

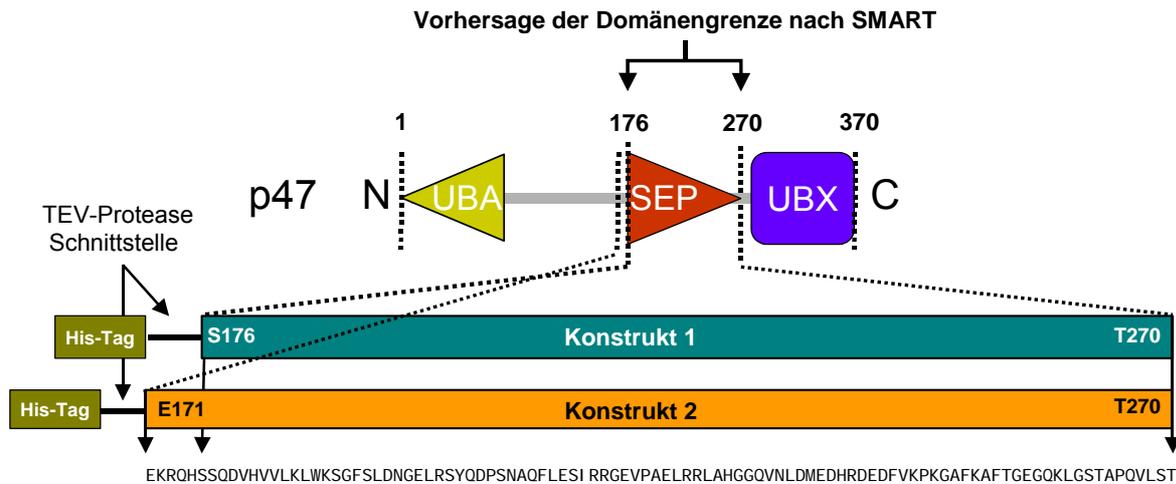
## 4 NMR-Struktur-Funktions-Bestimmung der humanen p47 SEP-Domäne

Zur NMR-Strukturbestimmung der humanen p47 SEP-Domäne wurden  $^{15}\text{N}$ - und  $^{15}\text{N}/^{13}\text{C}$ -markierte Proteinproben hergestellt. Mit diesen erfolgte die Aufnahme eines Standard NMR-Datensatzes, mit dem nahezu die vollständige Resonanzzuordnung des untersuchten Domänenkonstruktes möglich war. Aufbauend auf der abgeschlossenen Resonanzzuordnung wurden erste Strukturrechnungen mit dem Programm ARIA/CNS durchgeführt und das resultierende Strukturensamble manuell verfeinert und evaluiert, indem die durch ARIA (Ambiguos Restraints in Iterative Assignment) automatisch zugeordneten NOEs auf der Basis von manuell zugeordneten NOE-Sekundärstrukturmustern überprüft, sowie Strukturverletzungen durch korrigierte NOE-Zuordnungen ausgeräumt wurden. Durch weitere iterative Strukturrechnungen und Evaluierungsschritte ergab sich daraus das finale NMR-Strukturensamble der humanen p47 SEP-Domäne. Im weiteren Verlauf dieser Arbeit wurde für die p47 SEP-Domäne eine strukturbasierte Hypothese zur Funktion erarbeitet und experimentell bestätigt. Die p47 SEP-Domänenkonstrukte wurden mir freundlicherweise im Rahmen der Kooperation mit der Proteinstrukturfabrik Berlin zur weiteren Bearbeitung zur Verfügung gestellt. †

### 4.1 Proteinpräparation und Reinigung des humanen p47 SEP-Domänenkonstruktes G1-S2-p47 (171-270)

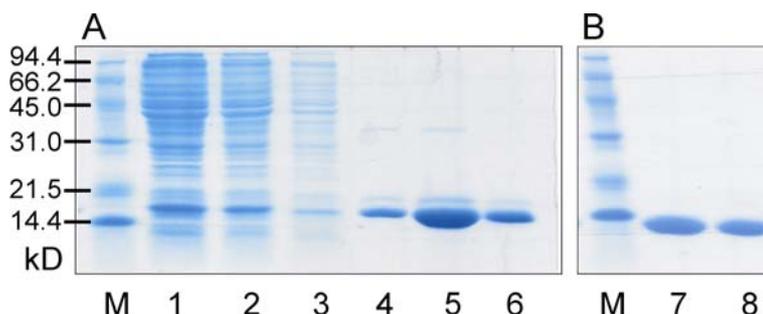
Laut Vorhersage der SMART-Datenbank (Letunic et al., 2004) wird die p47 SEP-Domäne durch die Reste S176 und T270 eingegrenzt. Entsprechend dieser Vorhersage wurde ein Konstrukt 1 (S176-T270) und ein weiteres um 5 Reste N-terminal verlängertes Konstrukt 2 (E171-T270) aus einer foetalen cDNA Expressionsbibliothek in den Expressionsvektor pQTEV kloniert (Bussow et al., 2000). Abbildung 4.1 gibt einen schematischen Überblick der klonierten Fusionsproteine wieder. Die Konstrukte sind jeweils N-terminal über eine spaltbare TEV-Protease-Schnittstelle an einen His-Tag gekoppelt. Der His-Tag ermöglicht die effiziente Aufreinigung des Fusionsproteins aus dem löslichen Zellüberstand nach Kultivierung und Aufschluß der *E. coli* (BL 21) Zellen.

†Teile dieses Kapitels wurden in der Publikation „The SEP domain of p47 acts as a reversible competitive inhibitor of cathepsin L“, Soukenik et al, FEBS Letters 576 (2004) 358-362 veröffentlicht.



**Abbildung 4.1:** Domänenorganisation von p47 (oben) und Sequenzen (E171-T270; SwissProt:Q9UNZ2) der klonierten SEP-Konstrukte (unten). Konstrukt 1 wurde in den Grenzen der SMART-Vorhersage und Konstrukt 2 um 5 Aminosäuren N-terminal verlängert kloniert. Beide Konstrukte verfügen über eine TEV-Protease-Schnittstelle zwischen His-Tag und Zielprotein.

In einem ersten Reinigungsschritt wird das His-Tag Fusionsprotein über Affinitätschromatographie an eine  $\text{Ni}^{2+}$ -Chelat Säule gebunden, von löslichen Fremdbestandteilen des Zellüberstandes getrennt und anschließend eluiert. Der His-Tag des Fusionsproteins wird über Nacht mittels TEV-Protease während eines Dialyseschrittes abgespalten. Das p47 SEP-Protein erhält man dann aus einem zweiten Reinigungsschritt über Größenausschlusschromatographie. Abbildung 4.2 zeigt die SDS-Proteingele der beiden Reinigungsschritte mit den entsprechenden Fraktionen der Proteinpräparation. Die Proteinprobe wird nach der Reinigung in den NMR-Puffer überführt, aufkonzentriert und vermessen. Beide Konstrukte zeigen ein gut aufgelöstes  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektrum aus dem gefolgert werden kann, dass beide Proteine nativ gefaltet sind. Hinsichtlich der proteinchemischen Eigenschaften zeigt das Konstrukt 2 günstigere Voraussetzungen für die NMR-Strukturbestimmung. So ergaben sich aufgrund einer 5-fach höheren Ausbeute an Tag-freien  $^{15}\text{N}$ -markiertem Protein für Konstrukt 2 (4.8 mg/100 ml Kultur), wohingegen Konstrukt 1 nur 1mg/100 ml Ausbeute zeigte. Zweitens verfügt das Konstrukt 2 über eine bessere Stabilität, da keine weiteren Proteinbanden von Verunreinigungen zu sehen sind (ohne Abbildung). Folglich wurden für die weitere NMR-Strukturbestimmung unmarkierte,  $^{15}\text{N}$ - und  $^{15}\text{N}/^{13}\text{C}$ -markierte Proben des Konstruktes 2 verwendet.



**Abbildung 4.2:** 15 % SDS-Page-Gele mit Proben nach beiden Metallchelate-Reinigungsschritten von Konstrukt 2 (A) nach primärer Affinitätschromatographie. M: Proteinmarker (kD), Spur 1-3: Durchfluß des nichtgebundenen Zellextraktes, 4-6: Säuleneluat: His-Fusionsprotein (B) nach proteolytischer Abspaltung des His-Tags und finaler Gelfiltration. M: Proteinmarker wie in (A), 7-8: reines G1-S2-p47 (171-270) Protein

## 4.2 Resonanzzuordnung

Für die Resonanzzuordnung der humanen p47 SEP-Domäne wurde das Konstrukt 2 mit den Resten G1-S2-p47(171-270) verwendet. An diesem Konstrukt wurden alle NMR-Messungen an  $^{15}\text{N}$ - und  $^{15}\text{N}/^{13}\text{C}$ -markierter Proben durchgeführt.

### 4.2.1 Strategie der Resonanzzuordnung

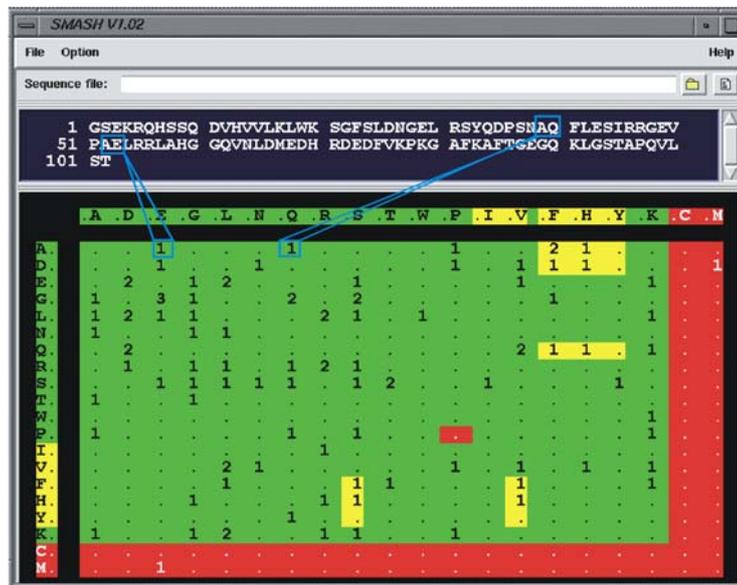
Die sequentielle Zuordnung der Signale des Peptidrückgrates erfolgte unter Einbeziehung neuer Aminosäure-selektiver  $^1\text{H}$ - $^{15}\text{N}$  Korrelationsexperimente (Schmieder et al., 1998, Schubert et al., 2000, Schubert et al., 2001, Schubert et al., 1999) in die klassische 3D-Tripelresonanz-Strategie. Damit sollte die mögliche Automatisierung der Resonanzzuordnung getestet werden. Tabelle 4.1 zeigt die aus den aufgenommenen NMR-Experimenten erhaltenen Informationen.

**Tabelle 4.1:** Überblick der erhaltenen Informationen aus den aufgenommenen NMR-Spektren für die Signalzuordnung und Strukturrechnung

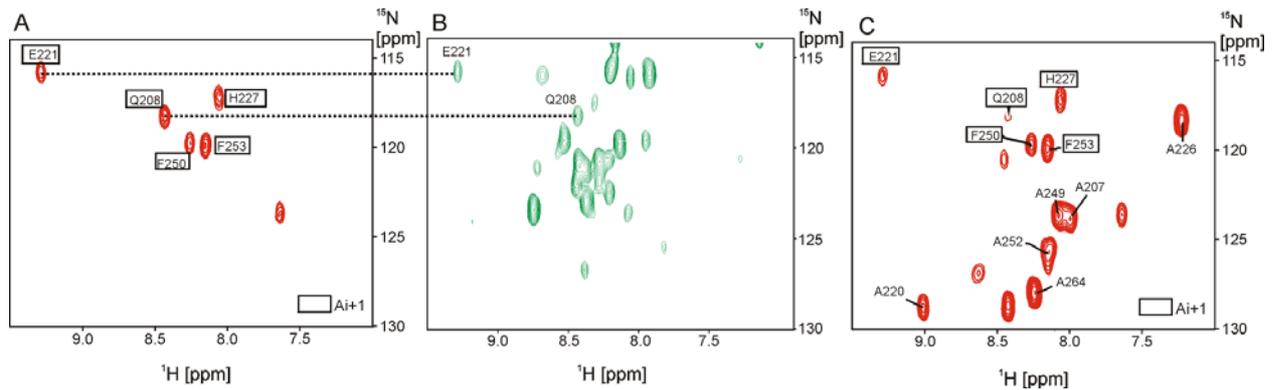
Zuordnung von	Zielstellung	NMR-Spektrum	Signale ( $^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$ )
<b>Peptid-rückgrat</b>	<i>fingerprint</i>	2D- $^1\text{H}$ - $^{15}\text{N}$ HSQC	$\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
	eindeutige Ankerpunkte	Selektive Aminosäure-Experimente	$\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
	Segmentbildung	3D-CBCA(CO)NNH/ CBCANNH	$\text{C}^{\alpha,\beta}(i-1)$ $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$ $\text{C}^{\alpha,\beta}(i,i-1)$ $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
	Carbonylzuordnung und Verifizierung von $\text{H}^{\text{N}}$	3D-HNCO	$\text{C}^{\prime}(i)$ , $\text{H}^{\text{N}}(i)$
	Brückenfunktion	Selektive Prolin-Experimente	$\text{N}(p-1,p+1)$ , $\text{H}^{\text{N}}(p-1,p+1)$
<b>Seitenkette</b>	aliphatisch	3D-CBCA(CO)NNH/ CBCANNH	$\text{C}^{\alpha,\beta}(i-1)$ $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$ $\text{C}^{\alpha,\beta}(i,i-1)$ $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
		3D-HBHA(CO)NH	$\text{H}^{\alpha,\beta}(i-1)$ , $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
		3D-H(CCCO)NH-TOCSY	$\text{H}^{\text{ali}}(i-1)$ , $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
		3D-(H)CC-(CO)NH-TOCSY	$\text{C}^{\text{ali}}(i-1)$ , $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
		3D-HCCH-COSY	$\text{H}^{\text{ali}}(i)$ , $\text{C}^{\text{ali}}(i)$ , $\text{H}^{\text{ali}}(i)$
		3D-(H)CCH-COSY	$\text{C}^{\text{ali}}(i)$ , $\text{C}^{\text{ali}}(i)$ , $\text{H}^{\text{ali}}(i)$
	Asn/Gln	2D-NH <sub>2</sub> -selektive COSY/TOCSY	$\text{C}^{\text{ali}}$ , $\text{H}^{\text{ali}}$ , $\text{NH}_2$
	Met	Selektive Experimente	$\text{C}^{\epsilon}\text{-CH}_3$
	aromatisch	3D- $^{13}\text{C}$ edited-HMQC NOESY	$\text{C}^{\text{aro}}$ , $\text{H}^{\text{aro}}$
<b>Struktur-rechnung</b>	Gewinnung von Abstandsinfos für die Strukturermittlung	3D- $^{13}\text{C}$ edited-HMQC NOESY (80ms)	$^1\text{H}^{\text{NOE}}$ , $\text{C}(i)$ , $\text{H}^{\text{C}}(i)$
		3D- $^{15}\text{N}$ edited -HSQC NOESY (80ms)	$^1\text{H}^{\text{NOE}}$ , $\text{N}(i)$ , $\text{H}^{\text{N}}(i)$
		2D- $^1\text{H}$ - $^1\text{H}$ NOESY (60ms)	$^1\text{H}$ - $^1\text{H}^{\text{NOE}}$

#### 4.2.2 Zuordnung der Signale des Peptidrückgrats

Die Sequenz des Konstruktes 2 wurde mit dem Softwareprogramm SMASH hinsichtlich geeigneter spezifischer Aminosäurepaare analysiert. Dabei identifizierte SMASH insgesamt 70 einzelne mögliche Aminosäurepaare (Ziffer 1), die durch den kompletten Satz Aminosäure-selektiver Experimente abgedeckt werden können (siehe Abbildung 4.3.). Darin sind einmalig vorkommende, 61 eindeutige (grün) und 9 mehrdeutige (gelb) Aminosäurepaare enthalten. Für Rot eingefärbte Paare stehen keine Aminosäure-selektiven Experimente vorhanden. Jedes dieser eindeutigen einmaligen Paare kann für die sequentielle Zuordnung als Startpunkt benutzt werden. Die türkisblau eingerahmten Aminosäurepaare AE und AQ kommen genau einmal in der Proteinsequenz vor. Mit Hilfe der Aminosäure-selektiven NMR-Experimente ist es möglich, eindeutige Paare zuzuordnen und sichere Ankerpunkte für die Zuordnung des Proteindrückgrats zu finden. Durch Kombination der selektiven Experimente A(i+1) und EQG(i,i+1), dargestellt in Abbildung 4.4 konnten die Aminosäurepaare A220/E221 und A207/Q208 eindeutig zugeordnet werden (Die SMASH Annotation 52/53 für AE in Abbildung 4.3 entspricht der p47 Zuordnung 220/221). Zur Unterscheidung der beiden Paare diente ein weiterer Vergleich des sich anschließenden Segmentsabschnitts mit der Aminosäuresequenz A220-A226 (siehe dazu Abbildung 4.5).



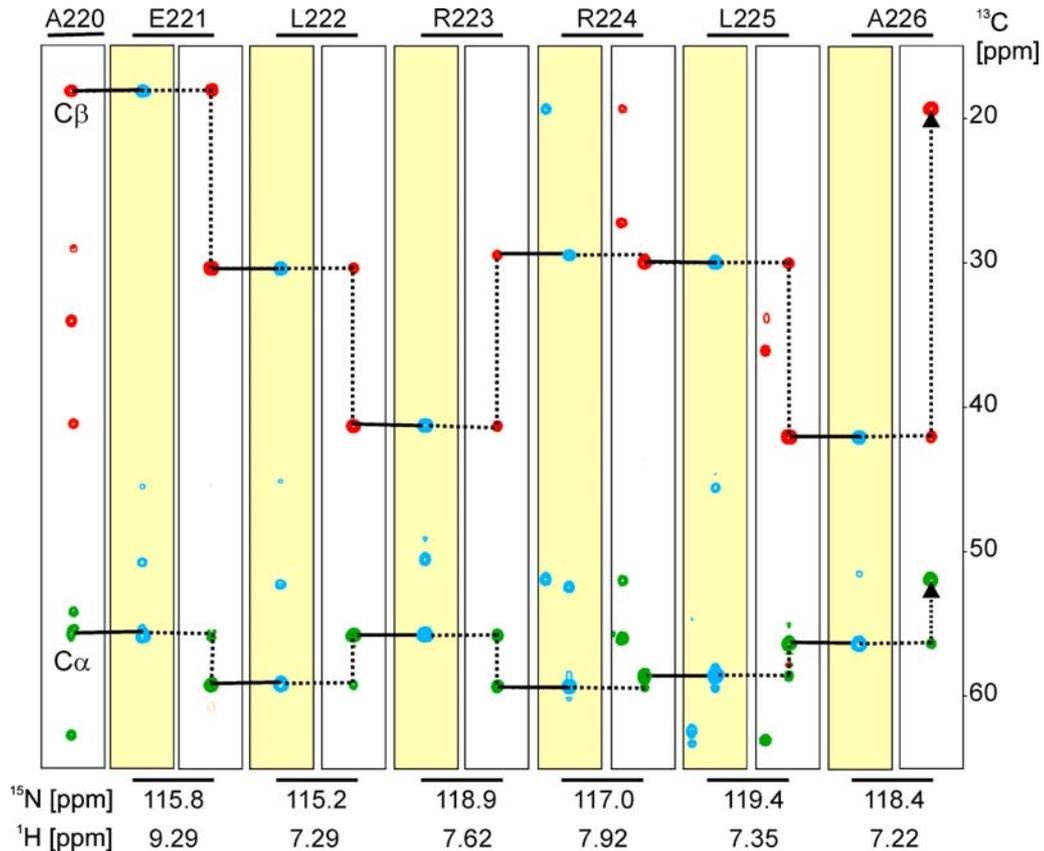
**Abbildung 4.3:** SMASH-Matrix für das verwendete Konstrukt 2. Jedes Aminosäurepaar ist entsprechend seiner Häufigkeit (1, 2, 3) in der Proteinsequenz dargestellt. Die hellblau eingerahmten Paare AE und AQ kommen genau einmal in der Sequenz vor. Durch die Aminosäure-selektiven Experimente sind die Paare AE und AQ eindeutig erfassbar und somit potentiell geeignete Ankerpunkte für die sequentielle Zuordnung des Peptidrückgrates. Durch geeignete Kombination der Aminosäure-selektiven Experimente identifizierbare Paare: grün-eindeutige, gelb-mehrdeutige, rot-nicht verfügbar



**Abbildung 4.4:** Identifizierung von Ankerpunkten. Die Aminosäurepaare A220/E221 und A207/Q208 kommen nur einmal in der Sequenz vor und durch den Vergleich der Aminosäureselektiven Spektren A(i+1) gezeigt in (A) und EQG(i,i+1) (B) zugeordnet werden. Spektren: (A) A(i+1), (B) EQG(i,i+1) und (C) A(i,i+1).

Im Verlauf der Rückgratzuordnung wurden insgesamt vier weitere Alanine (A226, A249, A252 und A264) sowie deren Nachfolgerreste (H227, F250, F253) vollständig zugeordnet. Alanin 264 besitzt als Nachfolgerrest ein Prolin. Somit entfällt ein Signal für den Nachfolgerrest im A(i,i+1) Spektrum (Abbildung 4.4 (C)).

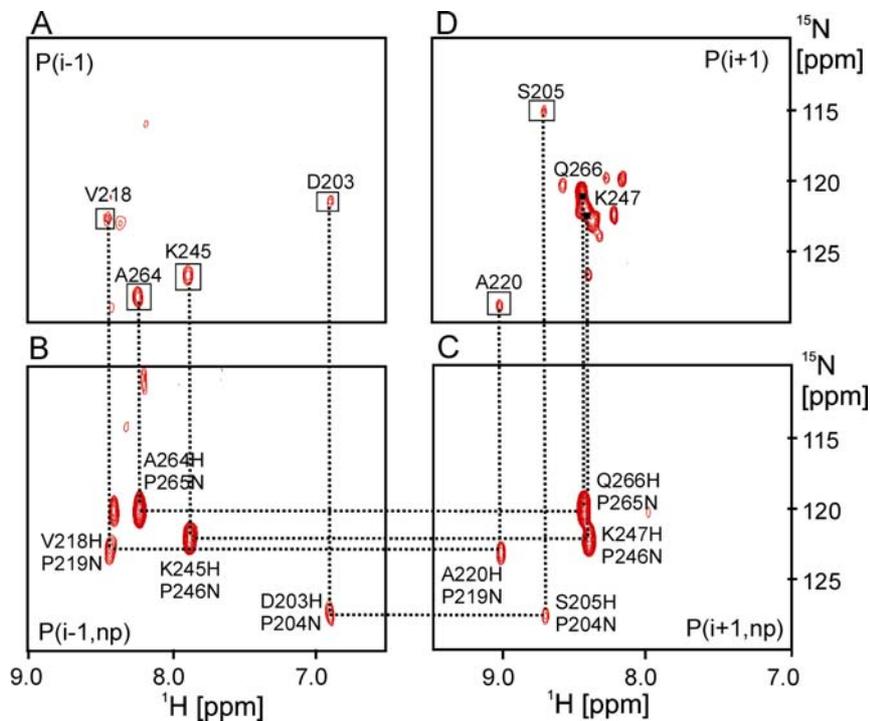
In Kombination mit dem 3D-Tripelexperimentepaar CBCA(CO)NNH/CBCANNH und spezifischen Ankerpunkten ist es möglich, über die Verknüpfung von C $^{\alpha}$ - und C $^{\beta}$ -chemischen Verschiebungen die restlichen Zuordnungen zu treffen (Abbildung 4.5).



**Abbildung 4.5:** Streifenausschnitte aus dem Spektrenpaar CBCANNH (weiß unterlegt; grün (negative Phase):C<sup>α</sup>- Signale; rot (positive Phase):C<sup>β</sup>-Signale) und CBCA(CO)NNH (gelb unterlegt; blau (gleiche Phase): C<sup>α</sup>- und C<sup>β</sup>-Signale) mit den erhaltenen sequentiellen Konnektivitäten für den Sequenzabschnitt A220-A226. Die Verknüpfung der Kette durch die C<sup>α</sup>- und C<sup>β</sup>-Signale ist mit durchgezogenen Linien markiert. Der gesamte Verknüpfungsweg ist gestrichelt dargestellt.

Ausgehend von gefundenen Ankerpunkten wurden Segmente der Proteinsequenz mit Hilfe von 3D-CBCA(CO)NNH und CBCANNH Spektren gebildet und spezifischen Positionen in der Proteinsequenz zugewiesen. C<sup>α</sup>- und C<sup>β</sup>-Signale erscheinen im 3D-CBCANNH Spektrum mit unterschiedlichen Phasen (-/+, grün/rot) und sind dadurch innerhalb dieses Spektrums unterscheidbar, während sie im 3D-CBCA(CO)NNH Spektrum die gleiche Phase (blau/blau) besitzen. Der Sequenzabschnitt A220-A226 in Abbildung 4.5 zeigt exemplarisch die Verknüpfung der Aminosäuren über die C<sup>α</sup> und C<sup>β</sup> chemischen Verschiebungen. Dadurch ist es möglich, sich in der Sequenz vom C- zum N-Terminus und vice versa zu bewegen, wodurch weiterführende Segmente entstehen. Anhand der C<sup>α</sup>/C<sup>β</sup> chemischen Verschiebung ist es auch möglich einen Hinweis auf den Typ der Aminosäure zu erhalten (Seavey et al., 1991). Unterbrochen werden die Segmente durch Prolin. Proline besitzen kein H<sup>N</sup>, weshalb eine Lücke in der Segmentbildung entsteht, die durch Prolin-selektive Experimente geschlossen werden

kann. Somit können die Ketten auch über Prolinreste hinweg verlängert werden. Abbildung 4.6 verdeutlicht den Lückenschluss durch selektive Prolin-Experimente.



**Abbildung 4.6:** Prolin-selektive Spektren. Die vier Ausschnitte (A,B,C,D) zeigen gegen den Uhrzeigersinn die Zuordnung der Prolin-Vorgänger und -Nachfolger. Das ermöglicht die Überbrückung der Segmentketten über das Prolin hinweg. Aufgeführte Experimente: (A) P(i-1), (B) P(i-1,np), (C) P(i+1,np) und (D) P(i+1) mit np =  $^{15}\text{N}$  ( $\text{P}_i$ ).

Für eine effiziente Zuordnung des Peptidrückgrates erwiesen sich die insgesamt 33 Aminosäure-selektiven Experimente zum größten Teil als vorteilhaft (siehe Tabelle 4.2: Verwendung der selektiven Experimente für den Zuordnungsprozeß). Insbesondere die 10 selektiven Experimente für Glycin, Serin, Alanin und Prolin zeigten die zu erwartenden Signale für die N- $\text{H}^{\text{N}}$ -Korrelationen und lieferten wertvolle Daten für die Bildung von geeigneten Ankerpunkten in der sequentiellen Zuordnung. 10 selektive Experimente erwiesen sich als nicht hilfreich für den Zuordnungsprozeß. Darunter sind die Experimente für Tryptophan, Lysin und Arginin, in denen keine Kreuzsignale auftraten. 13 Experimente konnten erst nach Abschluß der Zuordnung zur Verifizierung der Resonanzen der Proteinhauptkette verwendet werden, da entweder zusätzliche (DNG(i+1), DNG(i,i+1) u.a.) oder zu wenige Signale (LA(i+1), LA(i,i+1)) in den selektiven Spektren auftraten. Es ließen sich von den insgesamt 61 eindeutigen Aminosäurepaaren 8 Paare durch die neuen Experimente eindeutig zu ordnen.

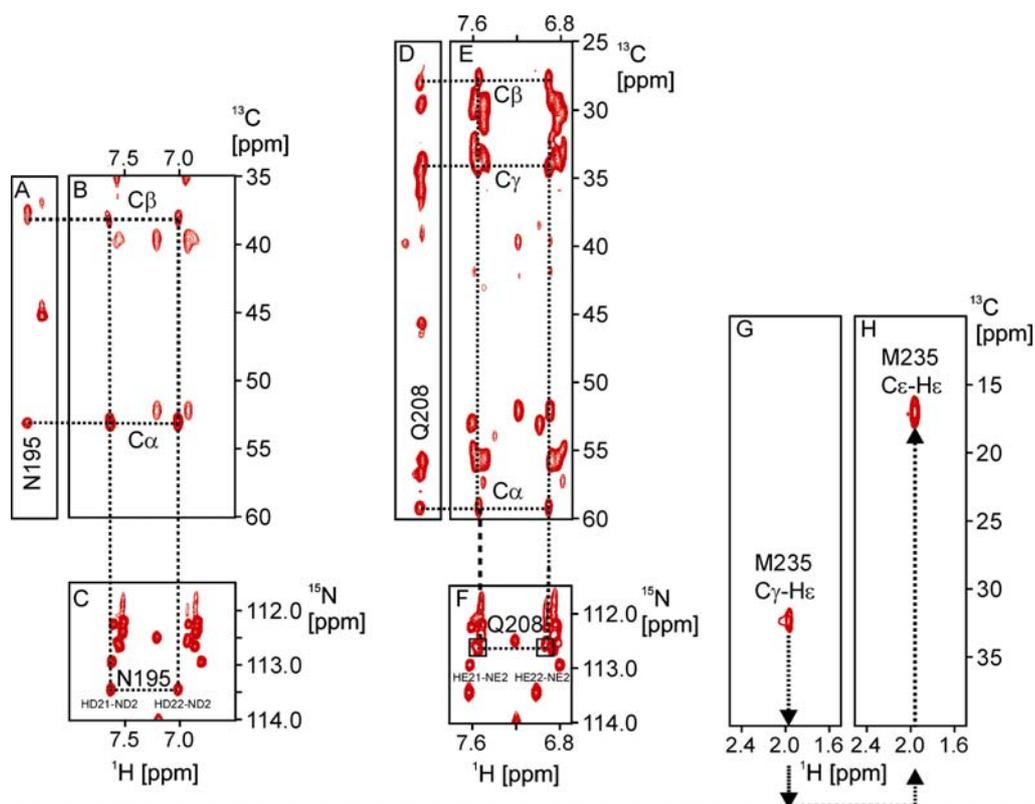
**Tabelle 4.2:** Überblick über die aufgenommenen Aminosäure-selektiven Experimente und deren Verwendung zur Zuordnung der Signale des Proteinerückgrats.

Selektives Experiment	Signale			Vollständigkeit (%)	Verwendung in der Rückgratzuordnung
	beobachtet	erwartet	zugeordnet		
G(i+1)	12	10	10	100	Ja
G(i,i+1)	21	19	18	95	Ja
A(i+1)	6	5	5	100	Ja
A(i,i+1)	16	11	11	100	Ja
S(i+1)	10	10	10	100	Ja
S(i,i+1)	20	19	18	95	Ja
P(i+1)	10	4	4	100	Ja
P(i-1)	5	4	4	100	Ja
P(i+1,np)	4	4	4	100	Ja
P(i-1,np)	5	4	4	100	Ja
TA(i+1)	7	7	5	71	Verifizierung
TA(i,i+1)	18	15	15	100	Verifizierung
VIA(i+1)	17	12	10	83	Verifizierung
VIA(i,i+1)	25	26	21	77	Verifizierung
LA(i+1)	24	15	12	80	Verifizierung
N(i+1)	7	3	3	100	Verifizierung
N(i,i+1)	6	6	4	67	Verifizierung
QN(i+1)	11	10	9	90	Verifizierung
QN(i,i+1)	14	20	13	65	Verifizierung
EQG(i,i+1)	40	43	18	42	Verifizierung
FYH(i+1)	11	10	8	80	Verifizierung
D(i+1)	5	6	5	83	Verifizierung
D(i,i+1)	10	13	8	62	Verifizierung
LA(i,i+1)	-	30	-	-	Nein
DNG(i+1)	40	19	-	-	Nein
DNG(i,i+1)	42	35	-	-	Nein
EQG(i+1)	-	25	-	-	Nein
FYH(i,i+1)	8	20	-	-	Nein
W(i+1)	-	1	-	-	Nein
W(i,i+1)	-	2	-	-	Nein
KR(i+1)	-	13	-	-	Nein
K(i,i+1)	-	13	-	-	Nein
E(i+1)	-	8	-	-	Nein

Neben den eindeutigen Zuordnungen ergaben sich auch mehrdeutige Zuordnungen, die durch den weiteren Zuordnungsprozeß geklärt werden konnten. Tabelle 4.2 gibt einen Überblick über die Ausschöpfung des Informationsgehaltes der Aminosäure-selektiven Experimente.

#### **4.2.3 Zuordnung der aliphatischen und aromatischen Seitenkettensignale**

Mit Abschluß der sequentiellen Zuordnung des Peptidrückgrates konnte die Resonanzzuordnung auf die aliphatischen und aromatischen Kerne der Seitenketten ausgedehnt werden. Aufbauend auf den erhaltenen Signalen für die einzelnen Aminosäuren im  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrum wurden die  $\text{H}^\alpha$ - und  $\text{H}^\beta$ -Resonanzen durch das 3D-HBHA(CO)NNH Spektrum bestimmt. Die Zuordnung der  $\gamma$ -Methyl oder  $\gamma$ -Methylengruppen der Seitenkettensignale der unverzweigten Reste Glutaminsäure, Glutamin und Threonin gelang durch den direkten Vergleich mit dem 3D-H(CCO)NNH/(H)CC(CO)NNH-TOCSY-Paar. Diese Experimente korrelieren die Protonen- bzw. Kohlenstofffrequenzen mit den  $\text{N-H}^{\text{N}}$ -Resonanzen der nachfolgenden Aminosäure. Die Zuordnung der Seitenkettensignale der verzweigten Aminosäuren Valin, Isoleucin und Leucin erfolgte unter Verwendung der 3D-HCCH/(H)CCH-COSY Spektren. Diese Spektren enthielten auch die notwendigen Informationen für die weitere Zuordnung der  $\gamma$ -,  $\delta$ - und  $\epsilon$ -Methylengruppen der Methionin-, Prolin- und Lysin-Reste. Durch die selektiven  $\text{NH}_2$ -Seitenkettenexperimente für Glutamin und Asparagin ( $\text{N/Q-C}^\alpha/\text{C}^\beta/\text{C}^\gamma$ ; unveröffentlichte Daten) ist es möglich, ausgehend von den gefundenen  $\text{C}^\alpha$ ,  $\text{C}^\beta$  und  $\text{C}^\gamma$  Resonanzen (für Glutamin) die zugehörige  $\text{NH}_2$ -Seitenkettengruppe zu annotieren (siehe dazu Abbildung 4.7). Darüber hinaus konnte die  $\epsilon$ -Methylgruppe des Methionin 235 durch die selektiven  $^{13}\text{C}/^1\text{H}$ - Experimente (unveröffentlichte Daten) bestimmt werden (Abbildung 4.7 (G, H)).



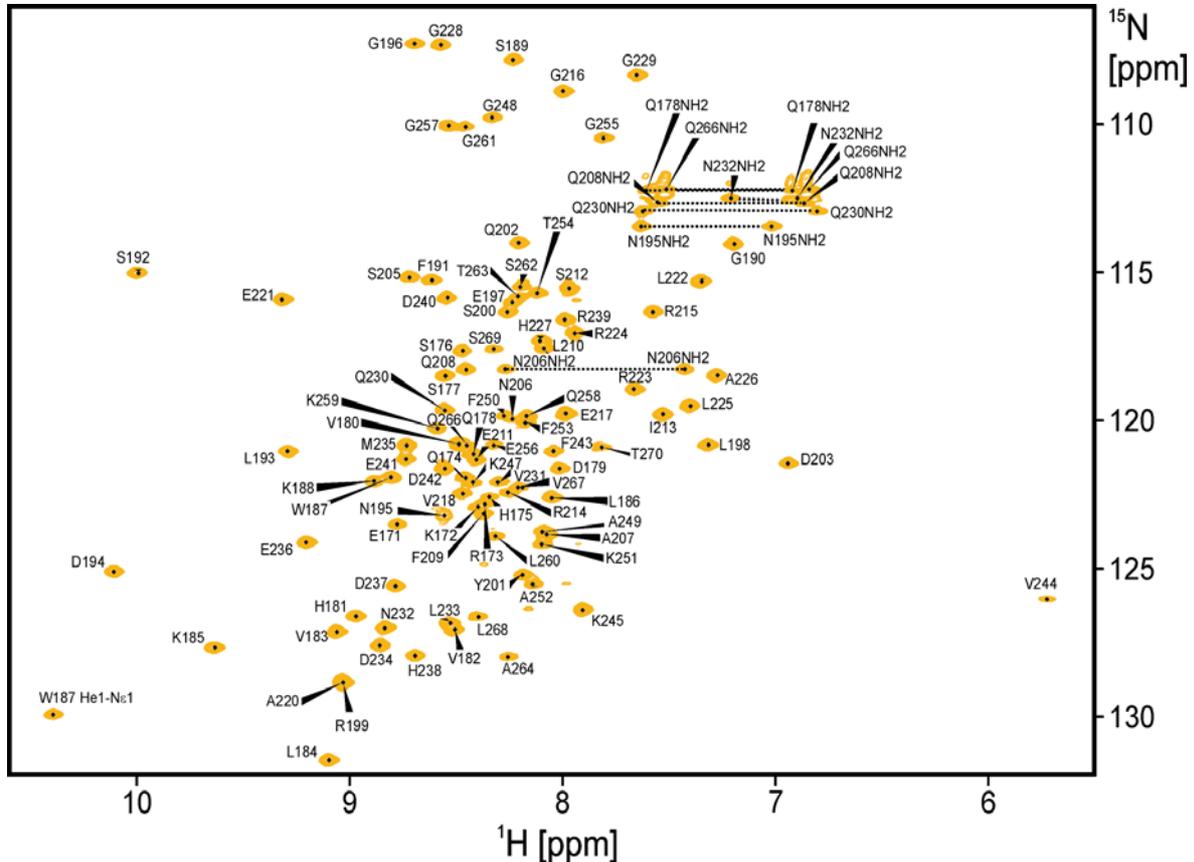
**Abbildung 4.7:** Die Spektren der N/Q-C1/C2/C3 Experimente (B, E) und der Methionin  $^{13}\text{C}/^1\text{H}$ -Experimente (G, H) vereinfachen die Seitenkettenzuordnung der charakteristischen Atomgruppen. Die Kombination von  $\text{C}^\alpha$ - und  $\text{C}^\beta$ -chemischen Verschiebung am Beispiel des N195 (A; Streifenausschnitt des CBCA(CO)NNH) und (B; selektives N-C1/C2-Experiment) ermöglicht die Zuordnung der  $\text{NH}_2$ -Seitenkettengruppe im  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektrum (C). In der Mitte (D, E, F) dargestellt ist in analoger Weise die Zuordnung für die Q208  $\text{NH}_2$ -Gruppe einschließlich der  $\text{C}_\gamma$ -chemischen Verschiebung. Die Spektren (G, H) zeigen die Zuordnung der M235  $\epsilon$ -Methylgruppe durch die selektiven Experimente für Methionin.

Die Zuordnung aromatischer Seitenkettenresonanzen geschah durch Verwendung intraresidualer  $\text{H}^\beta$ - $\text{H}^\epsilon$  NOE-Kontakte beobachtet im  $3\text{D-}^{13}\text{C}$ -NOESY Spektrum. So ist es möglich, die aromatischen Seitenketten von Tyrosin und Phenylalanin aufgrund der unterschiedlich starken Signalintensitäten der  $\text{H}^\delta$ - und  $\text{H}^\epsilon$ -Resonanzen innerhalb des Aromaten zu identifizieren. Für das umfangreichere Ringsystem des einzigen Tryptophans wurden noch zusätzlich  $\text{H}^{\epsilon 1}$ - $\text{N}^{\epsilon 1}$  NOE Kontakte aus dem  $3\text{D-}^{15}\text{N}$ -NOESY verwendet.

#### 4.2.4 Abschluß der Resonanzzuordnung

Abbildung 4.8 zeigt das zugeordnete  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektrum des untersuchten Konstruktes G1-S2-p47(171-270), aufgenommen mit einem DRX 600 MHz Spektrometer von Bruker, welches mit einem Kryo-Probenkopf ausgestattet ist. Die Resonanzen des Spektrums sind gut aufgelöst und liegen in einem typischen Bereich für ein gefaltetes

Protein. Jede einzelne Aminosäure außer Prolin wird durch ein  $N-H^N$  Signal im Spektrum repräsentiert. Zusätzlich treten die Signale der  $NH_2$ -Seitenkette für die Aminosäuren Glutamin und Asparagin sowie des  $N-H^N$  Signals des Tryptophan Seitenkettenindols auf.



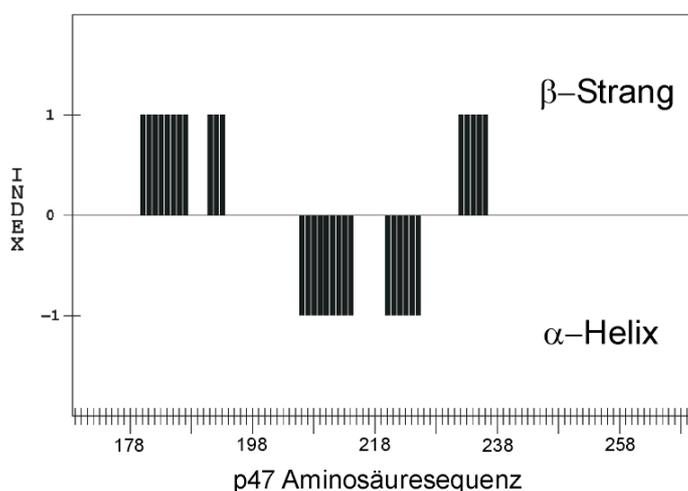
**Abbildung 4.8:**  $^1H$ - $^{15}N$  HSQC-Spektrum (*fingerprint*) des SEP-Konstruktes G1-S2-p47(171-270) aufgenommen auf einem DRX 600 MHz Spektrometer (Bruker) bei 300 K. Die  $NH_2$ -Seitenketten-Kreuzsignale der Q- und N-Reste sind mit gestrichelten Linien verbunden.

Ein Vergleich mit der BioMagResBank ([www.bmrb.wisc.edu](http://www.bmrb.wisc.edu)) ergibt, dass die Wasserstoffresonanzfrequenz des Restes Valin 244 eine anomale Verschiebung zu einer niedrigeren Verschiebung von 5.8 ppm zeigt. Der Standardwert für das Peptidrückgratproton  $N-H^N$  liegt im Bereich von 6.8-10.5 ppm (Cavanagh, 1996). Dieser außergewöhnliche chemische Verschiebungswert lässt sich aus der finalen NMR-Struktur erklären. Die Hochfeld-Verschiebung zu kleineren Resonanzfrequenzen wird durch die unmittelbare Nähe des Protons im entschirmenden Bereich der aromatischen Ebene des sequentiell benachbarten Phenylalanin 243 verursacht.

Anhand der verwendeten NMR-Experimente war die nahezu vollständige manuelle Resonanzzuordnung (96 %) der  $^{13}\text{C}$ ,  $^{15}\text{N}$  und  $^1\text{H}$  Frequenzen des humanen G1-S2 p47 (171-270) SEP-Konstruktes möglich. Lediglich die aromatischen Ringe der Reste F250 und F253 konnten nicht zugeordnet werden vermutlich aufgrund der erhöhten Flexibilität dieses Sequenzabschnittes (siehe dazu Abschnitt 4.4.2). Die gesamte Resonanzliste der zugeordneten Kerne ist dem Anhang A.1.1 Resonanzzuordnung des Konstruktes G1-S2 p47 (171-270) entnehmbar.

#### 4.2.5 Sekundärstrukturvorhersagen auf Basis der gefundenen chemischen Verschiebungen

Unter Verwendung der ermittelten chemischen Verschiebungen der  $\text{C}^\alpha$ - und  $\text{C}^\beta$ -Resonanzen wurde eine Sekundärstrukturvorhersage mit *CSI (Chemical Shift Index)* (Wishart & Sykes, 1994) durchgeführt. Zusätzlich wurden die Carbonyl-chemischen Verschiebungen aus einem zuvor zugeordneten 3D HNCO-Experiment verwendet. Die Vorhersage ergab für die Reste 181-189, 190-192 und 233-238 das Vorliegen von  $\beta$ -Strang Konformationen sowie einer  $\alpha$ -helikalen Konformation für die Reste 206-216 und 220-226.



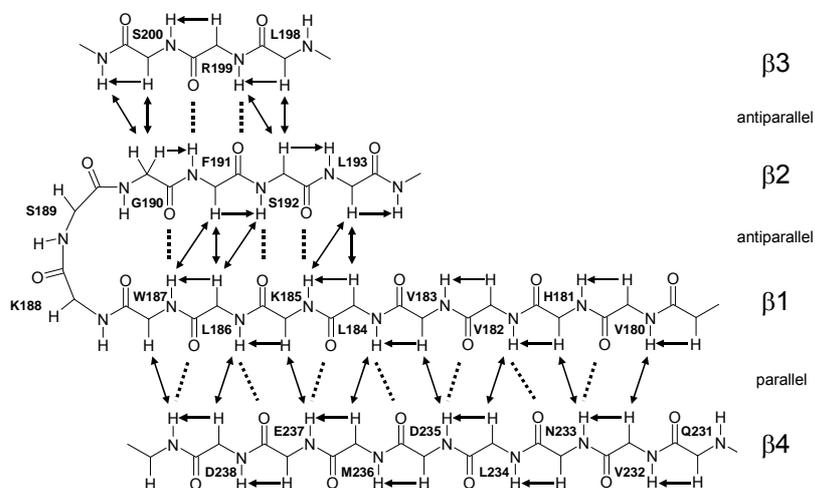
**Abbildung 4.9:** CSI-Sekundärstrukturvorhersage basierend auf  $\text{H}^{\text{N}}/\text{CO}/\text{C}^\alpha/\text{C}^\beta$  chemischen Verschiebungen. Positive Werte deuten eine  $\beta$ -Strang Konformation an. Negative Werte weisen auf einen  $\alpha$ -helikalen Bereich hin.

### 4.3 Berechnung der Proteinstruktur mit ARIA/CNS

Nach dem Abschluß der Resonanzzuordnung des Proteins wurde die 3-D Struktur errechnet. Zu diesem Zwecke wurden alle verfügbaren strukturelevanten Abstands-informationen (*restraints*) wie NOEs, Wasserstoffbrückenbindungen und Torsionswinkel in die Strukturberechnung mit einbezogen. Die wichtigsten Parameter sind die Proton-Proton-Abstände, die über Integration des Volumens der NOE-Signale (durchgeführt mit SPARKY) ermittelt wurden. Die integrierten Werte der NOE-Signale wurden während der automatischen Strukturrechnung in ARIA/CNS kalibriert.

#### 4.3.1 Ermittlung zusätzlicher Abstandsrandbedingungen

Erste Strukturrechnungen unter Verwendung ausschließlich nicht-zugeordneten NOE mit ARIA/CNS ergaben kein konvergierendes Strukturensamble. Allerdings konnte ein  $\beta$ -Faltblatt bestehend aus 4  $\beta$ -Strängen aufgrund seiner charakteristischen weitreichenden (*long-range*) NOE-Signale (siehe Abbildung 4.10) durch manuelle Zuordnung gefunden werden. Abbildung 4.10 zeigt das NOE-Netzwerk des  $\beta$ -Faltblattes, welches aus den Signalen der 3D- $^{13}\text{C}$ -editierten-NOESY und 3D-editierten- $^{15}\text{N}$ -NOESY Spektren erhalten wurde.



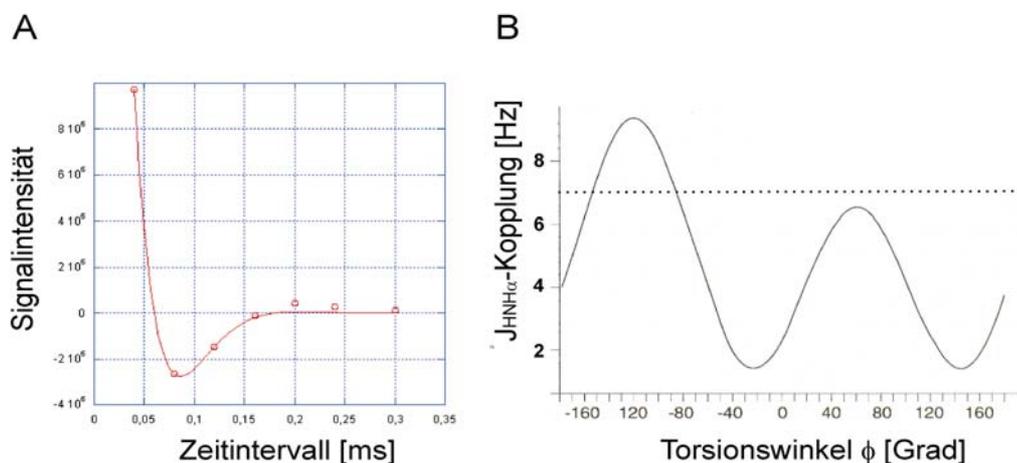
**Abbildung 4.10:** Manuell zugeordnetes Netzwerk der beobachteten weitreichenden NOE-Kreuzsignale (Doppelpfeile) und sequentiellen NOE-Kreuzsignale (Einfachpfeil) innerhalb und zwischen den  $\beta$ -Strängen 1-4. Die angenommenen Wasserstoffbrücken des  $\beta$ -Faltblattes sind in Form von gestrichelten Linien eingezeichnet.

Doppelpfeile in Abbildung 4.10 repräsentieren manuell zugeordnete NOE-Kreuzsignale mit großer Intensität zwischen  $\text{H}^{\alpha}_{(i)}\text{-H}^{\alpha}_{(j)}$  und  $\text{H}^{\text{N}}_{(i)}\text{-H}^{\alpha}_{(j)}$  Resonanzen. Einfachpfeile markieren die für  $\beta$ -Stränge gefundenen sequentiellen  $\text{H}^{\alpha}_{(i-1)}\text{-H}^{\text{N}}_{(i)}$  NOE-Signale mit großer

Intensität. Neben den durch CSI vorhergesagten drei  $\beta$ -Strängen konnte zusätzlich ein vierter  $\beta$ -Strang (L198-S200) durch die manuelle Zuordnung lokalisiert werden. Insgesamt ergab die Analyse des NOE-Netzwerkes geringe Änderungen bezüglich der Start- und Endpunkte der einzelnen  $\beta$ -Stränge im Vergleich zur CSI-Vorhersage.

In den folgenden Strukturberechnungen wurden die Liste der nicht-zugeordneten NOE-Signale durch die in Abbildung 4.10 gezeigten manuellen Zuordnungen ergänzt. Abgeleitet aus den verifizierten Sekundärstrukturelementen für die  $\beta$ -Stränge als auch für die  $\alpha$ -Helices (nicht gezeigt) wurde die Datengrundlage für die Strukturrechnung durch Wasserstoffbrückenbindungen und durch  $\phi/\psi$ -Torsionswinkel mit dem Standardwert  $-135 \pm 30^\circ$  für  $\beta$ -Faltblattreste ergänzt.

Weiter wurden experimentell Werte für die Torsionswinkel von  $\phi$  ermittelt. Dazu wurden aus einer Serie von 2D J-modulierten  $^{15}\text{N}$ - $^1\text{H}$  COSY-Experimenten die  $^3J_{\text{HNH}\alpha}$ -Kopplungskonstanten abgeleitet (Anhang C.1). Die Experimente wurden mit unterschiedlichen J-Evolutionszeiten aufgenommen.

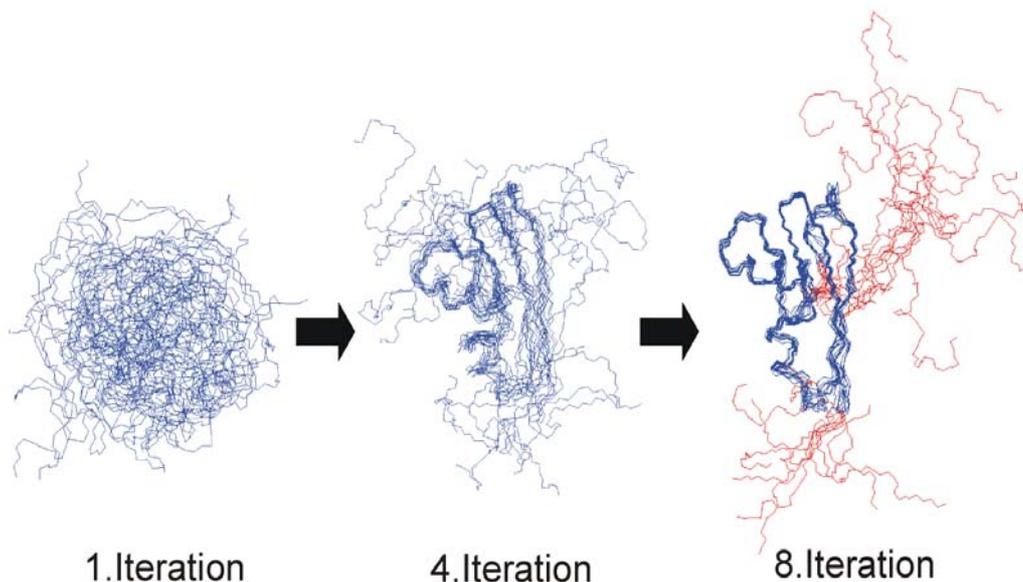


**Abbildung 4.11:** (A) Datenpunkte und Kurvenfit für V183, ( $^3J_{\text{HNH}\alpha} = 8.3 \text{ Hz}$ ), (B) Abschätzung der  $\phi$ -Torsionswinkel anhand der Karplus-Beziehung für die ermittelte Kopplungskonstante.

Durch Kurvenfit der Signalintensitäten (Abbildung 4.11(A)) wurden die  $^3J_{\text{HNH}\alpha}$ -Kopplungskonstanten bestimmt und über die Karplus-Beziehung  $^3J_{\text{HNH}\alpha} = A \cos^2\phi + B \cos\phi + C$  (siehe Abbildung 4.11 (B)) in  $\phi$ -Torsionswinkel übersetzt. Aufgrund der Mehrdeutigkeit der Karplus-Beziehung wurden nur Kopplungskonstanten berücksichtigt, die größer als 7 Hz waren. Für Kopplungskonstanten  $^3J_{\text{HNH}\alpha} > 7 \text{ Hz}$  gibt es nur 2 Lösungen, während für  $^3J_{\text{HNH}\alpha} < 7 \text{ Hz}$  insgesamt 4 Lösungsmöglichkeiten existieren.

### 4.3.2 Finale Strukturrechnung der humanen p47 SEP-Domäne

Unter Ausschöpfung aller strukturrelevanter Abstandsbedingungen ergänzt durch experimentell und abgeleitete  $\phi$ - $\psi$ -Torsionswinkel konnten nun konvergierende Strukturrechnungen durchgeführt werden. In Abbildung 4.12 ist der finale Verlauf der Strukturrechnung für die humane p47 SEP-Domäne gezeigt. Ausgehend von einem undefiniertem „Strukturhaufen“ nach Abschluß des ersten Rechenzyklus entwickelt sich über insgesamt 8 Rechenzyklen ein gut definiertes, „scharfes“ NMR-Strukturensemble.



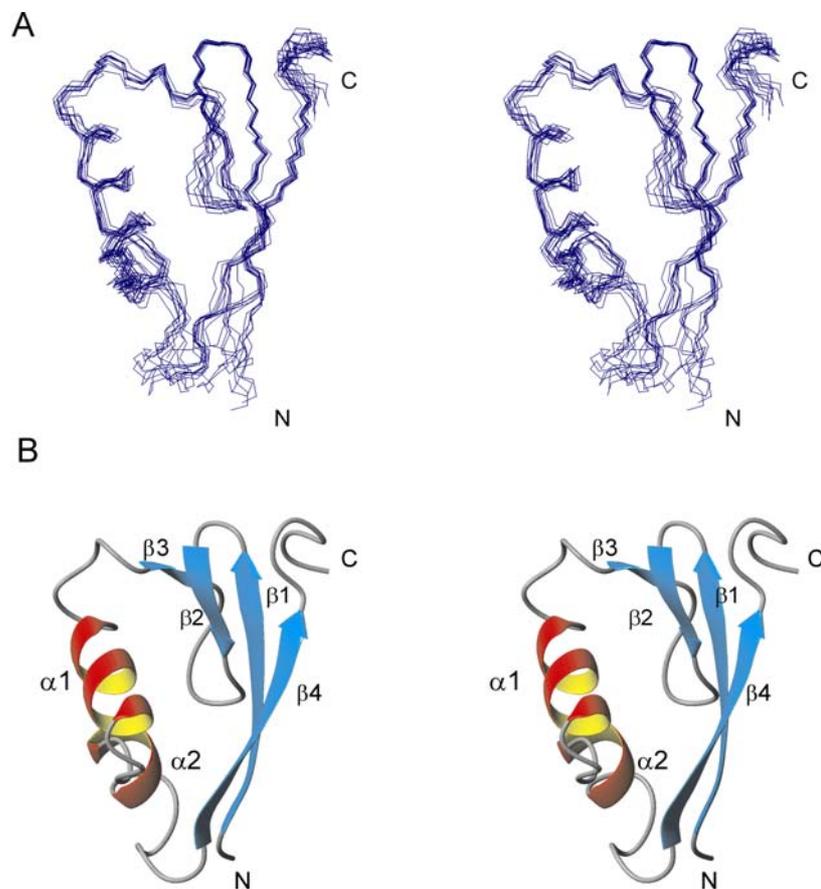
**Abbildung 4.12:** 8-stufiger iterativer Verlauf der ARIA/CNS Strukturrechnung. Iteration 1: Startensemble, Iteration 8: finales Strukturensemble. In der Rechnung ist (rechts) der unstrukturierte (rot) und der strukturierte (blau) Teil der Proteinkette farblich gegeneinander abgesetzt. Proteinrückgrat-RMS (*root mean square*) Abweichungen (Å) für die strukturierten Reste 179-244 sind von links nach rechts:  $8,58 \pm 1,58$ ,  $2,17 \pm 0,45$  und  $1,04 \pm 0,24$  (Koradi et al., 1996)

Während der iterativen Läufe optimiert ARIA die getroffenen NOE-Zuordnungen durch Rauschentfernung und Neuordnung unter Beibehaltung der manuell eingefügten fixierten NOE-Signalmuster und Torsionswinkel. Insgesamt konnten durch ARIA 1570 eindeutig NOE-Abstandsbedingungen während der 8. Iteration zugeordnet werden, darunter 165 mittellange (*medium-range*) Distanzlimits, die für die Bestimmung der Sekundärstrukturelemente wichtig sind. Nahezu ein Drittel der durch ARIA bestimmten eindeutigen 438 Distanzlimits sind weitreichende (*long-range*) Abstandsbedingungen. Diese sind maßgeblich entscheidend für die Bestimmung des Faltungstyps der Proteindomäne (siehe Tabelle 5.3).

## 4.4 Die NMR-Struktur der humanen p47 SEP-Domäne

### 4.4.1 Strukturbeschreibung

Die Struktur der humanen p47 SEP-Domäne, ermittelt an dem Konstrukt G1-S2-p47(171-270), besteht aus einem gemischten  $\beta$ -Faltblatt, das aus vier  $\beta$ -Strängen (Reste V180-W187 ( $\beta$ 1), F191-L193 ( $\beta$ 2), L198-S200 ( $\beta$ 3), V231-D237 ( $\beta$ 4)) gebildet wird, sowie aus zwei Helices  $\alpha$ 1 (Reste A207-R215) und  $\alpha$ 2 (Reste A220-R224). In der Abbildung 4.13 ist in (A) das NMR-Ensemble der 10 energieniedrigsten Strukturen aus insgesamt 200 berechneten und in (B) die Bänderdarstellung der energieniedrigsten Struktur zu sehen.



**Abbildung 4.13:** Stereodarstellungen der finalen Strukturen der humanen p47 SEP-Domäne (gezeigte Reste: D179-V244). (A) NMR-Ensemble der 10 energieniedrigsten Strukturen des Proteinrückgrates (B) Bänderdarstellung der energieniedrigsten Struktur

Eine Seite des  $\beta$ -Faltblatts liegt gegenüber den beiden Helices  $\alpha$ 1 und  $\alpha$ 2. Letztere Helix befindet sich direkt gegenüber der Strängen  $\beta$ 1 und  $\beta$ 4 am unteren Rand der Domäne und ist rechtwinklig versetzt in Bezug auf  $\alpha$ 1 angeordnet. Die beiden Loops zwischen  $\beta$ 1- $\beta$ 2 und  $\beta$ 3- $\alpha$ 1 nehmen einen kleinen Konformationsraum ein, was sich auch in einem kleinen *RMS* (*root mean square*)-Abweichungswert niederschlägt (vgl. Abbildung 4.14). Im

Gegensatz zu diesen beiden rigiden Loops befindet sich zwischen  $\alpha 2$  und  $\beta 4$  ein weniger gut charakterisierter Loop, der einen größeren Konformationsbereich einnimmt.

Der hydrophobe Kern der Domäne setzt sich aus den Resten in  $\beta 1$  (V180, L184, L186, W187),  $\beta 2$  (F191, L193), dem Rest L198 in  $\beta 3$  und den Resten von  $\alpha 1$  (F209, L210, I213) und  $\alpha 2$  (L223, L225) sowie von  $\beta 4$  (V231, L233, M235) zusammen. Die Reste Y201 und F243 tragen ergänzend zum hydrophoben Kern der SEP-Domäne bei. Der obere Teil der Abbildung 4.14 veranschaulicht die Verteilung der durch ARIA zugeordneten Distanzlimits auf die einzelnen Reste. Im Durchschnitt kommen auf jeden Rest 14,4 Distanzlimits für den Bereich D179-P246. Die aromatischen und hydrophoben Reste W187, F191, L198, Y201, F209, L210, I213, P219 und F243 tragen durch sehr viele *longrange* NOEs (schwarz eingerahmt) zur Definition der Struktur bei. Diese Reste mit Ausnahme von P219, sind Teil des hydrophoben Kernes, der strukturell gut definiert ist. Im Bereich von  $\alpha 1$ ,  $\alpha 2$  sind typischerweise viele *medium-range* NOEs (weiß eingerahmt) für die Reste F209, L210, P219 und L222 zu finden. Vor D179 und nach P246 wurden nur sehr wenige NOEs gefunden. Dies erklärt die schlecht definierte Struktur der N- und C-Termini des Ensembles.

Die *RMS*-Abweichung für die Reste D179-V244 beträgt  $1,04 \text{ \AA} \pm 0,24$  für das Peptidrückgrat. Die *RMS*-Werte des Ensembles, die zur gemittelten Struktur über die Proteinsequenz verteilt sind, sind im unterem Teil der Abbildung 4.14 aufgeführt. Es fällt auf, dass die *RMS*-Werte zwischen D179 bis P246 (durch gefundene NOE gut definierter Bereich) kleiner als  $1 \text{ \AA}$  sind, während N-Terminus und C-Terminus größere Abweichungen zeigen. Lediglich die Reste H227-G229 des  $\alpha 2$ - $\beta 4$  Loops unterliegen einer *RMS*-Abweichung zwischen  $1-2 \text{ \AA}$ . Die in Procheck-NMR (Laskowski et al., 1996) durchgeführte Ramachandran-Plot-Analyse zeigte, dass die  $\phi$ - und  $\psi$ -Winkel 96 % aller Aminosäuren sich innerhalb des favorisierten (*most favoured region*) und zusätzlich erlaubten (*additionally allowed region*) Bereichs befinden. Der erlaubten (*generously allowed*) Region wurden 2 % der Winkel zugeordnet. Lediglich 1,8 % der Torsionswinkel wurden dem nicht erlaubten Bereich (*disallowed region*) zugewiesen. Dies sind die Torsionswinkel der N- und C-terminalen Reste D179, H238 und D242.

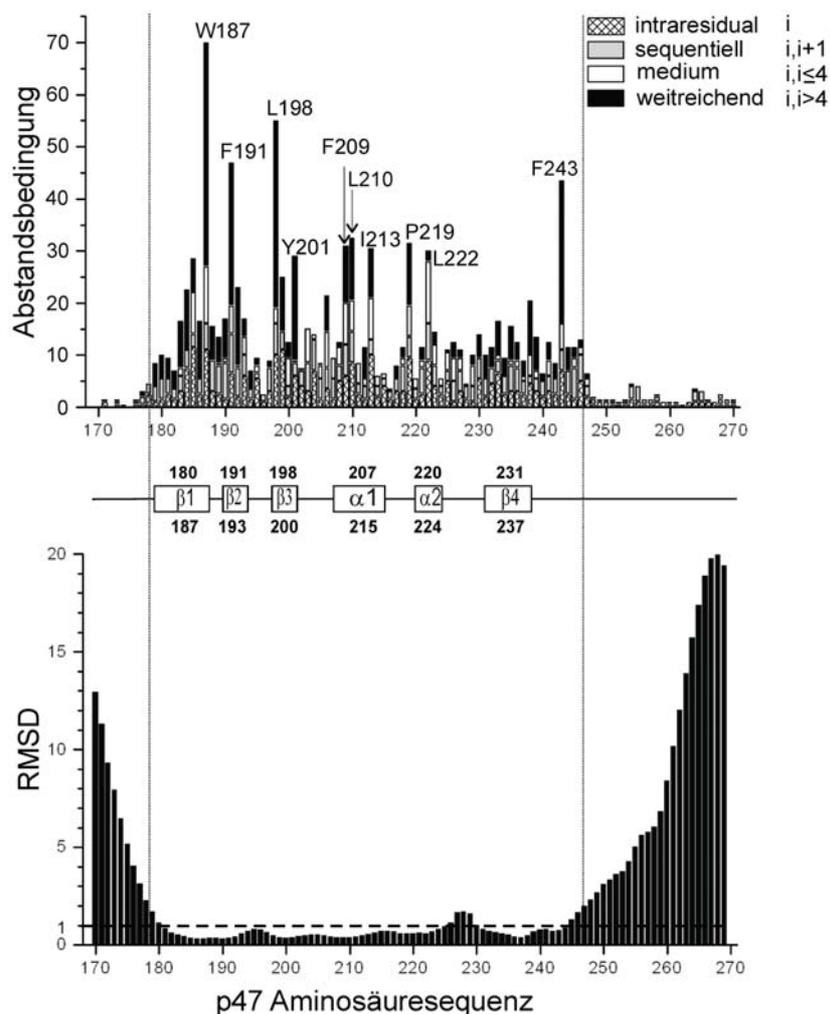
Eine DALI-Datenbank Suche (Holm & Sander, 1995) ergab für die Struktur der humanen p47 SEP-Domäne keine signifikant ähnliche Struktur (Stand:März 2004). Dies lässt den Schluss zu, dass es sich bei der SEP-Domäne um einen *novel fold* handelt. Die Raumkoordinaten der p47 SEP-Domäne sind in der PDB-Datenbank unter der Zugangsnummer 1SS6 abgelegt.

**Tabelle 3.3** Statistik des berechneten finalen NMR-Ensembles (20 Strukturen)

<b>ARIA-Zuordnung für Abstandsbedingungen (NOE restraints)</b>		
insgesamt gefundene eindeutige Zuordnungen		1570
intraresidual	(i)	637
sequentiell	( <i>sequentiel</i> ) (i, i+1)	330
mittellang	( <i>medium-range</i> ) (i, i≤4)	165
weitreichend	( <i>long-range</i> ) (i, i>4)	438
mehreutig		567
manuell		113
<b>weitere verwendete Strukturbedingungen</b>		
Wasserstoffbrückenbindungen (aus NOE-Mustern)		18
Torsionswinkel (aus $^3J_{\text{HNH}\alpha}$ -Experimenten und NOE-Mustern)		19
<b>Strukturverletzungen</b>		
Abstände > 0.4 Å		0
Torsionswinkel > 5°		0
<b>Mittlere Abweichung (RMSD) von experimentellen Strukturbedingungen</b>		
NOE (Å)		0,2 ± 0,03
Torsionswinkel (°)		4,3 ± 0,4
<b>CNS Potentialenergien (kcal/mol)</b>		
$E_{\text{total}}$		867 ± 11
$E_{\text{bonds}}$		8 ± 1
$E_{\text{angles}}$		65 ± 4
$E_{\text{impropers}}$		11 ± 1
$E_{\text{dihedral}}$		503 ± 35
$E_{\text{vdW}}$		105 ± 3
$E_{\text{NOE}}$		172 ± 5
$E_{\text{cdih}}$		1 ± 0,5
<b>Mittlere Abweichung (RMSD<sup>1</sup>) NMR-Ensemble zu gemittelter Struktur (Å)</b>		
Peptidrückgrat (N; C $\alpha$ ; C')		1,04 ± 0,24
Schweratome		1,75 ± 0,27
<b>Ramachandran Plot<sup>2</sup> (%)</b>		
favorisierte Bereiche ( <i>most favoured regions</i> )		80,9
zusätzlich erlaubte Bereiche ( <i>additionally allowed regions</i> )		15,2
erlaubte Bereiche ( <i>generously allowed regions</i> )		2,2
unerlaubte Bereiche ( <i>disallowed regions</i> )		1,8

<sup>1</sup> Die RMS-Abweichung des Ensembles wurde mit dem Programm Molmol für die Reste D179-V244 bestimmt

<sup>2</sup> Die Ramachandran-Plot Analyse erfolgte mit dem Programm Procheck-NMR für die Reste D179-V244

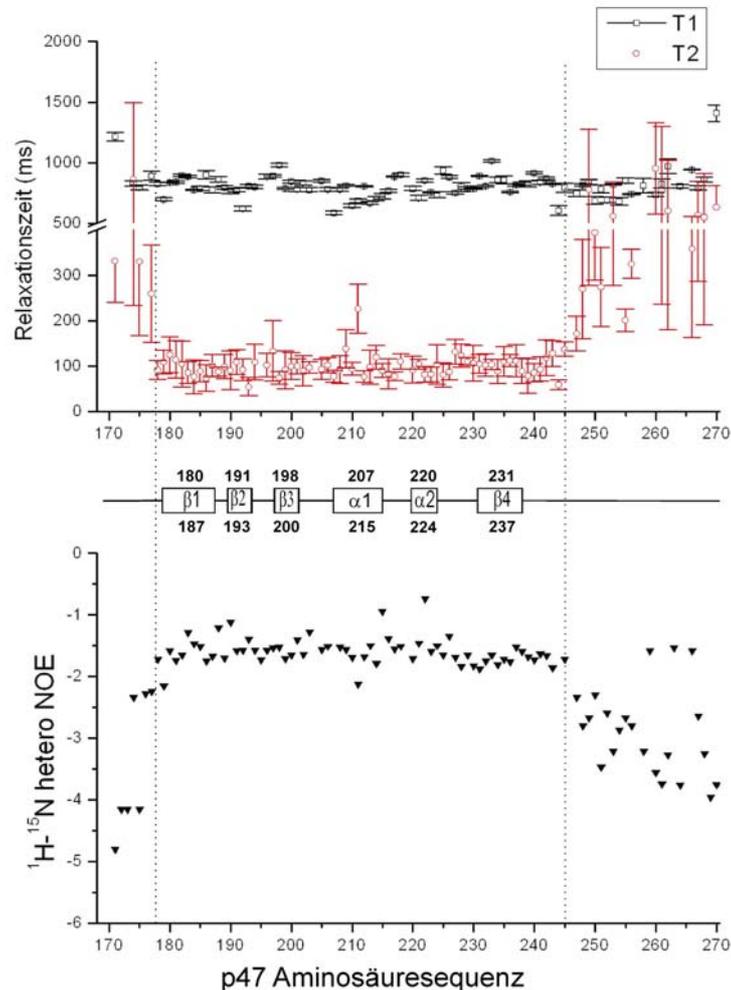


**Abbildung 4.14:** Strukturgüte der SEP-Domäne. (oben) Verteilung der durch die Strukturrechnung ARIA/CNS gefundenen Abstandsbedingungen pro Aminosäurerest. (Mitte) Sekundärstrukturelemente der p47 SEP-Domäne mit den gefundenen Domänengrenzen (gestrichelte vertikale Linien). (unten) RMS-Abweichung (Å) des Ensembles zur gemittelten Struktur des finalen Ensembles

#### 4.4.2 Dynamikuntersuchungen am G1-S2-p47 (171-270) Konstrukt

Die niedrige Anzahl an ermittelten Distanzlimits und die hohe RMS-Abweichung für den N- und C-Terminus wirft die Frage auf, ob es sich tatsächlich um flexible oder um wenig gut definierte Bereiche aufgrund mangelnder experimenteller Daten handelt. Zur Klärung dieser Frage wurden Untersuchungen zur Dynamik durchgeführt. Dazu wurden  $^{15}\text{N}$  Relaxationszeiten ( $T_1, T_2$ ) und heteronukleare  $^1\text{H}$ - $^{15}\text{N}$  NOE gemessen (Farrow et al., 1994). Die Ergebnisse dieser Dynamikmessungen sind in der Abbildung 4.15 zusammengefasst. Dabei zeigt sich, dass die Relaxationszeit  $T_2$  für die Amid-Stickstoffe der Reste D179 und K245 stark ansteigt. Die Werte für  $T_1$  bleiben nahezu konstant. Ein

Maß für die Flexibilität ist das Verhältnis T1/T2. Dieses Verhältnis fällt vor D179 und nach K245 signifikant ab, was eine erhöhte Flexibilität der N- und C-Termini anzeigt.

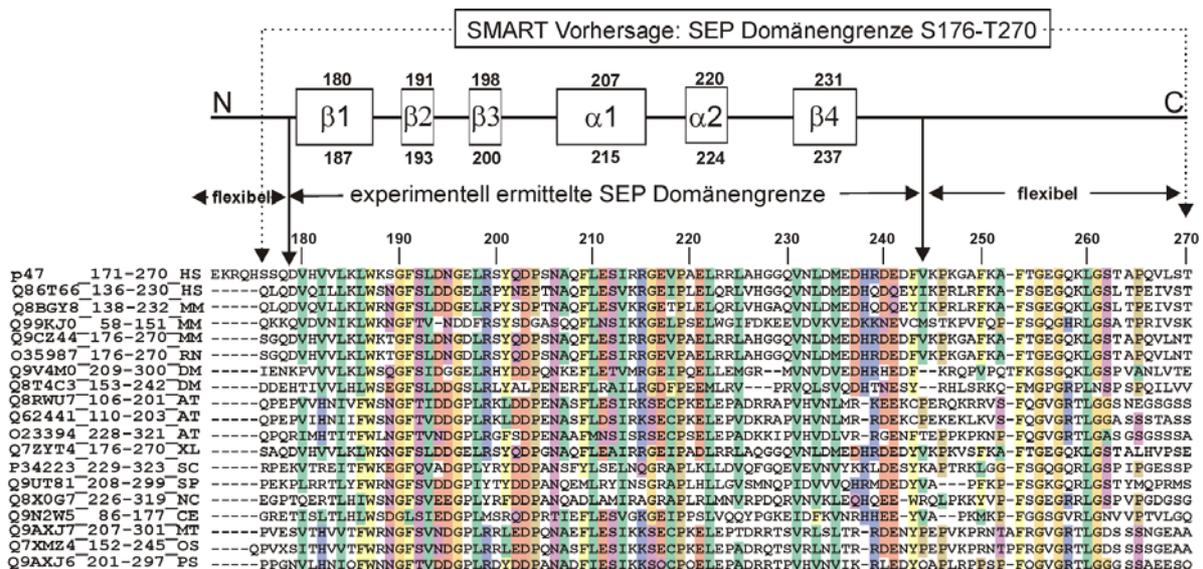


**Abbildung 4.15:** Dynamikmessungen der SEP-Domäne. (oben) T1 (schwarz) - und T2 (rot)-Relaxationszeiten. (Mitte) Sekundärstrukturelemente der p47 SEP Domäne mit den gefundenen Grenzen der Dömäne (gestrichelte Linien). (unten)  $^1\text{H}$ - $^{15}\text{N}$  NOE ( $\blacktriangledown$ ).

Diese Tatsache wird durch das Auftreten von stärker negativen  $^1\text{H}$ - $^{15}\text{N}$  hetero NOE vor D179 und nach K245 im Vergleich zu Resten innerhalb des genannten Bereiches bestätigt. Die Korrelationszeit  $\tau_c$  beschreibt die statistische Bewegung des Gesamtmoleküls und liegt für monomere Proteine in der Größenordnung von Nanosekunden (Gryk et al., 1998). Aus den gemittelten Werten für das Verhältnis von T1/T2 läßt sich die Korrelationszeit  $\tau_c$  bestimmen. Für das Konstrukt G1-S2 p47(171-270) ergab sich damit ein  $\tau_c$  -Wert von 6.4 ns. Dieser Wert entspricht einem monomeren Zustand für ein globuläres Protein dieser Größe.

#### 4.4.3 Proteinsequenzalignment und experimentelle Domänengrenzen

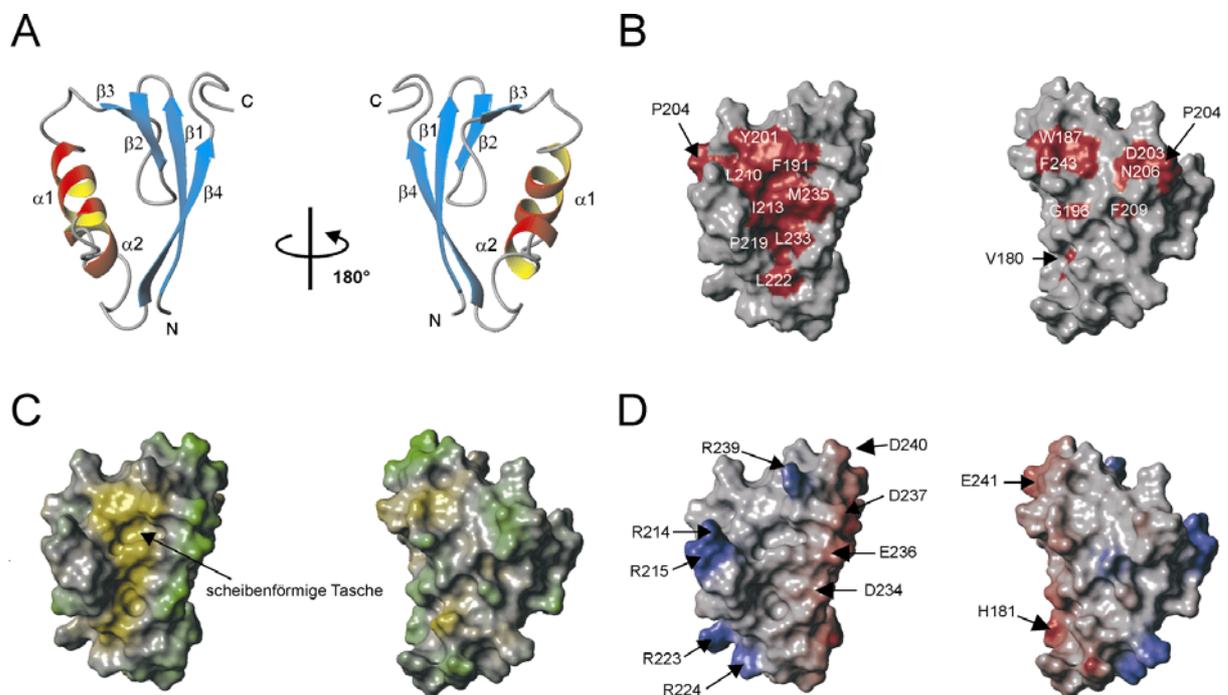
In der SMART Datenbank werden die Grenzen der SEP-Domäne für die Reste S176 und T270 vorhergesagt. Das multiple Sequenzalignment einer Auswahl dieser Proteine ist in Abbildung 4.16 dargestellt. Aufgrund der experimentellen Befunde der Dynamikmessungen, der Anzahl der zugeordneten Distanzlimits im strukturierten und unstrukturierten Bereich und der Befund, dass sämtliche Reste in den flexiblen N- und C-Termini einem schnellen H/D Austausch unterliegen (Daten nicht gezeigt), erlaubt eine neue Definition der SEP-Domänengrenze. Die SEP-Domäne ist zwischen dem N-terminalen Rest D179 bis zum C-terminalen Rest V244 strukturiert, während der Sequenzabschnitt K245-T270 unseres Konstruktes in Lösung flexibel ist.



**Abbildung 4.16:** Multiples Sequenzalignment von putativen SEP-Domänen (SwissProt Nummer für p47(171-270): Q9UNZ2). Die vorhergesagten und experimentell bestimmten Domänengrenzen für die p47 SEP-Domäne sind jeweils oberhalb des Sequenzalignments dargestellt. Die gefundenen Grenzen der Sekundärstrukturelemente für die  $\beta$ -Stränge ( $\beta$ 1-4) und  $\alpha$ -Helices ( $\alpha$ 1/2) sind ober- und unterhalb aufgeführt. Konservierte hydrophobe, aromatische, saure, basische und hydrophile Reste sind hellgrün, gelb, rot, blau und lila eingefärbt. Konservierte Proline und Glycine sind oliv und gelb gekennzeichnet. Verwendete Abkürzungen für die verschiedenen Spezies sind: HS (Homo sapiens), MM (*Mus musculus*), RN (*Rattus norvegicus*), DM (*Drosophila melanogaster*), AT (*Arabidopsis thaliana*), XL (*Xenopus laevis*), SC (*Saccharomyces cerevisiae*), SP (*Saccharomyces pombe*), NC (*Neurospora crassa*), CE (*Caenorhabditis elegans*), MT (*Medicago truncatula*), OS (*Oryza sativa*) und PS (*Pisum sativum*).

#### 4.4.4 Oberflächen der humanen p47 SEP-Domäne

Die Verteilung der konservierten hydrophoben Reste sowie der sehr hoch konservierten Reste (>80 %) D203, P204 und N206 auf der Oberfläche der humanen p47 SEP-Domäne zeigt Abbildung 4.17 (B). Dabei bedecken die Reste F191, Y201, P204, L210, I213, P219, L222, L233 und M235 eine umfangreiche zusammenhängende Region (links), während auf der anderen Seite zwei kleinere konservierte Bereiche durch die Aminosäure-Paare W187/F243 und D203/N206 auftreten. Die konservierten Reste V180, G196 und F209 bedecken einzelne kleine Regionen der Oberfläche und sind mit aufgeführt. (C) und (D) zeigen das berechnete lipophile/hydrophile und elektrostatische Potential der Domänenoberfläche. Die aufgeführten Oberflächeneigenschaften (B, C und D) sind in den gleichen Orientierungen gezeigt wie die Bänderdarstellung der Domäne in (A).



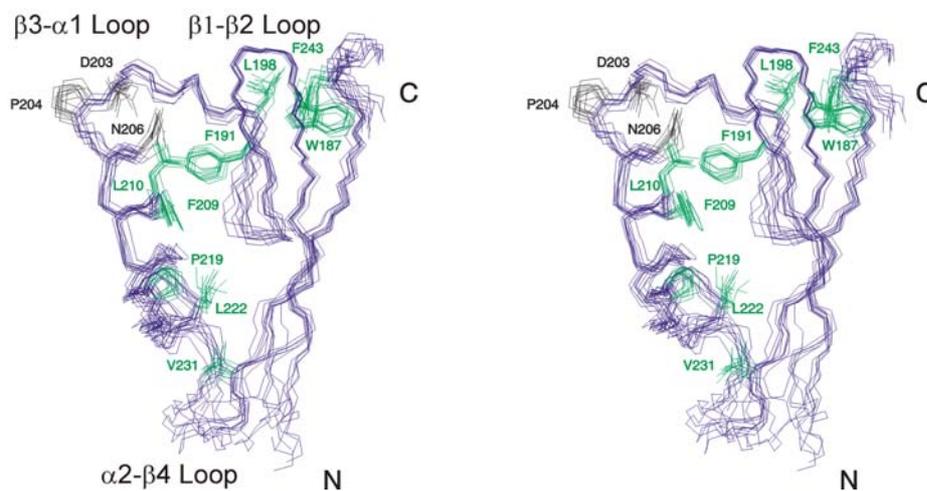
**Abbildung 4.17:** Oberflächencharakteristika der humanen p47 SEP-Domäne (A) Bänderdarstellung (B) konservierte hydrophobe Reste der Oberfläche (rot) sowie sehr hoch konservierte exponierte Reste DPxN des Loops β3-α1 (C) hydrophiles (grün) und lipophiles (gelb) Potential. (D) elektrostatisches Potential. Semikonservierte saure (rot) und basische (blau) Reste vermitteln einen Dipolcharakter der Domäne. Die Abbildungen (B, C und D) sind in den gleichen Orientierungen wie in (A) gezeigt.

Das kalkulierte lipophile Potential (Heiden et al., 1993) veranschaulicht hydrophile (grün) und hydrophobe (gelb) Bereiche. Die hydrophobe Region der Vorderseite weist scharfe Konturen und einige Kavitäten auf, die Seitenketten von potentiellen Interaktionspartnern aufnehmen können. Eine scheibenförmige Tasche, gebildet von den Resten L210, F191, I213 und M235, ist groß genug, um eine Tryptophan-Seitenkette aufzunehmen.

Die Berechnung des elektrostatischen Potentials für die freizugängliche Domänenoberfläche offenbart einen ausgeprägten Dipolcharakter. Dabei bilden die sauren Reste D234, E236 und D237 von  $\beta 4$  und die C-terminalen Reste D240 und E241 eine negativ geladene Fläche. Gegenüber erzeugen die basischen Reste R214, R215, R223, R224 und R239 ein positives Potential. Sowohl die basischen als auch die sauren Reste sind zwischen 40 und 60 % konserviert.

## 4.5 Entwicklung einer strukturbasierten Hypothese zur Funktion der p47 SEP-Domäne

Die detaillierte Strukturanalyse der humanen p47 SEP-Domäne eröffnet den Blick auf die auffällig „scharf“ definierten rigiden Loops  $\beta 1\text{-}\beta 2$  und  $\beta 3\text{-}\alpha 1$ . Diese Loops und der flexible, unstrukturierte C-Terminus befinden sich oberhalb einer vorstellbaren Ebene, die durch die hydrophoben konservierten Reste F209, L210, F191, L198, W187 und F243 gebildet wird. Abbildung 4.18 zeigt die Seitenketten der zu über 80 % konservierten Reste. Schwarz gekennzeichnet sind die zu 100 % konservierten Reste P204, D203 und N206 aus dem Loop  $\beta 3\text{-}\alpha 1$ , die sich oberhalb dieser Ebene befinden.



**Abbildung 4.18:** Stereodarstellung des Ensembles der 10 energieniedrigsten Strukturen der humanen p47 SEP-Domäne. In grün und schwarz sind die Seitenketten konservierter Reste (>80%) hervorgehoben. D203, P204 und N206 (schwarz) sind exponierte Reste des Loops  $\beta 3\text{-}\alpha 1$  oberhalb einer diagonalen Schnittebene, die durch die hydrophoben Reste F209, L210, F191, L198, W187 und F243 (grün) verläuft.

Die Konservierung und der exponierte Charakter dieser Reste sowie die Nähe des Loops  $\beta 1\text{-}\beta 2$  und des konservierten C-Terminus legt die Vermutung nahe, dass diese drei Strukturelemente Teil eines gemeinsamen Bindungsepitops sein könnten und aufgrund einer funktionellen Bedeutung zu über 80 % sind. Im Gegensatz dazu befinden sich die konservierten Reste P219, L222 und V231 unterhalb dieser Ebene und im Bereich des strukturell weniger „scharf“ definierten Loops  $\alpha 2\text{-}\beta 4$ . Die Konservierung dieser Reste deutet daraufhin, dass sie für die Stabilisierung der Domäne von großer Bedeutung sind.



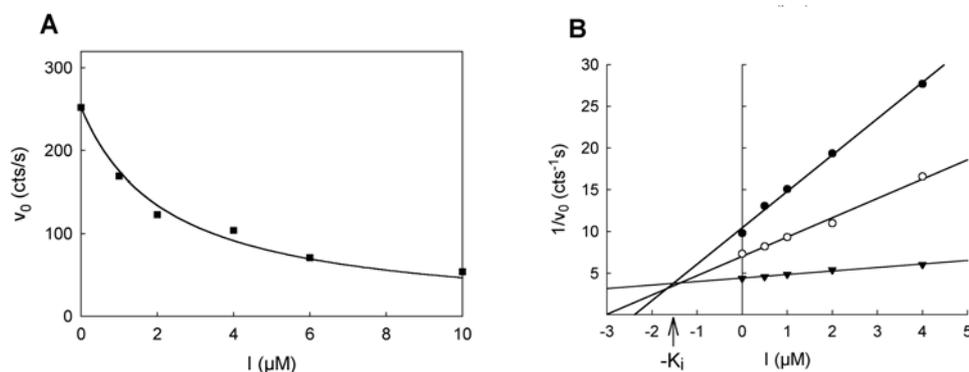
beiden Loops 1 und 2 des Cystatins zusammen zur spezifischen Inaktivierung der Proteaseaktivität bei.

Im direkten Vergleich fällt auf, dass die humane p47 SEP-Domäne eine analoge Topologie (s. Abbildung 4.19) besitzt. Die durch konservierte Reste charakterisierten, Loops  $\beta 1$ - $\beta 2$  und  $\beta 3$ - $\alpha 1$  entsprechen den Loops 2 und 1 des Cystatins und der C-Terminus nimmt die Rolle des N-Terminus ein.

Diese topologische Übereinstimmung der Strukturfragmente von Cystatin-Inhibitoren mit Strukturfragmenten der humanen p47 SEP-Domäne führte zu der Entwicklung der strukturbasierten Hypothese, dass die p47 SEP-Domäne ähnlich wie die Cystatine ein reversibler kompetitiver Inhibitor von Cysteinproteinasen ist. Dieser Vermutung wurde mit der auf diesem Tätigkeitsgebiet erfahrenen Arbeitsgruppe von Dr. Dorit K. Nägler (Ludwig-Maximilians-Universität München) nachgegangen.

#### **4.5.2 Bestimmung der Inhibition der Cathepsine L, B und X durch G1-S2 p47 (171-270) mittels eines Fluoreszenz-Aktivitätsassay**

Im Rahmen der Überprüfung dieser strukturbasierten Hypothese zur möglichen Interaktion und Hemmung von Cysteinproteasen durch die humane p47 SEP-Domäne wurden Inhibitionsstudien mit den drei rekombinanten humanen Cathepsinen L, B und X in Kollaboration mit Dr. Dorit K. Nägler am Institut für Klinische Chemie und Biochemie, der Ludwigs-Maximilians-Universität München, durchgeführt. Dabei zeigte sich in einem Fluoreszenzassay ((Nagler et al., 1997), siehe Abbildung 4.20), eine spezifische Hemmung der Aktivität von Cathepsin L unter Beobachtung der Hydrolyse des fluorogenen Substrates Cbz-FR-MCA (Carbobenzoxy-L-Phenylalanin-4-Methylcoumarinyl-7-amid) durch das Konstrukt G1-S2-p47(171-270). Die Inhibitionskonstante  $K_i$  wurde durch lineare Regression aus dem Schnittpunkt der Geraden des Dixon-Plots für drei unterschiedliche Substratkonzentrationen ermittelt. Für Cathepsin L ergab sich ein  $K_i$ -Wert von  $1,5 \pm 0,2 \mu\text{M}$ . Der Schnittpunkt der Geraden liegt oberhalb der X-Achse. Das bedeutet, dass das Konstrukt G1-S2-p47(171-270) als kompetitiver reversibler Inhibitor von Cathepsin L agiert.



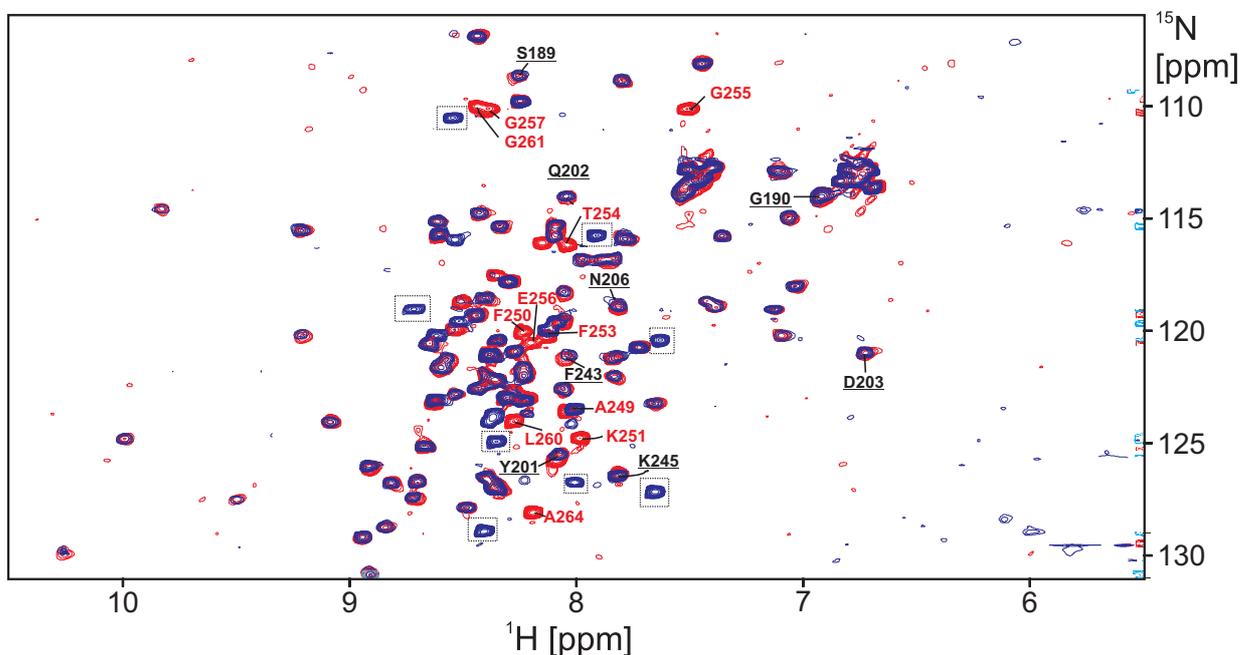
**Abbildung 4.20:** Reversible kompetitive Inhibition des humanen Cathepsin L durch G1-S2-p47(171-270) (I). (A) Abnahme der Reaktionsgeschwindigkeit  $v_0$  in Gegenwart der SEP-Domäne. (B) Dixon-Plots der Inhibition von Cathepsin L durch die SEP Domäne bei verschiedenen Substratkonzentrationen (Carbobenzoxy-L-Phenylalanin-4-Methylcoumarinyl-7-amid) 0,5  $\mu\text{M}$  ( $\bullet$ ), 1  $\mu\text{M}$  ( $\circ$ ) und 4  $\mu\text{M}$  ( $\blacktriangledown$ ).

Die Inhibitionsexperimente für Cathepsin B und X wurden in analoger Weise wie für Cathepsin L durchgeführt. Der  $K_i$ -Wert für Cathepsin B betrug 70  $\mu\text{M}$ , und für Cathepsin X wurde eine Inhibitionskonstante  $K_i > 100$   $\mu\text{M}$  ermittelt. Das heißt, Cathepsin B wurde wesentlich schwächer gehemmt als Cathepsin L, während für Cathepsin X keine signifikante Proteasehemmung durch die p47 SEP-Domäne gefunden werden konnte. Diese Versuche zeigten, dass die humane p47 SEP-Domäne als kompetitiver, reversibler Inhibitor mit der Inhibitionskonstante  $K_i=1,5$   $\mu\text{M}$  Cathepsin L verwandte Proteasen hemmt.

#### 4.5.3 NMR-Titration von deuteriertem $^{15}\text{N}$ markiertem G1-S2 p47(171-270) mit unmarkiertem Cathepsin L

Aufgrund der im Kapitel 4.5.2 festgestellten Tatsache der Inhibition der Proteaseaktivität von Cathepsin L durch die humane p47 SEP-Domäne schien es interessant die Hemmung mittels NMR-Spektroskopie näher zu charakterisieren. Dazu wurde versucht, die Bindungsstelle der SEP-Domäne zu bestimmen. Ausgehend von dem vollständig zugeordneten  $^1\text{H}$ - $^{15}\text{N}$  HSQC-Spektrum des SEP-Konstruktes ist es möglich, spezifische Aminosäurereste zu ermitteln, die bei Bindung eine Veränderung ihrer lokalen Umgebung erfahren. Das zeigt sich in einer veränderten chemischen Verschiebung der  $^1\text{H}$ - $^{15}\text{N}$  Kreuzsignale jener Reste. Für die SEP-Domäne alleine wurden zunächst vier  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektren bei 298 K, 293 K, 288 K und 278 K aufgenommen. Zunächst wurden die Signale der Aminosäuren bei allen Temperaturen neu zugeordnet durch Verfolgung und Vergleich der chemischen Verschiebung der aufgenommenen Spektren, die bei der Abkühlung von 298 K auf 278 K auftraten. Die Bestimmung wurde unter Verwendung

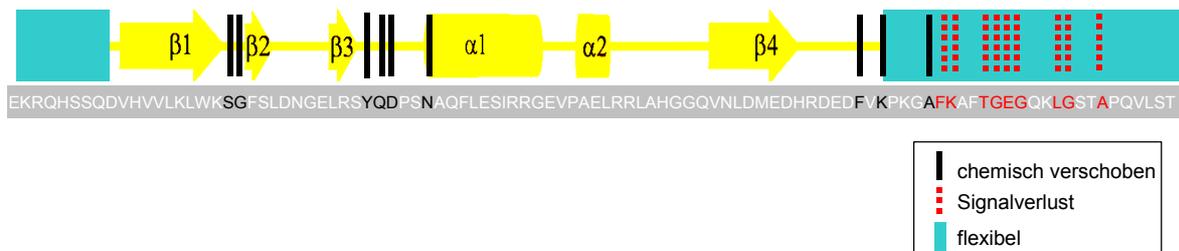
einer 60  $\mu\text{M}$   $^2\text{H}/^{15}\text{N}$ -markierten SEP Probe bei 278 K durchgeführt. Dabei wurde unmarkiertes Cathepsin L zur vorgelegten SEP Domänenprobe hinzugegeben. Die Endkonzentration für die Protease betrug 10  $\mu\text{M}$  und für die SEP-Domäne 50  $\mu\text{M}$ .



**Abbildung 4.21:** Überlagerung der  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektren von G1-S2 p47(171-270) ohne (rot) und mit Cathepsin L (blau). Das molare Verhältnis von Domäne zu Protease beträgt 5:1. Der Vergleich zeigt 9 Signale mit schwacher Änderung der chemischen Verschiebung (schwarz unterstrichen), 10 „Signalverluste“ (rote Schrift) und 8 „neu entstandene“ Signale (schwarz eingerahmt). Die Spektren wurden auf einem 600 MHz Spektrometer aufgenommen, ausgestattet mit einem Kryoprobekopf bei 278 K.

Daraus resultiert ein molares Verhältnis von 5:1. Sofort nach Zugabe der Protease erfolgte die Aufnahme des zweiten  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spektrums. In der Abbildung 4.21 ist die Überlagerung der beiden aufgenommenen Spektren ohne (rote Signale) und mit Cathepsin L (blaue Signale) dargestellt. Die Meßdauer für ein  $^1\text{H}$ - $^{15}\text{N}$ -HSQC Spektrum betrug 10 Minuten. Dabei zeigt sich, dass sich die Signale von 9 Aminosäurereste (in Abbildung 4.21 schwarz unterstrichen markiert) eine schwache chemische Verschiebung erfahren haben. Dies sind die Reste S189 und G190 des  $\beta 1$ - $\beta 2$  Loops sowie die im folgenden Loop  $\beta 3$ - $\alpha 1$  lokalisierten Reste Y201, Q202, D203 und N206. Eine Übersicht der resultierenden Veränderungen der chemischen Verschiebung durch die Titration mit der Protease Cathepsin L ist in der Abbildung 4.22 als Sekundärstrukturdiagramm und in der Sequenz wiedergegeben. Zusätzlich haben sich die Signale der Reste F243, K245 und A249, die sich in der Nähe befinden oder Teil des flexiblen C-Terminus sind, verschoben. Mit Rot beschriftet sind 10 Aminosäuren, deren Signale nach Zugabe der Protease verschwanden. Es sind die Reste F250, K251, F253, T254, G255, E256, G257,

L260, G261 und A264, die sich alle im flexiblen C-Terminus befinden. Das schwache Signal des Restes V244 konnte aufgrund der kurzen Meßzeit und der ungünstigen Wasserunterdrückung in beiden Spektren nicht detektiert werden. Weiterhin ist zu beobachten, dass 8 neue, nichtzuordnbare Signale erscheinen. Diese Signale sind in der Überlagerung (Abbildung 4.21) schwarz eingerahmt.



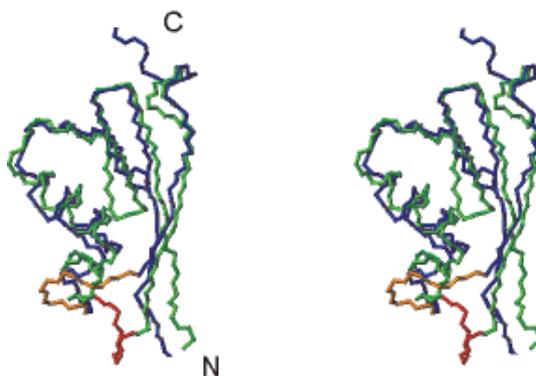
**Abbildung 4.22:** Sekundärstrukturdiagramm der p47 SEP-Domäne (171-270) mit den Veränderungen der  $^1\text{H}$ - $^{15}\text{N}$  chemischen Verschiebungen bei der Titration mit Cathepsin L. Mit schwarzen Balken sind die chemisch verschobenen Signale der Reste dargestellt. Rote gestrichelte Linien markieren Reste, deren Signale nach Zugabe der Protease verschwanden.

Zusammenfassend läßt sich sagen, dass aufgrund der experimentellen Befunde die p47 SEP-Domäne ein reversibler, kompetitiver Inhibitor der Protease Cathepsin L mit einer Inhibitionskonstante von  $1,5 \mu\text{M}$  ist. Die funktionelle Hypothese zur Interaktion der SEP-Domäne mit Cathepsin L wurde teilweise bestätigt, durch die Identifizierung vorhergesagter Reste der Bindungsseite.

## 4.6 Diskussion

### 4.6.1 Struktureller Vergleich der p47 SEP-Domänen aus *Rattus norvegicus* und *Homo sapiens*

Yuan et al. bestimmten letztes die NMR-Struktur der SEP-Domäne (PDB-Zugangsnummer: 1VAZ) des homologen Rattenproteins p47. Für diese SEP-Domäne wurde keine Funktion beschrieben (Yuan et al., 2004). Beide NMR-Strukturen zeigen die gleiche Faltung und weisen sehr ähnliche Domänengrenzen auf. In Abbildung 4.23 ist die Überlagerung der Proteinhauptkette der beiden Strukturen dargestellt. Die SEP-Domäne (Ratte) beginnt am gleichen Rest (D179), während ihre Domänengrenze am C-Terminus um die beiden Reste K245 und P246 im Vergleich zur humanen Domäne erweitert ist. Beide Strukturen unterscheiden sich jedoch geringfügig in der Orientierung des Loops  $\alpha 2$ - $\beta 4$ . Die abweichende Konformation dieses Loops in der hier bestimmten Struktur der humanen Domäne konnte durch manuell zugeordnete NOE-Signale zum  $\beta 4$ -Strang in dieser Arbeit verifiziert werden. Für die Reste D179-N195, L198-E217 und L233-E241 wurde eine *RMS*-Abweichung für die  $C\alpha$ -Atome von 1,74 Å für die beiden energieniedrigsten Strukturen festgestellt.



**Abbildung 4.23:** Stereodarstellung der überlagerten Proteinhauptkette der p47 SEP-Domänen aus *Rattus norvegicus* (blau) und *Homo sapiens* (grün). Der Loop  $\alpha 2$ - $\beta 4$  ist gelb (Ratte) und rot (Mensch) hervorgehoben.

Die Sequenz des Proteins aus der Ratte unterscheidet sich von der des humanen durch 3 Punktmutationen. Diese verursachen jedoch keine strukturellen Unterschiede. Die Punktmutation S198T zwischen  $\beta 1$  und  $\beta 2$  ist strukturell neutral, und die anderen beiden befinden sich im flexiblen N-Terminus. In einer von Yuan et al. (2004) durchgeführte DALI-Datenbank Suche (Holm & Sander, 1995) wurde keine signifikant ähnliche Struktur gefunden. Damit handelt es sich bei der p47 SEP-Domäne um einen „*novel fold*“.

#### **4.6.2 Welche Bedeutung hat der konservierte Sequenzabschnitt (250-262) für das p47 Protein ?**

Dynamikuntersuchungen und Linienbreitenanalysen an dem längerem p47(171-370) Konstrukt zeigten, dass die Reste der Linkersequenz (248-289) zwischen der SEP- und UBX-Domäne in Lösung flexibel und unstrukturiert sind (Yuan et al., 2004). Es konnte auch keine Interaktion zwischen den beiden Domänen festgestellt werden und demzufolge auch keine strukturelle Einschränkung des Linkers in seiner Flexibilität (Yuan et al., 2004). Auffällig ist jedoch der große Anteil konservierter Reste im C-terminalen Sequenzabschnitt 250-262 des Domänenlinkers (siehe dazu Abbildung 4.16, Seite 72). Dieser Linker könnte demnach eine funktionelle Eigenschaft aufweisen oder eine funktionelle Grenze für die SEP-Domäne darstellen.

#### **4.6.3 Analyse der Wechselwirkung zwischen der p47 SEP-Domäne und den Cathepsinen L,B und X**

In dieser Arbeit wurde eine Inhibition der Proteaseaktivität von Cathepsin L durch die p47 SEP-Domäne beschrieben. Dabei stellte sich heraus, dass die p47 SEP-Domäne ein kompetitiver reversibler Inhibitor mit einer Inhibitionskonstante  $K_i$  von 1,5  $\mu\text{M}$  für Cathepsin L ist. Dieser Wert entspricht einer schwachen Hemmung. Vorhergesagte Reste der Bindungsstelle konnten durch die durchgeführte NMR-Titration mit Cathepsin L bestätigt werden. Unklar bleibt der Mechanismus der Hemmung und inwiefern Reste des C-Terminus daran beteiligt sind. Aufgrund von Signalverlusten und der Beobachtung, dass sich einige Reste stark chemisch verschoben haben liegt der Schluß nahe, daß Teile des C-Terminus abgespalten wurden. Diese stark chemisch verschobenen Signale können von C-terminalen Peptidfragmenten der Domäne stammen. Das Phänomen der „substrattypischen Hemmung“ ist bislang bei der Hemmung von Papain durch Cystatinvarianten beobachtet worden. Dabei handelt es sich um eine temporäre Hemmung der Protease gefolgt von einer Spaltung des Inhibitors. Dieses Phänomen erklärt man sich durch die Konkurrenz des Enzyms um inhibitor- und substratähnliche Konformationen des Inhibitors. Dabei nähert sich die Peptidbindung des Inhibitors dem aktiven Cysteinrest soweit, daß sie spaltbar wird. Dies führt dann zu einer Verschiebung des Gleichgewichtes zu Gunsten des Substrat-Enzym-Komplexes und zum Verbrauch des Inhibitors (Nägler, 1996). Dies könnte die mögliche Abspaltung des C-Terminus der SEP-Domäne erklären.

Für die Cathepsine B und X wurde eine nur sehr schwache oder keine signifikante Hemmung ermittelt. Die geringe Affinität zu Cathepsin B kann dadurch erklärt werden, dass die Struktur von Cathepsin B eine Besonderheit aufweist. In Cathepsin B ist die *active site* von einem „occluding loop“ abgeschirmt, so dass Cystatin-ähnliche Inhibitoren eine geringere Affinität zeigen (Musil et al., 1991). Für Cathepsin X ergibt sich ein ähnlicher Fall. Vergleichbar dem „occluding loop“ in Cathepsin B zeigt die Kristallstruktur von Cathepsin B einen „Miniloop“ in der Nähe des aktiven Zentrums (Sivaraman et al., 2000). Diese sterische Einschränkung kann die Ursache für die sehr geringe Hemmung der Proteaseaktivität sein.

#### 4.7 Ausblick

In dieser Arbeit wurde eine Hypothese zur Funktion der isolierten p47 SEP-Domäne postuliert und in vitro Experimenten bestätigt. Es lässt sich aus der gefundenen Wechselwirkung der p47 SEP-Domäne mit Cathepsin L folgern, dass p47 ein spezifischer Adaptor für Cathepsin L sein kann und damit Teil eines bisher noch nicht beschriebenen zellulären Prozesses ist. Für die nähere Charakterisierung der Bindungsstelle zwischen Cathepsin L und der p47 SEP-Domäne ist es vorteilhaft, eine inaktive Protease ohne Restaktivität zu verwenden. Denkbar wäre es beispielsweise eine Cystein-Deletionsmutante herzustellen oder das aktive Cystein chemisch an der freien SH-Gruppe durch Carboxymethylierung zu modifizieren.

Es bleibt die Frage, ob das Protein p47 in voller Länge mit den drei Domänen UBA, SEP und UBX Cathepsine oder verwandte Proteasen inhibiert. Von weiterem Interesse ist es, ob sich die gefundenen Erkenntnisse in einer Untersuchung der Inhibition in vivo in zellbiologischen Versuchen nachvollziehen lassen.

#### 4.8 Referenzen

Bode, W., Engh, R., et al. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *Embo J* 7, 2593-2599

Bussow, K., Nordhoff, E., Lubbert, C., Lehrach, H. & Walter, G. (2000) A human cDNA library for high-throughput protein expression screening. *Genomics* 65, 1-8

- Cavanagh, J., Wayne, J., Fairbrother, J., Palmer, A.G., Skelton, N-J. (1996) *Protein NMR Spectroscopy Principles and Practice*, Academic Press
- Farrow, N. A., Muhandiram, R., et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry* 33, 5984-6003
- Gryk, M. R., Abseher, R., Simon, B., Nilges, M. & Oschkinat, H. (1998) Heteronuclear relaxation study of the PH domain of beta-spectrin: restriction of loop motions upon binding inositol trisphosphate. *J Mol Biol* 280, 879-896
- Heiden, W., Moeckel, G. & Brickmann, J. (1993) A new approach to analysis and display of local lipophilicity/hydrophilicity mapped on molecular surfaces. *J Comput Aided Mol Des* 7, 503-514
- Holm, L. & Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem Sci* 20, 478-480
- Koradi, R., Billeter, M. & Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *Journal of Molecular Graphics* 14, 51-&
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular Nmr* 8, 477-486
- Letunic, I., Copley, R. R., et al. (2004) SMART 4.0: towards genomic data integration. *Nucleic Acids Res* 32 Database issue, D142-144
- Machleidt, W., Thiele, U., Assfalg-Machleidt, I., Forger, D. & Auerswald, E. A. (1991) Molecular mechanism of inhibition of cysteine proteinases by their protein inhibitors: kinetic studies with natural and recombinant variants of cystatins and stefins. *Biomed Biochim Acta* 50, 613-620
- Musil, D., Zucic, D., et al. (1991) The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *Embo J* 10, 2321-2330
- Nägler, D. K. (1996) Herstellung und Charakterisierung von rekombinanten Hühnereiweiß-Cystatinvarianten zur Aufklärung von Struktur-Funktionsbeziehungen. In Medizinische Fakultät, Ludwig-Maximilians-Universität München, München
- Nägler, D. K. (1996) Herstellung und Charakterisierung von rekombinanten Hühnereiweiß-Cystatinvarianten zur Aufklärung von Struktur-Funktionsbeziehungen. In Medizinische Fakultät, Ludwigs-Maximilians-Universität, München
- Nagler, D. K., Storer, A. C., Portaro, F. C., Carmona, E., Juliano, L. & Menard, R. (1997) Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts. *Biochemistry* 36, 12608-12615
- Schmieder, P., Leidert, M., Kelly, M. & Oschkinat, H. (1998) Multiplicity-selective coherence transfer steps for the design of amino acid-selective experiments - A

- triple-resonance experiment selective for ash and gin. *Journal of Magnetic Resonance* 131, 199-202
- Schubert, M., Ball, L. J., Oschkinat, H. & Schmieder, P. (2000) Bridging the gap: A set of selective H-1-N-15-correlations to link sequential neighbors of prolines. *Journal of Biomolecular Nmr* 17, 331-335
- Schubert, M., Oschkinat, H. & Schmieder, P. (2001) Amino acid type-selective backbone H-1-N-15-correlations for Arg and Lys. *Journal of Biomolecular Nmr* 20, 379-384
- Schubert, M., Oschkinat, H. & Schmieder, P. (2001) MUSIC and aromatic residues: Amino acid type-selective H-1-N-15 correlations, III. *Journal of Magnetic Resonance* 153, 186-192
- Schubert, M., Oschkinat, H. & Schmieder, P. (2001) MUSIC, selective pulses, and tuned delays: Amino acid type-selective H-1-N-15 correlations, II. *Journal of Magnetic Resonance* 148, 61-72
- Schubert, M., Smalla, M., Schmieder, P. & Oschkinat, H. (1999) MUSIC in triple-resonance experiments: Amino acid type-selective H-1-N-15 correlations. *Journal of Magnetic Resonance* 141, 34-43
- Seavey, B. R., Farr, E. A., Westler, W. M. & Markley, J. L. (1991) A relational database for sequence-specific protein NMR data. *J Biomol NMR* 1, 217-236
- Sivaraman, J., Nagler, D. K., Zhang, R., Menard, R. & Cygler, M. (2000) Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine. *J Mol Biol* 295, 939-951
- Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. & Turk, V. (1990) The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *Embo J* 9, 1939-1947
- Turk, D. & Guncar, G. (2003) Lysosomal cysteine proteases (cathepsins): promising drug targets. *Acta Crystallogr D Biol Crystallogr* 59, 203-213
- Wishart, D. S. & Sykes, B. D. (1994) The <sup>13</sup>C chemical-shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J Biomol NMR* 4, 171-180
- Yuan, X., Simpson, P., et al. (2004) Complete backbone resonance assignments of p47: the 41kDa adaptor protein of the AAA ATPase p97. *J Biomol NMR* 28, 309-310
- Yuan, X., Simpson, P., et al. (2004) Structure, dynamics and interactions of p47, a major adaptor of the AAA ATPase, p97. *Embo J* 23, 1463-1473