

Department Chemie  
Lehrstuhl II für Organische Chemie der Technischen Universität München

**Moleküldynamik und Strukturaufklärung  
von Peptidmimetika und dem Parvulin *E. coli* Par10**

Georg Voll

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. A. Türler

Prüfer der Dissertation:

1. Univ.-Prof. Dr. H. Kessler
2. Univ.-Prof. Dr. W. Hiller

Die Dissertation wurde am 30. Juni 2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 9. September 2004 angenommen.

Die vorliegende Arbeit wurde während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Organische Chemie II im Department Chemie der Technischen Universität München in Garching in der Zeit von Juli 1998 bis Juni 2004 unter der wissenschaftlichen Betreuung von Prof. Dr. Horst Kessler angefertigt. Ihm gilt an dieser Stelle mein besonderer Dank für seine uneingeschränkte Unterstützung in allen Belangen, die Bereitstellung hervorragender Arbeitsbedingungen, die freie und ungezwungene Arbeitsatmosphäre und für das mir entgegengebrachte Vertrauen.

PD Dr. Gerd Gemmecker möchte ich ebenfalls für die immer kompetente Beratung, die vielen lehrreichen Diskussionen, das Korrekturlesen dieser Arbeit und nicht zuletzt für seine fast erschreckend ausnahmslose Freundschaftlichkeit danken.

Mein weiterer Dank gilt:

- meinen Kollegen Dr. Kay-E. Gottschalk, Dr. Rainer Gratias, Axel Meyer, Eckart Planker und Dr. Christian Rölz für das hervorragende Miteinander und die zahlreichen wissenschaftlichen Anregungen und Diskussionen,
- Dr. Rainer Haefner für unermüdliche Hilfe bei Computerproblemen aller Art und für die Einführung in die Welt der Systemadministration,
- Dr. Murray Coles für die Einführung in die Geheimnisse der Aufklärung von Proteinstrukturen,
- PD Dr. Gerd Gemmecker und Dr. Angelika Kühlewein für die interessante Zeit während der Zusammenarbeit bei der Strukturaufklärung des *E. coli* Parvulins Par10,
- Wirbelwind Dr. Sibylle Gruner und Hitzkopf Dr. Vincent Truffault für die spannende Zusammenarbeit bei der Strukturaufklärung der 12/10/12-Helix,
- Dr. Matthias Stöckle für die Zusammenarbeit im „Gum“-Projekt,
- Dr. Markus Born, Mandar Deshmukh, Dr. Dirk Gottschling, Dr. Luciana Marinelli und Dr. Christoph Riemer für die Zusammenarbeit in weiteren, zum Teil anderweitig veröffentlichten Projekten,

- Dr. Angelika Kühlewein für gelegentliches, unvermitteltes Überschäumen an guter Laune – immer dann, wenn man es am wenigsten erwartet, aber am meisten braucht –, sowie für ein Chamäleon mit Henkel,
- Dr. Claudia Lehotkay für den Beweis, dass ehrliche Bescheidenheit noch nicht ausgestorben ist,
- Puschi Ohr und den anderen zahmen Eichhörnchen des Arbeitskreises für das Begutachten der Abbildungen in dieser Arbeit, Aufmunterung, und gelegentliches „aus der Hand fressen“,
- Dr. Elsa Locardi für ihren messerscharfen Humor, der blitzartig zuschlägt, bevor man auch nur die Andeutung eines Themas hervorbringen kann,
- Dr. Manuel Tönnis, der sich wohl nicht zuletzt wegen seiner hochgradig ansteckenden Begeisterung für wissenschaftliche und andere Gebiete vornehmend im biologischen Sicherheitstrakt aufgehalten hat, für die Hilfe beim Betreuen von Praktika,
- Dipl. Biol. Alexander Frenzel für die Unterstützung bei logistischen Computerproblemen sowie bei biochemischen Fragestellungen,
- Dr. Mandar Deshmukh für anregende Diskussionen über Politik, Musik, und Geschichte,
- den Sekretärinnen Frau Bruckmaier und Frau Machule sowie Dekanatssekretärin Frau Delanoff für ihre Hilfe und ihre professionelle Arbeit,
- allen anderen derzeitigen und ehemaligen Mitarbeitern für tausende von Mücken und Elefanten, die die Arbeit zum Vergnügen gemacht haben,

... und meinen Eltern, denen ich alles verdanke, überhaupt erst in der Lage zu sein, danken zu können.

# Inhalt

<b>1</b>	<b>Einleitung</b> .....	<b>1</b>
<b>2</b>	<b>Methoden der Strukturberechnung</b> .....	<b>3</b>
2.1	Distanzgeometrie.....	3
2.1.1	Abstandsmatrizen.....	3
2.1.2	Erzeugung von Zufallsstrukturen (Metrisierung).....	5
2.1.3	Koordinatentransformation (Einbettung).....	5
2.2	Kraftfeldrechnungen.....	6
2.2.1	Das Kraftfeld.....	6
2.2.1.1	Distanzabhängiges Kraftfeld.....	7
2.2.1.2	CVFF.....	8
2.2.1.3	AMBER.....	10
2.2.1.4	CHARMM und PARALLHDG.....	11
2.2.1.5	Experimentelle Einschränkungen.....	12
2.2.2	Berechnung weitreichender Wechselwirkungen.....	14
2.2.3	Periodische Randbedingungen.....	15
2.2.4	Abbruchradien.....	17
2.2.5	Methoden zur Konformationsraumsuche.....	19
2.2.5.1	Minimierungsalgorithmen.....	20
2.2.5.2	Moleküldynamik.....	22
2.2.5.2.1	Grundlagen.....	22
2.2.5.2.2	Methoden zur Temperaturkontrolle.....	24
2.2.5.3	<i>Simulated Annealing</i> .....	26
2.2.5.4	<i>Time Averaged Restraints</i> -Verfahren.....	26
2.2.6	Automatisches, iteratives NOE-Assignment: ARIA.....	28
2.2.6.1	$r^6$ summierte Distanzen.....	31
2.2.6.2	Entfernen von Zuweisungsmöglichkeiten.....	33
2.2.7	Sortieren von Strukturen in Familien.....	34
<b>3</b>	<b>Struktur von <i>cyclo</i>(-Gum-Lys-Gum-D-Phe-) in Lösung</b> .....	<b>36</b>
3.1	Einleitung.....	36
3.2	Gum und andere Zuckeraminosäuren in cyclischen Peptiden.....	36
3.3	<i>Cyclo</i> (-Gum-Lys-Gum-D-Phe-) 4 als cyclisches Hexapeptid.....	40
3.3.1	NMR-Strukturdaten.....	41
3.3.2	Strukturrechnungen.....	42
3.3.3	Ergebnisse und Diskussion.....	44
3.4	<i>Cyclo</i> (-Gum-Lys-Gum-D-Phe-) 4 als Cyclohexan gemäß DUNITZ und WASER.....	46
3.4.1	Grundlagen.....	46
3.4.2	Ausweitung auf <i>cyclo</i> (-Gum-Lys-Gum-D-Phe-) 4.....	48
<b>4</b>	<b><math>\beta</math>-Peptide: Strukturbestimmung einer 12/10/12-Helix in Lösung</b> .....	<b>53</b>
4.1	Einleitung.....	53
4.2	Grundlagen.....	54
4.2.1	$\beta$ -Aminosäuren und aus ihnen aufgebaute Peptide.....	54
4.2.2	Quantenmechanische Vorhersage: Heterogene Helixtypen.....	56

---

4.3	Erstellung einer Solvenszelle für deuteriertes Acetonitril.....	59
4.4	Strukturbestimmung von Fmoc-[f-SAA- $\beta$ -hGly] <sub>3</sub> -OH.....	61
4.4.1	CD-spektroskopische Untersuchungen .....	62
4.4.2	NMR-Strukturdaten .....	64
4.4.3	Strukturrechnungen .....	65
4.4.4	Ergebnisse und Diskussion .....	66
<b>5</b>	<b>Das Parvulin <i>E. coli</i> Par10 .....</b>	<b>74</b>
5.1	Biologischer Hintergrund .....	74
5.1.1	Die Faltung von Proteinen .....	74
5.1.2	Die PPlase-Familien .....	77
5.2	NMR-spektroskopische Untersuchungen an Par10.....	80
5.3	Automatische NOE-Zuordnung mit ARIA .....	81
5.4	Strukturverfeinerung .....	85
5.5	Strukturdiskussion .....	87
5.6	Strukturvergleich mit anderen PPlasen .....	98
<b>6</b>	<b>Zusammenfassung und Ausblick .....</b>	<b>104</b>
<b>7</b>	<b>Literaturverzeichnis .....</b>	<b>107</b>
<b>8</b>	<b>Anhang .....</b>	<b>117</b>

## Abkürzungen

3D	dreidimensional
Å	Ångström
A	eine $\beta$ -Aminosäure mit Seitenkette an $C^\alpha$
AMBER	<i>assisted model building with energy refinement</i>
ARIA	<i>Ambiguous Restraints in Iterative Assignment</i>
B	eine $\beta$ -Aminosäure mit Seitenkette an $C^\beta$
$\beta$ -hGly	$\beta$ -Homoglycin, $H_2N-CH_2-CH_2-COOH$
CD	<i>circular dichroism</i>
CHARMM	<i>chemistry at Harvard molecular mechanics</i>
$C_i$	Wasserstoffbrücken-gebundener Ring mit $i$ Ringatomen
CNS	<i>chrystallography and NMR system</i>
COSY	<i>correlation spectroscopy</i>
CSI	<i>chemical shift index</i>
CVFF	<i>consistent valence forcefield</i>
$\delta$	Chemische Verschiebung
$\Delta$	Differenz
D	Debye
DG	Distanzgeometrie
DMSO	Dimethylsulfoxid
DNA	Desoxyribonucleinsäure
DQF	<i>double quantum filtered</i>
<i>E. coli</i>	Escherichia coli
E.COSY	<i>exclusive correlation spectroscopy</i>
F	Kraft
FKBP	FK bindendes Protein
fMD	Moleküldynamik ohne <i>restraints</i>
Fmoc	Fluorenylmethyloxycarbonyl-
f-SAA	eine furanoide Zuckeraminosäure (Abbildung 4-5)
$g(r)$	radiale Verteilungsfunktion
Gum	eine $\delta$ -Zuckeraminosäure (Abbildung 3-1)
$H_i$	Wasserstoffbrücken-gebundener Ring mit $i$ Ringatomen in einer Helix

---

HMBC	<i>heteronuclear multiple bond correlation</i>
HMQC	<i>heteronuclear multiple quantum correlation</i>
HSQC	<i>heteronuclear single quantum coherence</i>
J	Kopplungskonstante
K	Kraftkonstante
kDa	Kilo-Dalton
MD	Moleküldynamik
MO	Molekülorbital
MOLMOL	<i>molecule analysis and molecule display</i>
N	Atomanzahl
NMR	<i>nuclear magnetic resonance</i>
NOE	<i>nuclear OVERHAUSER enhancement</i>
NOESY	<i>nuclear OVERHAUSER enhancement spectroscopy</i>
NVT	Ensemble bei konstanter Teilchenzahl, Volumen und Temperatur
OG	Obergrenze
Par10	Parvulin 10 aus <i>Escherichia coli</i>
PASTA	<i>protein assignment by threshold accepting</i>
PBC	<i>periodic boundary conditions</i>
PPIase	Peptidyl-Prolyl- <i>cis/trans</i> -Isomerase
ppm	<i>parts per million</i>
<i>r</i>	(Atom-)Abstand
rMD	Moleküldynamik mit <i>restraints</i>
RMSD	<i>root mean square deviation</i>
ROESY	<i>rotating frame nuclear OVERHAUSER enhancement spectroscopy</i>
SA	<i>simulated annealing</i>
SAA	Zuckeraminosäure
TOCSY	<i>total correlation spectroscopy</i>
U	eine $\beta$ -Aminosäure ohne Seitenkette
[U- <sup>13</sup> C/ <sup>15</sup> N]	uniform <sup>13</sup> C und <sup>15</sup> N isotopenmarkiert
UG	Untergrenze
V	Potential
vdW	VAN DER WAALS
WW	Wechselwirkungs-
XAA	eine beliebige natürliche Aminosäure (in L- oder D-Form)

# 1 Einleitung

Die beiden wichtigsten Strategien in der modernen Arzneimittelentwicklung sind das rationale Design von Wirkstoffen und das Massen-*screening* von chemischen Bibliotheken. Während letzteres eher wahllos vorgeht, erfolgt beim rationalen Design eine äußerst zielgerichtete Suche nach neuen Wirkstoffen. Die Grundlage dazu bietet die molekulare Erkennung<sup>[1]</sup>, also die Fähigkeit von biologischen Ligand-Molekülen, zwischen strukturell und funktionell unterschiedlichen Rezeptoren zu unterscheiden und selektiv an ein Zielprotein zu binden. Durch das Verständnis der Vorgänge auf molekularer Ebene ist es daher möglich, gezielt in biochemische Prozesse einzugreifen und Pharmaka mit maßgeschneiderten Eigenschaften zu entwerfen.<sup>[2]</sup> Dazu ist die Aufklärung der räumlichen Strukturen von Rezeptor und Ligand sowie deren dynamischen Verhaltens erforderlich.

Ist im Idealfall die Struktur des Rezeptors bekannt, so kann durch Vergleich mit der Geometrie und den Ladungsverhältnissen in der Bindungstasche ein Ligand entworfen werden, der selektiv und mit hoher Affinität an den Rezeptor bindet (direktes Wirkstoffdesign)<sup>[3-5]</sup>. In vielen Fällen ist dessen Struktur allerdings nicht bzw. nur unzureichend bekannt. Mittels Struktur-Aktivitäts-Beziehungen, die sich durch Vergleich von Struktur und jeweiliger Aktivität möglichst vieler überprüfter Liganden ableiten lassen, können dann oftmals Rückschlüsse auf die Anforderungen des Rezeptors an den Liganden gezogen werden. Anhand der gewonnenen Kenntnisse kann anschließend eine Leitstruktur bezüglich Selektivität und Affinität, aber auch im Hinblick auf pharmakodynamische Eigenschaften wie Absorption, Verteilung und Metabolisierung optimiert werden.<sup>[2]</sup>

Diese Arbeit befasst sich sowohl mit Projekten des indirekten Wirkstoffdesigns als auch mit der Bestimmung einer Proteinstruktur als Grundlage zum direkten Wirkstoffdesign. In Kapitel 2 werden zunächst die verwendeten Methoden erläutert. Kapitel 3 beschreibt die dreidimensionale Struktur eines cyclischen Hexapeptid-Analogons, in das zwei Zuckeraminoacidsäuren eingebaut sind. Durch den Einbau von Zuckeraminoacidsäuren in Peptidmimetika ist es möglich, gezielt sowohl die Peptidrückgratstruktur als auch die Anordnung der Seitenketten zu beeinflussen und damit die Aktivität, Pharmakokinetik und metabolische Stabilität peptidischer Leitstrukturen zu verbessern. Cyclische Peptidmimetika dieser Art können als Gerüstmimetika Anwendung in der Templat-assoziierten Synthese von Proteinen (TASP) finden. Weiterhin wird in Kapitel 3 auf die Anwendung von Überlegungen



zum Vereinfachen cyclischer Hexapeptide auf Cyclohexan-Konformationen nach DUNITZ und WASER eingegangen. Kapitel 4 beschreibt die räumliche Struktur und Langzeit-Dynamiksimulation eines Hybridmoleküls aus Zuckeraminosäuren und  $\beta$ -Aminosäuren, sowie das Erstellen einer Lösungsmittelzelle für deuteriertes Chloroform, die für die Simulationen eingesetzt wurde. Derartige  $\beta$ -Peptide sind ebenfalls interessant als Template zur Verbesserung peptidischer Eigenschaften und bilden oftmals bemerkenswerte Strukturen aus. Kapitel 5 beschreibt die dreidimensionale Struktur des Proteins Par10 aus *Escherichia coli*. Par10 gehört zur Parvulin-Familie der Peptidyl-Prolyl-*cis/trans*-Isomerasen (PPIasen), die die Isomerisierung der Peptidbindung zwischen einer Aminosäure und Prolin katalysieren. Manche Vertreter dieser Familie, wie das in menschlichen Zellen vorkommende hPin1, spielen eine wichtige Rolle in pathologischen Prozessen wie Krebs oder ALZHEIMER. Das mit 10.1 kDa bisher kleinste Parvulin Par10 verkörpert den Urtyp der Parvuline und ist somit interessant, da es den Minimalsatz für die katalytische Domäne dieser Enzyme definiert.

## 2 Methoden der Strukturberechnung

Um Struktur und dynamisches Verhalten eines Moleküls mit Rechenverfahren zu untersuchen, bieten sich eine Reihe rein geometrischer sowie kraftfeldbasierter Ansätze an. Um dabei verwendete, experimentelle Daten aus NMR-Experimenten zu gewinnen, stehen weiterhin automatisierte Verfahren zur Verfügung. Die in dieser Arbeit eingesetzten Verfahren werden im Folgenden näher erläutert.

### 2.1 Distanzgeometrie

Die Distanzgeometrie (DG) ist ein mathematisches Verfahren zur Gewinnung räumlicher Strukturen aus Abständen.<sup>[6]</sup> Insbesondere bei der Strukturbestimmung in der Chemie hat sich die DG als rechenschnelle Methode zur Absuche des globalen Konformationsraums eines Moleküls bewährt. Dazu wird es oft mit kraftfeldbasierten Optimierungsverfahren kombiniert, wie im entsprechenden Kapitel näher erläutert wird. DG-Rechnungen werden in dieser Arbeit mit dem Programmpaket DISGEO<sup>[7-10]</sup> ausgeführt.

Das Verfahren der DG-Rechnung erfolgt in drei Schritten.<sup>[11-15]</sup> Zunächst werden Matrizen mit Abstandsobere- und unteren Grenzen aus der Topologie des Moleküls erzeugt, die durch experimentell ermittelte Abstände eingeschränkt werden können. Im zweiten Schritt, der Metrisierung (*metrization*), werden zufällig Abstände innerhalb der Ober- und Untergrenzen ausgesucht und in einer nun metrischen Matrix abgelegt. Der als Einbettung (*embedding*) bezeichnete letzte Schritt führt eine Projektion des Abstandsraums in den kartesischen Koordinatenraum durch. Die Einbettung ist der mathematische Schritt der DG, da hier die Konvertierung von „Distanz“ zur „Geometrie“ erfolgt.

#### 2.1.1 Abstandsmatrizen

Der Ausgangsschritt der DG-Rechnungen besteht in der Erzeugung einer holonomischen Abstandsmatrix, in der Unter- und Obergrenzen für alle Atomabstände im Molekül abgelegt sind. Diese holonomischen Abstände werden aus der Topologie einer Modellstruktur beliebiger Konformation ermittelt. Allerdings ist es wichtig, dass diese eine optimierte Bindungsgeometrie besitzt. Dies kann durch eine einfache Energieminimierung in einem der später beschriebenen Kraftfelder erreicht werden.

Die Unter- und Obergrenzen werden zunächst rein geometrisch unter der Annahme freier Drehbarkeit um vorher definierte Bindungen aus den exakten Abständen der Ausgangsstruktur bestimmt. Dabei wird meist eine zusätzliche Toleranz von üblicherweise  $\pm 1\%$  erlaubt. Die Summe der VAN DER WAALS-Radien eines Atompaares wird als Untergrenze verwendet, wenn eine ermittelte Untergrenze geringer als diese Summe ist.

Da die derartig bestimmten Abstandsbereiche in der Regel sehr weit sind, werden sie mit Hilfe der  $N$ -Eck-Ungleichungen weiter eingengt. Exemplarisch sei dazu die Anwendung der Dreiecksungleichung erläutert.

Die Summe der Längen zweier Seiten eines Dreiecks ist immer größer als die Länge der dritten Seite. Es gilt für drei Punkte  $A$ ,  $B$  und  $C$  im Raum:

$$|\vec{C} - \vec{A}| \leq |\vec{B} - \vec{A}| + |\vec{C} - \vec{B}| \quad [2-1].$$

Definiert man die Ober- und Untergrenzen als

$$OG_{AB} \geq |\vec{B} - \vec{A}| = d_{AB} \geq UG_{AB} \quad [2-2],$$

dann gilt:

$$d_{AC} \leq OG_{AB} + OG_{BC} \quad [2-3].$$

Es erfolgt die Zuweisung:

$$OG_{AC} := OG_{AB} + OG_{BC} \quad \forall \quad OG_{AC} > OG_{AB} + OG_{BC} \quad [2-4],$$

sonst

$$OG_{AC} \leq OG_{AB} + OG_{BC} \quad [2-5].$$

Anschließend können mittels trigonometrischer Beziehungen wie dem Kosinussatz,

$$|\vec{C} - \vec{A}|^2 = |\vec{C} - \vec{B}|^2 + |\vec{B} - \vec{A}|^2 - 2|\vec{C} - \vec{B}||\vec{B} - \vec{A}|\cos(\angle ABC) \quad [2-6],$$

die Abstandsbegrenzungen weiterhin eingeengt werden. An dieser Stelle können nun die experimentell ermittelten Daten integriert werden. Aus NMR-Experimenten gewonnene Abstände werden dazu in der Regel mit einer Toleranz von  $\pm 10\%$  versehen. Abstandsober- bzw. -untergrenzen eines Atompaares werden ersetzt, wenn dadurch eine weitere Einschränkung des bisher errechneten Abstandsintervalls erfolgt.

### 2.1.2 Erzeugung von Zufallsstrukturen (Metrisierung)

Der die Struktur erzeugende Schritt der DG ist die so genannte Metrisierung, die im DISGEO als Verfahren der *Zufallsmetrisierung* nach HAVEL implementiert ist.<sup>[8, 12]</sup> Dabei wird zunächst in der holonomischen Matrix ein Atompaar willkürlich gewählt. Diesem wird ein Abstandswert zugewiesen, der zufällig aus dem holonomischen Abstandsintervall gewählt wird. Mit Hilfe der Dreiecksungleichungen werden dann die holonomischen Grenzen der restlichen Abstände neu berechnet und eingeengt. Diese Prozedur wird solange wiederholt, bis sämtliche Atomabstände zugewiesen sind. Man gelangt so zu einer symmetrischen  $N \cdot N$ -Matrix ( $N$  = Atomanzahl) interatomarer Abstände.

An dieser Stelle sei darauf hingewiesen, dass durch die rein geometrische Einschränkung des Abstandsraumes sowie durch die rein zufällige Auswahl der Strukturen keine Einschränkung des dem Molekül zugänglichen Konformationsraums erfolgt (abgesehen von den experimentellen Einschränkungen).

### 2.1.3 Koordinatentransformation (Einbettung)

Durch das *embedding* werden die durch Metrisierung gewonnenen DG-Strukturen aus dem Abstandsraum in den Koordinatenraum transformiert (bzw. „eingebettet“). Dazu wird die Eigengleichung der symmetrischen  $N \cdot N$ -Matrix  $\hat{A}$ ,

$$\hat{A} \hat{X} = \hat{X} \hat{\Lambda} \quad [2-7],$$

gelöst. Die Dimensionalität des Koordinatenraumes kann beliebig gewählt werden, und es liegt nahe, die Transformation direkt in den dreidimensionalen Raum vorzunehmen.<sup>[16]</sup> Allerdings sind die durch *embedding* erzeugten Strukturen oft erheblich verzerrt, da die Wahl der Abstände innerhalb der Abstandsgrenzen rein zufällig geschieht. Eine anschließende Optimierung der Strukturen ist daher unumgänglich. Wie später beschrieben wird, erfolgen

die ersten Optimierungsschritte mittels eines rein Distanz-abhängigen Kraftfelds, das auch die Optimierung von Strukturen in höherdimensionalen Räumen erlaubt. Da durch Rotation und Translation in höhere Dimensionen Energiebarrieren in niederdimensionalen Räumen überwunden werden können, empfiehlt es sich somit, das *embedding* in einen Raum mit mehr als drei Dimensionen vorzunehmen.<sup>[17]</sup> Den Idealfall stellt ein Raum mit der Dimensionalität der Atomanzahl dar, da hier die Energiehyperfläche nur ein Minimum besitzt, das durch Minimierungsalgorithmen gefunden werden könnte. Andererseits führt die anschließend notwendige Reduktion in den dreidimensionalen Raum umso öfter zu physikalisch unmöglichen Strukturen, je höher die Dimensionalität des Ausgangsraums ist. Das in dieser Arbeit benutzte DISGEO-Programmpaket führt daher das *embedding* in den vierdimensionalen Raum aus.

## 2.2 Kraftfeldrechnungen

Die derzeit genaueste Methode zur Beschreibung von Struktur und Dynamik eines Moleküls stellt die Quantenmechanik zur Verfügung. Allerdings kann die zeitabhängige SCHRÖDINGER-Gleichung für größere Moleküle derzeit nur unter Verwendung von Näherungen und mit hohem zeitlichem Rechenaufwand durchgeführt werden. Daher ist für die Beschreibung größerer Systeme der auf der NEWTONSchen Mechanik beruhende, empirische Ansatz der Molekülmechanik<sup>[18-22]</sup> nach wie vor die Methode der Wahl. Atome werden dabei als Massepunkte behandelt, zwischen denen diverse Kräfte wirken, wie beispielsweise VAN DER WAALS-Wechselwirkungen oder elektrostatische Kräfte. Die Dynamik eines Moleküls entspricht der Bewegung der Atome auf einer Energiehyperfläche, und eine stabile Struktur entspricht einem Minimum dieser Energiefläche.

### 2.2.1 Das Kraftfeld

In der Molekülmechanik werden Energien durch ein Kraftfeld beschrieben. Ein Kraftfeld ist eine Sammlung von Parametern und Formeln, die sowohl atomare und interatomare Eigenschaften enthält, als auch die Potentiale auf einer Energiehyperfläche beschreibt. So kann ein Kraftfeld unter anderem beispielsweise Standard-Bindungslängen und Bindungswinkel, atomare Massen und Kraftkonstanten für Valenzschwingungen enthalten, sowie eine Gleichung, die diese Valenzschwingungen beschreibt. Die Parameter können entweder an quantenmechanisch berechneten, oder empirisch ermittelten Daten geeicht werden. Zum

Beispiel lassen sich Kraftkonstanten für Streck- und Valenzschwingungen aus Daten von Infrarotspektren entnehmen.

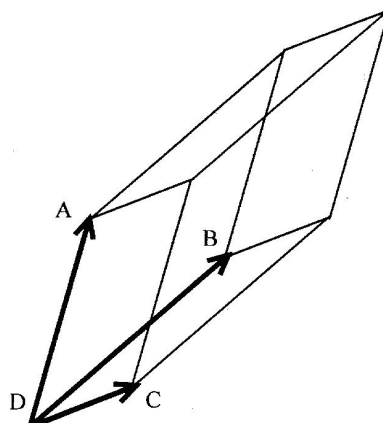
Die Optimierung eines Kraftfeldes ist ein multidimensionales Problem, da die Änderung eines Parameters alle weiteren Parameter beeinflusst. Außerdem muss ein Kraftfeld nach jedem Optimierungsschritt sorgfältig an Verbindungen bekannter Struktur getestet werden, was die Optimierung zu einem äußerst aufwendigen Unterfangen gestaltet. Daher ist ein Kraftfeld umso ungenauer, je weiter sein Anwendungsbereich ist. Hieraus erklärt sich die große Vielzahl der Kraftfelder, die heutzutage zur Anwendung kommen.

### 2.2.1.1 Distanzabhängiges Kraftfeld

Für die Optimierung der in der DG erzeugten Strukturen wird ein vom Atomabstand  $d_{AB}$  abhängiges Potential  $V_{\min}(r)$  verwendet, das sich aus zwei Haupttermen  $V_{Dist}$  und  $V_{Chi}$  zusammensetzt:

$$V_{\min}(r) = \underbrace{\sum_{d_{AB} > OG_{AB}} (d_{AB}^2 - OG_{AB}^2)^2 + \sum_{UG_{AB} > d_{AB}} (UG_{AB}^2 - d_{AB}^2)^2}_{V_{Dist}} + \underbrace{\sum_{chirals} (V_{ch} - V_{ch}^{ref})^2}_{V_{Chi}} \quad [2-8].$$

Wie bereits erwähnt, werden Strukturen in der DG aus einer holonomischen Matrix generiert, die Unter- und Obergrenzen für alle interatomaren Abstände enthält ( $UG$  und  $OG$  in Gleichung 2-8). Im harmonischen Potential  $V_{Dist}$  werden diese Abstandskriterien berücksichtigt. Der Term  $V_{Chi}$  gestattet es, vier Atome relativ zueinander im Raum zu fixieren. Werden drei Atome im Raum durch Vektoren mit einem vierten Atom verbunden, so spannen die Vektoren ein Spat auf, dessen Rauminhalt als chirales Volumen  $V_{ch}$  bezeichnet wird (Abbildung 2-1).



**Abbildung 2-1:** Zur Definition des chiralen Volumens  $V_{ch}$  verwendetes Spatprodukt.

Mittels des chiralen Volumens lässt sich die Chiralität an einem Stereozentrum fixieren, da bei Inversion am asymmetrischen Zentrum das Spatprodukt infolge Vektorvertauschung das Vorzeichen wechselt. Da das Volumen für planare Zentren Null ist, können auch planare Anordnungen wie die Amidbindung eben gehalten werden.

### 2.2.1.2 CVFF

Das *Consistent Valence Force Field*, CVFF,<sup>[23, 24]</sup> ist ein generisch für organische Moleküle entwickeltes Kraftfeld, das unter anderem für Aminosäuren bzw. Peptide parametrisiert ist. Die Potentialfunktion ist zusammengesetzt aus Bindungs- und Winkeltermen, sowie Termen für elektrostatische und VAN-DER-WAALS-Wechselwirkungen:

$$V_{CVFF} = V_{Bindung} + V_{Winkel} + V_{Im\ proper} + V_{Dieder} + V_{vdW} + V_{Coulomb} \quad [2-9].$$

Schwingungen einer Bindung werden dabei als Federschwingung um einen Nullpunkt  $b_0$  mit der Kraftkonstante  $K_b$  durch ein klassisches, harmonisches Potential beschrieben:

$$V_{Bindung} = \sum_{Bindungen} \frac{K_b (b - b_0)^2}{2} \quad [2-10].$$

Eine etwas genauere Beschreibung von Bindungsschwingungen kann durch ein MORSE-Potential erreicht werden, ist jedoch aufgrund der exponentiellen Abstandsabhängigkeit rechenintensiver (2-11).

$$V_{\text{Bindung}} = \sum_{\text{Bindungen}} \frac{K_b \left(1 - e^{-\alpha(b-b_0)}\right)^2}{2} \quad [2-11].$$

Ebenfalls durch ein harmonisches Potential werden Winkelschwingungen (2-12) und Pseudodiederwinkel-Deformationsschwingungen (2-13) behandelt:

$$V_{\text{Winkel}} = \sum_{\text{Winkel}} \frac{K_\theta (\theta - \theta_0)^2}{2} \quad [2-12],$$

$$V_{\text{Im proper}} = \sum_{\text{Im proper}} \frac{K(\xi - \xi_0)^2}{2} \quad [2-13].$$

Der Term 2-13 ist in seiner Wirkungsweise ähnlich dem Spatprodukt im distanzabhängigen Kraftfeld. Mit ihm können Chiralitäten und Bindungsplanaritäten konserviert werden. Das Potential  $V_{\text{Dieder}}$  beschreibt den periodischen Verlauf der Torsionsbarrieren bei Rotation um einen Diederwinkel:

$$V_{\text{Dieder}} = \sum_i F_{\varphi,i} [1 + \cos(n\varphi - \delta)] \quad [2-14].$$

VAN-DER-WAALSsche und elektrostatische Wechselwirkungen werden durch ein LENNARD-JONES- (2-15) bzw. durch ein COULOMB-Potential (2-16) beschrieben:

$$V_{\text{vdW}} = \sum_{i < j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad [2-15],$$

$$V_{\text{Coulomb}} = \sum_{i < j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad [2-16].$$



Um Kopplungen zwischen Schwingungen zu berücksichtigen, können im CVFF Kreuzterme eingesetzt werden, wie etwa der Term für die Kopplung der Schwingungen von benachbarten Bindungen:

$$V_{\text{Bindung, Bindung}} = \sum_{B_i} \sum_{B_j} F_{B_i B_j} (b1 - b1_0) (b2 - b2_0) \quad [2-17].$$

Die Verwendung von Kreuztermen erhöht allerdings drastisch die benötigte Rechenzeit und kann außerdem zu Instabilitäten<sup>[25]</sup> führen. In dieser Arbeit wurden daher keine Kreuzterme verwendet.

### 2.2.1.3 AMBER

Das AMBER-Kraftfeld (*Assisted Model Building with Energy Refinement*) wurde ursprünglich von KOLLMAN *et al.*<sup>[26,27]</sup> ausschließlich für Proteine und DNA parametrisiert. Durch etliche Modifikationen von zahlreichen Autoren ist auch AMBER mittlerweile für eine Vielzahl organischer Verbindungen, wie beispielsweise Polysaccharide,<sup>[28]</sup> einsetzbar. Das AMBER-Potential besitzt folgende Form:

$$V_{\text{AMBER}} = V_{\text{Bindung}} + V_{\text{Winkel}} + V_{\text{Im proper}} + V_{\text{vdW}} + V_{\text{Coulomb}} + V_{\text{H-Bonds}} \quad [2-18].$$

Die Bindungs- und Winkelterme werden auch im AMBER-Potential als klassisches, harmonisches Potential ausgedrückt (Gleichungen 2-10 und 2-12). Der Term für Pseudodiederwinkel-Deformationsschwingungen lautet:

$$V_{\text{Dieder}} = \sum_i F_{\varphi,i} [1 + \cos(n\varphi - \delta)] \quad [2-19].$$

Term 2-19 dient zum Fixieren von planaren Geometrien und Stereozentren. In der so genannten *united-atom*-Darstellung wird er benutzt, um die Chiralität und tetraedrische Natur von  $sp^3$ -Zentren zu korrigieren. Mit der *united-atom*-Darstellung kann erheblich Rechenzeit eingespart werden, indem unpolare Wasserstoffatome nicht explizit dargestellt werden, sondern in die Beschreibung der Schweratome eingebunden werden, an die sie gebunden sind. Wegen der damit verbundenen Ungenauigkeit und ständig steigenden Rechenkapazitäten rückt die *united-atom*-Darstellung heutzutage allerdings in den Hintergrund.

Ebenfalls analog den entsprechenden CVFF-Termen werden im ursprünglichen AMBER-Potential VAN-DER-WAALSSche und elektrostatische Wechselwirkungen beschrieben (Gleichungen 2-15 und 2-16). In den heutigen Versionen des in dieser Arbeit verwendeten Programms DISCOVER<sup>[25]</sup> wird der elektrostatische Term allerdings durch einen Term zur genaueren elektrostatischen Beschreibung von Wasserstoffbrücken ergänzt<sup>[27, 29]</sup>:

$$V_{H-Bonds} = \sum_{H-Bonds} \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \quad [2-20].$$

Der Wasserstoffbrücken-Term wurde aus zwei Gründen eingeführt. Einerseits ist bei starken H-Brücken eine repulsive Kraft nötig, um die Ausbildung von unrealistisch kurzen Brückenbindungsängen in Dynamiksimulationen zu vermeiden. Andererseits kann durch die 10-12-Form der Funktion eine Feinabstimmung der Länge und Energie einer Wasserstoffbrücke vorgenommen werden. Allerdings ändert Term 2-13 die Gesamtenergie einer H-Brücke nur um etwa  $0.5 \text{ kcal mol}^{-1}$ , so dass der Hauptanteil der Energie von der Dipol-Dipol-Wechselwirkung zwischen den Donor- und Akzeptorgruppen herrührt.

#### 2.2.1.4 CHARMM und PARALLHDG

Auf dem CHARMM-Kraftfeld<sup>[30]</sup> (*Chemistry at HARvard Molecular Mechanics*) basieren die Kraftfelder, die von den Programmen XPLOR<sup>[31]</sup> und CNS (Crystallography and NMR System)<sup>[32]</sup> verwendet werden. Bereits von Anfang an wurde es für eine weite Reihe von Verbindungen parametrisiert, wie beispielsweise Proteine, Nucleinsäuren, diverse prosthetische Gruppen und Substrate. Schon die ersten Versionen von CHARMM konnten sowohl isolierte Moleküle als auch Moleküle in Lösung oder in kristallinen Feststoffen behandeln. Das verwendete Potential ist in seiner Form sehr ähnlich dem CVFF-Potential:

$$V_{CHARMM} = V_{Bindung} + V_{Urey-Bradley} + V_{Winkel} + V_{Dieder} + V_{Im\ proper} + V_{VdW} + V_{Coulomb} \quad [2-21].$$

Während alle anderen Terme mit denselben Energiefunktionen wie im CVFF-Kraftfeld beschrieben werden, wird anstelle der im CVFF verwendeten Kreuzterme das UREY-BRADLEY-Potential  $V_{Urey-Bradley}$  eingeführt:

$$V_{Urey-Bradley} = \sum_{Urey-Bradley} K_{Urey-Bradley} (S - S_0)^2 \quad [2-22].$$

Das UREY-BRADLEY-Potential beschreibt (nichtbindende) sterische 1,3-Wechselwirkungen. Zusammen mit dem *improper*-Term dient es der besseren Anpassung an Daten, die aus Vibrationsspektren gewonnen wurden.

Neuere Versionen der Programme XPLOR und CNS verwenden TOPALLHDG, eine modifizierte Version von CHARMM. In der derzeitigen Version (TOPALLHDG5.2.PRO) werden nichtbindende Wechselwirkungen durch ein rein abstoßendes Potential  $V_{Repel}$  behandelt. Wie sich gezeigt hat<sup>[33]</sup>, führt die Verwendung der nichtbindenden Wechselwirkungen des ProLSQ *refinement* Programmes (*Protein Least Squares*; Gleichung 2-23)<sup>[34]</sup> zu besseren Ergebnissen gegenüber den in CHARMM gebräuchlichen Termen:

$$V_{Repel} = \begin{cases} 0 & , r > r_{min} \\ K_{vdW} (sr_{min} - r)^4 & , r \leq r_{min} \end{cases} \quad [2-23].$$

Dabei ist  $k_{vdW}$  die Energiekonstante,  $r$  der Abstand zwischen den Atomen,  $r_{min}$  die Summe der VAN DER WAALS-Radien der Atome, und  $s$  ein Skalierungsfaktor zur Feinabstimmung des Potentials.

### 2.2.1.5 Experimentelle Einschränkungen

Wie bereits erwähnt, ist die Suche nach dem globalen Energieminimum eines Moleküls in der Regel ein multidimensionales Problem. Beispielsweise besitzt ein Protein mit  $n$  Aminosäuren, unter der Annahme starrer Peptidbindungen,  $2n$  variable Rückgratwinkel. Für ein Protein mit 100 Aminosäuren entsteht dadurch ein mindestens 200-dimensionaler Konformationsraum. Auch die statistische Suche nach dem globalen Minimum in einem solchen Raum würde alle heutzutage verfügbaren Rechenkapazitäten sprengen. Daher ist die Einschränkung des Konformationsraums mit experimentell gewonnenen Strukturparametern unbedingt erforderlich. Hierzu wird das Kraftfeldpotential durch Strafterme erweitert.

Mittels der NMR-Spektroskopie lassen sich hauptsächlich Informationen über Atomabstände, Diederwinkel und die räumliche Orientierung von Substrukturen gewinnen. Abstandsinformationen werden dabei durch Ausnutzung des nuklearen OVERHAUSER-Effekts (NOE) erhalten.<sup>[35]</sup> Durch Anwendung der isolierten Zwei-Spin-Näherung (*isolated spin pair approximation*, ISPA)<sup>[36]</sup> kann aus den Volumina  $V_{ij}$  der Kreuzsignale eines NOESY- bzw. ROESY-Spektrums der Abstand  $r_{ij}$  errechnet werden:

$$r_{ij} = r_{ref} \left( \frac{V_{ref}}{V_{ij}} \right)^{\frac{1}{6}} \quad [2-24].$$

Zur Kalibrierung können ein oder mehrere bekannte Abstände  $r_{ref}$  verwendet werden, für die die entsprechenden Kreuzsignale im Spektrum gut aufgelöst und nicht überlagert sind. Geeignet ist beispielsweise der Abstand geminaler Protonen (1.78 Å). Aber auch feste Abstände in bekannten Strukturelementen, wie etwa  $\beta$ -Schleifen, können zur Eichung verwendet werden, wenn Sicherheit über deren Starrheit gegeben ist. Um aus den gemessenen interatomaren Abständen Einschränkungen für die Rechnungen zu gewinnen, werden sie mit einer Messtoleranz von üblicherweise 10 bis 30 % versehen. Die entstehenden Ober- und Untergrenzen werden dann meist als Grenzpunkte  $r_2$  und  $r_3$  in einem parabolischen *flat-bottom*-Potential folgender Form eingesetzt:

$$V_{dist} = \begin{cases} E_1 + F_1(r_1 - r_{ij}) & , r_{ij} < r_1 \\ K_2(r_{ij} - r_2)^2 & , r_1 < r_{ij} \leq r_2 \\ 0 & , r_2 < r_{ij} \leq r_3 \\ K_3(r_{ij} - r_3)^2 & , r_3 < r_{ij} \leq r_4 \\ E_4 + F_4(r_{ij} - r_4) & , r_4 < r_{ij} \end{cases} \quad [2-25].$$

Die linearen Bereiche am Rand des Potentials dienen dabei zur Vermeidung unerwünscht hoher Kräfte, die insbesondere bei fehlerhaften NOE-Zuweisungen sowie bei Hochtemperatur-Dynamiksimulationen auftreten können. Oftmals wird Gleichung 2-25 noch um einen weiteren Randbereich ergänzt, in dem die wirkende Kraft konstant auf einem Maximalwert gehalten wird.

Diederwinkel zwischen skalar koppelnden Kernen stehen nach KARPLUS in Beziehung zur  $^3J$ -Kopplungskonstante<sup>[37, 38]</sup>:

$$^3J = A \cos^2 \theta + B \cos \theta + C \quad [2-26].$$

Die Parameter A, B und C können an jeweils ähnlichen Systemen bekannter Struktur geeicht werden.<sup>[39-41]</sup> Allerdings besitzt die Funktion oftmals mehrere Lösungen für eine gemessene Kopplungskonstante. Daher wird während der Kraftfeldrechnung die Kopplungskonstante aus dem jeweiligen Diederwinkel berechnet und die Abweichung vom gemessenen Wert als harmonischer Strafterm berücksichtigt. Diesen Term benutzt auch die modifizierte Version von DISGEO<sup>[42]</sup>, die in dieser Arbeit verwendet wird:

$$V_J = K_J \left( J_{\text{exp}} - J_{\text{theo}} \right)^2 \quad [2-27].$$

Neben dem direkten Einsatz in Rechnungen können gemessene Strukturparameter auch zur Qualitätskontrolle einer berechneten Struktur bzw. eines Strukturensambles dienen. Beispielsweise liefert der CSI (*chemical shift index*) Informationen über die Sekundärstruktur eines Proteins.<sup>[43, 44]</sup> Eine qualitative Aussage über den Solvensaustausch von Amidprotonen kann entweder mit Hilfe des MEXICO-Experiments<sup>[45]</sup> getroffen werden, oder indem die Abhängigkeit der  $H^N$ -Verschiebung von der Temperatur gemessen wird.<sup>[46, 47]</sup> Weiterhin stehen heutzutage eine Reihe von Verfahren zur Verfügung, um die Dynamik eines Proteins direkt zu messen.

### 2.2.2 Berechnung weitreichender Wechselwirkungen

Insbesondere bei Kraftfeldrechnungen größerer Moleküle ist es nötig, einen Kompromiss zwischen erzielter Rechengenauigkeit und dem erforderlichen Zeitaufwand zu finden. Die Berechnung bindender Wechselwirkungen wie beispielsweise Bindungs- und Winkelschwingungen spielt dabei nur eine untergeordnete Rolle, da deren Anzahl linear mit der Anzahl der Atome steigt. Im Gegensatz dazu wächst die Anzahl VAN DER WAALSScher und elektrostatischer Wechselwirkungen, und somit auch der Rechenaufwand, quadratisch mit der Atomanzahl. Mit zunehmender Reichweite klingen diese Kräfte allerdings immer mehr ab. Rechenzeit kann daher eingespart werden, indem auf die Berechnung derartiger Wechselwirkungen ab einem gewählten Atomabstand verzichtet wird. Die Verwendung

derartiger Abbruchradien zur Berechnung weitreichender Wechselwirkungen wird zwangsläufig erforderlich, wenn Lösungsmittelleffekte mit dem Modell periodischer Randbedingungen berücksichtigt werden sollen. Dieses soll daher im folgenden Abschnitt näher erläutert werden.

### 2.2.3 Periodische Randbedingungen

Das Lösungsmittel, das ein gelöstes Molekül umgibt, hat meist einen entscheidenden Einfluss auf dessen Struktur. Beispielsweise begünstigt Trifluorethanol als Lösungsmittel für Proteine weit mehr die Ausbildung  $\alpha$ -helicaler Sekundärstrukturelemente, als Wasser dies tut.<sup>[48]</sup> Eine Simulation, die dem Solvens nicht Rechnung trägt, ist im Allgemeinen sehr ungenau, da Systeme im Vakuum eine ausgeprägte Neigung zur Minimierung ihrer Oberfläche besitzen. Daher wurden in der Vergangenheit verschiedene Verfahren zur Berücksichtigung von Lösungsmittelleffekten entwickelt.

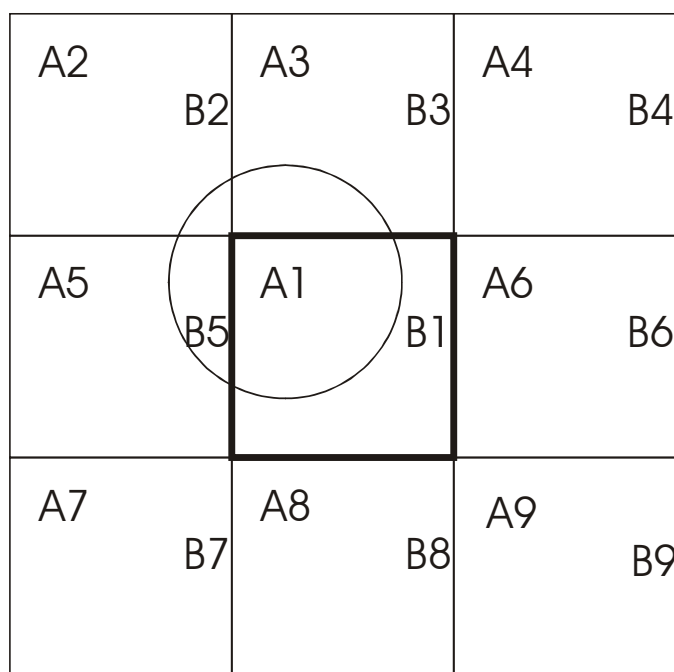
Lösungsmittel können entweder implizit durch die Dielektrizitätskonstante des Solvens oder explizit als Mehrteilchensystem behandelt werden. Simulationen im expliziten Lösungsmittel erfordern zwar den höheren Rechenaufwand, sind aber wegen der größeren Genauigkeit insbesondere für kleinere Moleküle die Methode der Wahl. Dabei müssen allerdings Vorkehrungen getroffen werden, um ein Auseinanderdriften des Molekülverbandes zu verhindern, was dem Verdunsten des Lösungsmittels entsprechen würde. Dies kann beispielsweise geschehen, indem die Zelle mit dem solvatisierten Molekül in eine Schale räumlich fixierter Lösungsmittelmoleküle eingebettet wird.<sup>[20]</sup> Eine andere Möglichkeit besteht darin, das Kraftfeld-Potential um Terme zu erweitern, die eine Rückstellkraft zur Solvenzelle hin bewirken. Alle diese Verfahren können allerdings zu unrealistischen Strukturen führen, wenn das zu berechnende Molekül nahe an den Rand der Solvenzelle und damit in den Einflussbereich des künstlich eingeführten Potentials kommt. Auch das Verfahren stochastischer Randbedingungen<sup>[49]</sup>, bei dem das Potential räumlich dem Solvatmolekül angepasst wird, kann hier nur partiell Abhilfe schaffen.

Die in dieser Arbeit angewandte Methode der periodischen Randbedingungen<sup>[19,20]</sup> vermeidet dahingegen grundsätzlich durch Randeffekte entstehende Artefakte. Hierzu wird zunächst eine Solvenzelle gewählt, deren Geometrie so beschaffen ist, dass mit ihr und ihren Abbildern ein unendlicher, dreidimensionaler Raum vollständig und lückenlos ausgefüllt werden kann. Hierzu eignen sich beispielsweise der Quader oder der abgekappte Oktaeder<sup>[50]</sup>.

Rechenaufwand kann dabei eingespart werden, indem eine Zelle gewählt wird, die der Geometrie des zu berechnenden Moleküls ähnelt.

Die Solvenszelle wird dann mit identischen Abbildern ihrer selbst umgeben. Bei jedem Simulationsschritt werden nun die Atome der Abbildzellen verwendet, um Kräfte zu berechnen, die auf die „echten“ Atome der zentralen Zelle wirken. Wechselwirkungen zwischen den Abbildatomen müssen dabei nicht berechnet werden, denn deren Koordinaten und Geschwindigkeiten können nach jedem Schritt durch Symmetrieoperationen mit den echten Atomen neu berechnet werden. Für die zentrale Zelle stellt dieses Verfahren eine wesentlich bessere Approximation einer großen Menge von umgebendem Lösungsmittel dar, da Moleküle an ihrer Oberfläche mit Solvensmolekülen in benachbarten Zellen wechselwirken können.

Die Methode der periodischen Randbedingungen offeriert zwei grundsätzliche Varianten, die sich erheblich in ihrer Rechengenauigkeit und ihrem zeitlichen Aufwand unterscheiden. Die einfache Variante wird als *minimum-image-Modell* bezeichnet.<sup>[20, 25]</sup> Dabei wechselwirkt jedes Atom der zentralen Zelle jeweils mit nur einem Abbild eines jeden Atoms, und zwar mit demjenigen, das ihm räumlich am nächsten ist. Abbildung 2-2 stellt diesen Sachverhalt, vereinfacht auf den zweidimensionalen Raum, dar.



**Abbildung 2-2:** Im *minimum-image-Modell* interagiert jedes Atom mit genau einem Abbild der anderen Atome.

Molekül A1, in der zentralen Zelle, ist umgeben vom realen Atom B1 und mehreren Abbildern von Atom B. Berechnet wird allerdings nicht die Wechselwirkung A1-B1, sondern nur diejenige zwischen A1 und dem Abbild B5, da dieses näher zu Atom A1 ist. Der Vorteil des Verfahrens liegt also in der hohen Rechengeschwindigkeit, da zwischen einem gegebenen Atompaar jeweils nur eine Wechselwirkung berechnet werden muss. Allerdings setzt dieses Vorgehen einen Abbruchradius voraus, außerhalb dessen keinerlei Wechselwirkungen zwischen zwei Atomen berechnet werden. Im *minimum-image*-Modell darf dieser nicht größer als die Hälfte der kürzesten Zelldimension sein. Dies führt zu einem Rechengenauigkeitsverlust, der nur mit dem Einsatz sehr großer Lösungsmittelzellen umgangen werden kann.

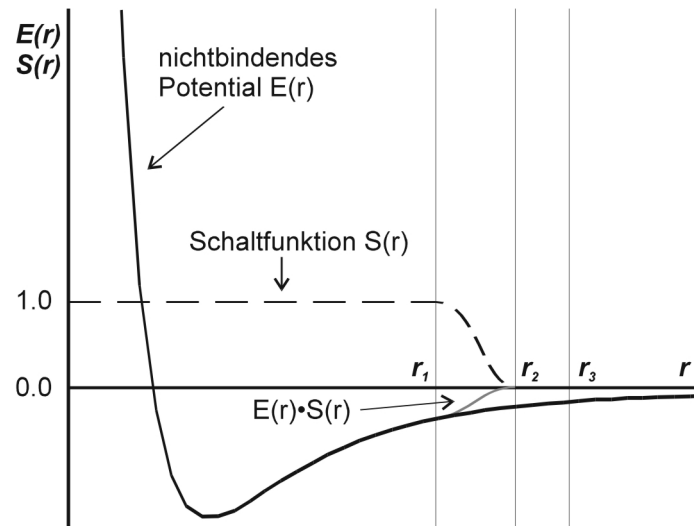
Das in dieser Arbeit angewendete *explicit-image*-Modell<sup>[25]</sup> erfordert ebenfalls die Verwendung eines Abbruchradius, der jedoch keinerlei Beschränkungen unterliegt. Hierbei kann ein Atom mit mehreren Abbildern seiner selbst und mit Abbildern anderer Atome wechselwirken. Da sich die zur Berechnung verwendeten Abbilder beim Ein- und Austritt der Zelle ändern können, wird das Führen einer Nachbarschaftsliste von Abbildatomen nötig. Dies führt zu einem Verlust an Rechengeschwindigkeit, der jedoch insbesondere bei der Berechnung kleinerer Moleküle durch die höhere Genauigkeit der Simulation aufgewogen wird.

#### 2.2.4 Abbruchradien

Wie bereits erwähnt, verursachen weitreichende, nichtbindende Wechselwirkungen einen Großteil des Rechenaufwands von Kraftfeldrechnungen. Das in dieser Arbeit verwendete DISCOVER-Programmpaket<sup>[25]</sup> stellt zu ihrer Behandlung verschiedene Verfahren zur Verfügung. Beim direkten Vergleich der Methoden schneidet allerdings derzeit die Verwendung von Abbruchradien<sup>[51]</sup> bezüglich Rechengenauigkeit und erzielter Rechenzeiterparnis am besten ab<sup>[52]</sup>.

Im einfachsten Fall wird das entsprechende nichtbindende Potential ab einem Atomabstand, der größer als der Abbruchradius ist, auf null gesetzt. Die dabei entstehende Unstetigkeit des Potentialverlaufs kann allerdings zu sprunghaften Änderungen von Atomgeschwindigkeiten und damit zu Strukturartefakten führen. Daher ist es üblich, die Wechselwirkungen durch Multiplikation mit einer Schalfunktion  $S(r)$  über einen gewählten Bereich langsam abklingen zu lassen. Abbildung 2-3 stellt den dadurch entstehenden Kurvenverlauf für VAN DER WAALSSche Wechselwirkungen dar.<sup>[25]</sup>

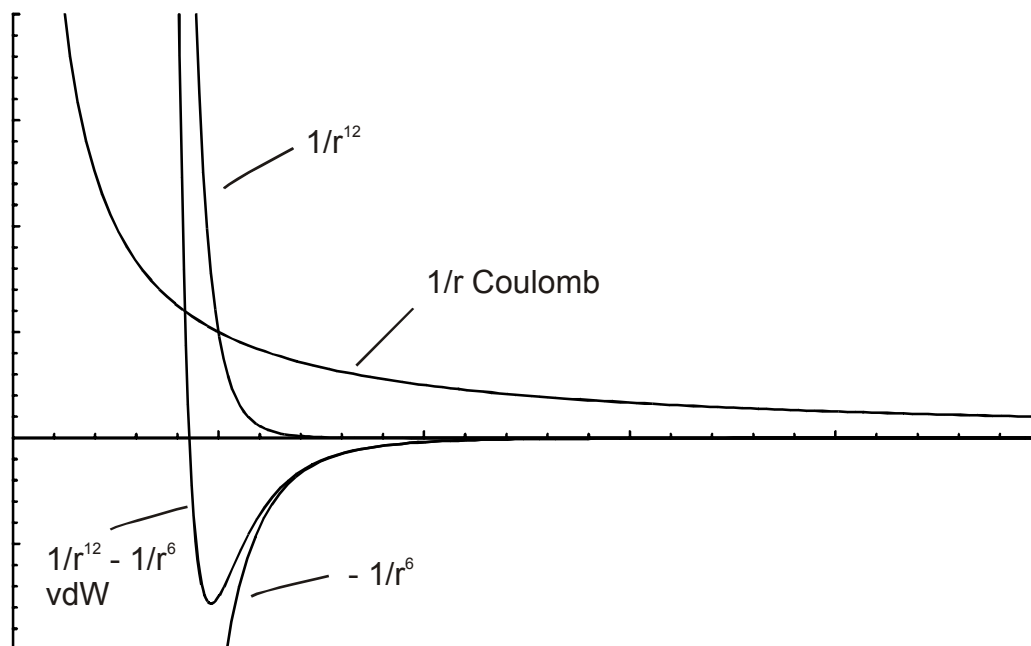




**Abbildung 2-3:** Allmähliches Abschalten des nichtbindenden Potentials.

In DISCOVER wird als Schaltfunktion ein Polynom fünften Grades verwendet, für den die Ableitung an den Rändern des Schaltbereichs (zwischen  $r_1$  und  $r_2$ ) null ist. Um weiterhin Rechenzeit einzusparen, verwaltet DISCOVER eine Nachbarschaftsliste von Atompaaaren, für die nichtbindende Wechselwirkungen berechnet werden müssen. Diese Liste wird allerdings nicht bei jedem Simulationsschritt neu erstellt. Sie enthält daher auch Atompaaare, die bis zur nächsten Erneuerung der Liste für die Energieberechnung relevant werden könnten. Diese Atompaaare werden aus einem Pufferbereich (zwischen  $r_2$  und  $r_3$  in Abbildung 2-3) gewählt. Um sicherzustellen, dass alle erforderlichen Wechselwirkungen berechnet werden, wird die Liste immer dann neu erstellt, wenn sich ein Atom mehr als die Hälfte der Pufferweite bewegt hat.

Da VAN DER WAALSSche Wechselwirkungen relativ schnell mit zunehmender Entfernung abklingen ( $V_{vdW} \propto 1/r^6$ ), können sie bereits in relativ geringen Distanzen (ca.  $10 \text{ \AA}$ ) abgeschaltet werden. Im Gegensatz dazu sind elektrostatische Wechselwirkungen wesentlich weitreichender ( $V_{Coulomb} \propto 1/r$ ) und können auch bei großen Entfernungen nicht vernachlässigt werden (Abbildung 2-4).



**Abbildung 2-4:** Nichtbindende Wechselwirkungen: VAN DER WAALS- und COULOMB-Potential.

Tatsächlich jedoch bestehen Moleküle oft zum größten Teil aus neutralen Fragmenten. Die resultierenden, elektrostatischen Wechselwirkungen sind somit keine COULOMBSchen Wechselwirkungen, sondern solche zwischen Dipolen bzw. Quadrupolen. Da diese bei zunehmender Entfernung mit  $1/r^3$  abklingen, können auch für sie Abbruchradien verwendet werden. Dabei besteht jedoch die Gefahr, dass Dipole durch den Abbruchradius in Monopole getrennt werden, was zur Vernachlässigung einer starken COULOMBSchen Wechselwirkung führen würde. DISCOVER vereint daher neutrale Fragmente zu so genannten Ladungsgruppen.<sup>[20, 25]</sup> Für jede dieser Ladungsgruppen wird ein so genanntes Schaltatom definiert, das möglichst zentral im Schwerpunkt der Gruppe liegen sollte. Sind sich zwei Schaltatome näher als der Abbruchradius, werden sämtliche Wechselwirkungen zwischen Atomen der beiden Ladungsgruppen berechnet, die in der Nachbarschaftsliste aufgeführt sind, andernfalls werden keine Wechselwirkungen berechnet.

### 2.2.5 Methoden zur Konformationsraumsuche

Viele Moleküle, wie Peptide oder Proteine, liegen in Lösung in mehreren Konformationen vor. Die einzelnen Konformationen liegen dabei zueinander in einem Verhältnis vor, das einer BOLTZMANN-Verteilung der potentiellen Energie bei der Temperatur des Systems entspricht. Gewöhnlich sind allerdings die höchstpopulierten Strukturen von Interesse. Das Ziel einer

kraftfeldbasierten Rechnung ist also primär, das globale Minimum bzw. gegebenenfalls Nebenkonformationen entsprechende, lokale Minima der Energiehyperfläche zu finden. Da in höherdimensionalen Räumen hierfür kein analytisches Verfahren existiert, müssen bereits für kleine Moleküle statistische Methoden angewendet werden. Die in dieser Arbeit verwendeten Verfahren werden im Folgenden erläutert.

### 2.2.5.1 Minimierungsalgorithmen

Um von einem beliebigen Punkt auf der Energiehyperfläche ausgehend das nächstgelegene Minimum aufzufinden, werden Minimierungen<sup>[53]</sup> eingesetzt.

In den Minima heben sich alle Kräfte, die auf ein jedes Atom wirken, in der Summe auf. Dabei besteht die Möglichkeit der Einführung künstlicher Kräfte in Form von *restraints*. Neben dem offensichtlichen Ziel, stabile Molekülkonformationen zu finden, bieten Minimierungen daher beim *molecular modeling* eine Reihe weiterer Vorteile. So ist es beispielsweise möglich, die Atome eines Moleküls auf eine Templatstruktur zu zwingen. Dadurch kann die erforderliche Energie bestimmt werden, um eine gewünschte Konformation, beispielsweise zum Andocken an einen Rezeptor, zu erzeugen.

Zur Minimierung eines Moleküls wird zunächst eine Zielfunktion erzeugt, die die Energie in Abhängigkeit der Atomkoordinaten beschreibt. Anschließend wird die Konformation so verändert, dass sich der Wert der Zielfunktion verringert. Eine Minimierung ist umso effizienter, je schneller dieser Wert berechnet wird, und je weniger Iterationen benötigt werden, um schließlich zum Minimum zu konvergieren.

Allgemein besteht ein Minimierungsalgorithmus aus zwei Komponenten. Der erste Teil dient dem Festlegen einer Suchrichtung auf der Energiefläche. Im zweiten, als Liniensuche (*line search*) bezeichneten Teil findet eine Suche nach dem Minimum einer eindimensionalen Funktion statt, die durch den Richtungsvektor vorgegeben wird. Erst hier erfolgt die eigentliche Änderung der Atomkoordinaten in Richtung einer Struktur niedrigerer Energie. Die eindimensionale Suche wird iterativ durchgeführt, indem vom Ausgangspunkt aus zunächst die Energie zweier Punkte in entgegengesetzter Richtung berechnet wird. Damit steht die Richtung zum Minimum fest, das in einer Reihe weiterer Iterationen gefunden werden kann. Der Vorteil des Verfahrens liegt in seiner Einfachheit. Allerdings sind Liniensuchen rechenaufwendig, da mit jeder Iteration die Zielfunktion neu evaluiert werden muss.

Um die Suchrichtung auf der Energiefläche festzulegen, kann im einfachsten Fall der lokale Gradient berechnet werden:

$$\bar{\mathbf{h}}_i = -\frac{dV(\bar{\mathbf{r}}_i)}{d\bar{\mathbf{r}}_i} \quad [2-28].$$

Die Liniensuche erfolgt somit in Richtung des größten Abstiegs, weswegen das Verfahren als *steepest-descent*-Minimierung<sup>[54]</sup> bezeichnet wird. Da ausschließlich Gradienten zur Anwendung kommen, ist das Verfahren sehr robust, es produziert also auch dann Strukturen niedrigerer Energie, wenn das Potential sehr anharmonisch ist.<sup>[25, 53]</sup> Andererseits besitzt es auch einen Nachteil, der besonders bei quadratischem, harmonischem Potentialverlauf zum Tragen kommt, wie beispielsweise bei MORSE-Potentialen nahe dem Gleichgewichtsabstand. In diesem Fall verläuft jeder Gradient, ausgehend vom Endpunkt einer Liniensuche, senkrecht zur vorherigen Linie. Dies führt dazu, dass der Suchweg oszilliert und somit die Konvergenz nahe dem Minimum zunehmend langsamer wird. Die *steepest-descent*-Minimierung wird daher meist zur Entzerrung stark gespannter Strukturen verwendet.<sup>[25, 53]</sup>

Wesentlich schneller zur Konvergenz führt das so genannte *conjugate-gradients*-Verfahren.<sup>[25, 53, 55]</sup> Hierbei werden die Oszillationen gedämpft, indem bei jeder Richtungs-suche die Richtung des vorherigen Minimierungsschritts mit einem Wichtungsfaktor  $\gamma$  mit berücksichtigt wird:

$$\bar{\mathbf{h}}_i = \left( -\frac{dV(\bar{\mathbf{r}}_i)}{d\bar{\mathbf{r}}_i} \right) + \gamma_{i-1} \bar{\mathbf{h}}_{i-1} \quad [2-29].$$

Durch Ausnutzung der in der zweiten Ableitung enthaltenen Information kann bei harmonischen Potentialen eine noch schnellere Konvergenz erreicht werden. Hiervon macht beispielsweise die NEWTON-RAPHSON-Methode Gebrauch. Jedoch steigt dabei der Rechenaufwand quadratisch mit der Atomanzahl. Da derartige Algorithmen außerdem oftmals instabil bei stark anharmonischen Potentialen sind, werden sie nur zur Minimierung kleinerer Moleküle nahe eines Minimums eingesetzt.<sup>[25]</sup>

Wird in dieser Arbeit von einer Minimierung bis zur Konvergenz gesprochen, so ist damit gemeint, dass die Minimierung abgebrochen wird, wenn nach einem Minimierungsschritt die maximale Ableitung der Energiefunktion in den einzelnen Dimensionen kleiner als  $0.001 \text{ kcal } \text{Å}^{-1}$  ist.

### 2.2.5.2 Moleküldynamik

Die im vorangegangenen Kapitel beschriebenen Minimierungen erlauben es, von einem beliebigen Punkt auf einer Energiefläche das nächstgelegene Minimum zu finden. Von Interesse sind jedoch die zeitliche Entwicklung eines Moleküls, sowie die höchstpopulierten, also tiefsten Minima. Um diese zu finden und zu untersuchen, stehen eine Reihe von Methoden zur Verfügung, die allgemein als moleküldynamische Rechenverfahren bezeichnet werden.

In der Moleküldynamik wird die Bewegungsenergie der Atome bei einer bestimmten Temperatur berücksichtigt. Durch Umwandlung von kinetischer Energie und potentieller Energie wird es dem System ermöglicht, Potentialbarrieren zu überwinden.

#### 2.2.5.2.1 Grundlagen

Im einfachsten Fall werden in der Moleküldynamik die Newtonschen Bewegungsgleichungen als Funktion der Ortskoordinaten  $\vec{\mathbf{r}}$  gelöst:

$$-\frac{dV(\vec{\mathbf{r}}_i(t))}{d\vec{\mathbf{r}}_i(t)} = [\vec{\mathbf{F}}_i(t) = m_i \vec{\mathbf{a}}_i(t) =] m_i \frac{d^2 \vec{\mathbf{r}}_i(t)}{dt^2} \quad [2-30],$$

wobei  $\vec{\mathbf{F}}_i$  die Kraft,  $m_i$  die Masse, und  $\vec{\mathbf{a}}_i$  die Beschleunigung eines Atoms  $i$  sind. Da die klassischen Bewegungsgleichungen deterministisch sind, können bei bekanntem Anfangszustand des Systems alle Atompositionen und Geschwindigkeiten zu jedem beliebigen, späteren Zeitpunkt ermittelt werden. Ein gesamter Satz an Koordinaten und Geschwindigkeiten, der schrittweise für einen bestimmten Zeitraum berechnet wurde, wird in der Moleküldynamik als Trajektorie bezeichnet.

Allerdings ist Gleichung 2-30 eine Differentialgleichung zweiter Ordnung, für die eine genaue, analytische Lösung wegen der Komplexität der eingesetzten Kraftfeldpotentiale nicht möglich ist. Sie muss daher auf numerischem Weg gelöst werden. Eine Möglichkeit dazu besteht darin, von einem Satz an Koordinaten und Geschwindigkeiten zu einem Zeitpunkt  $t$  ausgehend, den entsprechenden Satz zur Zeit  $t + \Delta t$  zu berechnen. Mit diesem kann dann der nächste Satz nach Ablauf eines weiteren Zeitschritts  $\Delta t$  berechnet werden. Um einen geeigneten Zeitschritt zu wählen, muss ein Kompromiss zwischen effizienter Nutzung der Rechenleistung und erwünschter Rechengenauigkeit gefunden werden, da zu große Zeitschritte zu Instabilitäten und Ungenauigkeiten bei der Integration führen.

Weil die Energieberechnung in Bezug auf die Rechenleistung sehr aufwendig ist, stellen sich eine Reihe von Anforderungen an einen guten Algorithmus zur Integration. Beispielsweise ist es von Vorteil, wenn ein Integrator schnell arbeitet, möglichst wenig Speicher benötigt, und die Verwendung langer Zeitschritte  $\Delta t$  erlaubt. Weiterhin ist es wichtig, dass die Erhaltung der Energie gewährleistet ist. Ein Algorithmus, der diese Kriterien gut erfüllt, ist der in der alten Version von DISCOVER (2.9.8) implementierte VERLET-*leapfrog*-Integrator.<sup>[25, 56]</sup> Die grundsätzliche Annahme der VERLET-Algorithmen<sup>[57]</sup> besteht darin, dass sich Geschwindigkeiten und Beschleunigungen während eines Zeitschritts nicht ändern. Im *leapfrog*-Algorithmus berechnen sich diese zu:

$$\bar{\mathbf{v}}_i(t + \frac{1}{2}\Delta t) = \bar{\mathbf{v}}_i(t - \frac{1}{2}\Delta t) + \Delta t \bar{\mathbf{a}}_i(t) \quad [2-31],$$

$$\bar{\mathbf{r}}_i(t + \Delta t) = \bar{\mathbf{r}}_i(t) + \Delta t \bar{\mathbf{v}}_i(t + \frac{1}{2}\Delta t) \quad [2-32],$$

$$\bar{\mathbf{a}}_i(t + \Delta t) = \frac{f(t + \Delta t)}{m_i} \quad ; \quad f(t + \Delta t) = \frac{-\partial V(\bar{\mathbf{r}}_i(t + \Delta t))}{\partial \mathbf{r}} \quad [2-33].$$

Wie alle Varianten des VERLET-Algorithmus besitzt die Methode den Vorteil, dass nur eine Auswertung der Energiefunktion pro Zeitschritt durchgeführt werden muss. Allerdings verlaufen die Atomkoordinaten und Geschwindigkeiten beim *leapfrog*-Integrator um einen halben Zeitschritt asynchron. Die dadurch entstehenden numerischen Instabilitäten bei der Energieberechnung treten bei dem in der neuen Version von DISCOVER (95.0) implementierten *velocity*-VERLET-Algorithmus<sup>[58]</sup> nicht auf.<sup>[19]</sup>

$$\bar{\mathbf{v}}_i(t + \Delta t) = \bar{\mathbf{v}}_i(t) + \frac{1}{2}\Delta t [\bar{\mathbf{a}}_i(t) + \bar{\mathbf{a}}_i(t + \Delta t)] \quad [2-34],$$

$$\bar{\mathbf{r}}_i(t + \Delta t) = \bar{\mathbf{r}}_i(t) + \Delta t \bar{\mathbf{v}}_i(t) + \frac{1}{2}\Delta t^2 \bar{\mathbf{a}}_i(t) \quad [2-35],$$

$$\bar{\mathbf{a}}_i(t + \Delta t) = \frac{f(t + \Delta t)}{m_i} \quad [2-36].$$

Um der Forderung nach konstanten Geschwindigkeiten und Beschleunigungen während eines Zeitschritts möglichst gerecht zu werden, ist es nötig, jede Bewegungsperiode im System in eine hinreichende Anzahl von Segmenten zu unterteilen. Die Wahl des Zeitschritts  $\Delta t$  wird daher durch die schnellste Bewegung im Molekül limitiert. Für die meisten Moleküle ist dies die C-H-Streckschwingung mit einer Periode von ca.  $10^{-14}$  s. Bei einer Unterteilung der

Schwingung in 10 Segmente ergibt sich damit für die VERLET-Algorithmen ein maximaler Zeitschritt von einer Femtosekunde.

Beim Beginn einer Simulation sind aus der Startstruktur die Ortskoordinaten  $\vec{r}_i(t_0)$  der Atome und die auf sie wirkenden Anfangsbeschleunigungen  $\vec{a}_i(t_0) = -1/m_i \cdot (dV(\vec{r}_i(t_0))/d\vec{r}_i)$  bekannt. Als Randbedingung zum Lösen der klassischen Bewegungsgleichung 2-30 werden jedoch auch die Anfangsgeschwindigkeiten  $\vec{v}_i(t_0)$  benötigt. Mittels eines Randomisierungsprozesses können diese so gewählt werden, dass sie gemäß der im thermodynamischen Gleichgewicht vorliegenden MAXWELL-BOLTZMANN-Verteilung der Geschwindigkeiten verteilt sind.

$$f(|\vec{v}|)d|\vec{v}| = \left(\frac{m}{2\pi k_b T}\right)^{\frac{3}{2}} e^{-\frac{m|\vec{v}|^2}{2k_b T}} 4\pi|\vec{v}|^2 d|\vec{v}| \quad [2-37].$$

Die kartesischen Geschwindigkeitskomponenten  $v_x$ ,  $v_y$  und  $v_z$  werden dabei von DISCOVER so zugewiesen, dass sie der naturgemäß vorliegenden GAUSS-Verteilung entsprechen:

$$g(|\vec{v}_x|)d|\vec{v}_x| = \left(\frac{m}{2\pi k_b T}\right)^{\frac{1}{2}} e^{-\frac{m|\vec{v}_x|^2}{2k_b T}} d|\vec{v}_x| \quad [2-38].$$

### 2.2.5.2.2 Methoden zur Temperaturkontrolle

Die Temperatur eines Systems bleibt während einer Moleküldynamik-Simulation keineswegs konstant, sondern ändert sich vielmehr aufgrund mehrerer Faktoren. Beispielsweise ist für eine manuell aufgebaute Startstruktur, zu Beginn einer Simulation, die Temperatur konsistent mit der MAXWELL-BOLTZMANN-Verteilung der Atomgeschwindigkeiten. Da die Struktur aber kaum dem thermodynamischen Gleichgewicht entsprechen wird, wird in ihr ein mehr oder weniger großer Anteil potentieller Energie enthalten sein, der sich recht schnell in einer Änderung der kinetischen Energie bemerkbar machen wird. Ebenso arbeiten die beschriebenen Integratoren niemals exakt, was sich in einer ständigen Fluktuation der Systemtemperatur äußert. Ohne geeignete Kontrollmechanismen ist es daher nicht möglich, ein kanonisches NVT-Ensemble, also ein System, das den gewöhnlich bei der Vermessung von Biomolekülen vorherrschenden Bedingungen entspricht, korrekt zu reproduzieren.

Die Temperatur ist eine makroskopische Größe, die gemäß der statistischen Thermodynamik mit der mikroskopischen Beschreibung durch die kinetische Energie  $E_{kin}$  eines Systems

verknüpft ist. Sie kann berechnet werden, wenn die Geschwindigkeitsvektoren aller Atome bekannt sind:

$$E_{kin} = \left\langle \sum_{i=1}^N \frac{m_i \vec{v}_i^2}{2} \right\rangle = \frac{N_f k_b T}{2} \quad [2-39].$$

Die Anzahl der Freiheitsgrade  $N_f$  ist dabei in der Regel  $3N-6$  in nichtperiodischen, sowie  $3N-3$  in periodischen Systemen.

Um die korrekte Temperatur aufrecht zu erhalten, müssen also die Atomgeschwindigkeiten entsprechend angepasst werden. Allerdings muss ein Regelmechanismus nicht nur die geforderte Temperatur  $T_{Ziel}$  einstellen, sondern auch das korrekte statistische Ensemble gemäß der MAXWELL-BOLTZMANN-Verteilung erzeugen. Eine einfache, jedoch drastische Methode ist die direkte Skalierung der Atomgeschwindigkeiten:<sup>[25]</sup>

$$\vec{v}_{i,skal} = \vec{v}_i \cdot \sqrt{\frac{T_{Ziel}}{T_{System}}} \quad [2-40].$$

Die Skalierung der Atomgeschwindigkeiten wird meist zur Äquilibration eines Systems eingesetzt. Während der eigentlichen Dynamiksimulation wird oftmals die BERENDSEN-Temperaturbadkopplung eingesetzt, die eine sanftere Temperatureinstellung erlaubt.<sup>[59]</sup> Die Atomgeschwindigkeiten werden dabei mit dem Faktor  $\lambda$  multipliziert:

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau} (T_{System} - T_{Ziel})} \quad [2-41].$$

Dabei ist  $\Delta t$  der Zeitschritt der Integration. Mit der Relaxationszeit  $\tau$  ist es möglich, den Grad der Kopplung an das Temperaturbad einzustellen.

Eine verbesserte Approximation eines NVT-Ensembles liefert die ANDERSEN-Methode zur Temperaturkontrolle, bei der Atomgeschwindigkeiten gemäß der BOLTZMANN-Verteilung randomisiert werden.<sup>[60]</sup> Hierzu kann auf zwei unterschiedliche Weisen vorgegangen werden. Einerseits können entweder alle Atomgeschwindigkeiten in einem regelmäßigen Intervall angepasst werden, oder nach jedem Zeitschritt jeweils die Geschwindigkeit eines einzelnen



Atoms. Bildlich gesehen entspricht diese Kopplung an das Temperaturbad einer Abfolge von stochastischen Stößen auf die Atome.

### 2.2.5.3 *Simulated Annealing*

*Simulated Annealing* (SA)<sup>[61-63]</sup> ist eines der am häufigsten angewendeten Verfahren zur globalen Absuche des Konformationsraums. Die Vorgehensweise entspricht dem Verfahren zur Erzeugung von perfekt gewachsenen Einkristallen durch äußerst langsames Abkühlen einer Schmelze.

Durch Aufheizen auf sehr hohe Temperaturen (beispielsweise 1000 K) wird es dem System zunächst ermöglicht, auch sehr hohe Potentialbarrieren zu überwinden, und sich somit frei auf der Energiefläche zu bewegen. Die Startstruktur kann dabei in einer beliebigen Konformation mit optimierter kovalenter Bindungsgeometrie vorliegen. Anschließend an die Hochtemperaturphase wird dann schrittweise abgekühlt, wobei nach jedem Schritt eine Äquilibrationsphase gesetzt wird. Bei hinreichend langsamer Abkühlung erfolgt somit eine Annäherung an die thermodynamisch stabilste Struktur, also an das globale Minimum der Energiefläche. Um das Überqueren von Potentialbarrieren in der Hochtemperaturphase zu erleichtern, werden oftmals VAN DER WAALS-Radien und Kraftkonstanten, ausgehend von geringen Werten, langsam hochskaliert.

Um die Unabhängigkeit des Verfahrens von der Startstruktur zu gewährleisten, wird das Verfahren meist mehrfach wiederholt, wobei die Atomgeschwindigkeiten für jeden Durchlauf neu randomisiert werden. Die erzeugten Strukturen können dann nach verschiedenen Kriterien, wie beispielsweise ihre Energie oder die Erfüllung experimenteller Daten, auf ihre Güte geprüft werden.

In dieser Arbeit wurde SA in Form eines Protokolls nach M. NILGES<sup>[62, 64]</sup> mit dem Programm X-PLOR<sup>[31]</sup> eingesetzt.

### 2.2.5.4 *Time Averaged Restraints-Verfahren*

Wie bereits in Kapitel 2.2.1.5 erläutert wurde, können Kraftfeldpotentiale durch zusätzliche Strafterme zur Einhaltung von Distanzen und Kopplungskonstanten erweitert werden, die aus NMR-Experimenten gewonnen wurden. Allerdings liegen gelöste Moleküle oft in mehr als einer Konformation vor. Erfolgt die Umwandlung der Konformationen ineinander langsam auf der NMR-Zeitskala, so besitzt auch jede Konformation ihren eigenen Signalsatz, der getrennt von anderen Konformationen zur Strukturberechnung verwendet werden kann.

Probleme treten jedoch dann auf, wenn ein schneller Konformationsaustausch stattfindet. Die NOESY- bzw. ROESY-Signale liegen dann als  $\langle r^{-6} \rangle^{-1/6}$  Mittel<sup>[65]</sup> der Distanzen der einzelnen Konformationen vor. Nur in den wenigsten solcher Fälle gelingt es, die einzelnen Signalbeiträge von Hand den jeweiligen Konformationen zuzuweisen.

Um dennoch aus gemittelten Signalen eines einzigen Datensatzes zu einem Ensemble von Strukturen zu gelangen, welches in seiner Gesamtheit die gemessenen Daten reproduziert, wurden in der Vergangenheit einige Methoden entwickelt. Beispielsweise berechnen Ensemble-Verfahren<sup>[66, 67]</sup> die Dynamik mehrerer Moleküle simultan. Nach jedem Dynamikschritt werden dann die aktuellen Werte der als *restraints* eingesetzten Parameter für jede Struktur berechnet und über das Ensemble gemittelt. Aus der Abweichung von den geforderten Werten kann dann das Strafpotential berechnet und auf jede Struktur im Ensemble angewendet werden. Einzelnen Strukturen ist es auf diese Weise erlaubt, die experimentellen Daten mehr oder minder stark zu verletzen, solange das Ensemble alle *restraints* erfüllt. Allerdings haben Ensemble-Verfahren den Nachteil eines oft immensen Speicher- und Rechenbedarfs. Abhängig von der Größe des simulierten Systems können daher meist nur recht einfache Kraftfelder bei kurzen Simulationszeiten eingesetzt werden.

Komplementär zu den Ensemble-Verfahren, die versuchen, alle *restraints* im Mittel über ein Ensemble zu erfüllen, streben *time-averaged-distance-restraints*-Verfahren<sup>[68-71]</sup> deren Einhaltung im Mittel über die Zeit der Simulation an. Der parabolische Teil des *flat-bottom*-Strafpotentials (Gleichung 2-25) wird dabei aus dem mittleren Abstand  $\langle r \rangle$  aller zum momentanen Zeitpunkt in der Trajektorie enthaltenen Strukturen berechnet:

$$V_{dist} = K(\langle r_{ij} \rangle - r_{exp})^2 \quad [2-42].$$

Bei zunehmender Simulationsdauer nimmt allerdings die aktuelle Konformation immer weniger Einfluss auf den arithmetischen Mittelwert. Das System würde daher im Laufe der Zeit immer träger auf Verletzungen der geforderten Abstandswerte reagieren, die somit erst nach unendlich langer Simulationszeit erfüllt werden könnten. Daher wird zur Berechnung des mittleren Abstands eine Funktion verwendet, in der zeitlich zurückliegende Konformationen mit zunehmendem Alter  $t$  exponentiell abfallend gewichtet werden.

$$\langle r \rangle = \left( \frac{1}{\tau} \int_0^t \exp\left(-\frac{t-t'}{\tau}\right) [r(t-t')]^{-3} dt' \right)^{\frac{1}{3}} \quad [2-43].$$

Die Abklingkonstante  $\tau$  regelt die Stärke des exponentiellen Abfalls und entspricht dem Gedächtnis des Systems an vergangene Konformationen.

Sowohl  $\tau$  als auch die Kraftkonstante  $K$  haben einen entscheidenden Einfluss auf den Verlauf der Simulation.<sup>[71]</sup> Wie sich herausgestellt hat, nimmt die Abklingkonstante fast keinen Einfluss auf die Übereinstimmung mit den experimentellen Daten, umso mehr jedoch auf die Dynamik des Systems. Der Wert  $\tau = 0$  Sekunden entspricht einer Simulation ohne zeitlich gemittelte *restraints*. Ein Wechsel zwischen einzelnen Konformationen ist dann nicht möglich, und das Molekül nimmt eine gemittelte Konformation an. Bereits kleine Werte von  $\tau$  erlauben die Ausbildung energetisch günstiger Konformationen, die sich allerdings in künstlicher Weise sehr schnell periodisch ineinander umwandeln. Erst bei hohen Werten, im Bereich der langsamsten Bewegung des Moleküls entspricht die Dynamik dem natürlichen Verlauf (bei Proteinen:  $\tau$  ca. 20 ps und länger). Bei einer großen, realistischen Verweildauer in den energetisch günstigen Zuständen ist andererseits eine lange Simulationsdauer erforderlich, um zu einer brauchbaren Statistik zu gelangen. Idealerweise wird daher ein möglichst großes  $\tau$  bei einer um mindestens eine Größenordnung darüber liegenden Simulationsdauer gewählt.

Mit steigender Kraftkonstante  $K$  sinken sowohl die Summe der Verletzungen der experimentellen Daten, als auch die insgesamt größte Einzelverletzung. Hohe Kraftkonstanten führen jedoch nicht nur in Richtung einer gemittelten Struktur, sondern erhöhen auch die potentielle Energie des Systems. Insbesondere bei kleinen Werten von  $\tau$  führt dies zur Erhöhung und zu starken Fluktuationen der Temperatur. Da die Fluktuationen obendrein noch unterschiedlich stark in verschiedenen Teilen des Moleküls sind, befindet sich das System dann keinesfalls im Gleichgewicht. Die Kraftkonstante sollte daher höchstens so groß gewählt werden, dass die experimentellen Daten erfüllt werden können.

### 2.2.6 Automatisches, iteratives NOE-Assignment: ARIA

Um zu sinnvollen Ergebnissen zu gelangen, benötigen die bereits beschriebenen Methoden zur Berechnung von Strukturen im Allgemeinen eine ausreichende Anzahl von Abstandseinschränkungen, die unter Ausnutzung des nuklearen OVERHAUSER-Effekts gewonnen wurden. Dabei ist meist eine mehr oder weniger eindeutige Zuordnung der Signale

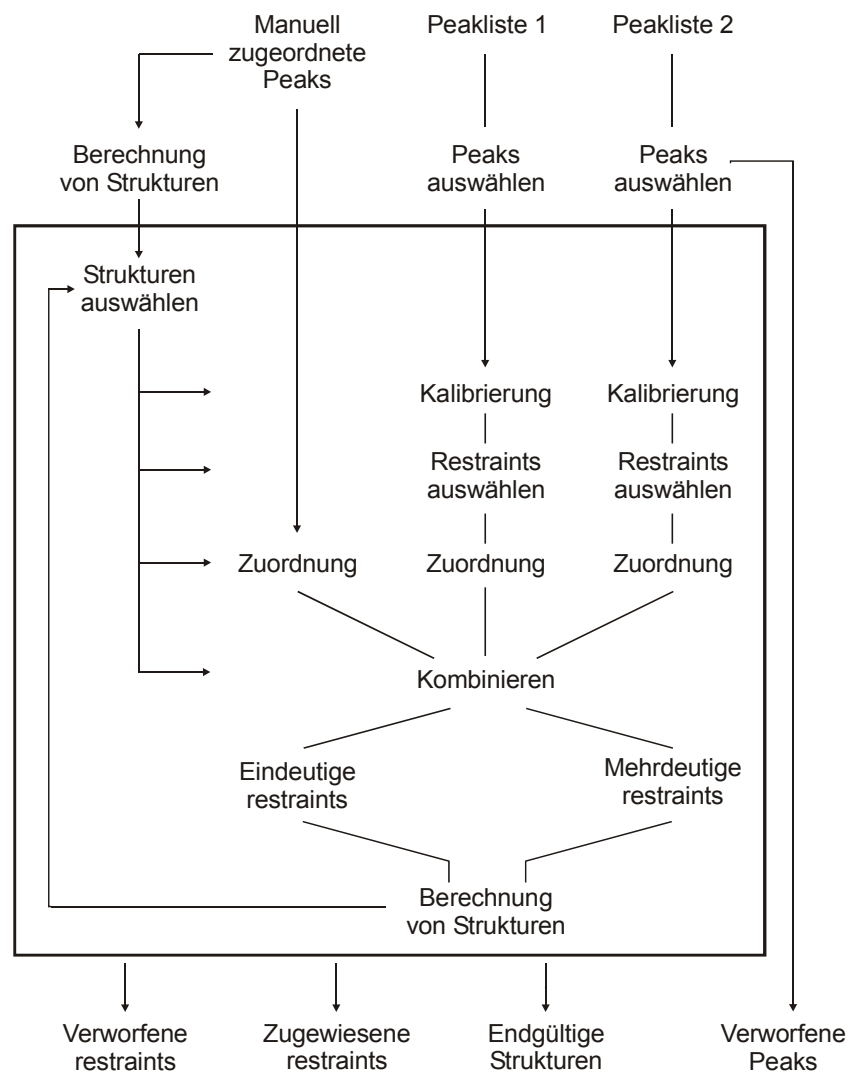
zu einzelnen Atompaaen erforderlich. In der Praxis haben verschiedene Protonen allerdings oftmals die gleiche chemische Verschiebung. Die eindeutige Zuweisung gelingt meist recht einfach, wenn nur wenige kovalente Bindungen zwischen zwei Protonen liegen, oder wenn sich die Protonen aufgrund der Sekundärstruktur in räumlicher Nähe zueinander befinden. Viele Signale können jedoch erst dann zugeordnet werden, wenn bereits eine gewisse Kenntnis der Struktur vorliegt. Mittels einiger, sicher und eindeutig zugeordneter Signale werden daher meist vorläufige Strukturen berechnet, anhand derer dann weitere Signale eindeutig zugeordnet werden können. Mit den daraus gewonnenen, neuen Abstandsbeschränkungen werden anschließend weiter verfeinerte Strukturen erzeugt und das Verfahren solange wiederholt, bis alle Signale, soweit möglich, eindeutig zugewiesen sind.

Bei diesem iterativen Verfahren der Strukturberechnung können allerdings einige Probleme auftreten. Beispielsweise kann es vorkommen, dass für die Bestimmung der Struktur eines Proteins durchaus genügend Signale vorhanden sind, um die Tertiärstruktur eindeutig festzulegen. Wenn jedoch die anfängliche Anzahl eindeutiger Signalzuweisungen nicht ausreicht, um vorläufige Strukturen mit konvergierender Faltung zu berechnen, wird eine weitere, iterative Verfeinerung naturgemäß unmöglich. Auch im ersten Iterationsschritt kann das Verfahren zum Stillstand kommen. So kann die Distanz zwischen zwei Protonen in den berechneten Strukturen weit mehr als 5 Å betragen, obwohl in der Realität ein NOE-Signal vorliegt. Wenn also zwar die richtige Faltung gefunden wurde, aber viele Seitenketten in allen Strukturen des berechneten Strukturensambles in falschen Konformationen vorliegen, kann die weitere Signalzuweisung unmöglich werden. Das Verfahren ist somit stark abhängig von der Vollständigkeit der Suche im Konformationsraum.

Während derartige Schwierigkeiten meist durch eine Verbesserung der experimentellen und rechnerischen Methoden behoben werden können, besteht jedoch eine grundlegende Gefahr in der iterativen Strukturberechnung. Wenn die vorläufigen Strukturen eine nicht der Realität entsprechende Faltung einnehmen, stellt sich der Fehler im Allgemeinen recht schnell in einer der anschließenden Iterationen heraus, da viele Signale dann unerklärbar bleiben. Es kann jedoch vorkommen, dass ein anfänglicher NOE-Datensatz nicht nur auf die korrekte Struktur zugewiesen werden kann, sondern mit nur wenigen unerklärbaren Signalen auch auf weitere Strukturen.<sup>[72]</sup> Ist in einem solchen Fall die Faltung der initialen Strukturen falsch, führen alle weiteren, iterativen Zuweisungen zwangsläufig zu einer Konformation, die nicht der Realität entspricht.

Zur Vermeidung dieser Gefahr wurde daher 1993 von M. NILGES<sup>[72, 73]</sup> ein Konzept vorgeschlagen, das später automatisiert im Programm ARIA<sup>[33, 74-80]</sup> implementiert wurde.

Wie üblich werden dabei die aus NOE-Spektren erhaltenen Distanzen in eine Zielfunktion eingesetzt. Während beim konventionellen Verfahren nach jeder Iteration neue, vormals mehrdeutige NOEs in die *restraint*-Liste aufgenommen werden, erlaubt es die Zielfunktion jedoch, von vornherein alle Zuweisungsmöglichkeiten anzugeben. Mittels eines speziellen *simulated-annealing*-Verfahrens wird dann die Zielfunktion minimiert. Abbildung 2-5 zeigt ein Flussdiagramm<sup>[74]</sup> der einzelnen Arbeitsschritte von ARIA, das im Folgenden kurz erläutert wird, und dessen Details in Unterkapiteln beschrieben werden:



**Abbildung 2-5:** Arbeitsschritte von ARIA.

Das Verfahren geht zunächst von einer manuell erstellten Liste von *restraints* aus, die zwar mehrdeutige Abstände enthalten darf, jedoch fehlerfrei sein muss. Als sicher zuweisbare Kontakte kommen dabei vor allem sequentielle, und Sekundärstruktur-NOEs in Betracht.

Mittels dieser Liste wird nun, durch Aufruf des Programms CNS<sup>[32]</sup>, ein initiales Ensemble von Strukturen berechnet. ARIA benötigt dann für jedes Spektrum eine Liste von Kreuzsignalen und ihren Integralen, die mittels üblicher, automatischer *peak-picking*-Verfahren gewonnen werden kann. Weiterhin muss eine möglichst vollständige Liste sämtlicher chemischen Verschiebungen vorhanden sein. Für alle Kreuzsignale, die einer Kombination vorhandener Verschiebungen entsprechen, werden zwei neue Listen von *restraints* für jedes Spektrum erstellt: Eine mehrdeutige, und eine eindeutige. Kreuzsignale, die keiner bekannten Verschiebung entsprechen, werden protokolliert. Abschließend werden alle *restraints* aus der manuell erzeugten Liste in die beiden neuen Listen überführt.

Zu diesem Zeitpunkt beginnt die eigentliche, erste Iteration, deren Verlauf sich für alle weiteren Iterationen wiederholt. Aus den Strukturen der vorangegangenen Iteration (bzw. im Falle der ersten Iteration, aus den initialen Strukturen) wird zunächst, durch Bewertung der Gesamtenergie, ein Ensemble ausgewählt. Ein erster Schritt ist die Einengung der Zuweisungsmöglichkeiten der mehrdeutigen *restraints*. Für jede Zuweisungsmöglichkeit wird dazu der entsprechende Abstand über das Ensemble gemittelt. Durch  $r^{-6}$  Summierung (siehe Kapitel 2.2.6.1) der Zuweisungsmöglichkeiten ergibt sich ein Gesamtabstand für den *restraint*. Trägt eine Zuweisungsmöglichkeit nur geringfügig zum Gesamtabstand bei, so wird sie entfernt (siehe Kapitel 2.2.6.2).

Nun werden ausgewählte Distanzen über das Ensemble gemittelt und als Referenzdistanzen zur Kalibrierung der Spektren verwendet. Mit den Integralen der Kreuzsignalliste werden damit Abstände berechnet, aus denen Ober- und Untergrenzen für die *restraint*-Listen erzeugt werden. Eine separate Routine führt jeweils alle eindeutigen und alle mehrdeutigen *restraint*-Listen der einzelnen Spektren zusammen (*merging*). Wird ein *restraint* kontinuierlich während mehrerer Iterationen verletzt, so wird er in einer Datei protokolliert und aus den Listen entfernt. Unter Verwendung der beiden *restraint*-Listen wird anschließend mit CNS ein neuer Satz von Strukturen berechnet und die nächste Iteration gestartet.

### 2.2.6.1 $r^{-6}$ summierte Distanzen

Um mehrdeutige *restraints* verwenden zu können, muss eine (weitgehend der Realität entsprechende) Summierung über alle Zuweisungsmöglichkeiten durchgeführt werden, die im Folgenden erläutert werden soll.

Das Integral eines mehrdeutigen NOE Kreuzsignals enthält Anteile aller Protonenpaare  $H_iH_j$ , deren Verschiebungen innerhalb eines Frequenzfensters ( $\{F1, \Delta1\}; \{F2, \Delta2\}$ ) liegen:

$$NOE_{F1,F2} = \sum_{i \in \{F1,\Delta1\}, j \in \{F2,\Delta2\}} NOE_{ij} \quad [2-44],$$

wobei  $\Delta 1$  und  $\Delta 2$  jeweils die Hälfte der Ausmaße des Signals in den entsprechenden Dimensionen betragen. Mit der Näherung isolierter Spinpaare (*Isolated Spin Pair Approximation*, ISPA)<sup>[36]</sup> ist bei gegebener Mischzeit  $\tau_m$  jeder Anteil am Integral eines Kreuzsignals, bzw. der NOE-Aufbaurrate, allein abhängig von der Distanz  $r_{ij}$ :

$$\frac{d}{d\tau_m} NOE_{ij} = c r_{ij}^{-6} \quad [2-45].$$

Unter der Annahme identischer Ordnungsparameter und Korrelationszeiten aller beteiligten Spinpaare gilt für das mehrdeutige Kreuzsignal:

$$\frac{d}{d\tau_m} NOE_{F1,F2} = c \left( \sum_{i \in \{F1,\Delta1\}, j \in \{F2,\Delta2\}} r_{ij}^{-6} \right) \quad [2-46].$$

Definiert man die „ $r^{-6}$  summierte“ Distanz  $\bar{r}$  als

$$\bar{r} = \left( \sum_{i \in \{F1,\Delta1\}, j \in \{F2,\Delta2\}} r_{ij}^{-6} \right)^{-1/6} \quad [2-47],$$

so vereinfacht sich Gleichung 2-46 zu

$$\frac{d}{d\tau_m} NOE_{F1,F2} = c \bar{r}_{F1,F2}^{-6} \quad [2-48].$$

Die  $r^{-6}$  summierte Distanz  $\bar{r}$  verhält sich ähnlich wie das  $\langle r^{-6} \rangle^{-1/6}$  Mittel der Distanzen<sup>[65]</sup>:

$$\langle r^{-6} \rangle^{-1/6} = \left( \sum_{i \in \{F1,\Delta1\}, j \in \{F2,\Delta2\}} \frac{r_{ij}^{-6}}{n} \right)^{-1/6} = n^{1/6} \left( \sum_{i \in \{F1,\Delta1\}, j \in \{F2,\Delta2\}} r_{ij}^{-6} \right)^{-1/6} = n^{1/6} \bar{r} \quad [2-49],$$

wobei  $n$  die Anzahl der Spinpaare in der Summe ist. Allerdings liegt das Mittel immer zwischen den beteiligten Atomen, während  $\bar{r}$  grundsätzlich kleiner als der kürzeste beteiligte Abstand ist. Bei Strukturberechnungen ist das Mittel daher wesentlich restriktiver. Die Summierung der Distanzen erlaubt viel eher räumliche Anordnungen von Atomen, in denen effektiv nur ein Abstand zum gemessenen Signal beiträgt. Bei einem hohen Grad von Mehrdeutigkeit ist somit die Summierung die Methode der Wahl, während die Mittelung bei bekannten, annähernd gleich starken Signalkomponenten (beispielsweise ein intraresidualer NOE mit einer Methylgruppe) mehr der Realität entspricht. Dabei ist die Anwendung der Summierung nicht auf Kontakte zwischen Protonen beschränkt. Beispielsweise lassen sich mehrdeutige *restraints* mittels Summierung auch auf Disulfidbrücken anwenden.

Die Definition der  $r^{-6}$  summierten Distanz erlaubt eine automatische Zuweisung von NOE-Abständen. Dazu werden die chemischen Verschiebungen der Kreuzsignale, die beispielsweise mittels einer automatischen *peak-picking* Routine gewonnen werden können, zunächst mit einer Toleranz  $\Delta_i$  versehen. Die Toleranz wird dabei im Allgemeinen entsprechend der Messgenauigkeit des Spektrums in den einzelnen Dimensionen  $i$  gewählt. Unter der erwähnten Voraussetzung einer möglichst vollständigen Liste aller Protonen-Verschiebungen können dann jeweils alle möglichen Protonenpaare, deren Verschiebungen in eines der dabei entstehenden Frequenzfenster fallen, zu einem mehrdeutigen *restraint* zusammengefasst werden.

### 2.2.6.2 Entfernen von Zuweisungsmöglichkeiten

Zuweisungsmöglichkeiten eines mehrdeutigen *restraints*, die nur geringfügig zur Gesamtdistanz beitragen, müssen im Verlaufe der einzelnen Iterationen entfernt werden. Im einfachsten Fall könnte dazu ein Schwellwert definiert werden: Ein Beitrag wird dann entfernt, wenn die entsprechende Distanz in allen Strukturen größer als der Schwellwert ist. Genauer ist es jedoch, die relativen Größen der Beiträge zueinander zu bewerten. In ARIA ist dies wie folgt implementiert.

Für jede Zuweisungsmöglichkeit  $k$  eines mehrdeutigen *restraints* wird die kleinste Distanz im Ensemble,  $D_{\min}^k$ , gesucht. Ihr Beitrag  $C^k$  zur Gesamtdistanz wird dann errechnet als

$$C^k = \frac{D_{\min}^k{}^{-6}}{\sum_i^{N(F1,F2)} D_{\min}^i{}^{-6}} \quad [2-50].$$



Die Beiträge  $C^k$  werden dann nach ihrer Größe sortiert und nur die  $N_p$  größten Beiträge beibehalten, so dass

$$\sum_i^{N_p} C^i > p \quad [2-51].$$

Der Parameter  $p$  kann dabei individuell gesetzt werden. Als Voreinstellung variiert ARIA  $p$  von 0.999 in der ersten, bis zu 0.8 in der letzten Iteration. Für den Fall eines *restraints* mit nur zwei Kontributionen vermittelt Tabelle 2-1 einen Eindruck, wie sich  $p$  auswirkt. Beispielsweise wird für  $p = 0.999$ , wenn die kürzere Distanz 2.5 Å ist, die zweite Komponente entfernt, wenn sie mehr als 7.9 Å beträgt.

**Tabelle 2-1:** Einfluss des Parameters  $p$  auf das Entfernen von Zuweisungsmöglichkeiten eines mehrdeutigen *restraints* mit zwei Komponenten. Distanzen, die länger als der angegebene Wert sind, werden als Zuweisungsmöglichkeit verworfen (alle Abstandswerte in Å).

Kürzere Distanz [Å]	Parameter $p$		
	0.999	0.95	0.8
2.5	7.9	4.1	3.3
4.4	12.6	6.6	5.2

Verbleibt nach dieser Prozedur nur noch eine einzige Zuweisungsmöglichkeit für den *restraint*, wird er von der Tabelle mehrdeutiger *restraints* in die Tabelle eindeutiger *restraints* überführt.

Der Einfachheit halber macht ARIA keinen Unterschied zwischen Mehrdeutigkeiten, die in einer einzelnen Struktur des Ensembles bestehen, und solchen, die zwischen verschiedenen Strukturen im Ensemble bestehen.

### 2.2.7 Sortieren von Strukturen in Familien

Oft stellt sich das Problem, eine sehr große Anzahl von Konformationen zu klassifizieren und in Familien zu sortieren. Da beispielsweise ein Ensemble von 1000 Strukturen eines Hexapeptids kaum manuell sortiert werden kann, werden so genannte *Cluster*-Algorithmen

eingesetzt. In dieser Arbeit wird zum Sortieren von Strukturen das Programm NMRCLUST<sup>[81]</sup> verwendet.

NMRCLUST benötigt eingangs eine Liste von Atomen zur Superpositionierung. Es führt dann zuerst die Superpositionierung jeder Struktur des  $N$  Strukturen enthaltenden Ensembles mit einer jeden anderen Struktur des Ensembles durch, und ermittelt für jede Superpositionierung die mittlere quadratische Abweichung (RMSD). Die dadurch erhaltene  $N \cdot N$  Matrix von RMSD-Werten wird dann so in Familien sortiert, dass einerseits der RMSD innerhalb einer Familie möglichst gering bleibt, andererseits aber auch die Anzahl von Familien so gering wie möglich bleibt. Auf diese gegensätzlichen Forderungen kann durch Angabe eines maximalen RMSDs Einfluss genommen werden. Für jede Familie gibt das Programm eine repräsentative Struktur an, die den geringsten maximalen RMSD zu allen anderen Strukturen der Familie hat. Familien, die aus nur einer einzelnen Struktur bestehen, werden als so genannte *outlier* klassifiziert.

Andere Algorithmen verwenden für die Superpositionierung beispielsweise Diederwinkel anstelle von Atomen, und auch im Verfahren zur Einteilung in Familien können Unterschiede bestehen. Wichtig ist es jedoch, auf ein grundlegendes Problem hinzuweisen. Bestehen abgestufte Übergänge, wie etwa zwischen Molekülen verschiedener Atommassen, ergibt sich stets die Möglichkeit der Einteilung in Familien. Immer dann jedoch, wenn ein fließender Übergang zwischen Zuständen möglich ist, wie beispielsweise zwischen Farben in einem Spektrum, wird das Ergebnis einer Einteilung in Familien stark abhängig von der Anzahl der Mitglieder im Ensemble. Die Einteilung wird mit zunehmender Anzahl von Farbtönen immer schwieriger. Im Extremfall, beim Vorhandensein aller Farbtöne, wird ein guter Algorithmus zwar immer noch in der Lage sein, Familien zu erzeugen und diese so zu wählen, dass vorwiegend vorkommende Farbtöne in separate Familien eingeteilt werden. Dennoch wird aber die Anzahl der nur einen Farbton enthaltenden Familien mit steigender Anzahl von Mitgliedern im Ensemble wachsen. Da zwischen den Konformationen eines Moleküls ebenfalls ein fließender Übergang besteht, sortieren Clusteralgorithmen oftmals Molekülkonformationen in verschiedene Familien, obwohl nach chemischem Verständnis die gleiche Faltung vorliegt. Daher empfiehlt es sich, nach Anwendung eines *Cluster*-Algorithmus stets auch die wenig populierten Familien zu betrachten und die einzelnen Familien gegebenenfalls manuell neu zu gruppieren.

## 3 Struktur von *cyclo(-Gum-Lys-Gum-D-Phe-)* in Lösung

### 3.1 Einleitung

In neuerer Zeit sind Zuckeraminosäuren<sup>[82-86]</sup> verstärkt in das Interesse der Wirkstoff- und Strukturforschung getreten. Durch ihren Einbau in Cyclopeptide kann gezielt die Peptidrückgratstruktur sowie die Seitenkettenanordnung beeinflusst werden. Auf diese Weise wird es möglich, die Pharmakokinetik, metabolische Stabilität und Aktivität von peptidischen Leitstrukturen zu modulieren und zu verbessern.

Eine solche Zuckeraminosäure, der „Gum“-Baustein, besitzt als  $\delta$ -Aminosäure die gleiche Länge wie ein Dipeptid aus natürlichen Aminosäuren. Gum wurde in unserem Arbeitskreis mehrmals erfolgreich einzeln als  $\beta$ -Schleifen-induzierendes Element eingesetzt.<sup>[87, 88]</sup> Das Ziel war es zunächst zu überprüfen, ob eine Verbindung der Art *cyclo(-Gum-XAA-Gum-XAA-)*, wie erwartet, die oftmals in cyclischen Hexapeptiden vorgefundene Anordnung zweier gegenüberliegender  $\beta$ -Schleifen einnimmt. Während sich diese Erwartung nicht bestätigte, stellte sich die berechnete Modellverbindung später als interessantes Studienobjekt zur Anwendung einiger Überlegungen zum Prinzip von DUNITZ und WASER<sup>[89]</sup> heraus.

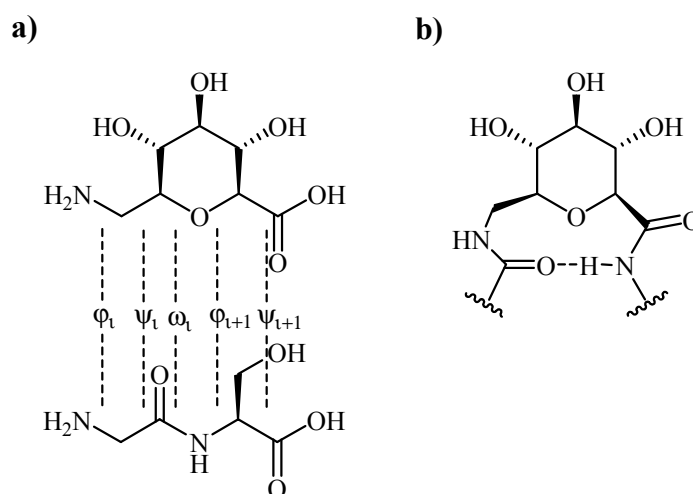
### 3.2 Gum und andere Zuckeraminosäuren in cyclischen Peptiden

Als Zuckeraminosäuren (engl. Sugar Amino Acids, SAAs) werden Pyranosen oder Furanosen bezeichnet, die mindestens eine Carboxyl- und eine Aminogruppe enthalten.<sup>[82-86]</sup> Sie werden als Schleifenmimetika, peptidomimetische und glycomimetische Gerüststrukturen sowie in der kombinatorischen Chemie eingesetzt.<sup>[82-86]</sup> Neben einigen natürlich vorkommenden Zuckeraminosäuren wurden bis heute unzählige derartige Verbindungen synthetisch dargestellt. Ihre Verwendung als Baustein in Cyclopeptiden bietet eine Reihe von Vorteilen gegenüber dem Einsatz konventioneller Aminosäuren. Neben der bereits erwähnten Möglichkeit der Steuerung von Rückgrat- und Seitenkettenanordnung lässt sich beispielsweise gezielt die Hydrophilie bzw. Lipophilie des Peptids durch eine geeignete Wahl von Schutzgruppen am Zucker beeinflussen.<sup>[85, 86]</sup> Dadurch ergibt sich die Möglichkeit, derartige Zyklen als Transportmoleküle für Wirkstoffe im Organismus zu verwenden.

Abgesehen von rein pharmakologischen Anwendungsmöglichkeiten lassen sich Zuckeraminosäuren enthaltende Cyclopeptide auch in anderen Gebieten der Chemie vorteilhaft einsetzen. Ein wichtiger Bereich ist beispielsweise die Templat-assoziierte Synthese von Proteinen (TASP).<sup>[90-92]</sup> Deren Ziel ist es, künstliche Proteine mit gewünschter räumlicher Struktur an einem topologisch maßgeschneiderten Schablonenmolekül zu erzeugen. Hierbei werden amphiphile, sekundärstrukturbildende Peptidblöcke kovalent an die Schablone gebunden. Die strukturinduzierende Wirkung der Schablone bewirkt zunächst eine Sekundärstruktur-Faltung der Peptidblöcke, und aufgrund der räumlichen Orientierung der Seitenketten resultiert dann die Tertiärstruktur des synthetischen Proteins.

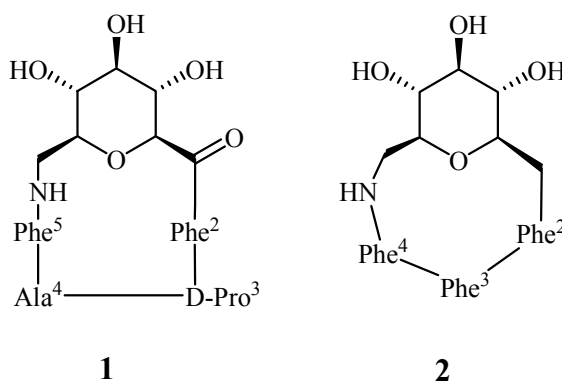
Pyranoide Zuckeraminosäuren, die ausschließlich äquatorial substituiert sind, besitzen wegen der relativen Rigidität des Zuckerrings eine besonders stark stabilisierende Wirkung auf die Rückgratkonformation von Cyclopeptiden. Ein derartige Verbindung ist das Glucosyluronsäure-methylamin (IUPAC-Name *7-Amino-2,6-anhydro-7-desoxy-L-glycero-L-guloheptonsäure*), das erstmals in unserem Arbeitskreis von E. GRAF VON ROEDERN dargestellt und mit dem Kürzel „Gum“ versehen wurde.<sup>[82, 83, 88]</sup> Der Gum-Baustein ist die erste Zuckeraminosäure, die in Peptide eingebaut wurde. Bezüglich der Anzahl von Rückgratatomen und der Raumstruktur verhält sich Gum analog zum Dipeptid Gly-Ser (Abbildung 3-1 a). Allerdings wurden auch die Sequenzen Gly-Gly, Phe-Pro, D-Phe-Val, Tyr-Gly und Trp-Phe durch Gum ersetzt.<sup>[87, 88]</sup>

Aufgrund der räumlichen Anordnung nimmt der Baustein, einzeln eingebaut in cyclische Hexapeptide, die Positionen  $i+1$  und  $i+2$  einer  $\beta$ II'-Schleife ein (Abbildung 3-1 b).



**Abbildung 3-1:** Der Gum-Baustein in Analogie zum Dipeptid Gly-Ser (a) und als  $\beta$ -Schleifen-Mimetikum (b).

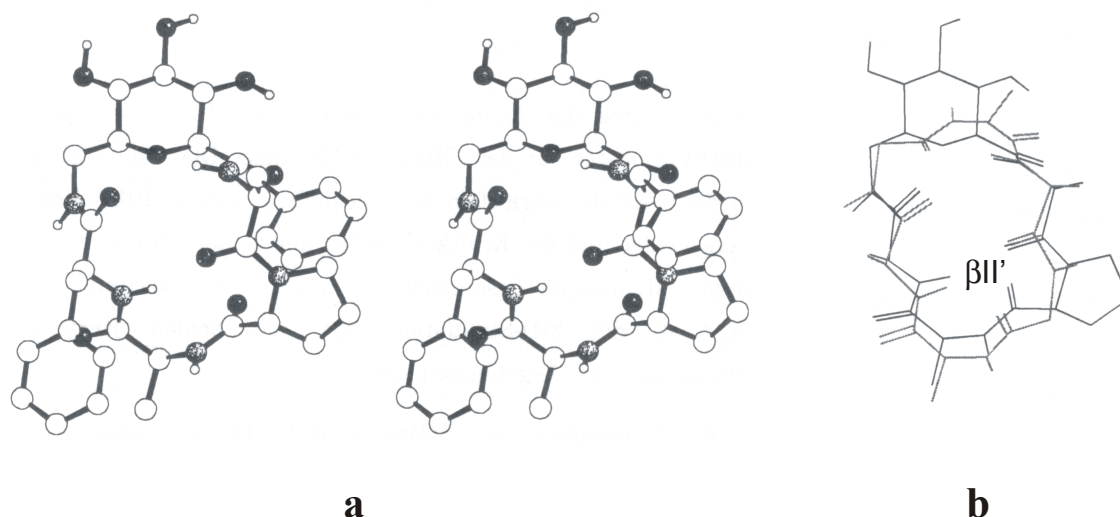
Erste systematische Untersuchungen der strukturellen Eigenschaften des Gum-Bausteins in Cyclopeptiden wurden von E. GRAF VON ROEDERN an den Modellpeptiden *cyclo(-Gum<sup>1</sup>-Phe<sup>2</sup>-D-Pro<sup>3</sup>-Ala<sup>4</sup>-Phe<sup>5</sup>-)* **1** und *cyclo(-Gum<sup>1</sup>-Phe<sup>2</sup>-Phe<sup>3</sup>-Phe<sup>4</sup>-)* **2** durchgeführt. Die Ringgrößen von **1** und **2** entsprechen dabei denen eines cyclischen Hexa- bzw. Pentapeptids.



**Abbildung 3-2:** Die cyclischen Modellpeptide **1** und **2**.

Cyclische Hexapeptide liegen oft in einer Hauptkonformation mit zwei gegenüberliegenden  $\beta$ -Schleifen vor.<sup>[93-97]</sup> Der Grundgedanke beim Entwurf von Verbindung **1** war es, dass das

stark  $\beta$ -Schleifen induzierende D-Pro<sup>3</sup> in der Position  $i+1$  eines  $\beta$ II'-turns liegen sollte. Dadurch müsste Gum<sup>1</sup>, falls möglich, die Positionen  $i+1$  und  $i+2$  einer gegenüberliegenden  $\beta$ -Schleife einnehmen. Diese Anordnung wurde in der Tat in der Struktur von **1** in Lösung beobachtet (Abbildung 3-3 a).<sup>[88]</sup>

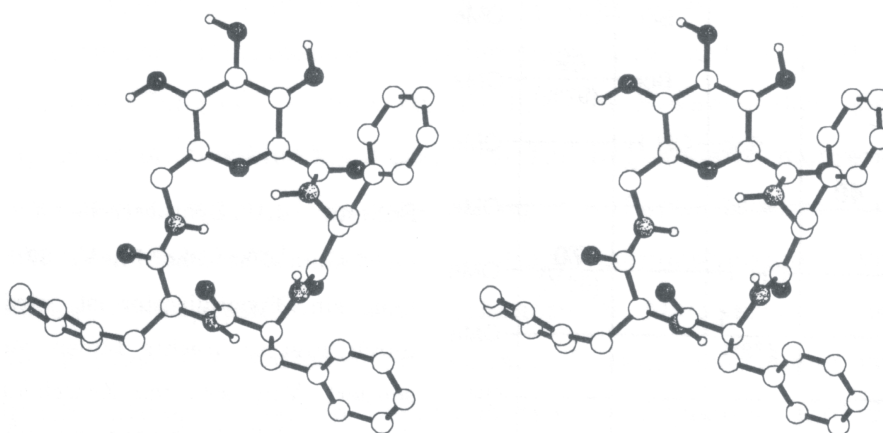


**Abbildung 3-3:** a) Stereodarstellung der Struktur von *cyclo(-Gum<sup>1</sup>-Phe<sup>2</sup>-D-Pro<sup>3</sup>-Ala<sup>4</sup>-Phe<sup>5</sup>-)* (**1**) in DMSO. b) Superpositionierung der Rückgrat-Atome von **1** auf *cyclo(-Gly-Ala-Ala-Gly-Ala-Ala-)* **3**.

Üblicherweise liegen die Ringatome eines cyclischen Hexapeptids, das in der Anordnung zweier gegenüberliegender  $\beta$ -Schleifen vorliegt, annähernd in einer Ebene. Die räumliche Abstoßung zwischen dem Ring-Sauerstoff von Gum und dem Carbonyl-Sauerstoff von Phe<sup>5</sup> führt jedoch bei Verbindung **1** zu einem relativ stark gewölbten Rückgrat. Ein Vergleich mit *cyclo(-Gly-Ala-Ala-Gly-Ala-Ala-)* **3**, das zwei gegenüberliegende  $\beta$ II'-Schleifen einnimmt, ist in Abbildung 3-3 b zu sehen.<sup>[88]</sup> Der Vergleich der Rückgratwinkel hat gezeigt, dass lediglich der frei drehbare Winkel  $\psi_i$  und der festgelegte Winkel  $\phi_{i+1}$  in Gum stark von den entsprechenden Winkeln in **3** abweichen. Diese Abweichungen gleichen sich allerdings über den gesamten Zyklus weitgehend aus, so dass **1** dennoch die beobachtete Konformation mit einer  $\beta$ II'- und einer Pseudo- $\beta$ II'-Schleife einnimmt.

Für das cyclische Pentapeptid-Analogon *cyclo(-Gum<sup>1</sup>-Phe<sup>2</sup>-Phe<sup>3</sup>-Phe<sup>4</sup>-)* **2** ergeben sich, wegen der verringerten Anzahl von Torsionsfreiheitsgraden in Gum, etwas andere Verhältnisse. Cyclische Pentapeptide sind trotz ihrer geringeren Ringgröße meist flexibler als

cyclische Hexapeptide, und in Lösung liegen daher oftmals mehrere Konformationen vor.<sup>[98, 99]</sup> Ein häufig anzutreffendes Strukturmotiv ist die Kombination einer  $\beta$ -Schleife mit einem  $\gamma$ -turn.<sup>[94, 100, 101]</sup> In **2** ist diese Anordnung allerdings nicht ohne Bindungsbruch realisierbar. Der Einbau von Gum führt zu einer erheblichen Erhöhung der Ringspannung und damit zu einem stark eingeschränkten Konformationsraum. In der Struktur von **2** in Lösung (Abbildung 3-4) ist eine  $\gamma$ -Schleife mit Phe<sup>2</sup> in Position *i* zu erkennen, die jedoch stark verzerrt ist und deren Wasserstoffbrücke während der Dynamiksimulation nur zu 12 % populiert blieb.<sup>[88]</sup>



**Abbildung 3-4:** Stereodarstellung der Struktur von *cyclo(-Gum<sup>1</sup>-Phe<sup>2</sup>-Phe<sup>3</sup>-Phe<sup>4</sup>-)* **2** in DMSO.

### 3.3 *Cyclo(-Gum-Lys-Gum-D-Phe-)* **4** als cyclisches Hexapeptid

Während cyclische Peptide mit einem einzelnen Gum-Baustein in unserem Arbeitskreis umfangreich untersucht worden sind, lagen bis vor kurzem noch keinerlei Daten über Verbindungen mit mehr als einem dieser Bausteine vor. Im Rahmen ihrer Dissertation synthetisierte und charakterisierte E. LOCARDI erstmalig die linearen und cyclischen Di-, Tri-, Tetra- und Hexamere von Gum.<sup>[102]</sup> Das cyclische Trimer *cyclo(-Gum<sub>3</sub>-)* entspricht bezüglich seiner Ringgröße wiederum einem cyclischen Hexapeptid aus  $\alpha$ -Aminosäuren. Diese Verbindung könnte zwar eine einzelne  $\beta$ -Schleife ausbilden, jedoch aus Symmetriegründen keine Anordnung zweier gegenüberliegender  $\beta$ -Schleifen. Allerdings zeigen einfache

kraftfeldbasierte Minimierungen, dass selbst eine Konformation mit einem einzelnen  $\beta$ -turn energetisch sehr ungünstig ist.

Die bisherigen Erkenntnisse über cyclische Hexa- und Pentapeptide mit jeweils einem Gum-Baustein legten es für M. STÖCKLE in unserem Arbeitskreis nahe, Verbindungen der Art *cyclo(-Gum-XAA-Gum-XAA-)* zu synthetisieren. Durch den alternierenden Einsatz von Gum in einem cyclischen Hexapeptid sollte es möglich sein, zusätzliche Rigidität in die bekannte Struktur zweier gegenüberliegender  $\beta$ -Schleifen einzuführen, falls diese Anordnung in solchen Verbindungen realisiert werden könnte. Um letzteres zu überprüfen, wurde im Rahmen dieser Arbeit von den 16 synthetisierten Verbindungen exemplarisch die NMR-Struktur von *cyclo(-Gum<sup>1</sup>-Lys<sup>2</sup>-Gum<sup>3</sup>-D-Phe<sup>4</sup>-)* **4** aufgeklärt.

### 3.3.1 NMR-Strukturdaten

Alle NMR-Experimente wurden von M. STÖCKLE durchgeführt und ausgewertet. Mittels Anwendung der Näherung isolierter Spinpaare (ISPA)<sup>[36]</sup> wurden die Integrale der Kreuzsignale eines *offset*-korrigierten ROESY-Spektrums (H<sub>2</sub>O/D<sub>2</sub>O 90:10, Mischzeit 200 ms) in Distanzen  $r_{NOE}$  umgewandelt. Hieraus wurden Ober- und Untergrenzen der *restraints* für die Strukturrechnungen gewonnen (1.15  $r_{NOE}$  bzw. 0.85  $r_{NOE}$ , siehe folgendes Kapitel). Wegen Mangels an aufgelösten Signalen geminaler Protonen erfolgte die Kalibrierung auf den Referenzabstand Gum<sup>1</sup>H<sup>3</sup>-Gum<sup>1</sup>H<sup>5</sup> (2.64 Å), der aufgrund der energetischen Bevorzugung der all-äquatorialen <sup>4</sup>C<sub>1</sub>-Sesselkonformation des Zuckerrings als hinreichend konstant angenommen werden kann. Für alle verwendeten Pseudoatome wurde die Obergrenze um die jeweilig erforderliche Standard-Pseudoatomkorrektur<sup>[103]</sup> erhöht. Homonukleare <sup>3</sup>J(H<sup>N</sup>H<sup>α</sup>) und <sup>3</sup>J(H<sup>α</sup>H<sup>β</sup>) Kopplungskonstanten wurden aus <sup>1</sup>H-NMR-Spektren, einem DQF-COSY-Spektrum und einem E.COSY-Spektrum entnommen. Heteronukleare <sup>3</sup>J(H<sup>β</sup>C<sup>γ</sup>)-Kopplungskonstanten wurden qualitativ aus den Intensitäten eines HMBC-Spektrums abgelesen. Die diastereotope Zuordnung der  $\beta$ -Protonen von Lys<sup>2</sup> gelang mittels der Kopplungen <sup>3</sup>J(H<sup>α</sup>H<sup>β</sup>) und <sup>3</sup>J(H<sup>β</sup>C<sup>γ</sup>). Gum<sup>3</sup>H<sup>C, pro S</sup> konnte durch eine ROESY-Korrelation mit Lys<sup>4</sup>H<sup>N</sup> eindeutig zugeordnet werden. Die Temperaturgradienten der Amidprotonen wurden aus einer Reihe von <sup>1</sup>H-NMR-Spektren bei den Temperaturen 293 K, 297 K, 300 K, 305 K, 310 K, 315 K und 320 K ermittelt.



### 3.3.2 Strukturrechnungen

Das ROESY-Spektrum von **4** deutete auf die Existenz einer auf der NMR-Zeitskala aufgelösten Nebenkonformation hin. Obschon diese nur relativ gering populiert ist (ca. 3 % bei 293 K, 5 % bei 320 K), schien es ratsam, eine besonders sorgfältige Suche im Konformationsraum durchzuführen. Unter der Verwendung von 18 dem ROESY-Spektrum entnommenen Abständen wurde daher mit dem DISGEO-Programmpaket<sup>[7-10]</sup>, ausgehend von einer Zufallskonformation Energie-optimierter, kovalenter Bindungsgeometrien, ein erstes Ensemble von 1000 Strukturen erzeugt. Wegen der geringen Population der Nebenkonformation war es allerdings nicht möglich, deren Signale eindeutig zuzuweisen. Manche der integrierten Kreuzsignale konnten somit Anteile der Nebenkonformation enthalten. Die Anwendung relativ strikter *restraint*-Grenzen könnte in einem solchen Fall das Auffinden von in der Realität vorkommenden Konformationen ausschließen. Daher wurde von der Standardprozedur abgewichen und die Ober- und Untergrenzen der *restraints* erzeugt, indem eine Toleranz von  $\pm 15\%$ , anstelle von  $\pm 10\%$ , auf die aus dem ROESY-Spektrum ermittelten Distanzen erlaubt wurde. Das Ensemble von DG-Strukturen wurde dann mit dem Programm NMRCLUST<sup>[81]</sup> in Familien sortiert. Für die Superpositionierung wurden dabei alle Rückgrat-Schweratome inklusive der im Peptidring enthaltenen Schweratome des Zuckerbausteins verwendet. Trotz der im ROESY-Spektrum beobachteten Nebenkonformation wurde für alle Strukturen im Ensemble im Wesentlichen die gleiche Rückgratfaltung gefunden. Die repräsentative Struktur des höchst populierten *clusters* diente dann, nach einer kurzen *steepest-descent*-Minimierung (300 Schritte) zur Beseitigung von lokalen Verzerrungen, als Startstruktur für die anschließende Verfeinerung.

Hierzu wurde zunächst eine *restrained* MD-Simulation (rMD) mit dem Programm DISCOVER 2.9.8.<sup>[25]</sup> durchgeführt. Dabei kam das *explicit image* periodischer Randbedingungen im CVFF-Kraftfeld zum Einsatz. Als *restraints* wurden ausschließlich die Abstandsbeschränkungen verwendet. Die Startstruktur wurde zunächst in eine kubische Solvenszelle mit einer Seitenlänge von 46 Å gegeben und die Zelle mit 3199 Wassermolekülen gefüllt. Das System wurde nun in zwei Stufen, unter Einwirkung der *restraints*, Energie-minimiert. Zuerst wurden die Atome der Startstruktur räumlich fixiert und die Zelle bis zur Konvergenz minimiert, um die Solvensmoleküle optimal an die Startstruktur zu arrangieren. Dann wurde auch den Atomen der Startstruktur Beweglichkeit erlaubt und nochmals bis zur Konvergenz minimiert. Nach langsamem Aufheizen des Systems auf die Messtemperatur von 300 K und einer Äquilibrationsphase von 25 ps wurden 150 ps Dynamik berechnet, wobei die Atomkoordinaten alle 100 fs in einer Trajektorie gespeichert wurden. Die Trajektorie wurde dann

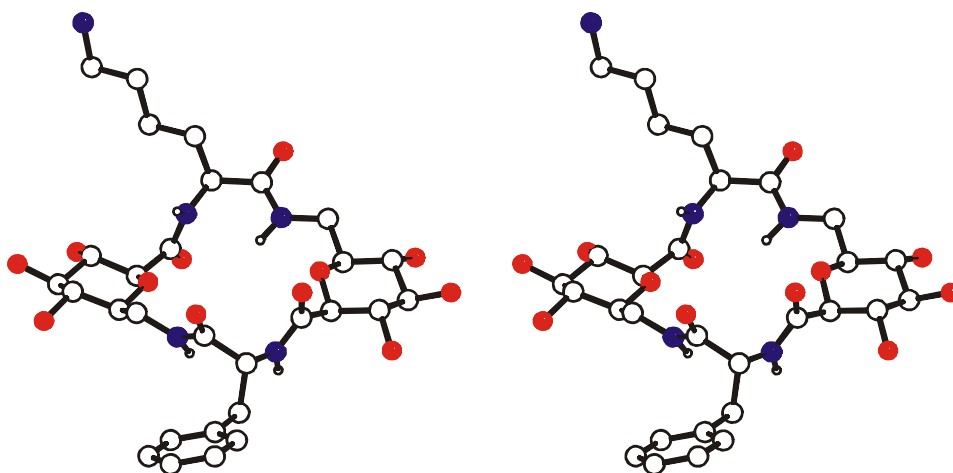
gemittelt und die resultierende Struktur schonend mit dem *steepest-descent*-Algorithmus Energie-minimiert (200 Schritte), um lokale Verzerrungen zu beseitigen, die durch die Mittelung hervorgerufen wurden.

Um die Stabilität der verfeinerten Struktur im Kraftfeld zu überprüfen, wurde eine Dynamiksimulation ohne Einwirkung von *restraints* durchgeführt (freie MD, fMD). Die Bedingungen (Kraftfeld, Dimensionen der Zelle) entsprachen dabei denen der rMD. Die rMD-Struktur wurde in die Solvenszelle gegeben und zunächst nur dem Solvens Beweglichkeit gestattet. Nach Energie-Minimierung bis zur Konvergenz wurde die Zelle, wie oben beschrieben, auf 300 K aufgeheizt und äquilibriert. Anschließend wurde dem gesamten System Beweglichkeit gestattet, allen Atomen Zufallsgeschwindigkeiten gemäß einer BOLTZMANN-Verteilung bei 300 K zugewiesen und wiederum eine Trajektorie von 150 ps Dauer aufgezeichnet. Die fMD-Struktur wurde dann, analog der rMD-Struktur, durch Mittelung und Minimierung gewonnen.

Zur Aufklärung der Struktur der Nebenkonformation bot das ROESY-Spektrum, wegen oftmals starker Überlagerung mit der Hauptkonformation, keine ausreichende Anzahl von Signalen.

### 3.3.3 Ergebnisse und Diskussion

Die Dynamiksimulation von Verbindung **4** zeigt eine gute Übereinstimmung mit den experimentellen Daten. Während der rMD-Simulation wurde keiner der 18 Abstands-*restraints* nennenswert verletzt (Tabelle im Anhang; die größte Abweichung,  $\langle r^{-3} \rangle^{-1/3}$  gemittelt über die Trajektorie, betrug 0.13 Å). Abbildung 3-5 zeigt die gemittelte und Energie-minimierte rMD-Struktur von **4**.



**Abbildung 3-5:** Stereoansicht der gemittelten und 300 Schritte steepest descent minimierten Struktur von cyclo(-Gum<sup>1</sup>-Lys<sup>2</sup>-Gum<sup>3</sup>-D-Phe<sup>4</sup>-) **4** in H<sub>2</sub>O, die sich während der 150 ps rMD-Simulation ergab. Sauerstoffatome sind in rot, Stickstoffatome in blau, Kohlenstoffatome in weiß, und Wasserstoffatome (kleine Kugeln) ebenfalls in weiß dargestellt. Zur Erstellung der Graphik wurde das Programm Schakal 99 verwendet.<sup>[104]</sup>

Wie die Struktur von **4** in H<sub>2</sub>O zeigt, bildet sich keine Anordnung zweier gegenüberliegender  $\beta$ -Schleifen aus. Entsprechende Wasserstoffbrücken zwischen Lys<sup>2</sup>CO und D-Phe<sup>4</sup>HN bzw. zwischen D-Phe<sup>4</sup>CO und Lys<sup>2</sup>HN fehlen vollkommen. Auch weitere Wasserstoffbrücken werden nicht ausgebildet. Es bestehen lediglich schwache elektrostatische Wechselwirkungen zwischen den Ring-Sauerstoffatomen der Gum-Reste und ihnen räumlich benachbarten Amidprotonen. Dieser Befund steht im guten Einklang mit den gemessenen Temperaturgradienten der Amidprotonen, die eine gute Solvenz Zugänglichkeit aller H<sup>N</sup> nahe legen (Tabelle 3-1).

**Tabelle 3-1:** Temperaturgradienten der Amidprotonen von **4**.

	Gum <sup>1</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>N</sup>	Gum <sup>3</sup> H <sup>N</sup>	D-Phe <sup>2</sup> H <sup>N</sup>
Temperaturgradient [ppb K <sup>-1</sup> ]	-8.08 ± 0.13	-5.35 ± 0.10	-9.16 ± 0.14	-7.60 ± 0.12

Bei allen durchgeführten Dynamiksimulationen verblieb der Rückgratwinkel  $\psi$  von Lys<sup>2</sup> in Bereichen, die einen gewissen Anteil an 1,3-Allylspannung<sup>[105, 106]</sup> nahe legen. Ein Zusammenhang mit den verwendeten *restraints*, die dies eventuell erzwingen würden, konnte allerdings nicht festgestellt werden. Interessanterweise unterscheidet sich der Temperaturgradient von Lys<sup>2</sup>H<sup>N</sup> leicht von den Gradienten der anderen Amidprotonen, was auf leichte Unterschiede zwischen den Umgebungen von Lys<sup>2</sup>H<sup>N</sup> und D-Phe<sup>4</sup>H<sup>N</sup> hindeuten würde. Dennoch dürfte die etwas unerwartete Stellung der Amidbindung zwischen Lys<sup>2</sup> und Gum<sup>3</sup> auf Ungenauigkeiten der Kraftfeldrechnungen zurückzuführen sein.

Die Superposition der gemittelten und Energie-minimierten rMD- und fMD-Strukturen zeigt, dass die ermittelte rMD-Struktur von **4**, bei einer Dynamiksimulation ohne Einwirkung von *restraints*, im CVFF-Kraftfeld hinreichend stabil bleibt. Der RMSD für die Superpositionierung aller Rückgrat-Schweratome, der Amidprotonen sowie der im Peptidring enthaltenen Schweratome des Zuckerbausteins beträgt 0.46 Å.

Wie bereits erwähnt, zeigt der Vergleich zwischen *cyclo(-Gum<sup>1</sup>-Phe<sup>2</sup>-D-Pro<sup>3</sup>-Ala<sup>4</sup>-Phe<sup>5</sup>-)* **1** und *cyclo(-Gly-Ala-Ala-Gly-Ala-Ala-)* **3**, dass die Gum-Winkel  $\psi_i$  (frei drehbar) und  $\phi_{i+1}$  (durch die Zuckerkonformation festgelegt) in **1** stark von den entsprechenden Winkeln in **3** abweichen (Kapitel 3.2, Seite 39). Diese Abweichungen lassen sich durch den Vergleich zwischen einer regulären  $\beta$ -Schleife und einer analogen Schleife mit Gum in den Positionen  $i+1$  und  $i+2$  veranschaulichen, der die Auswirkungen verdeutlicht, die sich durch den Einbau von Gum ergeben. Im der regulären  $\beta$ -Schleife ist die Peptid-Bindung zwischen den Aminosäuren in den Positionen  $i+1$  und  $i+2$  planar, und alle Winkel betragen annähernd 120°. Demgegenüber ist das der Peptidbindung entsprechende Fragment im Gum-Baustein eine aliphatische Kette mit annähernd tetraedrischen Winkeln. Insgesamt resultiert daraus eine merkliche Aufweitung der Wasserstoffbrücken-Distanz CO( $i$ )-HN( $i+3$ ) von ca. 1.9 Å im regulären Peptid auf ca. 2.4 Å in der Gum-Schleife. Während die sich daraus ergebenden Abweichungen in **1** noch hinreichend durch die Aminosäuren in den Positionen  $i+5$  und  $i+6$  ausgeglichen werden können, um eine Anordnung von zwei  $\beta$ -Schleifen zu erreichen, ist dies für **4** offensichtlich nicht mehr möglich. In Verbindung mit den früheren Ergebnissen über

Gum in cyclischen Pentapeptiden muss also festgestellt werden, dass der Einbau auch nur eines Gum-Bausteins eine sehr hohe, zusätzliche Ringspannung in cyclische Peptide einbringt.

An dieser Stelle soll kurz angemerkt werden, dass auf den ersten Blick eine Konformation von **4** denkbar wäre, in der D-Phe die übliche Position  $i+1$  eines  $\beta$ II'-turns einnimmt. Dies ist jedoch ausgeschlossen, da anstelle der stabilisierenden Wasserstoffbrücken zwei energetisch ungünstige Kontakte zwischen den Carbonyl-Sauerstoffatomen  $CO(i)$  und den entsprechenden Gum-Ring-Sauerstoffatomen  $O(i+3)$  rücken würden.

Die mit der hohen Ringspannung einhergehende Einschränkung des Konformationsraums kann durchaus von Vorteil sein. Beispielsweise ist es denkbar, bei einem vorliegenden Gleichgewicht zwischen zwei Konformeren in einem cyclischen Hexapeptid, durch Einbau von Gum eines der Konformere energetisch auszuschließen. Gepaart mit der erhöhten Rigidität des verbleibenden Konformers können auf diese Weise gezielt Informationen über Struktur-Wirkungsbeziehungen biologisch aktiver Verbindungen gewonnen werden.

### 3.4 *Cyclo(-Gum-Lys-Gum-D-Phe-)* 4 als Cyclohexan gemäß DUNITZ und WASER

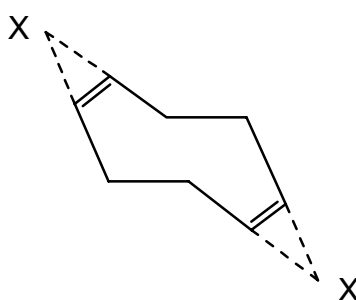
#### 3.4.1 Grundlagen

In realen Molekülen werden Konformation und Flexibilität meist durch eine große Anzahl von Faktoren, wie beispielsweise elektrostatische Wechselwirkungen, VAN DER WAALSsche Wechselwirkungen, sterische Abstoßung etc., bestimmt. Dennoch bestehen für cyclische Moleküle auch rein geometrische Aspekte bei der Ausbildung von Strukturen.

Bereits 1890 wurde von H. SACHSE<sup>[107]</sup> die Natur der Konformationen des Cyclohexans erkannt, und der Kontrast zwischen der Rigidität der Sesselform und der relativen Flexibilität der Wannenform beschrieben. Erst mehr als 70 Jahre später allerdings wurden von LAUWERIER und BOTTONA die mathematischen Eigenschaften des flexiblen, gleichseitigen, gleichwinkligen Hexagons untersucht.<sup>[108, 109]</sup> Dabei hat sich der Unterschied bezüglich der Flexibilität der beiden Konformationen bestätigt. Flexibilität im mathematischen Sinne bedeutet dabei, dass konformationelle Übergänge erfolgen können, ohne dass Bindungslängen oder Bindungswinkel deformiert werden müssen. Für die Wannenform von Cyclohexan ist dies der Fall, wobei der Übergang über die *twist*-Form verläuft. In der Sesselform dagegen können Variationen in der Konformation nur durch energetisch ungünstige Abweichungen von der idealen Bindungsgeometrie erkauft werden. In der Realität wird jedoch oft der leichte,

entropische Vorteil der Wannenform durch andere Faktoren überwogen. Beispielsweise kommt es in der Wannenform von Cyclohexan zu ekliptischen Anordnungen und ungünstigen, sterischen Wechselwirkungen.

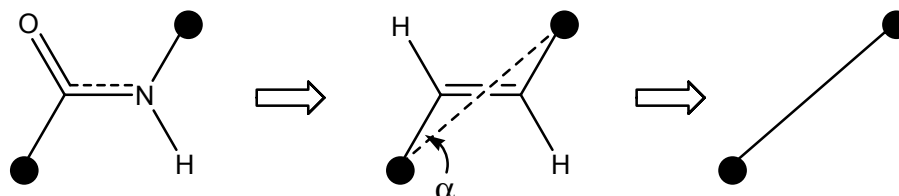
In Unkenntnis der Arbeiten von LAUWERIER und BOTTOMA wiederholten DUNITZ und WASER etwas später die theoretischen Überlegungen und wendeten sie erstmalig auch auf andere cyclische Systeme in der Chemie an.<sup>[89]</sup> Insbesondere neu war der Gedanke, Systeme einer gegebenen Ringgröße auf solche geringerer Ringgröße zu reduzieren, wenn ein oder mehrere Diederwinkel des Systems auf  $0^\circ$  (*cis*) festgelegt sind. So lässt sich beispielsweise 1,5-Cyclooctadien auf Cyclohexan „reduzieren“ (Abbildung 3-6).



**Abbildung 3-6:** Reduktion von 1,5-Cyclooctadien auf Cyclohexan nach DUNITZ und WASER.

Die Kohlenstoffe der C=C-Doppelbindung werden dazu durch ein Pseudoatom X ersetzt, das sich am Schnittpunkt der verlängerten C-C-Einfachbindungen an der Doppelbindung befindet. Eine wichtige Anwendungsmöglichkeit solcher Reduktionen besteht in der Qualitätskontrolle errechneter Strukturen. So sollte beispielsweise eine Struktur, die auf ein Cyclohexan reduziert wurde, einer der möglichen, energetisch günstigen Konformationen desselben entsprechen.

Die Methode wurde schließlich in unserer Arbeitsgruppe auf die Peptidbindung in cyclischen Peptiden übertragen und auch für auf  $180^\circ$  (*trans*) festgelegte Diederwinkel angewendet (Abbildung 3-7; zur Erklärung des Winkels  $\alpha$  siehe nächstes Kapitel).<sup>[110]</sup>



**Abbildung 3-7:** Reduktion von 1,5-Cyclooctadien auf Cyclohexan nach DUNITZ und WASER.

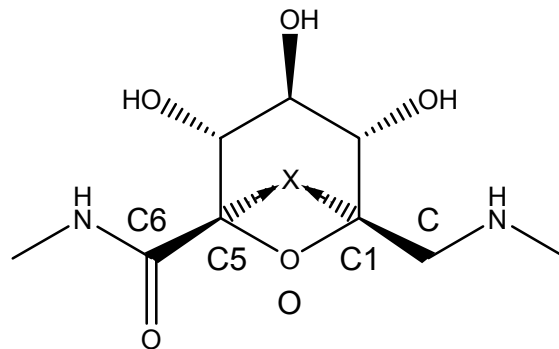
Während *cis*-Peptidbindungen wie *cis*-C=C-Doppelbindungen behandelt werden können, wurden *trans*-Peptidbindungen durch eine Einfachbindung zwischen den  $C^\alpha$ -Atomen durch einfaches Verbinden ersetzt. In diesem Zusammenhang beschrieben H. KESSLER *et al.* erstmalig die Idee, das Konzept für eine neuartige Methode der Suche nach hochpotenten Wirkstoffen, nämlich dem so genannten „räumlichen *screening*“, einzusetzen.<sup>[110]</sup>

### 3.4.2 Ausweitung auf *cyclo(-Gum-Lys-Gum-D-Phe-)* 4

Beim Verbinden der  $C^\alpha$ -Atome fällt auf, dass die Richtungsvektoren der Verbindungslinie und derjenigen der C-C-Einfachbindungen in eine etwas verschiedene Richtung deuten (Winkeldifferenz  $\alpha$  in Abbildung 3-7). Wird das Verfahren beispielsweise auf cyclische Hexapeptide angewendet, so ergeben sich, unter der Voraussetzung einer korrekten Struktur des Peptids, durchaus die erwarteten Konformationen des Cyclohexans. Obschon die Winkeldifferenzen recht gering sind ( $\angle C_i' C_i^\alpha C_{i+1}^\alpha \approx 20.5^\circ$ ,  $\angle C_i^\alpha C_{i+1}^\alpha N_{i+1} \approx 15^\circ$ ), könnten sie allerdings eine zusätzliche Erklärungsmöglichkeit für die oftmals beobachteten, leichten Verzerrungen liefern (neben naturbedingten Abweichungen von der „idealen“ Geometrie der jeweiligen Schleifentypen).

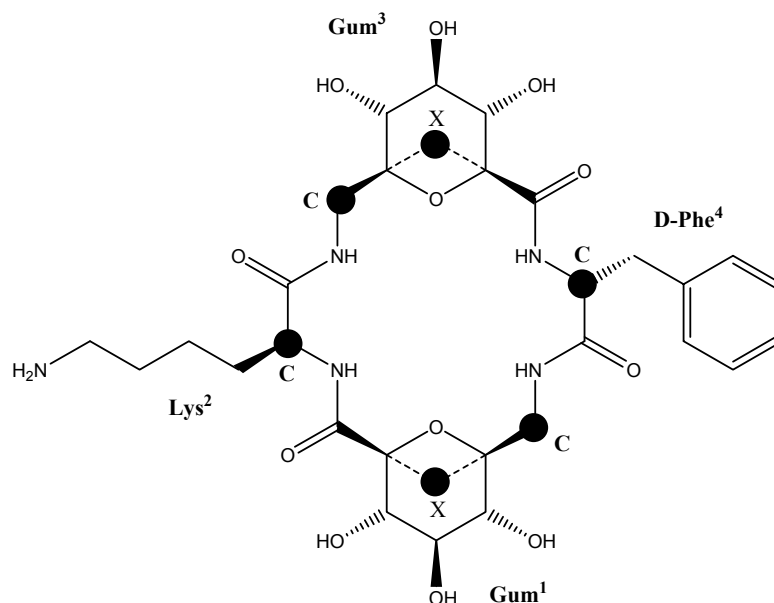
Als Analogon eines cyclischen Hexapeptids lässt sich auch Verbindung 4 auf ein Cyclohexan reduzieren. Voraussetzung dazu ist allerdings eine große Rigidität des Zuckerrings, die jedoch durch die fünf sterisch anspruchsvollen Substituenten hinreichend gewährleistet ist. Durch Verlängerung der Achsen C-C<sup>1</sup> und C<sup>6</sup>-C<sup>5</sup> im Gum-Rest könnte analog der Methode von DUNITZ und WASER ebenfalls ein Pseudoatom definiert werden, das sich im Schnittpunkt der

Verlängerungen befindet. Dieses würde dem  $C^\alpha$  in Position  $i+1$  des Gum entsprechen (Abbildung 3-8).



**Abbildung 3-8:** Pseudoatom im Gum-Baustein.

Wegen der nicht perfekt tetraedrischen Umgebung der Atome  $C^1$  und  $C^5$  würden sich allerdings die Verbindungslinien nicht schneiden. Es müsste also der Punkt berechnet werden, der geometrisch den geringsten Abstand zu beiden Verlängerungslinien besitzt. Viel wichtiger ist jedoch, dass die sechs Ecken der reduzierten Darstellung, im Gegensatz zu denen gewöhnlicher Hexapeptide, keineswegs äquidistant zu ihren jeweiligen Nachbarn wären:

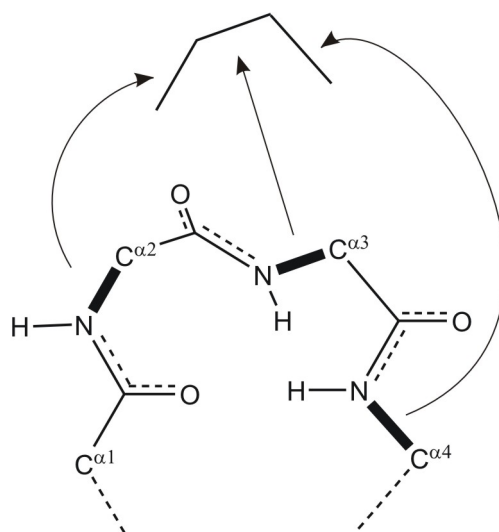


**Abbildung 3-9:** Für die Reduktion nach DUNITZ und WASER benötigte Atome in 4.



Es ist offensichtlich, dass beim Verbinden der Ecken in einem solchen Fall erheblich deformierte Strukturen entstehen würden. Um die Problematik der Bestimmung der Positionen für die Pseudoatome und die der anschließend nötigen Entzerrung zu umgehen, wurde daher in dieser Arbeit ein anderer Weg gewählt, der gleichzeitig die kleinen Ungenauigkeiten der Verbindungsmethode weiter vermindert.

Durch Parallelverschiebung der Bindung  $N(i+1)-C^{\alpha}(i+1)$  auf die Bindung  $C^{\alpha}(i)-C'(i)$  und Wiederholung des Verfahrens für die nächste Bindung  $N(i+2)-C^{\alpha}(i+2)$  entsteht ein Fragment von vier Atomen, in dem alle Richtungsvektoren erhalten sind (Abbildung 3-10).



**Abbildung 3-10:** Verfahren der Parallelverschiebung zur Gewinnung der Diederwinkel in 4.

Da genau genommen die Bindungen  $N(i+1)-C^{\alpha}(i+1)$  und  $C^{\alpha}(i)-C'(i)$  jedoch nicht exakt parallel zueinander sind, bietet es sich als gute Näherung an, zunächst jeweils neue Koordinaten für  $C^{\alpha}(i+1)$  zu berechnen, indem  $N(i+1)$  auf  $C^{\alpha}(i)$  verschoben wird, und die Koordinaten von  $C'(i)$  zu verwerfen. Die neuen Koordinaten von  $C^{\alpha}(i+1)$  können durch Vektoraddition erhalten werden:

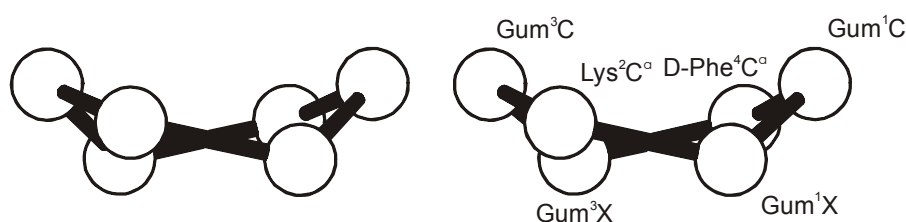
$$(x, y, z)_{C_{neu}^{\alpha}(i+1)} = (x, y, z)_{C^{\alpha}(i+1)} - (x, y, z)_{N(i+1)} + (x, y, z)_{C^{\alpha}(i)} \quad [3-1].$$

Genauso kann dann das nächste Bindungsfragment verschoben werden. Nach Verschieben von zwei Fragmenten kann anschließend der Diederwinkel wie üblich über den Winkel zwischen zwei Ebenen bestimmt werden. Auf entsprechende Weise können auch die

erforderlichen Diederwinkel für den Gum-Baustein gewonnen werden. Ein möglicher Weg ist es beispielsweise, das Fragment  $N(i+2)-C^{\alpha}(i+2)$  (das zur nächsten Aminosäure nach Gum gehört), an das Gum-Fragment  $C^{\delta}-C^{\epsilon}$  zu schieben.

Aus den berechneten Diederwinkeln kann anschließend auf schnelle Weise eine Struktur erzeugt werden, indem die Winkel als *restraints* (mit hoher Kraftkonstante) während einer Minimierung des entsprechenden Cycloalkans eingesetzt werden.

Bestimmt man auf diese Art die reduzierte Struktur von **4**, so ergibt sich eine nahezu ideale *twisted-boat*-Anordnung (Abbildung 3-11):



**Abbildung 3-11:** Durch Parallelverschiebung gewonnene, reduzierte Darstellung der rMD-Struktur von **4**.

Tabelle 3-2 zeigt den Vergleich der entsprechenden Diederwinkel  $\alpha_i$  mit denen einer im CVFF-Kraftfeld minimierten, idealen *twisted-boat*-Konformation von Cyclohexan:

**Tabelle 3-2:** Vergleich der berechneten Diederwinkel in **4** mit den Idealwinkeln in der *twisted-boat*-Konformation von Cyclohexan (beginnend bei  $Gum^1C-Gum^1X-D-Phe^2C^{\alpha}-Gum^3C$ ).

	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$
Ideal:	27.0	27.0	-54.8	27.0	27.0	-54.8
Berechnet:	17.7	43.9	-64.6	41.9	13.1	-53.1

Wie eingangs erwähnt, standen keine ausreichenden Daten zur Verfügung, um die gering populierte Nebenkonformation von **4** aufzuklären. Denkbar wäre beispielsweise eine Konformation mit *cis*-konfigurierten Peptidbindungen zwischen  $Lys^2CO-Gum^3H^N$  und  $D-Phe^4CO-Gum^1H^N$ . Obschon dafür im CVFF-Kraftfeld (bei implizitem Solvens) stabile Minima existieren, liegen diese energetisch jedoch ca.  $10 \text{ kcal mol}^{-1}$  höher als die gefundene

Konformation. Abgesehen davon waren entsprechende  $H^C(i+1)$ - $H^\alpha(i)$ -Kreuzsignale im ROESY-Spektrum von **4** abwesend. Ob die Nebenkonformation von **4** *cis*-Peptidbindungen enthält, oder ob sie doch die denkbare Sesselkonfiguration der all-*trans*-Konformation ist, bleibt daher ins Reich der Spekulationen verbannt.

## 4 $\beta$ -Peptide: Strukturbestimmung einer 12/10/12-Helix in Lösung

### 4.1 Einleitung

Eines der wichtigsten Ziele der Wirkstoffforschung ist die Verbesserung der pharmakologischen Eigenschaften biologisch aktiver Peptide. Aus natürlichen Aminosäuren aufgebaut, haben diese bekanntermaßen eine Reihe von Nachteilen. Beispielsweise besitzen sie geringe Stabilität gegenüber Proteasen, und aufgrund ihrer polaren Natur können sie nur schwierig Membranen oder die Blut-Hirn-Schranke überwinden. Überdies ist die Aktivität und Selektivität kleinerer, linearer Peptide wegen ihrer Flexibilität meist nur gering. Daher werden nach wie vor energische Anstrengungen unternommen, durch Veränderung des Rückgrats und der Seitenketten gezielt auf die Eigenschaften der Peptide einzuwirken.

Neben dem Einbau von Zuckeraminosäuren (Kapitel 3) bietet sich das Ersetzen natürlicher Aminosäuren durch  $\beta$ -Aminosäuren als interessante Variante zur Strukturmodifikation an. Während aus  $\alpha$ -Aminosäuren aufgebaute Peptide mittlerweile recht gründlich untersucht worden sind, bieten Peptide, die  $\beta$ -Aminosäuren enthalten, immer noch jede Menge unerforschtes Neuland, das es zu erkunden gilt.<sup>[111, 112]</sup>

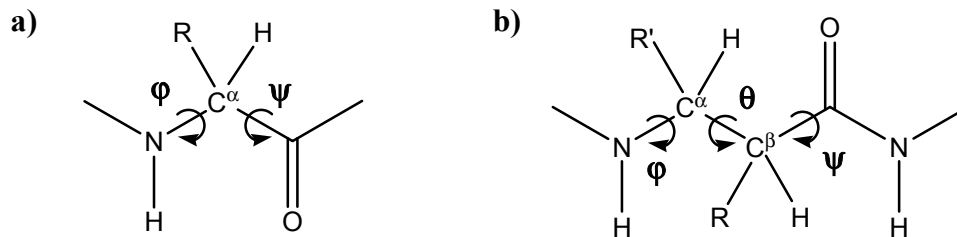
Das in dieser Arbeit berechnete Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH (**5**) ist ein Hybrid aus einem Zuckeraminosäuren-Peptid und einem  $\beta$ -Peptid. Für ähnliche Verbindungen konnte gezeigt werden, dass sie eine inhibierende Wirkung auf die Zelladhäsion Vitronektin-bindender Zellen und auf die Invasion von Tumorzellen besitzen.<sup>[113]</sup> Auf den ersten Blick mag es erstaunlich sein, dass Verbindung **5**, enthält sie doch das hoch flexible  $\beta$ -hGly (Synonym:  $\beta$ -Ala, H<sub>2</sub>N-CH<sub>2</sub>-CH<sub>2</sub>-COOH), eine geordnete Struktur einnehmen sollte. Dennoch zeigte sich, dass dies der Fall ist. In erfreulicher Weise stellte sich später heraus, dass sich die ermittelte Struktur mit quantenmechanischen Berechnungen deckt, die vormalig von R. GÜNTHER durchgeführt worden waren (der kurzfristig in unserem Arbeitskreis verweilte).

Nach einem kurzen Abschnitt über Grundlegendes zu  $\beta$ -Aminosäuren und aus ihnen aufgebauten Peptiden soll daher zunächst auf diese theoretischen Untersuchungen eingegangen werden, und erst danach die berechnete Verbindung beschrieben werden.

## 4.2 Grundlagen

### 4.2.1 $\beta$ -Aminosäuren und aus ihnen aufgebaute Peptide

Während  $\alpha$ -Aminosäuren nur zwei weitgehend frei rotierbare Rückgrat-Torsionswinkel  $\varphi$  und  $\psi$  besitzen, ist in  $\beta$ -Aminosäuren ein weiterer Winkel  $\theta$  variabel (Abbildung 4-1):



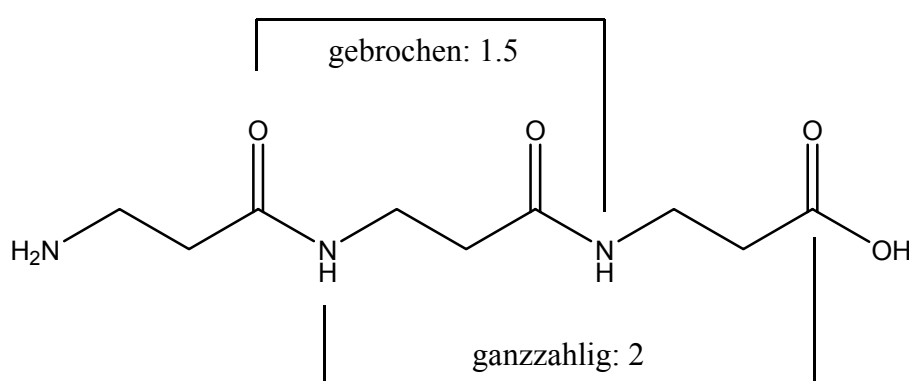
**Abbildung 4-1:** Variable Peptidrückgrat-Diederwinkel in  $\alpha$ -Aminosäuren (a) und in  $\beta$ -Aminosäuren (b).

Die durch den Winkel  $\theta$  zusätzlich eingeführte Flexibilität sollte eigentlich definierte Strukturen in  $\beta$ -Peptiden verhindern. Andererseits kann jedoch der Konformationsraum durch die Seitenketten  $R$  und  $R'$  so weit eingeeengt sein, dass es dennoch zur Ausbildung von Strukturmotiven kommt, die denen natürlicher Peptide und Proteine manchmal sehr ähnlich sind.

Während bereits 1965 die ersten Anstrengungen unternommen wurden, Verbindungen mit geordneten Strukturen zu erzeugen,<sup>[114]</sup> konnten erst etwa 20 Jahre später mittels Röntgenspektroskopie die ersten Beweise für starre Konformationen in  $\beta$ -Peptiden erbracht werden.<sup>[115, 116]</sup> Seither wurden neben einigen  $\beta$ -Faltblatt-artigen Strukturen verschiedene links- und rechtsgängige Helices, wie beispielsweise die  $3_1$ -,  $4_1$ -,  $5_2$ -,  $7_2$ - und  $13_4$ -Helices, postuliert und teilweise durch Röntgenspektroskopie oder mittels NMR bestätigt.<sup>[111, 117-127]</sup>

Um die betreffenden Helices in  $\beta$ -Peptiden zu klassifizieren, stehen mehrere Nomenklatorsysteme zur Verfügung. Eine Möglichkeit ist es, von der Nummer des Rests auszugehen, der dieselbe räumliche Position entlang der Helixachse hat wie der erste Rest.

Mit einem Index versehen, der die Anzahl der Windungen dazwischen angibt, kann eine Helix vollständig beschrieben werden.<sup>1</sup> In  $\beta$ -Peptid-Helices, wie der  $4_1$ -,  $5_2$ -, oder der  $13_4$ -Helix, liegen jeweils verschiedene Reihenfolgen von Wasserstoffbrücken vor. Eine andere Weise der Beschreibung geht daher von der Anzahl der Aminosäurereste in einem Wasserstoffbrücken-gebundenen Ring  $C_i$  aus und indiziert diese mit der Anzahl der im Ring enthaltenen Atome  $i$ . Beispielsweise besitzen die  $4_1$ -,  $5_2$ -, und  $13_4$ -Helices jeweils 18-, 12-, und 14-gliedrige Wasserstoffbrücken-Ringe ( $C_{18}$ ,  $C_{12}$ ,  $C_{14}$ ). In der alternativen Nomenklatur werden sie demnach als  $4_{18}$ -,  $2.5_{12}$ -, und  $3.25_{14}$ -Helices bezeichnet. Gebrochene „Hauptzahlen“ können dabei durchaus entstehen (Abbildung 4-2).



**Abbildung 4-2:** Zustandekommen von gebrochenen Indices bei Benennung von Helices nach der Anzahl der Aminosäurereste in einem Wasserstoffbrücken-gebundenen Ring.

Ein grundlegender Unterschied zwischen Helices in  $\alpha$ -Peptiden und denen in  $\beta$ -Peptiden besteht in der Ausrichtung des Dipolmoments. In allen Helixtypen in  $\alpha$ -Peptiden, die bisher gefunden wurden, sind die Wasserstoffbrücken in „rückwärtiger“ Weise, vom C- zum N-Terminus ausgebildet,<sup>2</sup> d. h. zwischen  $\text{NH}(i)$  und  $\text{CO}(i-n)$ . Beispielsweise liegen in  $\alpha$ -Helices Wasserstoffbrücken zwischen den CO-Gruppen in Position  $i-4$  und den NH-Gruppen in Position  $i$  vor, es existiert also eine  $5 \rightarrow 1$ -Interaktion<sup>[128]</sup>. In den meisten Helixtypen in  $\beta$ -Peptiden dagegen sind die Wasserstoffbrücken „nach vorne“ ausgebildet, also zwischen  $\text{NH}(i)$

<sup>(1)</sup> Dabei muss beachtet werden, dass eine solche Beschreibung immer idealisiert ist. In der Realität sind Helices keine starren Gebilde, weswegen ein Rest entlang der Helixachse nicht „exakt“ über einem anderen liegen kann. Weiterhin bedingt die Methodik der Bestimmung von Strukturen immer eine gewisse Unsicherheit bezüglich der Atompositionen. Die observierten bzw. berechneten Zahlenwerte werden daher immer gerundet.

<sup>(2)</sup> Eine Ausnahme ist das Gramacidin A, bei dem sich allerdings die helicale Struktur eher als Tertiärstruktur aus einer gewölbt  $\beta$ -Faltblatt-artigen Sekundärstruktur ergibt. Nach der im folgenden Text verwendeten Nomenklatur wäre es als  $20/22/20$ -helical zu beschreiben.

und  $\text{CO}(i+n)$ . Allerdings existieren auch Helices mit entgegengesetzter Orientierung. Zum Beispiel bestehen in der oben angeführten  $5_2$ -Helix Wasserstoffbrücken zwischen  $\text{NH}(i)$  und  $\text{CO}(i-3)$  ( $4 \rightarrow 1$ -Interaktion). Die rückwärtige Orientierung wurde immer dann beobachtet, wenn  $\text{C}^\alpha$  und  $\text{C}^\beta$  in einem Cyclopentanring gebunden sind.<sup>[118, 122, 126]</sup>

Von SEEBACH *et al.* wurde auch eine „heterogene“ 12/10/12-Helix beobachtet (s. u.), in der Wasserstoffbrücken-Ringe unterschiedlicher Größe bei alternierender Ausrichtung der Wasserstoffbrücken vorliegen.<sup>[121]</sup> In solchen Fällen kann das durch die Wasserstoffbrücken erzeugte Gesamt-Dipolmoment der Helix nahezu Null werden. Derartige Helices sollen im Folgenden gemäß ihrer Ringgrößen benannt werden. Eine Helix beispielsweise, in der der erste Ring die Größe  $C_{12}$  besitzt, und in der alternierend  $C_{10}$ - und  $C_{12}$ -Ringe auftreten, wird also als 12/10/12-Helix bezeichnet. Abbildung 4-3 zeigt abschließend die möglichen Wasserstoffbrückenbindungen in  $\beta$ -Peptiden.

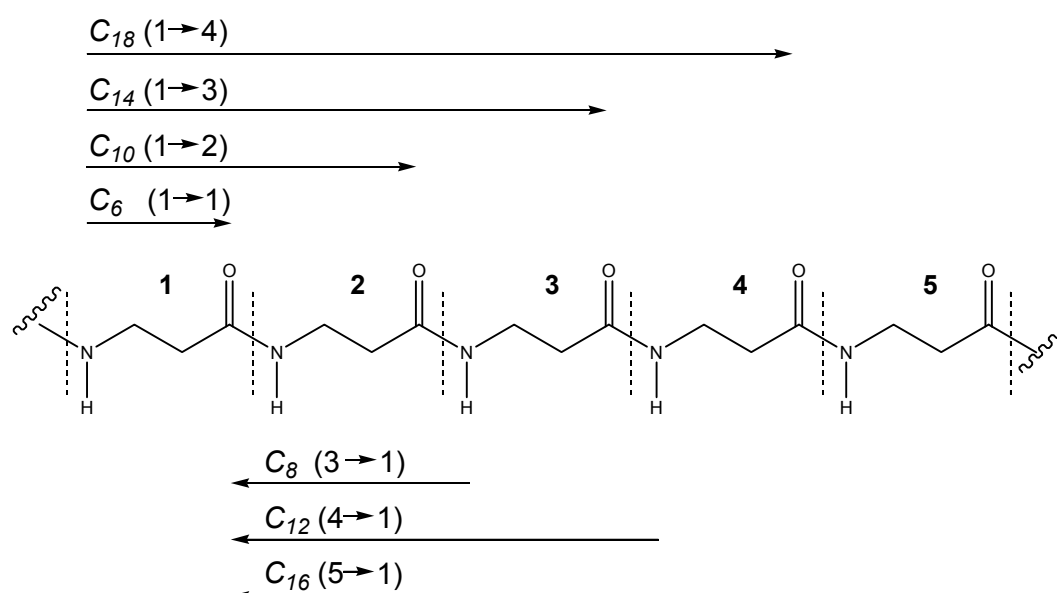


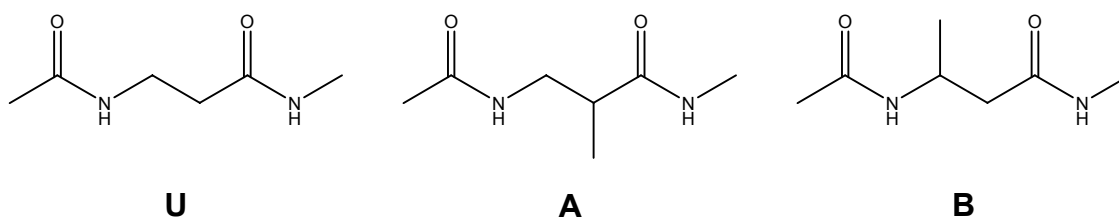
Abbildung 4-3: Möglichkeiten der Wasserstoffbrückenbindung in  $\beta$ -Peptiden.

#### 4.2.2 Quantenmechanische Vorhersage: Heterogene Helixtypen

In einer umfangreichen, theoretischen Studie untersuchte die Gruppe von HOFMANN *et al.* die grundlegenden Struktureigenschaften von  $\beta$ -Aminosäuren und deren Peptiden.<sup>[129]</sup> Hierzu berechneten sie eine Reihe kleinerer Moleküle meist mit quantenmechanischen Methoden.

Diese Untersuchungen<sup>1</sup> sollen im Folgenden kurz beschrieben werden, da sie die Richtigkeit der in dieser Arbeit ermittelten Struktur von Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH **5** nahe legen und hilfreich beim Ausschluss weiterer möglicher Konformationen waren.

Ausgangspunkt der Studie waren zunächst drei N- und C-terminal geschützte  $\beta$ -Aminosäuren-Monomere, die jeweils unsubstituiert (**U**), an C <sup>$\alpha$</sup>  substituiert (**A**), bzw. an C <sup>$\beta$</sup> -substituiert (**B**) sind (Abbildung 4-4).



**Abbildung 4-4:** Modellpeptide der Untersuchungen von HOFMANN et al.

Dabei wurden nur (S)-konfigurierte Substituenten eingesetzt, da sich die Ergebnisse der Rechnungen durch Symmetrie auf die jeweiligen Enantiomere übertragen lassen. Für jede der Verbindungen **U**, **A** und **B** wurden Startstrukturen erzeugt, indem die Winkel  $\theta$  jeweils in einem Raster von 60° variiert und für jedes  $\theta$  Permutationen mit ausgesuchten, wahrscheinlichen Kombinationen von  $\phi$  und  $\psi$  durchgeführt wurden (ebenfalls in einem Raster von 60°). Weitere Startstrukturen wurden molekülmechanisch durch eine Rastersuche nach Strukturen niedriger Energie im CHARMM-Kraftfeld<sup>[30]</sup> erzeugt, wobei  $\theta$ ,  $\phi$  und  $\psi$  in Schritten von 30° variiert wurden. Jede der Startstrukturen wurde dann mit dem Basissatz HF/3-21G der *ab initio* MO-Theorie berechnet. Die sich daraus ergebenden Strukturen wurden anschließend mit dem Basissatz HF/6-31G\* optimiert, wobei auch Solvenseffekte berücksichtigt wurden. Dabei hat sich herausgestellt, dass in allen drei Monomeren diejenigen Strukturen am stabilsten waren, in denen Wasserstoffbrücken-gebundene Ringe vorlagen. Das Energieminimum war jeweils ein C<sub>6</sub>-Konformer (also 6 Atome im Ring), gefolgt von einer Reihe anderer C<sub>6</sub>- und einiger weniger C<sub>8</sub>-Strukturen.

<sup>(1)</sup> Eine ähnliche Arbeit wurde von WU und WANG publiziert,<sup>[130]</sup> die zu den gleichen bzw. ähnlichen Ergebnissen über grundlegende Struktur Aspekte in  $\beta$ -Peptiden geführt hatte. Mittlerweile wurden eingehendere Untersuchungen über den Einfluss der Chiralität an C <sup>$\alpha$</sup>  und C <sup>$\beta$</sup>  durchgeführt.<sup>[131, 132]</sup> Eine neuere Zusammenfassung über heterogene Helices in  $\beta$ - als auch in  $\alpha$ -,  $\gamma$ - und  $\delta$ -Peptiden geben HOFMANN *et al.* in Referenz [133].



Daneben wurden auch Konformationen beobachtet, in denen die Winkelkombination zwar die Ausbildung von Wasserstoffbrücken-gebundenen Ringen  $H_i$  in einer Helix zulassen würde, letztere aber im Monomer nicht formiert werden kann (periodisch in Helices eingebundene Ringe  $C_i$  werden im Folgenden mit  $H_i$  bezeichnet). Weiterhin hatte sich gezeigt, dass die Substitution an  $C^\beta$  offenbar einen größeren Einfluss auf die Struktur hat als eine Substitution an  $C^\alpha$ .

Die Untersuchung oligomerer Einheiten der Modellpeptide **U**, **A** und **B** zeigte eine deutliche Abhängigkeit vom umgebenden Medium, die in ähnlicher Weise auch für Peptide beobachtet wird, die aus  $\alpha$ -Aminosäuren aufgebaut sind. Das tatsächliche Vorkommen von kleinen  $C_6$ - und  $C_8$ -Fragmenten der Monomere in Lösung (nicht jedoch im Kristall) ist allerdings fraglich.<sup>[123, 134]</sup> Dahingegen decken sich viele der Resultate von Berechnungen, die mit den Oligomeren durchgeführt wurden, mit experimentellen Beobachtungen. Ein gutes Beispiel dafür ist die periodische  $H_{12}$ -Helix, die ausschließlich aus  $C_{12}$ -Ringern aufgebaut ist. Diese Helix erwies sich bei den Berechnungen der Tri- und Tetramere von **B** in Lösung als äußerst stabil. Sie besitzt die rückwärtige Interaktion ( $4 \rightarrow 1$ ) und ist identisch mit der bereits erwähnten  $5_2$ -Helix, die in wässriger (d. h. polarer) Lösung immer dann beobachtet wird, wenn  $C^\alpha$  und  $C^\beta$  in einen Cyclopentanring eingebunden sind.<sup>[118, 122, 126]</sup> Für die Tri- und Tetrameren von **B** ergab sich weiterhin, dass auch  $H_{14}$ -Helices, welche unter weitgehender Konservierung der Rückgratwinkel aus  $C_{14}$ -Ringern aufgebaut sind, in wässriger Lösung ebenfalls sehr stabil sein sollten. Die treibende Kraft zur Ausbildung von Sekundärstruktur in den Oligomeren scheint nicht nur das umgebende Solvens, sondern auch die Optimierung der Ringgröße zu sein. Insgesamt ergab sich für die jeweiligen Oligomere eine große Anzahl möglicher Strukturen, die alle durch meist recht kleine Energiedifferenzen voneinander getrennt waren.

Wie bereits erwähnt, wurde bisher auch eine heterogene Helix beobachtet, die alternierend aus Wasserstoffbrücken-gebundenen Ringen unterschiedlicher Größe aufgebaut ist.<sup>[121]</sup> Für die gemischten Trimere Ac-**ABA**-NHMe und Ac-**BAB**-NHMe untersuchten HOFMANN *et al.* daher auch die Stabilität solcher inhomogener Helices und verglichen sie mit den jeweiligen homogenen Helixtypen. Für das Trimer Ac-**ABA**-NHMe ergab sich dabei eine deutliche Präferenz für die Anordnung  $C_{10}/C_{12}$ , also den Beginn einer 10/12/10-Helix, unabhängig davon, ob Solvenseinflüsse mitberücksichtigt wurden. Bereits wesentlich ungünstiger war die Anordnung  $C_{12}/C_{10}$ , deren Energie etwa  $18 \text{ kJ mol}^{-1}$  höher lag.

Diese Energiedifferenz wuchs auf  $46 \text{ kJ mol}^{-1}$  an, wenn Wasser als Solvens mitberücksichtigt wurde (implizit durch Einsatz des *self-consistent isodensity polarizable continuum model* SCI-PCM,<sup>[135]</sup> mit Dielektrizitätskonstante  $\epsilon = 78.4$ ). Das Trimer Ac-**BAB**-NHMe zeigte eine größere Vielfalt an energetisch günstigen Strukturen. Am stabilsten erwies sich die Anordnung  $C_{12}/C_{10}$ , die ohne Berücksichtigung von Solvenseinflüssen merklich stabiler als andere Strukturen war (Energiedifferenz größer als  $30 \text{ kJ mol}^{-1}$ ). Bei Berechnung mit Einfluss von Wasser als implizitem Lösungsmittel lagen die Strukturen  $C_{14}$  um  $4.5 \text{ kJ mol}^{-1}$ ,  $C_{12}/C_{12}$  um  $11.1 \text{ kJ mol}^{-1}$  und  $C_{10}/C_{12}$  um  $21.8 \text{ kJ mol}^{-1}$  über der Energie von  $C_{12}/C_{10}$ .

Die Ergebnisse der Untersuchungen von HOFMANN *et al.* wurden in dieser Arbeit herangezogen, um nach Nebenkonformationen der ermittelten Struktur von Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH (**5**) zu suchen.

### 4.3 Erstellung einer Solvenzelle für deuteriertes Acetonitril

Da Verbindung **5** (Abbildung 4-5, siehe weiter unten auf Seite 61) in gängigen Lösungsmitteln schlecht löslich war, wurden die NMR-Experimente in Acetonitril- $d_3$  durchgeführt. Für dieses Solvens (auch undeutert) bietet jedoch das zur Berechnung verwendete DISCOVER-Programmpaket keine Zelle an. In der *software*-Bibliothek unseres Arbeitskreises stand zwar eine  $C^1H_3CN$ -Zelle zur Verfügung, die allerdings nur in der *united-atom*-Darstellung vorlag und zudem undokumentiert war. Daher wurde im Rahmen dieser Arbeit eine explizite *all-atom*-Solvenzelle für  $CD_3CN$  aufgebaut. Das dabei angewendete Verfahren soll im Folgenden sowohl allgemein als auch im Hinblick auf  $CD_3CN$  beschrieben werden.

Zunächst muss der Platzbedarf eines einzelnen Lösungsmittelmoleküls berechnet werden, um eine Einheitszelle für ein einzelnes Molekül zu erstellen. Für  $CD_3CN$  war dies leicht möglich, da die Rotation um die  $H_3C-CN$ -Achse keine tatsächliche konformationelle Änderung bewirkt. Die einzelnen Abmessungen in den Dimensionen x, y und z hängen somit allein von der Orientierung im kartesischen Koordinatensystem ab. Das in dieser Arbeit verwendete Programm Insight II<sup>[136]</sup> stellt allerdings zur Abmessung der Dimensionen eine automatische Routine bereit. Für  $CD_3CN$  ergaben sich die Abmessungen  $1.71 \text{ \AA} \cdot 1.95 \text{ \AA} \cdot 3.06 \text{ \AA}$ . Der sich daraus ergebende Platzbedarf des Energie-minimierten Moleküls  $V_{Molekül,min.}$  ( $10.20 \text{ \AA}^3$  für  $CD_3CN$ ) ist jedoch geringer als derjenige eines realen Moleküls bei der Meßtemperatur,  $V_{Molekül,reell}$ . Dieser kann über die Dichte  $\rho$  und die molare Masse  $M_r$  ermittelt werden.

Für CD<sub>3</sub>CN ergab sich:

$$V_{\text{Molekül, reell}} = \frac{M_r}{\rho N_A} = \frac{44.07}{0.84 \cdot 6.023 \cdot 10^{23}} \text{ cm}^{-3} = 8.71 \cdot 10^{-23} \text{ cm}^{-3} = 87.11 \text{ \AA}^3 \quad [4-1].$$

Werden die Abmessungen des Energie-minimierten Moleküls mit dem Skalierungsfaktor  $(V_{\text{Molekül, reell}} / V_{\text{Molekül, min}})^{1/3}$  multipliziert, ergeben sich die tatsächlichen Abmessungen der Einheitszelle. Für CD<sub>3</sub>CN errechneten sich diese, mit dem Faktor von  $(87.11 / 10.20)^{1/3}$ , zu 3.50 Å · 3.99 Å · 6.25 Å. An dieser Stelle wurde in dieser Arbeit eine möglichst kubische Solvenszelle von etwa 30 Å Kantenlänge angestrebt<sup>1</sup>. Durch jeweiliges Teilen der Kantenlänge der Solvenszelle durch die Ausmaße der Einheitszelle in den einzelnen Dimensionen und Runden ergab sich als beste Lösung für das Arrangement von Einheitszellen:

$$\begin{aligned} 9 \cdot 6.25 \text{ \AA} &= 31.46 \text{ \AA} \\ 8 \cdot 3.99 \text{ \AA} &= 31.88 \text{ \AA} \\ 5 \cdot 3.50 \text{ \AA} &= 31.27 \text{ \AA} \end{aligned} \quad [4-2a-c].$$

Anschließendes Anordnen der Einheitszellen in den Quader der Abmessungen 31.46 Å · 31.88 Å · 31.27 Å führt zu einer Solvenszelle, die nun äquilibriert werden muss. Dies geschieht durch eine Dynamiksimulation bei hoher Temperatur, gefolgt von einer Abkühlphase auf die Solltemperatur, also im Allgemeinen auf die Messtemperatur. Dabei kann der Äquilibrierungszustand überwacht werden, indem das Gesamtdipolmoment der Zelle beobachtet wird. Eine perfekt äquilibrierte Zelle sollte theoretisch ein Gesamtdipolmoment von fast Null besitzen, was jedoch nur für unendlich große Zellen erreicht werden kann. Für Zellen wie der hier beschriebenen sind Werte bis etwa 10 D (Debye, 1 D = 3.33 · 10<sup>-30</sup> C m) akzeptabel, für stark polare Lösungsmittel auch Werte bis 20 D. Als Vergleich sei hier das Dipolmoment eines einzelnen CH<sub>3</sub>CN-Moleküls von 3.92 D angeführt.<sup>[137]</sup>

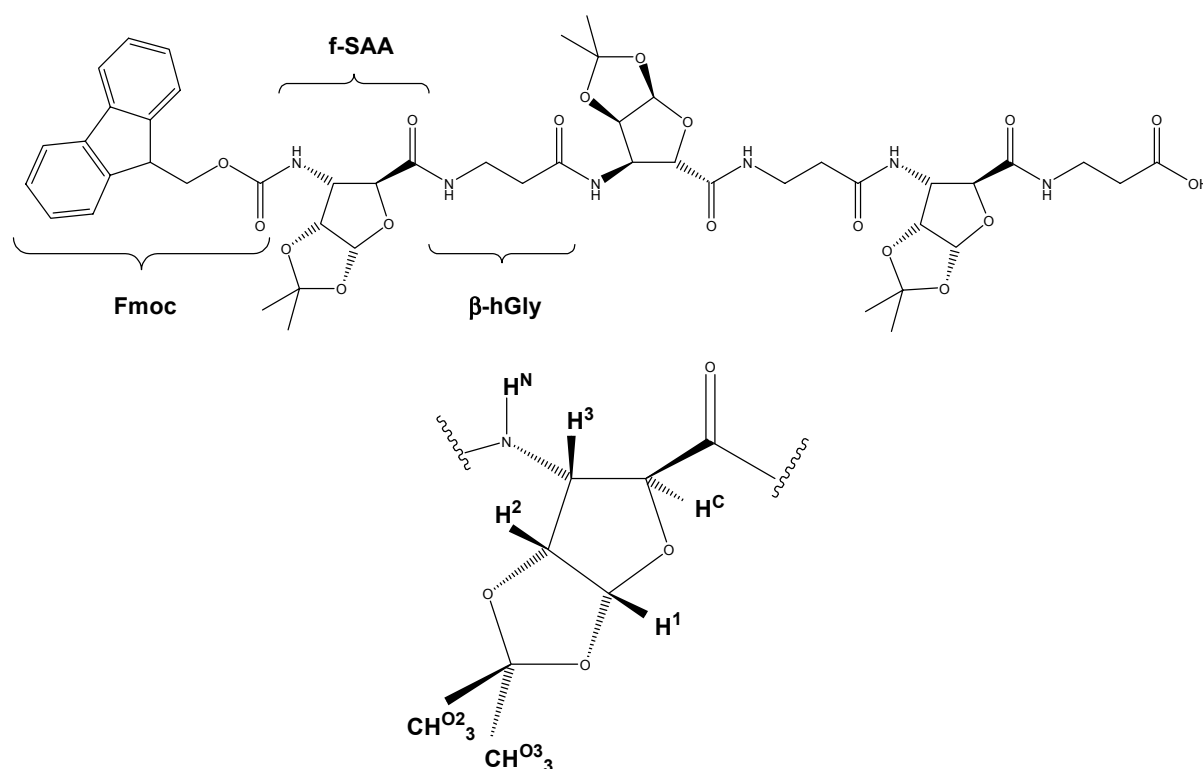
Vor der eigentlichen Äquilibrierung wurde zunächst eine kurze Energieminimierung durchgeführt, um Störstellen zu entfernen (3000 Schritte *steepest descent*, dann 10000 Schritte *conjugate gradients*, Abbruchkriterium:  $(\partial E / \partial r_i)_{\text{max}} < 0.2 \text{ kcal \AA}^{-1}$ ). Den Atomen wurden dann zufällig Geschwindigkeiten gemäß der BOLTZMANN-Verteilung bei 900 K

<sup>(1)</sup> Das zu berechnende Molekül erwies sich zwar recht bald als näherungsweise quaderförmig. Soll jedoch eine generisch anwendbare Solvenszelle erstellt werden, so ist im Allgemeinen eine möglichst kugel- bzw. würfelförmige Geometrie wünschenswert.

zugewiesen, und bei dieser Temperatur 150 ps Dynamiksimulation durchgeführt. Dabei sank das Gesamtdipolmoment der Zelle von 285.01 D auf 3.74 D. Die Temperatur wurde danach während 45 ps Simulationszeit exponentiell auf 298 K gesenkt. Dabei stieg das Dipolmoment leicht auf 4.98 D an, da bei tieferen Temperaturen ein höherer Ordnungszustand vorliegt. Um abschließend etwaige Störstellen zu entfernen, wurden weitere 100 Schritte *conjugate-gradients*-Minimierung angeschlossen. Das Gesamtdipolmoment der fertigen Solvenszelle betrug 7.75 D.

#### 4.4 Strukturbestimmung von Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH

Das in dieser Arbeit berechnete Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH (**5**) ist ein Trimer eines Bausteins, der peptidisch aus einer furanoiden  $\beta$ -Zuckeraminosäure und  $\beta$ -Homoglycin zusammengesetzt ist (Abbildung 4-5). Sowohl der Baustein als auch **5** wurden von S. GRUNER im Rahmen ihrer Doktorarbeit dargestellt.<sup>[138]</sup>



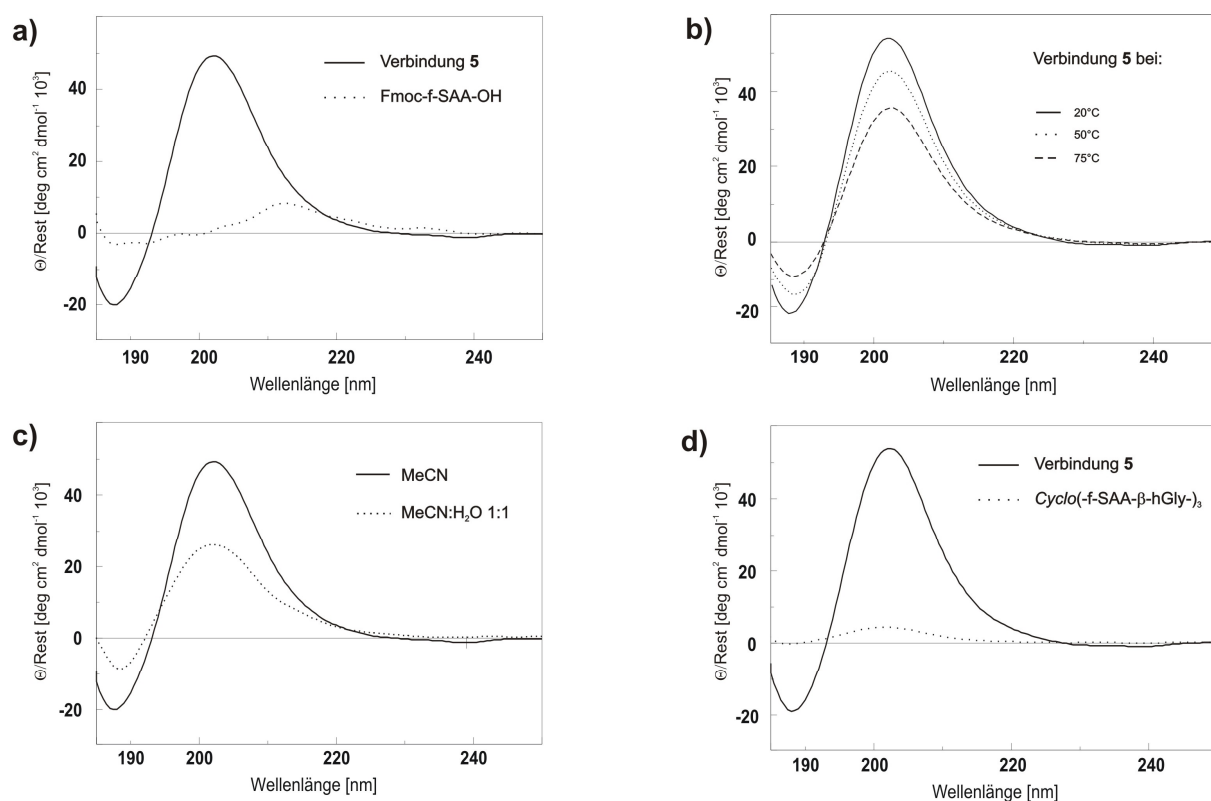
**Abbildung 4-5:** Strukturformel von Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH (**5**) mit den für f-SAA verwendeten Bezeichnungen der Protonen.

Wegen der zwei konsekutiven Methylengruppen in  $\beta$ -hGly und der sich daraus ergebenden Flexibilität wurde zuerst erwartet, dass Verbindung **5** keine geordnete Struktur einnehmen würde. Auch hatten Beobachtungen von SEEBACH *et al.* gezeigt, dass schon der Einbau eines einzigen  $\beta$ -hGly-Rests die Ausbildung von  $3_1$ -Helices in  $\beta$ -Hexapeptiden stören kann.<sup>[119]</sup> Es zeigte sich jedoch bald, dass **5** nicht vollkommen flexibel sein konnte.

#### 4.4.1 CD-spektroskopische Untersuchungen

Einen ersten Hinweis auf die Ausbildung einer definierten Struktur in Verbindung **5** lieferten von S. GRUNER durchgeführte CD-spektroskopische Untersuchungen.<sup>[138]</sup> Aus dem CD-Spektrum einer Verbindung kann nicht nur in  $\alpha$ -Peptiden, sondern auch in  $\beta$ - und  $\gamma$ -Peptiden auf das Vorkommen bestimmter Sekundärstrukturelemente geschlossen werden.<sup>[117, 126, 127, 139-145]</sup> Um einen direkten Vergleich von Oligomeren mit einzelnen Bausteinen zu ermöglichen, wurden die gemessenen molaren Elliptizitäten  $\Theta$  durch die Anzahl der Reste geteilt.

Das CD-Spektrum von Verbindung **5** bei 20 °C besitzt ein tiefes Minimum bei 188 nm, einen Nulldurchgang bei 193 nm und ein ausgeprägtes Maximum bei 203 nm. Der Vergleich mit dem CD-Spektrum von Fmoc-f-SAA-OH (Abbildung 4-6a) zeigt, dass die beobachtete residuale molare Elliptizität  $\Theta$  nicht von Chromophoren innerhalb der Zuckeraminosäure stammt, sondern kooperativ durch die Orientierung der Amid-Chromophore im Peptid-Rückgrat zustande kommt.



**Abbildung 4-6:** a) CD-Spektren von  $\text{Fmoc-}[\text{-f-SAA-}\beta\text{-hGly}]_3\text{-OH } \mathbf{5}$  (20 °C, 0.125 mM) und von  $\text{Fmoc-f-SAA-OH}$  (20 °C, 0.5 mM) in  $\text{CH}_3\text{CN}$ . b) CD-Spektrum von  $\mathbf{5}$  (0.125 mM) in  $\text{CH}_3\text{CN}$  bei 20, 50 und 75 °C. c) CD-Spektren von  $\mathbf{5}$  (20 °C, 0.125 mM bis 1 mM) in  $\text{CH}_3\text{CN}$  und in  $\text{CH}_3\text{CN}/\text{H}_2\text{O}$  (1:1). Das Spektrum verändert sich bei Temperaturänderung wie in b) gezeigt, bleibt aber konzentrationsunabhängig. d) CD-Spektren von  $\mathbf{5}$  (20 °C, 0.125 mM, in  $\text{CH}_3\text{CN}$ ) und seinem cyclischen Analogon  $\text{cyclo}(\text{-f-SAA-}\beta\text{-hGly-})_3$  bei gleichen Bedingungen.

Die Temperaturabhängigkeit des CD-Spektrums von  $\mathbf{5}$  legt ebenfalls eine charakteristische Sekundärstruktur nahe. Für eine flexible Verbindung ohne strukturelle Präferenzen sollte das Spektrum unabhängig von der Temperatur sein. Mit zunehmender Temperatur nimmt jedoch die Intensität der Extrema in der CD-Kurve von  $\mathbf{5}$  ab (Abbildung 4-6b). Interessanterweise führt der Wechsel vom recht unpolaren Acetonitril zum stärker polaren Lösungsmittelgemisch  $\text{CH}_3\text{CN}/\text{H}_2\text{O}$  (1:1) ebenfalls zu einem Rückgang in der Intensität der Extremstellen des Spektrums (Abbildung 4-6c). Dieser Befund schien darauf hinzudeuten, dass die von  $\mathbf{5}$  angenommene Sekundärstruktur von unpolaren Lösungsmitteln stabilisiert wird, in polaren Lösungsmitteln jedoch energetisch ungünstiger ist. Da die Kurven in  $\text{CH}_3\text{CN}$  bzw. im Solvengemisch bei Temperaturen zwischen 20 und 75 °C unabhängig von der Konzentration bleiben (zwischen 0.125 und 1 mM), kann davon ausgegangen werden, dass es zu keiner

Aggregation von **5** kommt. Einen letzten Hinweis auf eine stabile Sekundärstruktur lieferte der Vergleich mit dem CD-Spektrum des entschützten und zyklisierten **5** (Abbildung 4-6d). Unabhängig von der Wellenlänge des eingestrahlten Lichts ist die residuale, molare Elliptizität von *cyclo*(-f-SAA- $\beta$ -hGly-)<sub>3</sub> annähernd Null. Dies weist darauf hin, dass die Zyklisierung eine im großen Maßstab erfolgende Umorientierung der Amid-Chromophore bewirkt.

Das CD-Spektrum von **5** wurde bisher bei keinem anderen  $\beta$ -Peptid beobachtet. Allerdings ist bisher nicht bekannt, wie sich das Substitutionsmuster und chirale Zentren im Peptidrückgrat auf die molare Elliptizität  $\Theta$  auswirken. Während die CD-spektroskopischen Untersuchungen also einen eindeutigen Hinweis auf eine stabile helicale Sekundärstruktur in **5** lieferten, konnte an dieser Stelle keine der bekannten Strukturen von  $\beta$ -Peptiden postuliert oder ausgeschlossen werden.

#### 4.4.2 NMR-Strukturdaten

Die Aufnahme und Auswertung von TOCSY-, COSY-, HMQC-, HMBC- und NOESY-Spektren führte S. GRUNER im Rahmen ihrer Doktorarbeit durch. Das Auswerten des in dieser Arbeit für die Berechnung von Strukturen verwendeten NOESY-Spektrums wurde von V. TRUFFAULT vorgenommen.

Insgesamt wurden zunächst vier NOESY-Spektren mit Mischzeiten von 100, 200, 400 und 600 ms in CD<sub>3</sub>CN bei 300 K gemessen. Die Spektren bei 100 und 200 ms Mischzeit boten allerdings keine ausreichend intensiven Kreuzsignale zur Strukturbestimmung. Im Spektrum mit 400 ms Mischzeit befand sich die Aufbaurate der meisten NOEs noch im linearen Bereich. Dieses Spektrum wurde daher zur Berechnung der Struktur verwendet. Während der Verfeinerung wurden einige der 76 integrierten Kreuzsignale bei eindeutigen Fällen von Spindiffusion manuell ausgeglichen. Die Integrale wurden unter Anwendung der isolierten Zweispin-Näherung (ISPA)<sup>[36]</sup> gemäß Gleichung 2-24 in Distanzen umgerechnet. Als Referenz dienten die Kreuzsignale  $\beta$ -hGly<sup>2</sup>H <sup>$\alpha$ , pro R</sup>- $\beta$ -hGly<sup>2</sup>H <sup>$\alpha$ , pro S</sup> und  $\beta$ -hGly<sup>4</sup>H <sup>$\alpha$ , pro R</sup>- $\beta$ -hGly<sup>4</sup>H <sup>$\alpha$ , pro S</sup>, die miteinander überlagert, ansonsten jedoch separiert waren. Das Gesamt-Integral wurde daher durch 2 geteilt und dem Referenzabstand von 1.78 Å zugeordnet. Durch Versehen mit einer Toleranz von  $\pm 10\%$  wurden aus den Distanzen Ober- und Untergrenzen als *restraints* für die Strukturrechnungen gewonnen. Das NOESY-Spektrum ermöglichte die stereospezifische Zuordnung aller prochiralen Methylgruppen der f-SAA-Reste. Weiterhin konnten in allen  $\beta$ -hGly-Resten, außer im C-terminalen  $\beta$ -hGly<sup>6</sup>, alle H <sup>$\alpha$</sup> - und H <sup>$\beta$</sup> -Protonen

stereospezifisch zugeordnet werden. Für die Methylgruppen der f-SAA-Reste und für die H <sup>$\alpha$</sup> - und H <sup>$\beta$</sup> -Protonen in  $\beta$ -hGly<sup>6</sup> wurden Pseudoatome verwendet. Die Obergrenzen entsprechender *restraints* wurden mit der jeweils erforderlichen Standard-Pseudoatomkorrektur<sup>[103]</sup> versehen.

#### 4.4.3 Strukturrechnungen

Die globale Suche im Konformationsraum wurde mittels *simulated annealing* mit dem XPLOR-Programmpaket<sup>[31]</sup> unter Verwendung von Standard-Protokollen durchgeführt, wobei ein Ensemble von 10 Strukturen erhalten wurde.<sup>1</sup> Im zuletzt berechneten Ensemble zeigten 9 von 10 Strukturen gute Konvergenz (mittlerer RMSD zur gemittelten Struktur 0.37 Å). Die abweichende Struktur, die viele und hohe Verletzungen der *restraints* aufwies (bis 0.7 Å) und nach chemischem Verständnis sehr unwahrscheinlich erschien, wurde verworfen. Aus diesem Ensemble wurde die Struktur, die am besten im Einklang mit den experimentellen Daten stand, als Startstruktur für die weitere Verfeinerung ausgewählt.

Die Struktur besaß bereits eine merkliche, rechtsgängig helicale Tendenz mit einer alternierenden Anordnung 12- und 10-gliedriger Wasserstoffbrücken-gebundener Ringe. Allerdings waren die lokalen Geometrien teilweise noch recht stark verzerrt, wie es oft in SA-Strukturen beobachtet wird. Daher wurde eine rMD-Verfeinerung im expliziten Solvens angeschlossen, die im CVFF-Kraftfeld mit dem Programm DISCOVER in der Version 2.9.8<sup>[25]</sup> durchgeführt wurde. Als *restraints* wurden 76 Distanz-Informationen aus dem NOESY-Spektrum eingesetzt. Die SA-Struktur wurde zunächst zur Beseitigung stärkerer Verzerrungen kurz Energie-minimiert (300 Schritte *steepest descent*) und in eine kubische Solvenszelle der Kantenlänge 47 Å überführt. Die Zelle wurde dann mit 1039 CD<sub>3</sub>CN-Molekülen gefüllt, wobei die im Rahmen dieser Arbeit erstellte Solvenszelle (Kapitel 4.3) als Templat diente. Analog zu dem bereits in Kapitel 3 beschriebenen Verfahren wurde die gesamte Zelle in zwei Stufen minimiert, danach langsam auf die Temperatur von 300 K aufgeheizt und bei dieser Temperatur während 25 ps äquilibriert. Für diese und weitere Dynamiksimulationen kam das *explicit image* periodischer Randbedingungen zum Einsatz. Während der darauf folgenden Dynamik von 150 ps Dauer wurden im Zeitraster von 100 fs die Koordinaten in einer Trajektorie gespeichert. Die 1500 Strukturen der Trajektorie wurden

---

<sup>(1)</sup> Die erste Suche im Konformationsraum wurde von V. TRUFFAULT durchgeführt. Für ein Molekül in der Größenordnung des hier beschriebenen kann die Berechnung von nur 10 Strukturen durchaus zum Verfehlen einer tatsächlich eingenommenen Konformation führen. Im Rahmen der ersten Verfeinerung wurden allerdings mehr als 20 Iterationen durchgeführt, die im Wesentlichen das gleiche Ergebnis zeigten. Von einer ausreichenden Suche im Konformationsraum kann daher ausgegangen werden.

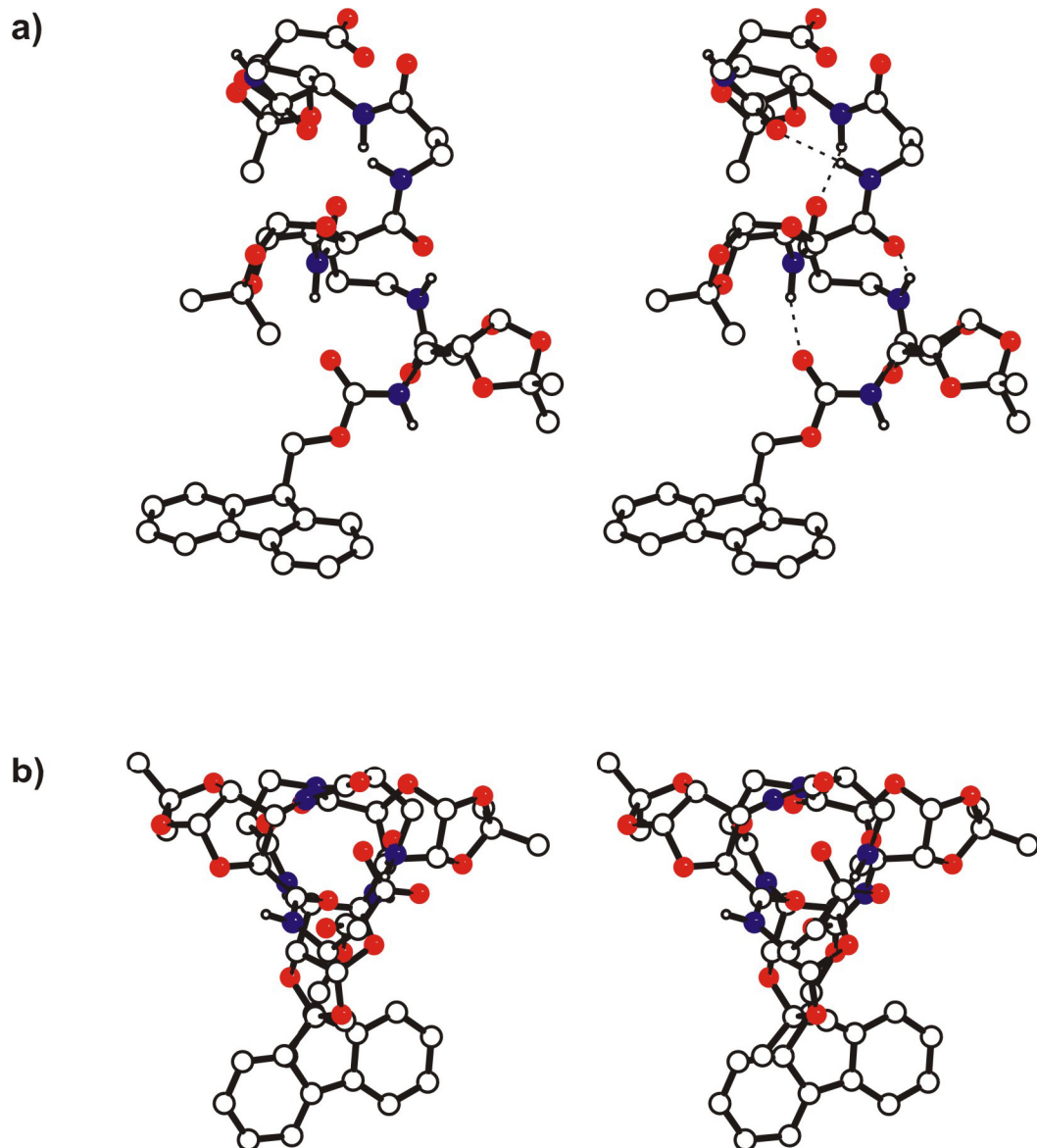


gemittelt und die daraus resultierende Struktur zur Beseitigung von sich aus der Mittelung ergebenden lokalen Verzerrungen vorsichtig Energie-minimiert (300 Schritte *steepest descent*).

Um die Stabilität der ermittelten Struktur zu überprüfen, wurde weiterhin eine Langzeit-fMD-Simulation von 3 ns Länge angeschlossen, die ebenfalls im expliziten Lösungsmittel durchgeführt wurde.

#### 4.4.4 Ergebnisse und Diskussion

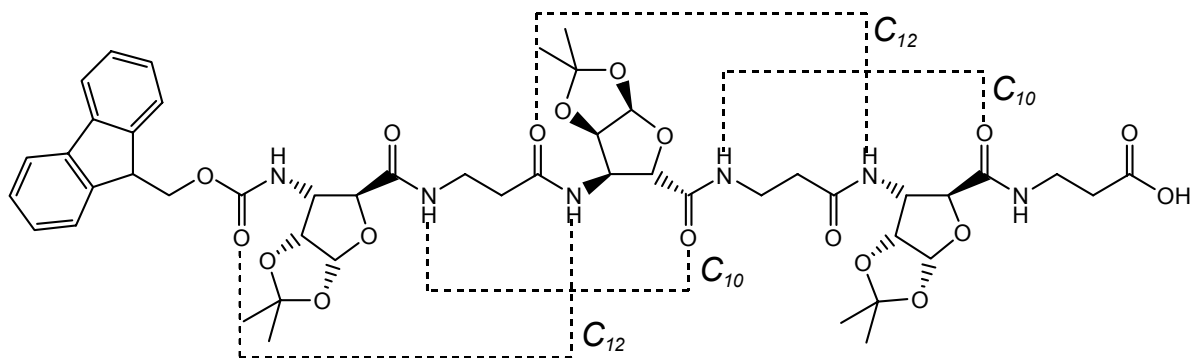
Der Einbau von  $\beta$ -hGly in  $\beta$ -Peptide hatte sich in der Vergangenheit als destabilisierend auf die Ausbildung von  $3_1$ -Helices erwiesen.<sup>[119]</sup> Wie bereits erwähnt, wäre es wegen der hohen Flexibilität dieses Aminosäurerests durchaus zu erwarten, dass Verbindung **5** keinerlei definierte Konformationen einnimmt. Interessanterweise zeigte sich jedoch, dass die Verbindung in CD<sub>3</sub>CN weitgehend in nur einer wohldefinierten Konformation vorliegt (Abbildung 4-7a,b).



**Abbildung 4-7:** Stereodarstellung der gemittelten und mit 300 Schritten steepest descent minimierten Struktur der 150 ps rMD-Simulation von Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH (5) in CD<sub>3</sub>CN: **a)** Seitenansicht, **b)** Draufsicht auf die Helixachse. Der N-Terminus ist jeweils unten bzw. hinten. Wasserstoffbrückenbindungen sind gestrichelt eingezeichnet. Sauerstoffatome sind in rot, Stickstoffatome in blau, Kohlenstoffatome in weiß und Amidprotonen (kleine Kugeln) ebenfalls in weiß dargestellt. Die Graphik wurde mit dem Programm Schakal 99 erstellt.<sup>[104]</sup>

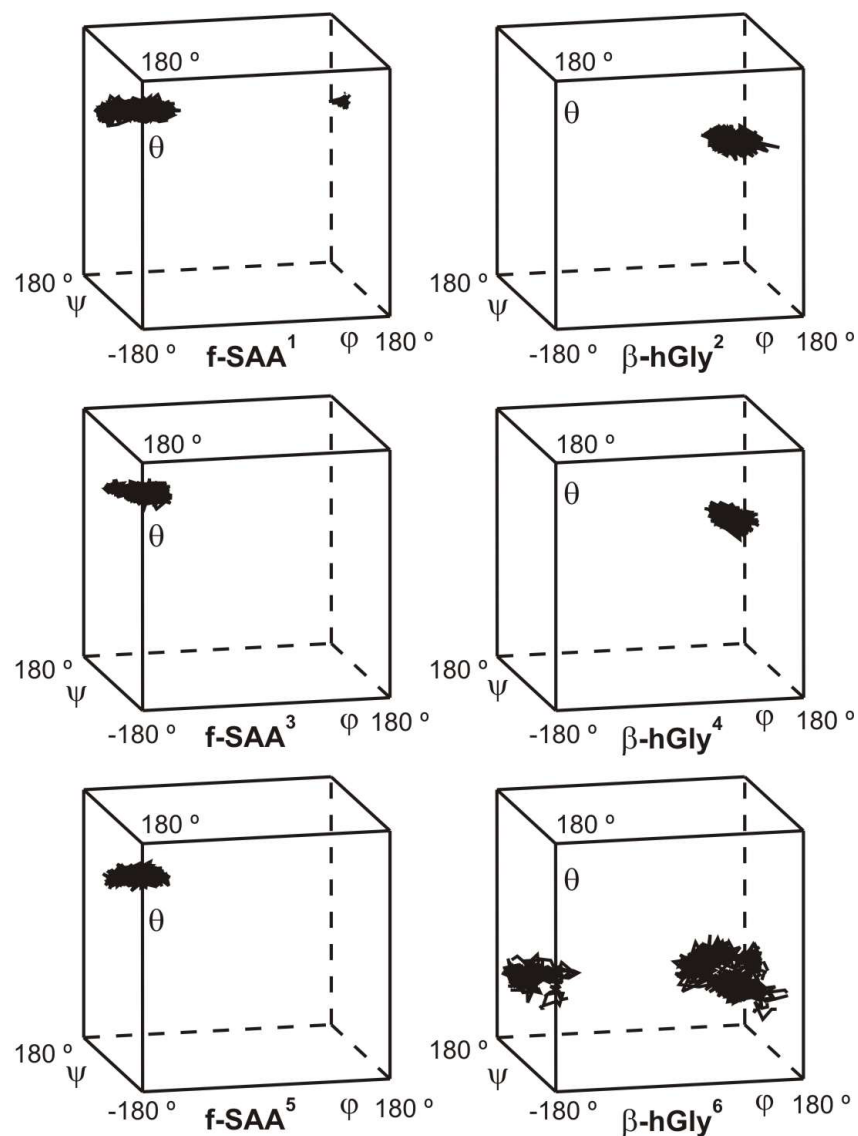
Die ermittelte Struktur ist rechtsgängig helical, wobei entlang der Helixachse die Seitenketten eines Rests  $i+3$  annähernd über denjenigen des Rests  $i$  liegen. Da zwischen diesen Resten eine Windung durchlaufen wird, könnte nach oben angeführter Nomenklatur ebenfalls die

Bezeichnung  $3_1$ -Helix verwendet werden (allerdings würde dies zur Verwechslung mit der bezüglich ihrer Wasserstoffbrücken-gebundenen Ringe homogen aufgebauten, bereits bekannten  $3_1$ -Helix führen). Die C=O- bzw. N-H-Bindungen sind in alternierender Reihenfolge entlang der Helixachse ausgerichtet. Beginnend bei der Carbonylgruppe des Fmoc-Rests sind ebenfalls alternierend Wasserstoffbrücken-gebundene Ringe der Größen  $C_{12}$  und  $C_{10}$  ausgebildet, es liegt also, anders benannt, die bereits angeführte 12/10/12-Helix vor. Insgesamt existieren 4 Wasserstoffbrücken, die schematisch in Abbildung 4-8 dargestellt sind. Die räumliche Nähe der Carboxylgruppe des C-terminalen  $\beta$ -hGly<sup>6</sup> zum Amidproton von  $\beta$ -hGly<sup>4</sup> (also eine eventuelle fünfte, gegabelte Wasserstoffbrücke) ergab sich dabei rein aus der Kraftfeldsimulation, da zwischen diesen beiden Aminosäuren keine *restraints* wirkten.



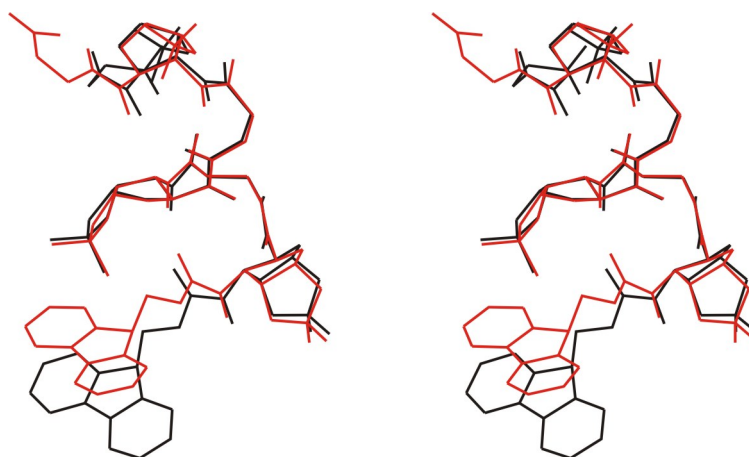
**Abbildung 4-8:** Alternierende Wasserstoffbrücken-gebundene Ringe  $C_{12}$  und  $C_{10}$  in der 12/10/12-helicalen Konformation von **5** in  $CD_3CN$ .

Während der 150 ps Dynamik-Simulationszeit blieben die verwendeten *restraints* weitgehend unverletzt. Alle Abweichungen lagen unterhalb 0.1 Å, bis auf einen *restraint* zwischen  $\beta$ -hGly<sup>6</sup>H<sup>N</sup> und f-SAA<sup>5</sup>H<sup>3</sup>, der um 0.31 Å verletzt wurde. Insgesamt scheint der C-Terminus eher etwas flexibel zu sein, was sich auch in den dreidimensional erweiterten RAMACHANDRAN-Diagrammen der einzelnen Aminosäurereste ausdrückt. Die Rückgratwinkel von  $\beta$ -hGly<sup>6</sup> zeigen dabei deutlich mehr Beweglichkeit als diejenigen der anderen Reste (Abbildung 4-9).



**Abbildung 4-9:** Räumlich erweiterte RAMACHANDRAN-Diagramme: Bewegung der Rückgratwinkel  $\phi$ ,  $\theta$  und  $\psi$  in den Aminosäuren von 5 während der 150 ps rMD-Simulation in  $CD_3CN$ . Die vordere linke untere Ecke der Würfel entspricht  $-180^\circ$  für alle drei Winkel.

Während der Langzeit-fMD-Simulation blieb die ermittelte Struktur bemerkenswert stabil. Allerdings kam es bereits während der ersten 150 ps zu deutlichen Verletzungen von fünf *restraints* (zwischen 0.14 und 0.46 Å). Drei der Verletzungen entstanden dabei durch die Einnahme einer kompakteren Faltung im Kernbereich der Helix (f-SAA<sup>1</sup>H<sup>1</sup> - f-SAA<sup>3</sup>H<sup>C</sup>; f-SAA<sup>1</sup>H<sup>3</sup> - f-SAA<sup>3</sup>H<sup>C</sup> und  $\beta$ -hGly<sup>2</sup>H<sup>N</sup> - f-SAA<sup>3</sup>H<sup>C</sup>). Dagegen legen der bereits in der rMD verletzte *restraint*  $\beta$ -hGly<sup>6</sup>H<sup>N</sup> - f-SAA<sup>5</sup>H<sup>3</sup> sowie ein verletzter Abstand zwischen  $\beta$ -hGly<sup>6</sup>H<sup>N</sup> und f-SAA<sup>5</sup>H<sup>3</sup> ein leichtes Entfalten am C-Terminus der Helix während der fMD nahe. Dies zeigt sich auch bei Superpositionierung der gemittelten und minimierten fMD- und rMD-Strukturen. Wird dazu als Atomsatz die Menge aller Rückgrat-Schweratome einschließlich der Carbonylsauerstoffatome, beginnend bei Fmoc-CO, bis zu  $\beta$ -hGly<sup>6</sup>CO, sowie der Seitenkettenatome aller f-SAA in  $\beta$ -Position zum Rückgrat verwendet, so ergibt sich ein RMSD von 0.90 Å. Dieser Wert sinkt drastisch auf 0.22 Å, wenn O, C <sup>$\alpha$</sup>  und C <sup>$\beta$</sup>  von  $\beta$ -hGly<sup>6</sup> ausgelassen werden. Die derartig superpositionierten Strukturen sind in Abbildung 4-10 dargestellt:



**Abbildung 4-10:** *Verbindung 5: Superpositionierung der über die ersten 150 ps Simulationszeit gemittelten und minimierten fMD-Struktur auf die gemittelte und minimierte rMD-Struktur von 5 in CD<sub>3</sub>CN. Schwarz: rMD-Struktur, rot: fMD-Struktur.*

Einen weiteren Hinweis auf die Stabilität der ermittelten Struktur liefert die Population der vier Wasserstoffbrücken in der helicalen Konformation von **5** in CD<sub>3</sub>CN (Tabelle 4-1):

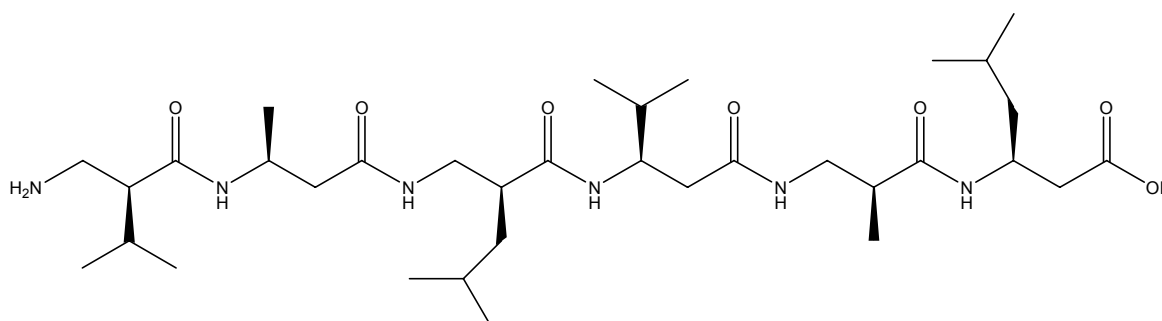
**Tabelle 4-1:** Durchschnittliche Abstände und Winkel der vier Wasserstoffbrücken von Verbindung **5** während der 150 ps rMD-Simulation in CD<sub>3</sub>CN, sowie deren Populationen.

Akzeptor–Donor-Paar	$\langle r_{N-O} \rangle$ [Å]	$\langle r_{H-O} \rangle$ [Å]	$\langle \angle_{N-O} \rangle$ [°]	Population [%]
Fmoc CO – f-SAA <sup>3</sup> HN	3.77	2.94	142.5	52.2
f-SAA <sup>3</sup> CO – $\beta$ -hGly <sup>2</sup> HN	3.70	2.86	142.6	41.0
$\beta$ -hGly <sup>2</sup> CO – f-SAA <sup>5</sup> HN	3.30	2.34	158.0	85.5
f-SAA <sup>5</sup> CO – $\beta$ -hGly <sup>4</sup> HN	3.52	2.62	149.9	65.5

Wässrige Lösungen von  $\alpha$ -Helices in  $\alpha$ -Peptiden bleiben in fMD-Simulationen im Allgemeinen nur kurz stabil und werden meist binnen weniger hundert Picosekunden entfaltet<sup>[146-149]</sup> (eine Ausnahme sind Polyalanine, die deutlich länger im helicalen Zustand verbleiben)<sup>[150]</sup>. Auch in unpolaren Lösungsmitteln werden diese Helices früher oder später aufgelöst. Beispielsweise wurde in unserem Arbeitskreis von R. GRATIAS eine Langzeit-fMD-Simulation der Helix A des Bakteriorhodopsins in einer von ihm entwickelten Methanol/Chloroform-Solvenczelle durchgeführt.<sup>[151]</sup> Dabei kam es, verglichen mit Simulationen im wässrigen Medium, nach ca. 700 ps relativ spät zu ersten Auflösungserscheinungen. Auch nach weiteren 550 ps, am Ende der Simulationszeit, verblieb die Tendenz in Richtung einer weiteren Zerstörung der Helix. Allerdings fanden die Auflösungserscheinungen im Bereich Gly<sup>21</sup>-Leu<sup>22</sup>-Gly<sup>23</sup> statt, womit sich die bekanntermaßen Helix-brechende Wirkung von Glycin<sup>[152]</sup> bestätigte. Wie oben bereits beschrieben, sollte sich die  $\beta$ -Aminosäure  $\beta$ -hGly erwartungsgemäß analog Helix-zerstörend zur  $\alpha$ -Aminosäure Glycin verhalten.<sup>[119]</sup> In Übereinstimmung mit den NMR-spektroskopischen Messungen, die eine hohe Stabilität der Konformation von **5** nahe legen, kam es jedoch erst nach ca. 1.7 ns fMD-Simulationszeit zu einer kompletten Auflösung der Wasserstoffbrücke zwischen Fmoc-CO und  $\beta$ -SAA<sup>3</sup>HN. Der Abstand wuchs dann schnell auf 9.4 Å an, was einer nahezu linearen Kettenanordnung und damit vollkommenen Entfaltung des Rückgrats zwischen Fmoc und f-SAA<sup>3</sup> entspricht. Bemerkenswerterweise fand jedoch bereits nach weiteren 400 ps Simulationszeit eine spontane Neuformierung der entsprechenden H-Brückenbindung statt.

Bis zum Ende der Simulationszeit von 3 ns blieb die Konformation anschließend rein 12/10/12-helical. Selbst während der kurzen N-terminalen Entfaltung blieb der Kernbereich der Helix intakt. Ähnliche Beobachtungen wurden von SEEBACH *et al.* bei der Langzeitsimulation von H- $\beta$ -hVal- $\beta$ -hAla- $\beta$ -hLeu-(S,S)- $\beta$ -hAla( $\alpha$ Me)- $\beta$ -hVal- $\beta$ -hAla- $\beta$ -hLeu-OH gemacht, das in Methanol eine  $3_1$ -helicale Konformation einnimmt.<sup>[120]</sup> Bei plötzlichem Erhöhen der Temperatur von 298 auf 350 K nach einer Simulationszeit von 200 ps geht zunächst die helicale Struktur verloren, allerdings bildet sich die Helix erneut nach weiteren 400 ps. Die Ergebnisse der Langzeitsimulation von Verbindung **5** legen also eine ähnliche, äußerst hohe Stabilität der  $3_1$ - und der 12/10/12-Helices in  $\beta$ -Peptiden nahe.

Bis zum Abschluss der Arbeiten an **5** war erst eine Verbindung mit ähnlicher Struktur bekannt. Die von SEEBACH *et al.* synthetisierte und mittels NMR aufgeklärte Verbindung **6**<sup>[121]</sup> ist jedoch recht unterschiedlich zu Verbindung **5** aufgebaut (Abbildung 4-11):



**Abbildung 4-11:** Strukturformel des von SEEBACH *et al.* dargestellten und NMR-spektroskopisch aufgeklärten  $\beta$ -Hexapeptids **6**.

In Verbindung **6** alternieren jeweils  $\beta$ -Aminosäurereste, die abwechselnd  $\alpha$ (S)- und  $\beta$ (S)-substituiert sind, welches dem Substitutionsmuster **BA**<sub>3</sub> nach der Nomenklatur von HOFMANN *et al.* entspricht.<sup>1 [129]</sup> Dahingegen ist die Sequenz von **5** abwechselnd aus  $\alpha$ , $\beta$ -disubstituierten und unsubstituierten Resten aufgebaut. Befreit von der Prioritätsreihenfolge der Nomenklatur nach IUPAC (durch Ersetzen jedes Substituenten durch eine Methylgruppe) ist f-SAA ( $\alpha$ S, $\beta$ S)-substituiert, woraus sich ein Substitutionsmuster [**S,S-U**]<sub>3</sub> ableitet. Wie bereits erwähnt, hatte sich jedoch in den Untersuchungen von HOFMANN *et al.* eine merkliche

<sup>(1)</sup> Verbindung **6** ist zwar N-terminal  $\alpha$ -substituiert, jedoch nicht geschützt. Im Gegensatz dazu waren die von HOFMANN *et al.* untersuchten Peptide N-terminal acyliert, was das Vertauschen von A und B nötig macht, um einen sinnvollen Vergleich ziehen zu können.

Dominanz des Einflusses des  $\beta$ -Substituenten gegenüber derjenigen des  $\alpha$ -Substituenten ergeben. Eine ( $\alpha$ S, $\beta$ S)-disubstituierte Aminosäure wie f-SAA sollte sich demnach eher wie eine ausschließlich **B**-substituierte Aminosäure verhalten, insbesondere auch, da durch die Cyclisierung zwischen  $C^\alpha$  und  $C^\beta$  eine Einschränkung der konformationellen Freiheit erfolgt (eine Analogie zum bekanntermaßen Helix-induzierenden Prolin ist für f-SAA durchaus gegeben). Im **U**-substituierten in  $\beta$ -hGly fehlt demgegenüber der in starkem Maße strukturbestimmende Einfluss eines  $\beta$ -Substituenten, womit sich dieses am ehesten mit einer **A**-substituierten Aminosäure vergleichen lässt. Insgesamt ergibt sich also für Verbindung **5** das zu Verbindung **6** analoge Substitutionsmuster **BA**<sub>3</sub>, wodurch erklärt werden kann, dass auch Verbindung **5** die Konformation einer 12/10/12-Helix annimmt.

Weiterhin hatte sich in den Untersuchungen von HOFMANN *et al.* gezeigt, dass die 12/10/12-Helix insbesondere in unpolaren Lösungsmitteln sehr stabil sein sollte. Dies kann darauf zurückgeführt werden, dass in dieser Helix die Carbonylgruppen alternierend zum N- bzw. zum C-Terminus ausgerichtet sind und die Helix somit ein nur geringes Dipolmoment besitzt. Verbindung **5** sollte daher in polaren Lösungsmitteln eine deutlich andere Struktur einnehmen, was sich auch tatsächlich aus CD- und NMR-spektroskopischen Untersuchungen an **5** in DMSO als Lösungsmittel ergab.<sup>[138]</sup> Dabei legten die CD-Spektren eine zumindest teilweise helicale Konformation nahe. Die genaue Bestimmung der Struktur von **5** in DMSO erwies sich allerdings als unmöglich, da im NMR-Spektrum viele chemische Verschiebungen der Zucker- und Alkylreste überlagert waren.

Um abschließend endgültige Sicherheit über die Richtigkeit der ermittelten Struktur von **5** in CD<sub>3</sub>CN zu erlangen, wurde das NOESY-Spektrum gründlich nach Signalen untersucht, die eine Unterscheidung von anderen Helixtypen erlauben. In Frage kommen dabei die 10/12/10-, 10/10-, 12/12-, und 14/14(also 3<sub>1</sub>)-Helices, die sich in den Untersuchungen von HOFMANN *et al.* als vergleichsweise stabil erwiesen hatten. Die entsprechenden Helices wurden basierend auf idealisierten Rückgratwinkeln aufgebaut und aus den Strukturen charakteristische Abstände abgeleitet, die die Unterscheidung ermöglichen. Auf diese Art konnte die 10/12/10-Helix anhand von 11, die 10/10-Helix mittels 6, sowie die 12/12- und 14/14-Helices durch 4 solcher Abstände eindeutig ausgeschlossen werden (siehe Tabelle 8-3 im Anhang).



## 5 Das Parvulin *E. coli* Par10

In dieser Arbeit wurde die Struktur des Parvulins Par10 aufgeklärt, einer Peptidyl-Prolyl-*cis/trans*-Isomerase (PPIase) mit 92 Aminosäureresten, die im Cytoplasma von *E. coli* vorkommt. Als bislang kleinstes Mitglied der Parvulin-Familie der PPIasen ist Par10 vor allem deshalb wichtig, da anhand seiner die für die enzymatische Aktivität der Parvuline essentiellen Struktur motive eingegrenzt werden können. Für die Medizin ist die Familie der Parvuline von hohem Interesse, da das menschliche *hPin1* funktional an einer Reihe von Krankheitsbildern wie beispielsweise Krebs oder Alzheimer beteiligt ist.<sup>[153-159]</sup> Anhand der NMR-Spektren konnte die komplizierte, autokatalysierte Faltung des Proteins auf das Vorliegen einer *cis*-Gly-Pro-Peptidbindung zurückgeführt werden. Im Folgenden soll zunächst auf den biologischen Hintergrund von *E. coli* Par10 und anderer PPIasen eingegangen werden. Anschließend wird die zur Strukturbestimmung verwendete Methodik beschrieben sowie die Struktur diskutiert und mit derjenigen anderer PPIasen verglichen.

### 5.1 Biologischer Hintergrund

#### 5.1.1 Die Faltung von Proteinen

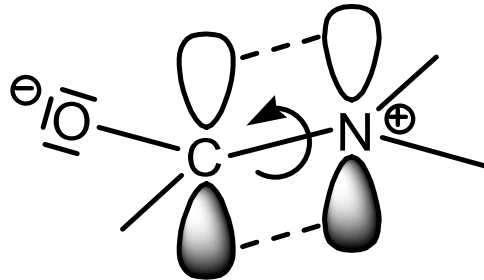
Bei der Synthese von Proteinen aus aktivierten Aminosäuren am Ribosom entstehen zunächst lineare Polypeptidstränge. Damit ein Protein in der Zelle die erforderliche Funktion erfüllen kann, muss es in einer charakteristischen Faltung (Sekundär- und Tertiärstruktur) vorliegen. Allerdings besitzen Proteine eine sehr hohe Anzahl von konformationellen Freiheitsgraden. Daher sollte zu erwarten sein, dass die spontane Faltung in den Energie-niedrigsten Zustand, ohne die Einwirkung äußerer Faktoren, nahezu unmöglich ist. Diese Annahme ist als LEVINTHAL-Paradoxon bekannt,<sup>[160]</sup> da sie im Gegensatz zu experimentellen Beobachtungen steht. In einer grundlegenden Arbeit konnten C. ANFINSEN *et al.* nämlich bereits im Jahre 1963 zeigen, dass der Faltungsprozess selbständig, also ohne Zufuhr von Energie oder die Einwirkung äußerer Faktoren stattfindet und zur korrekten Faltung führt.<sup>[161]</sup> Aus den Ergebnissen der *in vitro* an gereinigten Proteinen durchgeführten Arbeit, für die ANFINSEN 1972 den NOBEL-Preis für Chemie erhielt, wurde damals der nahe liegende Schluss gezogen, dass auch *in vivo* die spontane Faltung erfolgt.<sup>[162]</sup> Allerdings herrschen in der Zelle wesentlich ungünstigere Bedingungen zum spontanen Erreichen der korrekten Faltung. Beispielsweise liegen die am Ribosom erzeugten Proteine in recht hohen Konzentrationen

vor, wodurch die unerwünschte Aggregation in starke Konkurrenz mit der Faltung tritt (die vergleichsweise hohen Temperaturen in der Zelle begünstigen diesen Prozess zusätzlich). Mittlerweile weiß man jedoch, dass in der Zelle eine äußerst ausgereifte und komplexe Maschinerie die Faltung der Proteine unterstützt und überwacht. Insbesondere sind Faltungshelfer, so genannte molekulare *Chaperones*, in der Lage, die Aggregation zu verhindern. Dabei kommen in der Regel Mechanismen zum Zug, die das Abschirmen der hydrophoben Oberfläche des noch ungefalteten Proteins beinhalten, da im Allgemeinen die Aggregation dort stattfindet. Ein äußerst beeindruckendes Beispiel hierfür ist das *E. coli* GroE-System, das ungefaltete Proteine hydrophob bindet, diese in einer molekularen Zelle zur Faltung einschließt und nach einer Weile, zeitgesteuert durch die Hydrolyserate von gebundenem ATP, wieder freigibt. Selbst wenn ein Protein falsch gefaltet in einem lokalen Minimum der Energiehyperfläche gefangen ist, können molekulare *Chaperones* durch teilweise Entfaltung den Weg zur korrekten Konformation ebnen. In all diesen Fällen ist allerdings eine Energiequelle von Nöten (wie beispielsweise die Hydrolyse von ATP).

Eine weitere Art der Hilfestellung bei der Proteinfaltung erfolgt durch Enzyme, wobei zwei Enzymfamilien eine besondere Rolle spielen. Zum einen sind dies die von C. ANFINSEN entdeckten Protein-Disulfid-Isomerasen (PDI), welche die Bildung von Disulfidbrückenbindungen wesentlich beschleunigen können (interessanterweise hat sich dabei herausgestellt, dass die Disulfidbrücken dabei nicht willkürlich gebildet werden, sondern die Katalyse genau zu den im aktiven, korrekten Faltungszustand des Proteins erforderlichen Bindungen führt).<sup>[163]</sup>

Bereits seit längerer Zeit ist bekannt, dass in Proteinen die Peptidbindung zwischen einer Aminosäure und Prolin in der (energetisch prinzipiell ungünstigen) *cis*-Form vorliegen kann.<sup>[164-168]</sup> Auch wenn ältere Abschätzungen der Häufigkeit in letzter Zeit geringfügig von 6.5 % auf 5.2 % herunterkorrigiert wurden,<sup>[168]</sup> ist der Anteil von *cis*-XAA-Pro-Bindungen in Proteinen nicht zu unterschätzen. Am Ribosom werden Proteine zunächst in der all-*trans*-Form exprimiert, wobei prinzipiell die spontane Faltung in die korrekte Tertiärstruktur mit *cis*-Peptidbindungen erfolgen kann. Allerdings führen die für die Umwandlung zu überwindenden Aktivierungsbarrieren von ca. 70 bis 80 kJ mol<sup>-1</sup> bei beispielsweise 37 °C (menschlicher Körpertemperatur) immerhin zu mittleren Umwandlungszeiten im Minutenbereich. Aus den bereits erläuterten Gründen muss daher in der Zelle die Isomerisierung durch Enzyme katalysiert erfolgen. Diese zweite, bei der Proteinfaltung wichtige Enzymklasse wird als die Familie der Peptidyl-Prolyl-*cis/trans*-Isomerasen (PPIasen) bezeichnet.<sup>[169]</sup>

Damit die Isomerisierung von der *trans*- ( $\omega = 180^\circ$ ) zur *cis*-Konformation ( $\omega = 0^\circ$ ) stattfinden kann, ist es nötig, die  $\pi$ -Wechselwirkung zwischen dem p-Orbital des Carbonylkohlenstoffs C'(i-1) und dem freien Elektronenpaar des Amidstickstoffs N(i) zeitweilig aufzuheben ( $\omega = 90^\circ$ ) (Abbildung 5-1).



**Abbildung 5-1:** Zur Isomerisierung der Peptidbindung durch PPIasen muss der Doppelbindungscharakter, also die Überlappung der p-Orbitale, intermediär aufgehoben werden.

Die damit verbundene Aktivierungsbarriere wird durch die PPIasen herabgesetzt und somit die Isomerisierung beschleunigt. Für den entsprechenden Reaktionsmechanismus kommen verschiedene Varianten, sowohl ionischer als auch kovalenter Art, in Betracht. Beispielsweise könnte das Enzym als Nucleophil reagieren, durch Addition an den Carbonylkohlenstoff dessen Hybridisierung von  $sp^2$  nach  $sp^3$  ändern und somit die Peptid- $\pi$ -Bindung intermediär aufheben.<sup>[170, 171]</sup> Analog ist es denkbar, dass durch einen elektrophilen Angriff des Enzyms am Amidstickstoff dessen freies Elektronenpaar zur Bildung einer Enzym-Stickstoff-Bindung benutzt wird, wodurch der Doppelbindungscharakter ebenfalls zeitweilig aufgehoben wäre.<sup>[172]</sup> Auch ionische Mechanismen wären mit einem Angriff der Peptidbindung durch Übertragung einer (positiven oder negativen) Ladung und damit einer Schwächung der Doppelbindung verbunden.<sup>[172]</sup> Darüber hinaus wurde die Möglichkeit diskutiert, dass der Winkel  $\omega$  auch direkt durch das Enzym auf  $90^\circ$  verdreht werden könnte, wobei die Aktivierungsenergie durch den Transfer des Substrats vom wässrigen Medium in die hydrophobe Umgebung im Enzym bereitgestellt würde.<sup>[173-175]</sup> Es sind allerdings auch Kombinationen einzelner Mechanismen vorstellbar.

Für das Parvulin *hPin1* des Menschen wurde von RANGANATHAN *et al.* ein Mechanismus postuliert, bei dem eine nucleophile Addition an den Carbonylkohlenstoff stattfindet (siehe Kapitel 5.6).<sup>[176]</sup>

### 5.1.2 Die PPIase-Familien

Die PPIasen werden heutzutage in drei Familien eingeteilt, wobei man die Cyclophiline, die FK-bindenden Proteine (FKBP) und die Parvuline unterscheidet.

Die Familie der Cyclophiline umfasst eine Reihe von hochgradig konservierten Proteinen, die im gesamten phylogenetischen Baum, vom Menschen bis zurück zu Archaeobakterien vertreten sind. Sie sind in den verschiedensten Zellen eines jeweiligen Organismus allgegenwärtig und kommen oftmals in relativ großen Mengen vor. Wie der Name andeutet, binden Cyclophiline spezifisch an das cyclische, aus Pilzen stammende Undecapeptid Cyclosporin A (CsA), welches als Immunsuppressivum nach Organtransplantationen eingesetzt wird.<sup>[177, 178]</sup>

Die Zuordnung der Cyclophiline als Familie der PPIasen erfolgte durch eine zufällige Entdeckung. So wurde 1984 von G. FISCHER *et al.* erstmals ein Protein entdeckt, das in der Lage ist, die *cis-trans*-Isomerisierung der imidischen Peptidbindung zwischen einer natürlichen Aminosäure und Prolin zu katalysieren.<sup>[169]</sup> Bald zeigte sich, dass die gemäß dieser Eigenschaft als „Peptidyl-Prolyl-*cis/trans*-Isomerase“ (PPIase) bezeichnete Verbindung die Rückfaltung etlicher Proteine *in vitro* beschleunigt.<sup>[169, 179-181]</sup> Entsprechende PPIase-Aktivität wurde dabei in diversen Organismen wie beispielsweise Hefe, Insekten, *E. coli* und Säugetieren entdeckt. Interessanterweise stellte sich dann im Jahre 1989 heraus, dass die (aus Schweinenieren isolierte) PPIase und das von HANDSCHUMACHER *et al.* entdeckte Cyclophilin des Rinds<sup>[177]</sup> die gleiche Sequenz besitzen.<sup>[171, 182]</sup> Durch den Befund, dass Cyclosporin A ebenfalls die Aktivität der PPIase des Schweins hemmt, sowie durch eine Reihe weiterer Ergebnisse konnte gezeigt werden, dass die beiden Enzyme identisch sind.

Die zweite PPIase-Familie, die FK-bindenden Proteine (FKBP), wurden nach dem Makrolid<sup>1</sup> FK506 benannt, durch das sie inhibiert werden (und auch durch Rapamycin).<sup>[183]</sup> Die FKBP sind ebenfalls eine ubiquitäre Gruppe von Proteinen, die in allen Bereichen von Zellen vorkommen. Da auch FK506 als Immunsuppressivum eingesetzt wird,<sup>[184, 185]</sup> werden die Cyclophiline und die FKBP als Immunophiline bezeichnet. Neben ihrer Bedeutung als

---

<sup>(1)</sup> Unter Makroliden versteht man im Allgemeinen aus *Streptomyces*-Arten gewonnene, chemisch modifizierte Antibiotika, die jedoch auch rein synthetisch dargestellt sein können.

Zielrezeptoren für FK506 und andere Makrolide kommen den FKBP vielfältige Funktionen in der Zelle zu.<sup>[186, 187]</sup>

Das im Cytoplasma von *E. coli* vorkommende Par10, dessen Struktur in dieser Arbeit aufgeklärt wurde, gehört zur dritten Familie der PPIasen, den so genannten Parvulinen. Als erstes, 1994 durch G. FISCHER *et al.* aufgefundenes Parvulin<sup>[188, 189]</sup> ist Par10 mit einer molaren Masse von 10.1 kDa auch zugleich der bislang kleinste Vertreter der Familie. Mittlerweile haben DNA-Sequenzvergleiche zur Entdeckung etlicher weiterer Parvuline in diversen Organismen, sowohl in Pro- als auch in Eukaryoten, geführt. Von Par10 (SWISS-PROT: PPIC\_ECOLI, P39159, E.C. 5.2.1.8) wird vermutet, dass es an der zellulären oxidativen Stressantwort beteiligt ist.<sup>[190]</sup> Die PPIase-Aktivität der Parvuline wird durch die bekannten Immunsuppressiva der anderen beiden PPIase-Familien nicht beeinflusst.<sup>[189]</sup>

Die einzelnen Parvuline entfalten ihre katalytische Aktivität meist recht unspezifisch, können jedoch aber auch sehr gezielt wirken, wie beispielsweise das *hPin1*<sup>[191, 192]</sup> im Menschen. Dieses Enzym katalysiert ausschließlich die Isomerisierung von Peptidbindungen zwischen phosphoryliertem Serin bzw. Threonin und Prolin, wobei ein längerer Aminosäureeinschub für die Bindung an die phosphorylierte Aminosäure verantwortlich ist. Zusammen mit anderen derartigen Vertretern der Familie bildet das *hPin1* somit die Parvulin-Unterfamilie der *hPin1*-Homologen. Wie Abbildung 5-2 erkennen lässt, stellen die Parvuline eine hochkonservierte Familie von Proteinen dar, wobei auch der Einschub der *hPin1*-Homologen stark konserviert ist.

Parvulin box		motif I					
PPIC_ECOLI	1	--AKTAAALHILVKEEK-----LAL--DLLEQIKNGADFGKLAKKH-----					
SALMONELLA	1	--MAKMAAALHILVKEEK-----LAL--DLLEQIKNGGDFEKLAKKH-----					
CENARCHAEUM	1	--MADKIKCSHILVKQG-----EAL--AVQERIKAGEKFGKLAKELE-----					
ARABIDOPSIS	13	SASREILVQHLLVKNND-----VELFAELQKKFLDGEEMSDLAEEY-----					
PAR14	33	GGNAVKVRHILCEKHG-----KIM--EAMEKIKSGMRFNEVAAQY-----					
PIN1_HUMAN	50	QEPARVRCSHILVKHSQSRPSSWRQEKITRTKE--EALELLINGYIQIKKSGEED--FESLASQF					
DOD_DROMEIO	53	GAPDEVHCLHLLVKKHKSRRPSSWREANITRTKE--EAQLLEEVYRNKIVQQEAT--FDELARSY					
ESS1_YEAST	75	RHPVVRVRLHLLVKKHKSRRPASHRSENITISKQ--DATDELKTLITRLDDDSKTNSEALAKER					
<b>consensus</b>	<b>1</b>	<b>a lHiLvk ----- 1--dle ik g df kla -----</b>					

Parvulin box		motif II		motif III	
PPIC_ECOLI	38	SI-CPSGKRGDDLGEFRQGMVPAFDKVVFSQPVLEPTGPLHTQ-----FGYHIKVLRYRN			
SALMONELLA	39	SI-CPSGKKGGLGEFRQGMVPAFDKVVFSQPVLEPTGPLHTQ-----FGYHIKVLRYRK			
CENARCHAEUM	39	SIDGGSAKRDGSLGYFGRGKMVKPFEDAAFRLQVGEVSEPVKSE-----FGYHVIKRLG--			
ARABIDOPSIS	54	SI-CPSKKGGLGWLWVKLGOMVPEFEAAAFKAE-LNQVVRCTQ-----FGLHLLQVLSER			
PAR14	72	S--EDKARQGGDLGWMTRGSMVGPFOEAAAFALPVSMDKPVFTDPPVKTKFGYHIIMVEGRK			
PIN1_HUMAN	111	SD-CSSAKARGDLGAFSRGOMQKPFEDASALRTGEMSGPVFTD-----SGIHIILRTE--			
DOD_DROMEIO	114	SD-CSSAKRGDDLKGFGRGOMQAAFEDAAFKLNVNQLSGIVDS-----SGLHIILRKA--			
ESS1_YEAST	138	SD-CSSYKRGDDLGEFRGEMQPSFEDAAFQLKVGVEVSDIVESG-----SGVHVIKRVG--			
<b>consensus</b>		<b>Si-c sgkrgGdLG f GqMvpaFd Fs v e tgplhtq-----fGyHiikvl</b>			

**Abbildung 5-2:** Sequenzvergleich einiger Homologe der Parvulin-Familie. Grau unterlegte Felder kennzeichnen ähnliche, Schwarz unterlegte Felder identische Aminosäuren. Der Konsensus gibt an, welche Aminosäuren übereinstimmend in den Sequenzen vorkommen. Als Referenz für den Konsensus diente *E. coli* Par10 (PPIC\_ECOLI). Im hPin1 des Menschen, Dodo aus *Drosophila melanogaster* und Ess1 aus Hefe ist der Einschub zu erkennen, der die Bindung an phosphorylierte Aminosäuren ermöglicht. Der Sequenzvergleich wurde mit dem Programm FASTA<sup>[193, 194]</sup> erstellt.

Die Funktion des Parvulins hPar14 des Menschen<sup>[195]</sup> ist bisher noch nicht bekannt. Dahingegen hat sich mittlerweile herausgestellt, dass das ebenfalls in menschlichen Zellen vorkommende hPin1<sup>[191, 192]</sup> eine Reihe von Funktionen übernimmt. Beispielsweise wirkt hPin1 regulierend im G2/M-Übergang der Mitose. Fehlt hPin1, so kommt es zum Abbruch der Mitose und auch zur Apoptose der Zelle. Ebenso scheint hPin1 in neurodegenerativen Prozessen eine Rolle zu spielen, die letztendlich zum Krankheitsbild des ALZHEIMER-Syndroms führen, indem es die Funktion von phosphoryliertem Tau-Protein wiederherstellt.<sup>[153]</sup> Bei Brustkrebs kommt es zu einer Überexprimierung von hPin1 und damit zur Förderung der Onkogenese.<sup>[154-156]</sup> Weiterhin wird vermutet, dass Pin1 eine Antitumorwirkung besitzen könnte, indem es durch seine Interaktion mit dem Tumor-suppressorprotein p53 die Apoptose von Zellen nach genotoxischen Vorfällen einleitet.<sup>[157-159]</sup>

Interessanterweise wurde mittlerweile eine Unterfamilie von Parvulinen in Eukaryoten gefunden, die dem Par10 von *E. coli* durch Abwesenheit der sonst für eukaryotische Parvuline

typischen Wechselwirkungsdomäne ähneln.<sup>[196]</sup> Die in Zellen des Menschen aufgefundene Variante *hEPVH* dieser Unterfamilie der Parvuline (EPVH, engl. Eukaryotic Parvulin Homologues) findet sich vorwiegend in der mitochondrialen Matrix. Sie besitzt zwar eine messbare PPIase-Aktivität, ist jedoch merklich inaktiver und weniger selektiv als *hPin1*.

Durch das in Walnüssen vorkommende Juglon lassen sich Par10 und andere Parvuline selektiv bezüglich anderer PPIasen irreversibel inhibieren, da es an Cysteine des Enzyms bindet und somit vermutlich zu einer teilweisen Entfaltung des aktiven Zentrums führt.<sup>[197]</sup> Allerdings ist bisher noch kein Inhibitor gefunden worden, mit dem die selektive Inhibierung eines einzelnen Parvulins zu erreichen wäre.

## 5.2 NMR-spektroskopische Untersuchungen an Par10

Die NMR-spektroskopischen Messungen, Zuordnung der Resonanzen und die Bestimmung der Sequenz wurden in unserem Arbeitskreis durch A. KÜHLEWEIN und G. GEMMECKER vorgenommen. In Zusammenarbeit mit letzteren beiden habe ich die Auswertung der dreidimensionalen NOESY-Spektren durchgeführt.

Vermessen wurden Proben von *E. coli* Par10, das an der Max-Planck-Forschungsstelle für Enzymologie und Proteinfaltung in der Arbeitsgruppe von G. FISCHER exprimiert und gereinigt wurde.<sup>[188, 189]</sup> Zur Verfügung standen [U-<sup>15</sup>N]- und [U-<sup>13</sup>C/<sup>15</sup>N]-markierte Proben. Dabei kamen Spektrometer des Typs DMX600 und DMX750 der Firma BRUKER zum Einsatz.

Die sequentielle Zuordnung der <sup>1</sup>H-, <sup>13</sup>C- und <sup>15</sup>N-chemischen Verschiebungen wurde durch Zuhilfenahme des automatischen Zuordnungsprogramms PASTA durchgeführt.<sup>[198]</sup> Hierfür wurden die Daten aus <sup>15</sup>N-HSQC-, HNCACB-, CBCA(CO)NH-, HA(CACO)NH- und HNHA-Spektren verwendet. Die Zuordnung der Peptidrückgrat-Resonanzen gelang vollständig. Über 90 % aller Seitenkettensignale konnten durch die Experimente 15N-TOCSY-HSQC, HCCH-TOCSY, HCCH-COSY und (H)CCCONH zugeordnet werden. Die Zuordnung der H<sup>γ</sup>- und H<sup>δ</sup>-Protonen der Lysin-, Leucin- und Isoleucin-Seitenketten war wegen Signalüberlagerungen nur teilweise möglich. Mittels eines <sup>13</sup>C-NOESY-HSQC-Spektrums konnten manche überlagerte Signale allerdings aufgelöst werden. Die Zuordnung von aromatischen H<sup>δ</sup>- und H<sup>ε</sup>-Protonen erfolgte anhand von C<sup>β</sup>-H<sup>δ/ε</sup>-Korrelationsexperimenten<sup>[199]</sup>.

Die Bestimmung der Sekundärstruktur wurde durch die kombinierte Auswertung von vier voneinander unabhängigen Parametern durchgeführt. Betrachtet wurden die Abweichung der

chemischen Verschiebungen von denjenigen in ungeordneten Peptiden (engl. *chemical shift index*, CSI), die  $3J(\text{H}^{\text{N}}\text{H}^{\alpha})$ -Kopplungskonstanten (gewonnen aus einem 3D-HNHA-Spektrum), der Amidprotonen-Austausch mit dem Solvens, sowie für die Sekundärstruktur charakteristische, aus den NOESY-Spektren bestimmte Abstände. Für die Bestimmung des Amidprotonen-Austauschs mit dem Lösungsmittel wurden zwei MEXICO-Spektren<sup>[45]</sup> bei Mischzeiten von 50 und 200 ms gemessen.

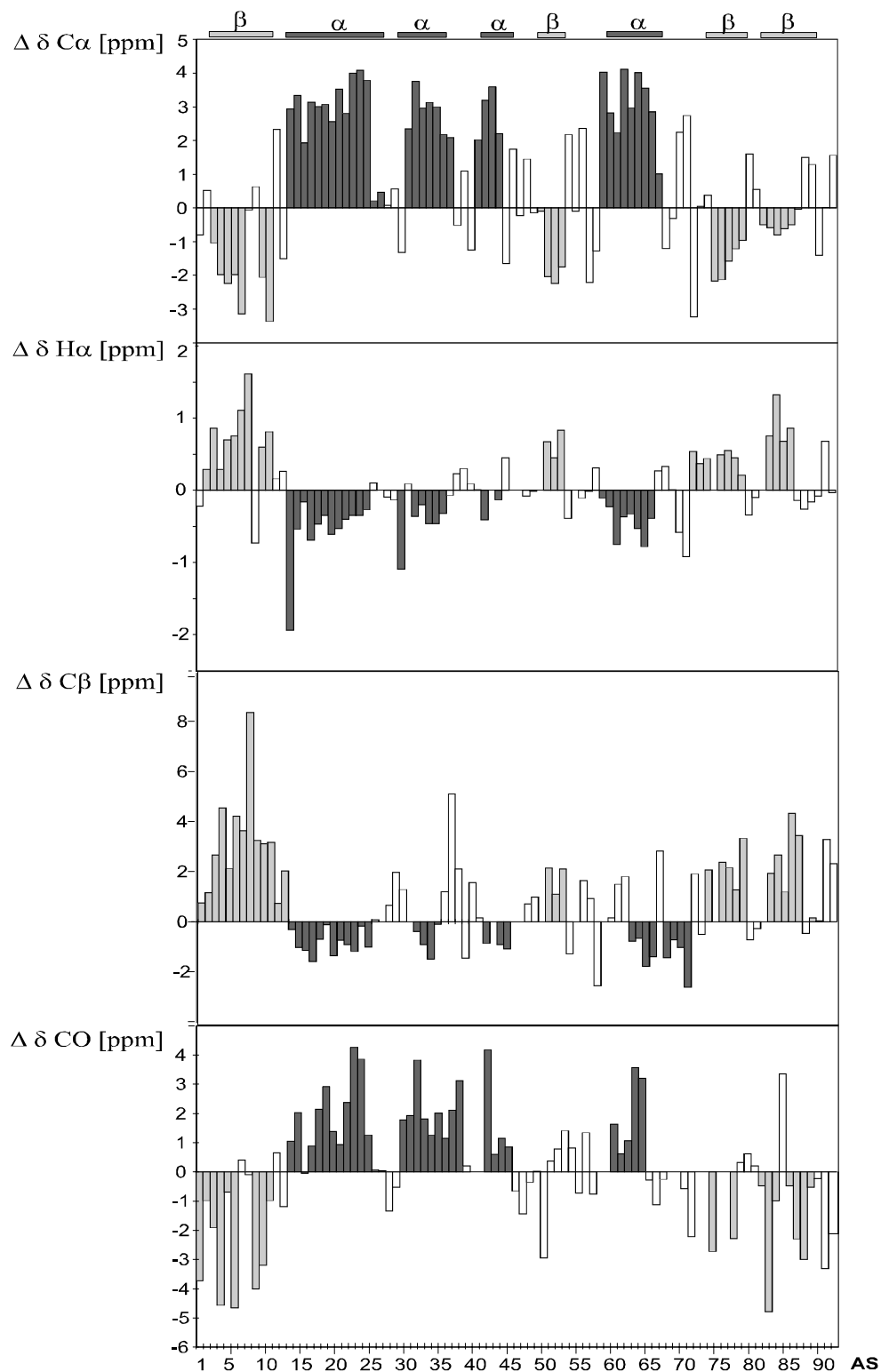
Zur Gewinnung von Abstandsdaten wurden bei einer Mischzeit von 80 ms jeweils ein 3D  $^{15}\text{N}$ -NOESY-HSQC-,  $^{13}\text{C}$ -NOESY-HSQC-, CNH-NOESY- sowie CCH-NOESY-Spektrum gemessen. Die mit dem Programm SYBYL gewonnenen Integrale der Kreuzsignale wurden mit Hilfe der automatischen Kalibrierungsmethode und wie oben beschrieben in Abstandsgrenzen für die Strukturrechnungen überführt.

### 5.3 Automatische NOE-Zuordnung mit ARIA

Die Erzeugung der ersten Strukturen von Par10 erfolgte mit Hilfe des automatischen Zuordnungsprogramms ARIA,<sup>[33, 74-80]</sup> wobei die Versionen 0.53 bzw. 0.9 verwendet wurden. Wie bereits in Kapitel 2.2.6 näher beschrieben, ist die Bestimmung der Struktur eines Proteins anhand von NMR-Daten stets von der Gefahr begleitet, von Anfang an eine falsche Konformation aufzufinden. ARIA benötigt daher einen Satz an strukturellen Informationen, die zwar mehrdeutig sein dürfen, aber fehlerfrei sein müssen. Daher wurden zunächst die Informationen, die sich aus dem CSI (*chemical shift index*) ergaben, sorgfältig mit denjenigen aus den Kopplungskonstanten, den Sekundärstruktur-relevanten NOEs und dem MEXICO-Experiment verglichen.

Mit dem CSI lassen sich erste Aussagen über die Sekundärstruktur eines Proteins treffen, da die Verschiebungen der Atome  $\text{C}^{\alpha}$ ,  $\text{C}^{\beta}$ ,  $\text{C}'$  und  $\text{H}^{\alpha}$  von der lokalen Sekundärstruktur abhängig sind. Durch den Vergleich der Verschiebungen dieser Atome in dem zu bestimmenden Protein mit denjenigen in ungeordneten Proteinen ist es möglich  $\alpha$ -helicale,  $\beta$ -faltblattartige und ungeordnete Bereiche zuzuordnen.<sup>[43, 44]</sup> In  $\beta$ -Faltblättern weichen die Verschiebungen von  $\text{C}^{\alpha}$ - und  $\text{C}'$ -Atomen gegenüber den *random-coil*-Werten positiv ab, wohingegen die Verschiebungen der Atome  $\text{H}^{\alpha}$  und  $\text{C}^{\beta}$  negativ abweichen. Für  $\alpha$ -Helices ist dies genau gegensätzlich. Abbildung 5-3 zeigt die für Par10 bestimmten Abweichungen der Verschiebungen von den *random-coil*-Werten und die sich daraus ergebenden Sekundärstruktur motive.





**Abbildung 5-3:** Graphische Darstellung der Abweichungen der chemischen Verschiebungen von Par10 von denjenigen einer ungeordneten Sequenz (CSI). Die  $\alpha$ -helicalen Bereiche sind dunkelgrau gefüllt, die  $\beta$ -Faltblattbereiche sind hellgrau gefüllt. Oben in der Grafik sind die tatsächlich eingenommenen Sekundärstrukturelemente in der verfeinerten Struktur von Par10 dargestellt (s. u.).

Gemäß dem CSI beginnt die Struktur von Par10 N-terminal mit einem Faltblattstrang, gefolgt von drei helicalen Bereichen. Auf einen weiteren, kurzen Faltblattstrang schließt sich dann ein helicaler Bereich an, wonach zwei C-terminale Faltblattstränge folgen. Allerdings ist die Bestimmung der Sekundärstrukturelemente mittels des CSI meist nicht genau genug, um ihre exakten Positionen ermitteln zu können. Ein genaueres Bild vermitteln NOE-Muster, die für die einzelnen Strukturelemente charakteristisch sind. In helicalen Bereichen sind insbesondere sehr starke  $H^N_i-H^N_{i+1}$ -Signale typisch. Etwas schwächer werden dann oft auch die Kontakte  $H^N_i-H^N_{i+3}$  und  $H^N_i-H^N_{i+4}$  beobachtet. Während diese NOEs in Faltblattsträngen wesentlich schwächer bis nicht vorhanden sind, treten dort starke  $H^N_i-H^\alpha_{i-1}$ -Signale auf. Vor den ersten Strukturberechnungen mit ARIA wurden diese Signale sorgfältig manuell zugeordnet. Dazu standen anfänglich zwei 3D NOESY-Spektren zur Verfügung ( $^{15}\text{N}$ -NOESY-HSQC und  $^{13}\text{C}$ -NOESY-HSQC), die dann um ein CNH-NOESY und ein CCH-NOESY-Spektrum ergänzt wurden.

Zur Unterstützung können weiterhin die skalaren  $^3\text{J-H}^N\text{H}^\alpha$ -Kopplungskonstanten herangezogen werden, wobei große Kopplungskonstanten auf Faltblattstränge und kleine Kopplungskonstanten auf helicale Bereiche hinweisen. Gemittelte Kopplungskonstanten zwischen 4 und 8 Hz deuten dahingegen entweder flexible oder auch starre, ungeordnete Bereiche an. In Abbildung 5-4 sind die Informationen dargestellt, die sich aus den NOE-Mustern und Kopplungskonstanten ergeben.



**Abbildung 5-4:** Graphische Darstellung der charakteristischen Abstände und der Kopplungskonstanten in Par10 und die sich daraus ergebenden Sekundärstrukturelemente (oberste Zeile). Die Höhe der einzelnen Kästchen repräsentiert die Größe des jeweiligen Werts. In der untersten Zeile ist die Sekundärstruktur angegeben, die sich aus dem Konsensus mit dem CSI ergibt.

Während eines ARIA-Laufs wurden jeweils 8 Iterationen durchgeführt (also das Maximum in den verwendeten Versionen), wobei in jeder Iteration 20 Strukturen erzeugt wurden. Die Berechnung der Strukturen wurde mit den Standardprotokollen von ARIA bzw. CNS<sup>[32]</sup> durchgeführt. Als Eingabe wurde jeweils eine Liste aller zugeordneten chemischen Protonenverschiebungen, der manuell zugeordneten Abstände und der  $^3\text{J-H}^{\text{N}}\text{H}^{\alpha}$ -Kopplungskonstanten eingesetzt. Ebenso wurde ARIA eine Liste von Kreuzsignalen und deren Integralen zur Verfügung gestellt, die zunächst mit Hilfe der *peak-picking*-Routine des SYBYL-Programmpakets automatisch erstellt wurde, dann manuell auf Korrektheit überprüft und nötigenfalls überarbeitet wurde. Da zwischen SYBYL und ARIA keine Konvertierungsroutine zur Verfügung stand, wurde diese Liste aus den Kreuzsignallisten von SYBYL mittels eines selbst geschriebenen JAVA-Programms konvertiert und gleichzeitig mit den für die einzelnen Protonen erforderlichen Frequenzfenstern (Toleranzen) versehen. Als Ausmaße der Frequenzfenster wurden, abhängig von der Auflösung des jeweiligen Spektrums, die in Tabelle 5-1 angegebenen Werte verwendet:

**Tabelle 5-1:** Ausmaße der Frequenzfenster für die automatische NOE-Zuordnung in den mit ARIA durchgeführten Strukturberechnungen von E. coli Par10 (alle Werte in ppm).

Dimension	<sup>13</sup> C-NOESY- HSQC	<sup>15</sup> N-NOESY- HSQC	CCH-NOESY	CNH-NOESY
<sup>1</sup> H (direkt)	± 0.04	± 0.03	± 0.03	± 0.03
<sup>1</sup> H (indirekt)	± 0.07	± 0.07	--	--
<sup>13</sup> C	± 0.70	--	± 0.80 / ± 1.20	± 0.70
<sup>15</sup> N	--	± 0.40	--	± 0.40

Eine Liste mit Peptidrückgrat-Diederwinkeln  $\phi$  und  $\psi$  diente dazu, den Sekundärstrukturelementen verstärkt Rechnung zu tragen, die sich aus den kombinierten Informationen aus Kopplungskonstanten, charakteristischen Abständen und CSI ergaben. Für Helices wurden die Winkel  $\phi = -65^\circ$  und  $\psi = -40^\circ$  sanft bei einer Kraftkonstante von  $1 \text{ kcal mol}^{-1} \text{ rad}^{-2}$  einem quadratischen *flat-bottom*-Potential mit der relativ weiten Bodenbreite von  $40^\circ$  unterworfen, um auch die Ausbildung von  $3_{10}$ -Helices zuzulassen. In den ersten Rechnungen wurde analog für  $\beta$ -Stränge vorgegangen, dann jedoch bald davon

abgesehen, um die Ausbildung einer (bei anderen Parvulinen beobachteten) Krümmung des  $\beta$ -Faltblatts nicht zu verhindern.

In der ersten Iteration eines ARIA-Laufs wurde für die 20 Startstrukturen jeweils eine gestreckte Konformation verwendet. Bei Beginn jeder neuen Iteration wurden jeweils die 10 energetisch günstigsten Strukturen, ausgewählt nach dem ARIA-Standardprotokoll, behalten und als Startstrukturen für die nächste Iteration verwendet. Zusätzlich wurden mit dem neu iterierten Satz von Abständen zehn weitere Strukturen berechnet, wobei wiederum die gestreckte Konformation als Startstruktur diente. Die Strukturen der letzten Iteration wurden anschließend einer kurzen Verfeinerung in Wasser als explizitem Solvens unterworfen.<sup>[33]</sup> Auf diese Weise wurden im Rahmen der initialen Strukturberechnungen etliche ARIA-Läufe durchgeführt, nach jedem Lauf die Ergebnisse vorsichtig anhand der NMR-Spektren überprüft und eindeutige Abstände in die ursprüngliche Liste sicher zugewiesener Abstände übernommen. Bereits nach wenigen ARIA-Läufen begannen die erzeugten Ensembles von Strukturen zu einer Faltung zu konvergieren, die deutliche Ähnlichkeit mit dem homologen Bereich von *hPin1* besaß (siehe folgende Kapitel).

Um auch den Informationsgehalt des MEXICO-Experiments zu berücksichtigen, wurde in späteren ARIA-Läufen zusätzlich eine Liste mit Wasserstoffbrücken-*restraints* zwischen den  $\beta$ -Strängen eingesetzt, sobald die entsprechenden Donoren und Akzeptoren eindeutig zugeordnet werden konnten. Als *restraints* wurden dazu die Abstände O'-H<sup>N</sup> (1.8 – 2.3 Å) und O'-N (2.8 – 3.3 Å) verwendet.

Nach jedem ARIA-Lauf, sowie nach jedem Rechnungslauf in der folgenden Strukturverfeinerung wurde die Qualität der errechneten Strukturen mit dem Programm PROCHECK<sup>[200]</sup>, Version 3.4, beurteilt.

## 5.4 Strukturverfeinerung

Wie bereits erwähnt konvergierten die ARIA-Rechnungen recht bald zu Strukturensamples, die der Faltung des homologen Bereichs von *hPin1* sehr ähnlich waren. Allerdings blieben die berechneten Strukturen grundsätzlich recht unkompakt. Wie in Kapitel 5.5 beschrieben wird, liegen drei der vier  $\alpha$ -Helices in Par10 über der konvexen Seite eines gewölbten, viersträngig antiparallelen  $\beta$ -Faltblatts und die vierte Helix auf dessen Innenseite. Insbesondere die Helices auf der konvexen Seite blieben bei allen ARIA-Rechnungen in einem recht

ungewöhnlich großen Abstand zum Faltblatt.<sup>1</sup> Um daher mehr Kontrolle über den Verlauf der Strukturberechnung zu bekommen, wurden die weiteren Rechnungen mit dem XPLOR-Programmpaket<sup>[31]</sup> in der Version 3.851 vorgenommen.

Als Eingabe für die XPLOR-Rechnungen dienten Tabellen mit den in den ARIA-Rechnungen eindeutig zugeordneten NOE-Abständen,  $^3\text{J-H}^{\text{N}}\text{H}^{\alpha}$ -Kopplungskonstanten und den bereits beschriebenen Diederwinkel-*restraints* für die  $\alpha$ -helicalen Segmente. Weiterhin wurde wiederum eine Tabelle mit Wasserstoffbrücken-*restraints* eingesetzt.

Mit diesen Daten wurde unter Verwendung eines in unserer Arbeitsgruppe von M. COLES modifizierten *simulated-annealing*-Standardprotokolls<sup>[62, 64]</sup>, das auch die Verwendung von Wasserstoffbrücken-*restraints* erlaubt, zunächst ein Ensemble von 10 Strukturen berechnet, wobei eine gestreckte Konformation als Startstruktur diente. Allerdings werden in dem verwendeten SA-Protokoll nichtbindende Wechselwirkungen nur durch einen einfachen, abstoßenden Term (*repel function*) repräsentiert. Da der Einsatz eines elektrostatischen Terms bei Simulationen im Vakuum im Allgemeinen zu unrealistisch starken Wechselwirkungen zwischen stark geladenen Gruppen führt und die Verwendung eines VAN-DER-WAALSSchen Terms für Proteine sehr rechenaufwendig ist, wurde von ihrem Einsatz in der weiteren Verfeinerung abgesehen. Anstatt dessen wurde jede der Strukturen in einem Potential weiter optimiert, das die Peptidrückgratwinkel  $\phi$  und  $\psi$  in energetisch günstige Bereiche der RAMACHANDRAN-Energiefläche drängt.<sup>[201, 202]</sup> Der Einsatz eines solchen Potentials bietet zudem den Vorteil, durch erhöhte Verletzungen auf problematische NOEs hinzuweisen. Die Strukturen wurden somit erneut einem SA-Protokoll unterworfen, wobei zunächst die Hochtemperaturphase bei 1000 K durchgeführt wurde. Die Kraftkonstante für den *improper*-Winkel, der Peptidbindungen planar hält, wurde mit  $500 \text{ kcal rad}^{-2}$  absichtlich hoch gewählt, um eine *cis/trans*-Isomerisierung zu vermeiden. Anschließend wurde bei 400 K eine weitere SA-Verfeinerung bei einer Kraftkonstante des *improper*-Winkels von  $50 \text{ kcal rad}^{-2}$  durchgeführt, wodurch die unrealistisch starke Planarität der Peptidbindungen aufgehoben wird und eine insgesamt bessere Faltung erreicht wird. Danach wurde die energetisch günstigste Struktur des Ensembles ausgewählt und durch Anwendung eines Distanzfilters von

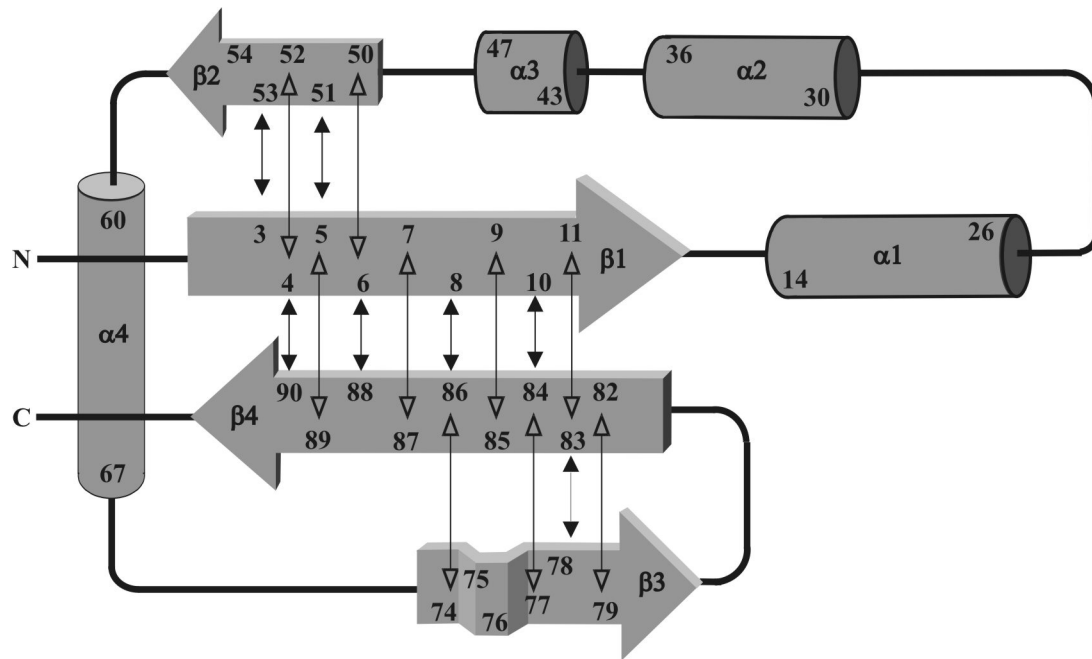
---

<sup>(1)</sup> Nicht zuletzt führte ein weiterer, wichtiger Grund zu dieser Entscheidung. Die ARIA-Rechnungen wurden jeweils auf Rechnern des Typs Origin 200 der Firma Silicon Graphics durchgeführt. Von diesen Rechnern mit je vier Prozessoren standen in unserem Arbeitskreis zwei Exemplare zur Verfügung, die sich (abgesehen von ihrem Arbeitsspeicher) nicht in ihrer *hard*- bzw. *software*-Ausstattung unterschieden. Dennoch lieferten die Rechnungen auf den beiden Computern insbesondere bezüglich der Qualität der Strukturen deutliche Unterschiede. Auch umfangreiche Tests konnten dies nicht erklärbar machen. Dazu kam erschwerend der Umstand, dass anfangs Informationen über die Funktionsweise von ARIA nur schwer zu erlangen waren. Während sich in letzter Zeit schnell ein lebendiges, rege kommunizierendes Forum für ARIA im Internet gebildet hat, waren zum Zeitpunkt der Arbeiten an Par10 nur äußerst rudimentär dokumentierte Anleitungen verfügbar.

10 Å, soweit möglich, vormals mehrdeutige NOEs eindeutig zugeordnet. Das Verfahren wurde dann solange iterativ wiederholt, bis sich keine weiteren NOEs mehr zuordnen ließen. Anschließend wurde nach selbigem Verfahren ein Ensemble von 50 Strukturen berechnet, wobei 1067 NOE-basierte Abstände, 30 Wasserstoffbrücken-*restraints*, 68 Diederwinkel-*restraints* und 42  $^3\text{J-H}^{\text{N}}\text{H}^{\alpha}$ -Kopplungskonstanten eingesetzt wurden. Von den 50 Strukturen wurden 17 (also ca. 1/3) aufgrund hoher Energie bezüglich des RAMACHANDRAN-Potentialterms verworfen. Von den verbleibenden Strukturen wurden diejenigen ausgewählt, für die keine NOE-Verletzung von mehr als 0.17 Å und keine Diederwinkel-*restraint*-Verletzung von mehr als 3° vorlag. Dabei wurden NOEs, die sich in flexiblen Bereichen befanden (Aminosäurereste 48-50 und 54-59 sowie flexible Seitenketten), nicht mit berücksichtigt. Um eine repräsentative Struktur aus dem verbleibenden Satz von 18 Strukturen zu gewinnen, wurden diese zunächst gemittelt. Anschließend wurde die gemittelte Struktur unter Einwirkung der experimentellen *restraints* optimiert, um durch die Mittelung entstandene Verzerrungen zu beseitigen. Dabei kam ein von M. COLES leicht modifiziertes XPLOR-Standardskript zum Einsatz, welches sonst zur Erzeugung von Startstrukturen idealer Geometrie dient (`generate_template.inp`; Änderungen: Einsatz experimenteller *restraints* wie im XPLOR SA-Standardprotokoll, Schleife über alle *input*-Strukturen). Als Bereich für die Superpositionierung der Ensemblestrukturen wurde die gesamte, wohl definierte Sekundärstruktur mit Ausnahme der dritten Helix (Aminosäurereste 43-47) gewählt.

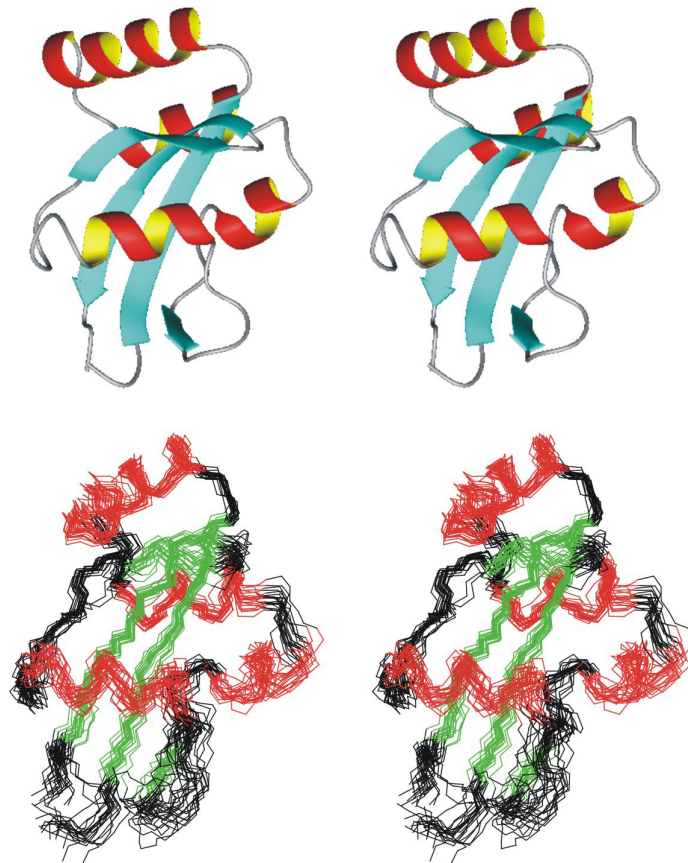
## 5.5 Strukturdiskussion

Die für Par10 ermittelte Struktur besteht aus vier  $\beta$ -Strängen sowie vier  $\alpha$ -Helices und zeigt ein Faltungsmuster, das demjenigen der homologen Bereiche von *hPin1* und *hPar14* sehr ähnlich ist. Sie ist charakterisiert durch ein stark gekrümmtes, antiparalleles  $\beta$ -Faltblatt der Strangabfolge 2143, das mit seiner Innenseite die Helix 4 leicht umschließt. Auf der konvexen Seite des Faltblatts befinden sich die Helices 1 bis 3. Abbildung 5-5 zeigt schematisch die Topologie von Par10.

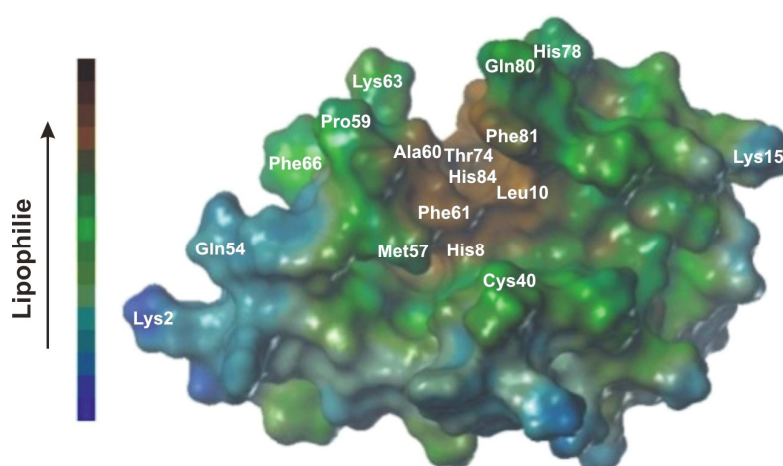


**Abbildung 5-5:** Schematische Darstellung der Topologie von E. coli Par10. Schwarz ausgefüllte Pfeile repräsentieren  $H^N-H^N$ -, nicht ausgefüllte Pfeile  $H^\alpha H^\alpha$ -NOE-Signale. Die konvexe Seite des gewölbten  $\beta$ -Faltblatts zeigt nach oben zum Betrachter. Im  $\beta$ -Strang 3 ist symbolisch ein Knick angedeutet, der durch das cis-konfigurierte Pro76 verursacht wird (siehe folgenden Text).

Die Krümmung des Falblatts ist deutlich in der Stereodarstellung der Struktur von Par10 zu erkennen. Das postulierte aktive Zentrum des Enzyms liegt zwischen der konkaven Seite des Falblatts und Helix 4. Die flexiblen Ankerregionen um Helix 4 sind mit 6 bzw. 7 Aminosäureresten recht lang. Um einem Substrat den Zugang zur lipophilen Bindungstasche zu gewähren, ist es denkbar, dass durch Konformationsänderungen in diesen Bereichen Helix 4 sich etwas vom Falblatt wegbewegen muss. Helix 3 liegt in etwa auf der Höhe der „Ebene“ des Falblatts (Abbildung 5-6). In Abbildung 5-7 ist deutlich eine Spalte zu sehen, in deren Inneren sich das aktive Zentrum des Enzyms mit vielen hochkonservierten Aminosäureresten befindet.



**Abbildung 5-6:** Stereodarstellung der gemittelten und optimierten Struktur sowie des auf das Peptidrückgrat superpositionierten Ensembles von Par10, erstellt mit dem Programm MOLMOL<sup>[203, 204]</sup>. Die konkave Seite des  $\beta$ -Faltblatts zeigt zum Betrachter.

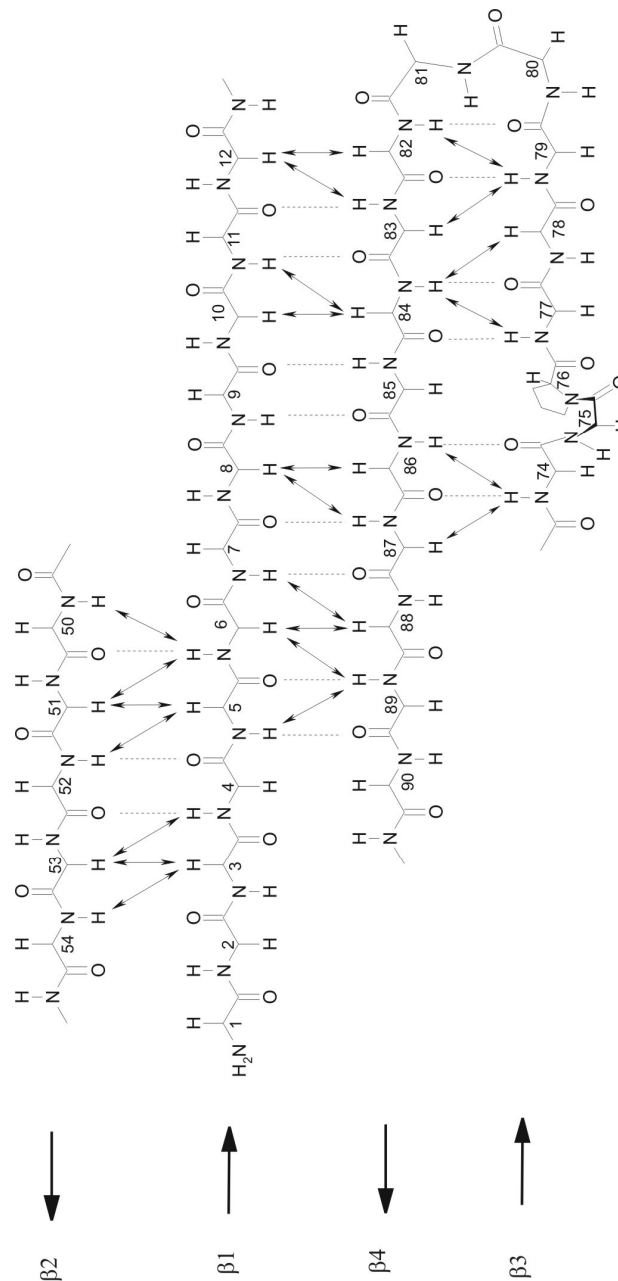


**Abbildung 5-7:** Lipophilie der Oberfläche von Par10. Die hydrophobe Bindungstasche oben in der Mitte des Bildes wird links von Helix 4 begrenzt. Die Reste H8, K49, F61, F81 und H84 sind in homologen Parvulinen hochkonserviert.



Wie später noch ausführlicher diskutiert wird, besitzen interessanterweise Par10 und andere Homologe eine charakteristische Knickstelle im Faltblattstrang 3, die in Par10 bei G75-P76 liegt. Für diese Sequenz ist wegen des Fehlens einer Seitenkette in Glycin die Tendenz zur Ausbildung einer *cis*-konfigurierten Amidbindung merklich höher als für andere XAA-Pro-Sequenzen. Da auch *hPin1* und Par14 an dieser Stelle die Sequenz G-P bzw. D-P-P besitzen, lag der Verdacht nahe, dass der Knick durch eine solche *cis*-Peptidbindung verursacht wird. In der Tat ist die chemische Verschiebung von P76 C<sup>β</sup> ( $\delta = 34.47$  ppm) typisch für ein *cis*-Prolin, diejenige von P76 C<sup>γ</sup> jedoch eher für ein *trans*-Prolin ( $\delta = 26.31$  ppm; wobei anzumerken ist, dass im Allgemeinen die Verschiebung von C<sup>β</sup> mehr aussagekräftig ist).<sup>[205]</sup> Eindeutigen Aufschluss gibt allerdings das 3D CCH-NOESY-Spektrum. So finden sich dort Signale zwischen G75 H<sup>α</sup> und P76 C<sup>α</sup>, nicht jedoch solche zwischen G75 H<sup>α</sup> und P76 C<sup>δ</sup>, welche für eine *trans*-konfigurierte Amidbindung zu erwarten wären. Ebenso fehlen Signale zwischen P76 H<sup>δ1</sup> bzw. H<sup>δ2</sup> und G75 C<sup>α</sup>, während ein Signal zwischen P76 H<sup>α</sup> und G75 C<sup>α</sup> die *cis*-Bindung bestätigt. Das Vorhandensein einer *cis*-konfigurierten Peptidbindung in Par10 liefert eine gute Erklärung für die Beobachtung, dass die Rückfaltung von mit Harnstoff denaturiertem Par10 mit einer sehr komplexen Kinetik bei zunächst sehr langsamen Raten autokatalysiert verläuft.<sup>[206]</sup>

Abbildung 5-8 zeigt eine schematische Darstellung des  $\beta$ -Faltblatts mit den darin ausgebildeten Wasserstoffbrücken und den in den Spektren beobachteten NOEs:



**Abbildung 5-8:** Schematische Darstellung des Peptidrückgrats des viersträngigen, antiparallelen  $\beta$ -Faltblatts in E. coli Par10. Gestrichelte Linien symbolisieren Wasserstoffbrückenbindungen, Pfeile deuten in den Spektren beobachtete NOEs an. Der Knick um das cis-konfigurierte Pro76 wird im folgenden Text näher beschrieben.

Als Maß für die Konvergenz im berechneten Strukturensamble ist in Tabelle 5-2 der RMSD-Wert der Schweratome bei Superpositionierung der Einzelstrukturen auf die gemittelte sowie auf die gemittelte und optimierte Struktur angegeben. Wie die widersprüchlichen NOE-Signale und die Daten des MEXICO-Experiments (s. u.) zeigen, ist die Struktur von Par10

insbesondere in den Ankerregionen um Helix 4, in einer Schleifenregion zwischen den Helices 1 und 2 sowie im Bereich von Helix 3 recht flexibel. Daher wurde der RMSD-Wert unter Ausschluss dieser Strukturelemente berechnet (Tabelle 5-2). Tabelle 5-3 vermittelt einen weiteren Eindruck über die Qualität des berechneten Ensembles.

**Tabelle 5-2:** RMSD-Werte für die Schweratome, bei Superpositionierung der 18 Einzelstrukturen SA des berechneten Strukturensambles von E. coli Par10 auf die gemittelte ( $\langle SA \rangle$ ) bzw. die gemittelte und optimierte Struktur  $\langle SA \rangle_r$ .

	SA gegen $\langle SA \rangle$		SA gegen $\langle SA \rangle_r$	
	Rückgrat	Alle	Rückgrat	Alle
Alle Reste	0.85 ± 0.14	1.39 ± 0.16	1.17 ± 0.20	1.87 ± 0.25
Sekundärstruktur	0.58 ± 0.11	1.15 ± 0.14	0.87 ± 0.09	1.64 ± 0.16
$\beta$ -Faltblatt, $\alpha 1$ , $\alpha 2$ , $\alpha 4$	0.50 ± 0.06	1.05 ± 0.13	0.79 ± 0.07	1.56 ± 0.15

**Tabelle 5-3:** Strukturstatistik für das berechnete Ensemble von E. coli Par10.

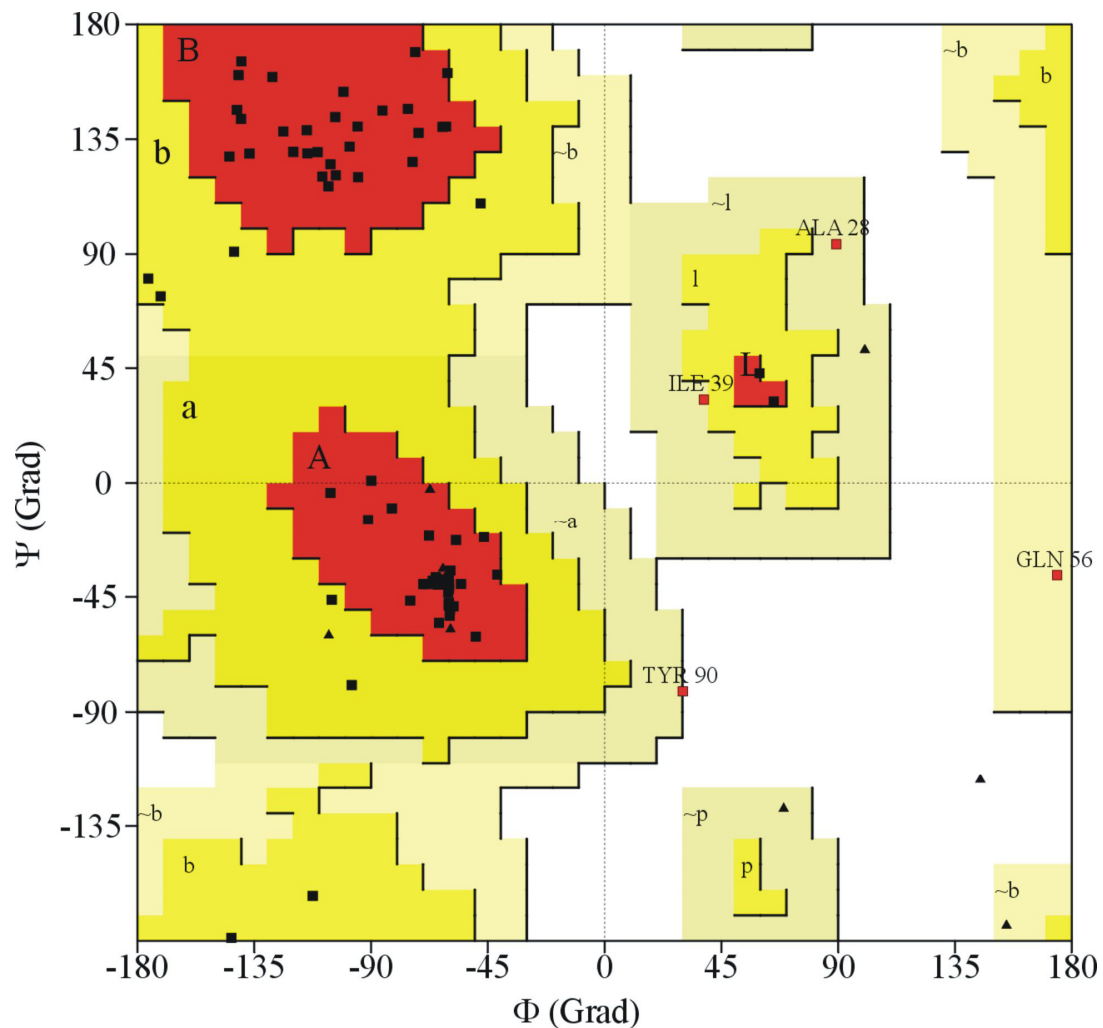
	SA 18 Strukturen [%]	$\langle SA \rangle_r$ optimierte Struktur [%]
RMSD-Werte:		
Distanzen [Å]		
Alle NOEs (1097)	0.023	0.018
Intraresiduale NOEs (494)	0.017	0.011
Sequentielle NOEs (272)	0.034	0.026
Nahbereichs-NOEs* (133)	0.023	0.023
Weitbereichs-NOEs* (168)	0.019	0.015
Wasserstoffbrücken (30)	0.006	0.012
Diederwinkel [°]	0.098	0.096
$^3\text{J-H}^{\text{N}}\text{H}^{\alpha}$ -Kopplungskonstanten [Hz] (42)	0.906	0.817
Abweichungen von idealer, kovalenter Geometrie		
Bindungslängen [Å]	0.002	0.002
Bindungswinkel [°]	0.561	0.532

(\*) Nahbereich: 2 bis 4 Aminosäuren, Weitbereich: > 4 Aminosäuren Entfernung.

Wie bereits beschrieben sind Ensemblerechnungen oder Rechnungen mit zeitlich gemittelten *restraints* die Methoden der Wahl, um die Struktur von flexiblen, kleinen Molekülen zu berechnen. Für Proteine kommen diese Verfahren jedoch (noch) aufgrund des hohen Rechenaufwands nicht in Betracht. Die Berechnung muss also „statisch“, d. h. ohne implizite Berücksichtigung der Dynamik erfolgen. Generell verbleiben somit zwei Vorgehensweisen zur Behandlung von stark flexiblen Regionen in Proteinen. Derartige Regionen sind meist durch äußerst widersprüchliche NOE-Muster gekennzeichnet, da sich die jeweiligen Distanzen der einzelnen, beitragenden Konformationen gemäß  $\langle r^{-6} \rangle^{-1/6}$  mitteln. Eine Möglichkeit ist es, für die flexible Region nur diejenigen NOEs einzusetzen, die auf eine einzelne der beitragenden Konformationen passen. Die Rechnung führt dann zu einer Struktur mit optimaler kovalenter Geometrie in der flexiblen Region, die allerdings nicht durch die experimentellen Daten gedeckt ist und somit nicht unbedingt der Realität entspricht. Die andere Möglichkeit ist es, alle beobachteten NOEs einzusetzen. Diese Methode bietet den Vorteil, ein unverfälschteres Bild der Wirklichkeit zu vermitteln und wurde daher für die Berechnung von Par10 gewählt. Allerdings entstehen dadurch zwangsweise verzerrte Geometrien in den flexiblen Regionen der berechneten Strukturen, die sich unter anderem im RAMACHANDRAN-Diagramm der gemittelten, optimierten Struktur bemerkbar machen. Tabelle 5-4 gibt die Verteilung der Aminosäurereste auf die einzelnen Bereiche des RAMACHANDRAN-Diagramms an. In Abbildung 5-9 ist zu erkennen, dass die in den flexiblen Regionen liegenden Reste Q56, I39 und A28 in geringfügig erlaubten Bereichen auftreten und das C-terminale Y90 sich im verbotenen Bereich befindet.

**Tabelle 5-4:** Verteilung der Aminosäurereste auf Bereiche im Ramachandran-Diagramm. Aufgeführt ist die Statistik für die Einzelstrukturen SA des berechneten Strukturensambles von E. coli Par10 sowie der gemittelten, optimierten Struktur  $\langle SA \rangle_r$ .

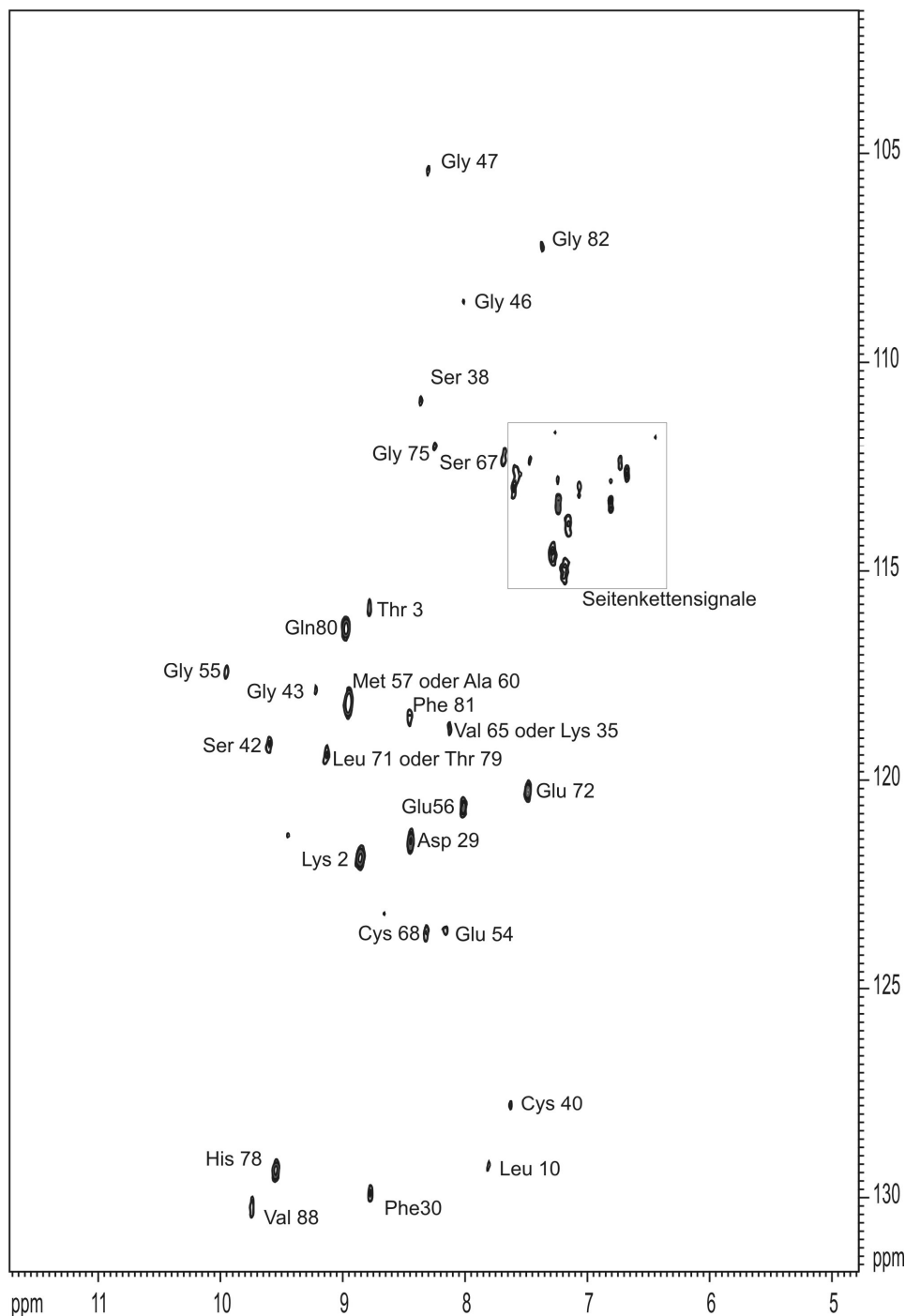
	SA (18 Strukturen) [%]	$\langle SA \rangle_r$ [%]
Bevorzugter Bereich	82.8	82.9
Erweitert erlaubter Bereich	9.6	13.2
Großzügig erlaubter Bereich	4.9	2.6
Verbotener Bereich	2.6	1.3



**Abbildung 5-9:** RAMACHANDRAN-Diagramm der gemittelten und optimierten Struktur von E. coli Par10, erstellt mit dem Programm PROCHECK<sup>[200]</sup>. Die Regionen A, B und L entsprechen Winkelkombinationen helicaler (A), Faltblatt-artiger (B) und linksgängig helicaler (L) Geometrie. Die Farbabstufung entspricht den Bereichen: Bevorzugt (rot), erweitert erlaubt (gelb), großzügig erlaubt (blassgelb) und verboten (weiß). Für Glycine gelten diese Bereiche nicht. A28, I39, Q56 und Y90 befinden sich in flexiblen Regionen der Struktur und daher in ungünstigen Bereichen im Diagramm. Gly: ▲, andere Aminosäuren: ■.

Wie bereits erwähnt, macht sich die Flexibilität einiger Regionen in Par10 auch im MEXICO-Experiment<sup>[45]</sup> bemerkbar. In diesem Spektrum entstehen starke Kreuzsignale immer dann, wenn ein Amidproton in starkem Austausch mit dem (protischen) Solvens steht. Ist dahingegen ein H<sup>N</sup> an einer stabilen Wasserstoffbrückenbindung beteiligt oder sterisch vom Lösungsmittel abgeschirmt (und liegt somit im meist rigiden Inneren des Proteins), bleibt das entsprechende Signal sehr schwach oder fehlt völlig. Insbesondere für die Amidprotonen der

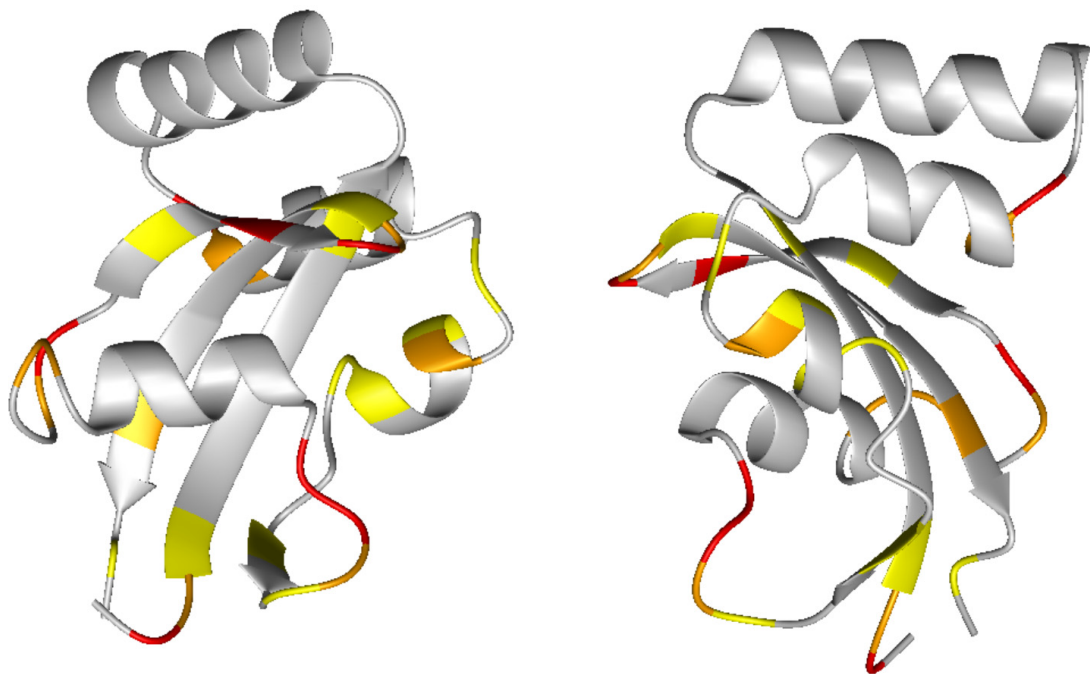
Aminosäurereste in den Termini und in den Schleifenregionen sind deutlich starke Signale im MEXICO-Spektrum von Par10 sichtbar (Abbildung 5-10).



**Abbildung 5-10:** MEXICO-Spektrum<sup>[45]</sup> von *E. coli* Par10 bei einer Mischzeit von 200 ms.

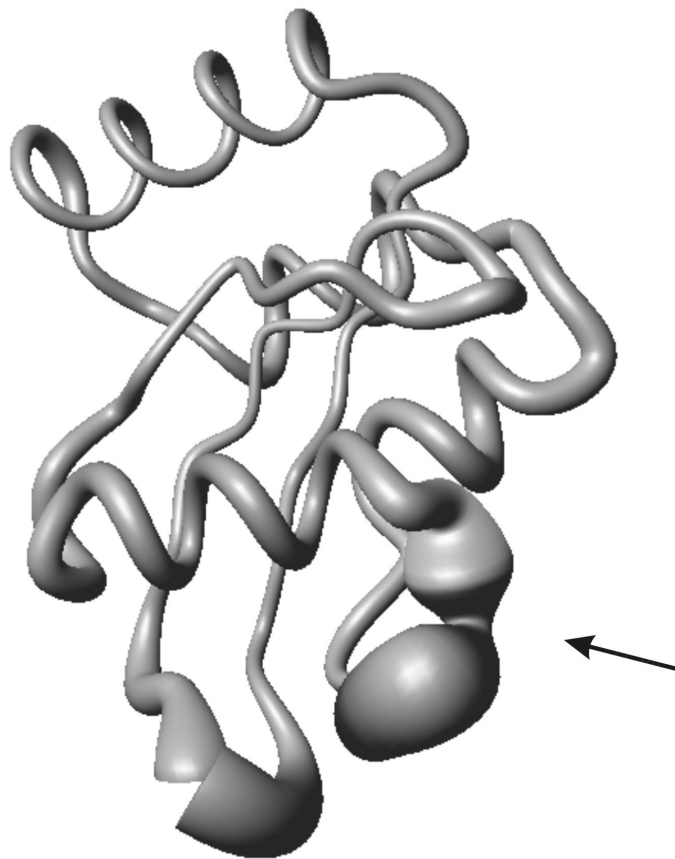
Für Aminosäuren in den Helices 1, 2 und 4 treten erwartungsgemäß keine Signale im MEXICO-Spektrum auf. Dahingegen zeigt sich an schwachen Signalen für G43, G46 und

G47, dass Helix 3 einigermaßen flexibel und nicht so gut definiert ist wie die anderen drei Helices. Wie zu erwarten, sind mit wenigen Ausnahmen auch keine Signale für Aminosäuren innerhalb des  $\beta$ -Faltblatts zu beobachten. Lediglich die zum Solvens ausgerichteten  $H^N$  der Aminosäurereste L10, H78 und V88 zeigen schwachen bis starken Austausch. Die mit starken Signalen im MEXICO-Spektrum vertretenen Reste Q80 und F81 liegen in Position  $i+1$  bzw.  $i+2$  einer  $\beta$ I-Schleife, womit ihre Amidprotonen ebenfalls zum Lösungsmittel zeigen. Alle weiteren Aminosäuren, für die Signale zu beobachten sind, liegen in flexiblen Schleifenregionen bzw. an den ebenfalls flexiblen Termini. Zur besseren Veranschaulichung zeigt Abbildung 5-11 die Struktur von Par10 mit den Informationen, die sich aus den MEXICO-Spektren ergeben:



**Abbildung 5-11:** Graphische Darstellung von *E. coli* Par10 mit dem Informationsgehalt der MEXICO-Spektren.<sup>[45]</sup> Das rechte Bild entsteht aus dem linken durch Drehung um die vertikale Achse. Gelb: Schwache, Orange: Mittelstarke, Rot: Starke Signale. Je stärker das Signal, umso mehr ist ein Amidproton dem Solvens zugänglich.

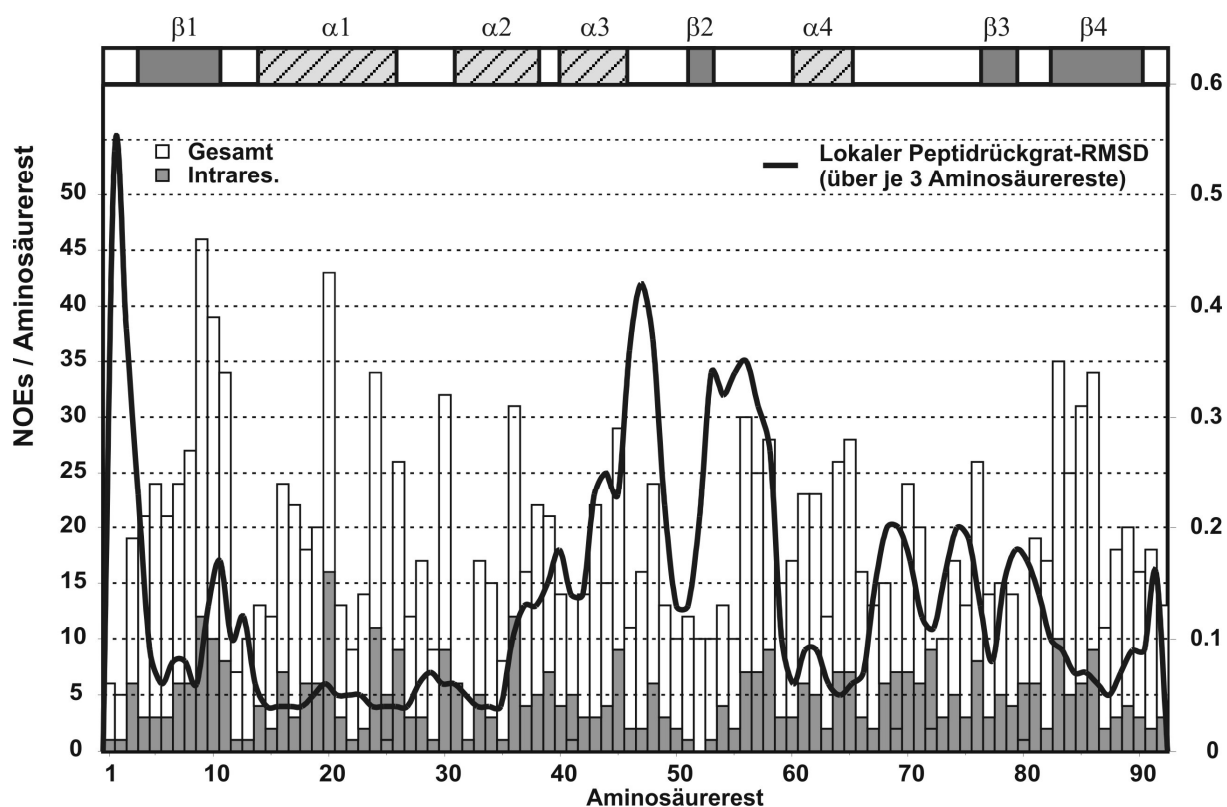
Abbildung 5-12 zeigt eine weitere graphische Präsentation der Flexibilität der einzelnen Regionen, indem der lokale RMSD-Wert auf die Struktur projiziert wird. Besonders deutlich wird dabei die Flexibilität im Bereich der flexiblen Ankerregion zwischen  $\beta$ -Strang 2 und Helix 4.



**Abbildung 5-12:** Lokale, über 3 Aminosäurereste berechnete RMSDs im Strukturenssemble von Par10, projiziert auf die gemittelte Struktur. Hohe Flexibilität zeigt sich insbesondere im Bereich zwischen  $\beta$ -Strang 2 und Helix 4 (Pfeil) sowie an den Termini. Die Grafik wurde von G. GEMMECKER mit dem Programm MOLMOL unter Verwendung des Makros sausage.mac erzeugt.

Wie aus dem Vergleich des lokalen RMSD mit der Anzahl der NOEs pro Aminosäurerest ersichtlich wird, nimmt die Flexibilität in einzelnen Bereichen trotz annähernd gleich bleibender Anzahl zu (Abbildung 5-13):





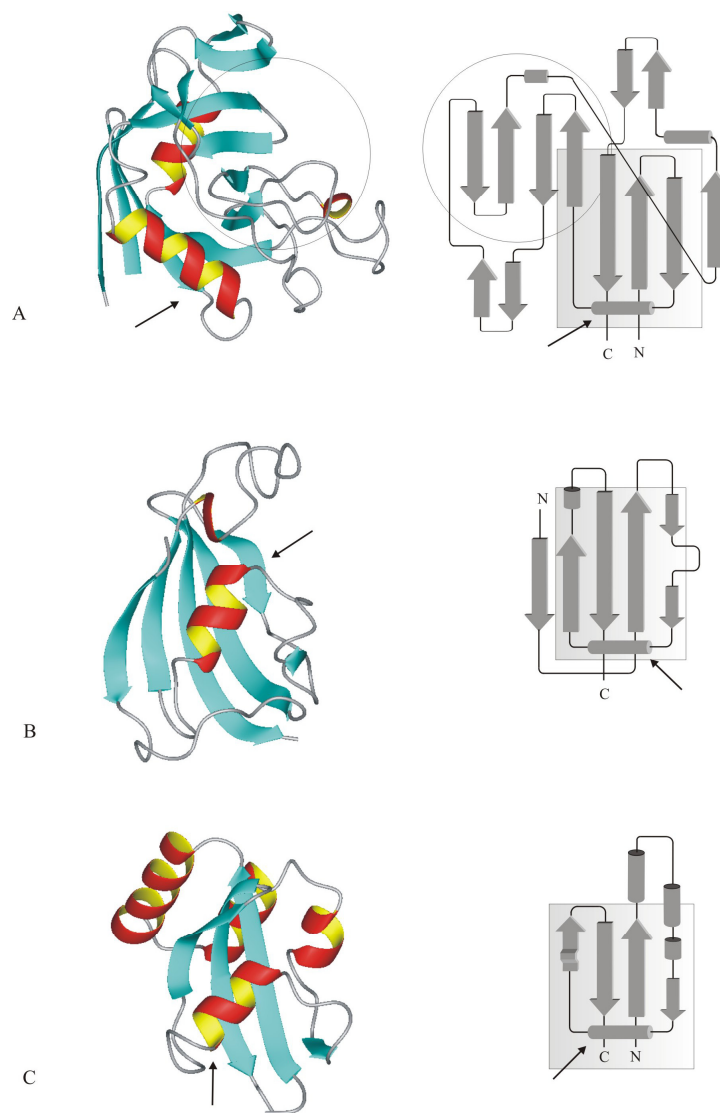
**Abbildung 5-13:** Vergleich der Anzahl von NOEs pro Aminosäurerest mit dem lokalen Peptidrückgrat-RMSD, der über jeweils 3 Aminosäuren berechnet wurde.

## 5.6 Strukturvergleich mit anderen PPIasen

Das Parvulin *hPin1*<sup>[191, 192]</sup> des Menschen spielt eine entscheidende Rolle in einer Reihe von Krankheitsbildern wie beispielsweise Krebs und Alzheimer.<sup>[153-159]</sup> Unter anderem fungiert *hPin1* als Peptidyl-Prolyl-*cis/trans*-Isomerase (PPIase). Im Gegensatz zu Parvulinen des Typs Pin1 (auch bezeichnet als „Dodo“-Parvuline, nach dem entsprechenden Parvulin von *Drosophila melanogaster*) fehlt in prokaryotischen Parvulinen wie *E. coli* Par10 eine N-terminale WW-Domäne, die für Wechselwirkungen mit Proteinen bzw. DNA zuständig ist. Nachdem mittlerweile auch Vertreter des Par10-Typs im Menschen und anderen Eukaryoten bekannt geworden sind,<sup>[196]</sup> gewinnt die Struktur von Par10 besonderes Interesse, da es den Minimalsatz an nötigen Strukturelementen besitzt, um PPIase-Aktivität zu entfalten.

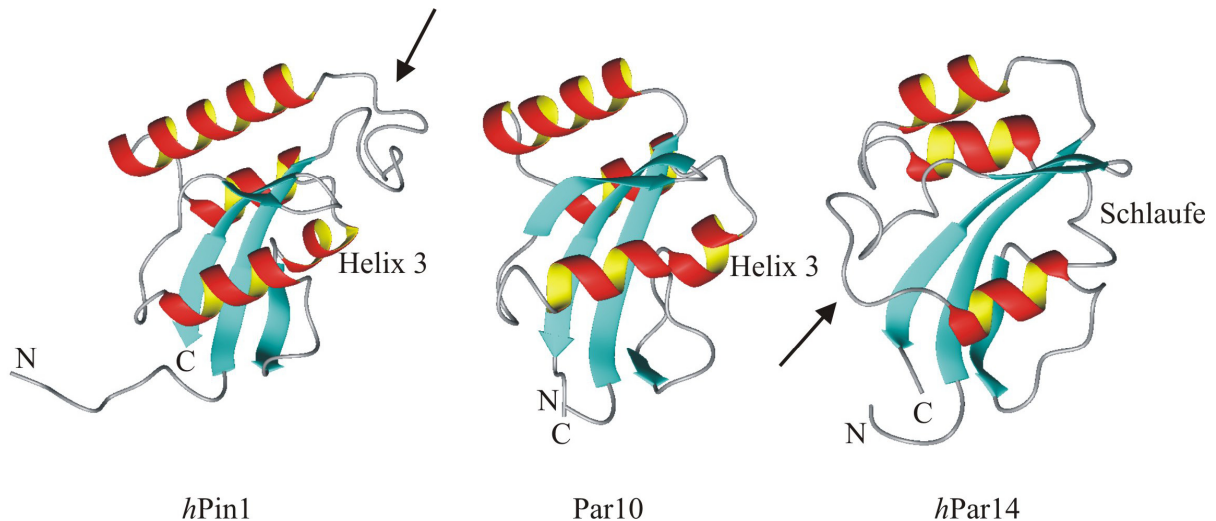
Beim Vergleich der PPIase-Familien zeigt sich, dass zwischen den Parvulinen und den Cyclophilinen bzw. den FKBP zwar keine Sequenzhomologie besteht, die aktiven Zentren jedoch eine äußerst ähnliche Struktur besitzen. Das gemeinsame Faltungsmotiv ist jeweils ein viersträngiges, antiparalleles  $\beta$ -Faltblatt, welches mehr oder minder stark gekrümmt ist und in

dessen konkaver Seite, leicht gewinkelt, eine  $\alpha$ -Helix liegt. Abbildung 5-14 zeigt den strukturellen und topologischen Vergleich zwischen Cyclophilin A, FKBP12 und Par10 als typischen Vertretern der drei PPIase-Familien. Dabei bildet Cyclophilin in gewisser Weise eine Ausnahme, da in ihm gleich zwei der beschriebenen Motive vorhanden sind, wobei die beiden Faltblätter gemeinsam ein  $\beta$ -Fass bilden. Das aktive Zentrum liegt in Cyclophilin A in dem zusätzlichen Strukturmotiv, welches zwischen den entsprechenden  $\beta$ -Strängen 3 und 4 von Par10 eingeschoben ist.



**Abbildung 5-14:** Strukturvergleich von Vertretern der PPIase-Familien.<sup>[207]</sup> A: Cyclophilin A (PDB: 1LOP), B: FKBP12 (PDB: 1FKT, 1FKR), C: Par10. Das grau hinterlegte Strukturmotiv findet sich in Vertretern aller PPIase-Familien und stellt im Allgemeinen das aktive Zentrum dar. In Cyclophilin A ist das mit einem Kreis versehene Strukturmotiv das aktive Zentrum. Die den Strukturen gemeinsame Helix ist mit einem Pfeil gekennzeichnet.

Während zwischen den Familien der PPIasen keine Sequenzhomologie vorliegt, sind viele Aminosäurereste zwischen einzelnen Mitgliedern einer Familie hochkonserviert. Wie bereits erwähnt, besitzen eukaryotische Parvuline allerdings meist eine N-terminale WW-Domäne sowie weitere Bereiche, die in prokaryotischen Parvulinen fehlen. Beispielsweise besitzt das in menschlichen Zellen vorkommende *hPar14*<sup>[195]</sup> einen basischen, unstrukturierten und Glycin-reichen N-Terminus mit 34 Aminosäureresten, der für die Erkennung von DNA zuständig ist und das Enzym in den Zellkern dirigiert.<sup>[208-210]</sup> Das Strukturmotiv des gekrümmten, viersträngigen  $\beta$ -Faltblatts mit darin liegender Helix ist jedoch bei allen Parvulinen mit bekannter Struktur besonders ähnlich, wie der Vergleich der Struktur von Par10 mit den Kristallstrukturen von *hPin1*<sup>[176]</sup> und *E. coli* SurA<sup>[211]</sup> sowie mit den NMR-Strukturen von *hPar14*<sup>[208, 212]</sup> und Pin1At<sup>[213]</sup> (aus *Arabidopsis thaliana*) in Lösung zeigt. Die folgenden zwei Abbildungen vermitteln einen Eindruck der hohen Sequenz- und Strukturhomologie innerhalb der Familie der Parvuline. Abbildung 5-15 zeigt einen Vergleich zwischen den drei Parvulinen *hPin1*, *hPar14* und *E. coli* Par10. In Abbildung 5-16 ist ein Sequenzvergleich zwischen diesen Proteinen dargestellt.



**Abbildung 5-15:** Strukturen der PPIase-Domänen der drei homologen Parvuline *hPin1*, *E. coli* Par10 und *hPar14*.<sup>[207]</sup> Die ungeordneten Einschübe in *hPin1* und *hPar14* sind mit einem Pfeil gekennzeichnet. Die in Par10 stark flexible Helix 3 tritt auch in der Struktur von *hPin1* auf, fehlt jedoch in *hPar14*.

```

PPIC_ECOLI      1  AKTAAALHILVKEEK-----LALDLLEQIKNGADFGKLAKKH-----
Par14           33  G-NAVKVRHILCEKHG-----KIMEAMEKLGKSGMRFNEVAAQY-----
PIN1_HUMAN     51  QEPAKVRCSHLLVKHSQSRPSSWRQEKITRSKEEALELINGYIQKIKSGEED-FESLASQ

PPIC_ECOLI     38  -SICPSGKRGGDLGFEFRQQMVPAFDKVVVFSCPV-L-E-P--T-GP[LHTQFGYHI IKVLYRN
Par14           70  -SEDK-ARQGGDLGWMTRGSMVGFQEAALPVS[SGMDKPVFTDP[VVKTKFGYHIIMVEGRK
PIN1_HUMAN    111  FSDCSSAKARGDLGPFPSRGQMOKPFEDASFALRT-G-E-M--S-G[VFTDSGIHIILRTE

```

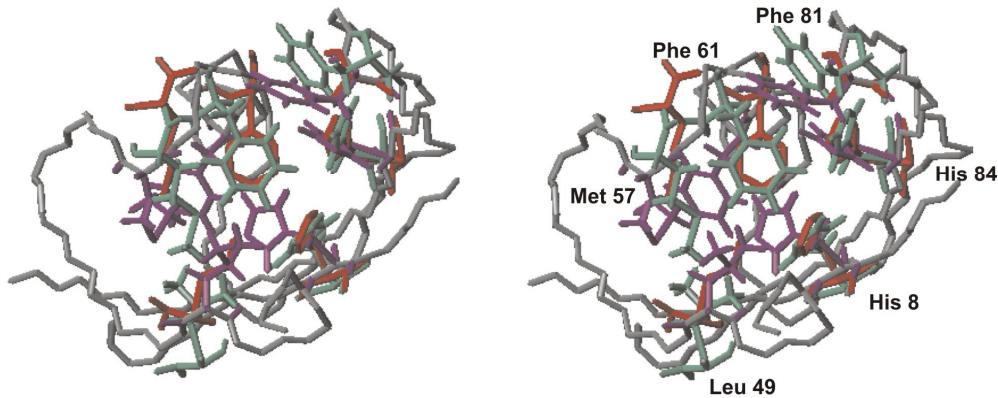
**Abbildung 5-16:** Sequenzvergleich der PPIase-Domänen der drei homologen Parvuline hPin1, E. coli Par10 und hPar14. Grau hinterlegte Felder sind strukturell ungeordnete Einschübe, die in Par10 fehlen. Die Felder mit schwarzem Hintergrund sind die konservierten Proline, für die eine cis-Konfiguration in Par10 gefunden wurde (siehe folgenden Text). Der Vergleich wurde von A. KÜHLEWEIN mit dem Programm FASTA erstellt.<sup>[207]</sup>

Wie bereits erwähnt, ist der Mechanismus der *cis-trans*-Isomerisierung bisher noch nicht aufgeklärt worden, obschon dazu einige Varianten diskutiert worden sind. Für hPin1 wurde von RANGANATHAN *et al.* ein hypothetischer Mechanismus vorgeschlagen, in dem die Reste L122, M130 und F134 die Bindungstasche bilden. Die Reste H59, H157, C113 und S154 sind demgemäß um die Rotationsachse der zu isomerisierenden Peptidbindung verteilt und wirken als katalytische Kaskade. Allerdings sind die Reste C113 und S154 des hPin1 in den anderen Homologen nicht konserviert, womit dieser Mechanismus für Par10 nicht in Frage kommt. Zur Übersicht sind in Tabelle 5-5 die Reste der Bindungstaschen von Par10, hPin1 und hPar14 aufgeführt:

**Tabelle 5-5:** Konservierte Aminosäurereste in den aktiven Zentren von hPin1, hPar14 und E. coli Par10. Fett hervorgehoben sind Aminosäurereste, die unterschiedlich sind.

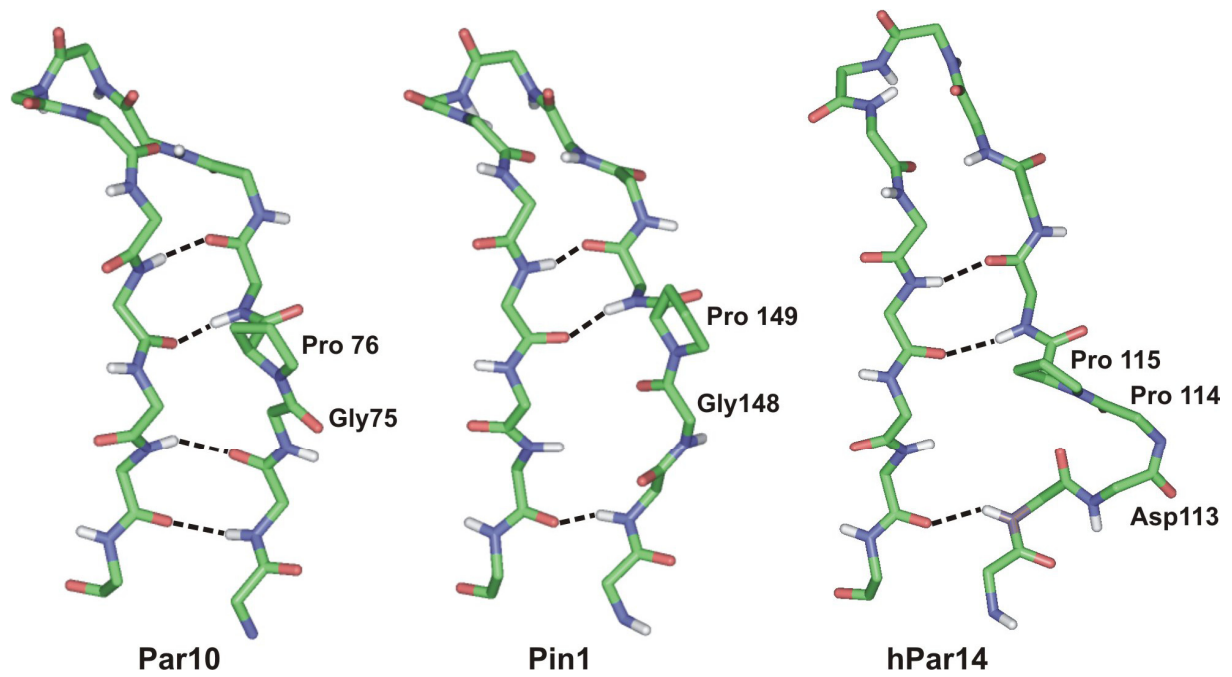
	<b>Par10</b>	<b>hPin1</b>	<b>hPar14</b>
Aktives Zentrum	Leu 49	Leu 122	Leu 82
	Met 57	Met 130	Met 90
	Phe 61	Phe 134	Phe 94
Katalytisch aktive Reste	His 8	His 59	His 42
	His 84	His 157	His 123
	<b>Cys 40</b>	<b>Cys 113</b>	<b>Asp 74</b>
	<b>Phe 81</b>	<b>Ser 154</b>	<b>Phe 120</b>

Abbildung 5-17 zeigt die entsprechenden Reste sowie die hohe strukturelle Ähnlichkeit der aktiven Zentren der drei Enzyme.



**Abbildung 5-17:** Stereodarstellung der superpositionierten aktiven Zentren von *E. coli* Par10 (violett), hPin1 (rot) und hPar14 (grün). Beschriftet sind einzelne Reste für Par10.

Wie schon erwähnt besitzt Par10 im  $\beta$ -Strang 3 einen charakteristischen Knick, der in den homologen Parvulinen hPin1 und hPar14 an analoger Stelle ebenfalls zu beobachten ist und im Falle von Par10 auf eine *cis*-konfigurierte Amidbindung zwischen G75 und P76 zurückgeführt werden konnte. Sowohl die Struktur von hPin1 als auch diejenige von hPar14 wurden mit einer *trans*-konfigurierten Bindung publiziert. Während eine *cis*-Bindung bei der Bestimmung der NMR-Struktur von hPar14 in Lösung<sup>[208, 212]</sup> vermutet, jedoch nicht mit Sicherheit bestätigt werden konnte, fehlen naturgemäß entsprechende Hinweise für die mit RÖNTGEN-Beugungsmethoden bestimmte Struktur von hPin1<sup>[176]</sup>. Die an der Stelle des Knicks mit Par10 identische Sequenz G-P (Abbildung 5-16, s. o.) lässt zumindest vermuten, dass die Bindung auch in hPin1 *cis*-konfiguriert ist. Wie in Abbildung 5-18 ersichtlich wird, erscheint eine *cis*-Bindung an dieser Stelle aus energetischer Sicht günstiger, da auf diese Weise trotz des Knicks ein nahezu komplettes Wasserstoffbrücken-Bindungsmuster ausgebildet werden kann.



**Abbildung 5-18:** Die *cis*-konfigurierte Peptidbindung zwischen G75 und P76 erlaubt in Par10 an der Knickstelle eine bessere Ausbildung des Wasserstoffbrückenmusters. Eine zusätzliche Wasserstoffbrückenbindung tritt gegenüber den Homologen hPin1 und hPar14 auf.

## 6 Zusammenfassung und Ausblick

Die vorliegende Arbeit beschäftigt sich mit strukturellen Fragestellungen, die im Rahmen des rationalen Designs von Wirkstoffen von Bedeutung sind. Dabei wurden die Vorzugskonformationen von Peptidmimetika berechnet, die für das indirekte Design von neuartigen Wirkstoffen mit verbesserten pharmakologischen Eigenschaften von Interesse sind. Darüber hinaus wurde die dreidimensionale Struktur eines Proteins in Lösung bestimmt. Räumliche Strukturen von Proteinen dienen als Ausgangspunkt zum Verständnis der molekularen Erkennung und ermöglichen die Modellierung kleinerer Moleküle, die in pathogene Prozesse eingreifen (direktes Wirkstoffdesign).

In Kapitel 2 werden zunächst die Methoden zur Aufklärung der Struktur und Dynamik von Peptiden und Proteinen in Lösung beschrieben. Daneben wird auf neuere Methoden zur automatisierten Auswertung von NMR-Spektren und Berechnung von Strukturen eingegangen. Mit entsprechenden Programmen kann das Risiko vermindert werden, zu Beginn der Strukturberechnung eines Proteins eine falsche Konformation aufzufinden und auf dieser basierend zu einer nicht der Realität entsprechenden Struktur zu gelangen.

Durch den Einsatz von Zuckeraminosäuren, also Zuckerbausteinen mit mindestens einer Carboxyl- und einer Aminogruppe, lassen sich Peptidmimetika mit veränderter Seitenketten- und Peptidrückgratanordnung erzeugen. Auf diese Weise können Wirkstoffe entwickelt werden, die günstigere pharmakologische Eigenschaften als ihre peptidischen Leitstrukturen besitzen. In Kapitel 3 wird gezeigt, dass der Einbau zweier solcher Zuckeraminosäuren (des „Gum“-Bausteins) in cyclische Hexapeptid-Analoga nicht wie erwartet zu einer Konformation zweier gegenüberliegender  $\beta$ -Schleifen führt. Weiterhin wird veranschaulicht, wie sich die oftmals zur Strukturkontrolle eingesetzte Vereinfachung cyclischer Peptide auf Cycloalkane nach DUNITZ und WASER auch auf Verbindungen wie *cyclo*(-Gum<sup>1</sup>-Lys<sup>2</sup>-Gum<sup>3</sup>-D-Phe<sup>4</sup>-) (**4**) anwenden lässt. Dabei hat sich herausgestellt, dass letzteres eine Konformation annimmt, die derjenigen des Cyclohexans in der „*twisted boat*“-Konformation entspricht.

Während  $\alpha$ -Peptide mittlerweile recht gründlich untersucht sind, besitzen aus  $\beta$ -Aminosäuren aufgebaute Peptide oftmals eine außergewöhnliche Struktur. Das in dieser Arbeit berechnete Fmoc-[f-SAA- $\beta$ -hGly]<sub>3</sub>-OH (**5**) ist ein Trimer eines Bausteins, der peptidisch aus einer furanoiden  $\beta$ -Zuckeraminosäure und  $\beta$ -Homoglycin zusammengesetzt ist. Aufgrund der

Flexibilität von  $\beta$ -hGly wurde zunächst erwartet, dass **5** keine geordnete Struktur in Lösung einnehmen würde. Es konnte jedoch gezeigt werden, dass die Struktur von **5** in  $\text{CD}_3\text{CN}$  einer helicalen Anordnung entspricht, in der alternierend 12- und 10-gliedrige Wasserstoffbrücken-gebundene Ringe auftreten. Die Ausrichtung der Wasserstoffbrücken alterniert dabei ebenfalls, woraus ein sehr geringes Gesamt-Dipolmoment der Helix resultiert. Damit wiederum ergibt sich eine sehr hohe Stabilität der Helix in unpolaren Lösungsmitteln, wie mittels einer Langzeit-Dynamiksimulation von 3 ns Dauer in einer zu diesem Zweck erstellten  $\text{CD}_3\text{CN}$ -Lösungsmittelzelle gezeigt werden konnte. Bisher wurde eine analoge Struktur nur in wenigen Fällen bei Verbindungen beobachtet, die gegenüber Verbindung **5** ein sehr unterschiedliches Substitutionsmuster am Peptidrückgrat besitzen. Wie sich herausgestellt hat, deckt sich die für **5** ermittelte Konformation durch Analogieschlüsse aus Ergebnissen quantenmechanischer Untersuchungen, die in der Gruppe von HOFMANN *et al.* durchgeführt worden waren.

Im Rahmen dieser Arbeit wurde die dreidimensionale Struktur des Proteins Par10 aus dem Cytoplasma von *Escherichia coli* aufgeklärt. Par10 gehört zur Parvulin-Familie der Peptidyl-Prolyl-*cis/trans*-Isomerasen (PPIasen), welche die Isomerisierung der Peptidbindung zwischen einer Aminosäure und Prolin katalysieren. Während eukaryotische Parvuline wie das hPin1 des Menschen meist eine N-terminale WW-Domäne besitzen, fehlt diese im Par10. Mit 10.1 kDa ist Par10 relativ klein und beinhaltet somit den Minimalsatz an Strukturelementen, der für die PPIase-Aktivität erforderlich ist. Menschliche Parvuline sind an einer Reihe von pathologischen Prozessen wie Krebs oder ALZHEIMER beteiligt. Die Kenntnis der Struktur von Par10 ist somit von Bedeutung, insbesondere zumal neuerdings auch Parvuline ohne WW-Domäne in menschlichen Zellen gefunden wurden.

Für Enzyme aus den anderen beiden PPIase-Familien, den Cyclophilinen und den FKBP, sind selektive Inhibitoren bekannt, die als Immunsuppressiva eingesetzt werden. Die Parvuline, für die ein solcher Inhibitor noch nicht gefunden wurde, könnten daher ebenfalls als potentielle Zielrezeptoren für Immunsuppressiva in Betracht kommen.

Für Par10 wurde eine  $\alpha\beta\alpha$ -sandwich-Struktur aus jeweils vier  $\alpha$ -Helices und  $\beta$ -Strängen ermittelt. Letztere bilden ein stark gewölbtes, antiparalleles  $\beta$ -Faltblatt der Strangordnung 2143, das mit seiner konkaven Seite die vierte Helix umschließt. Auf der konvexen Seite des Faltblatts liegen die Helices 1 bis 3. Zwischen dem Faltblatt und Helix 4 liegt das postulierte aktive Zentrum des Enzyms, in dem viele in Parvulinen hochkonservierte Reste liegen.



Die komplizierte, autokatalysierte Rückfaltungskinetik von denaturiertem Par10 konnte auf das Vorliegen eines *cis*-konfigurierten Prolins an einer auch in anderen Parvulinen vorkommenden, charakteristischen Knickstelle im  $\beta$ -Faltblatt zurückgeführt werden.

## 7 Literaturverzeichnis

- [1] J. M. Lehn, *Angew. Chem.* **1988**, *100*, 91-100.
- [2] H.-J. Böhm, G. Klebe, H. Kubinyi, *Wirkstoffdesign*, Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford **1996**.
- [3] C. L. M. J. Verlinde, W. G. J. Hol, *Structure* **1994**, *2*, 577-587.
- [4] M. Miller, J. Schneider, B. K. Sathyanarayana, M. V. Toth, G. R. Marshall, L. Clawson, L. Selk, S. B. H. Kent, A. Wlodawer, *Science* **1989**, *246*, 1149-1152.
- [5] C. Weber, G. Wider, B. von Freyberg, R. Traber, W. Braun, H. Widmer, K. Wüthrich, *Biochemistry* **1991**, *30*, 6563-6574.
- [6] L. M. Blumenthal, *Theory and Application of Distance Geometry*, Chelsea, Bronx, New York **1970**.
- [7] T. F. Havel, *DISGEO, Quantum Chemistry Exchange Program, Exchange No. 507*, Indiana University **1986**.
- [8] T. F. Havel, *Prog. Biophys. Molec. Biol.* **1991**, *56*, 43-78.
- [9] G. M. Crippen, T. F. Havel, *Distance Geometry and Molecular Conformation*, John Wiley & Sons, New York **1988**.
- [10] T. F. Havel, K. Wüthrich, *Bull. Math. Biol.* **1984**, *46*, 673-698.
- [11] C. M. Oshiro, J. Thomason, I. D. Kuntz, *Biopolymers* **1991**, *31*, 1049-1064.
- [12] Havel, *Biopolymers* **1990**, *29*, 1565-1585.
- [13] Kuszewski, M. Nilges, A. T. Brünger, *J. Biomol. NMR* **1992**, *2*, 33-56.
- [14] M. Billeter, T. F. Havel, K. Wüthrich, *J. Comp. Chem.* **1987**, *8*, 132-141.
- [15] W. Braun, N. Go, *J. Mol. Biol.* **1985**, *186*, 611-626.
- [16] G. M. Crippen, *J. Comp. Chem.* **1982**, *3*, 471-476.
- [17] T. C. Beutler, W. F. van Gunsteren, *J. Chem. Phys.* **1994**, *101*, 1417-1422.
- [18] J. M. Haile, *Molecular Dynamics Simulation: Elementary Methods*, John Wiley & Sons, New York **1992**.
- [19] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford **1987**.
- [20] W. F. van Gunsteren, H. J. C. Berendsen, *Angew. Chem.* **1990**, *102*, 1020-1055.
- [21] M. Karplus, G. A. Petsko, *Nature* **1990**, *347*, 631-639.
- [22] T. P. Lybrand, *Reviews in Computational Chemistry*, (pp. 295-317), VCH, Weinheim **1990**.

- [23] P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, A. T. Hagler, *Proteins: Structure, Function and Genetics* **1988**, 4, 31-47.
- [24] A. T. Hagler, *Molecular Dynamics and Protein Structure*, (pp. 133-138), Polycrystal Book Service, Western Springs, IL, USA **1985**.
- [25] *Discover 2.9.7/95.0/3.00 User Guide*, Biosym/MSI, San Diego **1995**.
- [26] P. K. Weiner, P. A. Kollman, *J. Comp. Chem.* **1981**, 2, 287-303.
- [27] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. K. Weiner, *J. Amer. Chem. Soc.* **1984**, 106, 765-784.
- [28] S. W. Homans, *Biochemistry* **1990**, 29, 9110-9118.
- [29] Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comp. Chem.* **1986**, 7, 230-252.
- [30] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comp. Chem.* **1983**, 4, 187-217.
- [31] A. T. Brunger, *X-PLOR, version 3.1. A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT **1992**.
- [32] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, N. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. and Warren, *Acta Cryst.* **1998**, D54, 905-921.
- [33] J. P. Linge, M. Nilges, *J. Biomol. NMR* **1999**, 13, 51-59.
- [34] J. H. Konnert, W. A. Hendrickson, *Acta Cryst.* **1980**, A36, 344-350.
- [35] D. Neuhaus, M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH **1989**.
- [36] A. M. Gronenborn, G. M. Clore, *Prog. Nucl. Magn. Reson. Spectrosc.* **1985**, 17, 1-32.
- [37] M. Karplus, *J. Chem. Phys.* **1959**, 30, 11-15.
- [38] M. Karplus, *J. Am. Chem. Soc.* **1963**, 85, 2870-2871.
- [39] V. I. Bystrov, *Prog. Nucl. Magn. Reson. Spectrosc.* **1976**, 10, 41-81.
- [40] M. T. Cung, M. Maraud, J. Neel, *Macromolecules* **1974**, 606-613.
- [41] C. A. G. Haasnoot, F. A. A. M. Leeuw de, C. Altona, *Tetrahedron* **1980**, 36, 2783-2792.
- [42] D. F. Mierke, H. Kessler, *Biopolymers* **1993**, 33, 1003-1017.
- [43] D. S. Wishart, B. D. Sykes, *J. Biomol. NMR* **1994**, 4, 171-180.
- [44] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, 5, 67-81.
- [45] G. Gemmecker, W. Jahnke, H. Kessler, *J. Am. Chem. Soc.* **1993**, 115, 11620-11621.
- [46] M. Llinas, M. P. Klein, *J. Am. Chem. Soc.* **1975**, 97, 4731-4737.

- [47] Y. A. Bara, A. Friedrich, H. Kessler, M. Molter, *Chem. Ber.* **1978**, *111*, 1045-1057.
- [48] D. S. Dwyer, *Biopolymers* **1999**, *49*, 635-645.
- [49] C. L. Brooks III, M. Karplus, *J. Chem. Phys.* **1983**, *79*, 6312-6325.
- [50] D. J. Adams, *Chem. Phys. Lett.* **1979**, *62*, 329-332.
- [51] H. Schreiber, O. Steinhauser, *Biochemistry* **1992**, *31*, 5856-5860.
- [52] R. Gratias, Dissertation, Technische Universität München, **1998**.
- [53] D. H. J. Mackay, A. J. Cross, A. T. Hagler, *Prediction of Protein Structure and the Principals of Protein Conformation*, (pp. 317-357), Plenum Press, New York, USA **1990**.
- [54] K. B. Wiberg, *J. Am. Chem. Soc.* **1965**, *87*, 1070-1078.
- [55] R. Fletcher, *Practical Methods of Optimization, Unconstrained Optimization*, John Wiley & Sons, New York, USA **1980**.
- [56] R. W. Hockney, *Methods Comput. Phys.* **1970**, *9*, 136-211.
- [57] L. Verlet, *Phys. Rev.* **1967**, *159*, 98-103.
- [58] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, *J. Chem. Phys.* **1982**, *76*, 637-649.
- [59] Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNiola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684-3690.
- [60] H. C. Andersen, *J. Chem. Phys.* **1980**, *72*, 2384-2393.
- [61] S. Kirkpatrick, J. C. D. Gelatt, M. P. Vecchi, *Science* **1983**, *220*, 671-680.
- [62] M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* **1988**, *239*, 129-136.
- [63] A. Nayeem, J. Vila, H. A. Scheraga, *J. Comput. Chem.* **1991**, *12*, 594-605.
- [64] M. Nilges, J. Kuszewski, A. T. Brunger, *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy*, (pp. 451-455), Plenum Press, New York **1991**.
- [65] H. Kessler, C. Griesinger, J. Lautz, A. Müller, W. F. van Gunsteren, H. J. C. Berendsen, *J. Am. Chem. Soc.* **1988**, *110*, 3393-3396.
- [66] R. M. Scheek, A. E. Torda, J. Kemmink, W. F. van Gunsteren, *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy*, (pp. 209-217), NATO ASI Series A225, Plenum Press **1991**.
- [67] J. Kemmink, C. P. M. van Mierlo, R. M. Scheek, T. E. Creighton, *J. Mol. Biol.* **1993**, *230*, 312-322.
- [68] A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *Chem. Phys. Lett.* **1989**, *157*, 289-294.

- [69] A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *J. Mol. Biol.* **1990**, *214*, 223-235.
- [70] D. A. Pearlman, P. A. Kollman, *J. Mol. Biol.* **1991**, *220*, 457-479.
- [71] A. P. Nanzer, W. F. van Gunsteren, A. E. Torda, *J. Biomol. NMR* **1995**, *6*, 313-320.
- [72] M. Nilges, *Proteins: Struct. Funct. Genet.* **1993**, *17*, 297-309.
- [73] M. Nilges, *J. Mol. Biol.* **1995**, *245*, 645-660.
- [74] M. Nilges, M. J. Macias, S. I. O'Donoghue, H. Oschkinat, *J. Mol. Biol.* **1997**, *269*, 408-422.
- [75] S. I. O'Donoghue, M. Nilges, *Biological Magnetic Resonance*, (pp. 131-161), Kluwer Academic/ Plenum, New York **1999**.
- [76] Z. Liu, M. J. Macias, M. J. Bottomley, G. Stier, J. P. Linge, M. Nilges, P. Bork, M. Sattler, *Structure* **1999**, *12*, 1557-1566.
- [77] R. Sprangers, M. J. Bottomley, J. P. Linge, J. Schultz, M. Nilges, M. Sattler, *J. Biomol. NMR* **2000**, *16*, 47-58.
- [78] J. P. Linge, S. I. O'Donoghue, M. Nilges, *Meth. Enzymol.* **2001**, *339*, 71-90.
- [79] J. P. Linge, *New Methods for automated NOE assignment and NMR structure calculation*, Book on demand Verlag, Norderstedt, Germany **2001**.
- [80] J. P. Linge, M. Habeck, W. Rieping, M. Nilges, *Bioinformatics* **2003**, *19*, 315-316.
- [81] L. A. Kelley, S. P. Gardner, M. J. Sutcliffe, *Protein Eng.* **1996**, *9*, 1063-1065.
- [82] E. Graf von Roedern, H. Kessler, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 687-689.
- [83] E. Graf von Roedern, E. Lohof, G. Hessler, M. Hoffmann, H. Kessler, *J. Am. Chem. Soc.* **1996**, *118*, 10156-10167.
- [84] E. Lohof, E. Planker, C. Mang, F. Burkhart, M. A. Dechantsreiter, R. Haubner, H.-J. Wester, M. Schwaiger, G. Hölzemann, S. Goodman, L. Simon, H. Kessler, *Angew. Chem. Int. Ed.* **2000**, *39*, 2761-2764.
- [85] S. A. W. Gruner, E. Locardi, E. Lohof, H. Kessler, *Chem. Rev.* **2002**, *102*, 491-514.
- [86] E. Lohof, F. Burkhart, M. A. Born, E. Planker, H. Kessler, *Adv. Amino Acid Mimetics Peptidomimetics* **1999**, *2*, 263-292.
- [87] E. Lohof, Dissertation, Technische Universität München, **1998**.
- [88] E. Graf von Roedern, Dissertation, Technische Universität München, **1994**.
- [89] J. D. Dunitz, J. Waser, *J. Am. Chem. Soc.* **1972**, *94*, 5645-5650.
- [90] M. Mutter, S. Vuilleumier, *Angew. Chem.* **1989**, *101*, 551-571.
- [91] G. Tuchscherer, *Tetrahedron Lett.* **1993**, *34*, 8419-8422.
- [92] G. Tuchscherer, D. Grell, M. Mathieu, M. Mutter, *J. Pept. Res.* **1999**, *54*, 185-194.
- [93] H. Kessler, *Angew. Chem. Int. Ed. Engl.* **1982**, *21*, 512-523.

- [94] M. Kurz, Dissertation, Technische Universität München, **1991**.
- [95] H. Kessler, M. Bernd, H. Kogler, J. Zarbock, O. W. Soerensen, G. Bodenhausen, R. R. Ernst, *J. Am. Chem. Soc.* **1983**, *105*, 6944-6952.
- [96] H. Kessler, U. Anders, M. Schudok, *J. Am. Chem. Soc.* **1990**, *112*, 5908-5916.
- [97] H. Kessler, H. Matter, G. Gemmecker, H. J. Diehl, C. Isernia, S. Mronga, *Int. J. Pept. Protein Res.* **1994**, *43*, 47-61.
- [98] D. F. Mierke, M. Kurz, H. Kessler, *J. Am. Chem. Soc.* **1994**, *116*, 1042-1049.
- [99] G. Müller, M. Gurrath, H. Kessler, R. Timpl, *Angew. Chem.* **1992**, *104*, 341-343.
- [100] D. Demel, H. Kessler, *Tetrahedron Lett.* **1976**, *32*, 2801-2804.
- [101] H. Kessler, V. Eiermann, *Tetrahedron Lett.* **1982**, *23*, 4689-4692.
- [102] E. Locardi, M. Stöckle, S. Gruner, H. Kessler, *J. Am. Chem. Soc.* **2001**, *123*, 8189-8196.
- [103] K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* **1983**, *169*, 949-961.
- [104] E. Keller, *Schakal 99*, Kristallographisches Institut der Universität Freiburg i.Br., **1999, 2000**.
- [105] F. Johnson, *Chem. Rev.* **1968**, *68*, 375-413.
- [106] R. W. Hoffmann, *Chem. Rev.* **1989**, *89*, 1841-1860.
- [107] H. Sachse, *Berichte* **1890**, *23*, 1363-1370.
- [108] H. A. Lauwerier, *Proc. Kon. Ned. Akad. Wetensch.* **1966**, *69*, 330.
- [109] Bottema, *Proc. Kon. Ned. Akad. Wetensch.* **1967**, *70*, 151.
- [110] H. Kessler, R. Gratias, G. Hessler, M. Gurrath, G. Müller, *Pure & Appl. Chem.* **1996**, *68*, 1201-1205.
- [111] D. Seebach, J. L. Matthews, *Chem. Commun.* **1997**, (21), 2015-2022.
- [112] R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, *101*, 3219-3232.
- [113] Y. Suhara, M. Izumi, M. Ichikawa, M. B. Penno, Y. Ichikawa, *Tetrahedron Lett.* **1997**, *38*, 7167-7170.
- [114] J. Kovacs, R. Ballina, R. L. Rodin, D. Balasubramanian, J. Applequist, *J. Am. Chem. Soc.* **1965**, *87*, 119-120.
- [115] J. M. Fernández-Santín, J. Aymamí, A. Rodríguez-Galán, S. Muñoz-Guerra, J. A. Subirana, *Nature* **1984**, *311*, 53-54.
- [116] J. M. Fernandez-Santin, S. Munoz-Guerra, A. Rodriguez-Galan, J. Aymami, J. Lloveras, J. A. Subirana, E. Giralt, M. Ptak, *Macromolecules* **1987**, *20*, 62-68.
- [117] D. Seebach, M. Overhand, F. N. M. Kühnle, B. Martinoni, L. O. U. Hommel, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 913-941.

- [118] S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173-180.
- [119] D. Seebach, P. E. Ciceri, M. Overhand, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amtstutz, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 2043-2066.
- [120] X. Daura, W. F. van Gunsteren, D. Rigo, B. Jaun, D. Seebach, *Chem. Eur. J.* **1997**, *3*, 1410-1417.
- [121] D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1998**, *81*, 932-982.
- [122] D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071-13072.
- [123] G. P. Dado, S. H. Gellman, *J. Am. Chem. Soc.* **1994**, *116*, 1054-1062.
- [124] J. J. Navas, C. Aleman, S. Munoz-Guerra, *J. Org. Chem.* **1996**, *61*, 6849-6855.
- [125] F. Lopez-Carrasquero, M. Garcia-Alvarez, J. J. Navas, C. Aleman, S. Muñoz-Guerra, *Macromolecules* **1996**, *29*, 8449-8459.
- [126] D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. Huang, J. J. Barchi Jr, S. H. Gellman, *Nature* **1997**, *387*, 381-384.
- [127] K. A. Bode, J. Applequist, *Macromolecules* **1997**, *30*, 2144-2150.
- [128] C. Toniolo, *CRC Crit. Rev. Biochem.* **1980**, *9*, 1-44.
- [129] K. Möhle, R. Günther, M. Thormann, N. Sewald, H. J. Hofmann, *Biopolymers* **1999**, *50*, 167-184.
- [130] Y.-D. Wu, D.-P. Wang, *J. Am. Chem. Soc.* **1998**, *120*, 13485-13493.
- [131] Y.-D. Wu, D.-P. Wang, *J. Am. Chem. Soc.* **1999**, *121*, 9352-9362.
- [132] T. A. Martinek, F. Fülöp, *Eur. J. Biochem.* **2003**, *270*, 3657-3666.
- [133] C. Baldauf, R. Günther, H.-J. Hofmann, *Angew. Chem.* **2004**, *116*, 1621-1624.
- [134] M. Marraud, J. Neel, *J. Polym. Sci. Polym. Symp.* **1975**, *52*, 271-282.
- [135] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, J. A. Pople, *Gaussian 94 (Revision D.4)*, Gaussian, Inc., Pittsburgh, PA, **1995**.
- [136] *Insight II User Guide*, Biosym/MSI, San Diego **1995**.
- [137] P. A. Steiner, W. Gordy, *J. Mol. Spectr.* **1966**, *21*, 291-301.

- [138] S. A. W. Gruner, Dissertation, Technische Universität München, **2002**.
- [139] D. Seebach, S. Abele, T. Sifferlen, M. Hänggi, S. Gruner, P. Seiler, *Helv. Chim. Acta* **1998**, *81*, 2218-2243.
- [140] H. E. Stanger, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 4236-4237.
- [141] X. Wang, J. F. Espinosa, S. H. Gellman, *J. Am. Chem. Soc.* **2000**, *122*, 4821-4822.
- [142] M. Brenner, D. Seebach, *Helv. Chim. Acta* **2001**, *84*, 1181-1189.
- [143] D. Seebach, K. Gademann, J. V. Schreiber, J. L. Matthews, T. Hintermann, B. Jaun, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1997**, *80*, 2033-2038.
- [144] J. Applequist, K. A. Bode, D. H. Appella, L. A. Christianson, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 4891-4892.
- [145] H. L. Schenck, S. H. Gellman, *Journal of the American Chemical Society* **1998**, *120*, 4869-4870.
- [146] V. Daggett, M. Levitt, *J. Mol. Biol.* **1992**, *223*, 1121-1138.
- [147] A. R. Van Buuren, H. J. Berendsen, *Biopolymers* **1993**, *33*, 1159-1166.
- [148] H. De Loof, L. Nilsson, R. Rigler, *J. Am. Chem. Soc.* **1992**, *114*, 4028-4035.
- [149] J. Tirado-Rives, W. L. Jorgensen, *Biochemistry* **1991**, *30*, 3864-3871.
- [150] F. M. DiCapua, S. Swaminathan, D. L. Beveridge, *J. Am. Chem. Soc.* **1991**, *113*, 6145-6155.
- [151] D. M. Korzhnev, V. Y. Orekhov, A. S. Arseniev, R. Gratias, H. Kessler, *J. Phys. Chem. B* **1999**, *103*, 7036-7043.
- [152] P. Y. Chou, G. D. Fasman, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1978**, *47*, 45-148.
- [153] P. J. Lu, G. Wulf, X. Z. Zhou, P. Davies, K. P. Lu, *Nature* **1999**, *399*, 784-788.
- [154] G. M. Wulf, A. Ryo, G. G. Wulf, S. W. Lee, T. Niu, V. Petkova, K. P. Lu, *Embo J.* **2001**, *20*, 3459-3472.
- [155] Y. C. Liou, A. Ryo, H. K. Huang, P. J. Lu, R. Bronson, F. Fujimori, T. Uchida, T. Hunter, K. P. Lu, *Proc. Natl. Acad. Sci. U S A* **2002**, *99*, 1335-1340.
- [156] G. Wulf, A. Ryo, Y. C. Liou, K. P. Lu, *Breast Cancer Res.* **2003**, *5*, 76-82.
- [157] G. M. Wulf, Y. C. Liou, A. Ryo, S. W. Lee, K. P. Lu, *J. Biol. Chem.* **2002**, *277*, 47976-47979.
- [158] P. Zacchi, M. Gostissa, T. Uchida, C. Salvagno, F. Avolio, S. Volinia, Z. Ronai, G. Blandino, C. Schneider, G. Del Sal, *Nature* **2002**, *419*, 853-857.
- [159] H. Zheng, H. You, X. Z. Zhou, S. A. Murray, T. Uchida, G. Wulf, L. Gu, X. Tang, K. P. Lu, Z. X. Xiao, *Nature* **2002**, *419*, 849-853.
- [160] C. Levinthal, *J. Chim. Phys.* **1968**, *65*, 44-45.



- [161] C. J. Epstein, R. F. Goldberger, C. B. Anfinsen, *Cold Spring Harb Sym* **1963**, 28, 439-449.
- [162] C. B. Anfinsen, *Science* **1973**, 181, 223-230.
- [163] D. Givol, R. F. Goldberger, C. B. Anfinsen, *J. Biol. Chem.* **1964**, 239, 3114-3116.
- [164] D. E. Stewart, A. Sarkar, J. E. Wampler, *J. Mol. Biol.* **1990**, 214, 253-260.
- [165] M. W. Macarthur, J. M. Thornton, *J. Mol. Biol.* **1991**, 218, 397-412.
- [166] M. S. Weiss, A. Jabs, R. Hilgenfeld, *Nat. Struct. Biol.* **1998**, 5, 676-676.
- [167] M. S. Weiss, R. Hilgenfeld, *Biopolymers* **1999**, 50, 536-544.
- [168] A. Jabs, M. S. Weiss, R. Hilgenfeld, *J. Mol. Biol.* **1999**, 286, 291-304.
- [169] G. Fischer, H. Bang, C. Mech, *Biomed. Biochim. Acta* **1984**, 43, 1101-1111.
- [170] G. Fischer, E. Berger, H. Bang, *FEBS Lett.* **1989**, 250, 267-270.
- [171] G. Fischer, B. Wittmannliebold, K. Lang, T. Kiefhaber, F. X. Schmid, *Nature* **1989**, 337, 476-478.
- [172] G. Fischer, *Angew. Chem.* **1994**, 106, 1479-1501.
- [173] R. K. Harrison, R. L. Stein, *Biochemistry* **1990**, 29, 1684-1689.
- [174] R. K. Harrison, R. L. Stein, *J. Am. Chem. Soc.* **1992**, 114, 3464-3471.
- [175] R. L. Stein, *Adv. Protein Chem.* **1993**, 44, 1-24.
- [176] R. Ranganathan, K. P. Lu, T. Hunter, J. P. Noel, *Cell* **1997**, 89, 875-886.
- [177] R. E. Handschumacher, M. W. Harding, J. Rice, R. J. Drugge, D. W. Speicher, *Science* **1984**, 226, 544-547.
- [178] M. W. Harding, R. E. Handschumacher, D. W. Speicher, *J. Biol. Chem.* **1986**, 261, 8547-8555.
- [179] G. Fischer, H. Bang, *Biochim. Biophys. Acta* **1985**, 828, 39-42.
- [180] K. Lang, F. X. Schmid, G. Fischer, *Nature* **1987**, 329, 268-270.
- [181] K. Lang, F. X. Schmid, *Nature* **1988**, 331, 453-455.
- [182] N. Takahashi, T. Hayano, M. Suzuki, *Nature* **1989**, 337, 473-475.
- [183] A. R. Marks, *Physiol Rev* **1996**, 76, 631-649.
- [184] S. L. Schreiber, M. W. Albers, E. J. Brown, *Acc. Chem. Res.* **1993**, 26, 412-420.
- [185] A. Stütz, M. Grassberger, K. Baumann, A. Edmunds, P. Hiestand, J. Meingassner, P. Nussbaumer, W. Schuler, G. Zenke, *Perspectives in Medicinal Chemistry*, (pp. 427-443), VCH, Verlag Helvetica Chimica Acta (Basel), Weinheim **1993**.
- [186] F. Schmid, *Molecular Chaperones in the Life Cycle of Proteins - Structure, Function and Mode of Action*, Marcel Dekker, Inc. **1998**.

- [187] G. Fischer, F. Schmid, *Molecular Chaperones and Folding Catalysts - Regulation, Cellular Function and Mechanisms*, Harwood Academic Publishers **1999**.
- [188] J. U. Rahfeld, A. Schierhorn, K. Mann, G. Fischer, *FEBS Lett.* **1994**, *343*, 65-69.
- [189] J. U. Rahfeld, K. P. Rücknagel, B. Schelbert, B. Ludwig, J. Hacker, K. Mann, G. Fischer, *FEBS Lett.* **1994**, *352*, 180-184.
- [190] B. Schelbert, Dissertation, Martin Luther Universität Wittenberg, **1998**.
- [191] K. P. Lu, S. D. Hanes, T. Hunter, *Nature* **1996**, *380*, 544-547.
- [192] D. G. Crenshaw, J. Yang, A. R. Means, S. Kornbluth, *Embo J.* **1998**, *17*, 1315-1327.
- [193] W. R. Pearson, D. J. Lipman, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2444-2448.
- [194] W. R. Pearson, *Methods Enzymol.* **1990**, *183*, 63-98.
- [195] T. Uchida, F. Fujimori, T. Tradler, G. Fischer, J. U. Rahfeld, *FEBS Lett.* **1999**, *446*, 278-282.
- [196] S. Rulten, J. Thorpe, J. Kay, *Biochem Biophys Res Commun* **1999**, *259*, 557-562.
- [197] L. Hennig, C. Christner, M. Kipping, B. Schelbert, K. P. Rucknagel, S. Grabley, G. Kullertz, G. Fischer, *Biochemistry* **1998**, *37*, 5953-5960.
- [198] M. Leutner, R. M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker, H. Kessler, *J. Biomol. NMR* **1998**, *11*, 31-43.
- [199] T. Yamazaki, J. D. Forman-Kay, L. E. Kay, *J. Am. Chem. Soc.* **1993**, *115*, 11054-11055.
- [200] R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283-291.
- [201] J. Kuszewski, A. M. Gronenborn, G. M. Clore, *Protein Sci.* **1996**, *5*, 1067-1080.
- [202] J. Kuszewski, A. M. Gronenborn, G. M. Clore, *J. Magn. Reson.* **1997**, *125*, 171-177.
- [203] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946-950.
- [204] R. Konradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51-55.
- [205] H. Kessler, *Angew. Chem.* **1982**, *94*, 509-520.
- [206] C. Scholz, J. Rahfeld, G. Fischer, F. X. Schmid, *J. Mol. Biol.* **1997**, *273*, 752-762.
- [207] A. Kühlewein, Dissertation, Technische Universität München, **2001**.
- [208] E. Sekerina, J. U. Rahfeld, J. Muller, J. Fanghanel, C. Rascher, G. Fischer, P. Bayer, *J. Mol. Biol.* **2000**, *301*, 1003-1017.
- [209] T. A. Surmacz, E. Bayer, J. U. Rahfeld, G. Fischer, P. Bayer, *J. Mol. Biol.* **2002**, *321*, 235-247.
- [210] T. Reimer, M. Weiwad, A. Schierhorn, P. K. Ruecknagel, J. U. Rahfeld, P. Bayer, G. Fischer, *J. Mol. Biol.* **2003**, *330*, 955-966.

- 
- [211] E. Bitto, D. B. McKay, *Structure* **2002**, *10*, 1489-1498.
- [212] T. Terada, M. Shirouzu, Y. Fukumori, F. Fujimori, Y. Ito, T. Kigawa, S. Yokoyama, T. Uchida, *J. Mol. Biol.* **2001**, *305*, 917-926.
- [213] I. Landrieu, J. M. Wieruszeski, R. Wintjens, D. Inze, G. Lippens, *J. Mol. Biol.* **2002**, *320*, 321-332.

## 8 Anhang

**Tabelle 8-1:** Für die Berechnung von cyclo(-Gum1-Lys2-Gum3-D-Phe4-) **4** in  $H_2O$  verwendete restraints und deren Verletzungen während der 150 ps Dynamiksimulation,  $\langle r^{-3} \rangle^{-1/3}$ -gemittelt über die Trajektorie.

Atom 1	Atom 2	Untergrenze	Obergrenze	Berechneter	Abstands-
				Abstand	verletzung
		[Å]	[Å]	[Å]	[Å]
Gum <sup>1</sup> H <sup>N</sup>	Gum <sup>3</sup> H <sup>5</sup>	4.50	6.00	4.64	0.00
Gum <sup>1</sup> H <sup>N</sup>	D-Phe <sup>4</sup> H <sup>α</sup>	2.86	3.87	3.61	0.00
Gum <sup>1</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>N</sup>	4.50	6.00	5.10	0.00
Lys <sup>2</sup> H <sup>N</sup>	Gum <sup>1</sup> H <sup>5</sup>	2.84	3.84	3.50	0.00
Lys <sup>2</sup> H <sup>N</sup>	Gum <sup>1</sup> H <sup>C, pro S</sup>	3.48	4.70	4.13	0.00
Lys <sup>2</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>β, pro S</sup>	3.17	4.28	3.78	0.00
Lys <sup>2</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>β, pro R</sup>	2.66	3.60	2.64	-0.02
Lys <sup>2</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>γ*</sup>	2.54	5.24	3.29	0.00
Lys <sup>2</sup> H <sup>β, pro S</sup>	Lys <sup>2</sup> H <sup>ε*</sup>	3.06	5.95	4.18	0.00
Gum <sup>3</sup> H <sup>N</sup>	D-Phe <sup>4</sup> H <sup>N</sup>	4.50	6.00	4.96	0.00
Gum <sup>3</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>N</sup>	2.72	3.68	3.07	0.00
Gum <sup>3</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>α</sup>	2.80	3.79	2.98	0.00
Gum <sup>3</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>β, pro S</sup>	3.61	4.89	4.21	0.00
Gum <sup>3</sup> H <sup>N</sup>	Lys <sup>2</sup> H <sup>β, pro R</sup>	3.54	4.79	3.81	0.00
D-Phe <sup>4</sup> H <sup>N</sup>	Gum <sup>3</sup> H <sup>4</sup>	3.30	4.46	4.32	0.00
D-Phe <sup>4</sup> H <sup>N</sup>	Gum <sup>3</sup> H <sup>5</sup>	2.37	3.21	2.24	-0.13
D-Phe <sup>4</sup> H <sup>N</sup>	D-Phe <sup>4</sup> H <sup>α</sup>	2.60	3.52	2.98	0.00
D-Phe <sup>4</sup> H <sup>N</sup>	Gum <sup>1</sup> H <sup>N</sup>	2.64	3.57	2.86	0.00

Für die mit „\*“ gekennzeichneten Protonen wurde in der Rechnung ein Pseudoatom eingesetzt und die Obergrenze um die entsprechende Standard-Pseudoatomkorrektur erhöht.<sup>[103]</sup>

**Tabelle 8-2:** Für die Berechnung von Fmoc-[f-SAA- $\beta$ -hGly]3-OH 5 in CD<sub>3</sub>CN verwendete restraints und deren Verletzungen während der 150 ps Dynamiksimulation,  $\langle r^{-3} \rangle^{-1/3}$ -gemittelt über die Trajektorie.

Atom 1		Atom 2		Unter- grenze [Å]	Ober- grenze [Å]	Abstand (berechnet) [Å]	Abstands- verletzung [Å]
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>2</sup>	3.51	4.29	3.47	-0.05
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>3</sup>	2.70	3.30	3.04	0.00
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>C</sup>	2.70	3.30	2.72	0.00
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>O2*</sup>	3.63	6.10	4.31	0.00
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>O3*</sup>	3.63	6.10	5.12	0.00
f-SAA <sup>1</sup>	H <sup>1</sup>	f-SAA <sup>1</sup>	H <sup>O2*</sup>	3.63	6.10	4.34	0.00
f-SAA <sup>1</sup>	H <sup>1</sup>	f-SAA <sup>1</sup>	H <sup>O3*</sup>	2.03	4.10	3.85	0.00
f-SAA <sup>1</sup>	H <sup>1</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	4.50	5.50	4.54	0.00
f-SAA <sup>1</sup>	H <sup>2</sup>	f-SAA <sup>1</sup>	H <sup>3</sup>	2.16	2.64	2.35	0.00
f-SAA <sup>1</sup>	H <sup>2</sup>	f-SAA <sup>1</sup>	H <sup>O2*</sup>	3.73	6.20	4.30	0.00
f-SAA <sup>1</sup>	H <sup>2</sup>	f-SAA <sup>1</sup>	H <sup>O3*</sup>	2.13	4.20	3.46	0.00
f-SAA <sup>1</sup>	H <sup>3</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	2.70	3.30	2.65	-0.05
f-SAA <sup>1</sup>	H <sup>3</sup>	f-SAA <sup>3</sup>	H <sup>O2*</sup>	2.83	5.10	4.89	0.00
f-SAA <sup>1</sup>	H <sup>C</sup>	f-SAA <sup>1</sup>	H <sup>O2*</sup>	2.13	4.20	3.42	0.00
$\beta$ -hGly <sup>2</sup>	H <sup>N</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>2</sup>	2.90	3.50	3.46	0.00
$\beta$ -hGly <sup>2</sup>	H <sup>N</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\beta</math>1</sup>	2.70	3.30	3.03	0.00
$\beta$ -hGly <sup>2</sup>	H <sup>N</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\beta</math>2</sup>	2.25	2.75	2.47	0.00
$\beta$ -hGly <sup>2</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>C</sup>	3.20	4.00	3.50	0.00
$\beta$ -hGly <sup>2</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	4.00	5.00	4.10	0.00
$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>1</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\beta</math>1</sup>	2.25	2.75	2.44	0.00
$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>1</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\beta</math>2</sup>	2.25	2.75	2.51	0.00
$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>1</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	3.63	6.10	4.03	0.00
$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>2</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\beta</math>1</sup>	2.25	2.75	2.52	0.00
$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>2</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\beta</math>2</sup>	2.70	3.30	3.09	0.00
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>3</sup>	2.70	3.30	3.14	0.00
f-SAA <sup>3</sup>	H <sup>N</sup>	$\beta$ -hGly <sup>2</sup>	H <sup><math>\alpha</math>2</sup>	2.25	2.75	2.16	-0.09

Tabelle 8-2: Teil II

Atom 1		Atom 2		Unter- grenze [Å]	Ober- grenze [Å]	Abstand (berechnet) [Å]	Abstands- verletzung [Å]
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>2</sup>	3.20	4.00	3.47	0.00
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	2.70	3.30	3.07	0.00
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	2.70	3.30	2.68	-0.02
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>O2*</sup>	2.63	5.10	4.22	0.00
f-SAA <sup>3</sup>	H <sup>1</sup>	f-SAA <sup>3</sup>	H <sup>2</sup>	2.43	2.97	2.36	-0.07
f-SAA <sup>3</sup>	H <sup>1</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	3.06	3.74	3.05	-0.01
f-SAA <sup>3</sup>	H <sup>1</sup>	f-SAA <sup>3</sup>	H <sup>O3*</sup>	2.23	4.30	3.85	0.00
f-SAA <sup>3</sup>	H <sup>2</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	2.43	2.97	2.35	-0.08
f-SAA <sup>3</sup>	H <sup>2</sup>	f-SAA <sup>3</sup>	H <sup>O2*</sup>	2.83	5.10	4.31	0.00
f-SAA <sup>3</sup>	H <sup>2</sup>	f-SAA <sup>3</sup>	H <sup>O3*</sup>	1.78	3.80	3.49	0.00
f-SAA <sup>3</sup>	H <sup>2</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	2.73	5.00	4.87	0.00
f-SAA <sup>3</sup>	H <sup>2</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	3.20	4.00	3.36	0.00
f-SAA <sup>3</sup>	H <sup>3</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	2.79	3.41	3.06	0.00
f-SAA <sup>3</sup>	H <sup>3</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	2.70	3.30	2.60	-0.10
f-SAA <sup>3</sup>	H <sup>3</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	3.48	5.85	4.61	0.00
f-SAA <sup>3</sup>	H <sup>C</sup>	f-SAA <sup>3</sup>	H <sup>O2*</sup>	1.93	4.00	3.35	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>1</sup>	3.24	3.96	3.49	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	2.52	3.08	2.59	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	3.15	3.85	3.51	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	H <sup>α2</sup>	3.25	3.95	3.39	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	H <sup>β1</sup>	2.70	3.30	3.04	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	H <sup>β2</sup>	2.25	2.75	2.48	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>N</sup>	2.70	3.30	3.29	0.00
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	4.00	5.00	4.02	0.00
β-hGly <sup>4</sup>	H <sup>α1</sup>	β-hGly <sup>4</sup>	H <sup>β1</sup>	2.25	2.75	2.42	0.00
β-hGly <sup>4</sup>	H <sup>α1</sup>	β-hGly <sup>4</sup>	H <sup>β2</sup>	2.25	2.75	2.53	0.00
β-hGly <sup>4</sup>	H <sup>α2</sup>	β-hGly <sup>4</sup>	H <sup>β1</sup>	2.25	2.75	2.55	0.00
β-hGly <sup>4</sup>	H <sup>α2</sup>	β-hGly <sup>4</sup>	H <sup>β2</sup>	2.70	3.30	3.10	0.00
β-hGly <sup>4</sup>	H <sup>α2</sup>	f-SAA <sup>5</sup>	H <sup>3</sup>	4.05	4.95	4.59	0.00

Tabelle 8-2: Teil III

Atom 1		Atom 2		Unter- grenze [Å]	Ober- grenze [Å]	Abstand (berechnet) [Å]	Abstands- verletzung [Å]
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	2.25	2.75	2.75	0.00
f-SAA <sup>5</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	H <sup>α1</sup>	3.15	3.85	3.34	0.00
f-SAA <sup>5</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	H <sup>α2</sup>	2.25	2.75	2.19	-0.06
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>2</sup>	3.20	4.00	3.56	0.00
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>3</sup>	2.70	3.30	3.07	0.00
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	2.61	3.19	2.64	0.00
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	3.63	6.10	4.17	0.00
f-SAA <sup>5</sup>	H <sup>1</sup>	f-SAA <sup>5</sup>	H <sup>3</sup>	2.88	3.52	3.10	0.00
f-SAA <sup>5</sup>	H <sup>1</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	3.51	4.29	3.80	0.00
f-SAA <sup>5</sup>	H <sup>1</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	3.63	6.10	4.35	0.00
f-SAA <sup>5</sup>	H <sup>1</sup>	f-SAA <sup>5</sup>	H <sup>O3*</sup>	1.93	4.00	3.80	0.00
f-SAA <sup>5</sup>	H <sup>2</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	3.42	4.18	3.82	0.00
f-SAA <sup>5</sup>	H <sup>2</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	3.63	6.10	4.32	0.00
f-SAA <sup>5</sup>	H <sup>2</sup>	f-SAA <sup>5</sup>	H <sup>O3*</sup>	2.13	4.20	3.52	0.00
f-SAA <sup>5</sup>	H <sup>C</sup>	f-SAA <sup>5</sup>	H <sup>O2*</sup>	1.93	4.00	3.25	0.00
f-SAA <sup>5</sup>	H <sup>C</sup>	f-SAA <sup>5</sup>	H <sup>O3*</sup>	2.93	5.20	4.82	0.00
β-hGly <sup>6</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>1</sup>	3.15	3.85	3.72	0.00
β-hGly <sup>6</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>3</sup>	3.20	4.00	2.89	-0.31
β-hGly <sup>6</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	3.20	4.00	3.44	0.00
β-hGly <sup>6</sup>	H <sup>N</sup>	β-hGly <sup>6</sup>	H <sup>α*</sup>	2.83	5.00	2.90	0.00
β-hGly <sup>6</sup>	H <sup>N</sup>	β-hGly <sup>6</sup>	H <sup>β*</sup>	2.13	4.10	2.21	0.00

Die mit „\*“ gekennzeichneten Protonen wurden in der Rechnung als Pseudoatom behandelt und die Obergrenze um die entsprechende Standard-Pseudoatomkorrektur erhöht.<sup>[103]</sup>

**Tabelle 8-3:** Distanzen, mit denen sich eine 10/12/10-, 10/10/10-, 12/12/12- bzw. 14/14/14-helicale (Haupt-)Konformation für Fmoc-[f-SAA-β-hGly]<sub>3</sub>-OH **5** in CD<sub>3</sub>CN ausschließen lässt. Pseudoatome im Falle der Methyl- bzw. Methylengruppen sind mit dem Buchstaben Q, und nicht beobachtete Signale mit „n. b.“ bezeichnet.

Atom 1		Atom 2		Abstand in jeweiliger Helix [Å]	Abstand in ermittelter Struktur [Å]	Gemessener Abstand [Å]
<b><u>10/12/10</u></b>						
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	Q <sup>QO</sup>	2.3	4.5	> 3.7*
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	3.5	4.5	n. b.
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	3.4	4.9	n. b.
β-hGly <sup>2</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>C</sup>	2.2	3.5	3.6
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	2.2	3.5	> 3.0*
β-hGly <sup>6</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	2.2	3.5	3.6
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>3</sup>	2.3	3.0	3.0
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	2.3	3.1	3.0
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>3</sup>	2.3	3.1	3.0
β-hGly <sup>6</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	Q <sup>β</sup>	3.4	> 5	n. b.
β-hGly <sup>6</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	3.2	> 5	n. b.
<b><u>10/10/10</u></b>						
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	Q <sup>QO</sup>	2.6	4.5	> 3.7*
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>3</sup>	2.3	3.0	3.0
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	2.3	3.1	3.0
f-SAA <sup>5</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>3</sup>	2.3	3.1	3.0
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	3.4	4.5	n. b.
f-SAA <sup>3</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	3.1	4.9	n. b.
<b><u>12/12/12</u></b>						
β-hGly <sup>2</sup>	H <sup>N</sup>	f-SAA <sup>1</sup>	H <sup>C</sup>	2.4	3.5	3.6
β-hGly <sup>4</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>C</sup>	2.2	3.5	> 3.0*
β-hGly <sup>6</sup>	H <sup>N</sup>	f-SAA <sup>5</sup>	H <sup>C</sup>	2.2	3.5	3.6
β-hGly <sup>6</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	Q <sup>β</sup>	3.5	> 5	n. b.



Tabelle 8-3: Teil II

Atom 1		Atom 2		Abstand in jeweiliger Helix [Å]	Abstand in ermittelter Struktur [Å]	Gemessener Abstand [Å]
<b><u>14/14/14</u></b>						
f-SAA <sup>1</sup>	H <sup>N</sup>	f-SAA <sup>3</sup>	H <sup>3</sup>	3.6	> 5	n. b.
f-SAA <sup>1</sup>	H <sup>N</sup>	β-hGly <sup>4</sup>	Q <sup>β</sup>	3.1	> 5	n. b.
β-hGly <sup>4</sup>	Q <sup>α</sup>	f-SAA <sup>1</sup>	QQO	2.4	> 5	n. b.
β-hGly <sup>4</sup>	Q <sup>β</sup>	f-SAA <sup>1</sup>	QQO	2.4	> 5	n. b.

(\*) Überlagertes Signal. Der tatsächliche Wert ist größer als der angegebene.