

Multidimensionale heteronukleare
NMR-Untersuchungen an der
MAR-bindenden Domäne von
cMeCP2



Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)
der naturwissenschaftlichen Fakultät III
- Biologie und Vorklinische Medizin
der Universität Regensburg

vorgelegt von

Björn Heitmann

Regensburg 11/2002

Promotionsgesuch eingereicht am: 21.11.2002

Die Arbeit wurde angeleitet von: Prof. Dr. Eike Brunner

Prüfungsausschuss:

Vorsitzender: Prof. Dr. Schmitt

Gutachter: Prof. Dr. Brunner

Prof. Dr. Dr. Kalbitzer

Drittprüfer: Prof. Dr. Kurtz

Inhaltsverzeichnis

1	Biologische und medizinische Bedeutung des Proteins cMeCP2 ARBD	1
1.1	Methylierte DNA	2
1.2	Das Rett-Syndrom	3
1.3	Das Protein	4
2	Materialien und verwendete NMR-Methoden	10
2.1	Bedeutung der NMR-Spektroskopie	10
2.2	Methodische Vorarbeiten	13
2.2.1	Überlegungen zur Linienbreite in HSQC- bzw. TROSY-Experimenten	14
2.2.1.1	Die ^{15}N -Linienbreite	16
2.2.1.2	Die $^1\text{H}_N$ -Linienbreite	20
2.2.1.3	Weiterführende Überlegungen zur Linienbreite	22
2.2.2	Deuterierung	25
2.3	Probenpräparation	29
2.4	Verwendete NMR-Experimente	30
2.4.1	Spektrenaufnahme und -auswertung	30
2.4.2	Zuordnung der chemischen Verschiebungen	31
2.4.2.1	Experimente zur Zuordnung der Signale der Kerne der Hauptkette	33

2.4.2.2	Experimente zur Zuordnung der Signale der Kerne der Seitenkette	36
2.4.3	Experimente zur Bestimmung von Kern-Kern-Abständen	37
2.4.3.1	NOESY-Experimente	37
2.4.3.2	Experimente zum Nachweis von Wasserstoffbrückenbindungen	38
2.4.4	Experimente zur Messung von Restdipolkopplungen	39
2.5	Berechnung der Proteinstruktur	44
2.5.1	Strukturberechnungen mit dem Computerprogramm CNS	45
2.5.2	Strukturberechnungen mit dem Computerprogramm DYANA	46
2.6	Rückrechnung von NOESY-Spektren	47
3	Zusammenfassung der Ergebnisse	49
3.1	Strukturbestimmung	49
3.1.1	Die sequentielle Zuordnung der Hauptkettenatome	50
3.1.2	Zuordnung der Seitenkettenatome	53
3.1.3	Sekundärstrukturvorhersagen	54
3.1.4	Diederwinkel	58
3.1.5	Restdipolkopplungen	61
3.1.6	Zuordnung der NOESY-Spektren	65
3.1.7	Numerische Berechnung der dreidimensionalen Struktur von cMeCP2 ARBD	69
3.2	Der Protein-DNA-Komplex	77
3.2.1	Methylierte DNA	78
3.2.2	Nicht methylierte DNA	83
4	Diskussion	87

4.1	Struktur des Proteins	87
4.2	Der Komplex aus Protein und DNA	91
	Zusammenfassung	96
	Literaturverzeichnis	97
	Anhang	102
A	Die chemischen Verschiebungen von cMeCP2 ARBD	103
B	<i>restraints</i>: Diederwinkel und Restdipolkopplungen	112
B.1	Diederwinkel als <i>restraints</i> bestimmt mit TALOS	112
B.2	Restdipolkopplungen als <i>restraints</i>	113
B.2.1	$^1\text{H}_N$ - ^{15}N -Restdipolkopplungen	113
B.2.2	$^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ -Restdipolkopplungen	114
B.2.3	^{15}N - $^{13}\text{C}'$ -Restdipolkopplungen	115

Kapitel 1

Biologische und medizinische Bedeutung des Proteins cMeCP2 ARBD

Die schnelle Entwicklung automatisierter Sequenzierungsverfahren ermöglichte die vollständige Entschlüsselung des über 3×10^9 Basenpaare umfassenden menschlichen Genoms (*human genome project*). Eine der Herausforderungen der nächsten Jahrzehnte wird es sein, neben der funktionellen Charakterisierung des durch die DNA codierten Proteins ein Verständnis der molekularen Mechanismen zu erlangen, welche die selektive Expression einzelner Gene steuern. Eng verknüpft mit dieser Fragestellung ist beispielsweise die enzymatisch gesteuerte Methylierung von DNA. Die Relevanz dieses Prozesses läßt sich an Fehlfunktionen des Methylierungsmechanismus erkennen, welche zu neurologischen Erkrankungen führen können. Eine dieser Erkrankungen ist das Rett-Syndrom [Ret66].

Zur biologischen und medizinischen Einordnung des in der vorliegenden Arbeit untersuchten Proteins wird in Abschnitt 1.1 eine kurze Einführung in das Vorkommen methylierter DNA und die involvierten Mechanismen gegeben, gefolgt von Abschnitt 1.2, in dem auf das Rett-Syndrom eingegangen wird. Im Anschluß (Abschnitt 1.3) erfolgt eine nähere Charakterisierung des untersuchten Proteins als Teil einer Familie funktionell ähnlicher Proteine.

1.1 Methylierte DNA

Methylierung von DNA ist ein weit verbreitetes Phänomen, unter anderem in Säugetierzellen. Sie tritt fast ausschließlich an der Position 5 von Cytosinresten auf, die sich in CpG-Dinucleotiden befinden.¹ Etwa 70% aller CpG-Motive sind methyliert, wobei weder die Verteilung der CpG-Dinucleotide noch das Vorliegen von methylierten Cytosinresten in der DNA zufällig ist [Coo89]. Auf die gesamte DNA bezogen, sind CpG-Motive unterrepräsentiert², was möglicherweise damit in Verbindung steht, daß sie besonders anfällig für Punktmutationen sind. CpG-Motive werden deshalb auch als *hotspots* bezeichnet. Wird 5-Methylcytosin desaminiert (spontaner Verlust einer Aminogruppe), so entsteht Thymin, eine in DNA natürlich vorkommende Base, die von Reparatursystemen im allgemeinen nicht erkannt wird. Bei Desaminierung von (nichtmethyliertem) Cytosin entsteht dagegen Uracil (kommt nicht natürlich in DNA vor), welches von Reparatursystemen gefunden und ersetzt werden kann [Leh94].

Bei der Methylierung von DNA interagieren drei voneinander unabhängige DNA-Methyltransferasen (DNMT1, DNMT3a, DNMT3b). Es wurde gezeigt, daß DNMT1 für die Aufrechterhaltung des DNA-Methylierungsmusters während der DNA-Replikation verantwortlich ist. Im Gegensatz dazu besteht die Funktion der anderen beiden DNA-Methyltransferasen unter anderem in der Methylierung vorher nicht methylierter Cytosinreste [Bes88].

Die Methylierung von DNA spielt eine wichtige Rolle bei der Aufteilung des Genoms in funktionell aktive und inaktive Bereiche; sie werden Euchromatin bzw. Heterochromatin genannt. Heterochromatin ist im Vergleich zu Euchromatin viel dichter gepackt. Diese kompakten Bereiche, wie man sie z.B. beim inaktivierten weiblichen X-Chromosom oder bei parasitärer DNA findet, sind transkriptionell stillgelegt [Bes00]. Defekte Methylierungsmechanismen, hervorgerufen durch Mutationen in den Methyltransferase-Enzymen führen bei Mäusen zu letalen Entwicklungsstörungen, wobei der Tod kurz vor oder nach der Geburt eintritt [Oka99], [Li92].

¹Dabei steht C für eine Cytosinnucleosid, p für die Phosphodiesterbrücke und G für ein Guaninnucleosid.

²Etwa 3×10^7 methylierte Cytosinreste, größtenteils in CpG-Motiven, liegen im Säugetiergenom vor.

Die Organisation des Chromatins und die Hemmung der Transkription erfolgen enzymatisch. Defekte in den involvierten Proteinen können zu Störungen des gesamten Organismus führen. Eine Krankheit, bei der dieser Zusammenhang gefunden wurde, ist das Rett-Syndrom.

1.2 Das Rett-Syndrom

Das von Andreas Rett zuerst beschriebene und nach ihm benannte Rett-Syndrom (MIM 312750) ist nach dem Down-Syndrom die häufigste Ursache für geistige Fehlentwicklungen bei Frauen. Typische Symptome sind unter anderem Demenz, Autismus, Abnahme des Schädelwachstums, Störungen der psychomotorischen Entwicklung sowie stereotype Bewegungen. Der Krankheitsverlauf läßt sich in vier zum Teil ineinander übergehende Phasen einteilen [Hag83], [Ret66]:

- *6.-18. Lebensmonat* (Verlangsamungsstadium) Desinteresse, verlangsamtes Kopfwachstum
- *1.-4. Lebensjahr* (schnelles Destruktiv-Stadium) rasante Entwicklungsregression
- *2.-10. Lebensjahr* (pseudostationäre Phase) Verminderung autistischer Züge, Kommunikationsfähigkeit verbessert sich
- *10. Lebensjahr bis zum Tod* (spätes motorisches Verschlechterungsstadium)

Die Identifizierung des defekten Gens (MECP2, welches das *methyl-CpG binding protein 2* codiert), das mit dem Rett-Syndrom im Zusammenhang steht, war schwierig, da die Krankheit nur sporadisch auftritt. Als Lage des Gens konnte das X-Chromosom ausgemacht werden, worin auch die Problematik liegt, vom Rett-Syndrom betroffene Patienten zu finden: Ist nämlich das MECP2-Gen mutiert und kein diesbezüglich intaktes homologes Chromosom vorhanden, wie bei männlichen Trägern der Erbkrankheit, so führt das Rett-Syndrom meist vor der Geburt zum Tod. Bei weiblichen Trägern der Krankheit ist eine phänotypische Ausprägung davon abhängig, welches der beiden X-Chromosomen in der Embryonalentwicklung abgeschaltet wurde. In [Lew92] und [Ami99] wurde MECP2

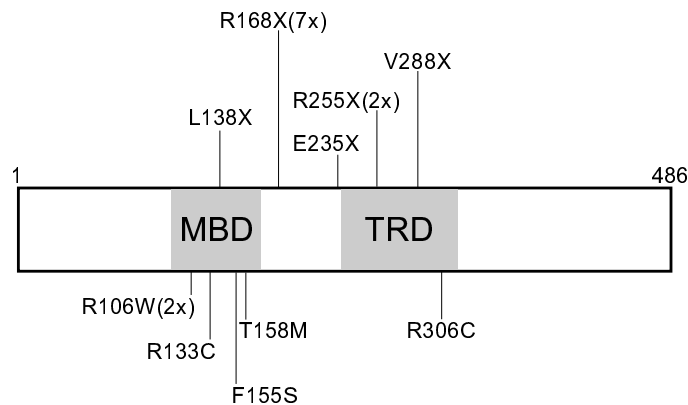


Abbildung 1.1: *Das Protein MeCP2*. Dargestellt ist die Sequenz mit den beiden funktionellen Domänen (siehe Abschnitt 1.3) und die Mutationen, die bei Patienten mit Rett-Syndrom gefunden wurden. In Klammern ist die Anzahl der Patienten vermerkt, bei denen diese Mutation auftrat, falls sie von eins abweichend war. Oberhalb der Sequenz sind die Mutationen eingezeichnet, die zur Verkürzung des Proteins geführt haben [Ami99], [Wan99].

identifiziert und der Zusammenhang zum Rett-Syndrom quantitativ untersucht. In Abbildung 1.1 ist die Sequenz von MeCP2 mit bei Rett-Patienten gefundenen Mutationen dargestellt.

1.3 Das Protein MeCP2

MeCP2 war das erste Protein aus der Familie der methylbindenden Proteine, dessen Bindung an methylierte CpG-Motive nachgewiesen werden konnte. Es ist ein häufig vorkommendes chromosomales Protein. Der Nachweis erfolgte durch Immunfluoreszenzmessungen an Maus- und Rattenzellen [Nan97]. MeCP2 besteht aus zwei funktionellen Domänen (siehe auch Abbildung 1.1). Die erste Domäne, MBD (*methyl binding domain*) genannt, ist für die Bindung des Proteins an methylierte CpG-Motive verantwortlich [Nan93]. Sie befindet sich im N-terminalen Bereich des Proteins und umfaßt etwa 80 Aminosäurereste. Die zweite, TRD (*transcriptional repression domain*), wird für die Transkriptionshemmung *in vivo* und *in vitro* benötigt [Nan97], [Jon98], [Kal00]. Diese Domäne umfaßt etwa 100 Aminosäurereste und befindet sich 75 Aminosäurereste hinter der MBD in der Mitte des Proteins. Durch Sequenzhomologiesuche in Datenbanken bezüglich der beiden funktionellen Domänen wurden andere Proteine gefunden, deren Wechselwirkung mit methylierter

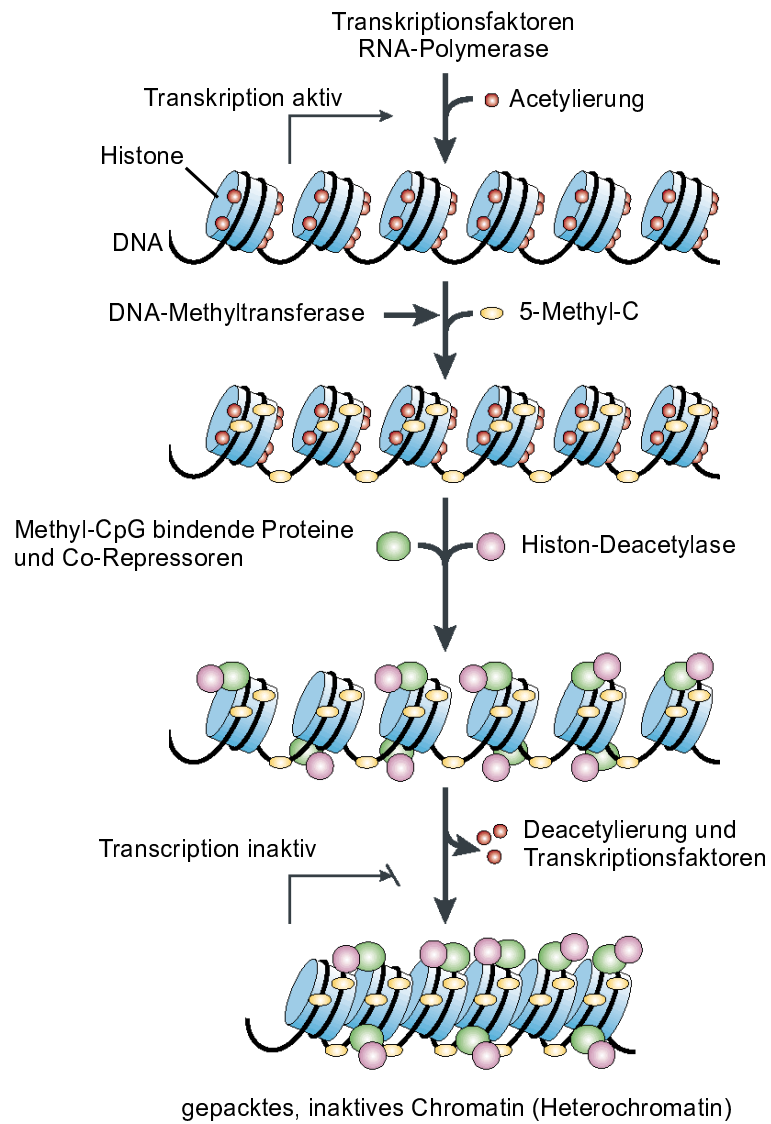


Abbildung 1.2: *Mechanismus der Transkriptionshemmung über DNA-Methylierung und Histon-Deacetylierung [Rob00].*

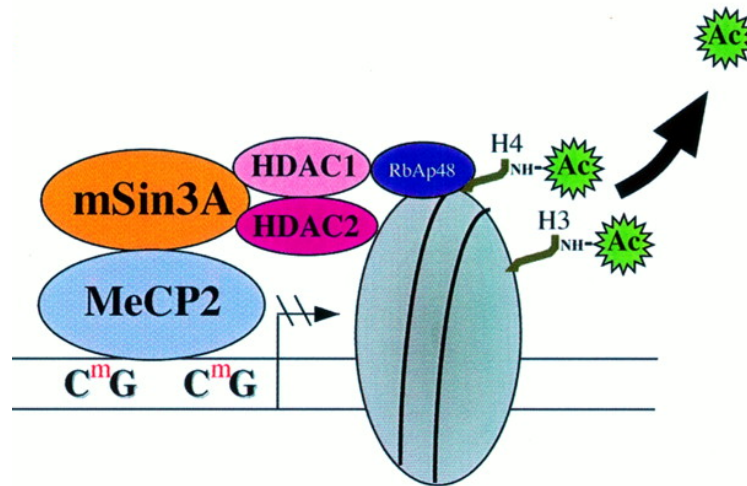


Abbildung 1.3: Die Transkriptionshemmung durch intaktes MeCP2 [Raz98].

DNA nachgewiesen werden konnte. Eine Liste dieser Proteine ist in Tabelle 1.1 und in Abbildung 1.4 gegeben. MBD4 nimmt hierbei eine Sonderstellung ein. Seine Funktion liegt im Austausch von (C^mpG)-(TpG)-Motiven zu (C^mpG)-(C^mpG)-Motiven, während die anderen Mitglieder der Familie an der Hemmung der Transkription beteiligt sind. Die Reparaturfunktion von MBD4 beruht auf der Wechselwirkung mit dem DNA *mismatch repair* Protein, MLH1.

Die dreidimensionale Struktur der methylbindenden Domäne von MeCP2 aus der Ratte [Wak99] sowie von menschlichem MBD1 [Ohk99] ist bekannt (siehe Abbildung 4.2). Bis auf geringe Unterschiede besitzen sie eine ähnliche Gestalt. Die Rolle von MeCP2 bei der Hemmung der Transkription beim Vorliegen von methylierter DNA läßt sich folgendermaßen beschreiben: Ist RNA- Polymerase unter Mitwirkung der Transkriptionsfaktoren aktiv, so liegt das Chromatin in einer nicht gepackten Form vor und die Termini der zentralen Histone H3 und H4 sind zu einem Großteil acetyliert. Unter Einwirkung von DNA-Methyltransferase entstehen methylierte CpG-Motive, an die methylbindende Proteine, wie MeCP2 binden. Unter dieser Voraussetzung kann das Protein mSin3A an die TRD binden, die Histon-Deacetylasen 1 und 2 (HDAC1,2) führen unter Mitwirkung des Proteins RbAp48 zur Deacetylierung der Histone H3 und H4. Dies hat eine Strukturveränderung hin zu einer dichteren Packung des Chromatins zur Folge, was die Inaktivie-

Name	Länge	Bindung	Aktivität	WW-Partner
MBD1 (PCM1)	636	(C ^m pG)-Paar	Repression	?
MBD2a/b	414/262	(C ^m pG)-Paar	Repression	HDAC1,2
MBD3	285	(C ^m pG)-Paar	Repression	HDAC1,2
MBD4 (MED1)	554	(C ^m pG)-(TpG)	Reparatur	MLH1
MeCP2	486	(C ^m pG)-Paar	Repression	HDACs, mSin3A, u.a.

Tabelle 1.1: *Die Familie der methylbindenden Proteine.* Die Länge, in Aminosäureresten gemessen, bezieht sich auf das im Menschen vorkommende Protein. MBD4 kann (C^mpG)-(TpG) in ein (C^mpG)-Paar überführen (→ Reparatur). Repression steht für die Hemmung der Transkription. Die Wechselwirkungspartner (WW-Partner) sind im Text näher erklärt.

rung der Transkription nach sich zieht. Diese Reaktionskette ist in Abbildung 1.2 und 1.3 schematisch dargestellt.

In diesem Zusammenhang ist ein anderes Protein, das ARBP (*attachment region binding protein*) zu nennen, welches die Fähigkeit besitzt, an sogenannte MARs bzw. SARs (*matrix attachment regions* bzw. *scaffold attachment regions*) zu binden [Kri91]. MARs bzw. SARs sind Bereiche der DNA, in denen eine Häufung von AT-Motiven vorliegt, die über Proteine wie ARBP an die Kernmatrix³ gebunden sein können.⁴ Die Länge der MARs kann bis zu mehreren Tausend Basenpaaren betragen. MARs liegen häufig in Bereichen der DNA-Sequenz, die an der Genregulation beteiligt sind. Durch Experimente an Mauszellen konnte gezeigt werden, daß in einigen Fällen SARs bei der Transkription ihnen benachbarter Gene eine Rolle spielen [Lod99]. MAR- bzw. SAR-bindende Proteine stellen somit neben den Histonen eine zweite Proteinklasse dar, die zum einen für die strukturelle Organisation der DNA verantwortlich sind, und zum anderen am Transkriptionsprozeß beteiligt sind. In [Wei97] wurde gezeigt, daß ARBP aus dem Huhn zu MeCP2 aus der Ratte (rMeCP2 -*rat* MeCP2) homolog ist.⁵ In [Wei97] wurde unter anderem das Bindungsverhalten von ARBP an verschiedenen DNA-Konstrukten untersucht. Es konnte nachge-

³Die Kernmatrix (*scaffold*) besteht aus Proteinen; die Struktur der Kernmatrix ist der des Chromosoms in der Metaphase ähnlich. Sie bleibt auch bei Wegfall der DNA (durch Verdauung durch Nukleasen) erhalten [Lod99].

⁴Befindet sich die Zelle in der Interphase bzw. Mitose so spricht man von MARs bzw. SARs.

⁵In [Str99] wurde ARBP aus dem Huhn cMeCP2 genannt.

	95	105	115	125	135	145	155	165								
MBD1	MAEDW	LDPCAL	GPGWKR	REVF	FRKSGAT	CGRSDT	YYQSPT	GDRIR	SKVEL	TRYLGP	ACDLTL	----	FDFK	-	QGILC	
MBD2a/b	ESGKR	MDCPAL	PPGWK	KEEV	IRKSGL	SAGKSD	VYFSP	SGKK	FRSKP	QLARYL	GNTVDL	SS----	FDFR	-	TGKMM	
MBD3	MERKR	WECPAL	PQGW	EREE	PRRSGL	SAGHRD	VFYFSP	SGKK	FRSKP	QLARYL	GGSMDL	ST----	FDFR	-	TGKML	
MBD4	ATAGTE	CRKSV	PCGWER	VVKQ	RFLFGK	TAGRFD	VYFYS	PQGLK	FRSKS	SLANYL	HKNGET	SLKPE	DFD	FTV	LSKRG	
MeCP2	DRGPM	YDDPTL	PEGW	TRK	LKQRK	SGRSAG	KYDVY	LINPQ	GKA	FRSKV	ELIAY	FEK	VGD	TSLD	PNDFD	FTVTGRGS

Abbildung 1.4: Sequenzen der Bindungsdomäne der verschiedenen menschlichen methylbindenden Proteine. Die Numerierung bezieht sich auf die in der vorliegenden Arbeit untersuchten Domäne von MeCP2 (cMeCP2 ARBD).

wiesen werden, daß eine Bindung zu einem Satelliten-DNA-Fragment (234 Basenpaare) der Maus, einem MAR-Fragment aus Hühner-Lysozym (235 Basenpaare) und mehreren repetitiven DNA-Fragmenten aus dem Huhn vorliegt. In diesen Bindungsstudien konnte eine minimale Domäne von 125 Aminosäureresten (*attachment region binding domain* (ARBD)) identifiziert werden, die zur Bindung an MARs nötig ist. Bei einer Verkürzung verliert die Domäne ihre Affinität zu DNA. Bei diesen Studien wurde auch das Bindungsverhalten der methylbindenden Domäne von MeCP2 allein untersucht. Das verwendete Konstrukt war am N- bzw. C-Terminus um sieben bzw. neun Aminosäurereste gegenüber der in [Wak99] beschriebenen Domäne (rMeCP2 MBD) verlängert. Es zeigte sich, daß diese Domäne nicht an MARs bindet. Bei Verwendung von ARBP und einem DNA-Fragment mit einem methylierten CpG-Motiv aus Satelliten-DNA der Maus wurden zwei bis fünf mal höhere Bindungsaffinitäten als für die nicht methylierte DNA gemessen. In Abbildung 1.5 ist die Sequenz von ARBP (Aminosäurerest 48 bis 232), von menschlichem MeCP2 (hMeCP2) und von rMeCP2 dargestellt. Ziel der vorliegenden Arbeit war die Bestimmung der Struktur der oben beschriebenen Domäne von 125 Aminosäureresten aus cMeCP2. Bei den in dieser Arbeit durchgeführten Bindungsstudien bezüglich des Protein-DNA-Komplexes kam ein Fragment der in [Wei97] verwendeten Satelliten-DNA der Maus zum Einsatz. Die Sequenz des verwendeten Konstrukts ist in Abbildung 2.3 gezeigt.

Anhand der hochaufgelösten Struktur von cMeCP2 ARBD sollten sich mögliche Interaktionsstellen mit DNA vorhersagen lassen. Ein Vergleich mit der Struktur von rMeCP2 MBD und MBD1 MBD ist in diesem Zusammenhang sinnvoll, um Gemeinsamkeiten innerhalb der Familie der methylbindenden Proteine aufzuzeigen, was Strukturvorhersagen für die anderen Mitglieder der Familie ermöglichen sollte. Weiterhin stellt die Struktur von cMeCP2 ARBD die Grundlage dafür dar, die Bindung des Proteins an DNA auf ato-

Kapitel 2

Materialien und verwendete NMR-Methoden

2.1 Bedeutung der NMR-Spektroskopie zur Aufklärung der Struktur von cMeCP2 ARBD als biologisches Makromolekül

Bedingt durch die großen Fortschritte bei der DNA-Sequenzierung ist die Primärstruktur vieler biologischer Makromoleküle mittlerweile bekannt. Die Funktion eines Protein hängt allerdings eng mit seinem dreidimensionalen Aufbau, d.h. seiner Sekundär- und Tertiärstruktur zusammen. Der Zusammenhang zwischen der Proteinstruktur und der Aminosäuresequenz des Proteins ist noch nicht hinreichend bekannt, um von der Sequenz allein auf die Struktur zu schließen. Die Struktur eines Proteins bleibt deshalb zwar nicht unzugänglich, erfordert aber die Untersuchung des Proteins auf atomarem Niveau. Zur Bestimmung der Struktur eines Proteins stehen im wesentlichen zwei unterschiedliche Methoden zur Verfügung. Zum einen ist dies die Röntgenkristallographie an Einkristallen, zum anderen die Kernspinresonanzspektroskopie (NMR-Spektroskopie) an Lösungen.

Andere Methoden, wie die Neutronenbeugung spielen eine untergeordnete Rolle und kommen nur in speziellen Fällen zum Einsatz. In diesem Zusammenhang ist darauf hinzuweisen, daß die Festkörper-NMR-Spektroskopie aufgrund neuerer methodischer Entwick-

lungen zunehmend an Bedeutung bei der Charakterisierung von Membranproteinen und Proteinaggregaten gewinnt [Hon99], [Les98], [Rei00].

Ein Großteil der bisher bekannten Strukturen biologischer Makromoleküle wurde röntgenkristallographisch bestimmt. Dies hat unter anderem folgende Gründe:

- Im Vergleich zur NMR-Spektroskopie ist die Strukturbestimmung von Proteinen mittels Röntgenkristallographie die wesentlich ältere Methode. Bereits 1959 wurde die dreidimensionale Struktur eines Proteins (Hämoglobin, Molekulargewicht: ≈ 60 kDa) Röntgenkristallographisch mit atomarer Auflösung bestimmt. Erst etwa 30 Jahre später wurde die Struktur von BPTI (Rinderpankreas-Trypsin-Inhibitor, Molekulargewicht: ≈ 8 kDa) mittels Flüssigkeits-NMR-Spektroskopie gelöst.¹
- Es gibt keine prinzipielle obere Grenze bezüglich der Molekülgröße für die Röntgenkristallographie im Gegensatz zur Flüssigkeits-NMR-Spektroskopie als Untersuchungsmethode. Dort lag die Obergrenze bis vor kurzem bei einem Molekulargewicht von etwa 20-30 kDa. Diese Grenze konnte durch Einführung der sogenannten *Transverse Relaxation Optimized Spectroscopy* (TROSY) [Per97] wesentlich nach oben verschoben werden.
- Ist die Zucht eines Einkristalls möglich, so ist der zeitliche und finanzielle Aufwand zur Bestimmung einer Struktur mit atomarer Auflösung wesentlich geringer als bei der Flüssigkeits-NMR-Spektroskopie.

Andererseits weist die Flüssigkeits-NMR-Spektroskopie einige wichtige Vorteile auf:

- Die Untersuchung von Biomolekülen in gelöster Form unter annähernd physiologischen Bedingungen (pH-Wert und Ionenkonzentration des Puffers) ist nur mit Flüssigkeits-NMR-Spektroskopie möglich.
- Die zeitlimitierende Größe bei der Strukturbestimmung mittels Röntgenkristallographie ist die Kristallisation des Proteins. Dieser rein empi-

¹Die routinemäßige Bestimmung von Proteinstrukturen mittels NMR-Spektroskopie erfordert einerseits mehrdimensionale NMR-Experimente und andererseits isotope markierte Proteine (die verwendeten Isotope sind meist ^{13}C und ^{15}N) in ausreichenden Mengen (vorzugsweise mehrere Hundert Microliter Lösung bei Konzentrationen im millimolaren Bereich).

rische Schritt gelingt häufig nur unter unphysiologischen pH-Bedingungen, was die Deutung der Struktur des Proteins hinsichtlich seiner Funktion unter anderen (physiologischen) Bedingungen erschwert.

- Dynamische Prozesse, insbesondere die eigentliche Faltung des Proteins, läßt sich mit Hilfe der Flüssigkeits-NMR-Spektroskopie im Gegensatz zur Röntgenkristallographie über weite Zeit- und Frequenzbereiche studieren.
- Wechselwirkungen verschiedener Moleküle können mit NMR-Spektroskopie relativ einfach auf atomarer Ebene untersucht werden. Dies findet besonders bei der pharmakologischen Wirkstoffentwicklung Anwendung.
- Durch neuere Entwicklungen auf dem Gebiet der Biochemie und Genetik ist es möglich eine Vielzahl von Proteinen isotonenmarkiert in einem hohen Reinheitsgrad herzustellen.

Hinsichtlich der Untersuchung der Struktur von cMeCP2 ARBD im nativen und im DNA-gebundenen Zustand erscheint daher die Flüssigkeits-NMR-Spektroskopie als die geeignetere Untersuchungsmethode. Röntgenkristallographische Untersuchungen an cMeCP2 ARBP sind bisher weder im nativen noch im DNA-gebundenen Zustand durchgeführt worden, über die Kristallisation des Proteins ist ebenfalls nichts bekannt. Mit einem Molekulargewicht von ca. 15,4 kDa im nativen Zustand und ca. 25 kDa im DNA-gebundenen Zustand mit einem geeignete Oligonucleotid von 15 Basenpaaren (siehe Abschnitt 2.3) ist dieses Protein Flüssigkeits-NMR-spektroskopischen Untersuchungen zugänglich. Dabei ist insbesondere beim DNA-gebundenem Komplex zu bemerken, daß wegen seines Molekulargewichtes die Anwendung von TROSY-Experimenten sinnvoll erscheint. Im Rahmen einer Kooperation mit Herrn Prof. W. H. Strätling vom Universitäts-Krankenhaus Eppendorf (Hamburg) war es möglich, cMeCP2 ARBD isotonenmarkiert in hoher Reinheit zu erhalten. Weiterhin bietet die Verwendung der Flüssigkeits-NMR-Spektroskopie als Untersuchungsmethode den Vorteil, daß sich die Untersuchung des Protein-DNA-Komplexes relativ einfach an die Untersuchung des Proteins anschliessen läßt.

Im Abschnitt 2.2 werden einige Arbeiten vorgestellt, die den eigentlichen Untersuchungen an cMeCP2 ARBD voran gingen. Dabei handelt es sich um allgemeine Betrachtungen zur Linienbreite (Abschnitt 2.2.1) und zur partiellen Deuterierung von Proteinen

(Abschnitt 2.2.2). Die sich anschließenden Abschnitte (2.3 und 2.4) behandeln die Probenpräparation, die Aufnahmebedingungen, die Spektren-Aufnahme und -Visualisierung. Daran schließt sich die Beschreibung der Methoden an, welche zur Weiterverarbeitung der gesammelten und vorverarbeiteten spektroskopischen Daten verwendet wurden (Abschnitt 2.5 und 2.6).

2.2 Methodische Vorarbeiten

Neben der ^{15}N - und ^{13}C -Isotopenmarkierung von Proteinen ist es auch möglich, ^1H -Atome im Protein durch ^2H -Atome zu ersetzen. Die Deuterierung von Proteinen beeinflusst in Abhängigkeit vom Deuterierungsgrad die Durchführbarkeit von Flüssigkeits-NMR-Experimenten. Einerseits sollte sich die Qualität der Spektren hinsichtlich der Linienbreite erhöhen, da ^2H -Kerne im Gegensatz zu ^1H -Kernen ein wesentlich geringeres gyromagnetisches Verhältnis aufweisen und dadurch der Einfluß der magnetischen Dipol-Dipol-Wechselwirkung auf die transversale Relaxationszeit minimiert wird. Andererseits wird die Anwendbarkeit vieler typischer Flüssigkeits-NMR-Experimente eingeschränkt, da die ^1H -Kerne häufig für den Magnetisierungstransfer bzw. zur Detektion der Signale benutzt werden. In diesem Zusammenhang sind daher Überlegungen zur Linienbreite (Abschnitt 2.2.1) mit besonderem Augenmerk auf den Einfluß des Deuterierungsgrades (Abschnitt 2.2.2) angestellt worden.

Eingangs sind noch einige Begriffe zu definieren, welche in den nachfolgenden Abschnitten Verwendung finden. Die hier besprochenen Flüssigkeits-NMR-Experimente basieren auf der sogenannten gepulsten *fourier transform*- (FT-)Spektroskopie. Im Gegensatz zur früher verwendeten *continuous wave*- (CW-)Spektroskopie erfolgt die Aufnahme des Spektrums nicht während der kontinuierlichen Variation des Magnetfeldes oder der Frequenz, sondern nach Einstrahlung von *radio frequency*- (rf-)Pulsen bei konstantem äußeren Magnetfeld. Das Anregungsprofil des Pulses ergibt sich durch Fouriertransformation der Funktion, die die rf-Amplitude in Abhängigkeit von der Zeit beschreibt. Im thermischen Gleichgewicht steht die makroskopische Magnetisierung in Richtung des äußeren Magnetfeldes (im allgemeinen als z -Richtung bezeichnet). Unter dem Einfluß eines rf-Pulses kommt es zum Aufbau eines effektiven Magnetfeldes unter dessen Einfluß die Magnetisie-

rung aus der z -Richtung heraus gedreht wird. Steht die Magnetisierung nach Einwirkung des Pulses senkrecht bzw. antiparallel zur z -Richtung, so bezeichnet man den Puls, der dies bewirkt hat als 90° - $(\pi/2)$ - bzw. 180° - (π) -Puls.

2.2.1 Überlegungen zur Linienbreite in HSQC- bzw. TROSY-Experimenten

Im sogenannten HSQC-Experiment (*Heteronuclear Single Quantum Coherence*) wird Magnetisierung vom Spin I mittels eines INEPT-Schrittes (*Insensitive Nuclei Enhanced by Polarization Transfer*) zum Spin S transferiert.² Dort entwickelt sich die Magnetisierung unter dem Einfluß des effektiven Magnetfeldes während der Zeit t_1 . In einem zweiten INEPT-Schritt wird die Magnetisierung wieder zurück auf Spin I transferiert und anschließend detektiert (t_2). Das HSQC-Experiment ist schematisch als Pulssequenz in Abbildung 2.2 dargestellt. Da ^1H -Kerne dem Betrag nach ein etwa zehnmal bzw. viermal größeres gyromagnetisches Verhältnis besitzen als ^{15}N - bzw. ^{13}C -Kerne, wird für den Spin I meist der ^1H -Kern und für den Spin S der ^{15}N - bzw. ^{13}C -Kern gewählt.³

Führt man ein Experiment aus, in dem nur die in Abbildung 2.2 schwarz dargestellten Pulse eingestrahlt werden, so ergibt sich für ein isoliertes, skalar gekoppeltes Zweispinsystem (z.B. $^1\text{H}_N$ und ^{15}N) ein Spektrum wie in Abbildung 2.3a. Verursacht durch die skalare Kopplung⁴ zwischen dem $^1\text{H}_N$ -Atom und dem ^{15}N -Atom kommt es zu einer Aufspaltung des Signals in der ^{15}N - ebenso wie in der ^1H -Dimension. Man spricht von einem nicht entkoppelten ^1H - ^{15}N -HSQC-Spektrum. Zur Veranschaulichung ist in Abbildung 2.1 das Termschema eines isolierten Zweispinsystems ($I, S=1/2$, z.B. ^1H und ^{15}N) unter dem Einfluß der Zeemann-Wechselwirkung (WW) und zusätzlich unter dem Einfluß der J -Kopplung zwischen den beiden Kernen dargestellt.

Führt man das gleiche Experiment durch, implementiert aber zusätzlich die grau unter-

²Die genaue Wirkungsweise des INEPT-Schrittes ist [Cav95] zu entnehmen.

³Für die gyromagnetischen Verhältnisse γ einiger für die NMR-Spektroskopie relevanter Kerne gilt: ^1H : $\gamma \approx 2,6752 \times 10^8$, ^2H : $\gamma \approx 4,107 \times 10^7$, ^{15}N : $\gamma \approx -2,712 \times 10^7$, ^{13}C : $\gamma \approx 6,728 \times 10^7$, ^{31}P : $\gamma \approx 1,0841 \times 10^8$ (Werte entnommen aus [Cav95]).

⁴häufig auch als J -Kopplung bezeichnet

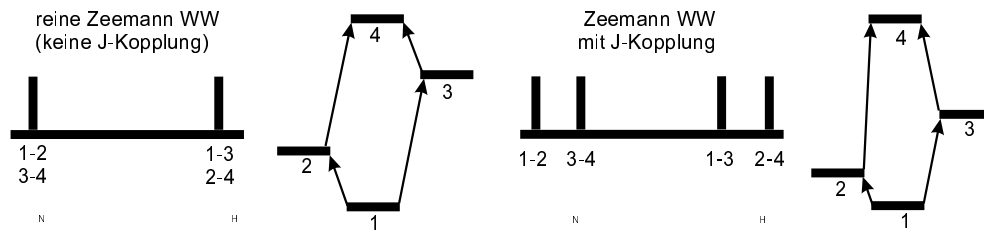


Abbildung 2.1: Darstellung des Termschemas zweier isolierter Spin-1/2-Kerne. Die Zeemann-Wechselwirkung (WW) bewirkt eine Aufspaltung der Energieniveaus (links), welche eine Verschiebung aufgrund der J -Kopplung (rechts) erfahren (Details sind [Cav95] zu entnehmen).

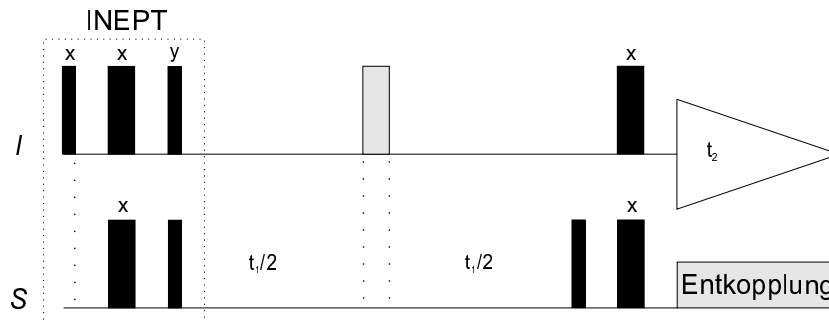


Abbildung 2.2: Schematische Darstellung des HSQC-Spektrums in Form einer Pulssequenz. $\pi/2$ -Pulse (π -Pulse) sind durch schmale (breite) Balken symbolisiert. $\tau = \frac{1}{4J_{IS}}$, mit der skalaren Kopplung J_{IS} zwischen Spin I und Spin S . Über den Pulsen sind die Phasen vermerkt; es gilt: $\phi_1 = (x, x)$, $\phi_2 = 2(x)$, $\phi_3 = 4(y)$, $4(-y)$. Weitere Informationen sind dem Text und [Cav95] zu entnehmen. (Der 2. INEPT-Schritt wurde nicht optisch hervorgehoben.)

legten Pulse⁵, so ergibt sich ein Spektrum, das dem in Abbildung 2.3b entspricht. Durch den grau unterlegten π -Puls zur Zeit $t_1/2$ wird Spin I vom Spin S entkoppelt. Eine weitere Entkopplung unterdrückt den Einfluß der J -Kopplung der beiden Spins während der Aufnahmezeit t_2 . Man spricht von einem entkoppeltem ^1H - ^{15}N -HSQC-Spektrum.

Betrachtet man Abbildung 2.3a näher, so erkennt man, daß die vier Signale in beiden Dimensionen unterschiedlich große Linienbreiten aufweisen (siehe z.B. [Gol84]). Ein Signal des Quartetts weist in beiden Dimensionen eine relativ geringe Linienbreite auf. Im Jahre 1997 wurde das TROSY-Experiment entwickelt, welches auf der Selektion dieses schmalen Signals beruht [Per97]. Es ergibt sich ein Spektrum, das dem in Abbildung 2.3c

⁵Die Entkopplung während der Aufnahme besteht meistens aus mehreren Pulsen. Typische Pulssequenzen für die Entkopplung sind z.B. waltz16 oder mlev.

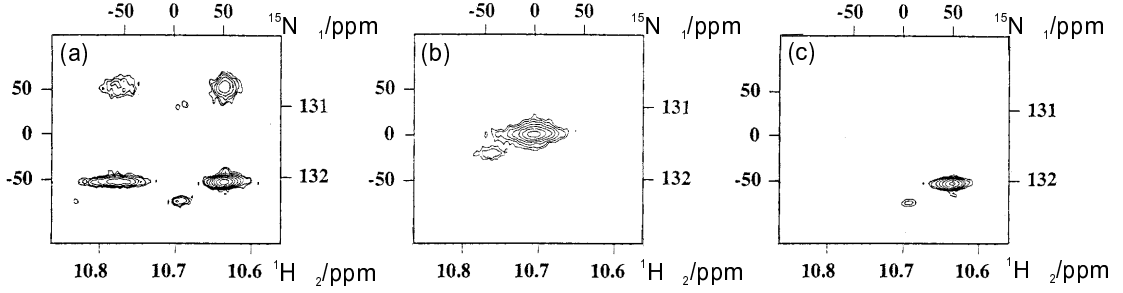


Abbildung 2.3: Ausschnitte aus einem (a) nichtentkoppelten ^1H - ^{15}N -HSQC-Spektrum, (b) entkoppelten ^1H - ^{15}N -HSQC-Spektrum und (c) ^1H - ^{15}N -TROSY-Spektrum (Abbildung entnommen aus [Per97]). Rechts (unterhalb) der Spektren sind die chemischen Verschiebungen für ^{15}N (^1H) in ppm aufgetragen. Links (oberhalb) der Spektren sind die Differenzen der chemischen Verschiebungen für ^{15}N (^1H) normiert auf das entkoppelte ^1H - ^{15}N -HSQC-Spektrum in Hertz aufgetragen.

entspricht. Die ^{15}N - sowie die $^1\text{H}_N$ -Linienbreite werden im wesentlichen durch transversale Relaxation bestimmt, welche mit der Molekülgröße zunimmt. Den Hauptanteil tragen die Dipol-Dipol-Wechselwirkung (DD) und die Anisotropie der chemischen Verschiebung (CSA) bei [Per97].

2.2.1.1 Die ^{15}N -Linienbreite

Für die ^{15}N -Linienbreite ($\Delta\nu_{\text{HSQC}}(^{15}\text{N})$) und ($\Delta\nu_{\text{TROSY}}(^{15}\text{N})$) eines isolierten, skalar gekoppelten Zweispinsystems ($^1\text{H}_N$ und ^{15}N) ergibt sich [Per97], [Fus99]:⁶

$$\Delta\nu_{\text{HSQC}}(^{15}\text{N}) = \frac{1}{\pi} \cdot (d^2 + c_N^2) \cdot (4J(0) + 3J(\omega_N)) \quad (2.1)$$

$$\Delta\nu_{\text{TROSY}}(^{15}\text{N}) = \frac{1}{\pi} \cdot (d^2 + c_N^2 + dc_N \cdot (3 \cos^2 \theta - 1)) \cdot (4J(0) + 3J(\omega_N)) \quad (2.2)$$

Der DD-Term d und der CSA-Term c_N sind gegeben durch:

$$d = \frac{\mu_0}{4\pi} \cdot \frac{h}{2\pi} \cdot \gamma_N \gamma_H \cdot \frac{1}{r_{HN}^3} \quad (2.3)$$

$$c_N = \gamma_N \cdot \frac{\Delta\sigma_N}{3} \cdot B_0 \quad (2.4)$$

⁶für $\Delta\sigma_N < 0$

Hierbei bezeichnet ω die Larmor-Frequenz, θ den Winkel zwischen der zz -Achse des Tensors der magnetischen Abschirmung (CSA-Tensor) von ^{15}N und dem Kernverbindungsvektor \vec{r}_{HN} sowie $\Delta\sigma_N$ die Anisotropie der chemischen Verschiebung (mit $\Delta\sigma = \sigma_{zz} - (\sigma_{yy} + \sigma_{xx})/2$; hierbei stehen die σ_{ii} für die Hauptwerte des CSA-Tensors). Weiterhin bezeichnet μ_0 die magnetische Permeabilität des Vakuums, h das Plancksche Wirkungsquantum und B_0 das äußere Magnetfeld in Tesla. Für die spektrale Dichtefunktion $J(\omega)$ eines starren, sphärischen Proteinmoleküls gilt:

$$J(\omega) = \frac{2\tau_c}{5(1 + (\tau_c\omega)^2)} \quad (2.5)$$

Dabei bezeichnet τ_c die Korrelationszeit für die isotrope Umorientierung des Moleküls in Lösung. Sie ist unter anderem von der Größe des betrachteten Moleküls abhängig und steigt mit wachsender Molekülgröße. In den gezeigten Simulationen wurde durchgehend $\tau_c = 10$ ns verwendet,⁷ was einem sphärischen Protein mit einem Molekulargewicht von etwa 25 kDa bei einer Temperatur von 298 K in Wasser entspricht. Für die Länge des Kernverbindungsvektors \vec{r}_{HN} wurde in allen Simulationen 0,102 nm verwendet. Der Einfluß von τ_c auf die ^{15}N -Linienbreite ist aus Abbildung 2.4 ersichtlich.⁸ Der Unterschied zwischen der ^{15}N -Linienbreite beim ^1H - ^{15}N -TROSY- und dem ^1H - ^{15}N -HSQC-Experiment nimmt mit steigender Molekülgröße kontinuierlich zu.

Die exakte Gleichung für die Linienbreite weist noch andere spektrale Dichteterme auf, welche von Kreisfrequenzen größer als ω_N abhängig sind; diese wurden auf Grund ihres geringen Einflusses auf die ^{15}N -Linienbreite in den Simulationen nicht berücksichtigt [Fus99].

In [Fus98] wurde die Anisotropie der chemischen Verschiebung von ^{15}N und der Winkel θ in Abhängigkeit von der Aminosäuresequenz für menschliches Ubiquitin bestimmt. Die Werte für die Anisotropie der chemischen Verschiebung von ^{15}N lagen zwischen -125 und -216 ppm mit einem Mittelwert von -157 ± 19 ppm. Für den Winkel θ werden Werte von 6 bis 26° (Mittelwert $15,7 \pm 5^\circ$) angegeben. Diese Werte dienen als Anhaltspunkt für

⁷mit Ausnahme von Abbildung 2.4

⁸Unter den gleichen Annahmen wie oben, entspricht $\tau_c = 20$ ns und $\tau_c = 40$ ns einem Molekulargewicht von 50 kDa bzw. 100 kDa.

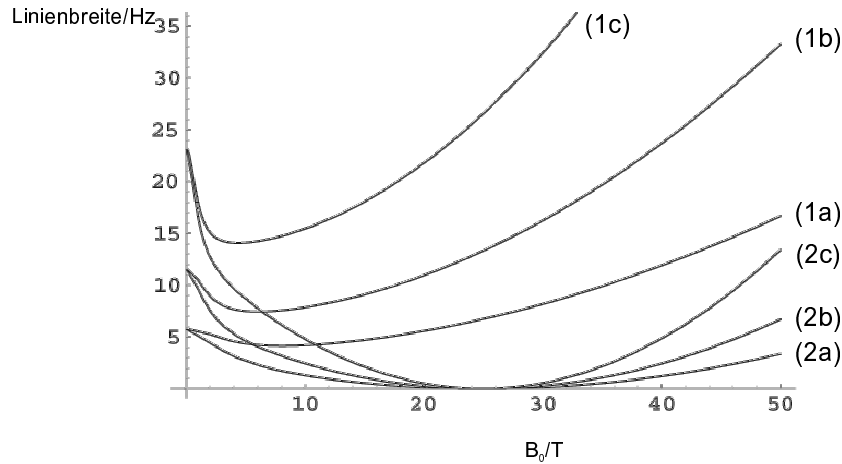


Abbildung 2.4: Simulation der ^{15}N -Linienbreite in Hertz in Abhängigkeit vom angelegten Magnetfeld B_0 in Tesla bei verschiedenen Werten für τ_c . (1) stellt die Magnetfeldabhängigkeit von Gleichung (2.1) und (2) die von Gleichung (2.2) dar. Für (1) und (2) wurde $\theta = 0^\circ$ angenommen. Bei (a), (b) und (c) wurde $\tau_c = 10$ ns, 20 ns bzw. 40 ns gewählt. Für $\Delta\sigma_N$ wurde in allen Rechnungen ein Wert von -160 ppm gewählt.

die Simulationen der ^{15}N -Linienbreite in diesem Abschnitt. Die Magnetfeldabhängigkeit der ^{15}N -Linienbreiten ist durch Gleichung (2.1) und (2.2) bestimmt und in Abbildung 2.5 graphisch dargestellt. Sind die zz -Achse des CSA-Tensors des ^{15}N und der Kernverbindungsvektor \vec{r}_{HN} parallel ($\theta = 0$), so ergibt sich bei einem Wert von -160 ppm für $\Delta\sigma_N$ im ^1H - ^{15}N -TROSY-Experiment eine ^{15}N -Linienbreite von 0 Hertz bei einem Magnetfeld von etwa 25 Tesla (T). Im entkoppelten ^1H - ^{15}N -HSQC-Experiment ergibt sich bei dieser Feldstärke eine ^{15}N -Linienbreite von etwa 6,5 Hz. Ein Minimum der Linienbreite im entkoppelten ^1H - ^{15}N -HSQC-Experiment ist bei einem Magnetfeld von etwa 9,4 T zu finden (in Abbildung 2.5 Graph (1) und (2)).⁹ Wird ein Wert von 17° für θ angenommen, so verschiebt sich das Minimum von (2) zu kleineren Feldern hin (von etwa 25 T auf 20,4 T) und der Wert für die minimale Linienbreite erhöht sich von 0 auf etwa 1,2 Hz (in Abbildung 2.5 Graph (3)). In Abbildung 2.6 ist die ^{15}N -Linienbreite beim ^1H - ^{15}N -TROSY-Experiment in Abhängigkeit vom äußeren Magnetfeld für unterschiedliche Werte von θ und $\Delta\sigma_N$ gezeigt. Für steigende Werte von θ bei konstantem $\Delta\sigma_N$ (-160 ppm) verschiebt sich die minimale Linienbreite zu niedrigerem Feld (von etwa 24 T bei $\theta = 5^\circ$ zu etwa 16 T bei

⁹zur Orientierung: Protonenresonanzfrequenzen von 800, 600 und 500 MHz entsprechen Magnetfeldern von ungefähr 18,8 T, 14,1 T und 11,7 T.

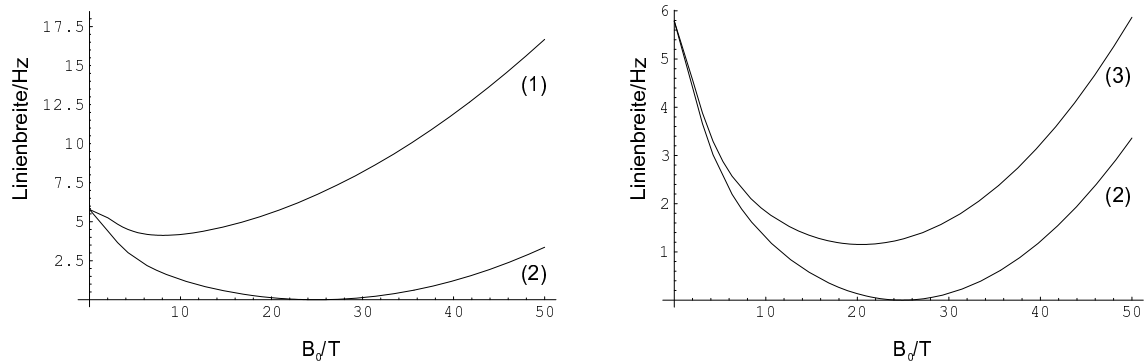


Abbildung 2.5: Simulation der ^{15}N -Linienbreite in Hertz in Abhängigkeit vom angelegten Magnetfeld B_0 in Tesla. (1) stellt die Magnetfeldabhängigkeit von Gleichung (2.1) und (2) die von Gleichung (2.2) dar. Für (1) und (2) wurde θ als Null angenommen. (3) stellt Gleichung (2.2) mit $\theta = 17^\circ$ dar. Für $\Delta\sigma_N$ wurde in allen Rechnungen ein Wert von -160 ppm gewählt.

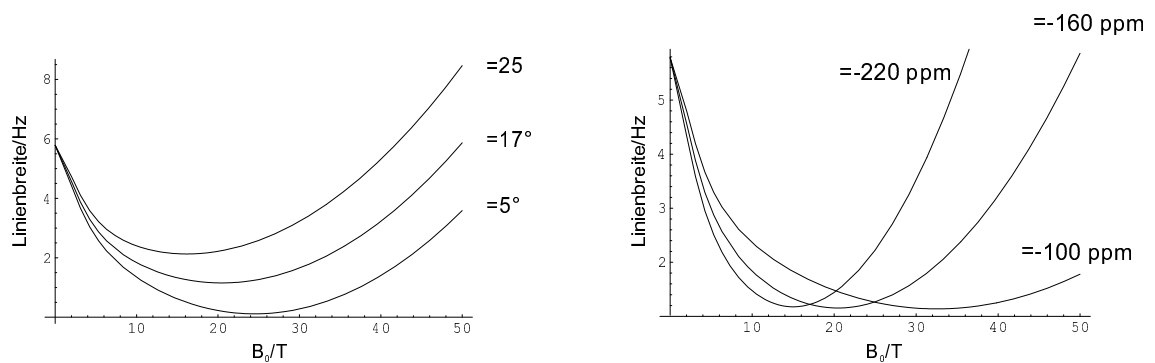


Abbildung 2.6: Simulation der ^{15}N -Linienbreite in Hertz in Abhängigkeit vom angelegten Magnetfeld B_0 in Tesla bei variierenden Werten für θ und $\Delta\sigma_N$ im ^1H - ^{15}N -TROSY-Experiment. Bei der Variation von θ (Links in der Abbildung) wurde für $\Delta\sigma_N$ ein Wert von -160 ppm gewählt. Bei der Variation von $\Delta\sigma_N$ (Rechts in der Abbildung) wurde für den Winkel θ ein Wert von 5° gewählt.

$\theta = 25^\circ$) und der Wert des Minimums steigt von unter 0,5 Hertz bei $\theta = 5^\circ$ auf über 2 Hertz bei $\theta = 25^\circ$ (linker Teil der Abbildung 2.6). Erhöht man $\Delta\sigma_N$ bei festem θ (5°), so verschiebt sich das Minimum zu niedrigeren Feldern (von etwa 32,5 T bei $\Delta\sigma_N = -100$ ppm zu etwa 14,7 T bei $\Delta\sigma_N = -220$ ppm) und die Variation der Linienbreite bei sich änderndem Magnetfeld nimmt zu, wobei sich der Wert für die minimale Linienbreite nicht ändert (rechter Teil der Abbildung 2.6).

2.2.1.2 Die $^1\text{H}_N$ -Linienbreite

Unter den gleichen Annahmen wie in Abschnitt 2.2.1.1 ergibt sich die $^1\text{H}_N$ -Linienbreite prinzipiell analog zur ^{15}N -Linienbreite keiner der Spektraldichteterme vernachlässigt werden darf [Per97]. Für die $^1\text{H}_N$ -Linienbreite im ^1H - ^{15}N -HSQC- und im ^1H - ^{15}N -TROSY-Experiment ($\Delta\nu_{\text{HSQC}}(^1\text{H})$ und $\Delta\nu_{\text{TROSY}}(^1\text{H})$) ergibt sich unter Annahme rein transversaler Relaxation nach [Per97] und [Fus99]:¹⁰

$$\begin{aligned} \Delta\nu_{\text{HSQC}}(^1\text{H}) = \frac{1}{\pi} \cdot & \left((d^2 + c_H^2) \cdot (4J(0) + 3J(\omega_H)) \right. \\ & \left. + d^2 (J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)) + 3c_N^2 J(\omega_N) \right) \quad (2.6) \end{aligned}$$

$$\begin{aligned} \Delta\nu_{\text{TROSY}}(^1\text{H}) = \frac{1}{\pi} \cdot & \left((d^2 + c_H^2 + dc_H \cdot (3\cos^2\theta - 1)) \cdot (4J(0) + 3J(\omega_H)) \right. \\ & \left. + d^2 (J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)) + 3c_N^2 J(\omega_N) \right) \quad (2.7) \end{aligned}$$

Der DD-Term d ist durch Gleichung (2.3) gegeben und der CSA-Term c_H ergibt sich analog aus Gleichung (2.4). Für Gleichung (2.6) und (2.7) gelten die gleichen Annahmen wie in Abschnitt 2.2.1.1. θ bezeichnet den Winkel zwischen der zz -Achse des CSA-Tensors des $^1\text{H}_N$ und dem Kernverbindungsvektor \vec{r}_{HN} .

Für die Anisotropie der chemischen Verschiebung von $^1\text{H}_N$ ($\Delta\sigma_H$) sind, wie für $\Delta\sigma_N$, Werte für menschliches Ubiquitin sequenzabhängig bestimmt worden [Tja97b]. Sie variieren von etwa 0,5 ppm bis etwa 14,5 ppm. In [Ger93] wurde unter anderem der Winkel θ an einer Modellaminosäure gemessen. Es ergab sich ein Winkel von ca. 0° bei einem Fehler von 9° . Diese Werte dienten als Ausgangspunkt für die Simulationen der $^1\text{H}_N$ -Linienbreite

¹⁰für $\Delta\sigma_H > 0$

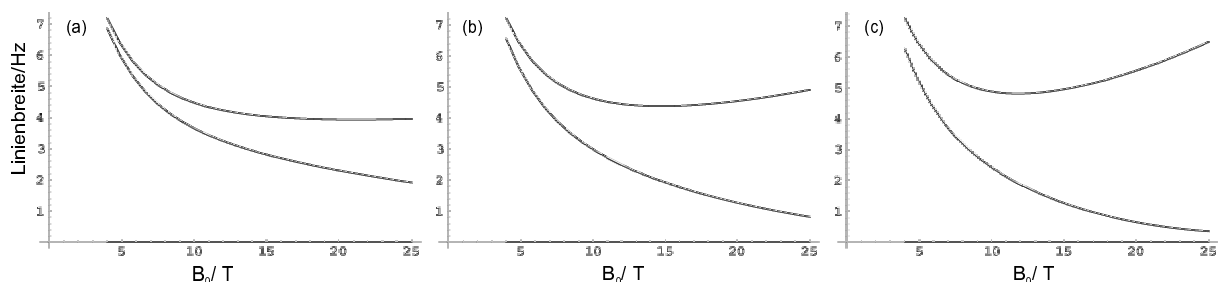


Abbildung 2.7: Simulation der ${}^1\text{H}_N$ -Linienbreite in Hertz in Abhängigkeit vom angelegten Magnetfeld B_0 in Tesla im ${}^1\text{H}$ - ${}^{15}\text{N}$ -HSQC-Experiment (obere Graphen) und im ${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiment (untere Graphen). Mit Werten für $\Delta\sigma_H$ von 5 ppm (a), 10 ppm (b) und 15 ppm (c).

in diesem Abschnitt. In Abbildung 2.7 sind die ${}^1\text{H}_N$ -Linienbreiten im ${}^1\text{H}$ - ${}^{15}\text{N}$ -HSQC- und im ${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiment für verschiedene Werte der Anisotropie der chemischen Verschiebung des ${}^1\text{H}$ -Atoms gegenübergestellt. Für eine Protonenresonanzfrequenz von 800 MHz ($\approx 18,8$ T) unterscheiden sich ${}^1\text{H}$ - ${}^{15}\text{N}$ -HSQC- und ${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiment bezüglich der ${}^1\text{H}_N$ -Linienbreite um 1,6 Hz, 3,1 Hz bzw. 4,7 Hz für $\Delta\sigma_H$ Werte von 5, 10 bzw. 15 ppm. In Abbildung 2.8 ist die ${}^1\text{H}_N$ -Linienbreite abgeleitet aus 2.7 (${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiment) abhängig vom äußeren Magnetfeld für verschiedene Werte für $\Delta\sigma_H$ und θ dargestellt.

Wie bei der Simulation des ${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiments bezüglich der ${}^{15}\text{N}$ -Linienbreite (Abbildung 2.6) zeigt sich bei der Simulation der ${}^1\text{H}_N$ -Linienbreite ein Minimum, dessen Lage von den Werten für $\Delta\sigma_H$ und θ abhängt. Erhöht man $\Delta\sigma_H$, so verschiebt sich das Minimum zu niedrigeren Feldern (von etwa 78 T bei $\Delta\sigma_H=5$ ppm zu etwa 26 T bei $\Delta\sigma_H=15$ ppm). Die Variation des Winkels θ hat für die verwendeten Werte (0° und 5°) nur einen geringen Einfluß auf die Lage und den Wert des Minimums der ${}^1\text{H}_N$ -Linienbreite, wobei analog zur ${}^{15}\text{N}$ -Linienbreite eine Erhöhung des Winkels zu einer Erhöhung des Wertes der minimalen Linienbreite und zu einer Verschiebung der Position des Minimums zu geringem Feld führt. Der Wert des Minimums liegt bei allen Simulationen der ${}^1\text{H}_N$ -Linienbreite im ${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiment unter 0,5 Hz. Es ist anzumerken, daß abweichend von den Simulationen der ${}^{15}\text{N}$ -Linienbreite auch für Simulationen mit angenommener Parallelität der zz -Achse des CSA-Tensors des ${}^1\text{H}_N$ -Kerns und des Kernverbindungsvektors \vec{r}_{HN} ($\theta=0$) die Linienbreite nicht Null wird. Dies liegt darin begründet, daß in Gleichung (2.7), anders

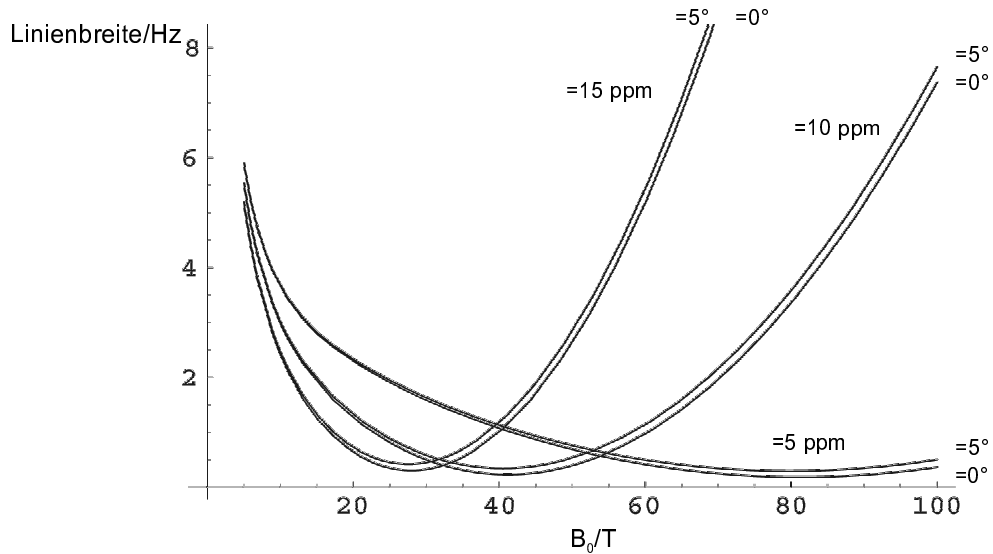


Abbildung 2.8: Simulation der ${}^1\text{H}_N$ -Linienbreite in Hertz in Abhängigkeit vom angelegten Magnetfeld B_0 in Tesla bei variierenden Werten für θ und $\Delta\sigma_H$ im ${}^1\text{H}$ - ${}^{15}\text{N}$ -TROSY-Experiment. Für $\Delta\sigma_N$ wurde in allen Simulationen ein Wert von -160 ppm gewählt.

als in Gleichung (2.2) alle spektralen Dichteterme einfließen.¹¹

2.2.1.3 Weiterführende Überlegungen zur Linienbreite

In den Abschnitten 2.2.1.1 und 2.2.1.2 wurde der Einfluß der transversalen Relaxation auf die ${}^{15}\text{N}$ - und die ${}^1\text{H}_N$ -Linienbreite für ein isoliertes Zweispinsystem betrachtet. In diesem Abschnitt soll der Einfluß benachbarter Spin-1/2-Kerne auf die ${}^{15}\text{N}$ - und ${}^1\text{H}_N$ -Linienbreite untersucht werden. Aufgrund der deutlich kleineren gyromagnetischen Verhältnisse von ${}^{15}\text{N}$ und ${}^{13}\text{C}$ im Vergleich zu ${}^1\text{H}$ wird nur der Einfluß benachbarter ${}^1\text{H}$ -Atome auf die ${}^1\text{H}_N$ -Linienbreite untersucht. Es läßt sich für die resultierende ${}^1\text{H}_N$ -Linienbreite schreiben:

$$\Delta\nu_{TROSY}^{res}({}^1\text{H}) = \Delta\nu_{TROSY}({}^1\text{H}) + \Delta\nu_1({}^1\text{H}) + \Delta\nu_2({}^1\text{H}) \quad (2.8)$$

Dabei steht der Linienbreitenbeitrag $\Delta\nu_1({}^1\text{H})$ bzw. $\Delta\nu_2({}^1\text{H})$ für die longitudinale bzw. transversale Relaxation, hervorgerufen durch dem ${}^1\text{H}_N$ -Kern benachbarte ${}^1\text{H}$ -Atome (${}^1\text{H}_k$). Durch die Verwendung deuterierter Proben wird der Einfluß dieser beiden Terme auf die Gesamtlinienbreite minimiert. Nach [Per97] spielt die DD-Wechselwirkung die

¹¹den Hauptanteil trägt der Term proportional zu c_N^2 bei

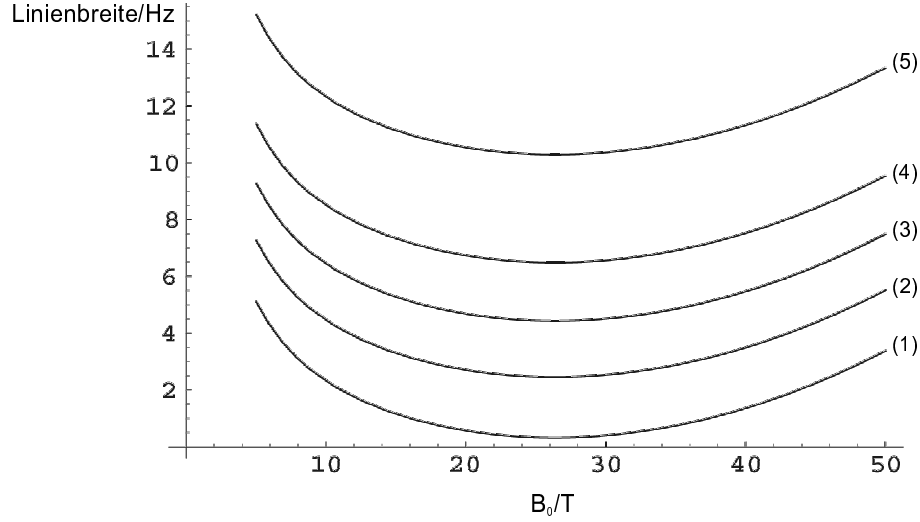


Abbildung 2.9: Simulation der ${}^1\text{H}_N$ -Linienbreite in Hertz in Abhängigkeit vom angelegten Magnetfeld B_0 in Tesla unter dem Einfluß benachbarter ${}^1\text{H}$ -Atome. (1) reines ${}^1\text{H}_N$ - ${}^{15}\text{N}$ -System; (2) mit zwei ${}^1\text{H}$ -Atomen im Abstand von 0,29 nm; (3) mit zwei ${}^1\text{H}$ -Atomen im Abstand von 0,26 nm; (4) mit drei ${}^1\text{H}$ -Atomen im Abstand von 0,26 nm; (5) mit drei ${}^1\text{H}$ -Atomen im Abstand von 0,24 nm. Es wurden $\Delta\sigma_N = -160$ ppm, $\theta = 0^\circ$ und $\Delta\sigma_H = 15$ ppm gewählt.

dominierende Rolle bezüglich der $\Delta\nu_1({}^1\text{H})$ - und $\Delta\nu_2({}^1\text{H})$ -Linienbreiten. [Pen92] liefert folgende Ausdrücke für die Beiträge zur resultierenden Linienbreite, hervorgerufen durch longitudinale und transversale Relaxation:

$$\Delta\nu_1({}^1\text{H}) = \frac{1}{\pi} \cdot \left(\sum_k \left(\mu_0 \gamma_H \hbar / 8\vec{r}_k^3 \right)^2 (J_k(0) + J_k(\omega_H) + J_k(2\omega_H)) \right) \quad (2.9)$$

$$\Delta\nu_2({}^1\text{H}) = \frac{1}{\pi} \cdot \left(\sum_k \left(\mu_0 \gamma_H \hbar / 8\vec{r}_k^3 \right)^2 \left(\frac{5}{2} J_k(0) + \frac{9}{2} J_k(\omega_H) + 3J_k(2\omega_H) \right) \right) \quad (2.10)$$

\vec{r}_k ist der Kernverbindungsvektor zwischen ${}^1\text{H}_N$ und ${}^1\text{H}_k$; J_k ist die spektrale Dichtefunktion des k-ten ${}^1\text{H}$ -Atoms; \hbar ist definiert als $h/2\pi$.

In Abbildung 2.9 ist $\Delta\nu_{TROSSY}^{res}({}^1\text{H})$ in Abhängigkeit vom äußeren Magnetfeld unter dem Einfluß unterschiedlich vieler ${}^1\text{H}$ -Atome mit unterschiedlichen Abständen zu ${}^1\text{H}_N$ aufgetragen. Für die spektralen Dichtefunktionen J_k wurde Gleichung (2.5) für die Simulationen verwendet; weiterhin wurde wie auch sonst in den Simulationen $\tau_c = 10$ ns gewählt.

Alle fünf Simulationen weisen ein Minimum der ${}^1\text{H}_N$ -Linienbreite bei etwa 26 T auf, wobei der Wert des Minimums stark von den benachbarten ${}^1\text{H}$ -Atomen abhängt: Der Wert

der minimalen Linienbreite steigt von unter 0,5 Hz für das isolierte $^1\text{H}_N$ - ^{15}N -System auf über 11 Hz für drei ^1H -Atome in einem Abstand von 0,24 nm vom $^1\text{H}_N$.

Faßt man die Ergebnisse der Linienbreitensimulationen dieses Abschnittes und der Abschnitte 2.2.1.1 und 2.2.1.2 zusammen, so läßt sich Folgendes sagen: Die ^{15}N - und die $^1\text{H}_N$ -Linienbreite gehorchen demselben funktionellen Zusammenhang. Der Unterschied zwischen ^1H - ^{15}N -TROSY- und ^1H - ^{15}N -HSQC-Experiment bezüglich der Linienbreite hängt von der Größe der Werte für $\Delta\sigma$ und θ ab und differiert für ^{15}N und $^1\text{H}_N$. Er ist am größten, wenn die Linienbreite im ^1H - ^{15}N -TROSY-Experiment ein Minimum aufweist. Für ^{15}N bzw. $^1\text{H}_N$ ist dies der Fall, wenn $d^2 + c_N^2 + dc_X \cdot (3 \cos^2 \theta - 1) \approx 0$ (wobei X für N bzw. H steht) ist.¹² Aus [Fus98] und [Tja97b] geht hervor, daß θ und $\Delta\sigma$ für jedes $^1\text{H}_N$ - ^{15}N -Paar innerhalb eines Proteins unterschiedlich sein können. Das bedeutet, daß das Minimum der Linienbreite für jedes einzelne $^1\text{H}_N$ und ^{15}N eines Proteins bei unterschiedlich hohem Magnetfeld liegen kann. Für menschliches Ubiquitin¹³ gilt, daß das Produkt aus dem gyromagnetischen Verhältnis γ und der Anisotropie der chemischen Verschiebung $\Delta\sigma$ für $^1\text{H}_N$ über alle Aminosäurereste gemittelt geringer ist als für ^{15}N . Das führt dazu, daß im ^1H - ^{15}N -TROSY-Experiment das äußere Magnetfeld, welches nötig ist für das Zustandekommen eines Minimums der Linienbreite, für ^1H , auf das gesamte Protein bezogen, höher liegt als für ^{15}N (siehe Abbildung 2.6 und 2.8).

Erweitert man die Fragestellung der Linienbreiten vom reinen $^1\text{H}_N$ - ^{15}N -System auf ein $^1\text{H}_N$ - ^{15}N -System mit Wasserstoffatomen in seiner Umgebung, so zeigt sich für die $^1\text{H}_N$ -Linienbreite ein deutlicher Effekt (siehe Abbildung 2.9). Nimmt man drei Protonen in einem Abstand von 0,24 nm vom $^1\text{H}_N$ an, so erhöht sich die $^1\text{H}_N$ -Linienbreite bei 800 MHz Protonenresonanzfrequenz um mehr als 10 Hz für ein kugelsymmetrisches, starres Protein einer Molekülgröße von etwa 25 kDa.¹⁴ Diese starke Abhängigkeit der $^1\text{H}_N$ -Linienbreite von seiner Umgebung, wird im folgenden Abschnitt (2.2.2) systematisch untersucht.

In allen hier gezeigten Simulationen der Linienbreite wurde Kugelsymmetrie des Proteins, aus dem das $^1\text{H}_N$ - ^{15}N -System stammt, angenommen. In [Fus99] wurde unter anderem

¹²Hierbei ist angenommen, daß $\Delta\sigma_N < 0$ und $\Delta\sigma_H > 0$ ist.

¹³Es ist anzunehmen, daß eine Verallgemeinerung auf andere Proteine sinnvoll ist.

¹⁴Die Annahme der Starrheit und Kugelsymmetrie ist unter anderem die Voraussetzung dafür, aus τ_c auf die Molekülgröße schließen zu können.

die ^{15}N -Linienbreite im ^1H - ^{15}N -TROSY-Experiment für einen axialsymmetrischen Diffusionstensor untersucht. Dort wurde gezeigt, daß der Einfluß von θ , mit $\theta > 0$, auf die ^{15}N -Linienbreite proportional zur Anisotropie der Diffusion ist.¹⁵

Neben dem hier untersuchten zweidimensionalen TROSY-Experiment gibt es viele der dreidimensionalen Experimente, bei denen der Magnetisierungstransfer mit dem $^1\text{H}_\text{N}$ -Kern beginnt (und häufig auch endet) als „TROSY-Versionen“. Im Bezug auf das in dieser Arbeit untersuchte Protein bedeutet dies, daß die Untersuchung des Protein-DNA-Komplexes (Molekulargewicht etwa 25 kDa) durchaus für die Untersuchung mit den „TROSY-Versionen“ der gängigen heteronuklearen dreidimensionalen Experimente in Frage kommt. Für die Untersuchungen des Proteins alleine (Molekulargewicht etwa 15,4 kDa) sollten die herkömmlichen Versionen ausreichend sein.

2.2.2 Deuterierung

Wie im Abschnitt 2.2.1.3 (siehe z.B. Abbildung 2.9) gezeigt wurde, besitzen dem $^1\text{H}_\text{N}$ -Kern benachbarte ^1H -Atome großen Einfluß auf die $^1\text{H}_\text{N}$ -Linienbreite. Im letzten Abschnitt wurde beispielhaft der Einfluß für einige Anordnungen benachbarter ^1H -Atome auf die Linienbreite simuliert. In diesem Abschnitt soll bei einer gegebenen Proteinstruktur der Linienbreitenzuwachs für einzelne $^1\text{H}_\text{N}$ - ^{15}N -Systeme, hervorgerufen durch die benachbarten ^1H -Atome, berechnet werden. Wie oben erwähnt, ist es möglich Proteine mit verschiedenen Deuterierungsgraden herzustellen. Glukose, die als Ausgangsstoff bakterieller Proteinbiosynthese bei der Überexpression des Zielproteins Verwendung findet, ist kommerziell meist mit Deuterierungsgraden von 50, 70 bzw. 100% erhältlich. Aufgrund der Komplexität der Aminosäurebiosynthese ist, selbst bei exakter Kenntnis des Deuterierungszustandes jeder H-Position der Glukose, nicht vorherzusagen, welchen Zustand eine bestimmte H-Position im Protein einnimmt. Um trotzdem quantitative Aussagen über den Einfluß der Deuterierung auf die $^1\text{H}_\text{N}$ -Linienbreite machen zu können, ist es nötig, weitere Vorüberlegungen bezüglich einer numerischen Behandlung der Fragestellung anzustellen.

Bei gegebener Anzahl von nicht mit dem Lösungsmittel austauschenden H-Atomen n und dem Grad der Deuterierung d ($d = 0$ für 100%-ige Besetzung mit ^1H -Atomen, $d = 1$

¹⁵Sie wird definiert als $D_{\parallel}/D_{\perp} - 1$.

für 100%-ige Deuterierung) gilt für die Anzahl der Möglichkeiten¹⁶ eine $(d \cdot n)$ -elementige Teilmenge¹⁷ aus der Grundmenge n auszuwählen:

$$\binom{n}{d \cdot n} = \frac{n!}{(d \cdot n)! (n \cdot (1 - d))!} \quad (2.11)$$

Mit der Stirlingschen Näherungsformel läßt sich $n!$ für $n \geq 18$ mit einem Fehler $< 0,001\%$ abschätzen. Es gilt:

$$n! \approx \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n \cdot \left(1 + \frac{1}{12n}\right) \quad (2.12)$$

Ein Protein mit beispielsweise 75 Aminosäureresten weist im allgemeinen über 400 nicht mit dem Lösungsmittel austauschende H-Atome auf. Damit und mit einem Deuterierungsgrad von 0,5 ergeben sich aus Gleichung (2.11) unter Zuhilfenahme der Näherung (2.12) mehr als 10^{119} Möglichkeiten, 400 Positionen zu jeweils 50% mit ^1H - und ^2H -Atomen zu besetzen.

Auf Grund dieser sehr großen Anzahl an Möglichkeiten scheidet ein numerischer Ansatz aus, in dem die Zunahme der $^1\text{H}_N$ -Linienbreite eines bestimmten Amidprotons für alle möglichen Konstellationen der Umgebung berechnet und anschließend gemittelt wird. Der hier verfolgte Ansatz macht sich die geringe Reichweite der DD-Wechselwirkung in den Gleichungen (2.9) und (2.10) zunutze. Unter der Annahme, daß alle H-Positionen mit gleicher Wahrscheinlichkeit d deuteriert vorliegen, werden in einem ersten Schritt die H-Positionen zufällig mit der Wahrscheinlichkeit d bzw. $1 - d$ mit ^2H - bzw. ^1H -Atomen besetzt. In einem zweiten Schritt wird der Linienbreitenzuwachs, gegeben durch die Gleichungen (2.9) und (2.10) für diese spezielle Konfiguration berechnet. Diese beiden Schritte werden nun N -mal wiederholt, und abschließend ergibt sich die Linienbreite durch Mittelung der N einzelnen Beiträge. Dies wurde mit Hilfe eines selbst entwickelten Computerprogramms realisiert. Als Modellprotein diente ralGEF, dessen Struktur gut bekannt ist [Gey97]. ralGEF ist wohlgefaltet: Es besteht aus einem fünfsträngigen β -Faltblatt und zwei α -Helices. Es wurden nun sechs Aminosäurereste betrachtet, drei aus α -Helices und drei aus dem β -Faltblatt. In Abbildung 2.10 ist die Linienbreite (Summe aus Gleichung

¹⁶unter der Annahme von gleicher Deuterierungswahrscheinlichkeit aller n

¹⁷Sollte $(d \cdot n)$ nicht ganzzahlig sein, so ist das Produkt durch die am nächsten liegende natürliche Zahl zu ersetzen.

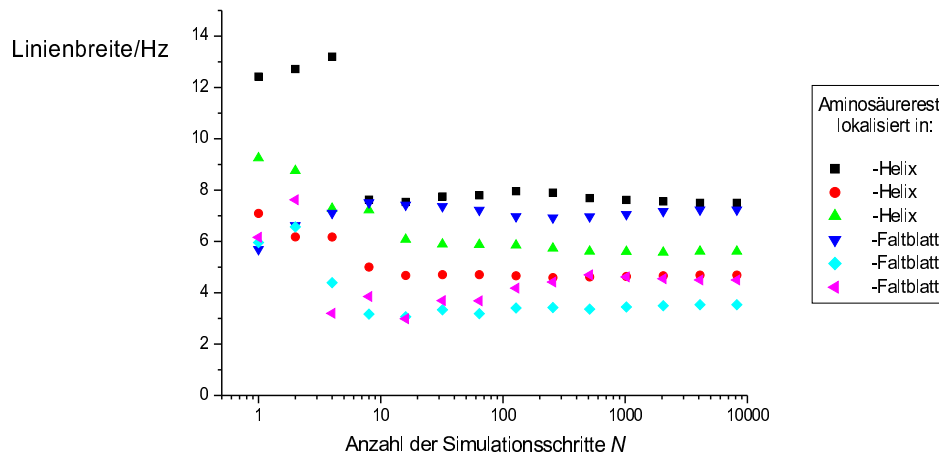


Abbildung 2.10: Simulation des Linienbreitenwachses (Summe aus Gleichung (2.9) und (2.10)) der 1H_N -Linienbreite in Hertz in Abhängigkeit von der Anzahl der Simulationsschritte für sechs Aminosäurereste des Proteins *ralGEF*. Für den Deuterierungsgrad d wurde 0,5 gewählt.

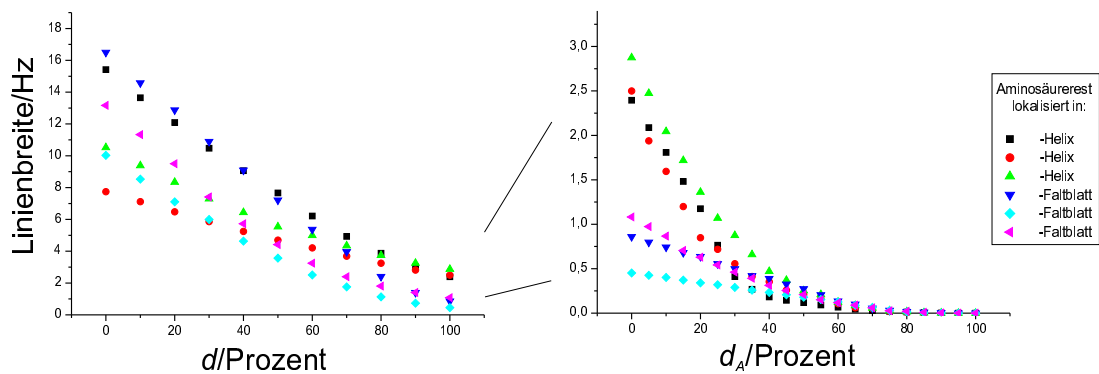


Abbildung 2.11: Simulation des Linienbreitenwachses (Summe aus Gleichung (2.9) und (2.10)) der 1H_N -Linienbreite in Hertz in Abhängigkeit vom Deuterierungsgrad für sechs Aminosäurereste des Proteins *ralGEF*. Links: Simulationen bei variiertem Deuterierungsgrad d der nicht mit dem Lösungsmittel (H_2O) austauschenden H-Positionen. Rechts: Simulationen bei variiertem Deuterierungsgrad d_A der mit dem Lösungsmittel austauschenden H-Positionen.

(2.9) und (2.10)) in Abhängigkeit von der Zahl der Simulationschritte N aufgetragen. Für den Deuterierungsgrad d wurde 0,5 gewählt. Wie bisher auch wurde für $\tau_c = 10$ ns gesetzt. Weiterhin wurde von einem kugelsymmetrischen, starren Protein ausgegangen und die DD-Wechselwirkung zwischen dem $^1\text{H}_N$ -Kern und benachbarten ^2H -Atomen, auf Grund des im Vergleich zum ^1H -Kerns etwa fünf mal geringeren gyromagnetischen Verhältnisses des ^2H -Kerns, vernachlässigt¹⁸. Bereits nach etwa 128 Simulationsschritten unterscheiden sich die Linienbreiten bei allen sechs Aminosäureresten nicht mehr als 5% vom Wert bei 8192 Simulationsschritten. Für die nachfolgenden Simulationen wurde $N = 1000$ gewählt.

Der linke Teil von Abbildung 2.11 zeigt die Linienbreite ($\Delta\nu_1(^1\text{H}) + \Delta\nu_2(^1\text{H})$) in Abhängigkeit vom Deuterierungsgrad d für die oben verwendeten sechs Aminosäurereste. Es zeigt sich, daß die Linienbreite für alle sechs Aminosäurereste unabhängig vom Wert der Linienbreite bei $d = 0$ auf unter 3 Hz für $d = 1$ sinkt. Es fällt weiterhin auf, daß bei den gewählten sechs Aminosäureresten für $d = 1$ die Linienbreiten der drei Aminosäurereste aus α -helikalen Bereichen etwa gleich und um mehr als einen Faktor 2 höher sind als die Linienbreiten der drei Aminosäurereste aus β -Faltblattstrukturen, die wiederum ähnliche Linienbreiten aufweisen. Die Restlinienbreite bei $d = 1$ rührt von den H-Atomen her, die mit dem Lösungsmittel austauschen. Unter der Annahme, daß nach erfolgtem Austausch keine weitere Wechselwirkung mit dem Lösungsmittel stattfindet, wurde die Linienbreite bei $d = 1$ für variierenden Deuterierungsgrad der mit dem Lösungsmittel austauschenden H-Atome (d_A) simuliert (Abbildung 2.11 rechts). Es zeigt sich, daß bei einem Deuterierungsgrad d_A von etwa 0,4 die Linienbreite auf unter 0,5 Hz sinkt und daß sich die Linienbreiten der sechs Aminosäurereste kaum noch unterscheiden.

Im Bezug auf das in dieser Arbeit untersuchte Protein bedeutet dies, daß bei der Untersuchung des Protein-DNA-Komplexes (Molekulargewicht etwa 25 kDa) durchaus die Verwendung einer zumindest teilweise deuterierten Probe sinnvoll ist. Für die Untersuchungen des Proteins allein (Molekulargewicht etwa 15,4 kDa) erscheint Deuterierung, besonders in Hinblick auf die Einschränkung der zur Verfügung stehenden Experimente bei Verwendung einer deuterierten Probe, weniger sinnvoll.

¹⁸Diese Annahmen wurden auch für die Simulationen mit variierendem Deuterierungsgrad (Abbildung 2.11) beibehalten.

2.3 Probenpräparation

Die Probenherstellung und -reinigung erfolgte am Universitäts-Krankenhaus Eppendorf (Hamburg) in der Arbeitsgruppe von Herrn Prof. W. H. Strätling. Die einzelnen Schritte der Herstellung (Überexpression in *E.coli*-Zellen) sind detailliert in [Kri91], [Wei97] und [Bru00b] beschrieben. Zur Aufreinigung über einer Nickel-Affinitätssäule diente ein Polyhistidin-*tag* der Form (MGHHHHHH-) am N-Terminus von cMeCP2 ARBD. Die vollständige Sequenz der Domäne ist in Abschnitt 1 aufgeführt. Bei allen Proben (nicht isotopenmarkiert, ^{15}N -, ^{15}N - ^{13}C - und ^2H - ^{15}N - ^{13}C -isotopenmarkiert)¹⁹ wurde eine Proteinausbeute von 1,8-2,6 mg pro Liter Wachstumsmedium erreicht. Damit war es möglich, die Messungen an Proben mit Konzentrationen größer 1 mM durchzuführen. Im allgemeinen wurden die NMR-Spektroskopie-Messungen am Protein in einem Probenvolumen von 0,5 ml durchgeführt. War nicht ausreichend Protein vorhanden, so wurde auf NMR-Röhrchen (Shigemi, Tokyo) mit einem Minimalfüllvolumen von 0,3 ml zurückgegriffen.

NMR-Messungen an cMeCP2 ARBD wurden, wenn dies nicht anders vermerkt ist, in einem wäßrigen Puffer folgender Zusammensetzung durchgeführt: 154 mM Natriumchlorid, 10 mM Natriumphosphat, 0,1 mM Natriumacid, 0,1 mM DSS.²⁰ Die gleichen Pufferbedingungen galten für die Messungen am cMeCP2 ARBD-DNA-Komplex, mit dem Unterschied, daß die Natriumchloridkonzentration bei 10 mM lag; bei späteren Messungen wurde sie auf 50 mM erhöht (siehe auch Abschnitt 3.2). Bei den Untersuchungen des Protein-DNA-Komplexes kamen drei verschiedene DNA-Konstrukte zum Einsatz (siehe Abbildung 2.12). Bei Probe 3 wurde aus Kostengründen auf den *hairpin* verzichtet. In Abschnitt 3.2 wird gezeigt, daß dies keinen Einfluß auf das Bindungsverhalten mit dem Protein hat. Die Konstrukte wurden von der Firma MWG AG BIOTECH (Ebersberg) bezogen, wobei für Probe 3 ausgehend von einzelsträngiger DNA durch mehrfaches Aufwärmen und langsames Abkühlen die Strangpaarung erzielt wurde. Der Nachweis des Doppelstrangs erfolgte elektrophoretisch an Polyacrylamidgelen. Ausgehend von einem 1:1 Komplex zwischen cMeCP2 ARBD und der beschriebenen DNA [Wei97] wurde ein Mischungsverhältnis von etwa 1:1,6 (Protein zu DNA) gewählt, um ein Überangebot an

¹⁹Bei der letztgenannten Probe betrug der Deuterierungsgrad 70%.

²⁰DSS - (4,4-Dimethyl-4-Silapentan-Sulfonsäure)

stellung der Spektrometer über eine externe Referenz (Ethylenglykol [Rai79]) kontrolliert wurde. Die Referenzierung der chemischen Verschiebungen der Protonen erfolgte unter Zuhilfenahme eines internen Standards.²³ Die Eichung der chemischen Verschiebungen der ¹⁵N- und ¹³C-Kerne erfolgte indirekt gemäß der IUPAC-Richtlinien [Mar98], [Wis95a]. Alle Spektren wurden mit dem Computerprogramm XWINNMR Version 2.6 aufgenommen und prozessiert²⁴. Die Auflösung in den indirekten Dimensionen wurde durch Extrapolation (*linear forward prediction, zero filling*) des freien Induktionsabfalls (*free induction decay* (FID)) gegenüber der Auflösung, die durch die Anzahl der aufgenommenen Datenpunkte gegeben ist, erhöht. Dreidimensionale Spektren wurden in der Regel mit 1024×64×128 Datenpunkten aufgenommen, abweichend davon waren nur die Aufnahmebedingungen beim HBHA(CBCACO)NH-Experiment, dort wurden 1024×64×194 Datenpunkte aufgenommen. Artefaktsignale wurden durch digitale Filterung des FIDs reduziert. Die Visualisierung und Auswertung der Spektren erfolgte mit den Programmen AURELIA 2.7.10 (Bruker, Karlsruhe) [Nei95] und AUREMOL (Bruker, Karlsruhe) [Gan00].

2.4.2 Zuordnung der chemischen Verschiebungen

Die Strukturbestimmung von cMeCP2 ARBD beruht hauptsächlich auf Abstandsinformationen zwischen ¹H-Kernen (ermittelt durch NOE-Kontakte (Abschnitt 2.4.3)) und der Kenntnis der Orientierung von Kernverbindungsvektoren bezüglich eines molekülfesten Koordinatensystems (ermittelt durch Restdipolkopplungen (Abschnitt 2.4.4)). In Abbildung 2.13 sind die beiden Methoden schematisch gegenübergestellt. Um diese Methoden anwenden zu können, ist die Kenntnis eines Großteils der chemischen Verschiebungen der Spin-1/2-Kerne des betrachteten Proteins Voraussetzung. Auf Grund der unterschiedlichen Umgebung eines jeden Kerns wird das von außen angelegte Magnetfeld unterschiedliche stark abgeschirmt bzw. verstärkt. Die chemischen Verschiebungen reagieren dabei sensitiv auf ihre elektronische Umgebung, so daß gleiche Kernspins mit unterschiedlicher Umgebung verschiedene chemische Verschiebungen aufweisen. Die Größe der chemischen Verschiebungen wird auf einen internen Standard referenziert und meist in *parts per mil-*

²³ Als interner Standard wurde DSS verwendet (siehe 2.3).

²⁴ Die Prozessierung beinhaltet neben der FT auch die richtige Phasierung und eine Basislinienkorrektur.

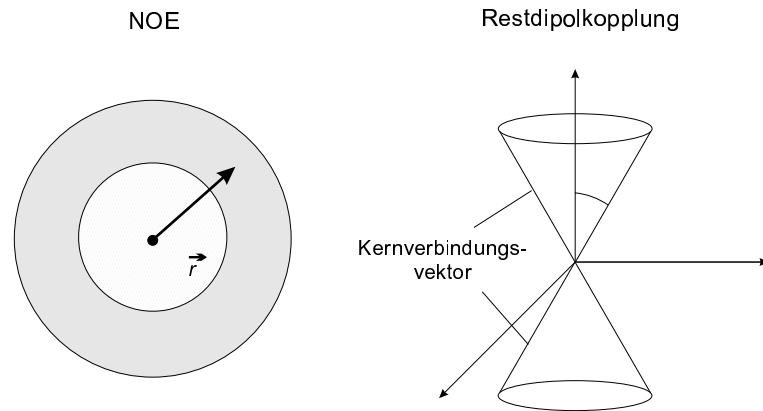


Abbildung 2.13: Schematische Darstellung von NOEs und Restdipolkopplungen als sogenannte *restraints* bzw. *Einschränkungen* für Strukturbestimmungen. NOE-*restraints* legen den Abstand zweier Protonen innerhalb einer Kugelschale fest, während *restraints* durch Restdipolkopplungen die Lage des Kernverbindungsvektors auf zwei Kegelmäntel beschränken.

tion (ppm) angegeben, für die chemischen Verschiebungen δ in ppm gilt:

$$\delta = \frac{\omega - \omega_{ref}}{\omega_0} \cdot 10^6 \quad (2.13)$$

ω ist die Resonanzfrequenz des betrachteten Kerns, ω_{ref} die Resonanzfrequenz der Referenzverbindung und ω_0 die Trägerfrequenz. Eine weitere Voraussetzung für die Bestimmung der Struktur eines Biomoleküls ist die schon in Abschnitt 2.2 erwähnte FT-Spektroskopie, welche die Grundlage der mehrdimensionalen Spektroskopie darstellt, die wiederum von entscheidender Bedeutung für die Auflösung der stark überlappenden Signale in eindimensionalen Spektren von Biomolekülen ist.

Die Strategie zur Bestimmung der Struktur von cMeCP2 ARBD einschließlich der Zuordnung der chemischen Verschiebungen der Kerne der Haupt- und Seitenkette wird in Abschnitt 3.1 beschrieben. Die dafür erforderlichen NMR-Experimente sind in diesem und den folgenden Abschnitten (2.4.2 - 2.4.4) prinzipiell beschrieben. Eine detailliertere Beschreibung vieler der zur Anwendung gekommen Experimente findet sich z.B. in [Cav95] und [Sat99].

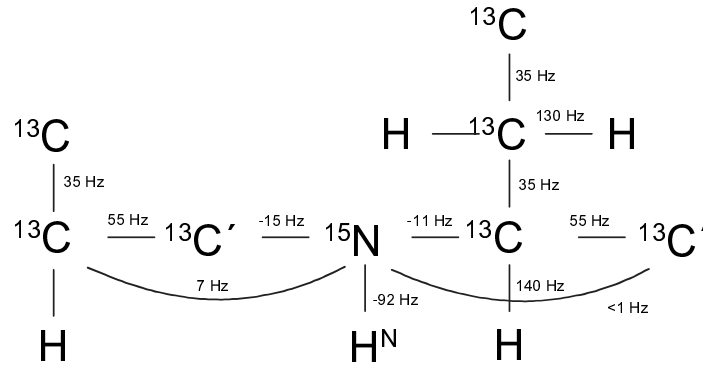


Abbildung 2.14: Skalare 1J - und 2J -Kopplungen im Rückgrat und der Seitenkette eines ^{15}N - und ^{13}C -isotopenmarkierten Proteins [Sat99].

2.4.2.1 Experimente zur Zuordnung der Signale der Kerne der Hauptkette

Die bis Anfang der 90er Jahre übliche Zuordnungsmethode, basierend auf der kombinierten Auswertung von TOCSY-, NOESY- und COSY-Experimenten²⁵ ist für Biomoleküle wie cMeCP2 ARBD auf Grund ihrer Größe nicht möglich. Es ist nötig, mehrdimensionale heteronukleare Experimente an ^{15}N - und ^{13}C -markierte Proben durchzuführen. Diese seit Anfang der 90er Jahre entwickelten Experimente korrelieren aufeinanderfolgende Aminosäurereste. Das Prinzip dieser Experimente beruht auf den großen 1J -Kopplungskonstanten zwischen ^1H -, ^{15}N - und ^{13}C -Kernen, die im Gegensatz zu den $^3J_{\text{HH}}$ -Kopplungen bei homonuklearen Messungen weitgehend unabhängig von der Sekundärstruktur sind. Die J -Kopplungen des Proteinrückgrats sind aus Abbildung 2.14 zu entnehmen.

Das in Abschnitt 2.2.1 näher beschriebene HSQC-Experiment stellt ein grundlegendes heteronukleares Korrelationsexperiment dar. Das ^1H - ^{15}N -HSQC-Spektrum von cMeCP2 ARBD ist in Abbildung 3.1 dargestellt. Die oben aufgeführten dreidimensionalen Experimente bauen im Prinzip meist auf dem HSQC-Experiment auf. Die Erweiterung auf die dritte Dimension wird durch einen weiteren INEPT-Schritt erreicht. Beispielhaft soll das

²⁵Dabei steht TOCSY, NOESY bzw. COSY für Total Correlation Spectroscopy, Nuclear Overhauser Effect Spectroscopy bzw. Correlation Spectroscopy.

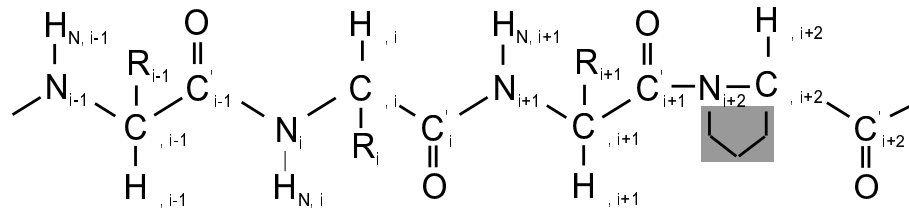


Abbildung 2.15: Schematische Darstellung eines Abschnitts des Rückgrats eines Proteins. R_i bezeichnet die Seitenkette des i -ten Aminosäurerests, Aminosäurerest $i+2$ stellt einen Prolinrest dar (die Seitenkette ist schematisch dargestellt und grau unterlegt), $i-1$, i und $i+1$ sind andere Aminosäurereste.

HNCA-Experiment näher betrachtet werden. Ausgangspunkt ist das in Abschnitt 2.2.1 beschriebene zweidimensionale ^1H - ^{15}N -HSQC-Experiment. Nachdem die Magnetisierung vom $^1\text{H}_N$ -Kern auf den ^{15}N -Kern transferiert wurde, erfolgt ein weiterer INEPT-Schritt mit $\tau = \frac{1}{4J}$, wobei J hier die -11 Hz große Kopplung zwischen dem ^{15}N -Kern und dem $^{13}\text{C}_\alpha$ -Kern ist. Dort entwickelt sich die Magnetisierung unter dem Einfluß des effektiven Magnetfeldes und wird mittels INEPT-Schritt wieder zurück auf den ^{15}N -Kern transferiert. Da die Kopplung des ^{15}N -Kerns zum $^{13}\text{C}_\alpha$ -Kerns des sequentiellen Vorgängers mit 7 Hz betragsmäßig nahe bei -11 Hz liegt, erfolgt neben dem Transfer zum eigenen $^{13}\text{C}_\alpha$ -Kern auch ein Transfer zum Vorgänger. Die Intensität dieses Signals ist allerdings schwächer als die des Signals des eigenen $^{13}\text{C}_\alpha$ -Kerns, da τ für die intraresiduale Kopplung von -11 Hz optimiert ist. Das bedeutet, daß man im Spektrum bei der chemischen Verschiebung des ^{15}N - und des $^1\text{H}_N$ -Kerns des i -ten Aminosäurerests die chemische Verschiebung des $^{13}\text{C}_\alpha$ -Kerns des i -ten Aminosäurerests und - mit einer schwächeren Signalintensität - die chemische Verschiebung des $^{13}\text{C}_\alpha$ -Kerns des $(i-1)$ -ten Aminosäurerests vorfindet.

Damit ist das Prinzip beschrieben, welches es erlaubt, benachbarte Aminosäurereste zu identifizieren. Dies ist beispielhaft für einen kurzen Sequenzabschnitt in Abbildung 3.2 dargestellt. Betrachtet man Abbildung 2.15, so erkennt man, daß durch Prolinreste die Sequenzierung unterbrochen wird, da am Amidstickstoff eines Prolinrests kein ^1H -Atom gebunden ist. Weiterhin kommt es auf Grund der Größe von cMeCP2 ARBD zu Signalüberlagerungen, die eine lückenlose sequentielle Zuordnung allein auf der Grundlage des HNCA-Experiments unmöglich machen. Die Zuhilfenahme weiterer heteronuklearer

dreidimensionaler Experimente war daher zwingend erforderlich. Dies waren im einzelnen das CBCA(CO)NH- und das HNCO-Experiment. Die Namen der Experimente spiegeln den Magnetisierungstransfer wider (eingeklammert werden die Kerne, über die Magnetisierung zwar transferiert wird, auf denen aber keine Entwicklung der Magnetisierung stattfindet).²⁶

Das HNCO-Experiment beginnt ähnlich wie das HNCA-Experiment, wobei anstelle des Magnetisierungstransfers vom ^{15}N -Kern des i -ten Aminosäurerests auf die $^{13}\text{C}_\alpha$ -Kerne ein Transfer der Magnetisierung auf den $^{13}\text{C}'$ -Kern des $(i-1)$ -ten Aminosäurerests stattfindet. Dort entwickelt sich die Magnetisierung unter dem Einfluß des effektiven Magnetfeldes und wird anschließend auf dem gleichen Weg zurück zum $^1\text{H}_N$ -Kern transferiert, auf welchem dann die Detektion stattfindet. Der Transfer der Magnetisierung zwischen den Kernen erfolgt wie beim HSQC- und HNCA-Experiment mittels INEPT-Schritten.

Der Magnetisierungstransfer beim CBCA(CO)NH-Experiment erfolgt wie bei den bisher beschriebenen Experimenten ebenfalls mittels INEPT, der Weg unterscheidet sich allerdings. Die Anregung erfolgt simultan auf den ^1H -Kernen, die am $^{13}\text{C}_\beta$ -Kern des $(i-1)$ -ten Aminosäurerests gebunden sind. Der Transfer erfolgt über den $^{13}\text{C}_\beta$ -Kern auf den $^{13}\text{C}_\alpha$ -Kern. Von dort wird die Magnetisierung über den $^{13}\text{C}'$ -Kern und den ^{15}N -Kern auf den $^1\text{H}_N$ -Kern übertragen, mit anschließender Detektion auf diesem Kern. Dieses Experiment stellt im Gegensatz zu den bisher beschriebenen Experimenten kein *out and back*-Experiment dar. Das Spektrum zeigt bei der chemischen Verschiebung des $^1\text{H}_N$ - und des ^{15}N -Kerns des i -ten Aminosäurerests die chemischen Verschiebungen des $^{13}\text{C}_\alpha$ - und des $^{13}\text{C}_\beta$ -Kerns des $(i-1)$ -ten Aminosäurerests.

Weiterhin war auch das HBHA(CBCACO)NH-Experiment nicht nur für die Zuordnung der chemischen Verschiebungen der Seitenkettenkerne,²⁷ sondern auch für die Absicherung der sequentiellen Zuordnung sehr hilfreich. Ein weiteres, bei der Zuordnung häufig eingesetztes Spektrum, das CBCANH-Experiment, wurde auf Grund seiner geringen Empfindlichkeit für Proteine größer als 15 kDa nicht verwendet [Sat99].

In diesem Abschnitt erfolgte eine prinzipielle Beschreibung der Experimente. Die in der

²⁶Bei sogenannten *out and back*-Experimenten, wie z.B. dem HNCA-Experiment gibt der Name nur den „Hinweg“ des Magnetisierungstransfers an.

²⁷In Abschnitt 2.4.2.2 wird das HBHA(CBCACO)NH-Experiment näher erklärt.

Praxis eingesetzten Experimente verwenden üblicherweise zusätzlich noch Techniken zur Artefakt- und Lösungsmittelunterdrückung. Details hierzu sind aus [Cav95] und [Sat99] zu entnehmen.

Die eigentlich Strategie der sequentiellen Zuordnung mit den resultierenden Ergebnissen ist in Abschnitt 3.1.1 dargestellt.

2.4.2.2 Experimente zur Zuordnung der Signale der Kerne der Seitenkette

Nach erfolgter Zuordnung der chemischen Verschiebungen der Hauptkettenkerne ist die Zuordnung der Seitenkettenkerne der letzte Schritt vor der Bestimmung von NOE-*restraints* mittels NOESY-Experimenten. Für diese Aufgabe eignen sich im allgemeinen TOCSY-Experimente. Ausgehend von den durch die sequentielle Zuordnung identifizierten $^1\text{H}_N$ - und ^{15}N -Kernen werden Protonen eines Aminosäurerests, die untereinander skalar gekoppelt sind, detektiert. Es zeigte sich bei Messungen an cMeCP2 ARBD, daß die Magnetisierung praktisch nur von den $^1\text{H}_\alpha$ -Kernen auf die $^1\text{H}_N$ -Kerne übertragen wurde. Eine Erhöhung der Mischzeit zwecks Verbesserung des Transfers der Seitenkettenprotonen hatte vornehmlich eine Verschlechterung des Signal/Rausch-Verhältnisses zur Folge. Insgesamt war der Informationsgehalt der aufgenommenen TOCSY-Spektren (homonukleares TOCSY und ^{15}N -editiertes TOCSY) sehr beschränkt. Diese Tatsache ist für Proteine von der Größe des cMeCP2 ARBD nicht ungewöhnlich. Zur weiteren Zuordnung der chemischen Verschiebungen der Seitenkettenkerne wurde ein HCCH-TOCSY-Experiment aufgenommen. In diesem Experiment wird die Magnetisierung durch einen INEPT-Schritt ausgehend von den ^1H -Kernen der Seitenkette über die große 1J -Kopplung zwischen den ^1H - und ^{13}C -Kernen transferiert. Durch isotropes Mischen wird dann die Magnetisierung entlang der ^{13}C -Kerne übertragen und anschließend zurück transferiert (INEPT). Die Detektion erfolgt dann auf den ^1H -Kernen. Zur Reduktion des Einflusses des Lösungsmittels wurde das HCCH-TOCSY-Experiment in 99% $^2\text{H}_2\text{O}$ aufgenommen. Das Spektrum ist ^{13}C -editiert und der Informationsgehalt ist überbestimmt, was besonders hilfreich ist bei der Zuordnung der chemischen Verschiebungen der zu langkettigen Aminosäureresten (wie z.B. Isoleucinresten) gehörigen Kerne. Die Diagonalsignale bei verschiedenen ^{13}C -Verschiebungen werden von ^1H -Kernen gebildet, die an das jeweilige ^{13}C -Atom gebunden sind. Kreuzsignale treten in der indirekten Dimension zu

über die ^{13}C -Kerne skalar gekoppelten ^1H -Kernen des selben Aminosäurerests auf. Auf Grund der nur teilweise zufriedenstellenden Auswertbarkeit dieses Spektrums²⁸ besonders für die Identifizierung der Seitenkettenkerne der Prolinreste wurden zusätzlich das HBHA(CBCACO)NH- und das CC(CO)NH-Experiment durchgeführt. Das erstgenannte Experiment ist dem CBCA(CO)NH-Experiment ähnlich, wobei nicht die $^{13}\text{C}_\alpha$ - und $^{13}\text{C}_\beta$ -Kerne, sondern deren Protonen detektiert werden. Dieses Spektrum liefert die chemischen Verschiebungen der $^1\text{H}_\alpha$ - und $^1\text{H}_\beta$ -Kerne und stellt eine weitere Möglichkeit dar, die sequentielle Zuordnung zu überprüfen und gegebenenfalls zu verbessern. Das zweitgenannte Spektrum liefert ähnlich dem CBCA(CO)NH-Experiment die chemischen Verschiebungen der $^{13}\text{C}_\alpha$ - und $^{13}\text{C}_\beta$ -Kerne. Zusätzlich lassen sich mit diesem Experiment aber auch die chemischen Verschiebungen der anderen Kohlenstoffe der Seitenkette (falls vorhanden) bestimmen.

2.4.3 Experimente zur Bestimmung von Kern-Kern-Abständen

2.4.3.1 NOESY-Experimente

Die Strukturbestimmung von cMeCP2 ARBD basiert im wesentlichen auf der Auswertung von NOESY-Spektren. Sind zwei Protonen weniger als etwa 0,5 nm voneinander entfernt, so läßt sich der Polarisationstransfer zwischen beiden Spins nachweisen. Dieser Magnetisierungstransfer (*Nuclear Overhauser Effect* (NOE)) erfolgt durch direkte magnetische Dipol-Dipol-Wechselwirkung und nicht über skalare Kopplungen wie in den bisher besprochenen Experimenten. Behandelt man die Spins als isoliert und geht nur von einer direkten Wechselwirkung jeweils zweier Spins aus,²⁹ so läßt sich ein mathematischer Ausdruck für den Magnetisierungsaufbau des zweiten Spins ableiten. Es zeigt sich, daß dieser Effekt stark vom Abstand der beiden beteiligten Protonen abhängt. Die Intensität der gemessenen Kreuzsignale I_{12} ist unter den oben gemachten Annahmen durch die Proportionalitätsbeziehung

$$I_{12} \propto r_{12}^{-6} \quad (2.14)$$

²⁸siehe auch Abschnitt 3.1.2

²⁹Sogenannte „Spindiffusion“ wird ausgeschlossen.

gegeben. Diese Proportionalität gilt für alle Protonenpaare eines Moleküls in gleicher Weise, wenn man von einem starren Molekül ausgeht, also keine interne Beweglichkeit vorliegt.³⁰ Mit dieser Beziehung können durch Intensitätsmessungen der jeweiligen Kreuzsignale den Protonenpaaren Abstände zugeordnet werden. Die Eichung dieser Abstände erfolgt z.B. anhand von Signalen mit bekanntem Kern-Kern-Abstand.³¹ Neben einem zweidimensionalen NOESY-Experiment, das alle Protonen des Proteins, sofern sie nicht zu weit räumlich voneinander getrennt liegen, miteinander korreliert, wurden auch ¹⁵N- bzw. ¹³C-editierte NOESY-Experimente durchgeführt. Die Erweiterung des zweidimensionalen NOESY-Experiments auf eine dritte Dimension durch Magnetisierungstransfer auf die Heterokerne und Entwicklung auf diesen führt zur Aufhebung vieler Signalüberlagerungen gegenüber dem zweidimensionalen Experiment. Die Strategie zur Zuordnung der NOE-Signale aus diesen Spektren ist in Abschnitt 3.1.6 dargestellt.

2.4.3.2 Experimente zum Nachweis von Wasserstoffbrückenbindungen

Für die Ausbildung der Proteinstruktur (Proteinfaltung) sind Wasserstoffbrückenbindungen von besonderer Bedeutung. Sie können mit Hilfe der NMR-Spektroskopie direkt nachgewiesen werden, da beispielsweise zwischen den an der Wasserstoffbrückenbindung beteiligten Atomen eine skalare Kopplung besteht. Diese ³J_{NC'}-Kopplung ist meist sehr klein (< 2 Hz), aber prinzipiell meßbar. Zum Nachweis der Kreuzsignale wird ein H(N)CO-Experiment mit auf die Größe der ³J_{NC'}-Kopplung angepaßten Transferzeiten aufgenommen. Diese Experimente scheiterten bei Messungen an cMECP2 ARBD; trotz über 1000 aufaddierten FIDs (*scans*) war es nicht möglich, Kreuzsignale zu identifizieren. Es ist anzunehmen, daß auf Grund der langen Transferzeit (66 ms) und der Größe des Proteins die Magnetisierung nahezu vollständig relaxiert.

Eine weitere Methode zum Nachweis von Wasserstoffbrückenbindungen mittels NMR-Spektroskopie beruht auf einem indirekten Verfahren: Eine in H₂O gelöste Probe wird lyophilisiert und anschließend in ²H₂O gelöst. Im direkten Anschluß wird eine Serie

³⁰Das bedeutet, daß für alle Protonen τ_c den gleichen Wert annimmt.

³¹Die Eichung der Abstände der in dieser Arbeit aufgenommenen NOESY-Spektren ist in Abschnitt 3.1.6 aufgeschlüsselt.

schneller³² ^1H - ^{15}N -HSQC-Experimente aufgenommen und der Austausch der Amidprotonen mit dem Lösungsmittel zeitlich verfolgt. Wasserstoffatome, die an Wasserstoffbrückenbindungen beteiligt sind, weisen geringere Austauschraten auf als nicht an Wasserstoffbrückenbindungen beteiligte Wasserstoffatome. Das heißt, die Signale der Wasserstoffatome, die an Wasserstoffbrückenbindungen beteiligt sind, sind im Vergleich zu den anderen nach der Zugabe von $^2\text{H}_2\text{O}$ länger beobachtbar. Dieses Experiment wurde an cMePC2 ARBD bei 298 K durchgeführt. Bereits nach einer Zeit von 30 Minuten waren sämtliche Amidprotonen mit dem Lösungsmittel ausgetauscht. Von einer Wiederholung des Experiments bei einer niedrigeren Temperatur³³ wurde abgesehen, da sämtliche anderen Experimente bei 298 K aufgenommen wurden und bei einer erniedrigten Temperatur die Vergleichbarkeit nicht mehr notwendigerweise gegeben ist.

2.4.4 Experimente zur Messung von Restdipolkopplungen

Neben den Einschränkungen des Konformationsraums des Proteins mittels NOE-Kontakten bilden Einschränkungen durch Restdipolkopplungen eine wichtige Ergänzung. Dieses moderne Verfahren in der Strukturbestimmung von Biomolekülen liefert Informationen über die relative Orientierung von Kernverbindungsvektoren. Räumlich anisotrope Wechselwirkungen (wie z.B. die magnetische Dipol-Dipol-Wechselwirkung) werden bei typischen NMR-Experimenten in isotropen Lösungen zu Null ausgemittelt. Gibt man der Probe Stoffe zu, die sich im äußeren Magnetfeld orientieren, so überträgt sich in Abhängigkeit von der Konzentration des zugegebenen Stoffes ein Teil der Orientierung auf das Protein, das Protein ist partiell orientiert. Seine Umorientierung ist im zeitlichen Mittel nicht mehr isotrop.³⁴ Deshalb mitteln sich räumlich anisotrope Wechselwirkungen wie die magnetische Dipol-Dipol-Wechselwirkung nicht mehr vollständig aus. Die Wechselwirkung des Proteins mit dem orientierenden Medium basiert sowohl auf sterischen wie auf elektrostatischen Wechselwirkungen [Zwe00]. Andere Möglichkeiten, eine partielle molekulare Orientierung in Magnetfeldern zu erzielen, sind:

³²D.h., es ist eine minimale Anzahl an Inkrementen und *scans* zu wählen, ohne die Auflösung bzw. das Signal/Rausch-Verhältnis so zu verschlechtern, daß eine Identifizierung der Signale unmöglich wird.

³³Eine Temperaturerniedrigung würde einen verlangsamten Austausch nach sich ziehen.

³⁴Unter der Voraussetzung, daß das Protein nicht vollständig kugelsymmetrisch ist.

- Externe elektrischer Felder (siehe z.B.[Buc62])
- Direkte Wechselwirkung des Probenmoleküls mit dem äußeren Magnetfeld³⁵ (siehe z.B. [Gay82])

In der vorliegenden Arbeit wurde die partielle Orientierung der Proteinmoleküle durch Einsatz von Bizellsystemen [Tja97] erreicht. Diese Systeme bestehen hauptsächlich aus einer Mischung zweier Phospholipide, welche scheibenförmige Aggregate (Bizellen) ausbilden. Die beiden Lipide weisen unterschiedliche Kettenlängen auf und bei günstig gewählten Mischungsverhältnissen bilden sich die erwähnten Aggregate.³⁶ In [Bru00a] wurde gezeigt, daß der Orientierungsgrad linear mit der Lipidkonzentration in der Probe ansteigt. Das in der vorliegenden Arbeit zum Einsatz gekommene System bestand aus DIODPC/CHAPSO/CTAB [Pet01] [Bam01] in einem molaren Verhältnis von 4,3/1/0,1, wobei der Anteil orientierendem Mediums an der Probe bei 7,5 % [M/V] lag.³⁷ Bei geeigneter Temperatur und höherer Konzentration bildet CTAB selbst eine flüssigkristalline Phase aus und orientiert sich im äußeren Magnetfeld [Rap99]. Die Zugabe von CTAB in niedriger Konzentration führt zu einer Stabilisierung des Bizellensystems [Ott99],[Los98]. Die orientierende Phase eines Systems aus DIODPC/CHAPSO ist bei 298 K und bei pH-Werten zwischen 1 und 6,5 stabil [Cav99]. Messungen an cMeCP2 ARBD wurden durchweg bei 298 K und pH 6,8 durchgeführt. Der pH-Wert liegt damit an der Grenze des Bereichs, in dem das DIODPC/CHAPSO Systems eine orientierende Phase ausbildet. Die Orientierung läßt sich anhand des ²H-NMR-Signals von ²H₂O überprüfen. Die residuale Quadrupolwechselwirkung führt zu einer Aufspaltung der Linie, die in isotroper Lösung nicht auftritt. Die Größe der Aufspaltung spiegelt den Orientierungsgrad wider

³⁵Voraussetzung (bei Messungen an derzeit zugänglichen Magnetfeldern von bis zu etwa 18,8 T) an die Probe ist, daß die zu untersuchenden Moleküle eine relativ große Anisotropie der magnetischen Suszeptibilität aufweisen.

³⁶Die kurzkettige Komponente lagert sich bevorzugt am Rand der Bizelle ein (siehe z.B. [San92]).

³⁷DIODPC steht für 1,2-di-O-Dodecyl-sn-glycero-3-phosphocholin (Avanti Polar Lipids, Alabaster, AL, USA), CHAPSO für 3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propansulfonat (Sigma, Deisenhofen) und CTAB für Cethyltrimethylammoniumbromid (Sigma, Deisenhofen).

[Sal72]. Bei Messungen mit diesem orientierenden Medium wurde unter anderem³⁸ auf Grund des kritischen pH-Bereichs vor und nach der eigentlichen Messung ein solches Deuteriumspektrum aufgenommen. Es zeigte sich, daß die orientierende Phase nicht immer über den Zeitraum der Messung stabil blieb. Allerdings war es möglich, durch mehrere Erwärmungs- und Abkühlungszyklen die Orientierung in der gewünschten Form wieder herzustellen, die dann über den Zeitraum der Messung unverändert blieb. Alternativ stand ein zweites System (DMPC/DHPC) zu Auswahl³⁹, von dem bekannt ist, daß die orientierende Phase in einem pH-Bereich von 2,3 bis 10,4 stabil ist [Tja97]. Allerdings kommt es zur Ausbildung der orientierenden Phase erst bei Temperaturen oberhalb von 303 K. Um temperaturabhängigen Veränderungen der Struktur vorzubeugen und die Struktur des Proteins nicht zu verändern, wurde von Messungen mit DMPC/DHPC abgesehen und mit DIODPC/CHAPSO/CTAB unter ständiger Kontrolle der Stabilität der orientierenden Phase gemessen.

Erfährt ein gelöstes Molekül eine partielle Orientierung, so gilt für die Restdipolkopplung zwischen Spin I und S nach [Bas87]:

$$\langle D_{IS} \rangle = -D_{IS} S_{zz} S \left(\left(\frac{3}{2} \cos^2 \theta - \frac{1}{2} \right) + \frac{\eta}{2} \sin^2 \theta \cos 2\varphi \right) \quad (2.15)$$

Die Winkel θ und φ sind die Polarwinkel, welche die Orientierung des Kernverbindungsvektors \vec{r}_{IS} im Hauptachsensystem des sogenannten molekularen Orientierungstensors S_{ij} festlegen. Die Hauptwerte des Tensors⁴⁰ sind S_{xx} , S_{yy} , S_{zz} , wobei gilt: $|S_{xx}| \leq |S_{yy}| \leq |S_{zz}|$ und $S_{xx} + S_{yy} + S_{zz} = 0$. Der Asymmetrieparameter η ist gegeben durch $\frac{S_{xx} - S_{yy}}{S_{zz}}$ und für die statische Dipolkopplungskonstante D_{IS} gilt:

$$D_{IS} = \frac{1}{\pi} \frac{\mu_0}{4\pi} \hbar \gamma_I \gamma_S \langle r_{IS}^{-3} \rangle \quad (2.16)$$

In Tabelle 2.1 sind einige Werte für D_{IS} gegeben. Die spitzen Klammern in Gleichung (2.15) und (2.16) bezeichnen eine zeitliche Mittelung. S steht für den Ordnungsparameter und beschreibt die interne Beweglichkeit⁴¹ unabhängig von der molekularen Umorientie-

³⁸Weiterhin benötigt man den Grad der Orientierung (Ordnungsparameter), wenn man Orientierungsmessungen an zwei verschiedenen Proben miteinander vergleichen will (siehe auch 3.1.5).

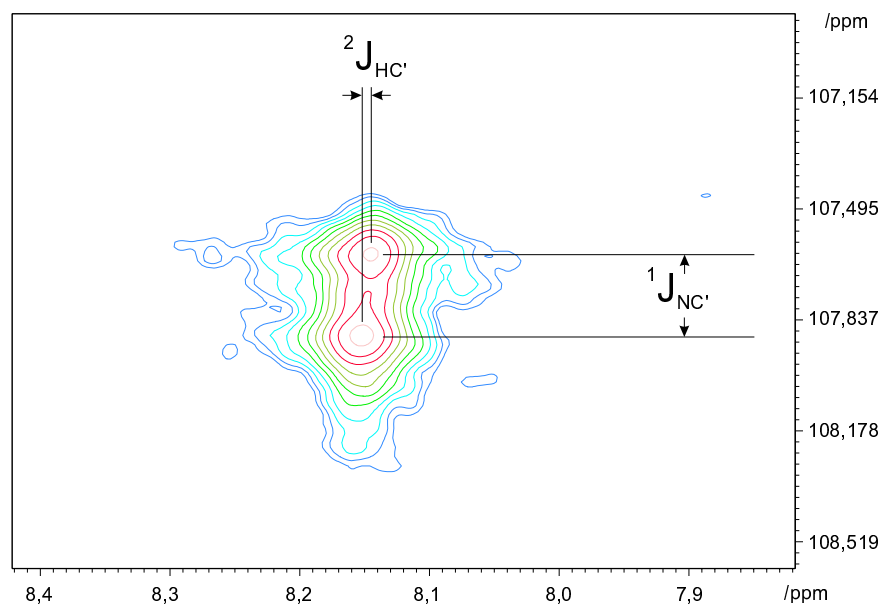
³⁹DMPC steht für Dimyristoylphosphorylcholin und DHPC steht für Dihexanoylphosphorylcholin.

⁴⁰Für isotrope Umorientierung des Moleküls gilt: $S_{xx} = S_{yy} = S_{zz} = 0$.

⁴¹siehe auch 3.1.5

statische Dipolkopplungskonstante	
$D_{C'N}$	= -2,0 kHz
$D_{C_\alpha C'}$	= 4,9 kHz
$D_{H_N N}$	= -24 kHz

Tabelle 2.1: Statische Dipolkopplungskonstanten entnommen aus [Bru01].

Abbildung 2.16: Dublettstruktur zur Messung der Restdipolkopplung zwischen ^{15}N - $^{13}\text{C}'$ - und der $^1\text{H}_N$ - $^{13}\text{C}'$.

rung, welche durch S_{ij} beschrieben wird.

Die Restdipolkopplung ergibt sich durch Differenzbildung aus der Kopplung im anisotropen Medium und im isotropen Medium, da sich die Restdipolkopplung im Grenzfall schwacher Kopplungen additiv mit der skalaren Kopplung zur effektiven Kopplung zusammensetzt. Eine Reihe von NMR-Experimenten bietet sich dazu an. Die Bestimmung der $^1\text{H}_N$ - ^{15}N -Restdipolkopplung erfolgte an einer 1,4 mM cMeCP2 ARBD-Probe unter Zuhilfenahme eines ^1H - ^{15}N -HSQC-Experiments, welches allerdings ohne den π -Puls im Protonenkanal in der Mitte der Entwicklungszeit t_1 aufgenommen wurde (siehe auch Abbildung 2.2). Damit ergibt sich ein Spektrum, welches für jeden Aminosäurerest ein

Dublett in indirekter Dimension aufweist, dessen Aufspaltung der effektiven Kopplung zwischen $^1\text{H}_N$ und ^{15}N entspricht.

Ausgehend von einer ^{13}C - und ^{15}N -isotopenmarkierten Probe (Konzentration: 1,6 mM) wurden Messungen vorgenommen, welche die Messung der Restdipolkopplungen zwischen $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$, der ^{15}N - $^{13}\text{C}'$ und der $^1\text{H}_N$ - $^{13}\text{C}'$ erlaubt. Zur Untersuchung der $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ -Kopplung fand ein modifiziertes H(N)CO-Experiment Verwendung, bei dem keine Entkopplung zwischen dem $^{13}\text{C}_\alpha$ - und dem $^{13}\text{C}'$ -Kern während der Entwicklungszeit t_1 stattfindet. Im Spektrum erscheint bei der chemischen Verschiebung des $^1\text{H}_N$ -Kerns des i -ten Aminosäurerests ein Dublett mit dem Signalschwerpunkt bei der chemischen Verschiebung des $^{13}\text{C}'$ -Kerns der $(i-1)$ -ten Aminosäure. Die eigentliche Restdipolkopplung läßt sich wie oben beschrieben aus zwei Spektren (unter orientierenden und isotropen Bedingungen gemessen) bestimmen. Mit einem herkömmlichen ^1H - ^{15}N -HSQC-Experiment⁴², in dem ein zusätzlicher, für die $^{13}\text{C}_\alpha$ -Kerne selektiver π -Puls während der Entwicklungszeit t_1 hinzugefügt wird, ist es möglich, die Restdipolkopplungen der ^{15}N - $^{13}\text{C}'$ - und der $^1\text{H}_N$ - $^{13}\text{C}'$ -Spinpaare zu messen.⁴³ Damit besteht während der Entwicklungszeit nur eine effektive Kopplung zum $^{13}\text{C}'$ -Kern des i -ten Aminosäurerests. Ein ähnliches Experiment wurde in [Wan99] vorgestellt. In Abbildung 2.16 ist die Dublettstruktur der Signale exemplarisch an einem Signal dargestellt. Bei der chemischen Verschiebung des ^1H - und des ^{15}N -Kerns des i -ten Aminosäurerests ist die Kopplung zum $^{13}\text{C}'$ -Kern des $(i-1)$ -ten Aminosäurerests zu erkennen.

Die Bestimmung der Kopplung erfolgt durch Abstandsbestimmung von zwei Signalen. Um möglichst genaue Werte für die Kopplungen zu erhalten, wurden in beiden Dimension 1024 Datenpunkte aufgenommen. Bei der Prozessierung wurde die Auflösung zusätzlich um einen Faktor zwei erhöht. Die Anzahl der *scans* lag bei allen Experimenten unter isotropen Bedingungen bei 32. Für die aufgenommenen Experimente unter anisotropen Bedingungen lag die Anzahl der *scans* bei 96. Diese Erhöhung der *scans* war, auf Grund der Linienverbreiterung als Folge der erhöhten Viskosität der Probe und der magnetischen Dipol-Dipol-Wechselwirkung, die sich auf Grund der partiellen Orientierung des Prote-

⁴²welches im allgemeinen an nur ^{15}N -markierten Proben Verwendung findet

⁴³Die Messung wurde am DRX600 Spektrometer an einem QXI-Probenkopf durchgeführt. Die Länge des selektiven rechteckigen π -Puls war 17,8 μs bei einer Dämpfung von 3 dB.

ins nicht ausmittelt, nötig. Für gut separierte Signale aus den oben in diesem Abschnitt beschriebenen Spektren wurden die Restdipolkopplungen bestimmt. Eine Liste der gemessenen Restdipolkopplungen befindet sich im Anhang. In Abschnitt 3.1.5 wird gezeigt, in welcher Form die gemessenen Restdipolkopplungen aufbereitet werden müssen, um als *restraints* in Strukturberechnungen einzufießen.

2.5 Berechnung der Proteinstruktur

Der natürliche Proteinfaltungsprozeß läuft im allgemeinen im Millisekundenbereich ab. Eine Simulation des Prozesses ist auf Grund des hohen Rechenaufwands mit den heute zur Verfügung stehenden Rechenkapazitäten nur eingeschränkt möglich. Infolgedessen wird die Strukturberechnung unter „beschleunigten“ Bedingungen in Form eines *simulated annealing*-Prozesses durchgeführt [Kir83]. Dabei werden Moleküldynamikrechnungen mit Nebenbedingungen, definiert durch experimentelle und empirische *restraints*, durchgeführt. Während dieser Rechnungen wird die zur Verfügung stehende kinetische Energie in Form einer Temperaturerniedrigung⁴⁴ verringert. Experimentelle und empirische *restraints*, gewonnen aus der Analyse z.B. von NOESY-Spektren bzw. aus den allgemeinen Kenntnissen über Molekülstrukturen (wie z.B. Bindungslängen) fließen in Form von Potentialen⁴⁵ in die Simulation ein. Das Ergebnis der Simulation ist eine Schar von Strukturen, deren Energien unter den gegebenen Nebenbedingungen minimiert sind.

Die endgültige Proteinstruktur ergibt sich aus vielen Zyklen, in denen die folgenden Prozesse durchlaufen werden: Definition der *restraints*, Strukturberechnung und Strukturbewertung. Zur Durchführung der Strukturberechnungen stehen zwei Computerprogramme zur Verfügung: CNS [Bru98] (*Crystallography- and NMR-System*) und DYANA [Gun97] (*Dynamics Algorithms for NMR Applications*). Dabei findet die Energieoptimierung der Moleküldynamikrechnungen hauptsächlich im Orts- bzw. Winkelraum statt.

⁴⁴Die Temperatur steht in keinem direkten Zusammenhang mit der beim realen Faltungsprozeß vorherrschenden Temperatur. Sie fungiert vielmehr als Energieparameter.

⁴⁵Potentiale, welche die experimentellen *restraints* repräsentieren, werden Pseudopotentiale genannt, da keine physikalischen Potentiale zu Grunde liegen.

2.5.1 Strukturberechnungen mit dem Computerprogramm CNS

Ausgehend von einer gestreckten Struktur wird die Dynamik der Atome unter Berücksichtigung der *restraints* hauptsächlich im Ortsraum simuliert; man spricht von rMD-Rechnungen (*restrained molecular dynamics*). Grundlage sind die Newtonschen Bewegungsgleichungen der involvierten Atome. Für das i -te Atom gilt:

$$\vec{F}_i(t) = -\vec{\nabla} E_{pot,i}(t) \quad (2.17)$$

$$\begin{aligned} E_{pot} &= E_{emp} + E_{exp} \\ &= (E_{bond} + E_{ang} + E_{imp} + E_{tor} + E_{vdW}) \\ &\quad + (E_{dihedral} + E_{NOE} + E_{coupl} + E_{shifts}) \end{aligned} \quad (2.18)$$

Dabei setzt sich die potentielle Energie E_{pot} aus unterschiedlichen Beiträgen zusammen, welche sich in zwei Klassen (E_{emp} und E_{exp}) einteilen lassen [Bru92]. Im empirischen Anteil des Potentials (E_{emp}) befinden sich Beiträge der kovalenten Bindungen (E_{bond}), der Bindungswinkel (E_{ang}), der starren Diederwinkel (E_{imp}), der rotationssymmetrischen Diederwinkel (E_{tor}) und der van-der-Waals-Wechselwirkung (E_{vdW}). Diese Terme werden mit Ausnahme von E_{vdW} durch harmonische Potentiale modelliert.⁴⁶ Der experimentelle Anteil des Potentials (E_{exp}) besteht aus Beiträgen bezüglich der Diederwinkel ($E_{dihedral}$), der Protonenabstände (E_{NOE}), der Kopplungskonstanten (E_{coupl}) und der chemischen Verschiebungen (E_{shifts}). Für die Terme, die in E_{exp} einfließen, werden im allgemeinen harmonische Potentiale angesetzt. Das Programm bietet aber auch die Möglichkeit andere Potentialformen zu Grunde zulegen.

Grundlage der Moleküldynamikrechnungen war das Standard-*annealing*-Protokoll der CNS-Version 1.0. Das Protokoll läßt sich grob in vier Stufen einteilen: Hochtemperaturphase, erste langsame Abkühlphase, zweite langsame Abkühlphase und abschließende Energieminimierung. Beim Durchlaufen dieser Stufen, die selbst wiederum mehrstufig sind, wird die Temperatur erniedrigt und die Gewichtung der den einzelnen Potentialen zugehörigen Kraftkonstanten erhöht. Dabei fällt besonders die Änderung der Gewichtung der zur van-der-Waals-Wechselwirkung gehörigen Kraftkonstante auf. In der Hochtempe-

⁴⁶Für E_{vdW} wurde ein Lennard-Jones-Potential verwendet.

raturphase wird die van-der-Waals-Wechselwirkung kaum berücksichtigt (Gewichtungsfaktor 0,1), am Ende der zweiten langsamen Abkühlphase beträgt der Gewichtungsfaktor 4,0. Das hat zur Folge, daß sich die Atome anfangs gut durchdringen können, was wiederum ermöglicht, daß sich fehlgefaltete Strukturen früh zurückorientieren können.

CNS weist jeder Struktur eine Energie zu, deren Größe davon abhängt, wie stark *restraints*, die in die Strukturberechnung eingeflossen sind, von der fertig berechneten Struktur verletzt werden. CNS protokolliert alle Schritte der Rechnung einschließlich der verletzten *restraints*. Allerdings ist das diese Informationen tragende *logfile* auf Grund seiner Größe und der Form der Ausgabe schwer auswertbar. Abhilfe verschafft das in Abschnitt 3.1.6 näher beschriebene Programm *cnsstat* [Ste01].

2.5.2 Strukturberechnungen mit dem Computerprogramm DYANA

Im Gegensatz zu CNS, das rMD-Rechnungen vorwiegend im Ortsraum ausführt, werden mit DYANA⁴⁷ TAD-Rechnungen (*torsion angle dynamics*) im Winkelraum durchgeführt. Dabei werden, um die kovalente Struktur des Proteins aufrecht zu erhalten, die Bindungslängen und die Bindungswinkel (außer den Torsionswinkeln) als fixiert angenommen.⁴⁸ Damit verringert sich die Zahl der Freiheitsgrade des Systems deutlich gegenüber rMD-Rechnungen im Ortsraum. Ziel ist es, die Bewegungsgleichungen der n Atome bzw. Atomgruppen des Proteins zu lösen. Für die n Lagrangegleichungen läßt sich schreiben (siehe z.B. [Gol63]):

$$\frac{d}{dt} \left(\frac{\partial(E_{kin} - E_{pot})}{\partial \dot{\Theta}_k} \right) - \frac{\partial(E_{kin} - E_{pot})}{\partial \Theta_k} = 0 \quad (2.19)$$

Dabei fungiert Θ_k als generalisierte Koordinate und bezeichnet den der k-ten Atomgruppe bzw. dem k-ten Atom zugehörigen Winkel. In [Gun97] werden Ausdrücke für die Energien geliefert. Ein Großteil des Terms für die potentielle Energie wird von der variablen Zielfunktion (*variable target function* (VTF)) [Bra85] ausgemacht, deren Wert sich

⁴⁷In der vorliegenden Arbeit wurde die Version 1.5 verwendet.

⁴⁸Der kovalente Charakter des Proteins wird bei CNS durch die Wirkung der unter E_{emp} zusammengefaßten Potentiale erzwungen.

für sinkende *restraint*-Verletzungen verringert. Am Ende der Strukturberechnungen werden die Strukturen mit wachsendem VTF-Wert ausgegeben. Der Rechenaufwand für die Lösung dieser n Differentialgleichungen kann von $\propto n^3$ durch schnelle rekursive Algorithmen [Jai93] auf $\propto n$ gesenkt werden.

Die Ausgabe der berechneten Struktur erfolgt geordnet nach steigender Größe der Summe der Verletzungen der *restraints*. Weiterhin erlaubt DYANA die Eingabe der *restraints* mit der selben Nomenklatur, wie sie bei CNS verwendet wird. Damit ist eine einfache Möglichkeit gegeben die Ergebnisse der Strukturberechnungen miteinander zu vergleichen (siehe dazu Abschnitt 3.1.7).

2.6 Rückrechnung von NOESY-Spektren

Die reale Struktur eines Proteins bestimmt über die Kern-Kern-Abstände die in den NOESY-Spektren auftretenden NOE-Kontakte. Liegt eine weitgehend vollständige Zuordnung der chemischen Verschiebungen der Kerne des Proteins sowie ein Strukturmodell des Proteins vor, so ist es möglich, die NOESY-Spektren zu simulieren. Die mathematische Behandlung dieser Fragestellung ist mit dem Relaxationsmatrixformalismus möglich, dessen Grundlage die Erweiterung der Solomon-Gleichung [Sol55] auf ein Vielteilchensystem ist [Cav95]. Die Vielteilchen-Solomongleichung beschreibt die zeitliche Entwicklung der Magnetisierung unter dem Einfluß von magnetischer Dipol-Dipol-Wechselwirkung eines Spins mit anderen Spins des Ensembles. In Vektorschreibweise läßt sich diese Gleichung wie folgt darstellen [Cav95]:

$$\frac{d\Delta\vec{M}_z(t)}{dt} = -\underline{\mathbf{R}} \cdot \Delta\vec{M}_z(t) \quad (2.20)$$

Dabei steht $\underline{\mathbf{R}}$ für die $N \times N$ Relaxationsmatrix, deren Diagonalelemente als Autorelaxationsraten und deren Nichtdiagonalelemente als Kreuzrelaxationsraten bezeichnet werden, welche die Informationen über die intramolekularen Bewegungsmoden der N beteiligten Spins tragen.

Das Computerprogramm RELAX (vorgestellt in [Gor97] und erweitert in [Gor99]), welches im Programm AUREMOL eingebettet ist, liefert auf der Grundlage des Relaxationsmatrixformalismus' die sogenannten rückgerechneten NOESY-Spektren. Das Programm

verlangt die Eingabe einer Struktur und der chemischen Verschiebungen der einzelnen Kerne. Neben der Mischzeit und der Larmofrequenz des ^1H -Kerns ist es außerdem möglich, verschiedene Bewegungsmodelle der einzelnen Wasserstoffatome anzunehmen.

Die Verwendung der Rückrechnung von NOESY-Spektren zur Vervollständigung der Zuordnung der NOE-Signale von cMeCP2 ARBD ist in Abschnitt 3.1.6 dargestellt.

Kapitel 3

Zusammenfassung der Ergebnisse

In diesem Kapitel wird beschrieben, wie die aufgenommenen NMR-Daten so aufbereitet werden können, daß biologisch relevante Schlüsse über die Proteinstruktur (Abschnitt 3.1) sowie die Bindung an DNA (Abschnitt 3.2) gezogen werden können.

3.1 Strukturbestimmung von cMeCP2 ARBD

Die NMR-spektroskopische Bestimmung der Struktur von cMeCP2 ARBD läßt sich in drei Schritte einteilen. Im ersten Schritt wurden Signale, die zu einem Aminosäurerest oder einem sequentiell benachbarten Aminosäurerest gehören, von Signalen getrennt, die zwischen Atomen verschiedener z.T. sequentiell weit auseinander liegender Aminosäurereste auftreten. Dieser erste Schritt war wiederum zweistufig, zuerst erfolgte die sequentielle Zuordnung der Hauptkettenatome (3.1.1) gefolgt von der Seitenkettenuordnung (3.1.2). Auf der Grundlage dieser Daten war es möglich, Vorhersagen über die Sekundärstruktur von cMeCP2 ARBD zu treffen (3.1.3). Diese Vorhersagen vereinfachten das weitere Vorgehen, denn sie ermöglichten unter bestimmten Bedingungen die Hinzunahme von Einschränkungen des Konformationsraums des Proteins (*restraints*) neben NOE-Signalen, wie Restdipolkopplungen (3.1.5) und vorhergesagten Werten von Diederwinkeln (3.1.4) aus einem Datenbankvergleich in die sich anschließenden Strukturberechnungen.

Nach dieser Arbeit war es möglich, in einem zweiten Schritt (3.1.6) *restraints* aus NOESY-Spektren zu finden, die im dritten Schritt (3.1.7) das Berechnen der möglichen

Struktur mittels *simulated annealing* erlaubt. Die Strukturbestimmung von cMeCP2 ARBD war ein iterativer Prozess; auf dem Weg zur endgültigen Struktur wurden mehrere Zyklen der letzten beiden Schritte durchlaufen. Dabei wurde der Satz der *restraints* immer wieder anhand der Qualität der erhaltenen Strukturen überprüft, verändert und erweitert.

Die Strukturbestimmung erfolgte derart, daß die Verletzung der verwendeten *restraints* möglichst gering ist. Die Information, welche die *restraints* tragen, ist teilweise fehlerhaft und in ihrer Gesamtheit überbestimmt. Fehler können zum einen in einer fehlerhaften Zuordnung der NOE-Signale zu finden sein, zum anderen in nicht ideal erfüllten Voraussetzungen, die eine Abweichung der r^{-6} -Abhängigkeit der Signalintensität zur Folge haben (siehe Abschnitt 2.4.3).

3.1.1 Die sequentielle Zuordnung der Hauptkettenatome

Der sogenannte „Fingerabdruck“ des Proteins ist in Abbildung 3.1 in Form eines ^1H - ^{15}N -HSQC-Spektrums gezeigt. Ziel der sequentiellen Zuordnung der Hauptkettenatome ist es, jedes Signal dieses Spektrums einem Aminosäurerest des Proteins zuzuordnen. Es ist zu erkennen, daß der Bereich zwischen 8,4 und 8,5 ppm ^1H -Verschiebung sowie zwischen 110 und 120 ppm ^{15}N -Verschiebung viele sich überlagernde Signale aufweist. Daraus leitet sich die Notwendigkeit ab, zur Verbesserung der Auflösung dreidimensionale heteronukleare Spektren zur sequentiellen Zuordnung der Hauptkette heranzuziehen.¹ Dies geschah bei cMeCP2 ARBD unter Verwendung der Spektren HNCA, CBCA(CO)NH und HNCO (mehr Informationen zu diesen Spektren in Abschnitt 2.4.2.1 und in [Cav95]). Beispielhaft ist die sequentielle Zuordnung mit Hilfe des HNCA-Spektrums für die Aminosäurereste 104 bis 108 in Abbildung 3.2 dargestellt. Das grundlegende Vorgehen bei der Signalzuordnung des Proteinrückgrats bestand im wesentlichen darin, in einem ersten Schritt zusammenhängende Abschnitte benachbarter Spinsysteme zu finden, um in einem zweiten Schritt diese Abschnitte der Primärsequenz zuzuordnen. Dies war möglich, da die verschiedenen Aminosäurereste teilweise charakteristische chemische Verschiebungen der $^{13}\text{C}_\alpha$ -

¹Für kleinere Proteine mit weniger überlagerten Bereichen ist auch eine Zuordnung mit homonuklearen TOCSY- und NOESY-Spektren möglich.

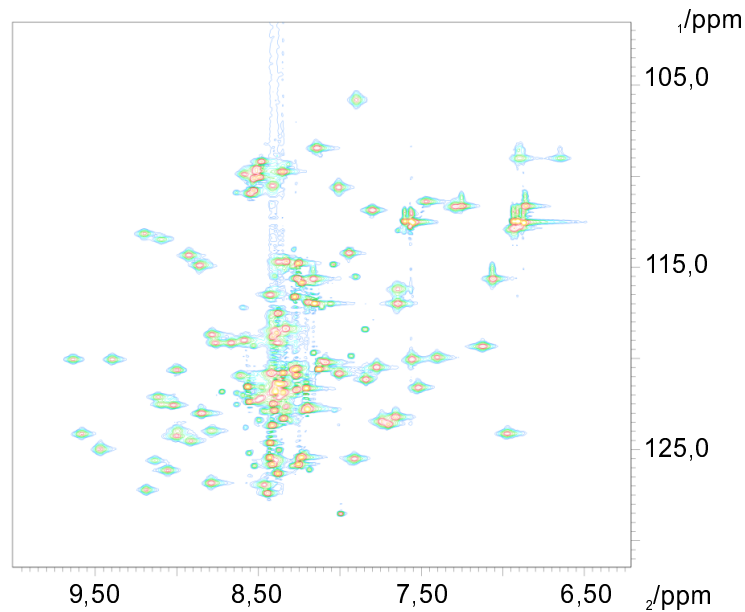


Abbildung 3.1: Das ^1H - ^{15}N -HSQC-Spektrum von *cMeCP2 ARBD*. Dargestellt sind die chemischen Verschiebungen der ^{15}N -Kerne (^{15}N -Verschiebungen) auf der Ordinate und die $^1\text{H}_\text{N}$ -Verschiebungen auf der Abszisse. Das Spektrum wurde bei einer Temperatur von 298 K und bei einer Protonenresonanzfrequenz von 600 MHz aufgenommen. Die Zusammensetzung der Probe ist in Abschnitt 2.3 aufgeführt.

und $^{13}\text{C}_\beta$ -Kerne aufweisen. Die $^{13}\text{C}_\beta$ -Verschiebungen ließen sich mit dem CBCA(CO)NH-Spektrum bestimmen. Dieses Spektrum ist ähnlich sensitiv, wie das HNCA-Spektrum und korreliert die ^{15}N - und die $^1\text{H}_\text{N}$ -Verschiebung des *i*-ten Aminosäurerests mit der $^{13}\text{C}_\alpha$ - und der $^{13}\text{C}_\beta$ -Verschiebung des (*i*-1)-ten Aminosäurerests. Mit dem HNCO-Spektrum lassen sich die chemischen Verschiebungen der $^{13}\text{C}'$ -Kerne ermitteln. In diesem Spektrum, es ist das empfindlichste der routinemäßig verwendeten dreidimensionalen Spektren, werden die chemischen Verschiebungen des ^{15}N - und des $^1\text{H}_\text{N}$ -Kerns des *i*-ten Aminosäurerests mit der $^{13}\text{C}'$ -Verschiebung des (*i*-1)-ten Aminosäurerests korreliert.

Die begrenzte Länge der zusammenhängend zugeordneten Abschnitte wird im allgemeinen durch das Fehlen von Signalen hervorgerufen, z.B. wegen Linienverbreiterungen oder durch das Auftreten von Prolinresten. Das Volumen eines Peaks ist direkt proportional zur Anzahl der NMR-aktiven Kerne in der Probe. Verbreitert sich ein Signal im Vergleich zu einem anderen, sinkt das Signal/Rausch-Verhältnis, da das Volumen des Si-

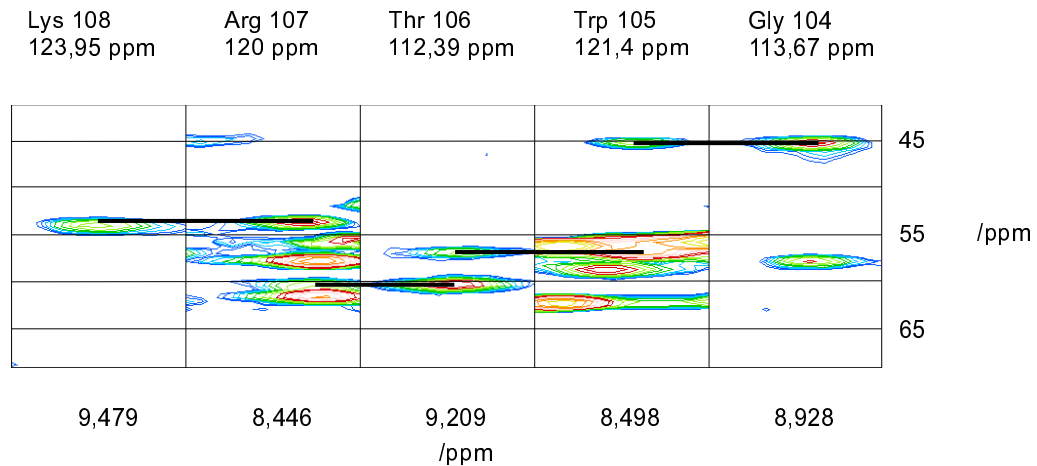


Abbildung 3.2: Ausschnitte aus dem HNCA-Spektrum von *cMeCP2 ARBD*. Dargestellt sind die Aminosäurereste 104 bis 108. Der schwarze Balken verbindet das $^{13}\text{C}_\alpha$ -Atom des *i*-ten Aminosäurerests bei der ^{15}N -Verschiebung des (*i*+1)-ten Aminosäurerests (linke Seite des Balkens) und seiner eigenen ^{15}N -Verschiebung (rechter Rand des Balkens). Die $^1\text{H}_\text{N}$ -Verschiebungen sind unten, die ^{15}N -Verschiebungen oben angegeben.

gnals konstant (unter Umständen bis zum Verschwinden des Signals im Rauschen). Bis auf zwei Aminosäurereste (Met95 und Val123) konnten alle Aminosäurereste, die im HNCA-Spektrum prinzipiell detektierbar sind, aufgefunden werden. Der hohe Anteil an Prolinresten (14 von 125 Aminosäureresten) stellte ein wesentliches Problem dar. Da Prolinreste kein $^1\text{H}_\text{N}$ -Atom besitzt, läßt sich mit Hilfe des HNCA-Spektrums keine Korrelation zwischen dem eigenen $^{13}\text{C}_\alpha$ - und dem $^{13}\text{C}_\alpha$ -Kern des sequentiellen Vorgängers herstellen, die „Kette reißt ab“. Für *cMeCP2 ARBD* ergeben sich 14 Bruchstücke², die unabhängig voneinander zugeordnet werden mußten. Je kürzer ein Abschnitt benachbarter Spinsysteme ist, um so weniger eindeutig läßt er sich einem Primärsequenzstück zuordnen. Besonders schwierig war die Signalzuordnung der C-terminalen Aminosäurereste; sieben der 14 Prolinreste verteilen sich auf die letzten 44 Aminosäurereste. Einen Überblick über den Grad der Zuordnung der chemischen Verschiebungen gibt Tabelle 3.1.

Das Vorkommen von Prolinresten in der Sequenz von *cMeCP2 ARBD* birgt eine weitere Schwierigkeit bei der Zuordnung. Abweichend von den restlichen Aminosäureresten besteht bei Prolinresten die Möglichkeit, daß neben der *trans*-Konformation auch die *cis*-

²14, da die Aminosäurereste 173 und 174 beides Prolinreste sind

Spin	Grad der Zuordnung
^{15}N	89%
$^1\text{H}_N$	98%
$^1\text{H}_\alpha$	94%
$^1\text{H}_\beta$	86%
$^{13}\text{C}_\alpha$	99%
$^{13}\text{C}_\beta$	81%
$^{13}\text{C}_\gamma$	74%
$^{13}\text{C}'$	83%

Tabelle 3.1: *Grad der Zuordnung* (^{15}N , $^1\text{H}_N$, $^1\text{H}_\alpha$, $^1\text{H}_\beta$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}_\gamma$, $^{13}\text{C}'$). Als Basis dient die Sequenz von Ala72 bis Gly196, d.h. der His-tag wurde nicht berücksichtigt. Eine Liste aller zugeordneten Signale befindet sich im Anhang.

Konformation in nicht zu vernachlässigender Häufigkeit eingenommen wird. So wurden bei cMeCP2 ARBD unter den hier gewählten experimentellen Bedingungen (siehe Abschnitt 2.3) in mehreren Fällen für Aminosäurereste in der sequentiellen Umgebung von Prolinresten mehr als eine Konformation nachgewiesen. Dabei handelt es sich ausschließlich um Prolinreste nahe der Termini (73, 76, 94, 166, 174, 181, 183). Für die sequentielle Zuordnung wurde die jeweils am stärksten populierte Konformation berücksichtigt.

3.1.2 Zuordnung der Seitenkettenatome

Im allgemeinen liefert eine Analyse des homonuklearen zweidimensionalen TOCSY-Spektrums, des ^{15}N -editierten TOCSY- und des HCCH TOCSY-Spektrums die chemischen Verschiebungen der Kerne in den Seitenketten (in Abschnitt 2.4.2.2 und in [Cav95] sind diese Spektren näher erklärt). Im Fall des cMeCP2 ARBD gestaltete sich die Auswertung der TOCSY-Spektren wenig erfolgreich. Die Signalintensitäten der Kreuzsignale beim homonuklearen TOCSY-Spektrum und beim ^{15}N -editierten TOCSY-Spektrums waren häufig sehr schwach und es war meist nicht möglich, den Signalen Spins zuzuordnen. Die Auswertung des HCCH TOCSY-Spektrums war erfolgreicher, wobei es für die Auswertung dieses Spektrums von Vorteil ist, eine der Protonenresonanzen der jeweiligen Seitenkette zu kennen, was häufig nicht gegeben war. Mit diesen drei Spek-

tren war es möglich, etwa 25% der Seitenkette zuzuordnen. Die Seitenketten der Prolinreste konnten so nicht zugeordnet werden, da durch das Fehlen des $^1\text{H}_N$ -Kerns und dem hohen Anteil an Prolinresten in der Sequenz ($> 11\%$) ein „Einstieg“ in das HCCH TOCSY-Spektrum unmöglich war. Dieser eher niedrige Grad der Zuordnung, der eine spätere Zuordnung der NOE-Signale fast unmöglich macht, war der Anlaß, andere, weniger häufig benutzte Spektren zur Seitenkettenuordnung heranzuziehen. Insbesondere wurde das HBHA(CBCACO)NH- und das CC(CO)NH-Spektrum aufgenommen. Im HBHA(CBCACO)NH-Spektrum wird die ^{15}N - und die $^1\text{H}_N$ -Verschiebung des i -ten Aminosäurerests mit der $^1\text{H}_\alpha$ - und der $^1\text{H}_\beta$ -Verschiebung (falls vorhanden) des sequentiellen Vorgängers korreliert. Im Fall des CC(CO)NH-Spektrums werden zusätzlich zu der $^{13}\text{C}_\alpha$ - und der $^{13}\text{C}_\beta$ -Verschiebung auch die übrigen Kohlenstoffe (falls vorhanden) mit der ^{15}N - und die $^1\text{H}_N$ -Verschiebung des sequentiellen Nachfolgers korreliert. Somit stand mit dem CC(CO)NH-Spektrum ein zweites Spektrum zur Verfügung, mit dem die durch das CBCA(CO)NH-Spektrum gewonnene Zuordnung der $^{13}\text{C}_\alpha$ - und $^{13}\text{C}_\beta$ -Verschiebungen überprüft und vervollständigt werden konnte. Mit diesen neu gewonnenen Informationen war es nun möglich, vorher nicht zu identifizierende Spinsysteme im HCCH TOCSY-Spektrum zuzuordnen und somit zu einem Zuordnungsgrad zu gelangen, der eine Zuordnung einer Vielzahl der NOE-Kontakte möglich machte (siehe Tabelle 3.1).

3.1.3 Sekundärstrukturvorhersagen

Aus den chemischen Verschiebungen der Kerne der einzelnen Aminosäurereste war es möglich, eine Vorhersage zur Sekundärstruktur von cMeCP2 ARBD zu treffen. Hierzu werden die gemessenen chemischen Verschiebungen mit sogenannten *random coil shifts*, d.h. chemischen Verschiebungen unstrukturierter Peptide verglichen. Diese *random coil shifts* sind aus Arbeiten, hauptsächlich von Wishart und Sykes [Wis91],[Wis94], an Hexapeptiden der Form Gly-Gly-X-Ala-Gly-Gly bekannt, wobei X für die jeweilig betrachteten Aminosäurerest steht. Man findet für $^{13}\text{C}_\alpha$ - und $^{13}\text{C}'$ -Kernen in β -Faltblättern charakteristische Hochfeldverschiebungen, in α -Helices Tieffeldverschiebungen. Für $^1\text{H}_\alpha$ - und $^{13}\text{C}_\beta$ -Kerne wurde der umgekehrte Effekt gefunden (Hochfeldverschiebungen für α -Helices und Tieffeldverschiebungen für β -Faltblätter). In [Wis95b] wurde gezeigt, daß es auch zu signifikanten Änderungen der *random coil shifts* kommt, wenn der sequentielle Nachfolger

ein Prolinrest ist. Die Tendenz der Verschiebung ist der einer β -Faltblatt-Konformation ähnlich. Alle anderen Aminosäurereste zeigen keine solche Tendenz; Alaninreste stehen also stellvertretend für die restlichen Aminosäurereste.

In Abbildung 3.3 sind die Differenzen der chemischen Verschiebungen verschiedener Kerne von den jeweiligen *random coil*-Werten dargestellt. β -Faltblattstrukturen liegen um Aminosäurerest 108, 120 und 131 und eine α -Helix liegt im Bereich von Aminosäurerest 140. Die genaue Lage der Sekundärstrukturelemente ließ sich allerdings auf der Grundlage der chemischen Verschiebungen allein nicht vorhersagen. Für genauere Analysen wurde in [Wis94] ein Programm vorgestellt, das zum einen die Verschiebungen von $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ und $^{13}\text{C}'$ in ihrer Gesamtheit betrachtet und wichtet sowie zum anderen dem kooperativen Charakter von Sekundärstrukturelementen Rechnung trägt. Der Einsatz dieses Programms führte zu dem in Abbildung 3.4 dargestellten Ergebnis.

Wie bereits in Kapitel 1 erwähnt, haben sich auch andere Arbeitsgruppen mit zu cMeCP2 ARBD sequenzhomologen Proteinen beschäftigt. Besondere Aufmerksamkeit wurde der Arbeit von Wakefield et al. [Wak99] gewidmet, da für den Bereich von Aminosäurerest Ala78 bis Aminosäurerest Ser165 eine 100%ige Sequenzidentität zu cMeCP2 ARBD besteht.³ In Abbildung 3.5 sind die Primärsequenz sowie die Vorhersagen für die Sekundärstruktur mittels CSI für beide Proteine vergleichend dargestellt. Es zeigte sich, daß eine prinzipielle Übereinstimmung besteht, mit Ausnahme des Bereichs um Aminosäurerest Asp98. Für diesen Bereich wurde von Wakefield et al. ein aus vier Aminosäureresten bestehender β -Strang vorhergesagt. Dieser Unterschied gab Anlaß, das gesamte β -Faltblatt mit Hilfe von NOE-Kontakten näher zu untersuchen. Typisch für Aminosäurereste in parallelen wie antiparallelen β -Faltblättern im Gegensatz zu Aminosäureresten in α -Helices sind Kontakte zwischen H_α des i -ten Aminosäurerests zum H_N des nachfolgenden Aminosäurerests ($i+1$), weiterhin sind Kontakte zwischen in der Sequenz z.T. weit auseinanderliegenden⁴ Aminosäureresten für das Vorliegen von Faltblattstrukturen charakteristisch. Die Kontakte weisen ein gewisses Muster auf: Liegt ein Kontakt zwischen dem i -ten und j -ten Aminosäurerest vor und sind zusätzlich auch Kon-

³Die Zählung bezieht sich auf cMeCP2 ARBD, die Zählung für MeCP2 MBD ist um eins nach unten verschoben im Vergleich zu cMeCP2 ARBD.

⁴meist als langreichweitige NOE-Kontakte bezeichnet

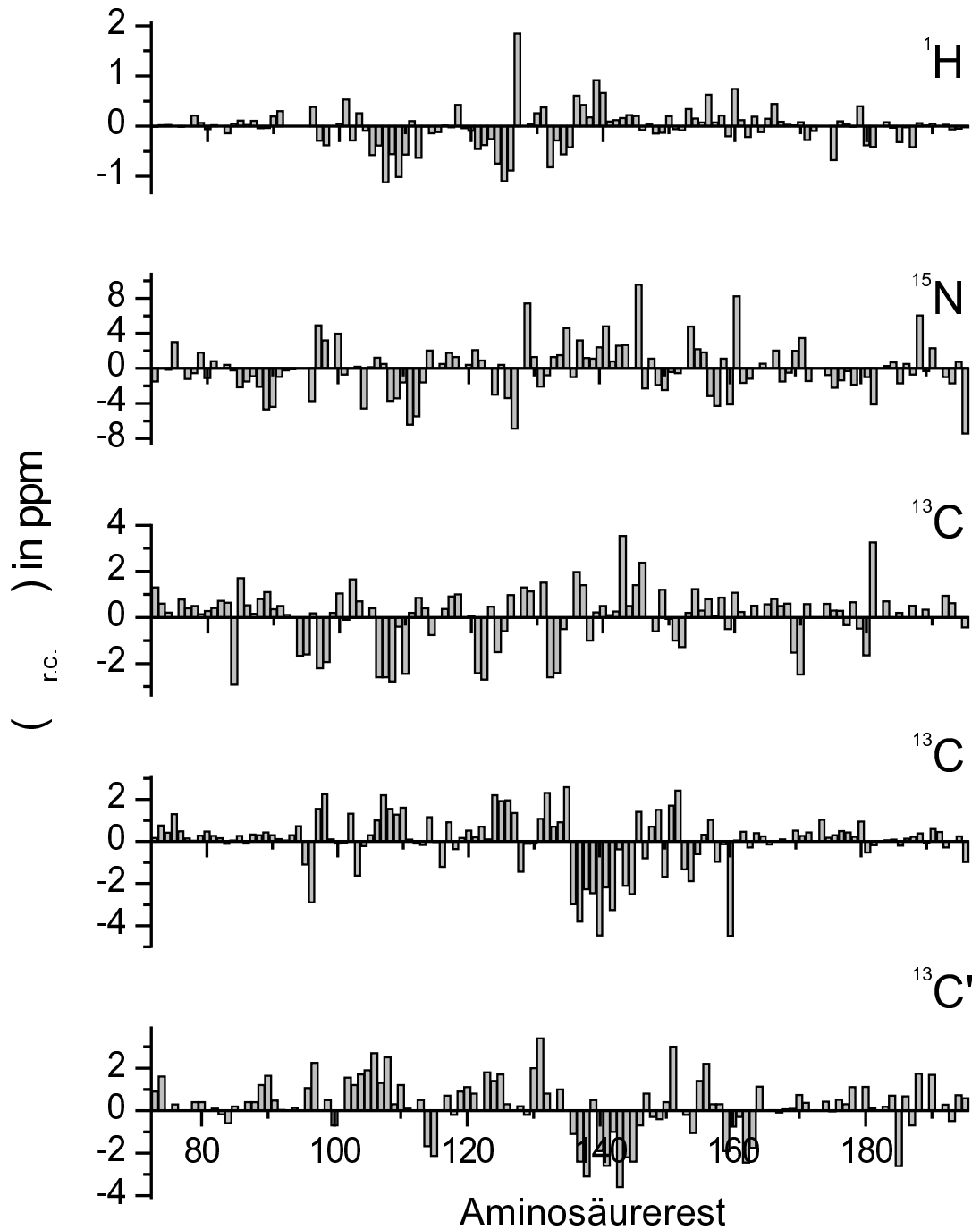


Abbildung 3.3: *Differenzen der chemischen Verschiebungen in Abhängigkeit der Sequenz.* Dargestellt sind die Änderungen der Werte für $^1\text{H}_\alpha$, ^{15}N , $^{13}\text{C}_\beta$, $^{13}\text{C}_\alpha$ und $^{13}\text{C}'$ zum jeweiligen *random coil (r.c.) shift*. Die gemessenen Werte sind mit DSS referenziert; die Referenzwerte stammen aus [Wis95b].

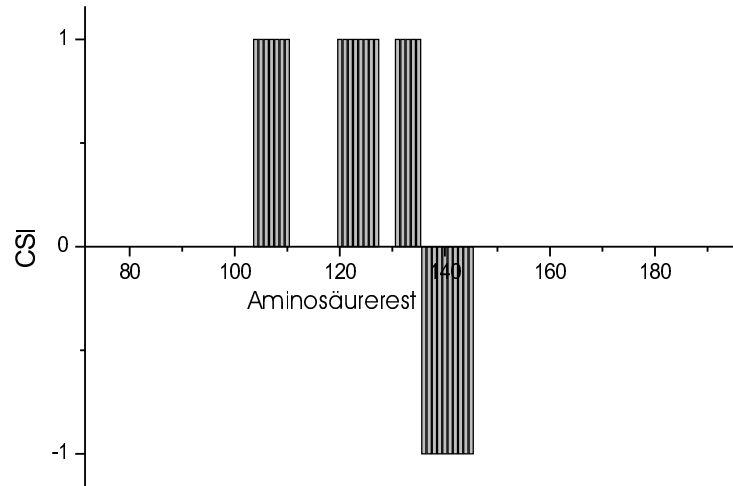


Abbildung 3.4: *CSI (chemical shift index)* für *cMeCP2 ARBD*. Es flossen die $\Delta\delta$ -Werte für $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ und $^{13}\text{C}'$ ein. Ein Wert von +1 steht für β -Faltblatt, -1 für α -Helix und 0 für nicht kanonische Sekundärstrukturelemente nach [Wis94], [Wis91].

takte zwischen dem (i+2)-ten und dem (j-2)-ten (für ein antiparalleles) oder zwischen dem (i+2)-ten und dem (j+2)-ten (für ein paralleles β -Faltblatt) vorhanden, so ist dies ein starker Hinweis auf das Vorliegen eines β -Faltblattes. Läßt sich dieses Muster über mehrere Aminosäurereste hinweg erkennen, so ist diese Sequenzabschnitt als β -Faltblattstrang identifiziert. Für *cMeCP2 ARBD* sind die NOESY-Spektren qualitativ auf NOE-Kontakte für den in Frage kommenden Bereich hin untersucht worden. Die Ergebnisse sind in Abbildung 3.6 dargestellt. Es zeigt sich, daß es für den Bereich um Aminosäurerest 98 keine für β -Faltblätter typischen NOE-Kontakte zum benachbarten Strang gibt. Faßt man die Ergebnisse der Vorhersage der Sekundärstrukturelemente zusammen, so ergibt sich ein antiparalleles β -Faltblatt (bestehend aus drei Strängen) direkt gefolgt von einer α -Helix:

Strukturelement	Aminosäurereste
β_1	104-110 (GWTRKCLK)
β_2	120-127 (KYDVYLIN)
β_3	131-135 (KAFRS)
α	136-145 (KVELIAYFEK)

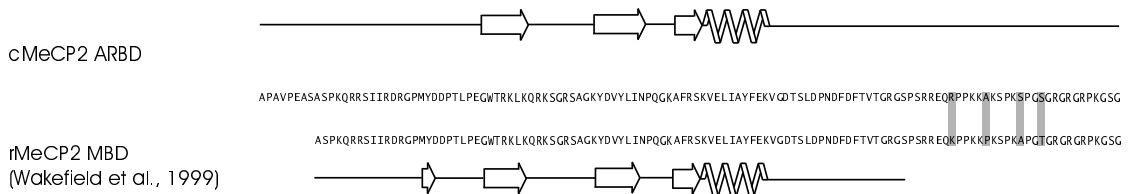


Abbildung 3.5: Vergleichsbild *cMeCP2 ARBP* (oben) und *rMeCP2 MBD* (unten). Zu sehen ist die Sekundärstruktur vorhergesagt mit CSI und die Sequenz der beiden Proteine. Grau unterlegt sind bezüglich der Aminosäuresequenz unterschiedliche Abschnitte.

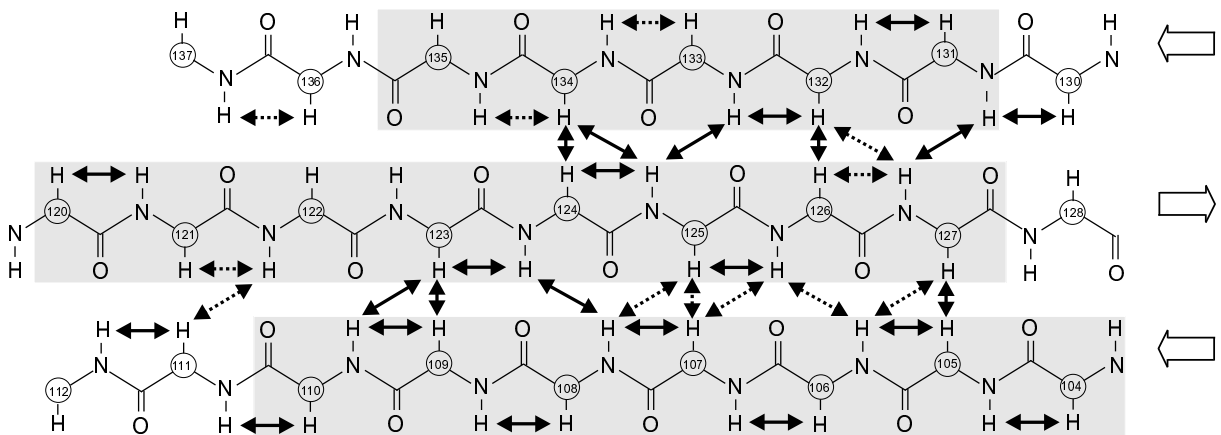


Abbildung 3.6: β -Faltblatt-Topologie von *cMeCP2 ARBD*, bestimmt durch eine qualitative Auswertung der NOESY-Spektren. Durchgezogene Pfeile symbolisieren gefundene NOE-Kontakte, gepunktete Pfeile stehen für NOE-Kontakte, die auf Grund von Signalüberlagerungen mehrdeutig sein können. Zum besseren Überblick sind die durch die Analyse der chemischen Verschiebungen mittels CSI vorhergesagten Sekundärstrukturelemente als graue Boxen eingezeichnet.

3.1.4 Diederwinkel

Eine weitere Methode, ohne vollständige Analyse der NOESY-Spektren Aussagen über die Sekundärstruktur zu treffen, wurde von Cornilescu et al. [Cor99] eingeführt. Die genannten Autoren entwickelten ein Computerprogramm (TALOS), welchem ein wissenschaftlicher Ansatz, kombiniert mit den experimentell bestimmten Werten für die chemischen Verschiebungen zu Grunde liegt, um Vorhersagen über die Diederwinkel Φ und Ψ zu treffen (siehe auch Abbildung 3.7). Nach Ramachandran [Ram63] nehmen die Winkel Φ und Ψ bestimmte Werte bevorzugt ein. Weiterhin gilt, daß bestimmte Winkelpaare typischerweise bei Aminosäureresten auftreten, die in β -Faltblättern liegen; andere Win-

kelpaare treten vermehrt bei Aminosäureresten in α -Helices auf. Im Programm TALOS wird ein drei Aminosäurereste langes Sequenzstück mit der Sequenz von 20 gut definierten Proteinstrukturen verglichen. Liegt Sequenzhomologie vor und sind die chemischen Verschiebungen für $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$ und ^{15}N ähnlich⁵, so wird angenommen, daß auch die Werte für die Diederwinkel vergleichbar sind. Werden nun mindestens 9 Tripel aus der Datenbank gefunden, die annähernd die gleiche Größe für die beiden Winkel voraussagen, so besteht nach [Cor99] eine Wahrscheinlichkeit von 97% dafür, daß der mit der beschriebenen Methode bestimmte Winkel dem experimentell bestimmten Wert innerhalb des angegebenen Fehlers entspricht.⁶ Das Programm vergibt dann die Wertung „gut“ für das Paar (Φ, Ψ) .

Die Ergebnisse der Analyse der chemischen Verschiebungen für cMeCP2 ARBD mittels TALOS sind in Abbildung 3.8 dargestellt. Auch diese Methode liefert die schon vorher gefundenen Sekundärstrukturelemente (siehe Abschnitt 3.1.3). Abweichend von den vorherigen Analysen wird neben dem Bereich um Aminosäurerest Asp98, der auch von Wakefield et al. als Sekundärstrukturelement vorhergesagt wurde (s.o.), der Bereich zwischen Asp152 und Arg163 als strukturiert erkannt. Für sieben dieser elf Aminosäurereste fand TALOS neun und mehr Werte für die beiden Diederwinkel in der benutzten Datenbank, die vom Bewertungsalgorithmus als „gut“ eingestuft werden. Betrachtet man die veröffentlichte Struktur von MBD1 MBD [Ohk99] in diesem Bereich, so zeigen sich auch dort von *random coil*-Bereichen verschiedene Strukturelemente. Nach [Ohk99] wird dem Bereich um Aminosäurerest 162 auch eine wichtige Rolle bei der Bindung des Proteins an methylierte DNA zugewiesen. In der Struktur von MeCP2 MBD ist der Bereich um Aminosäurerest Asp152 strukturiert (Helix), der sich anschließende Bereich allerdings nicht. Somit wurde diesem Bereich auch bei cMeCP2 ARBD bei der weiteren Bestimmung der Struktur besondere Aufmerksamkeit zuteil.

Es ist anzumerken, daß die gute Übereinstimmung von 97% nur für Aminosäurereste in Sekundärstrukturelementen bzw. in wenig flexiblen Bereichen erzielt wurde. Betrachtet man z.B. den Bereich um Asp98: TALOS erkannte hier Sekundärstrukturelemente, die

⁵nähere Informationen siehe[Cor99]

⁶Unter der Annahme, daß der betrachtete Aminosäurerest Teil kanonischer Sekundärstrukturelemente ist, bzw. in einem wenig flexiblen Bereich des Proteins liegt.

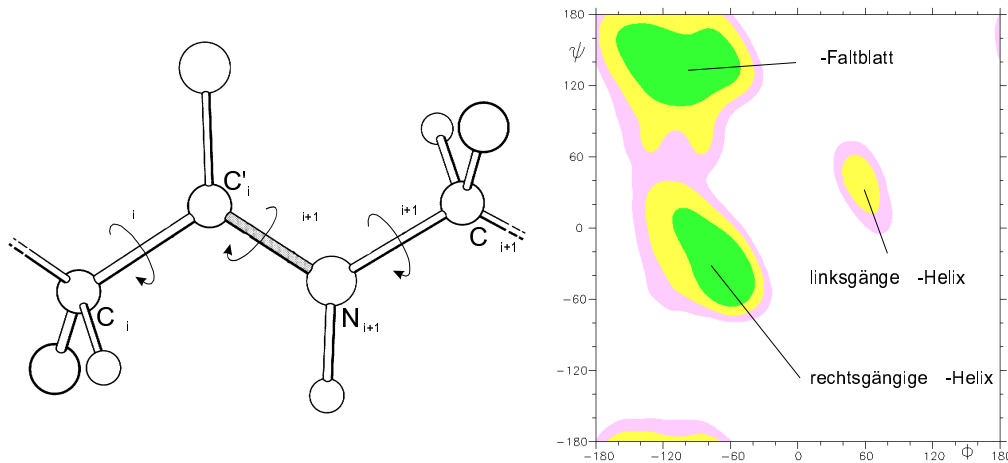


Abbildung 3.7: Die Diederwinkel und der Ramachandran Diagramm [Ram63]. Links in der Abbildung ist ein Stück aus dem Rückgrat eines Proteins eingezeichnet. Grau unterlegt ist die Peptidbindung; ω weitgehend fixiert bei 180° , Φ und Ψ sind flexibel, bevorzugen aber bestimmte Regionen, die in der Abbildung rechts (Ramachandran Diagramm) eingezeichnet sind. Farblich eingezeichnet sind die erlaubten Regionen für das Winkelpaar (Φ, Ψ) .

aber bei einer qualitativen Analyse der NOE-Kontakte in diesem Bereich (siehe Abbildung 3.6) keine für kanonische Sekundärstrukturelemente typischen Kontakte aufweisen. Eine Vereinfachung der Auswertung der TALOS-Ausgaben bezüglich der Vorhersage von Sekundärstrukturelementen würde sich ergeben, wenn ähnlich wie beim CSI einige Kriterien zur Weiterverarbeitung der Daten gegeben wären, die dem kooperativen Charakter von Sekundärstrukturelementen Rechnung tragen.

Neben der Möglichkeit, TALOS zur Vorhersage der Sekundärstruktur heranzuziehen, ergibt sich eine noch wesentlich nützlichere Verwertung der TALOS-Ausgaben für die Bestimmung der Struktur. Die Strukturberechnungsprogramme CNS und DYANA (siehe Abschnitt 2.5) bieten die Möglichkeit, die Diederwinkel Φ und Ψ als *restraints* für Strukturberechnungen einzubeziehen. Voraussetzung für dieses Einbringen der Winkelinformationen ist, daß Φ und Ψ aus Aminosäureresten in Sekundärstrukturelementen, bzw. aus wenig flexiblen Bereichen stammen. Die in die Strukturberechnung eingeflossenen *restraints* für die Diederwinkel sind im Anhang abgedruckt.

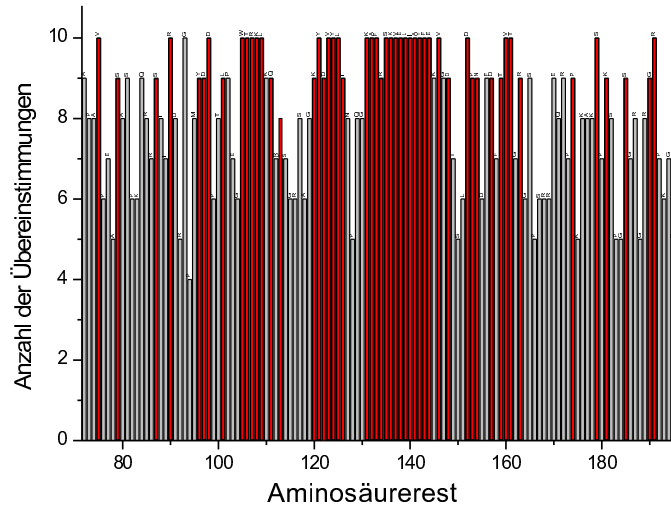


Abbildung 3.8: Vorhersagen der Diederwinkel mittels TALOS [Cor99] für *cMeCP2 ARBD*. Aufgetragen sieht man die Anzahl der Übereinstimmungen des Datenbankvergleichs der Diederwinkel Φ, Ψ in Abhängigkeit der Sequenz. Rot eingetragen, sind die Werte, die TALOS als „gut“ bewertet.

3.1.5 Restdipolkopplungen

Ein weitere Möglichkeit, den Konformationsraum eines Proteins einzuschränken, bieten Restdipolkopplungen. Wie in Abschnitt 2.4.4 näher beschrieben, läßt sich die Orientierung einer Bindung bezüglich eines molekülfesten Orientierungstensors messen. Voraussetzung für das Einbringen dieser Information in Strukturberechnungen ist die Kenntnis der Dynamik des Proteins im betrachteten Bereich, da diese in Form des Ordnungsparameters S die Größe der Restdipolkopplung beeinflusst. Ist S nicht bekannt, so muß man die Verwendung von Restdipolkopplungen in Strukturberechnungen auf die Bereiche beschränken, die erfahrungsgemäß einen Wert des Ordnungsparameters nahe 1 aufweisen (es gilt: $S \approx 0$ sehr flexibel; $S \approx 1$ starr). Dies gilt im allgemeinen für Sekundärstrukturelemente oder Bereiche, die auf Grund des Auftretens von interresidualen NOE-Kontakten auf eine definierte Struktur schließen lassen. Dies bedeutet für *cMeCP2 ARBD*, daß dem Bereich von Asp152 bis Arg163 besondere Aufmerksamkeit gewidmet werden muß, da der *CSI-plot* für diesen Bereich keine kanonischen Sekundärstrukturelemente vorhersagt, dies aber aus der Analyse der Berechnungen mit TALOS (siehe Abschnitt 3.1.4) nahe gelegt wird.

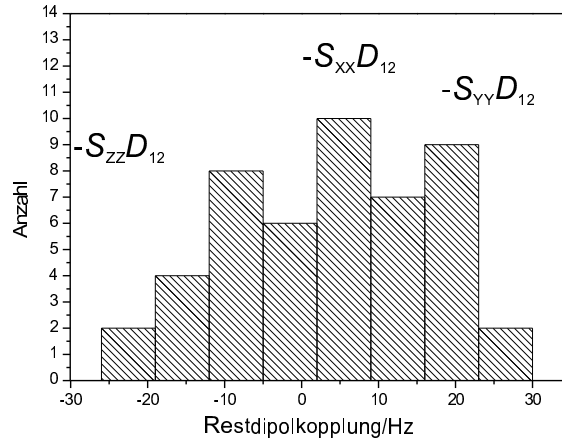


Abbildung 3.9: Histogramm der auf die $^1\text{H}_N$ - ^{15}N -Kopplung normierten Restdipolkopplungen. Es flossen die Restdipolkopplungen ($^1\text{H}_N$ - ^{15}N , $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ und ^{15}N - $^{13}\text{C}'$) für gut separierten Signale aus den durch CSI vorhergesagten Sekundärstrukturelementen ein. Die $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - und die ^{15}N - $^{13}\text{C}'$ -Kopplungen wurden auf die $^1\text{H}_N$ - ^{15}N -Kopplungen normiert (siehe Text).

Die Messung der Restdipolkopplungen ist in Abschnitt 2.4.4 näher beschrieben. Die Fehler ergeben sich aus den zwei Einzelmessungen (isotrop und orientiert) und addieren sich zum Gesamtfehler der Restdipolkopplung. Es ergaben sich im günstigsten Fall Fehler von 1,4 Hz. Für die $^1\text{H}_N$ - $^{13}\text{C}'$ -Kopplung war der Fehler auf Grund der großen Linienbreite in ^1H -Dimension für alle Aminosäurereste größer als die Kopplung selbst und somit nicht verwertbar. Es standen also Restdipolkopplungen bezüglich der $^1\text{H}_N$ - ^{15}N -, der $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - und der ^{15}N - $^{13}\text{C}'$ -Bindung zur Verfügung. Die Messungen der $^1\text{H}_N$ - ^{15}N -Kopplungen wurde an einer ^{15}N -markierten Probe vorgenommen. Das Deuteriumsignal (siehe Abschnitt 2.4.4) wies eine Aufspaltung von 18,4 Hz auf. Für die Messung der anderen beiden Kopplungen wurde eine ^{13}C - ^{15}N -markierte Probe eingesetzt, wobei sich eine Aufspaltung des Deuteriumsignals von 16,1 Hz einstellte.

Um diese drei Kopplungen miteinander vergleichen zu können, war es nötig, sie aufeinander zu normieren. Dies wurde in einem ersten Schritt durch Multiplikation der $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - und der ^{15}N - $^{13}\text{C}'$ -Restdipolkopplung mit $\frac{18,4\text{Hz}}{16,1\text{Hz}} \approx 1,14$ erreicht. In einem zweiten Schritt wurden die $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - und die ^{15}N - $^{13}\text{C}'$ -Restdipolkopplung auf die $^1\text{H}_N$ - ^{15}N -Kopplungen geeicht. Dazu wurden die Werte für $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ und ^{15}N - $^{13}\text{C}'$ mit dem

Quotienten der statischen Dipolkopplungskonstante aus $^1\text{H}_N$ - ^{15}N zu $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$, bzw. ^{15}N - $^{13}\text{C}'$ multipliziert.⁷ Dies geschah aus zweierlei Gründen: Zum einen ist die Normierung auf die $^1\text{H}_N$ - ^{15}N -Kopplung nötig für das Einbringen der $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - und der ^{15}N - $^{13}\text{C}'$ -Restdipolkopplungen in Strukturberechnungen mit dem Computerprogramm CNS (siehe Abschnitt 2.5). Zum anderen verleiht eine größere Anzahl einfließender Kopplungen dem Histogramm (Abbildung 3.9) eine bessere Interpretierbarkeit und höhere Aussagekraft.

Mit Hilfe des Histogramms und der in [Clo98] beschriebenen Methode ist es möglich, die Hauptwerte des Orientierungstensors abzuschätzen (ohne Kenntnis der Struktur). Hierbei gilt: $|S_{ZZ}| \geq |S_{YY}| \geq |S_{XX}|$. Diese Werte lassen sich einfach in die im Programm CNS verwendeten Größen umrechnen. Die in CNS verwendeten Größen heißen *rhombicity* und *axial* und für den Zusammenhang dieser beiden Größen mit den Hauptwerten des Orientierungstensors gilt:

$$axial = -\frac{1}{2} \cdot D_{12} \cdot S \cdot S_{ZZ} \quad (3.1)$$

$$rhombicity = -\frac{2}{3} \cdot \left(\frac{2 \cdot S_{YY}}{S_{ZZ}} + 1 \right) \quad (3.2)$$

Unter Berücksichtigung der Spurlosigkeit des Tensors läßt sich auch ohne weiteres S_{XX} alternativ zu S_{ZZ} oder S_{YY} in die Berechnung von *rhombicity* und *axial* einbeziehen.

Auf Grund der beschränkten Anzahl an gemessenen Restdipolkopplungen war die Bestimmung der Hauptwerte durch Vergleich mit einem idealen Histogramm nicht einfach und ließ Mehrdeutigkeiten zu. Ein weiteres Problem bezüglich der Auswertung des Histogramms bestand in der Tatsache, daß durch die Normierung auf die $^1\text{H}_N$ - ^{15}N -Kopplung die absoluten Fehler für die $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - und die ^{15}N - $^{13}\text{C}'$ -Restdipolkopplung mit der Normierung wachsen. Sowohl die statische Dipolkopplungskonstanten für $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ - als auch die für ^{15}N - $^{13}\text{C}'$ sind betragsmäßig kleiner als die statische Dipolkopplungskonstante für $^1\text{H}_N$ - ^{15}N . Für die Fehler der ^{15}N - $^{13}\text{C}'$ -Restdipolkopplung ergaben sich so Werte von z.T. mehr als 5 Hz. Dies war auch der Grund dafür, daß die ^{15}N - $^{13}\text{C}'$ -Restdipolkopplung zwischen Lys108 und Leu109 (28,5 Hz) nicht in die Betrachtungen zur Bestimmung der Hauptwerte mittels Histogramm einfloß. Da anzunehmen ist, daß für die Bestimmung von S_{XX} und

⁷Für die $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ -Restdipolkopplungen z.B. heißt das Multiplikation mit $\frac{-24\text{kHz}}{4,9\text{kHz}}$ (für die statischen Dipolkopplungskonstanten siehe Tabelle 2.1).

S_{YY} mit Hilfe des Histogramms der Einfluß der endlichen Anzahl an einfließenden Kopplungen geringer ist als bei S_{ZZ} , wurde S_{ZZ} nicht mit Hilfe des Histogramms bestimmt, sondern unter Ausnutzung der Spurlosigkeit des Tensors (siehe Abschnitt 2.4.4). Es ergab sich somit für die Hauptwerte des Tensors aus dem Histogramm: $S_{ZZ} \approx -1,04 \times 10^{-3}$, $S_{YY} \approx 8,3 \times 10^{-4}$ und $S_{XX} \approx 2,1 \times 10^{-4}$. Das entspricht Werten von *axial* und *rhomblcity* von -12,5 Hz bzw. 0,4. Durch leichte Variation der Auflösung des Histogramms kam es zu starken Schwankungen der Hauptwerte, die durchaus im Bereich von bis zu $\pm 1,5 \times 10^{-4}$ lagen.

Um die Hauptwerte des Orientierungstensors genauer zu bestimmen, wurde von jeweils verschiedenen Hauptwerten ausgehend eine Reihe CNS-Strukturberechnungen durchgeführt. In diesen Berechnungen wurden die mit TALOS vorhergesagten Diederwinkel Φ und Ψ , die $^1\text{H}_N$ - ^{15}N -Restdipolkopplungen und NOE-*restraints* (Anzahl interresidualer Kontakte 192) verwendet. In Abbildung 3.10 sind die Gesamtenergien aus den Strukturberechnungen in Abhängigkeit des jeweils verwendeten Wertepaar für *axial* und *rhomblcity* dargestellt. Es wurde jeweils ein Durchlauf durchgeführt, in dem 60 Strukturen berechnet wurden. Die Gesamtenergie der besten 20 Strukturen (hinsichtlich ihrer Gesamtenergie) wurde gemittelt und dieser Mittelwert ist in Abbildung 3.10 aufgetragen. Die Auswertung ergibt einen energetisch begünstigten Bereich bei etwa 0,5 für *rhomblcity* und -13 Hz für *axial*. Dies entspricht $S_{ZZ} \approx -1,08 \times 10^{-3}$, $S_{YY} \approx 9,4 \times 10^{-4}$ und $S_{XX} \approx 1,4 \times 10^{-4}$ für die Hauptwerte.

Da die Werte für S_{ZZ} und S_{YY} betragsmäßig nicht sehr verschieden sind, wurde das oben geschilderte Verfahren zur Bestimmung von *axial* und *rhomblcity* auch für den Fall angewandt, daß S_{ZZ} im Gegensatz zum oben angenommenen positiv ist. Dabei ergaben sich weit höhere Energien als im umgekehrten Fall, und es kann angenommen werden, daß S_{ZZ} negativ ist.

Es gibt eine weitere Möglichkeit, die Hauptwerte des Orientierungstensors abzuschätzen. Hierfür ist neben den experimentell bestimmten Restdipolkopplungen Voraussetzung, daß eine schon relativ gut definierte Struktur vorliegt. Für ein solches Vorhaben stehen eine Reihe Programme zur Verfügung, welche die Größe und Lage des Tensors in einem molekülfesten Koordinatensystems vorhersagen, wobei möglichst viele der experimentellen Restdipolkopplungen mit diesem Tensor erklärt werden sollen. In dieser

Arbeit kam hauptsächlich das Programm Module [Dos01] zum Einsatz. Dabei wurden unter Verwendung von über 650 NOE-Kontakten und den Diederwinkeln, vorhergesagt für Aminosäurereste in Sekundärstrukturelementen, mit dem Strukturberechnungsprogramm DYANA (siehe Abschnitt 2.5) 240 Strukturen berechnet. Die fünf energetisch besten Strukturen wurden für einen Fit mit Module herangezogen. Als Ergebnis hat sich -12 Hz für *axial* und 0,51 für *rhombicity*, was Hauptwerten von $S_{ZZ} \approx -1,0 \times 10^{-3}$, $S_{YY} \approx 9,1 \times 10^{-4}$ und $S_{XX} \approx 0,9 \times 10^{-4}$ entspricht, ergeben.

Vergleicht man alle drei Methoden die Hauptwerte des Orientierungstensors zu bestimmen⁸, so liefern sie im wesentlichen übereinstimmende Ergebnisse. Die mit dem Histogramm bestimmten Hauptwerte unterscheiden sich geringfügig von denen, die mit den beiden anderen Methoden bestimmt wurden. Auf Grund der geringeren Genauigkeit dieser Methode (s.o.) wurden für die anschließenden Strukturberechnungen mit CNS die mit den anderen beiden Methoden bestimmten Werte verwendet (*rhombicity*=0,51 und *axial*=-13 Hz)⁹ Die in die Strukturberechnung eingeflossenen Restdipolkopplungen sind im Anhang abgedruckt.

3.1.6 Zuordnung der NOESY-Spektren

Kern-Kern-Abstände, abgeleitet aus NOESY-Spektren, stellen bislang das wichtigste Hilfsmittel bei der Strukturbestimmung von biologischen Makromolekülen in Lösung dar. Im Fall von cMeCP2 ARBD war dies ähnlich, wobei außerdem Restdipolkopplungen und die Vorhersagen für die Werte der Diederwinkel mittels TALOS (siehe Abschnitt 3.1.5 und 3.1.4) eine wichtige Rolle bei der Strukturbestimmung von cMeCP2 ARBD spielten. Zusätzlich waren sie eine wertvolle Hilfe, um einen „Einstieg“ in die NOESY-Spektren zu erlangen.

Ein ¹³C- und ein ¹⁵N-editiertes NOESY-HSQC-Spektrum sowie ein zweidimensionales NOESY-Spektrum (das erste bei 600 MHz Protonenresonanzfrequenz, die letzten beiden bei 800 MHz Protonenresonanzfrequenz) standen zur Verfügung. In diesen Spektren

⁸Histogramm, Strukturberechnungen mit anschließender Energiebetrachtungen bei systematisch variierten Hauptwerten und das Anpassen des Tensors an eine gut definierte Struktur.

⁹In den Startdateien für CNS wird nur der Zahlenwert für *axial* eingegeben (ohne die Einheit).

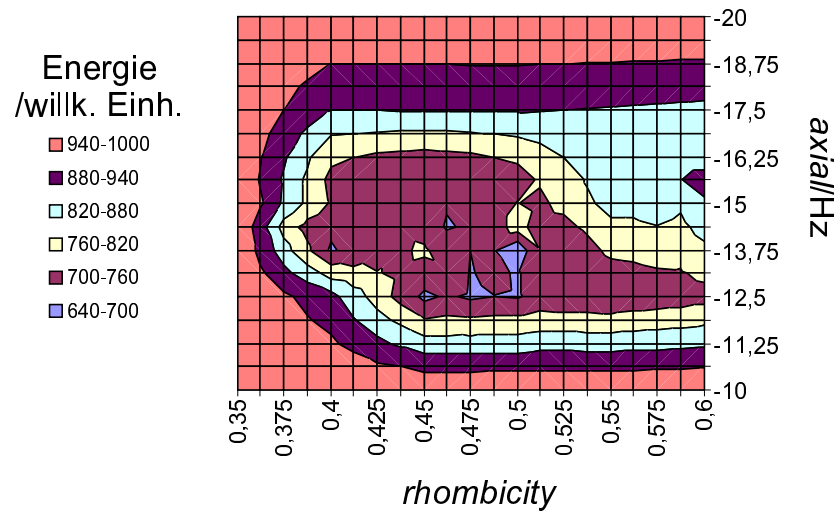


Abbildung 3.10: Gesamtenergie der einzelnen Strukturberechnungen in Abhängigkeit von axial und rhombicity. Im Text ist der Zusammenhang dieser beiden Größen mit den Hauptwerten des Orientierungstensors dargestellt.

kommt es zu sogenannten Kreuzsignalen, wenn zwei Protonen im Protein weniger als etwa 0,5 nm auseinanderliegen. Mit Hilfe der oben beschriebenen Zuordnung der Hauptketten- und Seitenkettensignale ist es möglich, diese beiden Protonen zu identifizieren. Das Volumen des Kreuzsignals läßt sich über eine bekannte Beziehung (abgeleitet aus dem Kern-Overhauser-Effekt) in Distanzen umrechnen. In Abschnitt 2.4.3 wird auf die verwendeten Spektren näher eingegangen, eine detaillierte Beschreibung der Experimente findet sich z.B. in [Cav95]. Auf Grund der Größe von cMeCP2 ARBD war die eindeutige Identifizierung der Signale wegen Signalüberlagerungen trotz der Verwendung von dreidimensionalen Experimenten oft kompliziert. Das Vorgehen zur Zuordnung der NOE-Signale war ein iterativer Prozeß: In einem ersten Schritt wurden in den NOESY-Spektren eindeutige NOE-Signale identifiziert. Die Abstandsinformationen aus diesen Signalen zusammen mit den Restdipolkopplungen und den Diederwinkeln (jeweils für Aminosäurereste aus mit Hilfe des CSI-plots vorhergesagten Sekundärstrukturelementen) bildeten die Grundlage für die ersten Strukturberechnungen. Für die Umrechnung der Signalintensitäten der Kreuzsignale auf Kern-Kern-Abstände wurden folgende Kreuzsignale verwendet:

Spektrum	Kreuzsignal	Distanz in nm
^{15}N -NOESY	H_N 125/ H_N 133	0,33
^{13}C -NOESY	H_α 137/ H_β 137	0,28
2D-NOESY	H_N 144/ H_N 143	0,28

Diese Kreuzsignale stammen alle aus Sekundärstrukturbereichen, für deren Abstände Literaturwerte bekannt sind [Wut86]. Die Eichung des ^{13}C -editierten NOESY-HSQC-Spektrum vollzog sich abweichend von den anderen beiden Spektren: In [Wut86] wird für den Abstand des H_α zum H_β eines Aminosäurerestes innerhalb einer α -Helix 0,22 bis 0,29 nm angegeben. Es wurde nun mit 0,26 nm geeicht, dabei kam es zu einer nicht zu vernachlässigenden Anzahl von unterschrittenen van-der-Waals-Radien. Daraufhin wurden die Distanzen solange in Schritten von 0,01 nm erhöht, bis nur noch wenige der berechneten Distanzen die Summe der van-der-Waals-Radien der beiden beteiligten Atome unterschritten. Auf diese Weise ergaben sich bei Betrachtung aller NOE-Kontakte maximale Distanzen von 0,5 nm, was mit den maximal durch NMR-Spektroskopie an Makromolekülen mittlerer Größe ermittelbaren Distanzen konform ist.

In einem zweiten Schritt wurden die aus dem ersten Schritt entstandenen Strukturen zur Rückrechnung der experimentellen Spektren (das Prinzip der Rückrechnung ist in Abschnitt 2.6 erklärt) verwendet. Mit dem Vorliegen einer groben Struktur verringerte sich die Anzahl der mehrdeutigen NOE-Kontakte und eine Zuordnung von vorher nicht identifizierbaren Signalen wurde möglich. Diese beiden Schritte wechselten sich mehrfach ab, bis es nicht mehr möglich war, den Grad der Zuordnung der NOESY-Spektren zu erhöhen. Zur Auswertung der Strukturberechnungen stand unter anderem das in der Arbeitsgruppe von Herrn Prof. Dr. Dr. Kalbitzer entwickelte Programm *cnsstat* [Ste01] zur Verfügung. Dieses Programm analysiert die *logfile*s einer CNS-Strukturberechnung und extrahiert unter anderem die am häufigsten verletzten NOE-*restraints*. Das *logfile* selbst ist auf Grund seiner Größe sehr unübersichtlich und eignet sich dadurch kaum zur Analyse der Rechnungen. Die verletzten NOE-*restraints* wurden in den Spektren überprüft und dann unter Umständen bei nachfolgenden Strukturberechnungen mit größeren Fehlern oder gar nicht mehr verwendet. Fehlerquellen waren zumeist vorher als solche nicht erkannte Signalüberlagerungen, selten auch Fehlzuordnungen der Signale. Andere Fehlerquellen sind (auf Grund von „Spindiffusion“) falsch ermittelte Distanzen zwischen zwei Atomen. Wa-

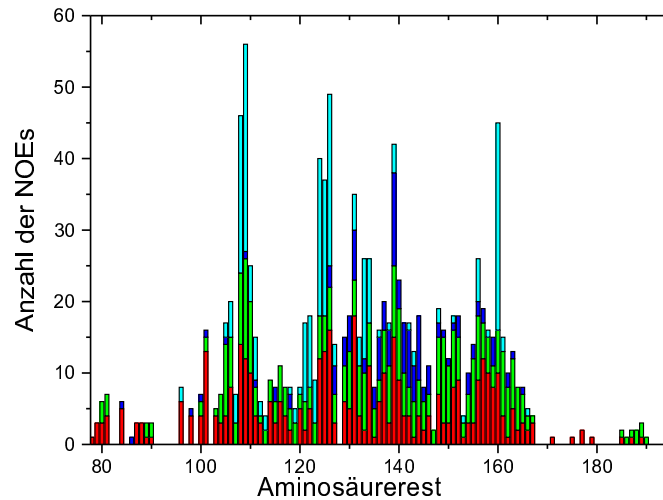


Abbildung 3.11: *Statistik der NOE-restraints*. Es sind die Summe der NOE-Kontakte (NOEs) pro Aminosäurerest dargestellt. Hellblau: langreichweitige NOEs (Abstand in der Sequenz ≥ 5), dunkelblau: NOEs (Sequenzabstand zwei bis vier Aminosäurereste), grün: sequentielle NOEs, rot: intraresiduelle NOEs.

ren solche Signale identifiziert, wurden sie in folgenden Strukturberechnungen nicht mehr berücksichtigt. Pauschal wurde ein Fehler von 30% bezüglich der Distanzinformationen aus den NOE-Kontakten angenommen, welcher in begründeten Fällen (s.o.) geändert wurde. Tabelle 3.2 und Abbildung 3.11 geben einen Überblick über die gefundenen NOE-Kontakte. In Abbildung 3.11 läßt sich erkennen, daß der Bereich von Aminosäurerest Gly104 bis Ser165 verhältnismäßig viele Kontakte aufweist,¹⁰ was für das Vorliegen wenig flexibler Strukturen in diesem Bereich spricht. Die Aminosäurereste der Termini des Proteins weisen hingegen selbst innerhalb der eigenen Seitenkette wenige bzw. gar keine Kontakte auf. Dies läßt auf eine hohe Flexibilität schließen, was wiederum auf das Vorliegen nicht gefalteter Bereiche hinweist. Diese Erkenntnis steht in Übereinstimmung mit den in den Abschnitten 3.1.3 und 3.1.4 dargestellten Resultaten, in denen unter anderem Vorhersagen über die Sekundärstruktur von cMeCP2 ARBD getroffen wurden. Die Anzahl der gefundenen NOE-Kontakte (siehe Tabelle 3.2) ist für ein Protein der Größe von cMeCP2 ARBD als relativ gering zu bewerten. Diese geringe Zahl an NOE-Kontakten

¹⁰mit Ausnahme des Bereichs von Aminosäurerest 110 bis 118

Art des Kontakts	Anzahl der Kontakte
intraresidual	447
sequentiell	186
mittelreichweitig	65
langreichweitig	131
Gesamt	829

Tabelle 3.2: *Überblick über die NOE-restraints.* Mittelreichweitig faßt Kontakte zwischen Atomen, deren Sequenzabstand zwei bis vier Aminosäurereste beträgt, zusammen. Ab fünf Aminosäureresten Abstand sind die Kontakte unter langreichweitig zusammengefaßt.

Art des <i>restraints</i>	Anzahl der <i>restraints</i> (CNS)	Anzahl der <i>restraints</i> (DYANA)
NOEs	829	829
Restdipolkopplungen	53	
Diederwinkel	62	62
Gesamtanzahl	944	891

Tabelle 3.3: *Überblick aller verwendeten restraints in DYANA und CNS.*

wird durch den hohen Anteil nicht gefalteter Bereiche des Proteins verursacht, die im Vergleich zu gefalteten Bereichen einen höheren Grad an Flexibilität aufweisen und somit wenig NOE-Kontakte aufweisen. Die flexiblen Bereiche (meist in der Nähe des C- und N-Terminus des Proteins) weisen zudem sehr ähnliche Werte für ihre chemischen Verschiebungen auf, was eine Zuordnung dieser Aminosäurereste besonders schwierig macht. Aus dieser geringen Zahl gefundener NOE-Kontakte leitete sich der hohe Stellenwert der Restdipolkopplungen als *restraints* in den durchgeführten Strukturberechnungen ab.

3.1.7 Numerische Berechnung der dreidimensionalen Struktur von cMeCP2 ARBD

Zur Berechnung der Struktur von cMeCP2 ARBD kamen die beiden in Abschnitt 2.5 näher beschriebenen Programme CNS und DYANA zum Einsatz. Die Gesamtheit der experimentellen *restraints*, die in den abschließenden Rechnungen verwendet wurden, sind

in der Tabelle 3.3 im Überblick aufgelistet. Die Aminosäurereste, die dem Polyhistidintag zugehörig sind (sie treten im nativen cMeCP2 ARBD nicht auf) wurden bei der Strukturberechnung nicht berücksichtigt.

Bei Moleküldynamikrechnungen besteht immer die Gefahr, daß der simulierte Faltungsweg in ein lokales Energieminimum führt, welches unter Umständen bis zum Ende der Simulation nicht mehr verlassen wird. Um so entstandene Strukturen von denen zu trennen, deren Faltungsweg näher am globalen Energieminimum endet, ist es nötig, eine große Anzahl von Strukturen (bei Verwendung der gleichen *restraints*) zu berechnen. Dann kann eine Auswahl betrachtet werden, deren Energien minimal sind. Aus diesem Grund wurden bei allen Rechnungen 240 Strukturen berechnet und die zehn Strukturen mit den niedrigsten Energien ausgewählt. Für diese zehn mit CNS berechneten Strukturen sind die Energiebeiträge (Mittelwert \pm Standardabweichung), welche durch die Verletzungen der *restraints* hervorgerufen werden, getrennt nach den einzelnen Beiträgen (siehe auch 2.5.1) in Tabelle 3.4 aufgeführt.¹¹ In dieser Tabelle werden zwei CNS-Rechnungen miteinander verglichen, denen die gleichen *restraints* bezüglich NOE-Kontakten und Diederwinkeln zu Grunde lagen. In einer Rechnung wurden zusätzlich Restdipolkopplungen als *restraints* in die Moleküldynamikrechnungen einfließen.

Die Beiträge aller Energien für die Berechnungen mit Restdipolkopplungen liegen deutlich höher als die entsprechenden Beiträge ohne Einbeziehung der Restdipolkopplungen. Auffällig ist, daß der Energiebeitrag direkt hervorgerufen durch verletzte Restdipolkopplungen (E_{coupl}) im Vergleich zu den anderen Energiebeiträgen gering ist (E_{coupl} macht weniger als 1% von der Gesamtenergie E_{pot} aus).

Die zehn energetisch besten mit DYANA berechneten Strukturen wiesen Werte für die VTF (siehe Abschnitt 2.5.2) von 34,78 für die energetisch beste Struktur bis 36,02 für die zehnte Struktur auf.

Auf Grund der Ergebnisse aus Abschnitt 3.1.3 und 3.1.4, die nicht gefaltete Elemente für Aminosäurereste in den Bereichen der Termini von cMeCP2 ARBD (Aminosäurerest 72-90 und 165-196) vorhersagen, und der Tatsache, daß nur wenige NOE-Kontakte in diesen Bereichen zu lokalisieren sind (siehe Abbildung 3.11), beschränkt sich die Dar-

¹¹Zugrundegelegt wurde hierbei die Standardeinstellung bezüglich der Kraftkonstanten der CNS-Version 1.0 (vgl. Abschnitt 2.5.1).

Art der Verletzung	Energie in kJ/mol (mit Kopplungen)	Energie in kJ/mol (ohne Kopplungen)
E_{vdW}	756 ± 74	158 ± 14
E_{bond}	146 ± 12	$51,8 \pm 1,4$
E_{imp}	284 ± 60	$28,2 \pm 1,8$
E_{ang}	1112 ± 218	324 ± 9
E_{NOE}	581 ± 97	239 ± 9
E_{coupl}	18 ± 9	
$E_{dihedral}$	167 ± 42	$4,3 \pm 0,7$
E_{pot}	3064 ± 303	805 ± 23

Tabelle 3.4: Vergleich der Energiebeiträge bei zwei CNS-Rechnungen (einmal mit, einmal ohne die Einbeziehung von Restdipolkopplungen als restraints).

stellung der Struktur in Abbildung 3.12 auf den Sequenzabschnitt von Aminosäurerest 82 bis 170. Die Darstellungsform ist das Bändermodell, bei dem insbesondere die Faltung der Hauptkette und die Anordnung der Sekundärstrukturelemente betont wird. Zur Darstellung wurde das Programm MOLMOL verwendet [Kon96]. Die Berechnung der Sekundärstrukturelemente erfolgte mit dem Kabsch-Sander-Algorithmus [Kab83], der in MOLMOL implementiert ist.

Vergleicht man die Ergebnisse der drei Strukturberechnungen (DYANA und CNS mit bzw. ohne Restdipolkopplungen als *restraints*) miteinander, ergibt sich folgendes Bild (siehe auch Tabelle 3.5): Übereinstimmend ergibt sich für die Struktur von cMeCP2 ARBD ein dreisträngiges antiparalleles β -Faltblatt, an das sich eine α -Helix anschließt. Die Längen und Positionen der einzelnen Sekundärstrukturelemente¹² wird in allen drei Rechnungen in guter Übereinstimmung wiedergegeben (siehe Tabelle 3.5).

Bei dem sich der α -Helix anschließenden Sequenzabschnitt lassen sich Unterschiede innerhalb der drei Strukturberechnungen erkennen: Bei den DYANA-Rechnungen ergibt sich eine fünf Aminosäurereste umfassende rechtsgängige Helix von Aminosäurerest 159 bis 163. Weiterhin auffällig ist der Bereich von Aminosäurerest 153 bis 155. Für neun der zehn

¹²bestimmt mit dem Kabsch-Sander-Algorithmus

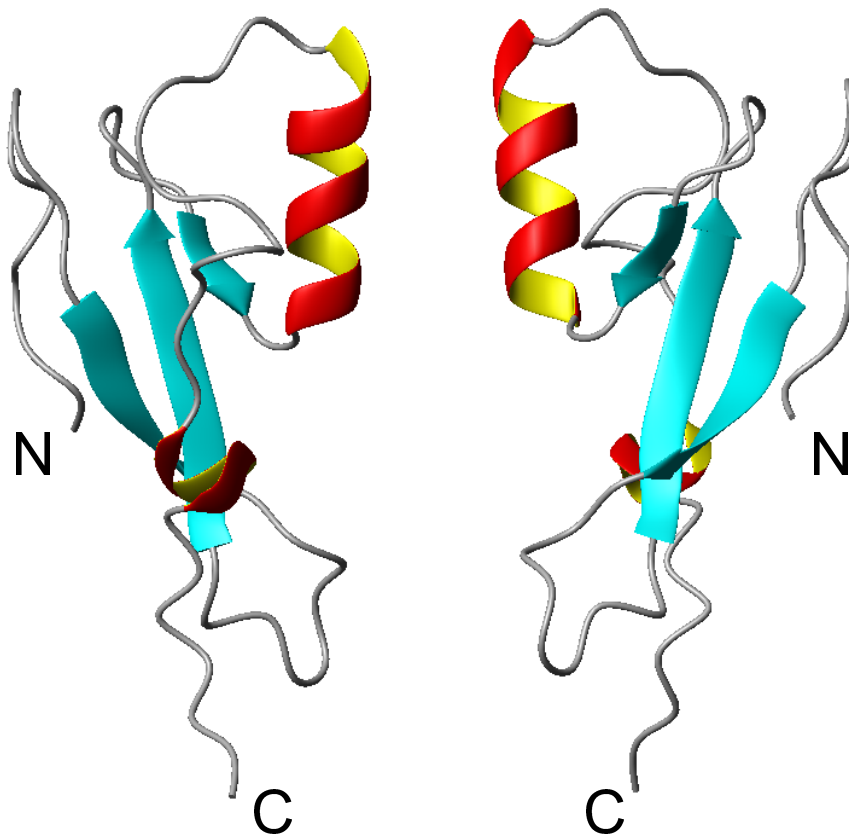


Abbildung 3.12: Struktur von cMeCP2 ARBD im Bändermodell in zwei verschiedenen Ansichten (Es ist die energetisch beste Struktur aus der Strukturberechnung mit DYANA dargestellt).

Sekundärstruktur (Sek.) nach [Kab83]		RMSD in nm					
		CSI		Sek.		95-170	
		RG	S	RG	S	RG	S
DYANA	FB: 106-110, 122-126, 132-133 Helix:136-144, 159-163	0,054	0,131	0,037	0,102	0,268	0,329
CNS	FB: 106-109, 122-126, 132-133 Helix:136-145	0,096	0,188	0,079	0,167	0,489	0,560
CNS oh. Koppl.	FB: 106-110, 122-126, 132-133 Helix:136-145	0,088	0,151	0,062	0,120	0,386	0,456

Tabelle 3.5: *RMSD-Werte und Sekundärstrukturelemente im Vergleich.* Schwere Atome (S) stehen dabei für N, C, S und O des gesamten Aminosäurerests; Hauptkettentome sind mit RG (Rückgrat) abgekürzt, FB steht für Faltblatt; Sek. steht für Sekundärstrukturelemente, bestimmt mit dem Kabsch-Sander-Algorithmus; CNS oh. Koppl. - es wurden keine Restdipolkopplungen als *restraints* verwendet. CSI steht für die Aminosäurereste, die mit Hilfe des CSI als Teile von Sekundärstrukturelementen vorhergesagt wurden (siehe Abbildung 3.4).

Strukturen wird mit dem Kabsch-Sander-Algorithmus eine von unstrukturiertem (*random coil*) Bereich abweichendes Strukturelement (*bend*) geliefert. Bei einer der zehn Strukturen lieferte der Algorithmus eine 3_{10} -Helix¹³ für den Sequenzbereich von 153 bis 155. Bei beiden Strukturberechnungen mit CNS werden nach der α -Helix (Aminosäurereste 136-145), bezogen auf alle zehn Strukturen, keine kanonischen Sekundärstrukturelemente durch den Kabsch-Sander-Algorithmus erkannt. Allerdings wird bei zwei Strukturen bzw. einer Struktur bei Verwendung von CNS mit bzw. ohne Restdipolkopplungen eine 3_{10} -Helix für den Sequenzabschnitt 153 bis 155 erkannt. In den übrigen acht bzw. neun Strukturen identifiziert der Kabsch-Sander-Algorithmus die Abschnitte um Aminosäurerest 154 und 161 nicht als *random coil*-Bereiche (*bend*).

Um die Güte der Struktur zu prüfen (speziell im Sequenzabschnitt 153-163), sind die Diederwinkel der einzelnen Aminosäurereste aufgetragen (Ramachandran Diagramm). In

¹³Ein im Vergleich zur α -Helix steiler gewundener Helixtyp (Es gilt: 3_{10} -Helix - 3 Aminosäurereste pro Windung, α -Helix - 3,6 Aminosäurereste pro Windung, π -Helix - 4,4 Aminosäurereste pro Windung [Cre93]).

Abbildung 3.13 sind im Ramachandran-Diagramm die Diederwinkel im Sequenzbereich der Aminosäurereste 82 bis 170 für die in Abbildung 3.12 gezeigte Struktur dargestellt.¹⁴ Es ist eine Häufung der (Φ, Ψ) -Paare der einzelnen Aminosäurereste in den farbig markierten Bereichen (sie werden aus sterischen Gründen bevorzugt besetzt) zu erkennen. Aminosäurereste, die nicht in den farbig markierten Bereichen liegen, stammen meist aus weniger gut definierten *loop*-Bereichen oder den Termini. Weiterhin auffällig sind die Glycinreste (in der Abbildung eingerahmt dargestellt). Diesen steht auf Grund der fehlenden Seitenkette ein größerer Konformationsraum offen. Auf Grund der unterschiedlichen Ergebnisse der drei Strukturberechnungen im Hinblick auf die Sekundärstruktur in dem Sequenzabschnitt von Aminosäurerest 153 bis 164 wurde dieser Bereich in Abbildung 3.14 in mehreren Ramachandran-Diagrammen im Vergleich dargestellt. Zur besseren Übersicht wurde nur eine (die energetisch Beste) der zehn Strukturen dargestellt.¹⁵ In allen drei Strukturberechnungen liegen die (Φ, Ψ) -Paare für die Aminosäurereste 153, 154, 160, 161, 163 in dem Winkelbereich, der gehäuft von Aminosäureresten aus helikalen Strukturelementen eingenommen werden. Das Winkelpaar von Aminosäurerest 155 liegt nur für die Strukturberechnung mit CNS unter Verwendung von Restdipolkopplungen nicht in diesem Bereich. Das heißt, bis auf diesen Aminosäurerest und Aminosäurerest 159 nehmen bei allen drei Rechnungen die Aminosäurereste im Bereich um 154 und 161, getrennt von einem *random coil*-Bereich (156-158), eine für helikale Strukturelemente typische Konformation ein (Aminosäurerest 162 kann bei dieser Betrachtung ausgeschlossen werden, da es sich um einen Glycinrest handelt (s.o.)). Es ergibt sich also für den Sequenzabschnitt 153 bis 163 eine Strukturelement der Form (*helix-turn-helix*). Dies ist im Hinblick auf die Funktion des Proteins (Bindung an DNA) ein interessantes Ergebnis, denn solche *helix-turn-helix*-Motive sind bekannte Interaktionsstellen zwischen Proteinen und DNA (siehe z.B. [Lod99]).

Ein Maß dafür, wie weit sich die zehn energetisch besten Strukturen voneinander unterscheiden, liefert der RMSD-Wert (*root mean square deviation*). In Tabelle 3.5 ist der RMSD-Wert (berechnet durch MOLMOL, bezogen auf eine mittlere Struktur) für Struk-

¹⁴Aus Abbildung 3.7 ist die Definition der Diederwinkel zu entnehmen.

¹⁵Es ist anzumerken, daß alle drei Strukturen Sekundärstrukturelemente wie in Tabelle 3.5 angegeben aufweisen, d.h., sie sind „gute“ Repräsentanten der jeweils zehn Strukturen.

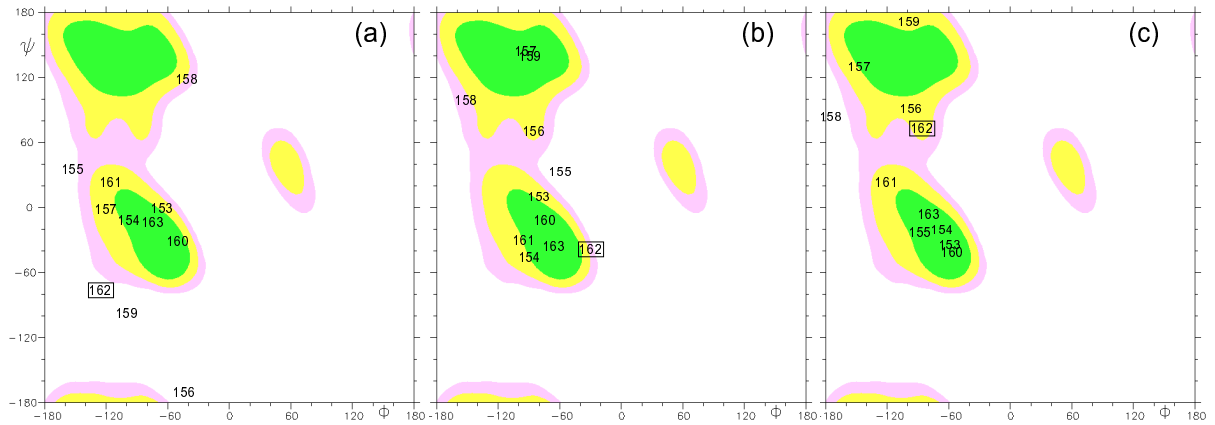


Abbildung 3.14: Der Sequenzabschnitt von Pro153 bis Arg163 im Ramachandran Diagramm. Die Strukturberechnungen wurden mit (a) DYANA, (b) CNS, mit und (c) CNS, ohne die Verwendung von Restdipolkopplungen als *restraints* durchgeführt. Glycinreste sind eingerahmt dargestellt.

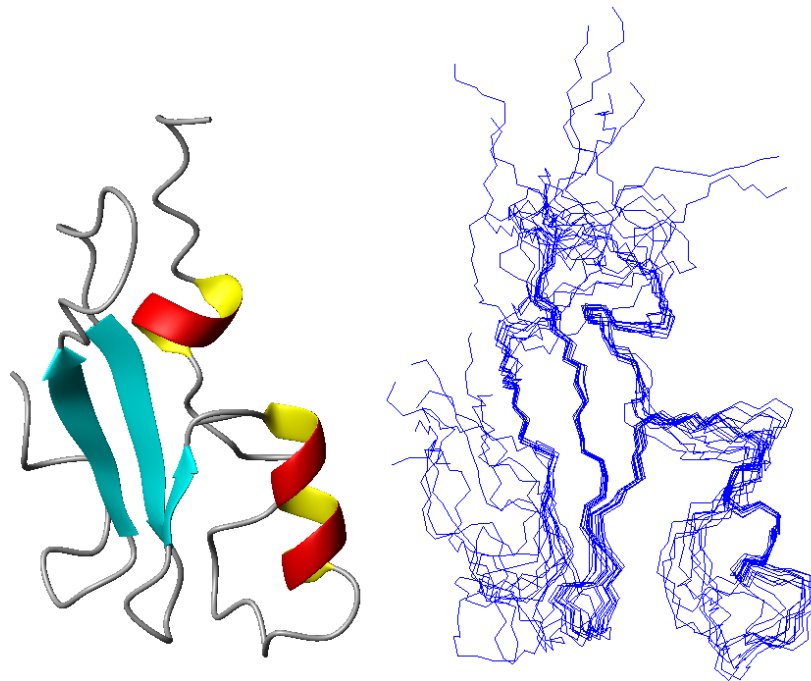


Abbildung 3.15: Überlagerung der Hauptkettenatome von zehn Strukturen niedrigster Energie. Zur besseren Veranschaulichung ist links eine der Strukturen des Bündels mit gleicher Orientierung im Bändermodell dargestellt (Strukturen wurden mit DYANA berechnet).

turberechnungen mit CNS und DYANA im Vergleich aufgetragen. Es wurde dabei zwischen Haupt- und Seitenkettenatomen unterschieden (siehe Tabelle 3.5). Weiterhin sind in dieser Tabelle die durch den Kabsch-Sander-Algorithmus für die jeweils zehn Strukturen erkannten Sekundärstrukturelemente festgehalten.¹⁶ Für verschiedene Sequenzabschnitte des Proteins wurde der RMSD-Wert bestimmt. Je niedriger der RMSD-Wert ist, um so weniger unterscheiden sich die zehn Strukturen (bezogen auf den zu Grunde gelegten Sequenzabschnitt) voneinander. Auffällig ist, daß die Strukturen, die mit DYANA berechnet wurden, für alle Sequenzabschnitte, sowie für alle in die Bestimmung des RMSD-Wertes einfließenden Atome geringere RMSD-Werte aufweisen als die mit CNS berechneten Strukturen. Vergleicht man die beiden CNS-Rechnungen miteinander, so ergeben sich für die Berechnungen mit Restdipolkopplungen durchweg höhere RMSD-Werte. Wird der Sequenzbereich von 95 bis 170 für die Berechnung des RMSD-Wertes herangezogen, so ergeben sich vergleichsweise hohe Werte. Das bedeutet, daß dieser Abschnitt viele nicht gut definierte Bereiche aufweist (vornehmlich der *loop* zwischen den ersten beiden Strängen des β -Faltblattes, der Bereich bis zum ersten β -Faltblattstang und der Bereich hinter Aminosäurerest 163). Die Überlagerung der zehn Strukturen niedrigster Energie ist in Abbildung 3.15 dargestellt. Die Positionen der Sekundärstrukturelemente sind sehr gut definiert. Die *loop*- sowie die terminalen Bereiche sind in ihrer Lage weniger gut definiert, was mit der Anzahl der gefundenen NOE-Kontakte in diesen Bereichen (siehe auch Abbildung 3.11) korreliert.

3.2 Der Protein-DNA-Komplex

Die Identifizierung von Interaktionsflächen verschiedener Moleküle mittels Flüssigkeits-NMR-Spektroskopie beruht oft auf der Beobachtung der Änderung der chemischen Verschiebung einer der beiden Wechselwirkungspartner im Vergleich zum Spektrum des ungebundenen Moleküls. Voraussetzung hierfür ist die Kenntnis der chemischen Verschiebungen der Signale eines der beiden Wechselwirkungspartner („sequenzielle Zuordnung“). Die nahezu vollständige Zuordnung der Hauptkettenatome (s.o.) gestattet die Untersuchung

¹⁶Trat bei mindestens acht der zehn Strukturen das gleiche Sekundärstrukturelement auf, so wurde dieses Element in Tabelle 3.5 festgehalten.

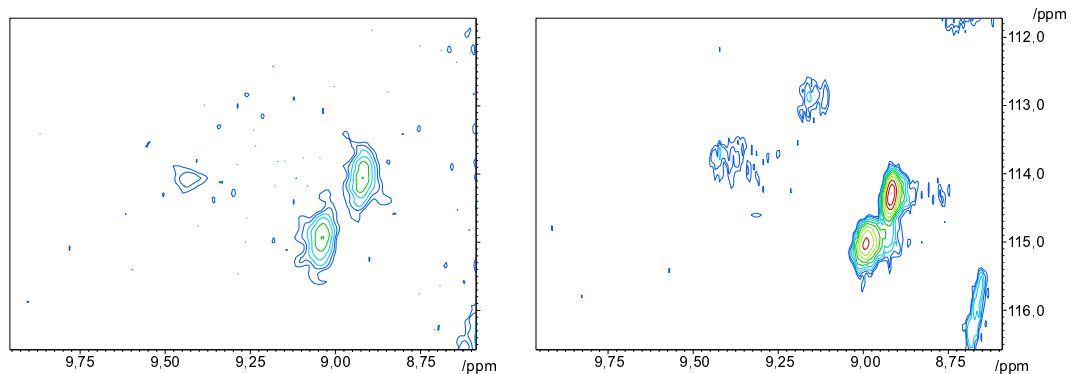


Abbildung 3.16: Ausschnitt aus dem (links) ^1H - ^{15}N -HSQC- und dem (rechts) ^1H - ^{15}N -TROSY-Spektrum vom cMeCP2 ARBD gebunden an methylierte DNA (Probe 2, siehe Abschnitt 2.3). Die Konzentration von cMeCP2 ARBD betrug 1,6 mM, die Aufnahmen erfolgten bei einer Protonenresonanzfrequenz von 600 MHz.

der Bindung von methylierter (Abschnitt 3.2.1) bzw. nicht methylierter (Abschnitt 3.2.2) DNA an cMeCP2 ARBD. Aus [Wei97] ist zu entnehmen, daß ein Protein-DNA-Komplex im Verhältnis 1:1 vorliegt. Mit einem Molekulargewicht des DNA von ca. 9,2 kDa ergibt sich für den Komplex ein Molekulargewicht von ca. 25 kDa. Wie in Abschnitt 2.2.1 diskutiert, erscheint bei der Größe des Komplexes die Verwendung von TROSY-Experimenten bereits sinnvoll. Um ein möglichst genaues Bild des Bindungsverhaltens zu erlangen, ist es nötig, kleine Änderungen der chemischen Verschiebung nachweisen zu können. Die Erhöhung der Auflösung der Spektren in den indirekten Dimensionen ist proportional zur Dauer der Messung. Weiterhin erhöht sich, um ein ähnliches Signal/Rausch-Verhältnis wie bei einem kleineren Molekül zu erhalten, die Anzahl der *scans*. Daraus erklärt sich die überwiegende Verwendung von zweidimensionalen Experimenten bei der Untersuchung der Bindung von DNA an cMeCP2 ARBD.

3.2.1 Methylierte DNA

In Abbildung 3.16 sind Ausschnitte aus dem ^1H - ^{15}N -HSQC- und dem ^1H - ^{15}N -TROSY-Spektrum des cMeCP2 ARBD-DNA-Komplexes gezeigt. Das Signal/Rausch-Verhältnis beider Spektren ist in Anbetracht der hohen Anzahl *scans* (64 bzw. 128 beim HSQC- bzw. TROSY-Experiment) überraschend schlecht. Die Linien sind stark verbreitert im Vergleich zu den Signalen des freien cMeCP2 ARBD. Diese Verbreiterung konnte durch

die Verwendung des TROSY-Experiments nur geringfügig verringert werden. Eine Erklärung für diesen Effekt könnte das Vorliegen von Komplexen sein, welche mehr als ein Proteinmolekül und ein DNA-Fragment umfassen (Aggregation). Eine Vergrößerung des zu untersuchenden Komplexes hat eine Erhöhung von τ_c zur Folge, was direkt zu einer Erhöhung der Linienbreite führt (siehe auch Abschnitt 2.2.1).

Es ist möglich, durch Messung der Diffusionskoeffizienten mit Hilfe der Flüssigkeits-NMR-Spektroskopie Aussagen über die Größe eines Moleküls zu treffen.¹⁷ Aus einem Vergleich des Diffusionsverhaltens von freiem cMeCP2 ARBD mit dem des cMeCP2 ARBD-DNA-Komplexes soll die relative Größe des Komplexes abgeschätzt werden. Im linken Teil von Abbildung 3.17 ist die Änderung der normierten Intensität in Abhängigkeit von der Gradientenstärke für reine cMeCP2 ARBD-Lösung (a) und den Komplex (c) dargestellt. Der mit (b) bezeichnete Graph ist für ein Molekül mit der Form von cMeCP2 ARBD und einem Molekulargewicht von 25 kDa berechnet worden. Der Komplex weist einen etwa viermal größeren Diffusionskoeffizienten als das Protein allein auf. Ein Rückschluß auf die Größe des Komplexes anhand dieses Faktors ist nur bedingt möglich, da die Form des Moleküls einen entscheidenden Einfluß auf sein Diffusionsverhalten hat. Allerdings ist es durchaus möglich eine grobe Abschätzung zu treffen: Demgemäß ist der Komplex 1,5 bis 2,5 mal größer als der zu erwartende 1:1 Komplex aus cMeCP2 ARBD und dem in Abschnitt 2.3 beschriebenen 15 Basenpaare umfassenden DNA-Fragment. Dies wiederum läßt darauf schließen, daß der Komplex in Form größerer Aggregate anstelle von 1:1-Komplexen vorliegt. Um den Einfluß der Ionenkonzentration des Puffers zu untersu-

¹⁷Das Prinzip der Diffusionsmessung beruht darauf, daß ein linearer Gradient dem äußeren Magnetfeld überlagert wird. Der Betrag des Gradient bleibt unverändert, aber das Vorzeichen kehrt sich nach der Hälfte der Entwicklungszeit um. Das Gradientenfeld wechselnden Vorzeichens macht sich nur bei Molekülen bemerkbar, die während der Messung ihre Lage im Raum geändert haben. Diffundiert das Molekül während der Messung, so wird es sich unter Annahme einer Brownschen Bewegung nach dem Vorzeichenwechsel des Gradienten im allgemeinen unter dem Einfluß eines anderen effektiven Magnetfeldes entwickeln als vor dem Wechsel. Das hat zur Folge, daß die einzelnen Beiträge zur Gesamtmagnetisierung (Vektorbild) zum Zeitpunkt der Detektion nicht die gleiche Phase haben und sich somit zu einer geringeren Gesamtmagnetisierung aufaddieren. Die Abnahme der Signalintensität bei wachsenden Gradientenfeldern ist ein Maß für die Diffusion des Moleküls. Mit der Stokes-Einstein-Gleichung ergibt sich die Beziehung zwischen der Größe des Moleküls und seinem Diffusionsverhalten. Für detailliertere Informationen zu Diffusionsmessungen sei auf z.B. [Kan00] verwiesen.

chen, wurde die Natriumchloridkonzentration in Schritten von 10 mM erhöht. Nach einer Inkubationszeit von mindestens vier Stunden bei Raumtemperatur wurden Diffusionsmessungen durchgeführt und ^1H - ^{15}N -TROSY-Spektren aufgenommen. Die Aufnahme der ^1H - ^{15}N -TROSY-Experimente erfolgte, um festzustellen, ob cMeCP2 ARBD noch an die DNA gebunden vorliegt. Ein Aufbrechen des Komplexes würde in einer Signalaufspaltung resultieren (ein Signal für DNA-gebundenes cMeCP2 ARBD und eines für freies cMeCP2 ARBD). Eine solche Aufspaltung konnte nicht beobachtet werden. Die Resultate der Diffusionsmessungen bei verschiedenen Natriumchloridkonzentrationen sind in Abbildung 3.17 (rechts) dargestellt. Bei einer Natriumchloridkonzentration von 50 mM zeigt der Komplex ein Diffusionsverhalten, welches dem von cMeCP2 ARBD mit einem hypothetischen Gewicht von 25 kDa nahe kommt (Graph (b)). Das ^1H - ^{15}N -TROSY-Spektrum bei dieser Pufferzusammensetzung zeigte keine Signalaufspaltung, d.h. es lag kein freies cMeCP2 ARBD in der Lösung vor (siehe dazu auch Abbildung 3.18) und das Signal/Rausch-Verhältnis der Spektren gegenüber den Spektren bei den vorangegangenen Pufferbedingungen stieg. Weitere Messungen wurden an einem anderen DNA-Fragment (Probe 3 aus Abschnitt 2.3) mit gleicher Sequenz und Länge wie Probe 2 durchgeführt, allerdings sind bei Probe 3 die beiden Stränge nicht über einen *hairpin* miteinander verbunden. Ein Vergleich zwischen Spektren des Komplexes mit Probe 2 und mit Probe 3 ergibt keine Unterschiede, was den Schluß zuläßt, daß der *hairpin* die Komplexbildung nicht beeinflußt und daß die Einzelstränge des DNA-Fragmentes über die Wasserstoffbrückenbindungen der Basenpaare stabil zusammengehalten werden. Zur Untersuchung des cMeCP2 ARBD-DNA-Komplexes wurde die „TROSY-Version“ des HN(C)O-, HNCO- und HNCA-Experiment sowie das ^1H - ^{15}N -TROSY-Experimentes verwendet. Eine Zuordnung, wie sie in Abschnitt 3.1.1 für freies cMeCP2 ARBD erfolgte, war für den Komplex nicht möglich. Das geringe Signal/Rausch-Verhältnis der dreidimensionalen Spektren machte eine Auswertung dieser Spektren schwierig. Auf Grund der Verwendung einer zu 70% deuterierten Probe konnte das bei der Untersuchung des freien cMeCP2 ARBD zum Einsatz gekommene CBCA(CO)NH-Experiment nicht verwendet werden (siehe Abschnitt 2.2.2). Die Alternative zu diesem Spektrum für deuterierte Proben (das CCC(CO)NH-Experiment) zeigte ein so geringes Signal/Rausch-Verhältnis, daß keine Auswertung möglich war. Daher mußte die Zuordnung vor allem auf der Grundlage zweidimensionaler Spektren durchgeführt werden.

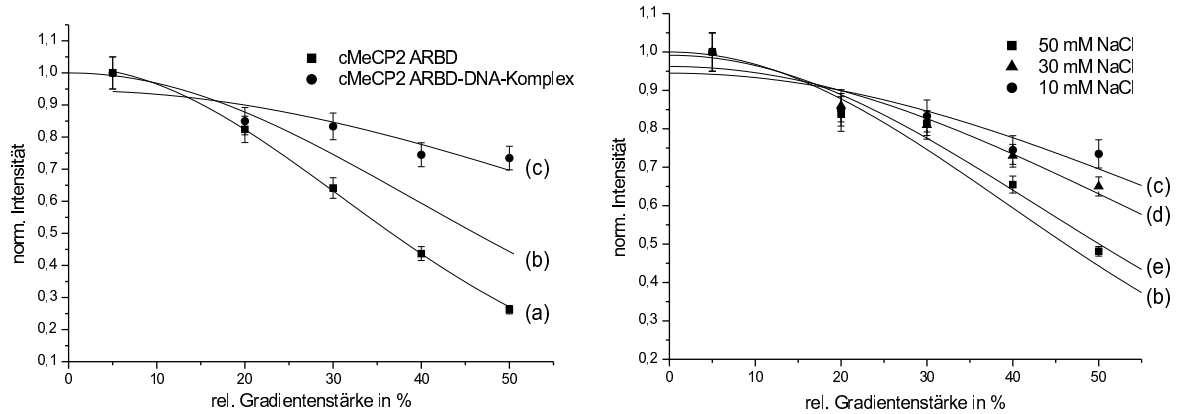


Abbildung 3.17: *Diffusionsmessungen.*(a) freies cMeCP2 ARBD, (b) simulierter 1:1 Komplex, (c)-(e) cMeCP2 ARBD-DNA-Komplex bei verschiedenen Natriumchloridkonzentrationen.

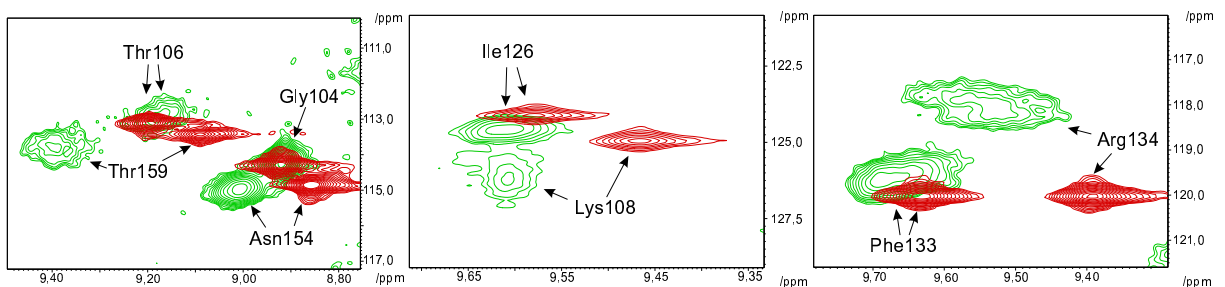


Abbildung 3.18: *Änderung der chemischen Verschiebung bei Komplexbildung von cMeCP2 ARBD mit methylierter DNA.* Grün: ^1H - ^{15}N -TROSY-Spektrum vom cMeCP2 ARBD-DNA-Komplex, Rot: ^1H - ^{15}N -HSQC-Spektrum von freiem cMeCP2 ARBD.

Um den Einfluß von Austauschereffekten zu untersuchen, wurde die Meßtemperatur in Schritten von 2 K von 292 K bis 302 K erhöht und bei jeder Temperatur ein ^1H - ^{15}N -TROSY-Spektrum aufgenommen. Es konnten weder eine Änderung der Linienbreite noch eine Änderung der Verschiebung der Signale beobachtet werden.

Es war möglich, etwa 25% der chemischen Verschiebungen der $^{13}\text{C}'$ -, ^{15}N - und $^1\text{H}_\text{N}$ -Kerne zuzuordnen. Ein höherer Grad der Zuordnung war auf Grund der teilweise großen Veränderungen der chemischen Verschiebungen in den stark überlagerten Bereichen (siehe Abbildung 3.1) nicht erreichbar.¹⁸ Um herauszufinden, ob Änderungen der chemischen

¹⁸Auf Grund des vergleichsweise schlechten Signal/Rausch-Verhältnisses war eine Bestimmung der

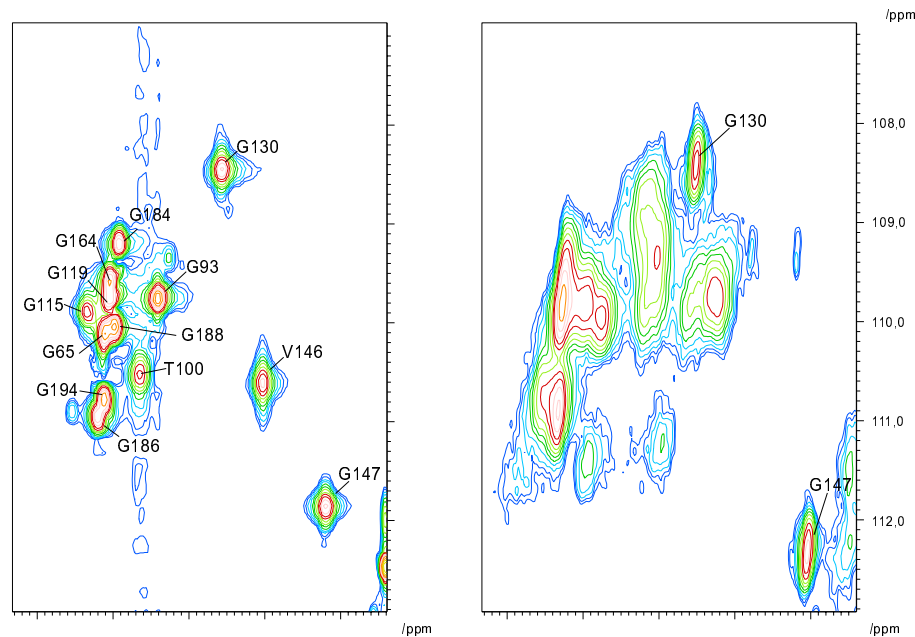


Abbildung 3.19: Ausschnitt aus dem ^1H - ^{15}N -TROSY-Spektrum vom cMeCP2 ARBD-DNA-Komplex (methylierte DNA) unter veränderten Pufferbedingungen (rechts) und dem ^1H - ^{15}N -HSQC-Spektrum vom freien cMeCP2 ARBD (links).

Verschiebungen mit den unterschiedlichen Pufferbedingungen (freies cMeCP2 ARBD - 154 mM Natriumchlorid, cMeCP2 ARBD-DNA-Komplex - 50 mM Natriumchlorid) zusammenhängen, wurde das freie Protein unter den Pufferbedingungen des Komplexes gemessen. Es zeigte sich, daß sich weder die Signalpositionen noch die Verhältnisse der Signalintensitäten innerhalb des Spektrums ändern (Pufferbedingungen siehe Abschnitt 2.3). Bindungsstudien von methylierter DNA an rMeCP2 MBD [Wak99] und MBD1 MBD [Ohk99], ergaben deutliche Änderungen der chemischen Verschiebungen in vier Abschnitten des Proteins. Bei cMeCP2 ARBD entspricht das dem Abschnitt von Aminosäurerest 88 bis 93, dem *loop* zwischen den ersten beiden Strängen des β -Faltblattes (Aminosäurereste 110 bis 121), dem Abschnitt zwischen dem dritten β -Faltblattstrang und der α -Helix (Aminosäurereste 134 bis 136) sowie dem Abschnitt von Aminosäurerest 158 bis 164. Die restlichen Aminosäurereste erfahren wesentlich geringere Änderungen ihrer chemischen

Änderung der chemischen Verschiebung bei Komplexbildung analog zu den Untersuchungen mit nicht methylierter DNA (Abschnitt 3.2.2) nicht sinnvoll. Das „Verschwinden“ eines Signals im Rauschen (was bei stark verbreiterten Signalen zwangsläufig erfolgt), hätte u.U. die Zuordnung einer zu großen Änderung der chemischen Verschiebung zur Folge.

Verschiebungen bei der Komplexbildung. In beiden Arbeiten wurden DNA-Fragmente verwendet, die dem in dieser Arbeit verwendeten DNA-Fragment ähnlich sind, so daß ein Vergleich der Resultate möglich ist. Für einen Teil der oben genannten Abschnitte konnten die chemischen Verschiebungen der Aminosäurereste von cMeCP2 ARBD im DNA-gebundenen Zustand zugeordnet werden. Beispielhaft ist der Einfluß der Komplexbildung auf die chemischen Verschiebungen für einige Aminosäurereste aus diesen Bereichen in Abbildung 3.18 dargestellt.

Der Anteil der Aminosäurereste in den terminalen Bereichen des Proteins (72 bis 95 und 165 bis 196), die bei der Bindung an methylierte DNA beteiligt sind, ist schwer abzuschätzen. Da wegen der oben geschilderten Empfindlichkeitsprobleme vorwiegend auf zweidimensionale Spektren zurückgegriffen werden mußte, war eine Identifizierung der im freien Protein in nicht gefalteten Bereichen liegenden Aminosäurereste in den Spektren des Protein-DNA-Komplexes nur selten möglich. Zur Veranschaulichung sei auf Abbildung 3.19 verwiesen, in der für das freie cMeCP2 ARBD mehrere Aminosäurereste des N- und C-Terminus dargestellt sind. Im Spektrum des Komplexes sind die Signale zum größten Teil nicht identifizierbar. Gründe hierfür waren das vergleichsweise niedrige Signal/Rausch-Verhältnis, die teilweise große Änderung der chemischen Verschiebung und der große Anteil nicht gefalteter Abschnitte des freien Proteins, was zu sehr ähnlichen chemischen Verschiebungen der Aminosäurereste in diesen Abschnitten des Proteins führt.

3.2.2 Nicht methylierte DNA

Neben Bindungsstudien mit methylierter DNA (s.o.) wurde auch die Wechselwirkung von cMeCP2 ARBD mit einem DNA-Fragment, dessen CpG-Motiv nicht methyliert ist, untersucht. Die Sequenz ist in Abschnitt 2.3 (Probe 1) dargestellt. Für die Pufferbedingungen wurde wie bei den ersten Messungen mit methylierter DNA eine Natriumchloridkonzentration von 10 mM gewählt (siehe Abschnitt 2.3 für die kompletten Puffer- und Aufnahmebedingungen). Das Verhältnis von Protein zu DNA betrug 1:2,4. Das Signal/Rausch-Verhältnis liegt bei vergleichbaren Aufnahmebedingungen deutlich höher als bei den Spektren an methylierter DNA. Die Änderungen der chemischen Verschiebungen sind ähnlich den Änderungen, die bereits bei den Messungen mit methylierter DNA auftraten. Da es nicht zu einer Signalaufspaltung, sondern zu einer vollständigen Verschiebung der Signale

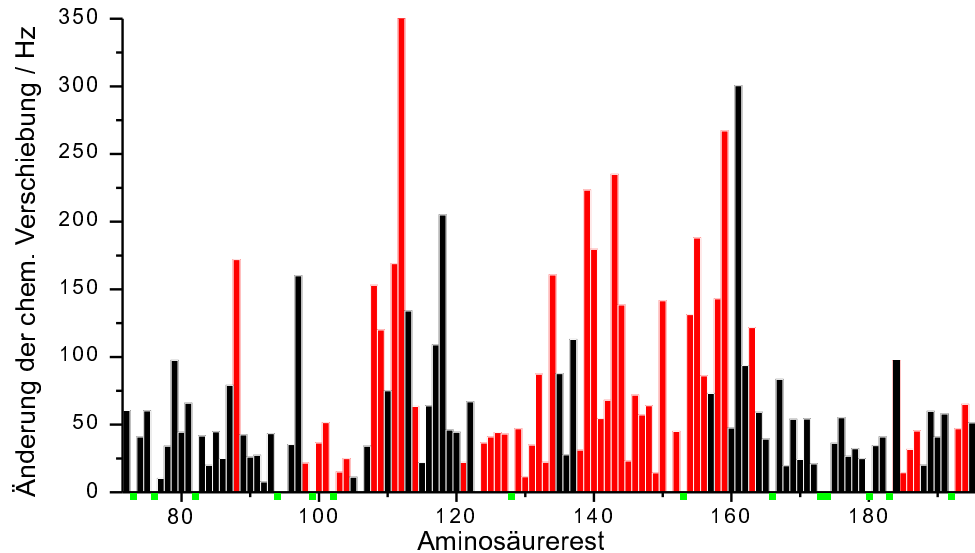


Abbildung 3.20: Änderung der chemischen Verschiebung der ^1H - und ^{15}N -Kerne der Hauptkette von cMeCP2 ARBP bei Komplexbildung mit methylierter DNA (nähere Erläuterung siehe Text).

kommt, liegt entweder schneller Austausch vor oder langsamer Austausch, wobei dann alle Proteinmoleküle im DNA-gebundenen Zustand vorliegen.¹⁹ Da ein ^1H - ^{15}N -HSQC-Spektrum bei einem Verhältnis von Protein zu DNA von 1:1,2 keine Signalverschiebungen gegenüber dem Spektrum mit einem Mischungsverhältnis von 1:2,4 aufweist, kann auf das Vorliegen von langsamem Austausch geschlossen werden.

Auf Grund des besseren Signal/Rausch-Verhältnisses war es möglich, die chemischen Verschiebungen des Komplexes mit den Verschiebungen des freien cMeCP2 ARBD zu vergleichen (siehe Abbildung 3.20). Grundlage ist ein ^1H - ^{15}N -HSQC-Experiment mit 128 scans. Für rot markierte Aminosäurereste war es möglich, dem Signal im Spektrum einen Aminosäurerest zuzuordnen. Der Wert für die Änderung der chemischen Verschiebung ergibt sich aus einer Differenzmessung zwischen dem jeweiligen Aminosäurerest im nativen und DNA-gebundenen Zustand für die beiden spektralen Dimensionen (^1H , ^{15}N) getrennt von einander, mit anschließender Addition der Beträge in Hz. Die schwarz markierten Aminosäurereste konnten nicht zugeordnet werden. Hier ergibt sich der Wert für

¹⁹langsam bzw. schnell im Sinne der NMR-Spektroskopie

die Änderung der chemischen Verschiebung folgendermaßen: Es wurde der Abstand zwischen dem Signal im Spektrum des freien cMeCP2 ARBD zu dem nächstliegenden, nicht zugeordneten Signal im Spektrum des DNA-gebundenen cMeCP2 ARBD gemessen. Unter der Annahme, daß das Signal im Spektrum des DNA-gebundenen cMeCP2 ARBD nicht so sehr verbreitert ist bzw. so schwach ist, daß es nicht vom Rauschen zu unterscheiden ist, stellt der so ermittelte Abstand eine untere Grenze für die Änderung der chemischen Verschiebung dar.²⁰ Für Prolinreste (grün markiert in Abbildung 3.20) konnte auf Grund des fehlenden $^1\text{H}_N$ -Kerns kein Vergleich zwischen den chemischen Verschiebungen angestellt werden. Für die Aminosäurereste Met95 und Val123 war ebenfalls kein Vergleich möglich, da für diese Aminosäurereste im Falle des freien cMeCP2 ARBD keine Zuordnung möglich war.²¹

Änderungen der chemischen Verschiebung, deren Beträge größer als 100 Hz sind, lassen sich gehäuft in drei Sequenzabschnitten erkennen. Dieses sind der *loop* (Aminosäurereste 111-121) zwischen den ersten beiden Strängen des β -Faltblattes, der Abschnitt vor und innerhalb der α -Helix (Aminosäurereste 134-143) sowie der Abschnitt von Aminosäurerest 154 bis 163. Weiterhin sind Änderungen der chemischen Verschiebung im Bereich der Termini zu beobachten (Aminosäurerest 88 und 184), wobei diese Änderungen weniger ausgeprägt sind als in den drei oben genannten Bereichen. Dazu ist anzumerken, daß nur wenige Aminosäurereste aus den Termini zugeordnet werden konnten, was bedeutet, daß die tatsächlichen Änderungen der chemischen Verschiebungen in diesen Bereichen unter Umständen größer sind als in Abbildung 3.20 dargestellt. Um einen besseres Bild vom Einfluß der Aminosäurereste, die in der Nähe der Termini liegen, auf die Bindung an nicht methylierte DNA zu erhalten, wurde in Abbildung 3.21 ein Bereich des ^1H - ^{15}N -HSQC-Spektrums des freien cMeCP2 ARBD und des DNA-gebundenen cMeCP2 ARBD, in dem hauptsächlich Signale von Aminosäureresten aus den Termini liegen, gegenüber gestellt. Beim freien cMeCP2 ARBD liegen viele Signale des in Abbildung 3.21 gezeigten Ausschnitts im Bereich typischer *random coil*-Werte. Liegt cMeCP2 ARBD gebunden an nicht methylierte DNA vor, so ändern viele dieser Signale ihre Position. Das bedeutet, daß

²⁰Im Gegensatz zu den Untersuchungen an der methylierten DNA (Abschnitt 3.2.1) ist ein solches Vorgehen sinnvoll, da die Spektrenqualität hier deutlich höher ist.

²¹Die Zuordnung der Hauptkettenatome des freien cMeCP2 ARBD ist in Abschnitt 3.1.1 dargestellt.

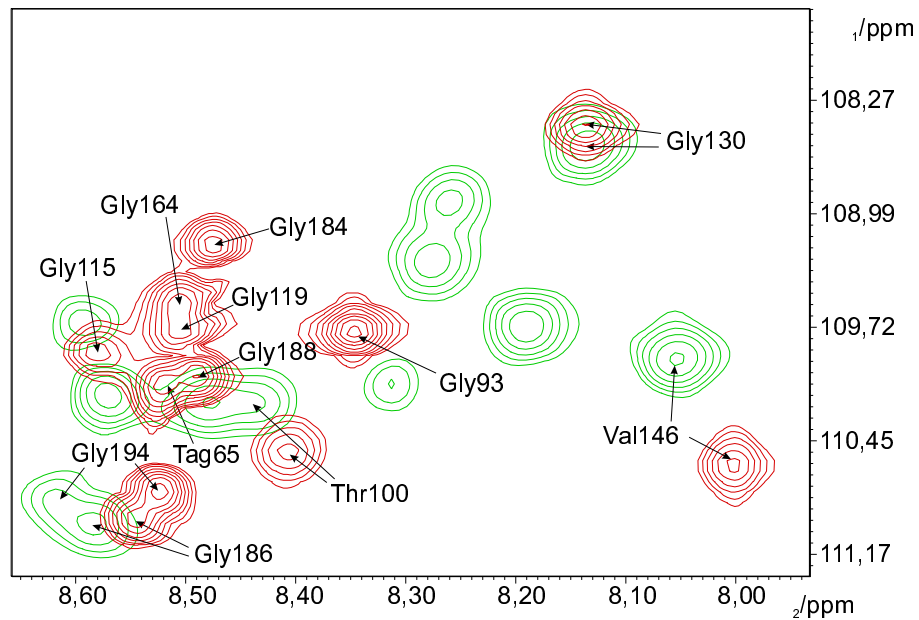


Abbildung 3.21: Ausschnitte aus den ^1H - ^{15}N -HSQC-Spektren von freiem cMeCP2 ARBD (rot) und an nicht methylierte DNA gebundenem cMeCP2 ARBD (grün).

sich die chemische Umgebung der ^1H - und ^{15}N -Kerne dieser Aminosäurereste geändert hat. Diese Änderung läßt auf eine Beteiligung dieser Aminosäurereste an der Bindung zwischen cMeCP2 ARBD und nicht methylierter DNA schließen.

Kapitel 4

Diskussion

In diesem Kapitel soll die für cMeCP2 ARBD berechnete Struktur diskutiert und mit den Strukturen von rMeCP2 MBD [Wak99] und MBD1 MBD [Ohk99] (Abschnitt 4.1) verglichen werden. Weiterhin wird die Bindung von cMeCP2 ARBD an methylierte und nicht methylierte DNA diskutiert und mit den Bindungsstudien, die in [Wak99] und [Ohk99] beschrieben sind, verglichen (Abschnitt 4.2).

4.1 Die Struktur von cMeCP2 ARBD

Ein wesentliches Problem bei der Strukturbestimmung von cMeCP2 ARBD war die Tatsache, daß große Teile des ungebundenen Proteins (Aminosäurereste 72-95 und 165-196) unstrukturiert vorliegen (siehe z.B. Abbildung 3.5 und 3.11). Diese Tatsache in Verbindung mit dem großen Anteil an Prolinresten erschwerte insbesondere die sequentielle Zuordnung der Hauptkettenatome. Wegen der starken Signalüberlagerungen war insbesondere auch die eindeutige Zuordnung der NOE-Kontakte kompliziert. Ungeachtet dieser Schwierigkeiten konnte trotzdem eine hinreichende Zahl von NOE-Kontakten zugeordnet werden, um insbesondere die Sekundärstrukturelemente gut zu definieren, was aus den niedrigen RMSD-Werten in Tabelle 3.5 ersichtlich ist. Die Verwendung zweier unterschiedlicher Strukturberechnungsprogramme (CNS und DYANA) lieferte in Bezug auf die Lage der Sekundärstrukturelemente und die Güte der Struktur bezogen auf den RMSD-Wert und die Diederwinkel (siehe Abbildung 3.14) sehr ähnliche Ergebnisse. Dies läßt den Schluß zu,

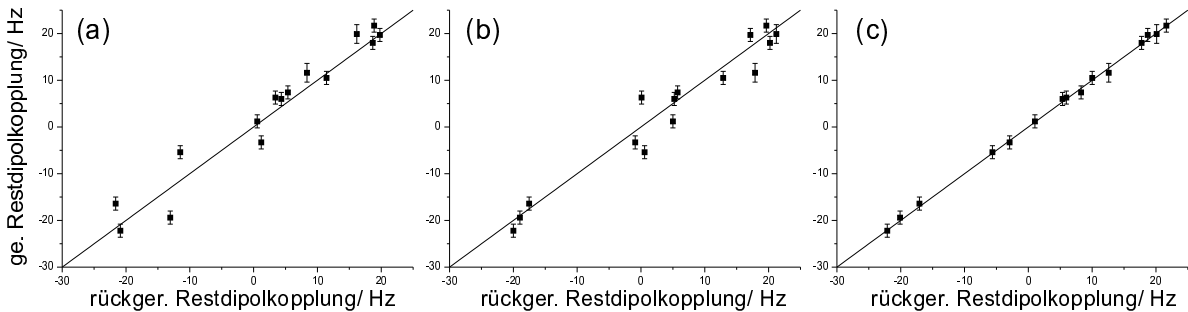


Abbildung 4.1: Gemessene $^1\text{H}_\text{N}$ - ^{15}N -Restdipolkopplungen von Aminosäureresten aus zentralen Bereichen der Sekundärstrukturelemente als Funktion der rückgerechneten $^1\text{H}_\text{N}$ - ^{15}N -Restdipolkopplungen. Die Strukturen wurden berechnet mit: (a) DYANA ohne Verwendung von Restdipolkopplungen, (b) CNS ohne Verwendung von Restdipolkopplungen und (c) CNS unter Verwendung von $^1\text{H}_\text{N}$ - ^{15}N -Restdipolkopplungen.

daß die berechnete Struktur der natürlichen Faltung des Proteins sehr nahe kommt und sie nicht eine Struktur darstellt, deren Konformation in einem lokalen Energieminimum liegt. Bei einem genaueren Vergleich zwischen DYANA und CNS erweist sich DYANA als das Programm, welches in kürzerer Zeit besser konvergierende Strukturen liefert. Im Hinblick auf die Verwendung von Restdipolkopplungen als *restraints* weist DYANA 1.5 allerdings einige Nachteile auf. So ist es beispielsweise nicht möglich, ^{15}N - $^{13}\text{C}'$ - und $^1\text{H}_\text{N}$ - $^{13}\text{C}'$ -Restdipolkopplungen als *restraints* zu verwenden.¹

Die Verwendung von Restdipolkopplungen als *restraints* hat sich als sehr hilfreich erwiesen. Besonders in der frühen Phase der Strukturbestimmung in Kombination mit der Spektrenrückrechnung waren Restdipolkopplungen sehr nützlich. So war es möglich, viele der vorher mehrdeutigen Signale eindeutig zuzuordnen. In der Endphase allerdings ergeben sich neben höheren Energien auch höhere RMSD-Werte bei der Verwendung von Restdipolkopplungen als *restraints* im Vergleich zu Strukturberechnungen ohne Restdipolkopplungen (siehe Tabelle 3.4 und 3.5). Alle Energiebeiträge (Tabelle 3.4) erhöhen sich bei Verwendung von Restdipolkopplungen, wobei die Energieverletzungen die direkt

¹Die Verwendung von $^1\text{H}_\text{N}$ - ^{15}N -Restdipolkopplung ist mit DYANA möglich. Allerdings zeigte sich, daß die mit Restdipolkopplungen berechneten Strukturen deutlich an Qualität einbüßten. Auf Grund dessen wurde die Verwendung von Restdipolkopplungen in Strukturberechnungen mit DYANA nicht weiter verfolgt.

durch Restdipolkopplungen (E_{coupl}) hervorgerufen werden vergleichsweise gering sind. Der Grund für dieses Verhalten ist wahrscheinlich, daß die Restdipolkopplungen im Vergleich zu den anderen *restraints*² deutlich stärker gewichtet werden und daß die Einschränkungen des Konformationsraums des Moleküls, die von Restdipolkopplungen ausgehen - im Gegensatz zu NOEs - auf Grund der geringeren Anzahl selten überbestimmt sind. Dies hat zur Folge, daß die Wirkung eines fehlerhaften *restraints* nicht durch andere Restdipolkopplungen korrigiert wird. Im Gegensatz dazu wird die Wirkung fehlerhafter NOE-*restraints* von anderen NOE-*restraints* zumindest teilweise korrigiert (unter der Voraussetzung, daß die Anzahl der fehlerhaften NOEs klein ist im Vergleich zur Anzahl aller NOEs und unter der Voraussetzung, daß Distanzinformationen überbestimmt vorliegen). Zur besseren Verdeutlichung sei auf Abbildung 4.1 verwiesen, in der gemessene mit rückgerechneten $^1\text{H}_N$ - ^{15}N -Restdipolkopplungen aus den zentralen Bereichen der Sekundärstrukturelemente verglichen werden.³ Diese Korrelation ist für drei Strukturen angegeben, die mit DYANA ohne Verwendung von Restdipolkopplungen als *restraints* (a), mit CNS ohne Verwendung von Restdipolkopplungen als *restraints* (b) und mit CNS unter Verwendung von $^1\text{H}_N$ - ^{15}N -Restdipolkopplungen als *restraints* (c) berechnet wurden. Bei den Strukturen, welche ohne Restdipolkopplungen bestimmt wurden, treten erwartungsgemäß Abweichungen zwischen berechneten und gemessenen Restdipolkopplungen auf. Im Gegensatz dazu stehen die gemessenen und die berechneten Restdipolkopplungen für die unter Verwendung von Restdipolkopplungen bestimmten Strukturen in sehr guter Übereinstimmung. Dies ist neben den niedrigen Energiebeiträgen (E_{coupl}) aus Tabelle 3.4 ein weiteres Indiz dafür, daß den Restdipolkopplungen ein vergleichsweise großes Gewicht in der Berechnung der Strukturen mittels CNS beigemessen wird.

In Abbildung 4.2 sind die Strukturen von Domänen der Proteine rMeCP2, cMeCP2 und MBD1 aus der Familie der methylbindenden Proteine dargestellt (vgl. Abschnitt 1.3). Es handelt sich dabei um die derzeit einzigen bekannten Strukturen von Proteinen dieser Familie. Erwartungsgemäß weisen alle drei das gleiche Faltungsmotiv auf, da die Sequenzen sich für den abgebildeten Kernbereich der Proteindomänen kaum unterscheiden

²empirischer wie experimenteller Natur

³Die Werte für die rückgerechneten Restdipolkopplungen wurden mit dem Computerprogramm MODULE [Dos01] berechnet.

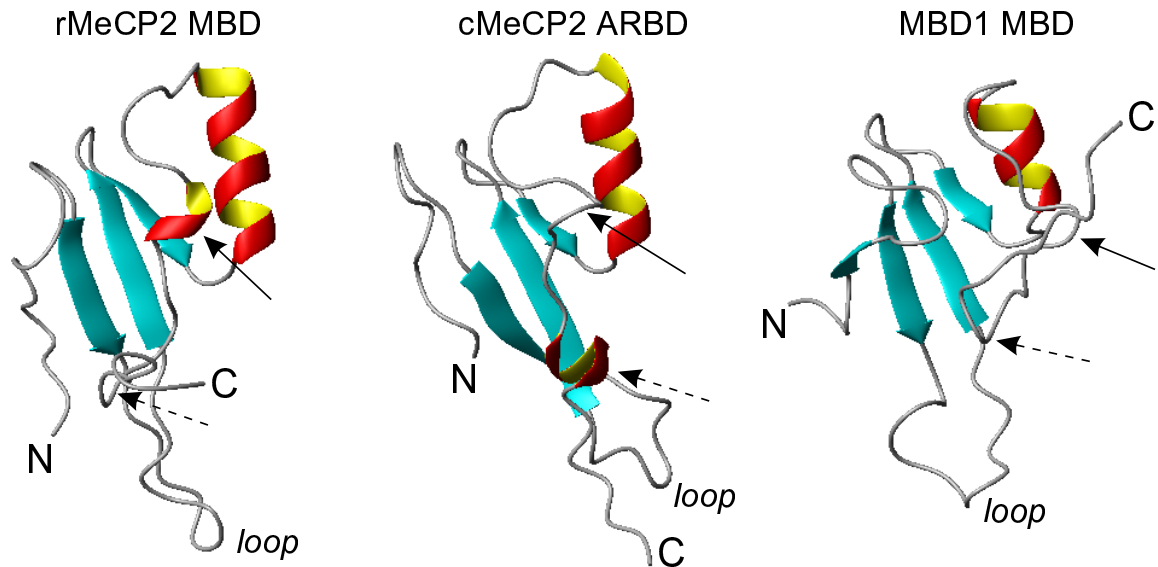


Abbildung 4.2: Die NMR-Strukturen von *cMeCP2 ARBD* (mitte), *rMeCP2 MBD* (links) [Wak99] und *MBD1 MBD* (rechts) [Ohk99]. Gestrichelte Pfeile markieren die Region von Aminosäurerest 160; durchgezogene Pfeile markieren die Region von Aminosäurerest 153 (nach der Zählung von *cMeCP2 ARBD*).

(siehe Abbildung 1.4).

Unterschiede ergeben sich für den Sequenzabschnitt, der sich an die α -Helix (Aminosäurereste 135-145)⁴ anschließt. Ohki et al. [Ohk99] sprechen von einem wohldefinierten Bereich, der mit einem *hairpin loop* (Aminosäurereste 156-162) endet. Die Orientierung dieses *loops* zum restlichen, wohlgefalteten Teil des Proteins wird als gut definiert beschrieben. Wakefield et al. [Wak99] berichten für den Abschnitt nach der α -Helix (Aminosäurereste 135-145) über eine Helix von nur einer Windung Länge (Aminosäurereste 151-153). Für den sich daran anschließenden Abschnitt wird kein definiertes Strukturelement erkannt. Die in der vorliegenden Arbeit bestimmte Struktur von *cMeCP2 ARBD* ergibt dagegen für den Bereich nach der α -Helix (Aminosäurereste 135-145) zwei helikale

⁴Nach der Zählung von *cMeCP2 ARBD* (siehe auch Abbildung 1.4 und 1.5). Auch für die weiter unten im Text folgenden Nummern von Aminosäureresten soll die Zählung von *cMeCP2 ARBD* gelten.

Strukturelemente (Aminosäurereste 153-155 und 159-163)⁵, die Ähnlichkeit mit den bekannten *helix-turn-helix*-Motiven besitzen. Diese Motive treten häufig bei DNA-bindenden Proteinen auf [Cre93]. Die Tatsache, daß in [Wak99] die fünf Aminosäurereste umfassende Helix (Aminosäurereste 159-163) nicht erkannt wird, ist verwunderlich, da es keine sequentiellen Unterschiede zwischen cMeCP2 ARBD und rMeCP2 MBD in diesem Bereich gibt. Eine wahrscheinliche Erklärung für diesen Umstand ist, daß die in [Wak99] untersuchte Domäne unmittelbar nach diesem Abschnitt endet und daß die letzten vier Aminosäurereste der dort untersuchten Domäne (166-169) GSGC sind, während die Sequenz im nativen Protein, ebenso wie bei cMeCP2 ARBD, mit den Aminosäureresten PSRR für die Positionen 166-169 fortgesetzt wird.⁶

4.2 Der Komplex aus Protein und DNA

Es wurde die Bindung von cMeCP2 ARBD an methylierte und nicht methylierte DNA untersucht. Die Untersuchungen zur Bindung an methylierte DNA haben sich als relativ schwierig herausgestellt. Erst nach Optimierung der Pufferbedingungen war es möglich, Spektren aufzunehmen, welche das Studium des Bindungsverhaltens erlaubten. Allerdings blieb die Spektrenqualität hinsichtlich des Signal/Rausch-Verhältnisses trotz dieser Verbesserung deutlich hinter der des freien cMeCP2 ARBD und auch der des Protein-DNA-Komplexes mit nicht methylierter DNA zurück. Der Vergleich mit den Bindungsstudien von Wakefield et al. [Wak99] und Ohki et al. [Ohk99] mit methylierter DNA zeigt, daß

⁵Diese beiden Strukturelemente werden, abhängig davon welches Strukturberechnungsprogramm verwendet wird, nicht immer vom Kabsch-Sander-Algorithmus erkannt, allerdings sprechen die Diederwinkel der Aminosäurereste dieses Abschnitts dafür, daß eine helikale Struktur vorliegt, unabhängig davon, welches Strukturberechnungsprogramm verwendet wurde.

⁶Ein weiterer vermeintlicher Unterschied zwischen der Struktur von rMeCP2 MBD und cMeCP2 ARBD besteht in einem vierten kurzen Faltblattstrang (Aminosäurereste 97-99), von dem in [Wak99] berichtet wird. Dieser Faltblattstrang ist in [Wak99] mittels CSI vorhergesagt worden und ist in einem Bändermodell von rMeCP2 MBD in [Wak99] abgebildet, zu erkennen. Betrachtet man allerdings die Struktur von rMeCP2 MBD im Bändermodell, erzeugt mit dem Programm MOLMOL [Kon96] (welches in der vorliegenden Arbeit verwendet wird), so wird der Sequenzabschnitt 97-99 nicht als Teil des β -Faltblattes erkannt (siehe Abbildung 4.2). Dies gab Anlaß, diesem Unterschied nicht weiter nachzugehen.

sich cMeCP2 ARBD ähnlich verhält wie rMeCP2 MBD und MBD1 MBD. Ein detaillierter Vergleich läßt sich allerdings nicht anstellen, da dies die Spektrenqualität nicht zuläßt. Dies läßt sich nicht allein durch den Unterschied in der Molekülgröße des zu erwartenden 1:1 Komplexes aus Protein und DNA erklären (cMeCP2 ARBD: 133 Aminosäurereste mit *tag*, DNA-Fragment: 15 Basenpaare, rMeCP2 MBD: 102 Aminosäurereste mit *tag*, DNA-Fragment: 14 Basenpaare, MBD1 MBD: 77 Aminosäurereste mit *tag*, DNA-Fragment: 12 Basenpaare). Auffällig ist, daß die Spektren vom cMeCP2 ARBD-Komplex mit nicht methylierter DNA geringere Linienbreiten aufweisen als die Spektren vom Komplex mit methylierter DNA. Dies könnte den Schluß nahelegen, daß ein Komplex vorliegt, der mehr als ein DNA-Fragment und ein cMeCP2 ARBD-Molekül beinhaltet. Gegen diese Vermutung sprechen allerdings die Diffusionsmessungen (siehe Abbildung 3.17), wobei dort für die Abschätzung der Größe des Komplexes angenommen wurde, daß der Komplex die gleiche Form besitzt wie ungebundenes cMeCP2 ARBD. Dies vermindert die Genauigkeit der Größenabschätzung.

Eine weitere Erklärung für das geringe Signal/Rausch-Verhältnis, welche z.T. mit der obigen Erklärung im Zusammenhang steht, könnte in der Wahl des untersuchten Sequenzabschnittes von cMeCP2 begründet liegen. Die untersuchte Domäne (ARBD) von cMeCP2 bindet nicht methylierte DNA und stellt gleichzeitig auch den minimalen Sequenzabschnitt von cMeCP2 dar, der die Fähigkeit zur Bindung an nicht methylierte DNA aufweist (siehe Abschnitt 3.2.2 und [Wei97]). Zur Bindung an methylierte DNA ist ein kleinerer Abschnitt von MeCP2, die methylbindende Domäne (MBD), ausreichend [Nan93], [Wak99] (einen Überblick über die beiden Domänen gibt z.B. Abbildung 3.5). So wäre es denkbar, daß ein Proteinmolekül mit mehr als einem DNA-Fragment wechselwirkt, da bereits der Kernbereich der Domäne zur Bindung an methylierte DNA ausreicht. Es wäre somit möglich, daß die Aminosäurereste in der Nähe der Termini mit einem anderen DNA-Fragment wechselwirken⁷ als der zentrale Bereich. Dies hätte eine Vergrößerung des Komplexes zur Folge.

Die Bindungsstudien mit nicht methylierter DNA liefern, wie oben erwähnt, Spektren mit besserem Signal/Rausch-Verhältnis. Die Änderungen der chemischen Verschiebungen, die in Abbildung 3.20 graphisch dargestellt sind, wurden in Abbildung 4.3 auf den

⁷welches unter Umständen wiederum an ein anderes Proteinmolekül gebunden ist

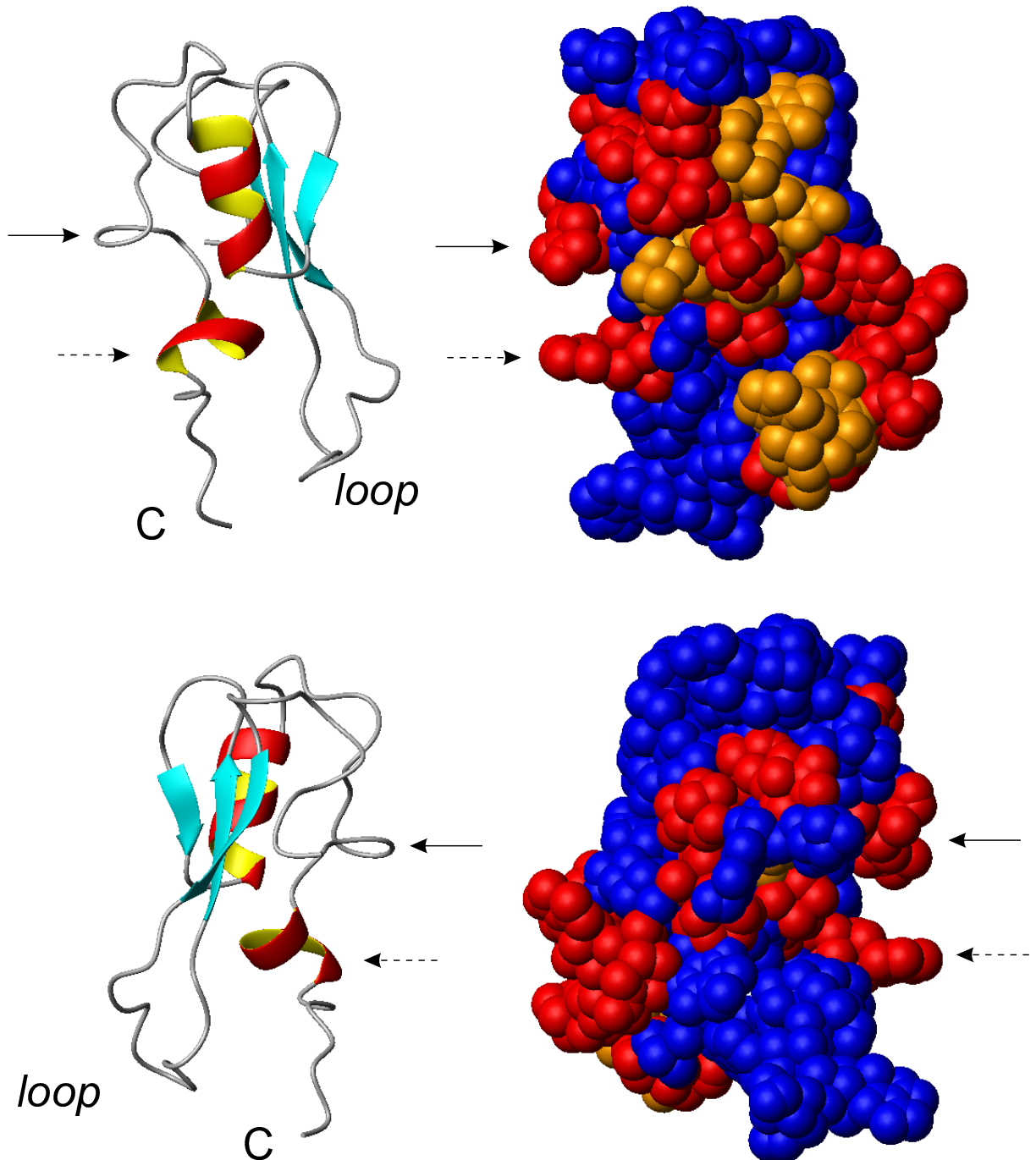


Abbildung 4.3: Änderung der chemischen Verschiebung bei Bindung an nicht methylierte DNA. Der Kernbereich von cMeCP2 ARBD ist in zwei Orientierungen dargestellt. Zur besseren Anschauung ist das Protein in jeweils der selben Orientierung als Bändermodell auf der linken Seite dargestellt. Blau: geringe Änderung der chem. Verschiebung ($< 100\text{ Hz}$) - Rot: Änderungen der chem. Verschiebung $\geq 100\text{ Hz}$ - Orange: Aminosäurereste, dessen Änderungen der chem. Verschiebungen $< 100\text{ Hz}$ betragen, welche aber Aminosäureresten benachbart sind, deren Änderung $\geq 100\text{ Hz}$ sind. Gestrichelte Pfeile markieren die Region von Aminosäurerest 160; durchgezogene Pfeile markieren die Region von Aminosäurerest 153.

Kernbereich der Struktur von cMeCP2 ARBD übertragen. Offensichtlich gibt es einen zusammenhängenden Bereich, in dem die Aminosäurereste deutliche Änderungen der chemischen Verschiebungen erfahren (≥ 100 Hz). Charakteristisch sind die Änderungen der chemischen Verschiebungen speziell im Bereich des *helix-turn-helix*-Motivs (mit Pfeilen gekennzeichnet). Es ist davon auszugehen, daß die Abschnitte des Moleküls, welche signifikante Änderungen der chemischen Verschiebungen erfahren, an der Wechselwirkung mit der DNA beteiligt sind. Vergleicht man den Einfluß der Titration von methylierter mit dem von nicht methylierter DNA auf die Änderung der chemischen Verschiebungen,⁸ so ergibt sich eine hoher Grad an Übereinstimmung. Das läßt darauf schließen, daß die Art der Bindung des Moleküls an methylierte DNA und nicht methylierte DNA sehr ähnlich ist. Unterschiede könnten in den Bereichen liegen, für die eine Zuordnung der chemischen Verschiebungen der Aminosäurereste des Komplexes nicht möglich war (vornehmlich in der Nähe der Termini). Die Beteiligung dieser Aminosäurereste an der Bindung läßt sich aus Abbildung 3.21 abschätzen.

Vergleicht man die Ergebnisse der Bindungsstudien von nicht methylierter DNA (Abschnitt 3.2.2, speziell Abbildung 3.20) mit den Ergebnissen der Bindungsstudien von methylierter DNA [Wak99], [Ohk99], so ergibt sich ein ähnliches Bild. Es fällt allerdings auf, daß bei [Wak99] und [Ohk99] die größten Änderungen der chemischen Verschiebungen beim cMeCP2 ARBD-DNA-Komplex im Bereich des *loops* (Aminosäurereste 111-121) liegen. Im Gegensatz dazu liegen bei cMeCP2 ARBD die größten Änderungen im Bereich des *helix-turn-helix*-Motivs (Aminosäurereste 153-163). Eine Erklärung hierfür könnte darin zu finden sein, daß die in [Wak99] und [Ohk99] untersuchten Domänen im direkten Anschluß an den Sequenzbereich, der bei cMeCP2 ARBD das *helix-turn-helix*-Motiv ausmacht, enden.⁹ In jedem Falle kann festgestellt werden, daß das *helix-turn-helix*-Motiv von Bedeutung für die Bindung des Proteins an DNA ist. In [Nan93] wird berichtet, daß MeCP2 MBD seine Bindungsaffinität komplett verliert, wenn die Domäne hinter Aminosäurerest 157 abgeschnitten wird. Die Bedeutung dieses Sequenzabschnittes für das

⁸soweit dies die unterschiedlichen Signal/Rausch-Verhältnisse zuließen

⁹Wobei die Möglichkeit besteht, daß die Änderung der chemischen Verschiebungen im Bereich des *loops* zwischen den ersten beiden Strängen des β -Faltblattes unterschätzt wurde, da eine Zuordnung für einige der Aminosäurereste in diesem Bereich nicht möglich war. Für diesen Fall steht die Änderung der chemischen Verschiebung in Abbildung 3.20 für eine untere Grenze.

Rett-Syndrom läßt sich aus [Ami99] und [Wan99] ablesen. Dort wird von mehreren Rett-Patienten berichtet, bei denen Mutationen genau im Bereich des *helix-turn-helix*-Motivs auftraten (siehe Abbildung 1.1).

In [Wak99] wie auch in [Ohk99] wurden neben Untersuchungen mit methylierter DNA auch Untersuchungen mit nicht methylierter DNA durchgeführt. Wakefield et al. berichten von Aggregationseffekten bei Zugabe von nicht methylierter DNA, die eine Spektrenaufnahme unmöglich machten. Ohki et al. berichten von einer nur schwachen Bindung an nicht methylierte DNA, die sich in geringen Änderungen der chemischen Verschiebung bei Titration der DNA widerspiegelt. Dies steht im Einklang mit den Ergebnissen aus [Wei97]; dort wurde cMeCP2 ARBD als minimale Domäne aus cMeCP2 für die Bindung an nicht methylierte DNA identifiziert. Diese Bindung konnte in der vorliegenden Arbeit bestätigt werden.

Damit ergeben sich unter Umständen Konsequenzen für die Interpretation der Wirkung von MeCP2 beim Prozeß der Transkriptionshemmung, da bisher davon ausgegangen wurde, daß MeCP2 nicht bzw. nur schwach an nicht methylierte DNA bindet. Da, wie in der vorliegenden Arbeit gezeigt wurde, cMeCP2 ARBD (als Teil von MeCP2) an nicht methylierte DNA bindet, sollte somit auch dort der Prozeß eingeleitet werden, der zur Hemmung der Transkription führt (siehe Abbildung 1.2).¹⁰ Da dies nicht der Fall ist (siehe z.B.[Rob00]), läßt sich die Hemmung der Transkription nicht allein durch die Bindung von MeCP2 an methylierte DNA erklären. Für eine Überprüfung dieser These wären Bindungsstudien von cMeCP2 ARBD mit verschiedenen DNA-Fragmenten, die alle ein nicht methyliertes CpG-Motiv enthalten, sinnvoll.

¹⁰Unter der Annahme, daß das gewählte DNA-Fragment außer durch das nicht methylierte CpG-Motiv nicht ausgezeichnet ist.

Zusammenfassung

Die in dieser Arbeit untersuchte Proteindomäne cMeCP2 ARBD des Proteins ARBP (*attachment region binding protein*, auch cMeCP2 genannt) weist mehr als 95% Sequenzidentität zu einem Abschnitt des Proteins MeCP2 (*methyl-CpG binding protein 2*) auf. Diese Proteine stehen u.a. im Zusammenhang mit der Genregulation, der strukturellen Organisation der DNA und der Methylierung von DNA. Neuere Arbeiten zeigen, daß Mutationen, die zu Veränderungen von MeCP2 führen, in Beziehung zum Rett-Syndrom (tödliche neurologische X-chromosomale Erkrankung) stehen.

In dieser Arbeit wurde mittels Flüssigkeits-NMR-Spektroskopie die Struktur der MAR-bindenden Domäne von cMeCP2 (MAR- *matrix attachment region*) bestimmt. Sie besteht aus einem dreisträngigen antiparallelen β -Faltblatt, gefolgt von einer α -Helix. Im Anschluß an diesen Bereich folgt ein *helix-turn-helix*-Motiv, welches im Rahmen der vorliegenden Arbeit erstmals nachgewiesen werden konnte. *Helix-turn-helix*-Motive treten häufig bei DNA-bindenden Proteinen auf.

Große Bereiche in der Nähe der Termini von cMeCP2 ARBD sind nicht strukturiert. Der Vergleich mit der Struktur der sequenzidentischen Domäne rMeCP2 MBD ergibt eine gute Übereinstimmung mit geringen Unterschieden im C-terminalen Bereich von rMeCP2 MBD.

Weiterhin konnte NMR-spektroskopisch die Bindung der Proteindomäne an methylierte sowie nicht methylierte DNA nachgewiesen werden. Die Ergebnisse der Bindungsstudien mit methylierter DNA stehen im Einklang mit bereits veröffentlichten Bindungsstudien anderer Vertreter der Familie der methylbindenden Proteine mit methylierter DNA. Zusätzlich wurde die Bindung von cMeCP2 ARBD an nicht methylierte DNA untersucht. Die beobachteten Änderungen der chemischen Verschiebungen bei Titration von nicht methylierter DNA legen nahe, daß bei Bindung der DNA weniger unstrukturierte Bereiche vorliegen als bei freiem cMeCP2 ARBD. Diese Ergebnisse, zusammen mit den Bindungsstudien an rMeCP2 MBD mit nicht methylierter DNA und Arbeiten unserer Kooperationspartner von der Universitätsklinik Eppendorf, Hamburg (Prof. Strätling und Mitarbeiter), bestätigen cMeCP2 ARBD als die minimal nötige Domäne des Proteins cMeCP2 für die Bindung an nicht methylierte DNA.

Literaturverzeichnis

- [Ami99] Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., Zoghbi, H. Y. (1999) *Nat. Genet.* **23**, 185-188
- [Bam01] Bamann, Christian (2001) Diplomarbeit, Universität Regensburg
- [Bas87] Bastiaan, E. W., MacLean, C., van Zijl, P. C. M., Bothner-By, A. A. (1987) *Annu. Rep. NMR Spectrosc.* **19**, 35-77
- [Bes88] Bestor, T., Laudano, A., Mattaliano, R., Ingram, V. (1988) *J. Mol. Biol.* **203**, 971-983
- [Bes00] Bestor, T., (2000) *Hum. Mol. Gen.* **9**, 2395-2402
- [Bra85] Braun, W., Go, N., (1985) *J. Mol. Biol.* **186**, 611-626
- [Bru92] Brünger, A. T. (1992) *X-PLOR Manual Version 3.0*, Yale University Press, New Haven
- [Bru98] Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., Warren, G. L. (1998) *Acta Cryst.* **D54**, 905-921
- [Bru00a] Brunner, E. Ogle, J., Wenzler M., Kalbitzer, H.R. (2000) *Biochem. Biophys. Res. Com.*, **272**, 694-698
- [Bru00b] Brunner, E., Weitzel, J., Heitmann, B., Maurer, T., Strätling, W. H., Kalbitzer, H.R. (2000) *J. Biomol. NMR* **17**, 175-176
- [Bru01] Brunner, E. (2001) *Concepts in Magnetic Resonance*, **13**, 238-259

- [Buc62] Buckingham, A. D., Lovering, E. G. (1962) **58**, 2077-2081
- [Cav95] Cavanagh, J., Fairbrother, W. J., Palmer III, A. G., Skelton, N. J. 1995, Academic Press
- [Cav99] Cavangero, S., Dyson, H. J., Wright, P. E., (1999) *J. Biom. NMR*, **13**, 387-391
- [Cre93] Creighton, T. E. (1993) *Proteins: structures and molecular properties* 2nd ed. by W.H. Freeman and Company
- [Clo98] Clore, G. M., Gronenborn, A. M., Bax, A. (1998) *J. Magn. Reson.* **133**, 216-221
- [Coo89] Cooper, D. N., Krawczak, M. (1998) *Hum. Genet.* **83**, 181-188
- [Cor99] Cornilescu, G., Delaglio, F., Bax, A. (1999) *J. Biomol. NMR* **13**, 289-302
- [Dos01] Dosset, P., Hus, J.-C., Marion, D., Blackledge, M. (2001) *M. J. Biomol. NMR* **20**, 223-231
- [Fus98] Fushman, D., Tjandra, N., Cowburn, D. (1998) *J. Am. Chem. Soc.* **120**, 10947-10952
- [Fus99] Fushman, D., Cowburn, D. (1999) *J. Biomol. NMR.* **13**, 139-147
- [Gan00] Ganslmeier, B., Gronwald, W., Neidig, K. P., Kalbitzer, H. R. (2000) *AUREMOL, a new program for the assignment of multidimensional NMR spectra*, Posterbeitrag, 22. Diskussionstagung der GDCh-Fachgruppe Magnetische Resonanzspektroskopie, Regensburg
- [Gay82] Gayathri, C., Bothner-By, A. A., van Zijl P. C. M., MacLean, C., (1982) *Chem. Phys. Letters* **87**, 192-196
- [Ger93] Gerald II, R., Bernhard, T., Haeberlen, U., Rendell, J., Opella, S. (1993) *J. Am. Chem. Soc.* **115**, 777-782
- [Gey97] Geyer, M., Herrmann, C., Wohlgemuth, S., Wittinghofer, A., Kalbitzer, H. R. (1997) *Nat. Struct. Biol.* **4**, 694-699
- [Gol84] Goldman, M. (1984) *J. Magn. Reson.* **60**, 437-452

- [Gol84] Goldman, M. (1984) *J. Magn. Reson.* **60**, 437-452
- [Gol63] Goldstein, H. (1963) *Klassische Mechanik*, Akademische Verlagsgesellschaft, Wiesbaden
- [Gor97] Görler, A., Kalbitzer, H. R. (1997) *J. Magn. Reson.* **124**, 177-188
- [Gor99] Görler, A., Gronwald, W., Neidig, K.-P., Kalbitzer, H. R. (1999) *J. Magn. Reson.* **137**, 39-45
- [Gun97] Güntert, P., Mumenthaler, C., Wüthrich, K., (1997) *J. Mol. Biol.* **273**, 283-298
- [Hag83] Hagberg, B., Aicardi, J., Dias, K., Ramos, O. *Ann. Neurol.* **14**, 471-479
- [Hen00] Hendrich, B., Bird, A. (2000) *Microbiol. Immunol.* **249**, 55-74
- [Hon99] Hong, M. J. (1999) *J. Biomol. NMR* **106**, 1-14
- [Jai93] Jain, A., Vaidehi, N., Rodriguez, G., (1993) *J. Comput. Phys.* **106**, 258-268
- [Jon98] Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., Wolffe, A. P. (1998) *Nat. Genet.* **19**, 187-191
- [Kab83] Kabsch, W., Sander, C. (1983) *Biopolymers* **22**, 2577-2637
- [Kan00] Kany, Harry (2000) Dissertation, Universität Kaiserslautern
- [Kal00] Kaludov, N. K., Wolffe, A. P. (2000) *Nucl. Acids Res.* **28**, 1921-1928
- [Kir83] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P. (1983) *Science* **220**, 671-680
- [Kon96] Konradi, R., Billeter, M., Wüthrich, K. (1996) *J. Mol. Graphics* **14**, 51-55
- [Kri91] von Kries, J. P., Buhrmester, H., Strätling, W. H. (1991) *Cell* **64**, 125-135
- [Leh94] Lehninger, A. L., Nelson, D. L., Cox, M. M., (1994) *Prinzipien der Biochemie 2. Auflage* dt. Übers. Spektrum, Akad. Verl.
- [Les98] Lesage, A., Sakellariou, D., Sternagel, S., Emsley, L. (1992) *J. Am. Chem. Soc.* **120** (50), 13194-13201

- [Lew92] Lewis, J. D., Meechan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppensen, P., Klein, F., Bird, A. (1992) *Cell* **69**, 905-914
- [Li92] Li, E., Bestor, T., Jaenisch, R., (1992) *Cell* **69**, 915-926
- [Lod99] Lohdsh, H. [et al.] (1999) *Molecular cell biology* 4. Auflage W. H. Freeman and Company
- [Los98] Losonczy, J. A., Prestegard, J. H. (1998) *J. Biomol. NMR* **12**, 447-451
- [Man97] Mandai, K., Nakanishi, H., Satoh, A., Obaishi, H., Wada, M., Nishioka, H., Itoh, M., Mizoguchi, A., Aoki, T., Fujimoto, T., Matsuda, Y., Tsukita, S., Takai, Y. (1997) *J. Cell. Biol.* **139**, 517-528
- [Mar98] Markley, J. L., Bax, A., Arata, A., Hilbers, Y., Kaptein, C. W., Sykes, B. D., Wright, P. E., Wüthrich, K. (1998) *Pure & Appl. Chem.* **70**, 117-142
- [Nan93] Nan, X., Meehan, R. R., Bird, A. (1993) *Nucl. Acids Res.* **21**, 4886-4892
- [Nan97] Nan, X., Campoy, F. J., Bird, A. (1997) *Cell* **88**, 471-481
- [Nei95] Neidig, K.-P., Geyer, M., Görler, A., Antz, C., Saffrich, R., Beneicke, W., Kalbitzer, H. R. (1995) *J. Biomol. NMR* **6**, 255-270
- [Ohk99] Ohki, I., Shimotake, N., Fujita, N., Nakao, M., Shirakawa, M. (1999) *EMBO J.* **18**, 6653-6661
- [Oka99] Okano, M., Bell, D. W., Haber, D. A., Li, W. (1999) *Cell* **99**, 247-257
- [Ott99] Ottiger, M., Bax, A. (1994) *J. Biomol. NMR* **13**, 187-191
- [Pen92] Peng, J. W., Wagner, G. (1992) *J. Magn. Reson.* **98**, 308-332
- [Pet01] Wolfgang Peti, persönliche Mitteilung (2001)
- [Per97] Pervushin, K., Riek, R., Wider, G., Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 12366-12371
- [Rai79] Raiford, D. S., Fisk, C. G., Becker, E. D. (1979) *Anal. Chem.* **51**, 2050-2051

- [Ram63] Ramachandran, G. N., Sasisekharan, V. (1963) *J. Mol. Biol.* **7**, 95-99
- [Rap99] Rapp, A., Ermolaev, K., Fung, B. M. (1999) *J. Phys. Chem. B* **103**, 1705-1711
- [Raz98] Razin, A. (1998) *EMBO J.* **17**, 4905-4908
- [Rei00] Reif, B., Hohwy M., Jaroniec, C. P., Rienstra, C. M., Griffin, R. G. (2000) *J. Magn. Reson.* **146**, 132-141
- [Ret66] Rett, A., (1966) *Wien Med. Wochensch.* **116**, 723-726
- [Rob00] Robertson, K. D., Wolffe, A. P., (2000) *Nat. Rev. Genet.* **1**, 11-19
- [Sal72] Salsbury, N. J., Darke, A., Chapman, D. (1972) *Chem. Phys. Lipids* **8**, 142-151
- [San92] Sanders, C. R., Schwonek, J. P. (1992) *Biochemistry* **31**, 8898-8905
- [Sat99] Sattler, M., Schleucher, J., Griesinger, C. (1999) *Prog. Nucl. Magn. Res. Spec.* **34**, 93-158
- [Sol55] Solomon, I. (1955) *Phys. Rev.* **99**, 559-565
- [Ste01] Steiner, G. (2001) Doktorarbeit, Universität Regensburg
- [Str99] Wolf H. Strätling (1999) persönliche Mitteilung
- [Tja97] Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M., Bax, A. (1997) *Nature Struct. Biol.* **4**, 732-738
- [Tja97b] Tjandra, N., Bax, A. (1997) *J. Am. Chem. Soc.* **119**, 8076-8082
- [Val91] Valencia, A., Kjeldgaard, M., Pai, E. F., Sander, C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5443-5447
- [Wak99] Wakefield, R. I. D., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. P., Barlow, P. N. (1999) *J. Mol. Biol.* **291**, 1055-1065
- [Wan99] Wan, M., Lee, S. S., Zhang, X., Houwink-Manville, I., Song, H. R., Amir, R. E., Budden, S., Naidu, S., Pereira, J. L., Zoghbi, H. Y., Schanen, N. C., Francke, U. (1999) *Am. J. Hum. Genet.* **65**, 1520-1529

-
- [Wei97] Weitzel, J. M., Buhrmester, H., Strätling, W. H. (1997) *Mol. Cell. Biol.* **17**, 5656-5666
- [Wis91] Wishart, D. S., Sykes, B. D., Richards, F. M. (1991) *FEBS Letters* **293**, 72-80
- [Wis94] Wishart, D. S., Sykes, B. D. (1994) *J. Biomol. NMR* **4**, 171-180
- [Wis95a] Wishart, D. S., Bigam, C. G., Yao, J., Abdilgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., Sykes, B. D. (1995) *J. Biomol. NMR* **6**, 135-140
- [Wis95b] Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., Sykes, B. D. (1995) *J. Biomol. NMR* **5**, 67-81
- [Wut86] Wüthrich, K. (1986) *NMR of Proteins and Nucleid Acids*, John Wiley and Sons, New York
- [Zwe00] Zweckstetter, M., Bax, A. (2000) *J. Am. Chem. Soc.* **122**, 3791-3792

Anhang A

Die chemischen Verschiebungen von cMeCP2 ARBD

ALA 72	N	126.5	PRO 76	CA	62.82	SER 79	CA	58.023
ALA 72	HN	8.409	PRO 76	HA	4.427	SER 79	HA	4.336
ALA 72	CA	50.33	PRO 76	CB	31.32	SER 79	CB	63.66
ALA 72	HA	4.627	PRO 76	HB1	2.318	SER 79	HB1	3.849
ALA 72	CB	16.80	PRO 76	HB2	1.905	SER 79	HB2	3.849
ALA 72	HB1	1.384	PRO 76	CG	27.13	SER 79	C	174.2
ALA 72	HB2	1.384	PRO 76	CD	50.73	ALA 80	N	124.89
ALA 72	HB3	1.384	PRO 76	C	177.01	ALA 80	HN	8.262
PRO 73	CA	62.54	GLU 77	N	121.4	ALA 80	CA	52.03
PRO 73	HA	4.402	GLU 77	HN	8.571	ALA 80	HA	4.399
PRO 73	CB	31.5	GLU 77	CA	56.46	ALA 80	CB	18.82
PRO 73	HB1	2.279	GLU 77	HA	4.432	ALA 80	HB1	1.401
PRO 73	HB2	1.888	GLU 77	CB	29.5	ALA 80	HB2	1.401
PRO 73	CG	27.12	GLU 77	HB1	1.810	ALA 80	HB3	1.401
PRO 73	CD	50.12	GLU 77	HB2	1.810	ALA 80	C	177.4
PRO 73	C	176.4	GLU 77	CG	36.24	SER 81	N	115.79
ALA 74	N	123.95	GLU 77	C	176.6	SER 81	HN	8.292
ALA 74	HN	8.426	ALA 78	N	124.36	SER 81	CA	56.14
ALA 74	CA	52.08	ALA 78	HN	8.441	SER 81	HA	4.759
ALA 74	HA	4.295	ALA 78	CA	52.47	SER 81	CB	62.89
ALA 74	CB	18.89	ALA 78	HA	4.109	SER 81	HB1	3.859
ALA 74	HB1	1.348	ALA 78	CB	18.6	SER 81	HB2	3.859
ALA 74	HB2	2.124	ALA 78	HB1	1.404	PRO 82	CA	63.15
ALA 74	C	176.2	ALA 78	HB2	1.404	PRO 82	CB	31.38
VAL 75	N	117.5	ALA 78	HB3	1.404	PRO 82	CG	27.07
VAL 75	HN	7.860	ALA 78	C	177.8	PRO 82	C	177.2
VAL 75	CA	58.5	SER 79	N	113.9	LYS 83	N	120
			SER 79	HN	8.273	LYS 83	HN	8.364

LYS	83	CA	56.31	ARG	86	HD1	3.185	ASP	91	N	121.38
LYS	83	HA	4.463	ARG	86	C	176.29	ASP	91	HN	8.355
LYS	83	CB	32.46	SER	87	N	116.60	ASP	91	CA	54.09
LYS	83	HB1	1.79	SER	87	HN	8.395	ASP	91	HA	4.34
LYS	83	HB2	1.79	SER	87	CA	57.97	ASP	91	CB	40.6
LYS	83	CG	25.14	SER	87	HA	4.364	ASP	91	HB1	2.638
LYS	83	HG1	1.44	SER	87	CB	63.64	ASP	91	HB2	2.540
LYS	83	HG2	1.44	SER	87	HB1	3.822	ASP	91	CG	180.064
LYS	83	CD	29.43	SER	87	HB2	3.822	ASP	91	C	175.83
LYS	83	HD1	1.67	SER	87	C	174.2	ARG	92	N	120.74
LYS	83	HD2	1.67	ILE	88	N	122	ARG	92	HN	8.355
LYS	83	CE	42.37	ILE	88	HN	8.224	ARG	92	CA	55.96
LYS	83	HE1	2.996	ILE	88	CA	60.8	ARG	92	HA	4.352
LYS	83	C	176.78	ILE	88	HA	4.210	ARG	92	CB	30.79
GLN	84	N	120.03	ILE	88	CB	38	ARG	92	HB1	1.75
GLN	84	HN	8.281	ILE	88	HB	1.844	ARG	92	HB2	1.269
GLN	84	CA	55.65	ILE	88	CG1	25.89	ARG	92	CG	26.96
GLN	84	HA	4.266	ILE	88	HG11	1.20	ARG	92	CD	43.22
GLN	84	CB	32.32	ILE	88	HG12	1.20	ARG	92	C	176.27
GLN	84	HB1	1.977	ILE	88	CG2	17.32	GLY	93	N	109.2
GLN	84	HB2	1.977	ILE	88	CD1	10.3	GLY	93	HN	8.357
GLN	84	CG	33.44	ILE	88	HD11	0.88	GLY	93	CA	44.2
GLN	84	HG1	2.340	ILE	88	HD12	0.88	GLY	93	HA1	3.973
GLN	84	HG2	2.340	ILE	88	HD13	0.88	GLY	93	HA2	3.973
GLN	84	C	176.59	ILE	88	C	176	PRO	94	CA	62.58
ARG	85	N	122.66	ILE	89	N	124.59	PRO	94	CB	33.76
ARG	85	HN	8.347	ILE	89	HN	8.251	PRO	94	C	177.17
ARG	85	CA	55.73	ILE	89	CA	60.67	MET	95	CA	56.5
ARG	85	HA	4.231	ILE	89	HA	4.208	MET	95	HA	4.40
ARG	85	CB	29.2	ILE	89	CB	37.7	MET	95	CB	34.5
ARG	85	HB1	1.75	ILE	89	HB	1.856	MET	95	HB1	2.286
ARG	85	CG	27.18	ILE	89	CG1	27.09	MET	95	HB2	1.947
ARG	85	HG1	1.580	ILE	89	HG11	1.49	TYR	96	N	124.63
ARG	85	HG2	1.580	ILE	89	HG12	1.20	TYR	96	HN	8.4
ARG	85	CD	43.5	ILE	89	CG2	17.32	TYR	96	CA	60.6
ARG	85	HD1	3.18	ILE	89	CD1	12.43	TYR	96	HA	4.169
ARG	85	HD2	3.18	ILE	89	HD11	0.88	TYR	96	CB	38.55
ARG	85	C	176.11	ILE	89	HD12	0.88	TYR	96	HB1	3.114
ARG	86	N	122	ILE	89	HD13	0.88	TYR	96	HB2	2.700
ARG	86	HN	8.339	ILE	89	C	175.2	TYR	96	C	174.84
ARG	86	CA	56.1	ARG	90	N	124.90	ASP	97	N	115.5
ARG	86	HA	4.316	ARG	90	HN	8.422	ASP	97	HN	7.644
ARG	86	CB	30.37	ARG	90	CA	55.7	ASP	97	CA	52.647
ARG	86	HB1	1.78	ARG	90	HA	4.147	ASP	97	HA	4.931
ARG	86	HB2	1.78	ARG	90	CB	30.54	ASP	97	CB	43.3
ARG	86	CG	27.36	ARG	90	CG	26.86	ASP	97	HB1	2.529
ARG	86	HG1	1.62	ARG	90	CD	43.43	ASP	97	C	174.06
ARG	86	CD	43.50	ARG	90	C	174.66	ASP	98	N	118.20

ASP 98 HN	8.788	GLU 103 CB	29.1	LYS 108 HB2	1.597
ASP 98 CA	50.01	GLU 103 HB1	1.986	LYS 108 CG	25.91
ASP 98 HA	5.313	GLU 103 HB2	1.986	LYS 108 HG1	1.390
ASP 98 CB	43.09	GLU 103 CG	35.55	LYS 108 HG2	1.310
ASP 98 HB1	3.02	GLU 103 HG1	2.32	LYS 108 CE	42.17
ASP 98 HB2	2.203	GLU 103 HG2	2.32	LYS 108 HE1	2.830
PRO 99 CA	63.2	GLU 103 C	175.4	LYS 108 HE2	2.830
PRO 99 HA	4.605	GLY 104 N	113.39	LYS 108 C	174.1
PRO 99 CB	31.9	GLY 104 HN	8.937	LEU 109 N	125.23
PRO 99 HB1	2.386	GLY 104 CA	45.32	LEU 109 HN	9.079
PRO 99 HB2	2.095	GLY 104 HA1	3.741	LEU 109 CA	53.6
PRO 99 CG	26.39	GLY 104 HA2	4.385	LEU 109 HA	5.384
PRO 99 C	176.8	GLY 104 C	173.2	LEU 109 CB	42.8
THR 100 N	109.64	TRP 105 N	121.15	LEU 109 HB1	1.499
THR 100 HN	8.414	TRP 105 HN	8.490	LEU 109 HB2	1.674
THR 100 CA	61.9	TRP 105 CA	57.2	LEU 109 CG	28.918
THR 100 HA	4.25	TRP 105 HA	5.233	LEU 109 HG	1.570
THR 100 CB	68.96	TRP 105 CB	29.2	LEU 109 CD1	25.521
THR 100 HB	4.11	TRP 105 HB1	3.46	LEU 109 HD11	0.468
THR 100 CG2	18.66	TRP 105 HB2	3.18	LEU 109 HD12	0.468
THR 100 HG21	1.139	TRP 105 C	176.64	LEU 109 HD13	0.468
THR 100 HG22	1.139	THR 106 N	112.39	LEU 109 CD2	25.197
THR 100 HG23	1.139	THR 106 HN	9.197	LEU 109 HD21	0.542
THR 100 C	175.4	THR 106 CA	60.8	LEU 109 HD22	0.542
LEU 101 N	123.3	THR 106 HA	4.736	LEU 109 HD23	0.542
LEU 101 HN	6.972	THR 106 CB	72.4	LEU 109 C	177.3
LEU 101 CA	53.152	THR 106 HB	4.107	LYS 110 N	122
LEU 101 HA	4.136	THR 106 CG2	21.70	LYS 110 HN	9.027
LEU 101 CB	41.8	THR 106 HG21	1.195	LYS 110 CA	54.6
LEU 101 HB1	1.65	THR 106 HG22	1.195	LYS 110 HA	4.885
LEU 101 HB2	1.36	THR 106 HG23	1.195	LYS 110 CB	35.04
LEU 101 CD1	26.67	THR 106 C	172	LYS 110 HB1	1.980
LEU 101 HD11	1.068	ARG 107 N	120	LYS 110 HB2	1.772
LEU 101 HD12	1.068	ARG 107 HN	8.430	LYS 110 CG	24.315
LEU 101 HD13	1.068	ARG 107 CA	53.8	LYS 110 HG1	1.387
LEU 101 CD2	24.89	ARG 107 HA	5.469	LYS 110 HG2	1.387
LEU 101 HD21	0.868	ARG 107 CB	33.5	LYS 110 CD	31.786
LEU 101 HD22	0.868	ARG 107 HB1	1.910	LYS 110 HD1	1.760
LEU 101 HD23	0.868	ARG 107 HB2	1.590	LYS 110 HD2	1.815
PRO 102 CA	61.985	ARG 107 CG	27.4	LYS 110 CE	41.39
PRO 102 HA	4.707	ARG 107 HG1	1.375	LYS 110 HE1	2.996
PRO 102 CB	30.45	ARG 107 C	175	LYS 110 HE2	3.038
PRO 102 CG	27.308	LYS 108 N	124.10	LYS 110 C	175.4
PRO 102 C	175.75	LYS 108 HN	9.471	GLN 111 N	126.23
GLU 103 N	120	LYS 108 CA	54.4	GLN 111 HN	9.192
GLU 103 HN	8.415	LYS 108 HA	4.899	GLN 111 CA	55.72
GLU 103 CA	58.22	LYS 108 CB	35.2	GLN 111 HA	4.24
GLU 103 HA	4.092	LYS 108 HB1	1.744	GLN 111 CB	29.2

GLN	111	HB1	1.827	ARG	116	HA	4.386	LYS	120	HE1	2.984
GLN	111	HB2	1.62	ARG	116	CB	29.66	LYS	120	HE2	2.984
GLN	111	CG	34.23	ARG	116	HB1	1.850	LYS	120	C	175.5
GLN	111	HG1	1.88	ARG	116	HB2	1.850	TYR	121	N	118.2
GLN	111	HG2	1.56	ARG	116	CG	27.17	TYR	121	HN	8.608
GLN	111	C	175.9	ARG	116	HG1	1.61	TYR	121	CA	57.7
ARG	112	N	125.97	ARG	116	HG2	1.61	TYR	121	HA	5.007
ARG	112	HN	8.806	ARG	116	CD	43.02	TYR	121	CB	40.8
ARG	112	CA	56.1	ARG	116	HD1	3.18	TYR	121	HB1	2.771
ARG	112	HA	4.970	ARG	116	HD2	3.18	TYR	121	HB2	2.671
ARG	112	CB	30.05	ARG	116	C	177.117	TYR	121	C	175.1
ARG	112	HB1	1.830	SER	117	N	113.93	ASP	122	N	119.5
ARG	112	HB2	1.830	SER	117	HN	8.394	ASP	122	HN	9.000
ARG	112	HG1	1.62	SER	117	CA	57.45	ASP	122	CA	53.5
ARG	112	CD	43.32	SER	117	HA	4.503	ASP	122	HA	5.015
ARG	112	HD1	3.066	SER	117	CB	63.05	ASP	122	CB	43.8
ARG	112	HD2	3.066	SER	117	HB1	3.908	ASP	122	HB1	2.675
LYS	113	N	122.0	SER	117	HB2	3.908	ASP	122	HB2	2.494
LYS	113	HN	8.831	SER	117	C	173.9	VAL	123	HN	7.56
LYS	113	CA	56.36	ALA	118	N	122.49	VAL	123	CA	62.1
LYS	113	HA	4.360	ALA	118	HN	7.752	VAL	123	HA	4.385
LYS	113	CB	32.7	ALA	118	CA	53.06	VAL	123	CB	32.43
LYS	113	HB1	1.860	ALA	118	HA	3.961	VAL	123	HB	1.85
LYS	113	HB2	1.690	ALA	118	CB	18.1	VAL	123	CG1	21.6
LYS	113	CG	24.96	ALA	118	HB1	1.260	VAL	123	HG11	0.74
LYS	113	HG1	1.400	ALA	118	HB2	1.260	VAL	123	C	174.5
LYS	113	CD	29.3	ALA	118	HB3	1.260	TYR	124	N	123.3
LYS	113	HD1	1.65	ALA	118	C	178	TYR	124	HN	9.011
LYS	113	CE	42.3	GLY	119	N	108.8	TYR	124	CA	55.7
LYS	113	HE1	2.921	GLY	119	HN	8.522	TYR	124	HA	5.295
LYS	113	C	176.1	GLY	119	CA	44.93	TYR	124	CB	39.7
SER	114	N	113.67	GLY	119	HA1	4.13	TYR	124	HB1	3.080
SER	114	HN	7.94	GLY	119	HA2	3.84	TYR	124	HB2	2.640
SER	114	CA	57.15	GLY	119	C	174	TYR	124	C	174.5
SER	114	HA	4.637	LYS	120	N	120	LEU	125	N	121.4
SER	114	CB	64.56	LYS	120	HN	8.006	LEU	125	HN	9.125
SER	114	HB1	3.844	LYS	120	CA	55.68	LEU	125	CA	53.3
SER	114	HB2	3.810	LYS	120	HA	4.45	LEU	125	HA	5.45
SER	114	C	176.27	LYS	120	CB	33.05	LEU	125	CB	42.99
GLY	115	N	108.8	LYS	120	HB1	2.25	LEU	125	HB1	1.610
GLY	115	HN	8.596	LYS	120	HB2	1.9	LEU	125	HB2	1.920
GLY	115	CA	45.12	LYS	120	CG	24.71	LEU	125	CG	24.38
GLY	115	HA1	4.188	LYS	120	HG1	1.315	LEU	125	HG	0.850
GLY	115	HA2	3.88	LYS	120	HG2	1.315	LEU	125	CD1	23.36
GLY	115	C	177.03	LYS	120	CD	28.59	LEU	125	HD12	0.758
ARG	116	N	120	LYS	120	HD1	1.66	LEU	125	HD12	0.758
ARG	116	HN	8.631	LYS	120	HD2	1.66	LEU	125	HD13	0.758
ARG	116	CA	57.21	LYS	120	CE	42.5	LEU	125	CD2	22.769

LEU	125	HD21	0.812	GLN	129	C	176.2	ARG	134	CG	26.65
LEU	125	HD22	0.812	GLY	130	N	107.50	ARG	134	HG1	1.455
LEU	125	HD23	0.812	GLY	130	HN	8.147	ARG	134	HG2	1.455
LEU	125	C	175.9	GLY	130	CA	45.2	ARG	134	CD	42.6
ILE	126	N	123.3	GLY	130	HA1	4.09	ARG	134	HD1	3.055
ILE	126	HN	9.590	GLY	130	HA2	3.36	ARG	134	HD2	3.007
ILE	126	CA	59.15	GLY	130	C	173.4	ARG	134	C	175.3
ILE	126	HA	5.054	LYS	131	N	122.49	SER	135	N	111.1
ILE	126	CB	37.83	LYS	131	HN	7.731	SER	135	HN	7.295
ILE	126	HB	1.910	LYS	131	CA	55.3	SER	135	CA	55.716
ILE	126	CG1	27.46	LYS	131	HA	3.98	SER	135	HA	4.895
ILE	126	HG11	1.374	LYS	131	CB	31.59	SER	135	CB	63.55
ILE	126	HG12	1.052	LYS	131	HB1	1.236	SER	135	HB1	3.723
ILE	126	CG2	17.26	LYS	131	HB2	1.236	SER	135	C	174.6
ILE	126	HG21	0.768	LYS	131	CG	24.86	LYS	136	N	121.4
ILE	126	HG22	0.768	LYS	131	HG1	1.011	LYS	136	HN	8.511
ILE	126	HG23	0.768	LYS	131	HG2	0.691	LYS	136	CA	59.18
ILE	126	HD11	0.710	LYS	131	CD	28.72	LYS	136	HA	3.517
ILE	126	HD12	0.710	LYS	131	HD1	1.424	LYS	136	CB	31.13
ILE	126	HD13	0.710	LYS	131	HD2	1.280	LYS	136	HB1	1.770
ILE	126	C	176.1	LYS	131	CE	41.54	LYS	136	HB2	2.73
ASN	127	N	125.88	LYS	131	HE1	2.900	LYS	136	CG	25.04
ASN	127	HN	8.464	LYS	131	HE2	2.678	LYS	136	HG1	1.458
ASN	127	CA	50.8	LYS	131	C	173.2	LYS	136	HG2	1.458
ASN	127	HA	5.052	ALA	132	N	124.59	LYS	136	CD	29.0
ASN	127	HB1	2.194	ALA	132	HN	7.926	LYS	136	HD1	1.77
ASN	127	HB2	1.719	ALA	132	CA	50.2	LYS	136	HD2	1.77
ASN	127	CG	176.48	ALA	132	HA	5.139	LYS	136	CE	42.83
ASN	127	ND2	108.2	ALA	132	CB	21.7	LYS	136	HE1	2.986
ASN	127	HD21	6.926	ALA	132	HB1	1.214	LYS	136	HE2	2.986
ASN	127	HD22	6.655	ALA	132	HB2	1.214	LYS	136	C	177.7
PRO	128	CA	64.74	ALA	132	HB3	1.214	VAL	137	N	116.01
PRO	128	HA	4.239	ALA	132	C	177	VAL	137	HN	7.653
PRO	128	CB	30.8	PHE	133	N	119.0	VAL	137	CA	66
PRO	128	HB1	2.372	PHE	133	HN	9.641	VAL	137	HA	3.694
PRO	128	HB2	2.050	PHE	133	CA	57	VAL	137	CB	31.1
PRO	128	CG	27.5	PHE	133	HA	4.903	VAL	137	HB	1.946
PRO	128	C	177.1	PHE	133	CB	42.3	VAL	137	CG1	19.77
GLN	129	N	112.39	PHE	133	HB1	3.111	VAL	137	HG11	1.03
GLN	129	HN	6.938	PHE	133	HB2	3.111	VAL	137	HG12	1.03
GLN	129	CA	55.81	PHE	133	C	175.8	VAL	137	HG13	1.03
GLN	129	HA	4.265	ARG	134	N	119.0	VAL	137	CG2	18.65
GLN	129	CB	27.77	ARG	134	HN	9.402	VAL	137	HG21	0.98
GLN	129	HB1	2.008	ARG	134	CA	55.09	VAL	137	HG22	0.98
GLN	129	HB2	2.270	ARG	134	HA	4.900	VAL	137	HG23	0.98
GLN	129	CG	34.58	ARG	134	CB	31.4	VAL	137	C	178.7
GLN	129	HG1	2.285	ARG	134	HB1	1.818	GLU	138	N	119.0
GLN	129	HG2	2.285	ARG	134	HB2	1.763	GLU	138	HN	7.415

GLU 138 CA	58.87	ALA 141 HB1	1.536	VAL 146 HB	1.803
GLU 138 HA	4.189	ALA 141 HB2	1.536	VAL 146 CG1	15.31
GLU 138 CB	30	ALA 141 HB3	1.536	VAL 146 HG11	0.774
GLU 138 HB1	2.41	ALA 141 C	180.4	VAL 146 HG12	0.774
GLU 138 HB2	2.41	TYR 142 N	119.51	VAL 146 HG13	0.774
GLU 138 CG	36.5	TYR 142 HN	7.780	VAL 146 CG2	20.86
GLU 138 HG1	2.39	TYR 142 CA	61.16	VAL 146 C	177
GLU 138 HG2	2.39	TYR 142 HA	4.43	GLY 147 N	111.1
GLU 138 C	179.7	TYR 142 CB	38.54	GLY 147 HN	7.802
LEU 139 N	120.7	TYR 142 HB1	3.383	GLY 147 CA	46.46
LEU 139 HN	7.528	TYR 142 HB2	2.984	GLY 147 HA1	4.000
LEU 139 CA	57.56	TYR 142 C	176.9	GLY 147 HA2	3.856
LEU 139 HA	3.423	PHE 143 N	117.7	GLY 147 C	174.1
LEU 139 CB	41.99	PHE 143 HN	8.791	ASP 148 N	119.28
LEU 139 HB1	1.404	PHE 143 CA	58.069	ASP 148 HN	8.107
LEU 139 HB2	1.38	PHE 143 HA	4.457	ASP 148 CA	53.495
LEU 139 CG	25.35	PHE 143 CB	36.36	ASP 148 HA	4.785
LEU 139 HG	0.82	PHE 143 C	179.4	ASP 148 CB	41
LEU 139 CD1	24.87	GLU 144 N	117.53	ASP 148 HB1	2.659
LEU 139 HD11	-0.01	GLU 144 HN	8.333	ASP 148 HB2	2.419
LEU 139 HD12	-0.01	GLU 144 CA	58.7	ASP 148 CG	179.03
LEU 139 HD13	-0.01	GLU 144 HA	4.128	ASP 148 C	176.6
LEU 139 CD2	24.4	GLU 144 CB	28.8	THR 149 N	115.5
LEU 139 HD21	0.89	GLU 144 HB1	2.15	THR 149 HN	8.447
LEU 139 HD22	0.89	GLU 144 HB2	2.15	THR 149 CA	60.66
LEU 139 HD23	0.89	GLU 144 CG	36.82	THR 149 HA	4.5
LEU 139 C	177.1	GLU 144 HG1	2.483	THR 149 CB	68.6
ILE 140 N	117.5	GLU 144 HG2	2.483	THR 149 HB	4.517
ILE 140 HN	8.403	GLU 144 C	178.8	THR 149 CG2	19.21
ILE 140 CA	65.56	LYS 145 N	120.35	THR 149 HG21	1.1
ILE 140 HA	3.506	LYS 145 HN	7.851	THR 149 HG22	1.1
ILE 140 CB	37.79	LYS 145 CA	58.7	THR 149 HG23	1.1
ILE 140 HB	1.762	LYS 145 HA	4.116	THR 149 C	175.1
ILE 140 CG1	27	LYS 145 CB	31.7	SER 150 N	118.167
ILE 140 HG11	1.25	LYS 145 HB1	2.01	SER 150 HN	8.674
ILE 140 HG12	1.45	LYS 145 CG	24.46	SER 150 CA	60.05
ILE 140 CG2	16.83	LYS 145 HG1	1.51	SER 150 HA	4.286
ILE 140 HG21	0.964	LYS 145 HG2	1.51	SER 150 CB	63.94
ILE 140 HG22	0.964	LYS 145 CD	28.93	SER 150 HB1	3.852
ILE 140 HG23	0.964	LYS 145 CE	42.38	SER 150 HB2	3.852
ILE 140 CD1	11.1	LYS 145 HE1	2.966	SER 150 C	174.2
ILE 140 HD11	0.88	LYS 145 HE2	2.966	LEU 151 N	122.22
ILE 140 C	178.3	LYS 145 C	179	LEU 151 HN	7.668
ALA 141 N	119.0	VAL 146 N	109.64	LEU 151 CA	53.4
ALA 141 HN	7.56	VAL 146 HN	8.016	LEU 151 HA	4.399
ALA 141 CA	54.68	VAL 146 CA	61.67	LEU 151 CB	43.4
ALA 141 HA	4.229	VAL 146 HA	4.29	LEU 151 HB1	1.123
ALA 141 CB	17	VAL 146 CB	31	LEU 151 HB2	0.966

LEU	151	CG	26.5	ASP	157	CA	53.25	THR	161	HG23	1.04
LEU	151	HG	1.600	ASP	157	HA	4.565	THR	161	C	175
LEU	151	CD1	26.16	ASP	157	CB	40.65	GLY	162	N	110.45
LEU	151	HD11	0.982	ASP	157	HB1	2.836	GLY	162	HN	7.472
LEU	151	HD12	0.982	ASP	157	HB2	2.427	GLY	162	CA	44.64
LEU	151	HD13	0.982	ASP	157	C	176	GLY	162	HA1	3.945
LEU	151	CD2	22.78	PHE	158	N	124.59	GLY	162	HA2	4.333
LEU	151	C	174.6	PHE	158	HN	9.148	GLY	162	C	177.34
ASP	152	N	121.96	PHE	158	CA	58.66	ARG	163	N	121.68
ASP	152	HN	8.864	PHE	158	HA	4.407	ARG	163	HN	9.116
ASP	152	CA	50.2	PHE	158	CB	39.04	ARG	163	CA	56.28
ASP	152	HA	4.98	PHE	158	HB1	3.508	ARG	163	HA	4.102
ASP	152	CB	42.18	PHE	158	HB2	2.798	ARG	163	CB	30.39
ASP	152	HB1	3.070	PHE	158	C	175.5	ARG	163	HB1	1.738
ASP	152	HB2	2.450	THR	159	N	112.5	ARG	163	HB2	1.738
PRO	153	CA	64.63	THR	159	HN	9.109	ARG	163	CG	26.6
PRO	153	HA	4.084	THR	159	CA	61.93	ARG	163	CD	43.22
PRO	153	CB	31.9	THR	159	HA	4.576	ARG	163	HD1	3.143
PRO	153	HB1	2.064	THR	159	CB	70.3	ARG	163	HD2	3.143
PRO	153	CG	27.48	THR	159	HB	4.4	ARG	163	C	178.05
PRO	153	CD	51.42	THR	159	CG2	15.87	GLY	164	N	108.57
PRO	153	C	177.5	THR	159	HG21	1.375	GLY	164	HN	8.522
ASN	154	N	113.92	THR	159	HG22	0.77	GLY	164	CA	44.7
ASN	154	HN	8.872	THR	159	C	176.6	GLY	164	HA1	3.9
ASN	154	CA	54.99	VAL	160	N	123.3	GLY	164	HA2	4.161
ASN	154	HA	4.552	VAL	160	HN	8.99	GLY	164	C	173.77
ASN	154	CB	37.67	VAL	160	CA	66.69	SER	165	N	116.06
ASN	154	HB1	2.788	VAL	160	HA	3.36	SER	165	HN	8.205
ASN	154	HB2	2.650	VAL	160	CB	30.36	SER	165	CA	56.16
ASN	154	C	176.25	VAL	160	HB	2.113	SER	165	HA	4.632
ASP	155	N	118.2	VAL	160	CG1	19.77	SER	165	CB	62.73
ASP	155	HN	8.400	VAL	160	HG11	0.786	SER	165	HB1	3.878
ASP	155	CA	54.84	VAL	160	HG12	0.786	PRO	166	CA	63.44
ASP	155	HA	4.55	VAL	160	HG13	0.786	PRO	166	HA	3.979
ASP	155	CB	39.7	VAL	160	CG2	21.99	PRO	166	CB	31.3
ASP	155	HB1	2.544	VAL	160	HG21	1.16	PRO	166	HB1	1.756
ASP	155	HB2	2.544	VAL	160	HG22	1.16	PRO	166	CG	27.25
ASP	155	C	174.9	VAL	160	HG23	1.16	PRO	166	C	177.28
PHE	156	N	118.47	VAL	160	C	177.05	SER	167	N	113.67
PHE	156	HN	7.121	THR	161	N	104.83	SER	167	HN	8.339
PHE	156	CA	57.375	THR	161	HN	7.912	SER	167	CA	58.3
PHE	156	HA	3.994	THR	161	CA	61.75	SER	167	HA	4.384
PHE	156	CB	38.57	THR	161	HA	4.229	SER	167	CB	63.3
PHE	156	HB1	2.138	THR	161	CB	69.56	SER	167	HB1	3.818
PHE	156	HB2	2.840	THR	161	HB	4.360	SER	167	HB2	3.818
PHE	156	C	173.6	THR	161	CG2	27.8	SER	167	C	174.69
ASP	157	N	123.56	THR	161	HG21	1.04	ARG	168	N	122
ASP	157	HN	8.937	THR	161	HG22	1.04	ARG	168	HN	8.459

ARG 168 CA	55.9	LYS 175 C	176.63	SER 182 HN	8.012
ARG 168 HA	4.31	LYS 176 N	122.6	SER 182 CA	56.57
ARG 168 CB	30.3	LYS 176 HN	8.348	SER 182 HA	4.78
ARG 168 CG	26.93	LYS 176 CA	55.9	PRO 183 CA	63.3
ARG 168 CD	43.25	LYS 176 HA	4.221	PRO 183 HA	4.342
ARG 168 C	176.24	LYS 176 CB	32.813	PRO 183 CB	31.4
ARG 169 N	121	LYS 176 HB1	1.750	PRO 183 HB1	1.832
ARG 169 HN	8.4	LYS 176 CG	24.42	PRO 183 CG	27.0
ARG 169 CA	56	LYS 176 CD	28.8	PRO 183 C	177.12
ARG 169 HA	4.327	LYS 176 CE	41.76	GLY 184 N	108.53
ARG 169 CB	32.42	LYS 176 C	176.09	GLY 184 HN	8.487
ARG 169 HB1	2.66	ALA 177 N	125.170	GLY 184 CA	45.04
ARG 169 CG	24.39	ALA 177 HN	8.395	GLY 184 HA1	3.99
ARG 169 C	176.22	ALA 177 CA	52.007	GLY 184 HA2	3.99
GLU 170 N	118.2	ALA 177 HA	4.289	GLY 184 C	174.2
GLU 170 HN	8.401	ALA 177 CB	19.43	SER 185 N	114.73
GLU 170 CA	56.08	ALA 177 HB1	1.381	SER 185 HN	8.278
GLU 170 HA	4.271	ALA 177 HB2	1.381	SER 185 CA	58.21
GLU 170 CB	32.38	ALA 177 HB3	1.381	SER 185 HA	4.765
GLU 170 HB1	1.784	ALA 177 C	177.54	SER 185 CB	63.6
GLU 170 CG	24.41	LYS 178 N	120.7	SER 185 HB1	3.919
GLU 170 C	175.86	LYS 178 HN	8.394	SER 185 C	177.21
GLN 171 N	116.34	LYS 178 CA	55.77	GLY 186 N	109.91
GLN 171 HN	8.121	LYS 178 HA	4.33	GLY 186 HN	8.547
GLN 171 CA	55.44	LYS 178 CB	32.44	GLY 186 CA	45.3
GLN 171 HA	4.616	LYS 178 HB1	1.813	GLY 186 HA1	3.984
GLN 171 CB	28.82	LYS 178 C	175.5	GLY 186 HA2	3.984
GLN 171 HB1	2.301	SER 179 N	118.475	GLY 186 C	174.23
GLN 171 HB2	1.962	SER 179 HN	8.546	ARG 187 N	119.54
GLN 171 CG	33.58	SER 179 CA	56.18	ARG 187 HN	8.285
GLN 171 C	175.63	SER 179 HA	4.400	ARG 187 CA	55.87
ARG 172 N	122.75	SER 179 CB	63.78	ARG 187 HA	4.76
ARG 172 HN	8.434	SER 179 HB1	3.82	ARG 187 CB	30.39
ARG 172 CA	53.57	SER 179 HB2	3.82	ARG 187 HB1	2.370
ARG 172 HA	4.744	PRO 180 CA	62.49	ARG 187 HB2	2.070
PRO 174 CA	62.27	PRO 180 HA	4.723	ARG 187 CG	26.6
PRO 174 CB	31.51	PRO 180 CB	34	ARG 187 CD	43.22
PRO 174 CG	26.81	PRO 180 HB1	2.369	ARG 187 C	177
PRO 174 C	176.88	PRO 180 HB2	2.126	GLY 188 N	109.1
LYS 175 N	121.15	PRO 180 CG	24.42	GLY 188 HN	8.510
LYS 175 HN	8.364	PRO 180 C	176.18	GLY 188 CA	44.88
LYS 175 CA	56.02	LYS 181 N	120.89	GLY 188 HA1	3.809
LYS 175 HA	4.994	LYS 181 HN	8.730	GLY 188 HA2	3.809
LYS 175 CB	32.80	LYS 181 CA	56.72	GLY 188 C	173.17
LYS 175 HB1	1.757	LYS 181 HA	4.734	ARG 189 N	114.46
LYS 175 HB2	2.668	LYS 181 CB	29.84	ARG 189 HN	7.912
LYS 175 CG	24.51	LYS 181 C	176.48	ARG 189 CA	55.61
LYS 175 CD	28.97	SER 182 N	120.7	ARG 189 HA	4.578

ARG 189 CB	30.56	PRO 192 CG	27.13	GLY 194 CA	45.1
ARG 189 HB1	1.407	PRO 192 CD	50.477	GLY 194 HA1	4.01
ARG 189 C	176.28	PRO 192 C	177.02	GLY 194 HA2	4.01
GLY 190 N	109.11	LYS 193 N	121.4	GLY 194 C	174.18
GLY 190 HN	8.239	LYS 193 HN	8.55	SER 195 N	114.96
GLY 190 CA	45.18	LYS 193 CA	56.48	SER 195 HN	8.243
GLY 190 HA1	3.911	LYS 193 HA	4.384	SER 195 CA	58.06
GLY 190 HA2	3.911	LYS 193 CB	32.48	SER 195 HA	4.769
GLY 190 C	173.22	LYS 193 HB1	1.841	SER 195 CB	64.23
ARG 191 N	119.01	LYS 193 CG	24.49	SER 195 HB1	3.988
ARG 191 HN	7.942	LYS 193 CD	28.88	SER 195 HB2	3.988
ARG 191 CA	53.41	LYS 193 HD1	1.65	SER 195 C	174.02
PRO 192 CA	62.85	LYS 193 CE	41.84	GLY 196 N	116.07
PRO 192 HA	4.387	LYS 193 HE1	3.198	GLY 196 HN	8.168
PRO 192 CB	31.16	LYS 193 C	177.09	GLY 196 CA	46.07
PRO 192 HB1	1.899	GLY 194 N	110.5		
PRO 192 HB2	2.310	GLY 194 HN	8.529		

Anhang B

restraints: Diederwinkel und Restdipolkopplungen

B.1 Diederwinkel als *restraints* bestimmt mit TALOS

Aminosäurerest	Φ	$\Delta\Phi$	Ψ	$\Delta\Psi$
T 106	131	25	153	15
R 107	119	20	127	19
K 108	128	20	138	22
L 109	104	9	132	10
Q 111	-88	11	124	12
K 120	-90	15	136	16
D 122	118	27	137	12
V 123	100	18	131	16
Y 124	109	22	147	13
L 125	114	15	127	10
I 126	108	26	127	21
K 131	-99	26	118	25
A 132	106	19	137	14
F 133	115	20	141	10
S 135	114	26	163	13

V 137	-66	6	-35	10
E 138	-65	6	-39	6
L 139	-67	6	-41	5
I 140	-65	4	-38	7
A 141	-64	6	-39	5
Y 142	-71	14	-39	9
F 143	-68	14	-40	10
E 144	-64	6	-41	8
V 146	-96	9	1	12
D 152	107	23	114	26
P 153	-59	3	-20	9
N 154	-89	14	-13	23
T 159	-92	16	158	17
V 160	-71	12	-24	15
T 161	102	13	5	12
R 163	-97	9	0	15

B.2 Restdipolkopplungen als *restraints*

B.2.1 $^1\text{H}_N$ - ^{15}N -Restdipolkopplungen

Aminosäurerest	Kopplung /Hz	Fehler /Hz
106	1,2	1,4
107	6,3	1,4
109	-19,4	1,4
110	-22,2	1,4
111	-18,5	2
122	-16,4	1,4
125	-5,4	1,4
126	6	1,4
133	-3,3	1,4

134	7,4	1,4
137	19,9	2
138	21,7	1,4
139	10,5	1,4
141	18	1,4
142	19,7	1,4
144	11,6	2
145	19,7	1,4
146	17,2	1,4
154	14,1	1,4
155	16,7	2
159	9	1,4
161	-11,5	2
162	-3,6	1,4
163	6,1	1,4

B.2.2 $^{13}\text{C}_\alpha$ - $^{13}\text{C}'$ -Restdipolkopplungen

Aminosäurerest	Kopplung /Hz	Fehler /Hz
106	0,4	1,4
107	3,7	2
109	1,9	1,4
110	-2,2	1,4
121	-1,1	1,4
123	2,2	1,4
124	0,1	1,4
133	-1,8	1,4
137	1,2	1,4
138	-2,9	2
140	-1,9	1,4
141	3,2	1,4
142	-0,6	1,4

143	0,3	1,4
145	2,7	1,4
153	-1,3	1,4
155	-1,3	1,4
161	0,8	1,4
163	2,9	1,4

B.2.3 ^{15}N - $^{13}\text{C}'$ -Restdipolkopplungen

Aminosäurerest	Kopplung /Hz	Fehler /Hz
106	-1	1,4
108	0,66	2
109	3	1,4
111	0	1,4
125	1,5	1,4
133	1,5	1,4
134	0	1,4
137	-0,9	1,4
142	0,3	1,4
145	0,6	1,4
154	1,3	1,4

Danksagung

An dieser Stelle möchte ich mich bei allen, die zum Gelingen dieser Arbeit beigetragen haben, bedanken.

Besonders möchte ich Prof. Dr. Eike Brunner und Prof. Dr. Dr. Hans Robert Kalbitzer für die Vergabe dieses interessanten Themas bedanken.

Prof. Dr. Eike Brunner danke ich herzlich für die exzellente Betreuung in allen Belagen meiner Arbeit. Er fand immer Zeit und hat nie die Geduld verloren, mir bei meinen Fragen zu helfen, was bestimmt nicht immer einfach war. Dies geschah alles trotz der zeitlichen Mehrbelastung infolge seiner Berufung. Prof. Dr. Dr. Hans Robert Kalbitzer sei dafür gedankt, mich in seiner sympathischen Arbeitsgruppe aufgenommen und mir die Möglichkeit gegeben zu haben einen Einblick in die NMR-Spektroskopie zu bekommen.

Prof. Dr. Wolf H. Strätling und seinen Mitarbeitern sei gedankt für die Bereitstellung der Proteinproben und für die geduldige Behandlung eines Physikers im Neuland der Biochemie und Medizin.

Der DFG und der Universität Regensburg danke für die finanzielle Unterstützung (Br 1278/8-1 und Br 1278/8-2).

Dr. Till Maurer danke ich neben der netten persönlichen Betreuung für die Einführung in die Spektrometerhandhabung und in die Kniffe der Spektrenauswertung.

PD Dr. Wolfram Gronwald und Dr. Werner Kremer möchte ich für ihre Hilfsbereitschaft in allen fachlichen, spektroskopischen und computertechnischen Belangen danken. Besonders sei PD Dr. Wolfram Gronwald und seinen Mannen für die Arbeit im Softwarebereich gedankt - Sie waren immer für Anregungen offen das leistungsstarke Softwarepaket AUREMOL dem Benutzer optimal anzupassen.

Müßte ich diese Danksagung an dieser Stelle mit einem letzten Namen abbrechen, so wäre dies Michael Wenzler. Mein langjähriger Zimmergenosse, der sich für keine meiner Fragen zu schade war, dem keines meiner Probleme kalt ließ und der immer ein offenes Ohr hatte bei der Diskussion von fachlichen wie weniger fachlichen Dingen. Durch ihn weiß ich, wie man sich gegebenenfalls an eine Eisenbahnstrecke kettet, daß meine Risikobereitschaft beim Langlaufen noch ausbaufähig ist und daß es ^{15}N -markierte Glucose gar nicht gibt.

Jetzt mal im Ernst - Danke.

Meinen anderen Zimmermitinsassen sei gedankt: Dr. Guido Steiner für die Hilfestellungen bei meinen ersten Gehversuchen in TEX und perl, Sonja Lorenz für das Unter-die-Arme-greifen in allen Belangen des geschriebenen Wortes und des Schöngesties, Fritz Huber für seine besonnene Art mit der sich der nötige Abstand zur Materie wieder finden ließ, Adrina Campeanu für meinen ersten Satz auf Rumänisch.

Hubert Riepl, Dr. Harry Kany, Dr. Bernhard Ganslmeier, Dr. Michael Spörner, Renate Kirchhöfer, Barbara Wisotzki, Dr. Rolf Döker, Frau Dr. Gudrun Horn und Ralf Elsner danke ich für ihre kollegiale Art und Hilfsbereitschaft.

Frau Ingrid Cuno, Frau Emmi Fuchs, Frau Ingrid Kulbartz, Herrn Gunnar Schmitt, Herrn Johann Krauß und Herrn Kurt Schindler sei gedankt, daß sie ein Nordlicht wie mich nicht gleich von der Biergartenbank geschubst haben. Besonders gedankt sei Frau Cuno für die Durchsicht des Manuskripts dieser Arbeit.

Sämlichen Mitgliedern und ehemaligen Mitgliedern (Jörg, Sebastian, Uli - schee wars) des Lehrstuhls, die mir die Zeit in Regensburg verschönt haben, sei gedankt.

Gesonderten Dank verdienen meine Eltern, die mir mein Studium ermöglicht haben und mich in Ruhe haben machen lassen.

Ganz besonders möchte ich Andrea danken, die so brav nicht gedrängelt hat mit der Fertigwerderei. Dafür einen Kuß.

Björn

Erklärung

Hiermit erkläre ich, daß ich die vorliegende Arbeit selbständig angefertigt und keine außer den angegebenen Hilfsmitteln verwendet habe.

Regensburg, im November 2002

(Björn Heitmann)