mutants of the WW23 tandem, in which the binding activity of one domain was deleted by a point mutation in the "xP"-groove, were generated. ITC binding studies with the WW23-mutants were used to determine the binding affinities of mono- and bivalent interactions and allowed direct comparison. In the ITC consistently weak affinities were detected.

Finally, two mechanisms of bivalent interactions could be detected by interaction analysis. The first mechanism is based on the combination of specific and non-specific interactions that results in a significant increase of affinity. The second one underlies two specific interactions and did not lead to an increase of affinity. Rather the nature of interaction changes in the way that in the bivalent binding situation the tri- or bivalent proline-rich ligand attaches other surface regions of at least one WW domain, in contrast to the monovalent binding.

By combining the data of the interaction analyses also the binding preferences of the WW domains could be identified and previously unknown proline-rich motifs could be uncovered.

Thereby, it could be shown as well, to which positions of the N-terminal CA150<sub>PRS</sub> the WW domains bind. Based on this information, hypothetical models could be derived for bivalent intramolecular or tetravalent intermolecular CA150<sub>WW</sub>/CA150<sub>PRS</sub>-interactions.

---

---

## **LITERATUR**

1. Stumpf MPH, Thorne T, Silva E De, Stewart R, An HJ, et al. (2008) Estimating the size of the human interactome. *PNAS* 105: 6959–6964.
  2. Ryan DP, Matthews JM (2005) Protein-protein interactions in human disease. *Current opinion in structural biology* 15: 441–446.
  3. Strynadka NCJ, James MNG (1991) Towards an understanding of the effects of calcium on protein structure and function. *Curr Opin Struc Biol* 1: 905–914.
  4. Clore GM, Bax A, Ikura M, Gronenborn AM (1993) Structure of calmodulin-target peptide complexes. *Curr Opin Struc Biol* 3: 838–845.
  5. Benesch R, Benesch RE (1967) The Effect of organic Phosphates from the human Erythrocyte. *Biochem Biophys Res Commun* 26: 162–167.
  6. Van Wijk R, Van Solinge WW (2005) The energy-less red blood cell is lost: erythrocyte enzyme abnormalities of glycolysis. *Blood* 106: 4034–4042.
  7. Ginty DD, Kornhauser JM, Thompson M a, Bading H, Mayo KE, et al. (1993) Regulation of CREB phosphorylation in the suprachiasmatic nucleus by light and a circadian clock. *Science (New York, NY)* 260: 238–241.
  8. Richardson JS (1981) The anatomy and taxonomy of protein structure. *Advances in protein chemistry* 34: 167–339.
  9. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.
  10. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. (2001) The sequence of the human genome. *Science (New York, NY)* 291: 1304–1351.
  11. Rose GD (1979) Hierarchic organization of domains in globular proteins. *J Mol Biol* 134: 447–470.
  12. Patthy L (1994) Introns and exons. *Curr Opin Struc Biol* 4: 383–392.
  13. Krishnamurthy VM, Estroff LA, Whitesides GM (2006) Multivalency in Ligand Design. *Fragment-based Approaches in Drug Discovery*. pp. 11–54.
  14. Sriram SM, Banerjee R, Kane RS, Kwon YT (2009) Multivalency-assisted control of intracellular signaling pathways: application for ubiquitin-dependent N-end rule pathway. *Chemistry & Biology* 16: 121–131.
  15. Choi S-K (2004) *Synthetic Multivalent Molecules: Concepts and Biomedical Applications*. John Wiley & Sons, Inc. p.
  16. Sabatini DM, Erdjument-Bromage H, Lui M, Tempst P, Snyder SH (1994) RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell* 78: 35–43.
  17. De Graffenried CL, Laughlin ST, Kohler JJ, Bertozzi CR (2004) A small-molecule switch for Golgi sulfotransferases. *Proc Natl Acad Sci USA* 101: 16715–16720.
-

- 
18. Haruki H, Nishikawa J, Laemmli UK (2008) The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Mol Cell* 31: 925–932.
  19. Mallet VO, Mitchell C, Guidotti J-E, Jaffray P, Fabre M, et al. (2002) Conditional cell ablation by tight control of caspase-3 dimerization in transgenic mice. *Nature biotechnology* 20: 1234–1239.
  20. Janse DM, Crosas B, Finley D, Church GM (2004) Localization to the proteasome is sufficient for degradation. *J Biol Chem* 279: 21415–21420.
  21. Schwartz EC, Saez L, Young MW, Muir TW (2007) Post-translational enzyme activation in an animal via optimized conditional protein splicing. *Nature chemical biology* 3: 50–54.
  22. Mammen M, Choi S, Whitesides GM (1998) Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angewandte Chemie (International ed in English)* 37: 2754–2794.
  23. Kitov PI, Bundle DR (2003) On the nature of the multivalency effect: a thermodynamic model. *JACS* 125: 16271–16284.
  24. Kiessling LL, Gestwicki JE, Strong LE (2000) Synthetic multivalent ligands in the exploration of cell-surface interactions. *Curr Opin Chem Biol* 4: 696–703.
  25. Luigi, Mandolini (1987) Intramolecular Reactions of Chain Molecules. *Advances in Physical Organic Chemistry* 22: 1–111.
  26. Kramer RH, Karpen JW (1998) Spanning binding sites on allosteric proteins with polymer-linked ligand dimers. *Nature* 395: 710–713.
  27. Rao J (1998) A Trivalent System from Vancomycin-d-Ala-d-Ala with Higher Affinity Than Avidin·Biotin. *Science* 280: 708–711.
  28. Rao J, Lahiri J, Weis RM, Whitesides GM (2000) Design, Synthesis, and Characterization of a High-Affinity Trivalent System Derived from Vancomycin and L-Lys- D-Ala-D-Ala. *J Am Chem Soc* 122: 2698–2710.
  29. Kiessling LL, Gestwicki JE, Strong LE (2006) Synthetic Multivalent Ligands as Probes of Signal Transduction. *Angewandte Chemie (International ed in English)* 45: 2348–2368.
  30. Huskens J (2006) Multivalent interactions at interfaces. *Curr Opin Chem Biol* 10: 537–543.
  31. Basha S, Rai P, Poon V, Saraph A, Gujraty K, et al. (2006) Polyvalent inhibitors of anthrax toxin that target host receptors. *Proc Natl Acad Sci USA* 103: 13509–13513.
  32. Whitesides GM, Mathias JP, Seto CT (1991) Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures. *Science (New York, NY)* 254: 1312–1319.
  33. Lehn J-M (2002) Toward complex matter: supramolecular chemistry and self-organization. *Proc Natl Acad Sci USA* 99: 4763–4768.
  34. Lehn J-M (2002) Toward self-organization and complex matter. *Science (New York, NY)* 295: 2400–2403.
  35. Whitesides GM (2005) Nanoscience, nanotechnology, and chemistry. *Small* 1: 172–179.
  36. Philp D, Stoddart JF (1996) Self-Assembly in Natural and Unnatural Systems. *Angewandte Chemie International Edition in English* 35: 1154–1196.
-

- 
37. Spaltenstein A, Whitesides GM (1991) Polyacrylamides bearing pendant  $\alpha$ -sialoside groups strongly inhibit agglutination of erythrocytes by influenza virus. *J Am Chem Soc* 113: 686–687.
  38. Choi SK, Mammen M, Whitesides GM (1996) Monomeric inhibitors of influenza neuraminidase enhance the hemagglutination inhibition activities of polyacrylamides presenting multiple C-sialoside groups. *Chemistry & biology* 3: 97–104.
  39. Qu L, Dai L, Stone M, Xia Z, Wang ZL (2008) Carbon Nanotube Arrays with Strong Binding-On and Easy Normal Lifting-Off. *Science* 322: 238–242.
  40. Chen H, Privalsky ML (1995) Cooperative formation of high-order oligomers by retinoid X receptors: an unexpected mode of DNA recognition. *Proceedings of the National Academy of Sciences of the United States of America* 92: 422–426.
  41. Yu H, Chen JK, Feng S, Dalgarno DC, Brauer a W, et al. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* 76: 933–945.
  42. Sparks AB, Ridert JE, Hoffmant NG, Fowlkest DM, Quilliam LA, et al. (1996) Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cartactin, p53bp2, PLC $\gamma$ , Crk, and Grb2. *Proc Natl Acad Sci U S A* 93: 1540–1544.
  43. Reinhard M, Rüdiger M, Jockuschb BM, Walter U (1996) VASP interaction with vinculin: a recurring theme of interactions with proline-rich motifs. *FEBS Letters* 399: 103–107.
  44. Reinhard M, Giehl K, Abel K, Haffner C, Jarchau T, et al. (1995) The proline-rich focal adhesion and microfilament protein VASP is a ligand for profilins. *The EMBO journal* 14: 1583–1589.
  45. Reinhard M, Jouvenal K, Tripier D, Walter U (1995) Identification, purification, and characterization of a zyxin-related protein that binds the focal adhesion and microfilament protein VASP (vasodilator-stimulated phosphoprotein). *Proc Natl Acad Sci USA* 92: 7956–7960.
  46. Freund C, Dötsch V, Nishizawa K, Reinherz EL, Wagner G (1999) The GYF domain is a novel structural fold that is involved in lymphoid signaling through proline-rich sequences. *Nature structural biology* 6: 656–660.
  47. Nishizawa K, Freund C, Li J, Wagner G, Reinherz EL (1998) Identification of a proline-binding motif regulating CD2-triggered T lymphocyte activation. *Proc Natl Acad Sci USA* 95: 14897–14902.
  48. Pornillos O, Alam SL, Rich RL, Myszkka DG, Davis DR, et al. (2002) Structure and functional interactions of the Tsg101 UEV domain. *The EMBO journal* 21: 2397–2406.
  49. Pornillos O, Alam SL, Davis DR, Sundquist WI (2002) Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. *Nature structural biology* 9: 812–817.
  50. Mahoney NM, Janmey P a, Almo SC (1997) Structure of the profilin-poly-L-proline complex involved in morphogenesis and cytoskeletal regulation. *Nat Struct Biol* 4: 953–960.
  51. Schutt CE, Myslik JC, Rozycki MD, Goonesekere NCW, Lindberg U (1993) The structure of crystalline profilin- beta-actin. *NATURE* 365: 810–816.
  52. Sudol M, Chen HI, Bougeret C, Einbond A, Bork P (1995) Characterization of a novel protein-binding module - the WW domain. *FEBS letters* 369: 67–71.
  53. Macias MJ, Hyvönen M, Baraldi E, Schultz J, Sudol M, et al. (1996) Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* 382: 646–649.
  54. Huang X, Poy F, Zhang R, Joachimiak A, Sudol M, et al. (2000) Structure of a WW domain containing fragment of dystrophin in complex with  $\beta$ -dystroglycan. *Nature structural biology* 7.
-

- 
55. Ball LJ, Kühne R, Schneider-Mergener J, Oschkinat H (2005) Recognition of Proline-Rich Motifs by Protein-Protein-Interaction Domains. *Angewandte Chemie (International ed in English)* 44: 2852–2869.
  56. Zarrinpar A, Bhattacharyya RP, Lim WA (2003) The Structure and Function of Proline Recognition Domains. *Science's STKE* 2003: RE8.
  57. Cowan PM, McGavin S (1955) Structure of Poly-L-Proline. *Nature* 176: 501–503.
  58. Fraser RDB, Macrae TP, Suzuki E (1979) Chain Conformation in the Collagen Molecule. *J Mol Biol* 129: 463–481.
  59. Williamson MP (1994) The structure and function of proline-rich regions in proteins. *The Biochemical Journal* 297: 249–260.
  60. Darbon H, Bernassau JM, Deleuze C, Chenu J, Roussel A, et al. (1992) Solution conformation of human neuropeptide Y by 1H nuclear magnetic resonance and restrained molecular dynamics. *FEBS* 209: 765–771.
  61. Blundell TL, Pitts JE, Tickle IJ, Wood SP, Wu CW (1981) X-ray analysis (1.4-Å resolution) of avian pancreatic polypeptide: Small globular protein hormone. *Proc Natl Acad Sci USA* 78: 4175–4179.
  62. Adzhubei A, Sternberg MJ (1993) Left-handed polyproline II helices commonly occur in globular proteins. *J Mol Biol* 229: 472–493.
  63. Hagerman AE, Butler LG (1981) The specificity of proanthocyanidin-protein interactions. *J Biol Chem* 256: 4494–4497.
  64. Murray NJ, Williamson MP, Lilley TH, Haslam E (1994) Study of the interaction between salivary proline-rich proteins and a polyphenol by 1H-NMR spectroscopy. *FEBS* 219: 923–935.
  65. Zamminer J, Brockmann C, Huy P, Opitz R, Reuter C, et al. (2010) Addressing protein-protein interactions with small molecules: a Pro-Pro dipeptide mimic with a PPII helix conformation as a module for the synthesis of PRD-binding ligands. *Angewandte Chemie (International ed in English)* 49: 7111–7115.
  66. Freund C, Schmalz H-G, Sticht J, Kühne R (2008) Proline-rich sequence recognition domains (PRD): ligands, function and inhibition. *Handbook of experimental pharmacology*: 407–429.
  67. Bork P, Sudol M (1994) The WW domain: a signalling site in dystrophin? *Trends Biochem Sci* 19: 531–533.
  68. André B, Springael J-Y (1994) WWP, a New Amino Acid Motif Present in Single or Multiple Copies in Various Proteins Including Dystrophin and the SH3-Binding Yes-Associated Protein YAP65. *Biochem Biophys Res Commun* 205: 1201–1205.
  69. Hofmann K, Bucher P (1995) The rsp5-domain is shared by proteins of diverse functions. *FEBS letters* 358: 153–157.
  70. Garnier L, Verderame MF, Sudol M (1996) WW domains and retrovirus budding. *Nature* 381: 744–745.
  71. Garcel A, Gout E, Timmins J, Chroboczek J, Fender P (2006) Protein transduction into human cells by adenovirus dodecahedron using WW domains as universal adaptors. *J Gene Med* 8: 524–531.
-

- 
72. Itani O a, Campbell JR, Herrero J, Snyder PM, Thomas CP (2003) Alternate promoters and variable splicing lead to hNedd4-2 isoforms with a C2 domain and varying number of WW domains. *Am J Physiol Renal Physiol* 285: F916–929.
  73. Debonneville C, Flores SY, Kamynina E, Plant PJ, Tauxe C, et al. (2001) Phosphorylation of Nedd4-2 by Sgk1 regulates epithelial Na<sup>+</sup> channel cell surface expression. *The EMBO Journal* 20: 7052–7059.
  74. Huang X, Poy F, Zhang R, Joachimiak A, Sudol M, et al. (2000) Structure of a WW domain containing fragment of dystrophin in complex with  $\beta$ -dystroglycan. *America* 7: 634–638.
  75. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic acids research* 40: D302–5.
  76. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc Natl Acad Sci USA* 95: 5857–5864.
  77. Kato Y, Nagata K, Takahashi M, Lian L, Herrero JJ, et al. (2004) Common mechanism of ligand recognition by group II/III WW domains: redefining their functional classification. *J Biol Chem* 279: 31833–31841.
  78. Lu P (1999) Function of WW Domains as Phosphoserine- or Phosphothreonine-Binding Modules. *Science* 283: 1325–1328.
  79. Sudol M, Hunter T (2000) NeW Wrinkles for an Old Domain. *Cell* 103: 1001–1004.
  80. Otte L, Wiedemann URS, Schlegel B, Pires JR, Beyermann M, et al. (2003) WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains. *Protein Science* 12: 491–500.
  81. Macias MJ, Wiesner S, Sudol M (2002) WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Letters* 513: 30–37.
  82. Aasland R, Abrams C, Ampe C, Ball LJ, Bedford MT, et al. (2002) Normalization of nomenclature for peptide motifs as ligands of modular protein domains. *FEBS letters* 513: 141–144.
  83. Bednarek AK, Laflin KJ, Daniel RL, Liao Q, Hawkins KA, et al. (2000) WWOX, a Novel WW Domain-containing Protein Mapping to Human Chromosome 16p23.3-24.1, a Region Frequently Affected in Breast Cancer. *Cancer Res* 60: 2140–2145.
  84. Nunez MI, Ludes-Meyers J, Abba MC, Kil H, Abbey NW, et al. (2005) Frequent loss of WWOX expression in breast cancer: correlation with estrogen receptor status. *Breast Cancer Research and Treatment* 89: 99–105.
  85. Schild L, Lu Y, Gautschi I, Schneeberger E, Lifton RP, et al. (1996) Identification of a PY motif in the epithelial Na channel subunits as a target sequence for mutations causing channel activation found in Liddle syndrome. *The EMBO journal* 15: 2381–2387.
  86. Kanelis V, Rotin D, Forman-Kay JD (2001) Solution structure of a Nedd4 WW domain – ENaC peptide complex. *Nat Struct Biol* 8: 407–412.
  87. Holbert S, Denghien I, Kiechle T, Rosenblatt A, Wellington C, et al. (2001) The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: Neuropathologic and genetic evidence for a role in Huntington’s disease pathogenesis. *PNAS* 98: 1811–1816.
  88. Lubs H, Abidi FE, Echeverri R, Holloway L, Meindl A, et al. (2006) Golabi-Ito-Hall syndrome results from a missense mutation in the WW domain of the PQBP1 gene. *Journal of medical genetics* 43: e30.
-



- 
89. Russo T, Faronio R, Minopoli G, De Candia P, De Renzis S, et al. (1998) Fe65 and the protein network centered around the cytosolic domain of the Alzheimer's  $\beta$ -amyloid precursor protein. *FEBS Letters* 434: 1–7.
  90. Suné C, Garcia-Blanco MA (1999) Transcriptional cofactor CA150 regulates RNA polymerase II elongation in a TATA-box-dependent manner. *Mol Cell Biol* 19: 4719–4728.
  91. Sánchez-Alvarez M, Goldstrohm AC, Garcia-Blanco M a, Suñé C (2006) Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through its amino and carboxyl regions. *Mol Cell Biol* 26: 4998–5014.
  92. Suné C, Hayashi T, Liu Y, Lane WS, Young RA, et al. (1997) CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. *Mol Cell Biol* 17: 6029–6039.
  93. Bedford MT, Leder P (1999) The FF domain: a novel motif that often accompanies WW domains. *Trends in biochemical sciences* 24: 264–265.
  94. Bohne J, Cole SE, Sune C, Lindman BR, Ko VD, et al. (2000) Expression analysis and mapping of the mouse and human transcriptional regulator CA150. *Mammalian Genome* 11: 930–933.
  95. Murphy JM, Hansen DF, Wiesner S, Muhandiram DR, Borg M, et al. (2009) Structural Studies of FF Domains of the Ttranscription Factor CA150 Provide Insights into the Organization of FF Domain Tandem Arrays. *J Mol Biol* 393: 409–424.
  96. Ramirez-Espain X, Ruiz L, Martin-Malpartida P, Oschkinat H, Macias MJ (2007) Structural Characterization of a New Binding Motif and a Novel Binding Mode in Group 2 WW Domains. *J Mol Biol* 373: 1255–1268.
  97. Przedziak J, Tremmel S, Kretzschmar I, Beyermann M, Bienert M, et al. (2006) Probing the Ligand-Binding Specificity and Analyzing the Folding State of SPOT-Synthesized FBP28 WW Domain Variants. *ChemBioChem* 7: 780–788.
  98. Hu H, Columbus J, Zhang Y, Wu D, Lian L, et al. (2004) A map of WW domain family interactions. *Proteomics* 4: 643–655.
  99. Macias MJ, Gervais V, Civera C, Oschkinat H (2000) Structural analysis of WW domains and design of a WW prototype. *Nat Struct Biol* 7: 375–379.
  100. Ferguson N, Johnson CM, Macias M, Oschkinat H, Fersht A (2001) Ultrafast folding of WW domains without structured aromatic clusters in the denatured state. *Proc Natl Acad Sci USA* 98: 13002–13007.
  101. Petrovich M, Jonsson AL, Ferguson N, Daggett V, Fersht AR (2006) Phi-Analysis at the Experimental Limits: Mechanism of beta-Hairpin Formation. *J Mol Biol* 360: 865–881.
  102. Ferguson N, Becker J, Tidow H, Tremmel S, Sharpe TD, et al. (2006) General structural motifs of amyloid protofilaments. *Proc Natl Acad Sci USA* 103: 16248–16253.
  103. Jahn TR, Radford SE (2005) The Yin and Yang of protein folding. *FEBS J* 272: 5962–5970.
  104. Selkoe DJ (2003) Folding proteins in fatal ways. *Nature* 426: 900–904.
  105. Shimada M, Saito M, Katakai T, Shimizu A, Ichimura T, et al. (1999) Molecular cloning and splicing isoforms of mouse p144, a homologue of CA150. *Journal of biochemistry* 126: 1033–1042.
-

- 
106. DeMarco R, Oliveira KC, Venancio TM, Verjovski-Almeida S (2006) Gender biased differential alternative splicing patterns of the transcriptional cofactor CA150 gene in *Schistosoma mansoni*. *Molecular and biochemical parasitology* 150: 123–131.
  107. Yu L-R, Zhu Z, Chan KC, Issaq HJ, Dimitrov DS, et al. (2007) Improved Titanium Dioxide Enrichment of Phosphopeptides from HeLa Cells and High Confident Phosphopeptide Identification by Cross-Validation of MS/MS and MS/MS/MS Spectra. *Journal of Proteome Research* 6: 4150–4162.
  108. Arango M, Holbert S, Zala D, Brouillet E, Pearson J, et al. (2006) CA150 Expression Delays Striatal Cell Death in Overexpression and Knock-In Conditions for Mutant Huntingtin Neurotoxicity. *The Journal of neuroscience* 26: 4649–4659.
  109. Chattopadhyay B, Ghosh S, Gangopadhyay PK, Das SK, Roy T, et al. (2003) Modulation of age at onset in Huntington's disease and spinocerebellar ataxia type 2 patients originated from eastern India. *Neurosci Lett* 345: 93–96.
  110. Neubauer G, King A, Rappsilber J, Calvio C, Watson M, et al. (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nature genetics* 20: 46–50.
  111. Zhou Z, Licklider LJ, Gygi SP, Reed R (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature* 419: 182–185.
  112. Goldstrohm AC, Albrecht TR, Suné C, Bedford MT, Garcia-Blanco MA (2001) The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. *Mol Cell Biol* 21: 7617–7628.
  113. Selenko P, Gregorovic G, Sprangers R, Stier G, Rhani Z, et al. (2003) Structural Basis for the Molecular Recognition between Human Splicing Factors U2AF65 and SF1/mBBP. *Molecular cell* 11: 965–976.
  114. Goldstrohm AC, Greenleaf AL, Garcia-Blanco M a (2001) Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing. *Gene* 277: 31–47.
  115. Lin K-T, Lu R-M, Tarn W-Y (2004) The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo. *Mol Cell Biol* 24: 9176–9185.
  116. Carty SM, Goldstrohm AC, Suñé C, Garcia-Blanco MA, Greenleaf AL (2000) Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II. *Proc Natl Acad Sci U S A* 97: 9015–9020.
  117. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, et al. (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic acids research* 39: D561–8.
  118. Shimkets RA, Warnock DG, Bositis CM, Nelson-williams C, Hansson JH, et al. (1994) Liddle's Syndrome: Heritable Human Hypertension Caused by Mutations in the  $\beta$  Subunit of the Epithelial Sodium Channel. *Cell* 79: 407–414.
  119. Passani LA, Bedford MT, Faber PW, McGinnis KM, Sharp AH, et al. (2000) Huntingtin's WW domain partners in Huntington's disease post-mortem brain fulfill genetic criteria for direct involvement in Huntington's disease pathogenesis. *Human molecular genetics* 9: 2175–2182.
  120. Wiesner S, Stier G, Sattler M, Macias MJ (2002) Solution structure and ligand recognition of the WW domain pair of the yeast splicing factor Prp40. *J Mol Biol* 324: 807–822.
-

- 
121. Fedoroff OY, Townson SA, Golovanov AP, Baron M, Avis JM (2004) The Structure and Dynamics of Tandem WW Domains in a Negative Regulator of Notch Signaling , Suppressor of Deltex \*. 279: 34991–35000.
  122. Jennings MD, Blankley RT, Baron M, Golovanov AP, Avis JM (2007) Specificity and autoregulation of Notch binding by tandem WW domains in suppressor of Deltex. *J Biol Chem* 282: 29032–29042.
  123. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, et al. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins-Structure Function and Bioinformatics* 59: 687–696.
  124. Palmer AG (1998) Program CurveFit: General fitting program for batch use.
  125. Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 89: 392–400.
  126. Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* 32: W668–W673.
  127. Lapid C, Gao YG (2003) PrimerX - Automated design of mutagenic primers for site-directed mutagenesis.
  128. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, et al. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31: 3784–3788.
  129. Delano WL (2009) PyMOL molecular viewer: Updates and refinements. *Abstracts of Papers of the American Chemical Society* 238.
  130. Delano WL, Lam JW (2005) PyMOL: A communications tool for computational models. *Abstracts of Papers of the American Chemical Society* 230: U1371–U1372.
  131. Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. *Biophysical Journal* 82: 1096–1111.
  132. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal* 78: 1606–1619.
  133. Hayes D, Laue T, Philo J (1995) Sednterp: Sedimentation Interpretation Program.
  134. Balbo A, Minor KH, Velikovskiy CA, Mariuzza RA, Peterson CB, et al. (2005) Studying multiprotein complexes by multisignal sedimentation velocity analytical ultracentrifugation. *Proc Natl Acad Sci USA* 102: 81–86.
  135. Houtman JCD, Brown PH, Bowden B, Yamaguchi H, Appella E, et al. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling. *Protein Science* 16: 30–42.
  136. Schuck P (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Analyt Biochem* 320: 104–124.
  137. Vistica J, Dam J, Balbo A, Yikilmaz E, Mariuzza RA, et al. (2004) Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Analyt Biochem* 326: 234–256.
-

- 
138. Schuck P (2005) Diffusion-deconvoluted sedimentation coefficient distributions for the analysis of interacting and non-interacting protein mixtures. In: Scott DJ, Harding SE, Rowe AJ, editors. *Modern Analytical Ultracentrifugation: Techniques and Methods*. The Royal Society of Chemistry, Cambridge, UK.
  139. Dam J, Velikovskiy CA, Mariuzza RA, Urbanke C, Schuck P (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modeling and sedimentation coefficient distributions  $c(s)$ . *Biophysical Journal* 89: 619–634.
  140. Dam J, Schuck P (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: Sedimentation coefficient distributions  $c(s)$  and asymptotic boundary profiles from Gilbert-Jenkins theory. *Biophysical Journal* 89: 651–666.
  141. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular Nmr* 44: 213–223.
  142. Dosset P, Hus JC, Marion D, Blackledge M (2001) A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *Journal of Biomolecular Nmr* 20: 223–231.
  143. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25: 1605–1612.
  144. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor laboratory press. Cold Spring Harbor. p.
  145. Weiner MP, Costa GL, Schoettlin W, Cline J, Mathur E, et al. (1994) Site-directed mutagenesis of double-stranded chain reaction DNA by the polymerase. *Gene* 151: 119–123.
  146. Studier FW, Moffatt BA (1986) Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes. *J Mol Biol* 189: 113–130.
  147. Sambrook J, Russell DW (2001) *Molecular cloning-A laboratory manual*. Vol 1. 2nd ed. Cold Spring Harbor Laboratory Press, U.S. p.
  148. Laemmli UK (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227: 680–685.
  149. Schagger H, Jagow G (1987) Tricine-sodium dodecylsulfate-polyacrylamide gelelectrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analyt Bioch* 166: 368–379.
  150. Frank R (1992) SPOT-Synthesis: An Easy Technique for the Positionally Addressable, Parallel Chemical Synthesis on a Membrane Support. *Tetrahedron* 48: 9217–9232.
  151. Piotto M, Saudek V, Sklenar V (1992) Gradient-Tailored Excitation for Single-Quantum Nmr-Spectroscopy of Aqueous-Solutions. *Journal of Biomolecular Nmr* 2: 661–665.
  152. Bodenhausen G, Ruben DJ (1980) Natural Abundance N-15 Nmr by Enhanced Heteronuclear Spectroscopy. *Chemical Physics Letters* 69: 185–189.
  153. Ikura M, Kay LE, Bax A (1990) A Novel-Approach for Sequential Assignment of H-1, C-13, and N-15 Spectra of Larger Proteins - Heteronuclear Triple-Resonance 3-Dimensional Nmr-Spectroscopy - Application to Calmodulin. *Biochemistry* 29: 4659–4667.
  154. Grzesiek S, Bax A (1992) Improved 3D Triple-Resonance Nmr Techniques Applied to A 31-Kda Protein. *Journal of Magnetic Resonance* 96: 432–440.
-

- 
155. Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE (1994) A Suite of Triple-Resonance Nmr Experiments for the Backbone Assignment of N-15, C-13, H-2 Labeled Proteins with High-Sensitivity. *Journal of the American Chemical Society* 116: 11655–11666.
  156. Morin S (2011) A practical guide to protein dynamics from N-15 spin relaxation in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy* 59: 245–262.
  157. Farina B, Pirone L, Russo L, Viparelli F, Doti N, et al. (2010) NMR backbone dynamics studies of human PED/PEA-15 outline protein functional sites. *Febs Journal* 277: 4229–4240.
  158. Dosset P, Hus JC, Blackledge M, Marion D (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *Journal of Biomolecular Nmr* 16: 23–28.
  159. d’Auvergne EJ, Gooley PR (2008) Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. *Journal of Biomolecular Nmr* 40: 121–133.
  160. d’Auvergne EJ, Gooley PR (2007) Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm. *Molecular Biosystems* 3: 483–494.
  161. d’Auvergne EJ, Gooley PR (2006) Model-free model elimination: A new step in the model-free dynamic analysis of NMR relaxation data. *Journal of Biomolecular Nmr* 35: 117–135.
  162. d’Auvergne EJ, Gooley PR (2008) Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *Journal of Biomolecular Nmr* 40: 107–119.
  163. d’Auvergne EJ, Gooley PR (2003) The use of model selection in the model-free analysis of protein dynamics. *Journal of Biomolecular Nmr* 25: 25–39.
  164. Janes RW (2005) Bioinformatics analyses of circular dichroism protein reference databases. *Bioinformatics (Oxford, England)* 21: 4230–4238.
  165. Ong S-E (2002) Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol Cell Proteomics* 1: 376–386.
  166. Zhu H, Pan S, Gu S, Bradbury EM, Chen X (2002) Amino acid residue specific stable isotope labeling for quantitative proteomics. *Rapid Commun Mass Spectrom* 16: 211521–211523.
  167. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 26: 1367–1372.
  168. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analyt Biochem* 287: 252–260.
  169. Mao D, Wachter E, Wallace BA (1982) Folding of the Mitochondrial Proton Adenosine-Triphosphatase Proteolipid Channel in Phospholipid-Vesicles. *Biochemistry* 21: 4960–4968.
  170. Hilger M, Mann M (2012) Triple SILAC to determine stimulus specific interactions in the Wnt pathway. *J Proteome Res* 11: 982–994.
  171. Parthier C (2005) Entwicklung eines Protein-Kopplungsverfahrens basierend auf WW-Domänen Martin-Luther-Universität Halle-Wittenberg.
  172. Ingham RJ, Colwill K, Howard C, Dettwiler S, Lim CSH, et al. (2005) WW Domains Provide a Platform for the Assembly of Multiprotein Networks. *Mol Cell Biol* 25: 7092–7106.
-

- 
173. Behzadnia N, Golas MM, Hartmuth K, Sander B, Kastner B, et al. (2007) Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *The EMBO journal* 26: 1737–1748.
  
  174. Śledź P, Stubbs CJ, Lang S, Yang Y-Q, McKenzie GJ, et al. (2011) From Crystal Packing to Molecular Recognition: Prediction and Discovery of a Binding Site on the Surface of Polo-Like Kinase 1. *Angewandte Chemie (International ed in English)* 50: 4003–4006.
-

---

## **EIDESSTATTLICHE ERKLÄRUNG**

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Berlin, .....

.....

(Anne Wartenberg)

---





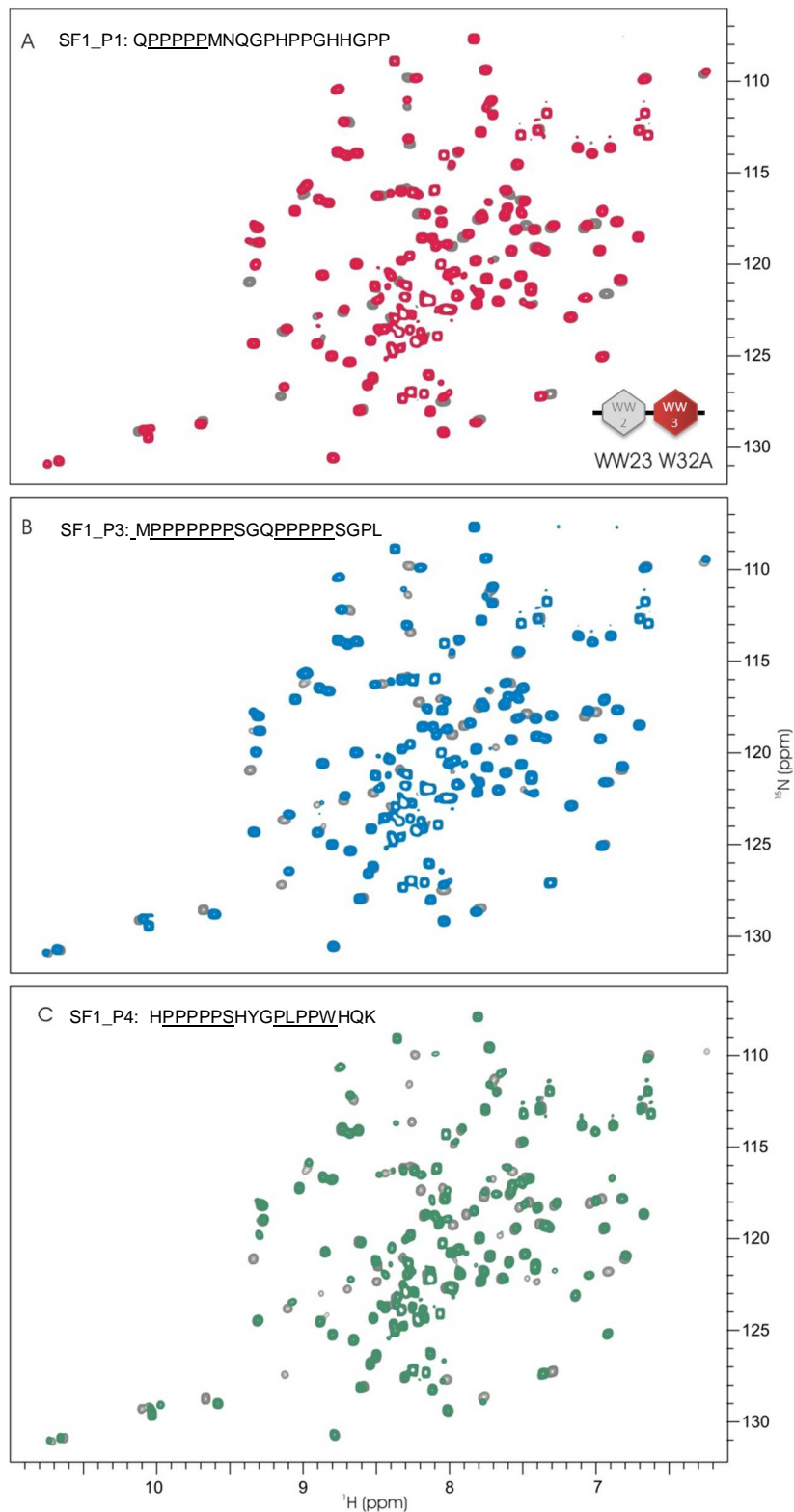
## 6.2 NMR-DATEN

**Tabelle 6.2.I: Zugeordnete  $^1\text{H}$ -,  $^{15}\text{N}$ - und  $^{13}\text{C}$  chemische Verschiebungen des WW23-Tandems**

Rest	$\omega \text{ } ^1\text{H}_\text{N}$ (ppm)	$\omega \text{ } ^{15}\text{N}$ (ppm)	$\omega \text{ N}_\epsilon$ (ppm)	$\omega \text{ H}_\epsilon$ (ppm)	$\omega \text{ C}_\alpha$ (ppm)	$\omega \text{ C}_\beta$ (ppm)	$\omega \text{ C}_\text{O}$ (ppm)
2 Ala	8,55	123,61	-	-	52,47	19,1	178,13
3 Gly	8,43	108,88	-	-	44,8	-	173,71
4 Ala	8,13	123,93	-	-	52,36	19,16	177,97
5 Thr	8,09	114,01	-	-	61,55	69,67	174,07
6 Ala	8,22	127,04	-	-	52,22	19,15	177,43
7 Val	8,1	119,91	-	-	61,99	32,6	176,05
8 Ser	8,3	119,5	-	-	57,99	63,8	174,91
9 Glu	8,81	124,98	-	-	57,45	29,6	175,4
10 Trp	8,11	118,32	129,42	10,1	55,99	32,01	176,53
11 Thr	9,25	118,67	-	-	61,68	71,76	171,97
12 Glu	8,38	126,36	-	-	55,19	-	174,05
13 Tyr	8,41	126	-	-	55,83	42,59	173,34
14 Lys	8,02	116,74	-	-	53,88	35,72	178,52
15 Thr	9,11	114,03	-	-	60,84	71,22	177,04
16 Ala	9,1	124,46	-	-	54,77	18,03	178,24
17 Asp	7,88	112,04	-	-	52,76	39,58	176,23
18 Gly	7,82	107,56	-	-	45,83	-	174,2
19 Lys	7,75	121,61	-	-	55,75	33,37	175,85
20 Thr	8,59	120,72	-	-	62,97	70,14	173,19
21 Tyr	8,57	123,07	-	-	55,25	39,5	170,66
22 Tyr	8,7	117	-	-	56,34	40,55	174,15
23 Tyr	9,2	124,35	-	-	56,13	43,23	173,4
24 Asn	8,06	129,66	-	-	51,19	38,53	174,56
25 Asn	8,35	121,75	-	-	-	32,73	176,37
26 Arg	8,39	119,73	-	-	57,45	30,51	177,34
27 Thr	7,78	109,34	-	-	61,51	69,28	175,81
28 Leu	7,64	116,9	-	-	-	37,15	175,67
29 Glu	6,97	119,16	-	-	58,04	31,67	174,55
30 Ser	8,3	116,1	-	-	58,03	63,65	174,29
31 Thr	9,22	117,91	-	-	59,09	69,73	171,95
32 Trp	8,29	127,71	129,05	9,94	55,84	32,01	174,01
33 Glu	8,21	123,85	-	-	54,19	29,3	174,1
35 Pro	-	-	-	-	57,49	32,58	177,52
36 Gln	7,88	119,87	-	-	57,44	29,64	177,4
37 Glu	9,11	117,09	-	-	58,64	29,57	177,99
38 Leu	6,92	117,68	-	-	55,65	41,2	177,98
39 Lys	7,51	121,27	-	-	58,42	32,14	178,33
41 Lys	7,57	120,66	-	-	-	32,62	174,94
42 Glu	8,15	119,03	-	-	57,3	29,42	177,63
43 Lys	7,82	120,81	-	-	-	32,37	177,27
44 Leu	7,88	122,18	-	-	55,58	41,89	177,81

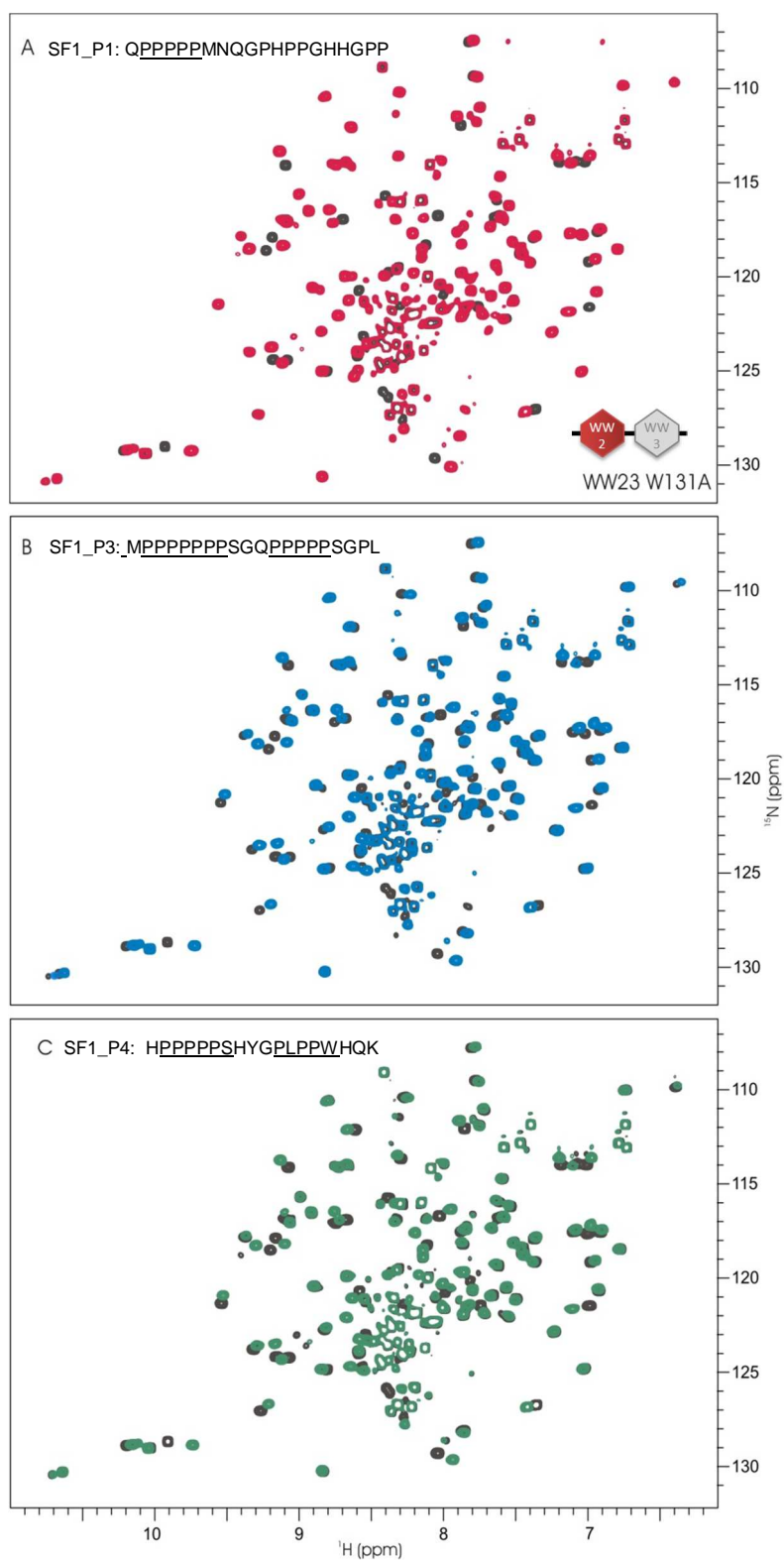
Rest	$\omega$ $^1\text{H}_\text{N}$ (ppm)	$\omega$ $^{15}\text{N}$ (ppm)	$\omega$ $\text{N}_\epsilon$ (ppm)	$\omega$ $\text{H}_\epsilon$ (ppm)	$\omega$ $\text{C}_\alpha$ (ppm)	$\omega$ $\text{C}_\beta$ (ppm)	$\omega$ $\text{C}_\text{o}$ (ppm)
45 Glu	8,04	120,44	-	-	56,55	29,98	176,56
47 Lys	8,47	121,26	-	-	56,12	32,93	176,52
48 Ile	8,01	122,43	-	-	-	38,06	176,02
49 Lys	8,12	126,46	-	-	55,98	32,82	174,89
50 Glu	8,04	122,99	-	-	53,1	31,4	174,2
53 Lys	8,55	120,22	-	-	55,76	32,9	175,93
54 Glu	8,43	124,49	-	-	54,06	29,31	174,19
56 Ser	8,37	109,82	-	-	47,18	56,71	174,78
67 Glu	8,3	122,74	-	-	55,92	30,3	175,99
68 Asp	8,43	123,43	-	-	52,32	40,23	174,07
76 Glu	8,4	123,72	-	-	56,02	30,28	176,09
77 Ile	8,27	124,18	-	-	60,62	38,24	175,25
84 Glu	8,91	120,66	-	-	59,45	28,67	179,47
85 Glu	8,69	120	-	-	59,72	28,81	179,43
86 Met	7,73	121,97	-	-	58,71	29,36	179,57
87 Thr	8,76	112,19	-	-	60,4	70,71	176,93
92 Ala	7,86	121,6	-	-	54,51	17,65	180,16
93 Ala	7,66	121,04	-	-	54,25	17,77	180,03
95 Lys	7,64	116,06	-	-	51,07	38,47	174,35
96 Ala	8,3	121,54	-	-	47,95	15,39	176,29
98 Pro	-	-	-	-	61,15	38,11	-
100 Ala	7,51	122,18	-	-	52,03	21,5	175
101 Thr	8,32	113,32	-	-	60,09	71,95	173,47
102 Ala	9,2	126,96	-	-	49,29	21,56	176,58
106 Gly	8,81	110,43	-	-	45,1	-	173,46
107 Thr	7,77	111,82	-	-	58,88	70,43	173,2
108 Pro	-	-	-	-	62,33	30,98	175,25
109 Trp	7,84	117,26	130,9	10,8	57,3	33,63	176,14
110 Cys	9,38	117,85	-	-	56,49	31,19	173,26
111 Val	8,78	122,6	-	-	61,42	31,39	173,81
112 Val	8,11	127,4	-	-	60,51	32,61	181,98
113 Trp	8,35	127,23	129,1	10,14	55,71	32,99	176,3
114 Thr	8,68	114,04	-	-	60,9	70,87	175,39
115 Gly	8,34	111,23	-	-	-	-	177,93
116 Asp	7,87	117,55	-	-	52,03	38,86	174,57
117 Glu	7,79	111,24	-	-	57,41	25,86	175,57
118 Arg	7,61	118,23	-	-	55,14	33,16	175,01
119 Val	8,04	118,99	-	-	59,41	34,94	174,46
120 Phe	8,4	120,93	-	-	54,85	38,73	-
121 Phe	8,53	116,25	-	-	56,3	41,16	174,49
122 Tyr	9,19	123,67	-	-	55,53	44,3	172,71
123 Asn	7,85	128,5	-	-	47,85	37,82	173,31
125 Thr	8	113,86	-	-	65,76	67,8	175,56
126 Thr	6,72	109,82	-	-	61,15	69,13	174,83
127 Arg	8,12	117,14	-	-	57,33	26,1	174,61

Rest	$\omega$ $^1\text{H}_\text{N}$ (ppm)	$\omega$ $^{15}\text{N}$ (ppm)	$\omega$ $\text{N}_\epsilon$ (ppm)	$\omega$ $\text{H}_\epsilon$ (ppm)	$\omega$ $\text{C}_\alpha$ (ppm)	$\omega$ $\text{C}_\beta$ (ppm)	$\omega$ $\text{C}_\text{O}$ (ppm)
128 Leu	6,89	120,92	-	-	54,5	43,87	175,03
129 Ser	8,27	117,33	-	-	56,62	65,58	173,7
130 Met	9,42	120,88	-	-	55,04	36,83	174,44
131 Trp	8,58	122,19	128,68	9,74	58,92	-	176,71
132 Asp	7,55	117,83	-	-	52,61	41,79	173,51
133 Arg	7,95	118,52	-	-	53,46	29,73	174,19
135 Asp	8,3	122,29	-	-	55,82	40,01	176,01
136 Asp	8,76	114,08	-	-	55,23	39,43	176,2
137 Leu	7,6	114,53	-	-	53,01	41,72	177,32
138 Ile	7,02	125,07	-	-	63,37	37,29	176,3
139 Gly	9,02	115,65	-	-	45,14	-	171,45
140 Arg	6,78	118,54	-	-	54,93	29,79	176,21
141 Ala	8,84	130,6	-	-	54,2	18,31	179,33
142 Asp	8,94	116,45	-	-	55,67	37,74	178,57
143 Val	7,23	122,86	-	-	65,97	31,64	176,34
144 Asp	7,47	118,11	-	-	57,51	39,68	178,03
145 Lys	7,36	117,99	-	-	59	32,43	-
146 Ile	7,42	119,31	-	-	64,67	38,65	175,69
147 Ile	-	-	-	-	55,64	33,01	176,72
148 Gln	8,48	123,44	-	-	56,76	29,99	176,66
149 Glu	8,52	121,84	-	-	56,31	30,07	176,29
152 His	8,18	121,9	-	-	61,02	38,26	176,24
153 Lys	8,17	126,07	-	-	55,86	32,89	174,95
154 Lys	7,88	121,3	-	-	54,38	30,38	174,37



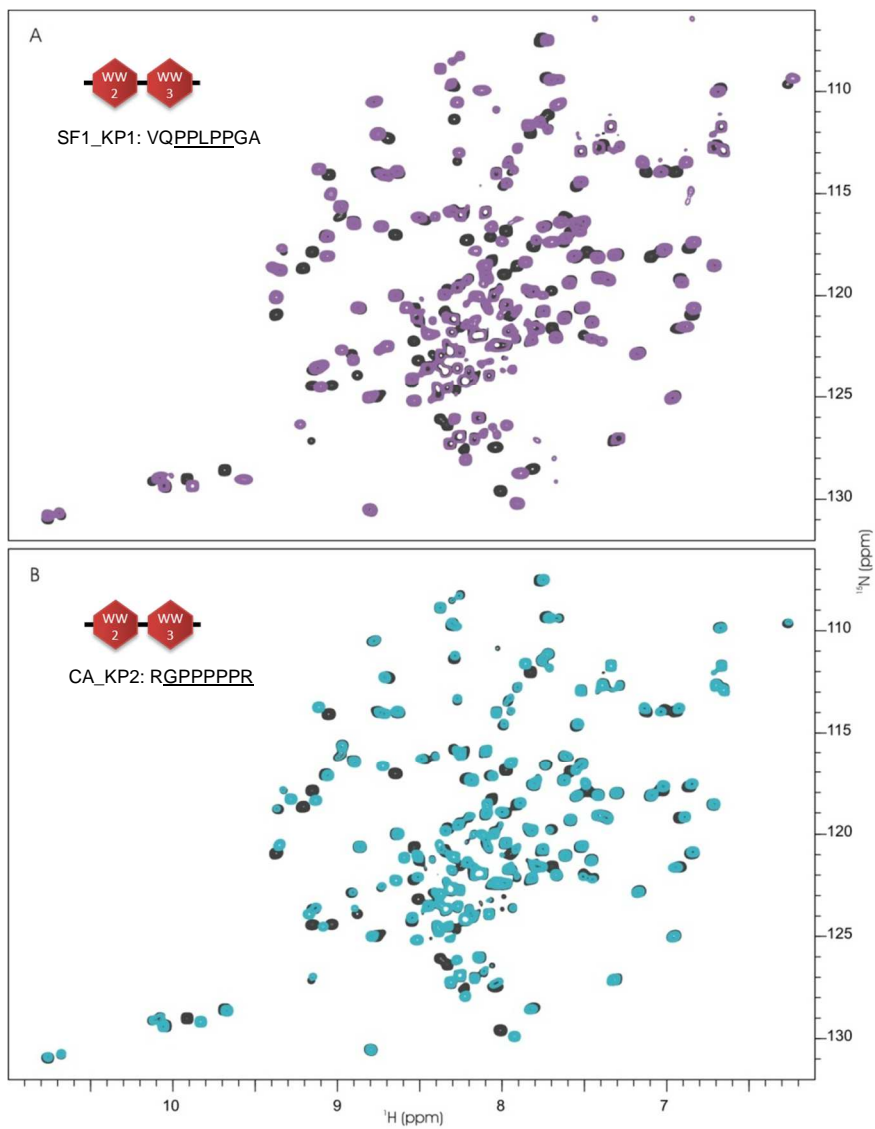
**Abbildung 52:  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von WW23\_W32A in An- und Abwesenheit verschiedener SF1 Peptide**

Überlagerung von  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren aus NMR-Titrationsexperimenten. WW23\_W32A ( $c = 50 \mu\text{M}$ ) in Abwesenheit vom Peptidliganden (grau), WW23\_W32A in Anwesenheit eines 10-fachen Überschusses von SF1\_P1 (rot), SF1\_P3 (blau), SF1\_P4 (grün). Aufgenommen an einem 600 MHz Spektrometer bei 280K.



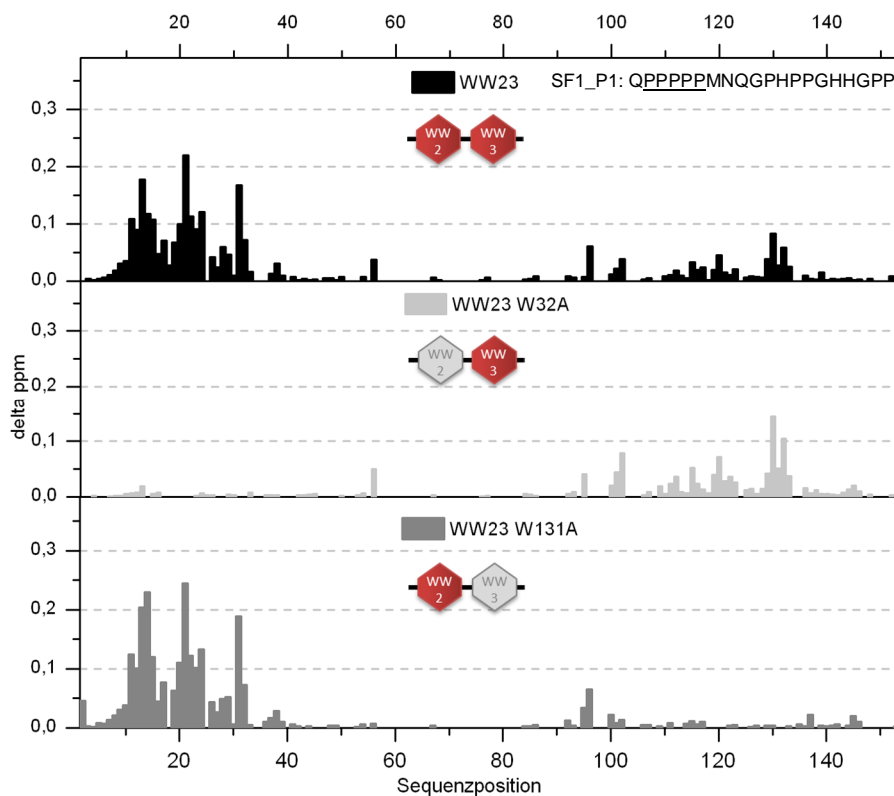
**Abbildung 53:**  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von WW23\_W131A in An- und Abwesenheit verschiedener SF1 Peptide

Überlagerung von  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren aus NMR-Titrationsexperimenten. WW23\_W131A ( $c = 50 \mu\text{M}$ ) in Abwesenheit vom Peptidliganden (schwarz), WW23\_W131A in Anwesenheit eines 10-fachen Überschusses von SF1\_P1 (rot), SF1\_P3 (blau), SF1\_P4 (grün). Aufgenommen an einem 600 MHz Spektrometer bei 280K.



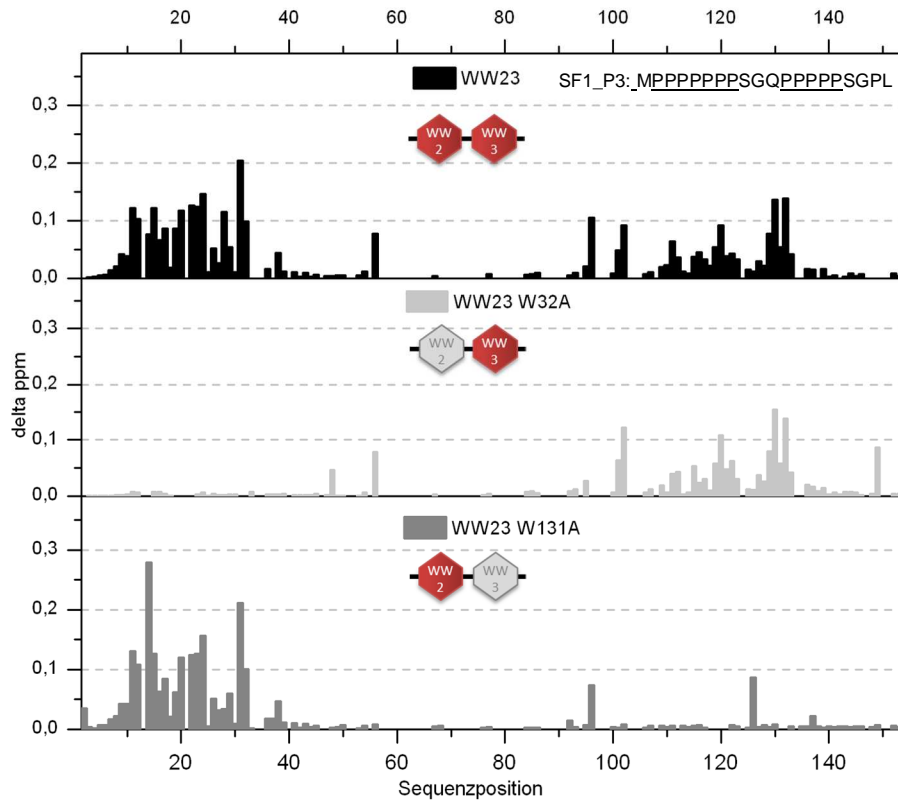
**Abbildung 54:**  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von WW23 in An- und Abwesenheit verschiedener monovalenter PP-Peptide

Überlagerung von  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren aus NMR-Titrationsexperimenten. WW23 ( $c = 50 \mu\text{M}$ ) in Abwesenheit vom Peptidliganden (schwarz), WW23 in Anwesenheit eines 10-fachen Überschusses von SF1\_KP1 (violett), CA\_KP2 (türkis). Aufgenommen an einem 600 MHz Spektrometer bei 280K.



**Abbildung 55: Analyse der Änderung der  $^{15}\text{N}$  chemischen Verschiebungen von WW23, WW23\_W32A und WW23\_W131A nach SF1\_P1-Zugabe**

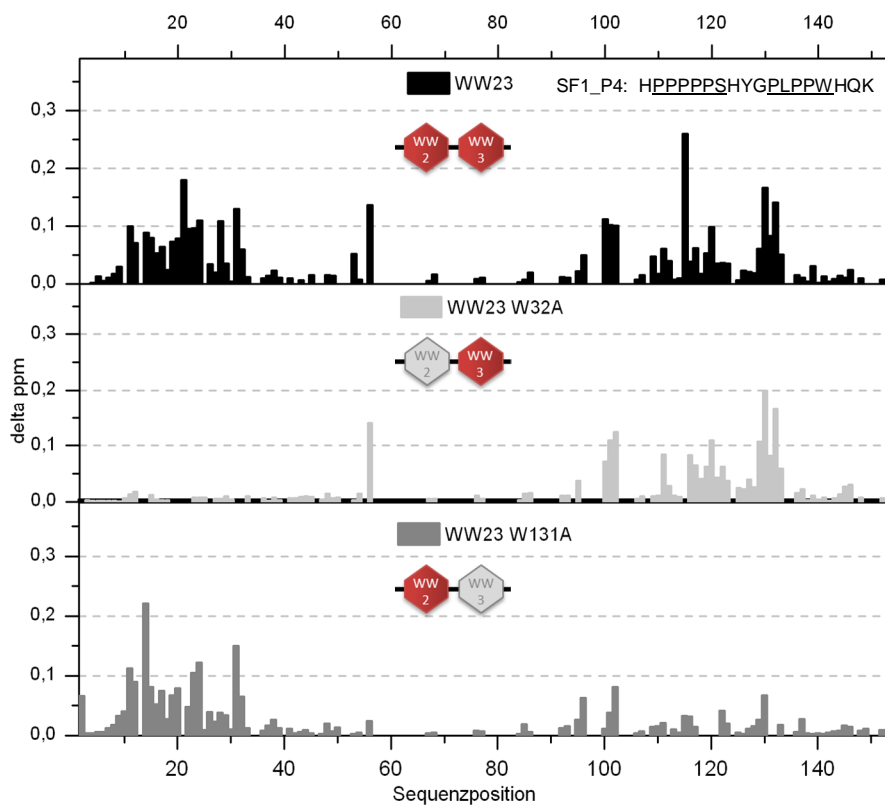
Die peptidinduzierten Differenzen der  $^{15}\text{N}$  chemischen Verschiebungen ( $\Delta\sigma^{15}\text{N}$ ) von WW23 (schwarz), WW23\_W32A (hellgrau) bzw. WW23\_W131A (dunkelgrau) Kreuzsignalen wurden gegen die entsprechende Sequenzposition aufgetragen.



**Abbildung 56: Analyse der Änderung der  $^{15}\text{N}$  chemischen Verschiebungen von WW23, WW23\_W32A und WW23\_W131A nach SF1\_P3**

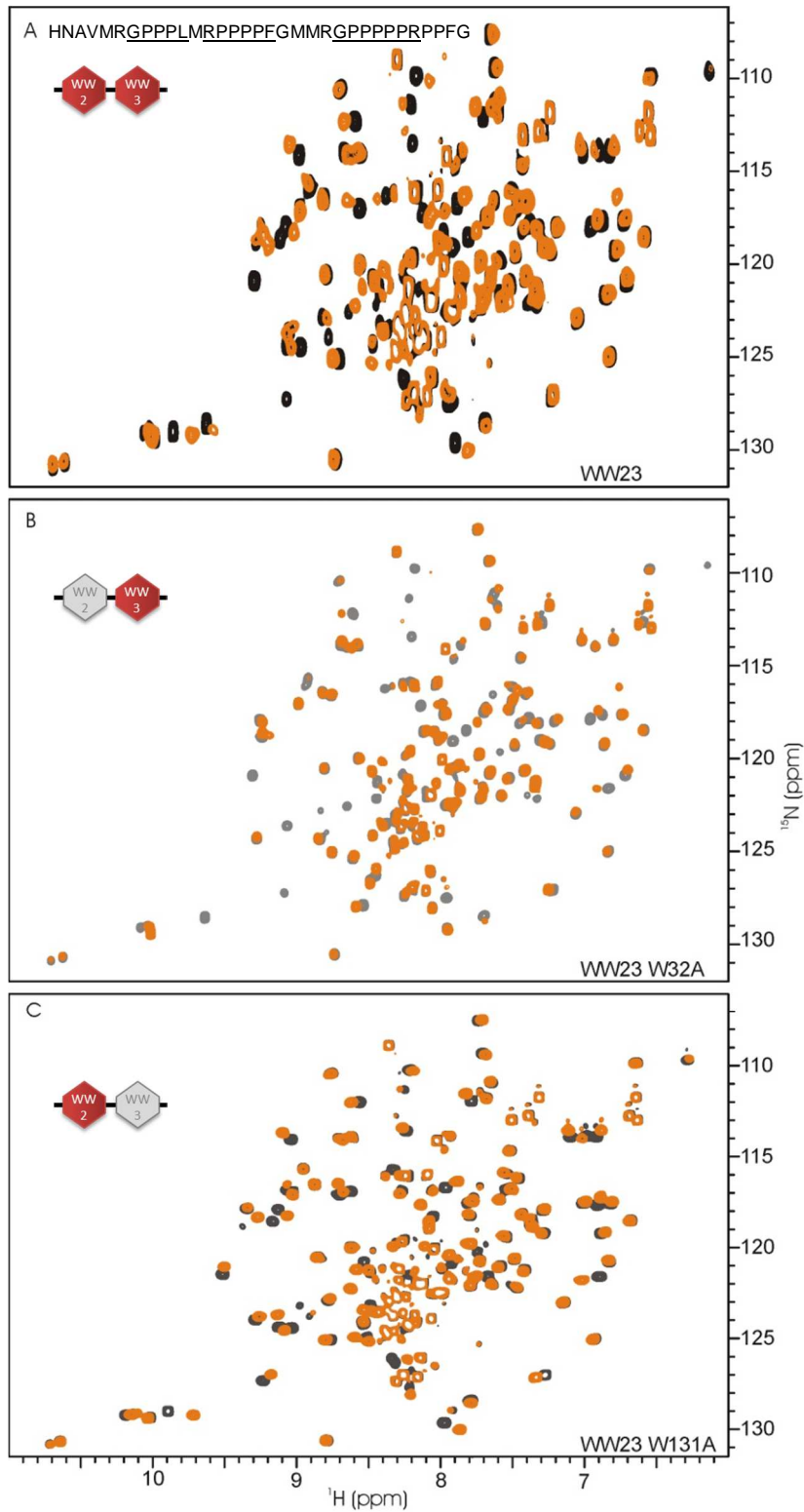
Die peptidinduzierten Differenzen der  $^{15}\text{N}$  chemischen Verschiebungen ( $\Delta\sigma^{15}\text{N}$ ) von WW23 (schwarz), WW23\_W32A (hellgrau) bzw. WW23\_W131A (dunkelgrau) Kreuzsignalen wurden gegen die entsprechende Sequenzposition aufgetragen





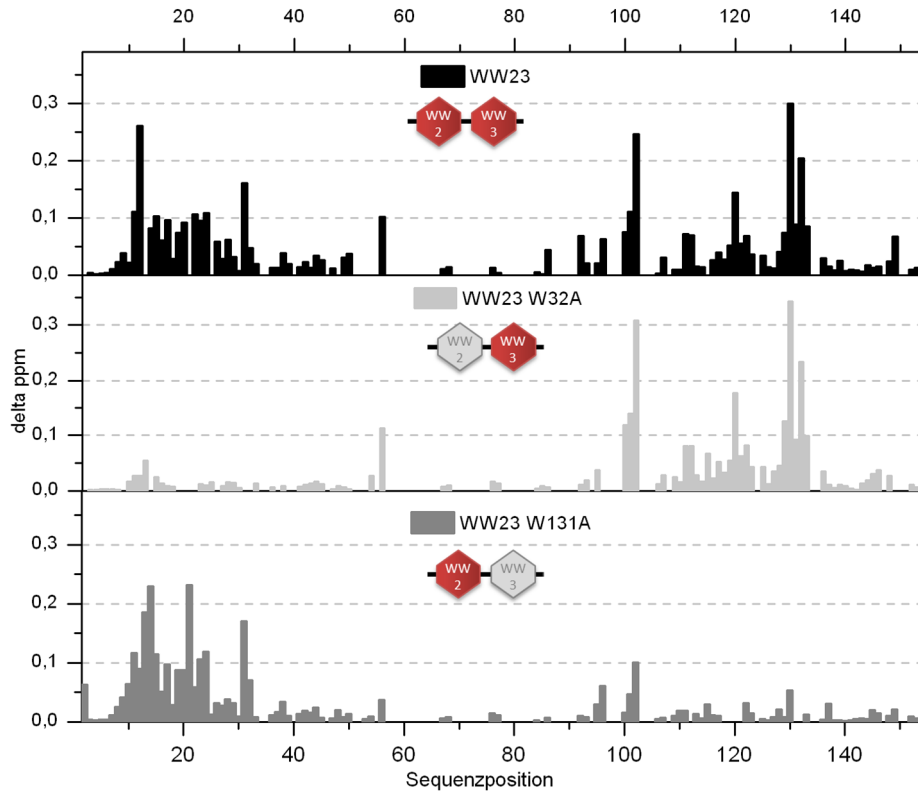
**Abbildung 57: Analyse der Änderung der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen von WW23, WW23\_W32A und WW23\_W131A nach SF1\_P4-Zugabe**

Die peptidinduzierten Differenzen der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen ( $\Delta\sigma^{1\text{H}^{15}\text{N}}$ ) von WW23 (schwarz), WW23\_W32A (hellgrau) bzw. WW23\_W131A (dunkelgrau) Kreuzsignalen wurden gegen die entsprechende Sequenzposition aufgetragen.



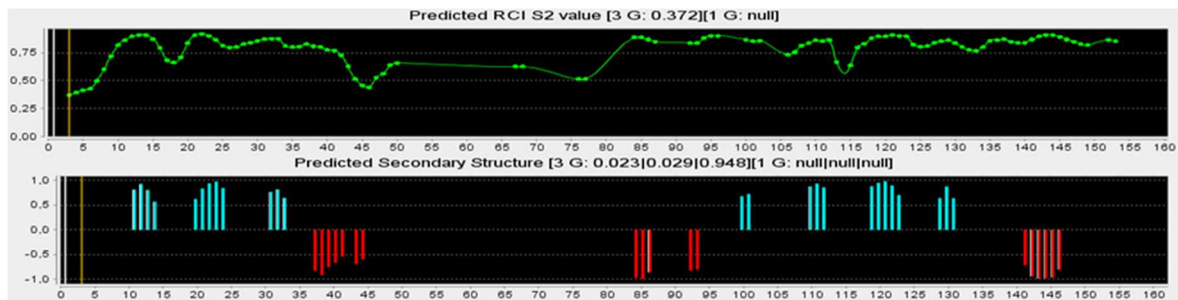
**Abbildung 58:**  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren von WW23, WW23\_W32A und WW23\_W131A in An- und Abwesenheit verschiedener SF1 Peptide

Überlagerung von  $^1\text{H}^{15}\text{N}$ -HSQC-Spektren aus NMR-Titrationsexperimenten. WW23 (schwarz), WW23\_W32A (hellgrau) und WW23\_W131A (dunkelgrau) ( $c = 50 \mu\text{M}$ ) in Abwesenheit vom Peptidliganden, WW23-Tandem-Variante in Anwesenheit eines 10-fachen Überschusses von CA\_P1 (orange). Aufgenommen an einem 600 MHz Spektrometer bei 280K.



**Abbildung 59: Analyse der Änderung der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen von WW23, WW23\_W32A und WW23\_W131A nach CA\_P1-Zugabe**

Die peptidinduzierten Differenzen der  $^1\text{H}^{15}\text{N}$  chemischen Verschiebungen ( $\Delta\sigma^{1\text{H}^{15}\text{N}}$ ) von WW23 (schwarz), WW23\_W32A (hellgrau) bzw. WW23\_W131A (dunkelgrau) Kreuzsignalen wurden gegen die entsprechende Sequenzposition aufgetragen. CA\_P1 : HNAVMRGPPPPLMRPPPPFEGMMRGGPPPPRPPFG



**Abbildung 60: Ergebnis der Talos+ Strukturdynamik- und Sekundärstrukturvorhersage**

Oben ist die strukturelle Dynamik des Proteinrückgrates (Orderparameter) und darunter die Sekundärstrukturstrukturelemente (blau:  $\alpha$ -Helix; rot:  $\beta$ -Strang) des WW23-Tandems aus CA150 abgebildet. Die Werte sind gegen die Sequenzposition aufgetragen. Die Vorhersage wurde mit Talos+ erstellt und basiert auf den sequentiell zugeordneten  $^{15}\text{N}$ - und  $^{13}\text{C}$  chemischen Verschiebungen.