
Konformationsdynamik lichtschtbarer Peptide: Molekulardynamiksimulationen und datengetriebene Modellbildung

Heiko Carstens



München 2004

Konformationsdynamik lichtschtbarer Peptide: Molekulardynamiksimulationen und datengetriebene Modellbildung

Heiko Carstens

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Heiko Carstens
aus München

München, den 19.04.2004

Erstgutachter: Prof. Paul Tavan
Zweitgutachter: Prof. Wolfgang Zinth
Tag der mündlichen Prüfung: 16.06.2004

für Nadia

Zusammenfassung

Die Faltung und die Funktionsdynamik von Proteinen basieren auf schnellen Prozessen, die zum Teil im Zeitbereich der Pikosekunden bis Nanosekunden ablaufen. Zur Untersuchung dieser Dynamiken und der mit ihnen verbundenen strukturellen Änderungen werden häufig Molekulardynamik (MD)-Simulationen eingesetzt, die auf empirisch parametrisierten molekularmechanischen (MM) Kraftfeldern basieren.

Die vorliegende Arbeit stellt einen Ansatz zur Validierung solcher MM-Kraftfelder vor, der darin besteht, die Relaxationsdynamik kleiner lichtschtbarer Modellpeptide zu simulieren und die dabei auftretenden Kinetiken mit Ergebnissen der Femtosekunden-Spektroskopie zu vergleichen. Erste Simulationen dieser Art zeigen eine überraschende Übereinstimmung zwischen den simulierten und den gemessenen Kinetiken [S. Spörlein, H. Carstens, H. Satzger, C. Renner, R. Behrendt, L. Moroder, P. Tavan, W. Zinth, J. Wachtveitl, *PNAS* **99**, 7998-8002 (2003)]. Weitere Untersuchungen, bei denen einzelne Details des eingesetzten Kraftfelds variiert werden, lassen jedoch erkennen, dass diese Übereinstimmung auf einer zufälligen Kompensation von Fehlern beruht. Es wird gezeigt, dass die simulierten Kinetiken sehr empfindlich auf Änderungen am MM-Kraftfeld reagieren und damit als Maßstab für die Güte seiner Parametrisierung dienen können. Besonders die Modellierung des Lösungsmittels DMSO hat einen entscheidenden Einfluss auf die beobachteten Kinetiken, und zwar nicht nur auf die Kühlzeiten der Wärmedissipation, sondern auch auf die Relaxationsdynamik des Peptidteils der Modellsysteme.

Als Vorarbeit für die Simulation der Modellpeptide wird ein flexibles und explizites DMSO-Modell aus ersten Prinzipien abgeleitet und dessen thermodynamische und strukturelle Eigenschaften mit denen existierender Modelle verglichen. Ferner wird das eingesetzte Kraftfeld um Parameter für den in die Modellpeptide integrierten Farbstoff Azobenzol erweitert und dessen lichtinduzierte Isomerisierungsreaktion modelliert. Darüber hinaus werden neuartige Methoden zur statistischen Auswertung von MD-Trajektorien vorgestellt, die dazu dienen, eine strukturelle Klassifikation der Peptidgeometrien zu ermöglichen. Mit Hilfe dieser Klassifikation kann ein vertiefter Einblick in die während der Relaxation der Modellpeptide auftretenden Konformationsübergänge gewonnen werden. Ferner ermöglichen es die statistischen Auswertungsverfahren, aus Langzeitsimulationen der Modellpeptide deren Gleichgewichtskonformationen zu bestimmen. Der Vergleich dieser Konformationen mit Daten der NMR-Spektroskopie zeigt schließlich die Leistungsfähigkeit der Methode der MD-Simulation für die Vorhersage von Peptidstrukturen.

Inhaltsverzeichnis

| | |
|---|-------------|
| Abbildungsverzeichnis | xi |
| Tabellenverzeichnis | xv |
| Abkürzungsverzeichnis | xvii |
| 1 Einführung | 1 |
| 1.1 Struktur und Dynamik von Peptiden und Proteinen | 2 |
| 1.2 Experimentelle Methoden zur Untersuchung von Biomolekülen | 10 |
| 1.3 Theoretische Beschreibungen | 15 |
| 1.4 Lichtschaltbare Modellpeptide als kombinierter Ansatz | 23 |
| 1.5 Ziele und Gliederung der Arbeit | 28 |
| 2 Standardkraftfelder und Simulationsmethoden | 31 |
| 2.1 Grundlagen der Molekulardynamik | 31 |
| 2.2 Das eingesetzte MM-Kraftfeld | 34 |
| 2.3 Die strukturadaptierte Multipolmethode | 38 |
| 3 DMSO als explizites Lösungsmittel | 43 |
| 3.1 Parametrisierung des neuen DMSO-Modells | 44 |
| 3.1.1 DFT-Berechnungen im Vakuum | 45 |
| 3.1.2 Hybride DFT/MM-Simulationen von DMSO in DMSO | 52 |
| 3.1.3 Optimierung der Van der Waals Parameter | 59 |
| 3.2 Evaluation der DMSO-Modelle | 61 |
| 3.2.1 Thermodynamische und kinetische Observablen | 61 |
| 3.2.2 Radiale Paarverteilungsfunktionen | 68 |
| 3.2.3 Diskussion | 71 |
| 4 Modellierung der lichtschaltbaren Peptide | 73 |
| 4.1 Kraftfeldparameter für die Azobenzolfarbstoffe | 74 |
| 4.2 Ein MM-Modell der Isomerisierung | 82 |
| 5 Datenanalyse und Modellbildung | 89 |
| 5.1 Dimensionsreduktion | 92 |
| 5.2 Dichteschätzung | 97 |
| 5.3 Unüberwachte Klassifikation und Prototypen | 106 |
| 5.4 Zirkuläre Statistik | 111 |

| | | |
|----------|---|------------|
| 6 | Validierung des Peptid-Kraffteldes: das Alanin-Dipeptid | 117 |
| 6.1 | Minimalstrukturen im Vakuum | 118 |
| 6.2 | Energieflächen im Vakuum und in Lösung | 121 |
| 7 | Die Modellpeptide im Gleichgewicht | 129 |
| 7.1 | Details der Simulationen | 130 |
| 7.2 | Die Energielandschaften der <i>cis</i> - und <i>trans</i> -Isomere | 135 |
| 7.2.1 | Ramachandran-Plots | 135 |
| 7.2.2 | Projektion der Energieflächen auf ihre Hauptkomponenten | 139 |
| 7.3 | Konformationsanalyse durch Klassifikation der MD-Trajektorien | 143 |
| 7.3.1 | Die Anzahlen der bestimmten Konformationen | 144 |
| 7.3.2 | Repräsentative Strukturen | 146 |
| 7.4 | Erfüllung der NMR-Constraints | 154 |
| 7.4.1 | Die beispielhafte Diskussion einiger NOEs | 155 |
| 7.4.2 | Die Erfüllung der NOEs durch die Konformationen des <i>cis</i> -cAPBs | 159 |
| 7.4.3 | Die Protonenabstände der anderen Modellpeptide | 162 |
| 7.4.4 | Vergleich der ³ J-Kopplungskonstanten | 164 |
| 7.5 | Diskussion | 167 |
| 8 | Die lichtinduzierte Relaxationsdynamik der Modellpeptide | 169 |
| 8.1 | Details der Simulationen | 172 |
| 8.2 | Energetische und geometrische Observablen | 175 |
| 8.3 | Ergebnisse und Diskussion | 179 |
| 8.3.1 | Übereinstimmung mit den Experimenten: Azobenzol und cAPB | 179 |
| 8.3.2 | Der Einfluss des DMSO-Modells auf die Kühlzeit des Azobenzols | 184 |
| 8.3.3 | Die Reproduzierbarkeit der Simulationen des cAPBs | 187 |
| 8.3.4 | Veränderungen an λ_{14} und den Dihedralpotentialen | 189 |
| 8.3.5 | Das cAPB in unterschiedlichen DMSO-Modellen | 196 |
| 8.3.6 | Ein Einblick in die lichtinduzierten Konformationsübergänge | 199 |
| 9 | Zusammenfassung und Ausblick | 209 |
| A | Optimierung von VDW-Parametern mit der Simplex-Methode | 213 |
| B | Korrektur der langreichweitigen VDW-Beiträge | 215 |
| C | Geometrien der Konformationen | 219 |
| D | Tabellen des MD-NMR Vergleichs | 221 |
| | Literaturverzeichnis | 229 |
| | Danksagung | 251 |

Abbildungsverzeichnis

| | | |
|------|---|----|
| 1.1 | Ausschnitt aus einer Peptidkette | 2 |
| 1.2 | Die Peptidbindung und Definition der Torsionswinkel | 3 |
| 1.3 | Seitengruppen von Aminosäuren | 4 |
| 1.4 | Ramachandran-Plots | 5 |
| 1.5 | α -Helix und β -Faltblatt als Sekundärstrukturen | 5 |
| 1.6 | Die 3D-Struktur der Thioredoxin Reduktase | 7 |
| 1.7 | Energielandschaften der Proteinfaltung | 9 |
| 1.8 | Das Ergebnis einer Röntgenstrukturanalyse | 11 |
| 1.9 | Eine klassische Illustration des NMR-Prinzips | 13 |
| 1.10 | Raumfüllende Simulationsvolumina | 20 |
| 1.11 | Behandlung der Elektrostatik unter periodischen Randbedingungen | 21 |
| 1.12 | Die Geometrieänderung von Azobenzol durch Photoisomerisierung | 24 |
| 1.13 | Die chemische Struktur von Azobenzol und den Modellpeptiden | 25 |
| 1.14 | Je eine Konformation des <i>cis</i> - und <i>trans</i> -cAPBs | 27 |
| 2.1 | Skizze der internen Koordinaten | 35 |
| 2.2 | SAMM, die strukturadaptierte Multipolmethode | 38 |
| 2.3 | Der Überlapp der Hierarchieebenen in FAMUSAMM | 40 |
| 3.1 | Dimethylsulfoxid | 43 |
| 3.2 | Die IR-aktiven Normalmoden des DMSO | 48 |
| 3.3 | Die Wahl der Gitterpunkte zur Bestimmung der ESP-Ladungen des DMSOs | 50 |
| 3.4 | Die Abhängigkeit der ESP-Ladungen und des Dipolmoments von $r_{\text{cov}}(\text{H})$ | 51 |
| 3.5 | Hybride Simulation von DMSO in DMSO | 53 |
| 3.6 | Das molekülfeste Koordinatensystem für DMSO | 54 |
| 3.7 | Das Dipolmoment des DMSO-Moleküls in der 5-ps Hybridsimulation | 55 |
| 3.8 | Die Verteilung der internen Koordinaten während der Hybridtrajektorie | 58 |
| 3.9 | Die Paarverteilungsfunktionen der schweren Atome des DMSO | 69 |
| 3.10 | Partielle Strukturfaktoren aus der Simulation und der Neutronenstreuung | 71 |
| 4.1 | Die Struktur und die Atomnamen der Azobenzolfarbstoffe | 75 |
| 4.2 | Eine farbkodierte Darstellung der MM-Partialladungen | 77 |
| 4.3 | Die Potentialflächen der Bindungen und Winkel des Azobenzols | 79 |
| 4.4 | Die Potentialkurven der Dihedralwinkel | 83 |
| 4.5 | Die diskutierten Reaktionspfade der Isomerisierung des Azobenzols | 84 |
| 4.6 | Das Zusatzpotential der erzwungenen Inversion | 85 |

| | | |
|------|---|-----|
| 4.7 | Das lokale Koordinatensystem der Isomerisierung | 86 |
| 5.1 | Beispiel einer MD-Trajektorie | 90 |
| 5.2 | Hauptkomponentenanalyse | 94 |
| 5.3 | Die Histogrammschätzung | 98 |
| 5.4 | Die Partitionierung des Merkmalsraumes | 100 |
| 5.5 | Der UNIVAR-Algorithmus | 103 |
| 5.6 | Der Verlauf eines UNIVAR-Trainings | 104 |
| 5.7 | Die Entwicklung der Orte der Kodebuchzentren | 104 |
| 5.8 | Das Ergebnis einer UNIVAR Dichteschätzung | 105 |
| 5.9 | Die skalierte Modelldichte und ihre Maxima | 107 |
| 5.10 | Das Ergebnis des ASSO-Verfahrens | 109 |
| 5.11 | Die Hierarchie der Prototypen als Baumstruktur | 109 |
| 5.12 | Der Mittelwert zweier Winkelvariablen | 112 |
| 5.13 | Die Zirkuläre Normalverteilung und ihre Darstellungen | 113 |
| 5.14 | Fehler in den Darstellungen der zirkulären Normalverteilung | 114 |
| 5.15 | Dichteschätzung im Fall zirkulärer Variablen | 115 |
| 6.1 | Das Alanin-Dipeptid | 118 |
| 6.2 | Drei stabile Strukturen des Alanin-Dipeptids im Vakuum | 119 |
| 6.3 | Zeitreihen der Dihedralwinkel des Alanin-Dipeptids in Lösung | 123 |
| 6.4 | Die Energieflächen des Alanin-Dipeptids | 125 |
| 7.1 | Relaxation einer NMR-Struktur während der MD-Simulation | 134 |
| 7.2 | Ramachandran-Plots der freien Energie der Modellpeptide | 136 |
| 7.3 | Die Motivation der Anwendung der Hauptkomponentenanalyse | 139 |
| 7.4 | Die mit den Hauptkomponenten verbundenen Varianzen | 141 |
| 7.5 | Die freie Energie in Projektion auf die ersten beiden Hauptkomponenten | 142 |
| 7.6 | Die Anzahlen der Konformationen | 144 |
| 7.7 | Die beobachteten Konformationen der Modellpeptide | 148 |
| 7.8 | Die NMR-Strukturen | 149 |
| 7.9 | Die Konformationen des cAPBs | 152 |
| 7.10 | Einige NOEs des <i>cis</i> -cAPBs | 156 |
| 7.11 | Detaillierte Ergebnisse der Simulation zu einigen der NOEs | 157 |
| 7.12 | Abweichung der <i>cis</i> -cAPB Konformationen von den NOEs | 160 |
| 7.13 | Das Ansteigen der mittleren NOE-Verletzungen mit ΔG | 162 |
| 7.14 | 3J -Kopplungskonstanten und die Dihedralwinkel | 165 |
| 8.1 | Energierelaxation: Azobenzol (R1) und cAPB (R5) | 180 |
| 8.2 | Energierelaxation: Azobenzol in verschiedenen DMSO-Modellen (R2-R4) | 186 |
| 8.3 | Energierelaxation: alte (R5) und neue (R6) Simulation des cAPBs | 188 |
| 8.4 | Energierelaxation: Veränderung der Dihedralpotentiale (R6 und R8) | 190 |
| 8.5 | Energierelaxation: Änderung des Skalierungsfaktors λ_{14} (R6 und R7) | 191 |
| 8.6 | Klassifikation der Peptidgeometrien: cAPB, R6 bis R9 | 193 |

| | | |
|------|--|-----|
| 8.7 | Energierelaxation: Veränderung der Dihedralpotentiale II (R7 und R9) . . . | 194 |
| 8.8 | Energierelaxation: cAPB in diversen DMSO-Modellen (R9, R11 und R12) | 197 |
| 8.9 | Klassifikation der Peptidgeometrien: cAPB, R9, R11 und R12 | 198 |
| 8.10 | Geometrische Klassifikation einiger Einzeltrajektorien | 200 |
| 8.11 | Hauptkomponentenprojektion einiger Relaxations-Trajektorien | 201 |
| 8.12 | Schnappschüsse der Molekülstrukturen | 204 |
| 8.13 | Die Einzeltrajektorien der Dihedralwinkel | 205 |
| A.1 | Der Simplex-Optimierungsalgorithmus in zwei Dimensionen | 213 |
| B.1 | Die optimale Korrektur der VDW-Wechselwirkung | 216 |

Tabellenverzeichnis

| | | |
|-----|---|-----|
| 1.1 | Zeitskalen der Dynamik von Proteinen | 8 |
| 1.2 | Die Anwendungsbereiche verschiedener theoretischer Methoden | 16 |
| 3.1 | Mit DFT optimierte Geometrien und Dipolmomente von DMSO <i>in vacuo</i> | 46 |
| 3.2 | Wellenzahlen der IR-aktiven Normalmoden des DMSOs | 48 |
| 3.3 | Zunahme des Dipolmoments beim Übergang in die kondensierte Phase | 57 |
| 3.4 | Dipolmomente der DMSO-Modelle | 57 |
| 3.5 | Die Geometrie des DMSO-Moleküls in der Hybridtrajektorie | 59 |
| 3.6 | Bindungsparameter der DMSO-Modelle | 62 |
| 3.7 | Partiellladungen und VDW-Parameter der DMSO-Modelle | 63 |
| 3.8 | Thermodynamische und dynamische Observablen der DMSO-Modelle | 65 |
| 3.9 | Experimentelle Gewichtungsfaktoren der Paarverteilungsfunktionen | 70 |
| 4.1 | Kraftfeldparameter der Bindungslängen und -winkel | 76 |
| 4.2 | Atomtypen und Partiellladungen der Azobenzolfarbstoffe | 78 |
| 4.3 | Kraftfeldparameter der Torsionswinkel | 80 |
| 6.1 | Die Energien und Minimalstrukturen des Alanin-Dipeptids im Vakuum | 119 |
| 7.1 | Mittlere RMSDs der Peptid-Konformationen | 147 |
| 7.2 | Mittlere RMSDs der cAPB-Konformationsensembles | 151 |
| 8.1 | Übersicht über die Relaxations-Simulationen | 172 |
| 8.2 | Zeitkonstanten: Azobenzol (R1) und cAPB (R5) | 180 |
| 8.3 | Zeitkonstanten: Azobenzol in verschiedenen DMSO-Modellen (R2-R4) | 186 |
| 8.4 | Zeitkonstanten: alte (R5) und neue (R6) Simulation des cAPBs | 188 |
| 8.5 | Zeitkonstanten: cAPB, R6 bis R9 | 190 |
| 8.6 | Zeitkonstanten: cAPB, R9, R11 und R12 | 197 |
| B.1 | Effektive Gyrationradien zur VDW-Korrektur | 217 |
| C.1 | Die Geometrien der Peptidkonformationen | 220 |
| D.1 | <i>cis</i> -cAPB: Protonenabstände und ^3J -Kopplungen | 222 |
| D.2 | <i>trans</i> -cAPB: Protonenabstände und ^3J -Kopplungen | 223 |
| D.3 | <i>cis</i> -cAMPB: Protonenabstände und ^3J -Kopplungen | 224 |
| D.4 | <i>trans</i> -cAMPB: Protonenabstände und ^3J -Kopplungen | 225 |

Tabellenverzeichnis

| | | |
|-----|---|-----|
| D.5 | <i>cis</i> -bcAMPB: Protonenabstände und ^3J -Kopplungen | 226 |
| D.6 | <i>trans</i> -bcAMPB: Protonenabstände und ^3J -Kopplungen | 227 |

Abkürzungsverzeichnis

| | |
|-------------|---|
| ϕ/ψ | Torsionswinkel des Peptidrückgrats, Seite 3 |
| AMPB | (4-Aminomethylphenyl)azobenzoesäure, Seite 25 |
| APB | (4-Aminophenyl)azobenzoesäure, Seite 25 |
| bcAPB | bizyklisches Peptid mit AMPB als Schalter, Seite 25 |
| cAMPB | monozyklisches Peptid mit AMPB als Schalter, Seite 25 |
| cAPB | monozyklisches Peptid mit APB als Schalter, Seite 25 |
| CPMD | Car-Parinello Molecular Dynamics, Seite 17 |
| DFT | Dichtefunktionaltheorie, Seite 16 |
| DK | Dielektrizitätskonstante, Seite 123 |
| DMSO | Dimethylsulfoxid, C_2H_6SO , Seite 28 |
| HF | Hartree-Fock Theorie oder Verfahren, Seite 16 |
| IR | Infrarot, Seite 14 |
| MD | Molekulardynamik, Seite 18 |
| MIC | engl. Minimum Image Convention, Seite 20 |
| MM | Molekularmechanik, ein effektives Molekülmodell, Seite 17 |
| NMR | engl. Nuclear Magnetic Resonance: Kernspinresonanz, Seite 12 |
| NOE | Nuklearer Overhauser Effekt, Seite 12 |
| PBC | engl. Periodic Boundary Conditions, periodische Randbedingungen, Seite 19 |
| PCA | engl. Principal Component Analysis, Hauptkomponentenanalyse, Seite 94 |
| PMF | engl. Potential of Mean Force, freie Energiefunktion, Seite 90 |
| PP | Pseudopotential, Seite 17 |

Abkürzungsverzeichnis

| | |
|------|--|
| RF | Reaktionsfeld, Seite 22 |
| RMSD | engl. Root Mean Square Deviation, mittlere quadratische Abweichung der Atomkoordinaten, Seite 92 |
| VDW | Van der Waals-Wechselwirkung, Seite 19 |
| WND | engl. Wrapped Normal Distribution, eine der Normalverteilungen zirkulärer Variablen, Seite 113 |

1 Einführung

Hofstadter's Law:

*It always takes longer than you expect, even when you take into account
Hofstadter's Law.*

Douglas Hofstadter, *Gödel, Escher, Bach*, 1979 [1]

Auf molekularer Ebene sind die Proteine die wichtigsten Bauteile und Werkzeuge der Lebewesen [2]. Um die vielfältigsten Aufgaben in den Organismen wahrnehmen zu können, wurden sie durch die Evolution fortwährend weiterentwickelt und optimiert. Sie dienen als Strukturelemente der Zellen, als Botenstoffe und als Rezeptoren für solche Stoffe, ermöglichen die Signalleitung in Nervenzellen oder sind Transportmittel, wie das Hämoglobin für Sauerstoff. Eine wichtige Aufgabe von Proteinen ist die Katalyse biochemischer Reaktionen. So spalten sie beispielsweise als Enzyme große Moleküle der Nahrung in kleine, verdaubare Teile oder bauen Giftstoffe ab, die dem Organismus schaden könnten. Die in den letzten Jahrzehnten entwickelten Computersimulationen [3–8] ermöglichen einen detaillierten Einblick in die Funktionsweise von Proteinen auf atomarer Ebene.

Die vorliegende Arbeit soll zur Beantwortung der Frage beitragen, unter welchen Bedingungen solche Computersimulationen geeignet sind, eine „virtuelle Realität“ dieser Biomoleküle zu erzeugen, d.h. experimentell verifizierbare Vorhersagen für die Werte physikalischer Observablen zu treffen. Im Zentrum steht dabei die Beschreibung der Dynamik von Proteinen und Peptiden auf Zeitskalen im Bereich von Femto- bis Nanosekunden.

Statt großer und sehr komplexer Proteine sollen dabei kleine lichtsichtbare Modellpeptide in Lösung betrachtet werden. Dabei besteht die, wie es sich zeigen wird, begründete Hoffnung, aus solchen Minimalmodellen grundlegende Einsichten über den Ablauf jener photobiologischen Prozesse zu generieren, vermittels derer in Proteinen Licht in Signale oder chemische Energie umgewandelt wird.

Zur Einführung in die Physik solcher Systeme wird daher zuerst die Struktur und Dynamik von Peptiden und Proteinen vorgestellt. Es folgt eine Erläuterung zugehöriger experimenteller Methoden und theoretischer Ansätze. Darauf aufbauend werden Ideen präsentiert, die im Rahmen des SFB 533 „Lichtinduzierte Dynamik von Biopolymeren“ entwickelt wurden und zur dort durchgeführten Synthese, experimentellen Charakterisierung und theoretischen Modellierung lichtsichtbarer Modellpeptide führten. Schließlich werden

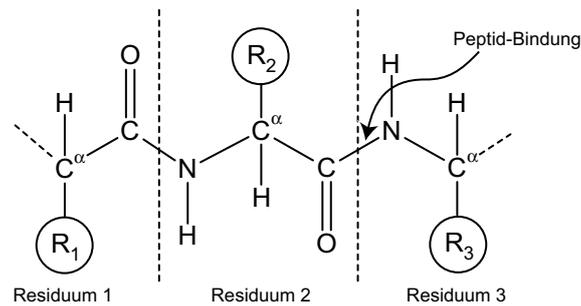


Abbildung 1.1: Ausschnitt aus einer Peptidkette. Eine Aminosäure ist mit ihren zwei Nachbarn nach Abspaltung von Wassermolekülen jeweils über eine Peptidbindung verbunden und bildet in der Peptidkette einen Aminosäure-Rest, der auch als *Residuum* bezeichnet wird. Die Seitengruppen der Aminosäuren werden als R_1 , R_2 und R_3 abgekürzt und die Fortsetzung des Polypeptids durch gestrichelte Bindungen angedeutet.

die für eine Simulation dieser Modellpeptide zu schaffenden Voraussetzungen skizziert und die in dieser Arbeit angestrebten Ziele genauer vorgestellt.

1.1 Struktur und Dynamik von Peptiden und Proteinen

Die Vielfalt von Funktionen, die Proteine als molekulare Maschinen und Werkstoffe wahrnehmen, resultiert aus der kombinatorischen Vielfalt, mit der ihre chemischen Bausteine dreidimensional angeordnet werden können. Abgesehen von Metallionen oder kleinen funktionalen Gruppen sind Proteine vor allem lineare Polymere, die in den Zellen aus 20 verschiedenen natürlichen *Aminosäuren* aufgebaut werden. Die so genannten α -Aminosäuren werden durch ein zentrales Kohlenstoffatom C_{α} charakterisiert, an das vier weitere Gruppen gebunden sind. Neben einem Wasserstoff $-H$ sind dies die Seitengruppe $-R$, durch die sich die Aminosäuren unterscheiden, eine Carboxylgruppe $-COOH$ und eine Aminogruppe $-NH_2$. Bis auf den Fall der einfachsten Aminosäure (Glyzin) ist das C_{α} asymmetrisch koordiniert, so dass jeweils zwei *Isomere*, also Moleküle gleicher Zusammensetzung, aber unterschiedlicher Struktur, existieren. In der Natur treten die Aminosäuren dennoch stets nur als L-Isomere auf, da eine der beiden Händigkeiten im Laufe der Evolution selektiert wurde.

Unter Abspaltung von Wassermolekülen und Ausbildung von *Peptidbindungen* können Aminosäuren zu *Polypeptiden* polymerisieren. Abbildung 1.1 zeigt einen Ausschnitt eines solchen Polypeptids mit drei seiner Monomere. Üblicherweise bezeichnet man mit dem Begriff *Peptid* eine kurze Kette dieses Typs, in der die Reihenfolge der Monomere, nicht unbedingt aber ihre dreidimensionale Anordnung definiert ist. Proteine hingegen sind natürlich vorkommende Polypeptide. Die Reihenfolge ihrer Aminosäuren, die so genannte

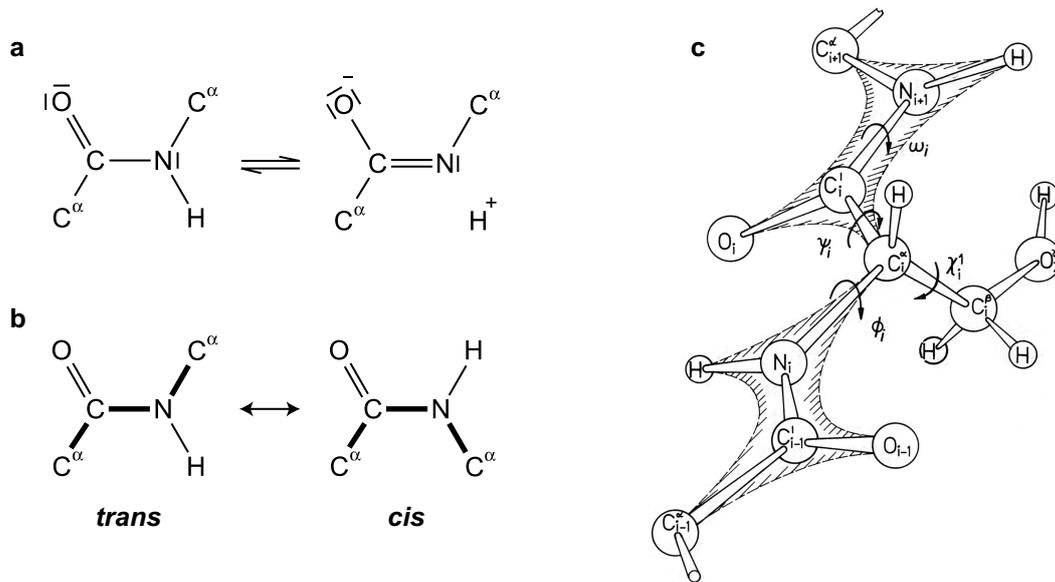


Abbildung 1.2: Die Peptidbindung und Definition der Torsionswinkel. a) Durch Resonanz der beiden dargestellten Valenzstrukturen hat die Peptidbindung einen partiellen Doppelbindungscharakter und die Peptidgruppe ein starkes Dipolmoment. b) *Trans*- und *cis*-Isomere der Peptidbindung bezüglich der C_{α} -Atome. c) Perspektivische Darstellung der Peptidplättchen und Definition der Torsionswinkel ϕ und ψ (Erläuterung im Text, Abbildung c entnommen aus [9]).

Primärstruktur, ist im genetischen Code der DNS verschlüsselt. Dabei gehört in der DNA zu jeder Aminosäure mindestens ein Triplet aus den vier Nukleinsäuren Adenin, Cytosin, Guanin und Thymin. Anders als Peptide besitzen Proteine in ihrer natürlichen Umgebung üblicherweise eine genau festgelegte Struktur, durch die ihre Funktionsweise bestimmt ist.

Wie Abbildung 1.2 a verdeutlicht, beeinflusst der durch Mesomerie begründete partielle Doppelbindungscharakter der Peptidbindung entscheidend die Struktur des Peptidrückgrats. Der rechte Grenzfall der Elektronenverteilung ist sowohl für den auftretenden Dipol als auch für die Länge der C–N Bindung verantwortlich, die mit 1,36 Å deutlich kürzer ist als eine C_{α} –N Einfachbindung mit 1,47 Å. In dieser Grenzstruktur ist die Rotation um die Peptidbindung erschwert, es entsteht ein starres *Peptidplättchen*, das nur über eine Energiebarriere von etwa 20 kcal/mol von seinem energetisch günstigen *trans*- in das ungünstigere *cis*-Isomer überführt werden kann (vgl. Abbildung 1.2 b). Durch die starren Peptidplättchen wird der Abstand zwischen zwei benachbarten C_{α} -Atomen auf etwa 3,8 Å festgelegt. Die räumliche Struktur oder *Konfiguration* der Kette von C_{α} -Atomen und damit des Peptidrückgrats kann somit vollständig durch die Angabe der *Torsions*- oder *Dihedralwinkel* ϕ (phi) und ψ (psi) beschrieben werden^a. Zum Beispiel erhält man für den

^aDie in der vorliegenden Arbeit verwendete Definition der Dihedralwinkel ϕ und ψ richtet sich nach der Vorgabe der IUPAC-IUB Commission on Biochemical Nomenclature [10], in welcher sowohl der Drehsinn (vgl. Abbildung 1.2 c) als auch der Wertebereich ($-180^{\circ} \leq \phi, \psi \leq 180^{\circ}$) der Winkel festgelegt wird.

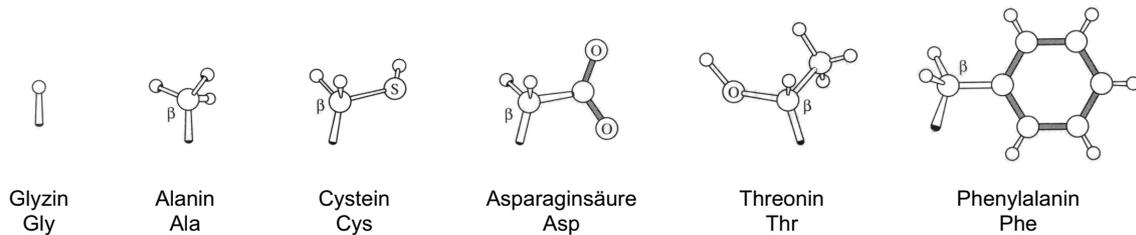


Abbildung 1.3: Seitengruppen von Aminosäuren. Die in dieser Arbeit behandelten Peptide setzen sich ausschließlich aus den sechs Aminosäuren zusammen, deren Seitengruppen hier gezeigt werden. Unter der chemischen Struktur ist jeweils der ausgeschriebene Name der Aminosäure und die dreibuchstabile Kurzform angegeben (Abbildung nach einer Vorlage aus [2]).

Dihedralwinkel, der in Abbildung 1.2 b durch die vier Atome $C_\alpha-C-N-C_\alpha$ gebildet wird, in der skizzierten *trans*-Konfiguration den Wert $\pm 180^\circ$ und in der *cis*-Konfiguration den Wert 0° .

Neben der Peptidbindung bestimmen die *Seitengruppen* der Aminosäuren die räumliche Struktur von Proteinen und Peptiden. Abbildung 1.3 zeigt die Seitengruppen derjenigen sechs Aminosäuren, aus denen sich die in dieser Arbeit untersuchten Peptide zusammensetzen. Beim einfachsten Beispiel, dem Glyzin, ist dies ein einzelnes Wasserstoffatom, welches so klein ist, dass die Peptidkette in der Umgebung sehr flexibel ist. Alle anderen Aminosäuren besitzen zumindest ein weiteres Kohlenstoffatom C_β , im Fall des Alanins in Form einer Methylgruppe. Wie auch die übrigen aliphatischen Aminosäuren ist Alanin chemisch inert und hydrophob, weshalb es häufig im Inneren von gefalteten Proteinen zu finden ist, deren Struktur es als Baustein stabilisiert. Reaktiver sind Cysteine, die über ihr Schwefelatom mit anderen Cysteinen Disulfid-Brücken formen und auf diese Weise kovalente Bindungen zwischen Peptidketten ausbilden können. Die Asparaginsäure liegt unter physiologischen Bedingungen deprotoniert vor, trägt also eine negative Ladung, wogegen die Hydroxyl-Gruppe des Threonins nur sehr schwach reaktiv ist, aber Wasserstoffbrücken ausbilden kann. Dagegen ist das aromatische Phenylalanin hydrophober.

Nach Abbildung 1.2 c können die Dihedralwinkel ϕ und ψ nicht alle möglichen Wertekombinationen einnehmen. Schon die Sperrigkeit der Peptidplättchen führt zu *sterischen Hinderungen*, so dass Konfigurationen, die eine Überlappung von Atomen zur Folge hätten, energetisch ausgeschlossen sind. Bei allen Aminosäuren, die ein C_β -Atom aufweisen, ist, unabhängig von der sonstigen chemischen Struktur der Seitenkette, der Bereich der zugänglichen Torsionswinkel zusätzlich eingeschränkt. Modelliert man die Atome einer Aminosäure als harte Kugeln und hält man sowohl die Länge der Bindungen als auch die Bindungswinkel konstant, so kann man durch Variieren der Torsionswinkel die Winkelbereiche sterisch erlaubter Konfigurationen bestimmen. Eine Darstellung der auf diese Weise gefundenen (ϕ, ψ) -Bereiche, wie sie Abbildung 1.4 a zeigt, wird als *Ramachandran-Plot* bezeichnet [13]. Ein Vergleich zwischen den Ergebnissen dieses Modells mit Werten der Dihedralwinkel, die in Kristallstrukturen von Proteinen beobachtet wurden (Abbildung 1.4

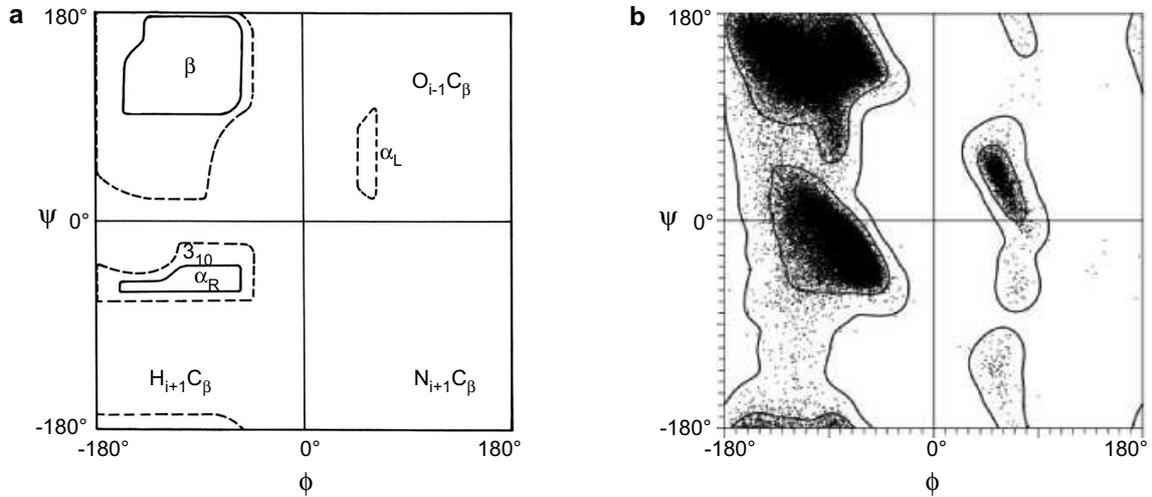


Abbildung 1.4: Ramachandran-Plots. a Erlaubte und sterisch verbotene Dihedralwinkel (ϕ , ψ) nach einem Modell, das feste Bindungslängen und -winkel im Peptidrückgrat annimmt. Durchgezogene Linien umfassen erlaubte Winkelwerte bei der Wahl größerer Radien für die beteiligten Atome, gestrichelte Linien gelten für etwas kleinere Atomradien. Die Paare von Bezeichnungen (Nomenklatur aus Abbildung 1.2 c) benennen die Atome, die sich in einem verbotenen Bereich überlappen. Ferner sind Winkelwerte angegeben, die einer links- (α_L) oder rechtshändigen (α_R) α -Helix, einer 3_{10} -Helix sowie einem β -Faltblatt entsprechen (aus [9]). b Winkelpaare von etwa 81.000 Residuen, die aus 500 hochaufgelösten Kristallstrukturen gewonnen wurden [11]. Es sind keine Glyzine oder Proline enthalten, da diese eine veränderte Winkelverteilung aufweisen.

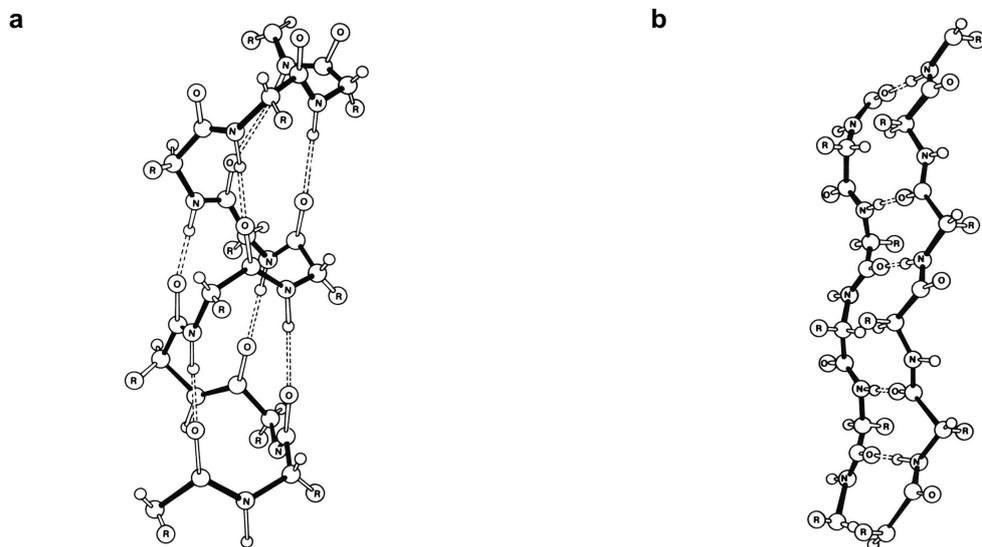


Abbildung 1.5: α -Helix und β -Faltblatt als Sekundärstrukturen. a Eine spiralförmige α -Helix, die stabilisierenden Wasserstoff-Brücken sind gestrichelt eingezeichnet. b Zwei oder mehr gestreckte Peptidketten können ein β -Faltblatt bilden (Abbildungen aus [12]).

b), weist eine gewisse Übereinstimmung auf. Allerdings sind, offensichtlich durch Auslenkung der elastischen Bindungslängen und -winkel, weitere Bereiche zugänglich. So ist vor allem die Barriere zwischen den mit α_R und β bezeichneten Gebieten nicht so ausgeprägt, wie es das starre Modell suggeriert.

Abbildung 1.4 b zeigt besonders in den beiden soeben genannten Gebieten eine hohe Dichte von Datenpunkten. Diese Regionen werden in der Legende zu Abbildung 1.4 a der rechtshändigen Helix und dem β -Faltblatt zugeordnet. Abbildung 1.5 zeigt die zugehörigen Molekülstrukturen. Eine Abfolge von Residuen, deren ϕ/ψ -Winkel in einer dieser beiden Regionen liegen, bildet eine energetisch günstige Nahordnung aus, die *Sekundärstruktur* eines Peptids oder Proteins. So wird die spiralförmige α -Helix (Abbildung 1.5 a) durch Wasserstoffbrücken zwischen einem Residuum und seinem viertnächsten Nachbarn stabilisiert. Ferner ordnen sich in dieser Struktur die Dipole der einzelnen Peptidbindungen parallel zueinander an, was zu einer weiteren Stabilisierung und zu einem Nettodipol der α -Helix führt. Ähnliche Eigenschaften, aber eine dichtere Packung, weisen die 3_{10} -Helizes auf, bei denen die Brücken zum drittnächsten Nachbarn ausgebildet werden und eine Windung genau 10 Atome umfasst. Verglichen mit den helikalen Sekundärstrukturen ist die Peptidkette in einem β -Faltblatt lang gestreckt (Abbildung 1.5 b), bildet aber zusammen mit anderen Teilsträngen eine reguläre Struktur, die wiederum durch Wasserstoffbrücken zwischen den Ketten gefestigt wird. Der Begriff „Faltblatt“ bezieht sich dabei auf die wellblechartige Faltung, die diese Sekundärstruktur in der Seitenansicht aufweist.

Abbildung 1.6 stellt die Thioredoxin Reduktase des Bakteriums *Escherichia coli* vor, ein Protein, das hier nicht nur als Anschauungsbeispiel für die Bildung von Sekundärstrukturen dienen, sondern auch deren Anordnung in der Tertiärstruktur verdeutlichen soll. Ferner wurde die Thioredoxin Reduktase als Beispiel ausgewählt, da die Aminosäuresequenz der in dieser Arbeit untersuchten Modellpeptide dem aktiven Zentrum dieses Proteins entnommen wurde [14]. Wie der Name andeutet, besteht die Aufgabe dieses Flavoproteins in der Reduktion eines anderen Proteins, des Thioredoxins, das dazu dient, eine Vielzahl weiterer Substrate zu reduzieren [15]. So ist es am Metabolismus der Deoxyribonukleotide, der Bausteine der DNA, beteiligt. Dabei werden die Redoxreaktionen durch die Übertragung von Wasserstoffatomen ermöglicht, die an die Cysteine der aktiven Zentren dieser Proteine gebunden sind. Um einen Einblick in den Aufbau der Thioredoxin Reduktase zu erhalten, wurde ihre Struktur der Protein Data Bank (PDB) (Code 1TDE) [16] entnommen und mit Hilfe des Programms VMD [17] visualisiert. In Teilabbildung 1.6 a sind alle Atome des Proteinkomplexes in einem Kugel-Stäbchen-Modell dargestellt, um die Komplexität dieses 316 Residuen umfassenden Proteins zu verdeutlichen. Erst durch das Weglassen der Seitengruppen und die vereinfachte Darstellung 1.6 b des Proteinrückgrats kann die dreidimensionale Struktur des Proteins erkannt werden. In Übereinstimmung mit der üblichen Praxis wurden Sekundärstrukturen durch spezielle geometrische Formen gekennzeichnet: Zylinder für Helizes und flache Bänder für β -Faltblätter.

Mit diesem Beispiel eines Enzyms will ich die Vorstellung des Aufbaus der Peptide und Proteine abschließen. Vertiefungen finden sich in vielen gängigen Lehrbüchern [2, 9, 18].

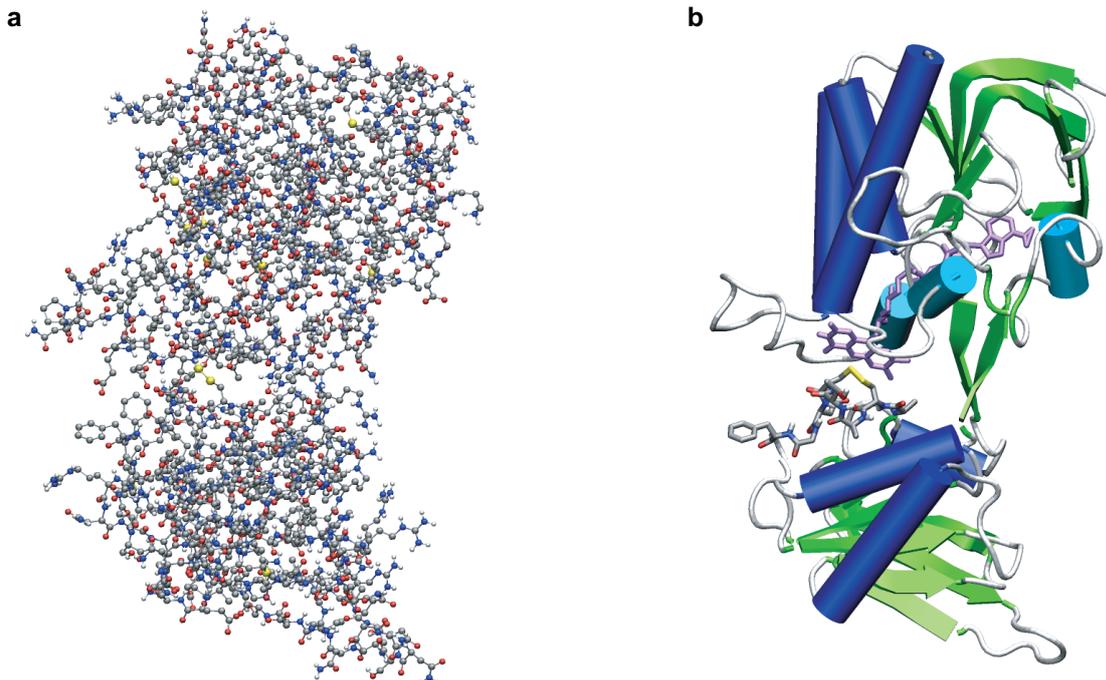


Abbildung 1.6: Die 3D-Struktur der Thioredoxin Reduktase. **a** Dargestellt als Kugel-Stäbchen-Modell der einzelnen Atome und ihrer Bindungen, und **b** als „Cartoon“ der Sekundärstrukturen. Hierbei sind α -Helizes dunkelblau, 3_{10} -Helizes hellblau, β -Faltblätter gelb und ungeordnete Teile des Rückgrats grau eingefärbt. Die FAD-Gruppe ist violett hervorgehoben. Die Residuen 134-141 der Bindungstasche, deren Aminosäuresequenz als Vorlage für die in dieser Arbeit untersuchten Modellpeptide diente, werden explizit dargestellt.

Die Dynamik und das Faltungsproblem

Im Folgenden soll die Hierarchie der in Proteinen auftretenden dynamischen Prozesse skizziert werden, die sich aus ihrem oben diskutierten hierarchischen Aufbau ergibt. Tabelle 1.1 bietet einen Überblick über die wichtigsten Zeitskalen. Die in Proteinen stattfindenden Prozesse und Dynamiken umfassen:

(i) Kleinskalige Schwingungen und thermische Fluktuationen von Atomen oder kleinen Atomgruppen um ihre durch die lokalen Bindungen gegebenen Gleichgewichtslagen. Diese Schwingungen finden im Zeitbereich einiger Femtosekunden (fs, 10^{-15} s) statt [12].

(ii) Rotationen von Seitengruppen und Flips einzelner Peptidplättchen laufen auf Zeitskalen von Pikosekunden (ps, 10^{-12} s) bis Nanosekunden (ns, 10^{-9} s) ab [12].

(iii) Kollektive Bewegungen ganzer Sekundärstrukturen, die z.B. mit scharnierartigen Verformungen der Polypeptidkette verbunden sein können, überdecken den Zeitbereich von 100 ns bis zu Millisekunden (ms, 10^{-3} s). Derart langsame Kinetiken wurden beispielsweise kürzlich an der Natrium-Kalium-ATPase beobachtet [19].

| Bewegung/Prozess | Zeitskala |
|---|--------------------------|
| Bindungs- und Winkelschwingungen | 10 fs ... 100 fs |
| Bewegungen von Peptid- und Seitengruppen | 10 ps ... 1 ns |
| Kollektive Bewegungen | 10 ns ... 1 ms |
| Bildung von α -Helizes in Peptiden | 10 ns ... 100 ns |
| Aufbau von β -Faltblättern („Hairpins“) | 1 μ s ... 10 μ s |
| Proteinfaltung | 4 μ s ... 1000 s |

Tabelle 1.1: Zeitskalen der Dynamik von Proteinen. Die unterschiedlichen Bewegungen und Prozesse erstrecken sich zusammengenommen über einen Zeitbereich von etwa 17 Zehnerpotenzen. Die Referenzen finden sich im Text.

(iv) Der Aufbau helikaler Sekundärstrukturelemente wird in den Nanosekunden-Bereich eingeordnet [20]. Schnellere Prozesse auf der ps-Zeitskala können jedoch bei der Nukleation von α -Helizes eine entscheidende Rolle spielen.

(v) Der erste Schritt hin zu einem β -Faltblatt besteht in der Anordnung zweier Ketten zu einem *Hairpin* (vgl. Abbildung 1.5 b). Dieser Vorgang ist entropisch sehr ungünstig, weswegen er mit einer hohen Barriere in der *freien Energie* verbunden ist und dementsprechend mit einer Zeitkonstante im Bereich einiger Mikrosekunden (μ s, 10^{-6} s) recht langsam verläuft [20].

(vi) Die *Proteinfaltung* als der gesamte Vorgang, in dem ein Protein von ungeordneten, ausgedehnten Konfigurationen zu seiner biologisch funktionalen Struktur findet, ist als zusammengesetzter Prozess in Tabelle 1.1 mit aufgenommen. Als Rekordhalter schnellster Faltungsprozesse gelten momentan kleine Proteine, die mit einer Zeitkonstante von 4 μ s ihre native Form annehmen [21]. Große Eiweiße benötigen dagegen deutlich länger für ihre Faltung.

Neben einer Trennung der zeitlichen Skalen unterscheiden sich die aufgeführten dynamischen Prozesse auch hinsichtlich der Größenordnung der mit ihnen verbundenen Geometrieänderungen eines Proteins. Während der schnellen Fluktuationen (i) durchläuft ein Protein schnell viele, geometrisch sehr ähnliche Konfigurationen. Dagegen führen die übrigen Prozesse teilweise zu Änderungen großer Teile der Struktur und damit der so genannten *Konformation* eines Proteins. Mit dem Begriff Konformation wird dabei ein metastabiler Zustand eines Moleküls bezeichnet, in dem es eine gewisse Zeitspanne lang verweilt und thermisch fluktuiert, bevor es relativ plötzlich in einen anderen – meist auch geometrisch deutlich verschiedenen – Zustand übergeht. Damit entspricht eine Konformation einem lokalen Minimum der freien Energie, das von anderen Konformationen durch Barrieren getrennt ist [22, 23].

Da die Energielandschaft eines Proteins äußerst komplex ist und viele lokale Minima aufweist, existieren auch viele mögliche Konformationen eines solchen Moleküls. Jedoch stellten Anfinsen et al. [24] schon vor mehr als 40 Jahren die Hypothese auf [25], dass

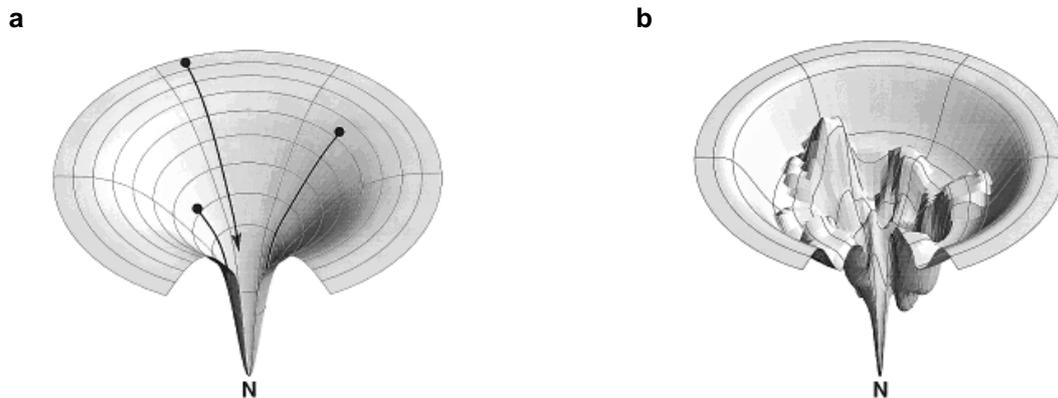


Abbildung 1.7: Energielandschaften der Proteinfaltung. **a** Die „New View“ erklärt die Dynamik der Proteinfaltung durch Energielandschaften, die trichterförmig auf den nativen Zustand N zulaufen. und **b** Obwohl in diesem Modell keine ausgezeichneten Reaktionspfade existieren, können Unebenheiten in der Energielandschaft zu intermediären Zuständen und komplexen Reaktionsraten führen (Abbildungen entnommen aus [29]).

Proteine reversibel zu genau einer, nur durch die Sequenz der Aminosäuren bestimmten, nativen Struktur oder Konformation falten, die dem globalen Minimum der freien Energie entspricht. Die Unmöglichkeit, dieses Minimum durch ein zufälliges Absuchen des *Konformationsraumes* in endlicher Zeit zu finden, wurde durch Levinthal [26] formuliert und später als „Levinthals Paradoxon“ bekannt [27]. Eine Überschlagsrechnung zeigt dies: wenn man für die Torsionswinkel der 100 Aminosäuren eines fiktiven Proteins nur zwei erlaubte Winkelkombinationen annimmt, ergibt dies schon $2^{100} \approx 10^{30}$ mögliche Konfigurationen. Nimmt man jede dieser Konfigurationen als mögliche Konformation an und geht nun davon aus, dass das Protein in einer Nanosekunde eine dieser Konfigurationen testen kann, so dauert das Absuchen des gesamten Konformationsraumes etwa 10^{13} Jahre und damit entschieden länger als die beobachteten Zeiten der Proteinfaltung.

Als relativ einfache Modelle zur Erklärung der Proteinfaltung wurde später die Existenz von Reaktionspfaden angenommen, die von einem denaturierten Zustand über Sequenzen von Intermediaten zum nativen, gefalteten Zustand führen [28]. Durch die Annahme solcher Reaktionspfade konnte die Beobachtung multi-exponentieller Kinetiken in Experimenten zur Proteinfaltung erklärt werden, problematisch ist aber die in diesem Modell auftretende Abhängigkeit der Endstruktur von den Startbedingungen, d.h. von der Wahl des denaturierten Zustands.

Die Vorstellung von einzelnen Molekülen, die sich entlang definierter Reaktionspfade falten, wurde in den letzten Jahren durch die so genannte „New View“ [30] abgelöst. Diese spricht von Ensembles von Proteinen, die sich auf einer trichterförmigen Energielandschaft bewegen [31, 32]. Die Zuspitzung der hochdimensionalen Energiefläche hin zur nativen Struktur, wie sie Abbildung 1.7 a in zwei Dimensionen schematisch zeigt, ist dabei eine Folge verschiedenster Prozesse, die sukzessiv den Konformationsraum des Mole-

küls einschränken und seine Freie Energie minimieren. Beispiele hierzu könnten der *hydrophobe Kollaps* sein, in dem sich die Seitengruppen der wasserlöslichen Aminosäuren nach außen richten, während die wasserabweisenden Residuen einen dichter gepackten Kern des Proteins bilden, oder auch die oben skizzierte Ausbildung von Sekundärstrukturelementen. Auf diese Weise umfasst die „New View“ nicht nur die besprochene Mehrskaligkeit der Proteine, sondern löst auch Levinthals Paradoxon, da sie unabhängig von der Ausgangsstruktur eine schnelle Faltung hin zu einer einzigen nativen Konformation beschreibt.

1.2 Experimentelle Methoden zur Untersuchung von Biomolekülen

Das bisher vorgestellte Verständnis der Struktur und Dynamik von Proteinen und Peptiden ist das Ergebnis von einer Vielzahl experimenteller und theoretischer Untersuchungen, die in den letzten 50 Jahren an solchen Systemen vorgenommen wurden. Nun sollen zunächst einige der dabei eingesetzten experimentellen Methoden etwas näher beschrieben werden, vor allem jene, die bei der experimentellen Charakterisierung der in dieser Arbeit untersuchten Modellpeptide eingesetzt wurden (vgl. Abschnitt 1.4). Dies ist nötig, weil meine eigenen theoretischen Untersuchungen darauf abzielen, die mit diesen Methoden gewonnenen Daten zu verstehen und zu interpretieren. Anschließend folgt eine Vorstellung theoretischer Ansätze und auch hier liegt das Gewicht auf jenen, die ich bei meinen Untersuchungen eingesetzt habe.

Röntgenstreuung

Die Röntgenstreuung ist die meist verwendete Methode zur Strukturaufklärung von Biomolekülen: etwa 85% der bisher in der PDB Datenbank hinterlegten Strukturen sind auf diese Weise bestimmt worden [16]. Abbildung 1.8 zeigt, dass dieses Verfahren es ermöglicht, auch im Fall großer Proteine die Orte aller schweren Atome relativ direkt und mit sehr hoher Auflösung zu bestimmen. Somit können nicht nur die Orientierungen von Seitenketten oder der Aufbau von Sekundärstrukturen, sondern insbesondere auch die Tertiärstruktur eines Proteins oder die Quartärstruktur ganzer Proteinkomplexe charakterisiert werden.

So stark die Röntgenstreuung bei der Strukturuntersuchung ist, so wenig kann sie bisher über dynamische Eigenschaften von Peptiden aussagen. Allenfalls kann der gemessene Temperaturfaktor Einsicht in die Größe der Fluktuationen einzelner Atome oder kleiner Gruppen geben [34]. Viel versprechend sind deshalb erste zeitaufgelöste Messungen an Proteinkristallen [35]. Die dabei angewendete Technik gehört zu den *Pump-Probe* Methoden, bei denen zuerst ein dynamischer Prozess durch einen ultraschnellen „Pump“-Laserpuls initiiert wird, um dann nach einer variablen Verzögerungszeit t den Zustand des

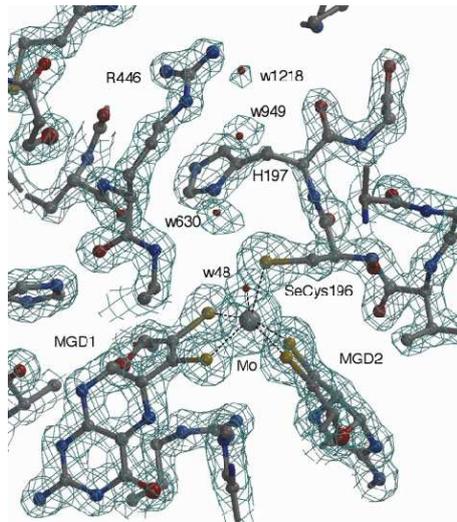


Abbildung 1.8: Das Ergebnis einer Röntgenstrukturanalyse. Gezeigt ist die mit $1,6 \text{ \AA}$ aufgelöste Elektronendichte der Bindungstasche des Membranproteins Fdh-N aus *Escherichia coli*. Sie wurde von Jormakka et al. [33] an der ESRF in Grenoble vermessen.

Systems mit einem „Probe“-Puls, in diesem Fall Röntgenstrahlung, abtasten zu können. Moderne Synchrotronstrahlungsquellen können hierzu breitbandige Röntgenpulse von etwa 100 ps Dauer erzeugen. Diese Abtastpulse ermöglichen es, mit der Laue-Methode die sich dynamisch ändernden Proteinstrukturen mit einer Sub-Nanosekunden-Auflösung über Zeitbereiche von Nanosekunden bis Millisekunden zu verfolgen. Allerdings steht die Datenauswertung in diesem Gebiet vor der Aufgabe, sich zeitlich überlagernde Prozesse zu trennen und robust gegenüber dem niedrigen Signal-zu-Rausch-Verhältnis zu sein.

Neutronenstreuung

Während Wasserstoffatome im Fall der Röntgenstreuung nur eine sehr kleine Streulänge aufweisen und damit nicht oder nur sehr schwer detektiert werden können, gilt dies nicht für die Streuung von Neutronen [34, 36]. Deshalb kann die Neutronenstreuung Aussagen zu biologisch wichtigen Fragestellungen liefern, die z.B. die Bildung von Wasserstoffbrücken in Bindungstaschen von Proteinen betreffen. Dabei ist es möglich, durch inelastische Streuung Informationen über dynamische Prozesse in Proteinen zu erhalten. Diesen Vorteilen stehen aber große Hürden gegenüber, da solche Experimente einen hohen Neutronenfluss benötigen, wie ihn nur moderne Reaktoren oder Spallationsquellen liefern können [36]. Eine Messung von zeitaufgelösten Strukturen mittels Neutronenstreuung scheint deswegen erst durch in Planung befindliche Großgeräte, wie eine gepulste Spallationsquelle, erreichbar zu sein. Von einer gewissen Relevanz für die vorliegende Arbeit ist die Möglichkeit, mit der Neutronenbeugung an ungeordneten Systemen Aussagen über radial gemittelte Verteilungsfunktionen zu machen. Da solche Verteilungsfunktionen beispielsweise für Lösungsmittel auch aus Computersimulationen ermittelt werden kön-

nen, bieten sie sich für einen Vergleich zwischen der Theorie und dem Experiment an. In Kapitel 3 werde ich derartige Vergleiche anstellen.

Kernspinresonanz

Von größerer Bedeutung für meine Untersuchungen ist jedoch die Strukturaufklärung durch Messung der Kernspinresonanz (NMR, nuclear magnetic resonance), da mit dieser Methode auch die Strukturen der Modellpeptide bestimmt wurden [37, 38]. NMR-Spektren von Proteinen und Peptiden können in Lösung und damit in der natürlichen Umgebung dieser Biomoleküle aufgezeichnet werden [39, 40]. An einem Protein wurde dies zuerst 1957 von Saunders et al. [41] gezeigt. Im Gegensatz zu den oben beschriebenen Beugungstechniken wird bei der NMR-Strukturuntersuchung auf Grund der Verwendung einer wässrigen Lösung die biologische Funktion der Proteine nicht durch Kristallisationseffekte eingeschränkt.

Das am besten mit NMR zu beobachtende, auf natürliche Weise in Biopolymeren enthaltene Atom ist der Wasserstoff (^1H), dessen Kern einen halbzahligen Spin aufweist. Die Wechselwirkung der Kernspins einer Probe mit einem von außen angelegten Feld muss quantenmechanisch beschrieben werden; Abbildung 1.9 zeigt jedoch ein anschauliches und klassisches Modell des zu Grunde liegenden Prinzips. Aktuellste NMR-Instrumente erreichen eine Magnetfeldstärke von mehr als 21 Tesla, in der die ^1H -Spins mit einer Frequenz von $\nu_0 = 900\text{ Mhz}$ präzedieren. Diese Bewegung des magnetischen Moments innerhalb der x,y-Ebene induziert in einer Detektorspule einen messbaren Strom, der die Observable der Messung darstellt. Während das System in das Gleichgewicht zurückkehrt und die Spins ihre Kohärenz verlieren, nimmt die Induktion mit der Netto-Magnetisierung ab. Aus dem so aufgenommenen Signal kann das NMR-Spektrum durch Fourier-Transformation errechnet werden. In einem solchen Spektrum wird deutlich, dass die Protonen durch ihre chemischen Umgebung beeinflusst werden. Die kovalent angebondenen Nachbaratome oder auch Kernspins in räumlicher Nähe des betrachteten Wasserstoffatoms ändern das lokale Magnetfeld und bewirken so eine Verschiebung der gemessenen Präzessionsfrequenz. Diese „chemische“ Verschiebung δ wird üblicherweise unabhängig von der verwendeten Magnetfeldstärke als Abweichung von einer Referenz in ppm (parts per million) angegeben und nimmt Werte von 1 bis 10 an. Die durch kovalente Bindungen vermittelte Wechselwirkung wird als „J coupling“ bezeichnet. Für 3J , den Fall von drei intermediären Bindungen, gilt die Karplus-Beziehung [42]

$$^3J = A \cos^2 \phi + B \cos \phi + C \quad (1.1)$$

zwischen der Stärke der Wechselwirkung und dem durch die Atome gebildeten Dihedralwinkel ϕ . Dabei sind A , B und C Konstanten, deren Werte für die verschiedenen auftretenden Torsionswinkel bekannt und tabelliert sind.

Einen direkten Einblick in die Struktur eines Proteins ermöglicht der *Nukleare Overhauser Effekt* (NOE), eine per Dipol-Dipol-Kopplung durch den Raum vermittelte Wechselwirkung [43]. Prinzipiell zeigt der NOE eine r_{ij}^{-6} -Abhängigkeit von r_{ij} , dem Abstand zweier

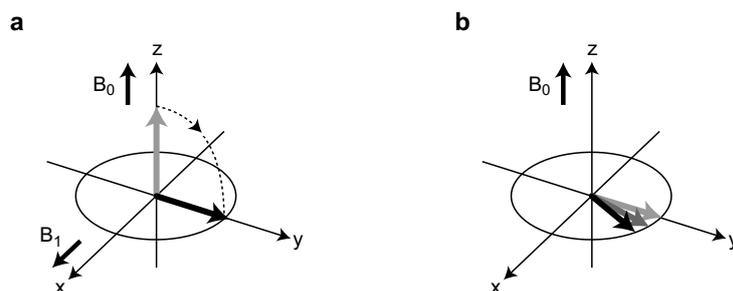


Abbildung 1.9: Eine klassische Illustration des NMR-Prinzips. a Durch ein starkes Magnetfeld B_0 werden die Kernspins (grau) polarisiert und in Richtung der z-Achse ausgerichtet: es entsteht eine makroskopische Magnetisierung der Probe. Ein elektromagnetischer Puls passender Dauer B_1 (ein so genannter 90° -Puls) übt anschließend ein Drehmoment auf die Spins aus, das diese in die x,y-Ebene dreht. **b** Die Spins präzedieren nun unter dem Einfluss von B_0 mit der Lamor-Frequenz ν_0 um die z-Achse.

Wasserstoffkerne i und j . Zur Messung finden hierbei verschiedene Methoden der 2D-NMR Spektroskopie Anwendung (z.B. NOESY, nuclear Overhauser enhancement and exchange spectroscopy), bei denen mit einem zusätzlichen Feld B_2 sukzessiv einzelne Resonanzen angeregt und deren Einflüsse auf die Spektren der anderen Spins gemessen werden. Die so ermittelten NOEs liefern obere Grenzen für den Abstand der zwei Wasserstoffkerne, wobei Distanzen kleiner als etwa 5 \AA beobachtet werden können [43].

Die Zuordnung der Resonanzen im 2D-Spektrum zu einzelnen Atomen geschieht anschließend in zwei Schritten. Zuerst kann der Typ des Residuums durch einen Vergleich mit bekannten Resonanzmustern bestimmt werden, bevor mit Hilfe der Primärstruktur des Peptids und den Inter-Residue-NOEs zwischen den Aminosäuren ihre genaue Identität festgelegt wird. Auf diese Weise resultiert die Analyse der NMR-Daten in einer Menge von Abständen zwischen spezifischen Paaren von Wasserstoffatomen, den so genannten *Restraints*. Um aus diesen Informationen eine dreidimensionale Struktur des Moleküls zu erhalten, ist eine Modellbildung unter Zuhilfenahme der Topologie der Aminosäuren mit ihren angenommenen Bindungslängen und -winkeln notwendig. Durch Verfahren wie das „simulated annealing“ wird das Optimierungsproblem gelöst, einen Satz von Backbone-Dihedralwinkeln zu finden, der die Restraints erfüllt [44]. Das Ergebnis dieser Berechnungen ist üblicherweise keine einzelne Struktur, sondern ein Ensemble von gleichermaßen gültigen Modellen.

In den letzten Jahren wurden zusätzlich zu den etablierten Methoden der Strukturbestimmung durch NMR solche zur Messung von Spin-Relaxationen entwickelt [45, 46], die auf der ps-ns Zeitskala einen Einblick in die Konformationsfluktuationen von Proteinen ermöglichen. Sie können dabei Aussagen über die Rotationsdiffusionskonstanten der Seitengruppen oder ganzer Proteindomänen machen. Allerdings ist zur Interpretation dieser Ergebnisse stets eine Modellbildung oder der Vergleich mit Simulationen unerlässlich.

Schnelle spektroskopische Techniken

In den bisher besprochenen experimentellen Techniken stand die Bestimmung der Gleichgewichtsstrukturen nativ gefalteter Proteine im Vordergrund. Um Einblicke in die Funktions- und Faltungsdynamik dieser Makromoleküle zu erhalten, sind Methoden zur schnellen Auslösung einer geeigneten Reaktion und zur Beobachtung aussagekräftiger Observablen notwendig. Einen guten Überblick zu diesem Thema bietet der Review-Artikel von Eaton et al. [20].

Die ersten Experimente auf diesem Gebiet verwendeten zur schnellen Auslösung der Faltung die „Stopped Flow“-Technik, bei der chemisch denaturierte Proteine möglichst schnell mit einer Pufferlösung vermischt werden, die das Denaturant verdünnt und damit die Faltung einleitet. Umgekehrt kann auch eine Entfaltungsreaktion durch das Zumischen des Denaturants eingeleitet werden. Die Rate der mikroskopischen Vermischung der Flüssigkeiten limitiert jedoch bei den „Stopped Flow“-Methoden die Zeitauflösung auf einige Millisekunden [20]. Der Zeitbereich der Nanosekunden wurde erst durch Temperatursprungexperimente untersuchbar, bei denen eine Probe mit einem kurzen Laserpuls über die „Schmelztemperatur“ der Entfaltung erhitzt wird [47–50]. Auch kann ein durch Piezokristalle erzeugter Drucksprung die Faltung auslösen, da sie stets mit einer Volumenänderung des Proteins einhergeht [51]. Für Proteine, die wie das Cytochrom Metallatome beinhalten, wurden Reaktionen erfolgreich durch einen photoinduzierten Elektronentransfer gestartet [52], allerdings ist die Auswahl an untersuchbaren Systemen bei dieser Methode gering. Ferner ist eine direkte photochemische Beeinflussung durch das lichtinduzierte Aufbrechen von Bindungen erreichbar, z.B. durch das Verbinden einer Cysteinbrücke mit einem photolabilen Disulfid [53]. Schließlich kann eine ultraschnelle Auslösung von Konformationsänderungen durch den Einbau von Farbstoffen in das zu untersuchende Molekül erreicht werden. Hierzu geeignete Farbstoffe zeigen nach Anregung durch einen Laserpuls Geometrieänderungen im Femtosekundenbereich. Ein weiterer Vorteil dieses Ansatzes ist, dass der gleiche Farbstoff auch als Sonde für die ablaufenden Konformationsänderungen dienen kann. Der Einbau eines Farbstoffs in ein Biomolekül ist die experimentelle Grundlage der vorliegenden Arbeit und wird deshalb in Abschnitt 1.4 detailliert dargestellt.

Wie schon im Fall der zeitaufgelösten Röntgenstreuung erwähnt wurde, ist die Kombination aus einem Laserpuls zum Auslösen einer Reaktion mit einem variabel zeitlich verzögerten Folgepuls zur Beobachtung als „Pump-Probe“-Experiment bekannt. Während die Wellenlänge des Pump-Pulses durch die zu erzielende Anregung der Probe bestimmt wird, z.B. Infrarot (IR) für Temperatursprünge und im sichtbaren Spektrum oder UV für elektronische Übergänge, stehen für den Probe-Vorgang eine große Anzahl an spektroskopischen Techniken zur Verfügung [34].

Beginnend bei niedrigen Energien gibt die *Infrarotspektroskopie* Auskunft über die Normalmoden der Molekülschwingung im elektronischen Grundzustand. Besonders die *Amidbanden* der Peptidgruppen bei Wellenzahlen von etwa 1600 cm^{-1} sind empfindliche Observablen für die Konfiguration eines Peptids. Auf diese Weise konnten Williams et al. [47] zeigen, dass ein helikales Peptid aus 21 Alanin-Residuen nach einem Temperatur-

sprung mit Zeitkonstanten im Bereich von 20 ns ent- und refaltet. Auch Huang et al. [50] beobachteten mit dieser Methode den „Helix-Coil“ Übergang von Peptiden und fanden eine Kinetik, die sich nur durch einen gedehnten exponentiellen Zerfall^a modellieren lassen kann. Solch ein Verhalten kann durch einen Faltungsprozess erklärt werden, der auf einer Energielandschaft abläuft, wie sie Abbildung 1.7 b zeigt. Seit einiger Zeit werden zudem Verfahren der 2D IR-Spektroskopie entwickelt [54] und auf kleine Systeme wie das Trialanin angewendet [55, 56]. Die Dipol-Dipol-Wechselwirkung zwischen den einzelnen Peptidgruppen trägt dabei prinzipiell ähnliche Informationen wie die NOEs der NMR.

Erfolgt die Abtastung der Probe mit Licht im sichtbaren oder ultravioletten Bereich (UV-vis), so werden in der *transienten Spektroskopie* von elektronischen Übergängen herührende Absorptionsbanden vermessen. Zeigen die untersuchten Farbstoffe in Folge struktureller Umwandlungen, der Anregung von Schwingungen oder chemischer Reaktionen ein geändertes Absorptionsspektrum, so können diese Prozesse mit hoher Zeitaufösung beobachtet werden. Beispielhaft seien hier Untersuchungen genannt, die schon 1988 von Dobler et al. [57] am Bakteriorhodopsin und kürzlich von Lutz et al. [58] am bakteriellen Sensorhodopsin durchgeführt wurden. Allerdings sind solche lichtschtbaren Proteine so groß und kompliziert, dass sie sich einer komplementären theoretischen Beschreibung bisher entzogen haben.

1.3 Theoretische Beschreibungen

Bei einer Vielzahl experimentell gewonnener Daten zur Proteinstruktur und -dynamik können erst vermittels adäquater theoretischer Beschreibungen Schlüsse auf die molekularen Strukturen und ihre Veränderungen gezogen werden (vgl. z.B. die oben skizzierte Konstruktion von NMR-Strukturen). Angesichts der Komplexität der zu behandelnden Systeme müssen dabei rechnergestützte Verfahren eingesetzt werden. Unterstützt durch die ständig anwachsende Leistungsfähigkeit der verfügbaren Computer, nimmt deshalb die Bedeutung theoretischer Ansätze stetig zu. Dabei macht es nicht nur die besprochene Vielskaligkeit der Dynamik von Biopolymeren, sondern auch das Vorliegen eines Vielteilchenproblems (N Teilchen mit $O(N^2)$ Wechselwirkungen) notwendig, für jede gegebene Fragestellung oder zu erfassende Eigenschaft eine Wahl der zu verwendenden theoretischen Methode unter Abwägung von Rechenaufwand und zu erzielender Genauigkeit zu treffen. Tabelle 1.2 zeigt einen Überblick über die Anwendungsbereiche gegenwärtig verfügbarer Verfahren, die ich, beginnend mit den quantenmechanischen Methoden, im Folgenden vorstellen werde.

^aEin „stretched exponential“ hat die Form $a(t) = A \cdot [1 - B \cdot \exp(-t/\tau)^\beta]$, Huang et al. [50] bestimmten Werte von $\beta \approx 0.8$ und $\tau \approx 250$ ns.

| Methoden | Eigenschaften/Systeme | Atome | Zeitskala |
|--|--|-------------------------------|-----------|
| Post-HF Methoden der Quantenmechanik | - Exakte Minimalstrukturen - 1D Potentialflächen - Angeregte Zustände | < 50 | - |
| Dichtefunktionaltheorie | - Minimalstrukturen - 2D Potentialflächen - Schwingungsmoden | < 100 | ps |
| QM/MM-Hybridrechnungen | - Aktive Zentren in Proteinen - Kleine Moleküle in Lösung | QM < 100 MM $\approx 10^5$ | 10 ps |
| Molekularmechanik mit explizitem Lösungsmittel | - Proteine in nativer Umgebung - Konformationsdynamik - Thermodynamische Ensembles | $\approx 10^5$ | ns |

Tabelle 1.2: Die Anwendungsbereiche verschiedener theoretischer Methoden. Nach abnehmender Genauigkeit sortiert, werden die Eigenschaften und mögliche Einsatzgebiete verfügbarer Methoden aufgeführt. Ihre jeweilige Anwendbarkeit wird sowohl durch die Zahl der zu behandelnden Atome als auch durch die zu erreichende Zeitskala der Simulation beschränkt.

Quantenmechanische *ab initio*-Verfahren

Die exaktesten Verfahren zur Beschreibung eines Moleküls sind quantenmechanischer Natur, beinhalten keine freien Parameter zur Anpassung an Messergebnisse und sind als „*ab initio*“-Verfahren der Quantenmechanik (QM) bekannt. Ihre Aufgabe besteht darin, für eine gegebene Konfiguration der Atomkerne die elektronischen Wellenfunktionen der Grund- und der angeregten Zustände zu bestimmen. Das bekannte Hartree-Fock-Verfahren (HF) konstruiert von Atomorbitalen ausgehend einen Satz von Molekülorbitalen für den Grundzustand. Die Wechselwirkung der Elektronen wird dabei durch ein mittleres Feld genähert, dessen Gestalt durch ein iterativ zu lösendes Eigenwertproblem gegeben ist. Die im HF-Verfahren verwendete *mean-field Näherung* vernachlässigt allerdings Effekte der Elektronenkorrelation, die z.B. besonders in den π -Elektronensystemen von Farbstoffen einen großen Beitrag liefern [59]. So genannte *post-HF*-Theorien können zwar diese Korrelationseffekte berücksichtigen, ihr Rechenaufwand liegt jedoch um Größenordnungen über dem der HF-Verfahren (vgl. z.B. das von Strodel und Tavan [60] angegebene Skalierungsverhalten so genannter Konfigurations-Wechselwirkungs-Verfahren), weshalb sie bis auf weiteres für die Beschreibung biologischer Systeme nicht verwendet werden können.

Dichtefunktionaltheorie

Schon besser geeignet für eine Beschreibung größerer Biomoleküle ist das Verfahren der Dichtefunktionaltheorie (DFT), welches auch im Rahmen dieser Arbeit häufig eingesetzt wurde. Die DFT hat seit etwas mehr als einem Jahrzehnt eine herausragende Bedeutung in der theoretischen Chemie gewonnen, die durch die Verleihung des Nobelpreises für Che-

mie im Jahr 1998 an Walter Kohn gewürdigt wurde [61]. Sie basiert auf den Hohenberg-Kohn-Theoremen [62], nach denen zur Beschreibung des Grundzustands eines Moleküls die Elektronendichte $\rho^0(\mathbf{x})$ genügt. Letztere lässt sich mit einem zum HF-Verfahren vergleichbaren Rechenaufwand durch Anwendung der Kohn-Sham Gleichungen [63] mittels eines selbstkonsistenten Verfahrens berechnen. Durch die Verwendung von empirisch parametrisierten Austausch- und Korrelationsfunktionalen ist die DFT zudem besser als das HF-Verfahren in der Lage, die Elektronenkorrelation zu erfassen. In dieser Arbeit finden neben dem einfachsten Funktional, der nur auf der lokalen Elektronendichte beruhenden *local-density-approximation* (LDA), auch gradientenkorrigierte Funktionale wie BLYP [64, 65] oder das hybride B3LYP [66] Verwendung (siehe dazu die Monographie von Parr und Yang [67]).

Die DFT-Methode ist in einer Vielzahl von Programmen implementiert worden, von denen in der vorliegenden Arbeit sowohl GAUSSIAN [68] als auch CPMD [69] Verwendung finden. GAUSSIAN als der Klassiker unter den quantenmechanischen Programmpaketen zeichnet sich durch eine einfache Bedienung aus und eignet sich Dank einer Vielzahl von eingebundenen Optimierungsalgorithmen besonders zur Bestimmung molekularer Minimalstrukturen und zur Untersuchung der Form von Potentialflächen.

Während in GAUSSIAN die Molekülorbitale nach einem Satz von Gaußfunktionen entwickelt werden (in der vorliegenden Arbeit wurde stets der Basissatz 6-31G* eingesetzt), verwendet CPMD eine Entwicklung nach ebenen Wellen. Die Anzahl der ebenen Wellen in der Basis muss dabei durch die Angabe einer größten Wellenzahl oder maximalen kinetischen Energie^a angegeben werden. Dieser Energie-Cutoff bestimmt somit durch die Fourier-Transformation die maximale Auflösung des realen Raums, in dem die Elektronendichte dargestellt wird.

Große Dichtegradienten entstehen besonders durch die eng gebundenen Rumpfelektronen der Atome, deren Einfluss und Abschirmung der Kernladung deshalb bei CPMD durch *Pseudopotentiale* beschrieben wird. So können in der *frozen-core*-Näherung allein die Valenzelektronen betrachtet werden. Dies führt zu einer Glättung der Elektronendichte und einer Verringerung des Rechenaufwandes. Von Bedeutung für die vorliegende Arbeit sind die Pseudopotentiale von Troullier und Martins [70] (MT) und die von Goedecker et al. [71] (SG), deren Parametrisierung den zitierten Arbeiten entnommen werden kann.

Zwar ist CPMD durch den *Car-Parinello*-Ansatz [72] in der Lage, auf effiziente Weise gleichzeitig die Zeitentwicklung der Kernkoordinaten und der Elektronendichte zu behandeln, doch sind, wie Tabelle 1.2 ausweist, die mit der DFT-Methode behandelbaren Systeme zu klein für die Beschreibung der hier betrachteten Modellpeptide in Lösung [73]. Trotzdem spielen DFT-Berechnungen im Folgenden eine wichtige Rolle bei der Erstellung effektiver Molekülmodelle, denen wir uns nun zuwenden wollen.

^aCPMD verwendet hier Energie-Einheiten von Rydberg (Ry). Es gilt $1 \text{ Ry} = 2 \text{ Hartree} \approx 54,42 \text{ eV}$.

Molekularmechanische Modelle

Die Entwicklung so genannter molekularmechanischer (MM) Molekülmodelle oder *Kraftfelder* hat es ermöglicht, die Dynamik von Peptiden und Proteinen mit *Molekulardynamiksimulationen* (MD-Simulationen) [74, 75] auf atomarer Auflösung zu verfolgen.

Die Ergebnisse dieser Entwicklung liegen in Form von MM-Standardkraftfeldern vor, welche jeweils eine mögliche empirische Parametrisierung des durch die Elektronen erzeugten und auf die Atomkerne wirkenden Potentials darstellen. Häufig werden die Kraftfelder AMBER [76], CHARMM [77], GROMOS [78] oder OPLS [79] eingesetzt, die zu meist zusammen mit einem gleichnamigen Programmpaket verfügbar sind. Dagegen wurde von mir für die Simulationen der Modellpeptide ein eigenes, in unserer Gruppe entwickeltes Simulationsprogramm (EGO-MMII) [80] verwendet, das eine erweiterte Form des CHARMM22-Kraftfeldes [81] implementiert. In einer MD-Simulation werden die Atome der untersuchten Moleküle durch Punktmassen repräsentiert und die auf sie wirkenden Kräfte aus dem MM-Potential abgeleitet. Durch eine zeitdiskrete Integration der klassischen Newtonschen Bewegungsgleichungen wird in den Simulationen dann die Zeitentwicklung des molekularen Systems verfolgt [74, 75].

Bei solchen MM-MD-Simulationen werden häufig Systeme betrachtet, die eine große Anzahl von Atomen umfassen. Dies ist nicht nur durch die Größe der untersuchten Peptide oder Proteine selbst begründet, sondern auch dadurch, dass diese Biomoleküle ihre Struktur und Funktionsdynamik nur in ihrer jeweiligen physiologischen Umgebung zeigen [2], die im einfachsten Fall durch Wasser als Lösungsmittel gegeben ist. Daher ist es notwendig, in ein Simulationsystem zusätzlich zum Biomolekül eine Vielzahl von Atomen des Lösungsmittels mit aufzunehmen.

Zwar wurde 1977 bei den ersten MD-Simulationen eines Proteins [6], des *Bovine Pancreatic Trypsin Inhibitors* (BPTI), das Lösungsmittel weggelassen, aber schon einige Jahre später konnten van Gunsteren und Berendsen [82] die erste MD-Simulation eines Proteins in wässriger Lösung vorstellen. Seit diesen Anfängen wurde die Methode der MD-Simulation stetig weiterentwickelt und eine Vielzahl von Untersuchungen an Biomolekülen durchgeführt [7, 8, 83–91]. Manche dieser Arbeiten zielten sogar darauf ab, experimentell beobachtbare Größen zu berechnen, um so die Validität des MD-Verfahrens zu überprüfen [84, 87, 89, 90]. So haben beispielsweise Daura et al. [86, 87] die Gleichgewichtsfluktuationen eines Beta-Peptids^a in Methanol über 100 ns simuliert und mit NMR-Daten verglichen. Ferner konnten z.B. Zhou et al. [90] mit einer geeigneten Kombination aus MD- und Monte-Carlo-Simulationen die beobachtete Temperaturabhängigkeit der Kinetik einer β -Hairpin-Bildung reproduzieren.

In allen MM-Kraftfeldern werden die chemischen Bindungen der Atome durch kurzreichweitige Modellpotentiale dargestellt, welche die elastischen Deformierbarkeiten der Bindungslängen, Bindungswinkel und Torsionswinkel erfassen [76–79]. Dabei wird die Bin-

^aBeta-Peptide unterscheiden sich von den natürlich vorkommenden Alpha-Peptiden durch den zusätzlichen Einbau einer Methylengruppe in das Rückgrat und eine damit verbundene deutlich höhere Flexibilität.

dungsstruktur als unveränderlich angenommen. Auf die detaillierte Form dieser kurzreichweitigen Modellpotentiale, werde ich Abschnitt 2.2 vertieft eingehen.

Behandlung der langreichweitigen Wechselwirkungen in MM-Kraftfeldern

Neben den chemischen (engl. bonded) Wechselwirkungen müssen in einem Kraftfeld auch langreichweitige Kräfte zwischen den Atomen berücksichtigt werden. Üblicherweise wird dabei die Pauli-Abstoßung der Elektronenhüllen und die attraktive van der Waals-Wechselwirkung (VDW) zusammen durch ein Lennard-Jones-Potential erfasst [74]. Die numerische Auswertung der VDW-Wechselwirkung ist unproblematisch, da sie sehr schnell (mit r^{-6}) abfällt, so dass nur Kräfte zwischen Atomen in einer räumlichen Nahzone berücksichtigt werden müssen. Weiterreichende Energieanteile können, wie in Abschnitt 2.2 genauer erklärt werden wird, im Rahmen einer Kontinuumsnäherung behandelt werden.

Weitaus schwieriger ist es, eine angemessene Beschreibung der elektrostatischen Wechselwirkungen im Simulationssystem zu erzeugen. In den verfügbaren nicht-polarisierbaren Kraftfeldern wird die Ladungsverteilung der untersuchten Moleküle durch statische Partialladungen beschrieben, die an den Orten der Atome liegen [81]. Die Annahme statischer Partialladungen stellt das größte Defizit gegenwärtiger MM-Kraftfelder dar, da durch sie die Effekte der elektronischen Polarisierbarkeit lediglich in einer Molekularfeldnäherung erfasst werden können. In inhomogenen und anisotropen Medien – Proteine und Peptide sind Prototypen solcher Materialien – ist diese Näherung jedoch nicht begründbar [92–94]. Es wurde gezeigt, dass die Vernachlässigung der Polarisationsfluktuationen bei der Beschreibung von Peptiden und Proteinen zu signifikanten Fehlern führt [94, 95].

Aber auch für nicht-polarisierbare Kraftfelder wirft die Auswertung der Coulomb-Wechselwirkung der Partialladungen ein Problem auf, denn die bekannte $1/r$ -Abhängigkeit des Coulomb-Potentials führt dazu, dass die Wechselwirkung jedes geladenen Atoms mit jeder anderen Ladung des notwendig endlich dimensionierten Simulationssystems auszuwerten ist. Der Aufwand zur Berechnung der Coulomb-Summe steigt somit quadratisch mit der Anzahl der Atome an.

Zur Lösung des Problems der Behandlung der langreichweitigen Wechselwirkungen existieren eine Reihe von Ansätzen, die für Simulationssysteme mit periodischen Randbedingungen (engl. periodic boundary conditions, PBC) formuliert wurden [94, 97–100]. Solche Randbedingungen sind notwendig, um für die Simulation von gelösten Proteinen oder Peptiden ein gut definiertes thermodynamisches Ensemble zu erzeugen [74, 83]. Bei Simulationssystemen mit PBCs wird das begrenzte Simulationsvolumen periodisch fortgesetzt. Auf diese Weise wird z.B. die Topologie eines Würfels zu der eines dreidimensionalen Toruses erweitert. Abbildung 1.10 zeigt jene drei möglichen raumfüllenden Simulationsvolumina, die von Gerald Mathias [96] in EGO-MMII implementiert wurden. Ein solches periodisches System besitzt keine Grenzflächen und weist deshalb kontrollierbare thermodynamische Parameter wie Temperatur, Druck und Volumen auf, die mit einem Thermostaten bzw. Barostaten geregelt werden können [101].

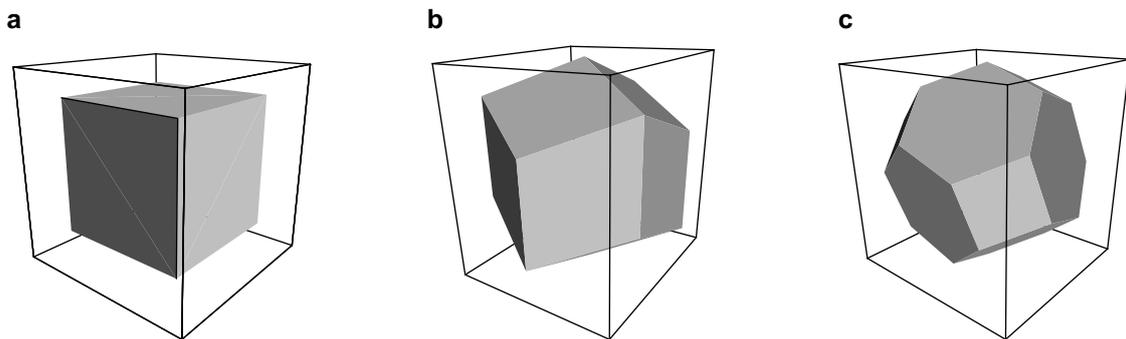


Abbildung 1.10: Raumfüllende Simulationsvolumina. In EGO-MMII wurden von Gerald Matthias [96] für eine Simulation unter periodischen Randbedingungen drei verschiedene Formen des Simulationsvolumens implementiert: **a** der Quader bzw. Würfel, **b** der Rhombendodekaeder und **c** der abgeschnittene Oktaeder. Mit jedem dieser Körper kann der dreidimensionale Raum lückenlos gefüllt werden. Während der Simulation können einzelne Lösungsmittelmoleküle das Simulationsvolumen durch eine der Seitenflächen verlassen und in eines der periodisch fortgesetzten Bilder des Systems übertreten. Der Vorteil der Körper **b** und **c** liegt darin, dass diese, verglichen mit dem Würfel, besser die Form einer Kugel annähern. Auf diese Weise müssen – bei einem gegebenen oder erwünschten Cutoff-Radius – weniger in den Ecken des Volumens liegende, und damit zu den Wechselwirkungen des gelösten Moleküls kaum beitragende, Lösungsmittelmoleküle behandelt werden. Beim Übergang vom würfelförmigen Volumen zu dem des Rhombendodekaeders entspricht dies einer Rechenzeiterparnis von etwa 30%.

Durch die Einführung einer Periodizität in das Simulationssystem wird zwar dessen Thermodynamik beherrschbar, jedoch tritt das Problem der artifiziellen Selbst-Wechselwirkung eines geladenen Teilchens mit seinen Spiegelbildern auf. Die so genannte *Minimum-Image Konvention* (MIC) [74] fordert deshalb, alle Wechselwirkungen zwischen Ladungen zu vernachlässigen, die weiter als die halbe Kantenlänge des Simulationssystems voneinander entfernt sind. Ein einfaches Abschneiden der Wechselwirkungen an dem von der MIC vorgegebenen Radius, d.h. die Anwendung eines so genannten *Cutoffs*, führt aber zu strukturellen und energetischen Artefakten [100, 102, 103]. Nichtsdestotrotz wurden bis vor einigen Jahren viele MD-Simulationen [85, 86, 104] unter der Verwendung eines einfachen Cutoffs durchgeführt, der jeweils sogar noch deutlich unter dem von der MIC vorgegebenen Größe lag, da eine vollständige Auswertung der Coulomb-Summe innerhalb des von der MIC angegebenen Radius bei größeren Systemen zu rechenaufwändig war.

Abbildung 1.11 veranschaulicht nicht nur den zweidimensionalen Fall von periodischen Randbedingungen, sondern skizziert auch zwei aktuelle Ansätze zur Berechnung der langreichweitigen Wechselwirkungen in einem solchen System. Der erste Ansatz, der in Teilbild a verdeutlicht wird, sind Gittermethoden [97–99], die auf der Methode der Ewald-Summation [105] beruhen und von einer Periodizität nicht nur der in den Simulationen beschriebenen Materie, sondern auch der des elektrostatischen Potentials ausgehen. Unter

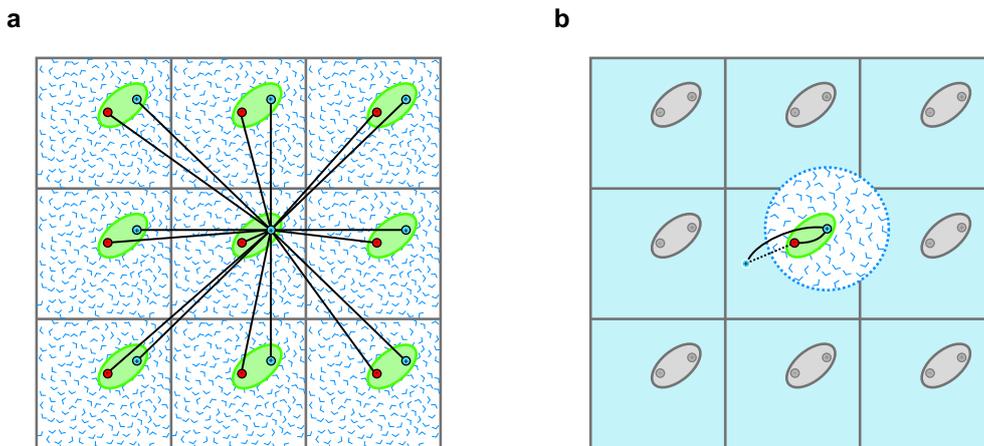


Abbildung 1.11: Behandlung der Elektrostatik unter periodischen Randbedingungen. In diesen zweidimensionalen Skizzen ist die zentrale Simulationsbox von acht periodisch fortgesetzten Kopien umgeben. **a** Die Ewald-Summationsmethode nimmt ein periodisches elektrostatisches Potential an. Dadurch wechselwirkt ein gelöstes Protein mit allen seinen Kopien. Dies ist am Beispiel der positiven Ladung dargestellt. **b** Durch die in dieser Arbeit verwendete strukturadaptierte Multipolmethode werden die Wechselwirkungen bis zu einem durch die Minimum-Image Konvention [74] gegebenen Abstand (der halben Boxlänge) explizit berechnet. Ein Reaktionsfeld (hellblau) beinhaltet die Beiträge eines dielektrischen Kontinuums für größere Abstände. Auf diese Weise erfährt das simulierte Protein keine Kraftbeiträge seiner periodischen Kopien. Abbildung nach einer Vorlage aus [100].

Verletzung der MIC wechselwirkt bei den Gittermethoden ein gewähltes Atom nicht nur mit einem anderen Atom, sondern auch mit unendlich vielen Kopien, deren Abstände in jeder Dimension um diskrete Vielfache der Boxgröße ansteigen. Bei Gittermethoden kann die Coulomb-Summe im Fourier-Raum mit hoher Genauigkeit und unter vertretbarem numerischen Aufwand berechnet werden. Zwar führt dabei die Missachtung der MIC zu artifiziellen langreichweitigen Korrelationen, welche die Struktur eines simulierten Peptids oder Proteins stark beeinflussen können [106], jedoch ist zu erwarten, dass diese Artefakte in großen Simulationssystemen mit großen Periodizitäten vernachlässigbar klein werden [100].

Das dieser Arbeit zu Grunde liegende und in unserer Gruppe entwickelte Programmpaket EGO-MMII [80] ermöglicht es, als alternativen Ansatz zur Behandlung der langreichweitigen Wechselwirkungen die Coulomb-Summe durch die Anwendung einer strukturadaptierten Multipolmethode [107, 108] bis zu dem von der MIC vorgegebenen Radius näherungsweise auszuwerten. Dazu nähert diese SAMM genannte Methode die Coulomb-Beiträge weiter entfernter Atomgruppen durch die führenden Momente ihrer Ladungsverteilung (Nettoladungen, Dipole und evtl. Quadrupole) und das Potential am Ort der betrachteten Ladung durch eine entsprechende Reihenentwicklung. Dieses Verfahren führt zu einem linearen Skalierungsverhalten mit der Anzahl der Ladungen, dessen Vorfaktor

noch durch einen Mehrschrittalgorithmus erniedrigt wird. Zusammengenommen spricht man dabei vom FAMUSAMM-Algorithmus [108]. Abbildung 1.11 b verdeutlicht, wie auf diese Weise die Beiträge zur elektrostatischen Wechselwirkung von allen Atomen im hell gezeichneten Umkreis um die betrachtete Ladung berücksichtigt werden können.

Mathias et al. [100] haben das FAMUSAMM-Verfahren um ein Kontinuumsmodell erweitert, welches die Beiträge des außerhalb des MIC-Radius liegenden Lösungsmittels in Form eines dielektrischen Kontinuums mit passender Dielektrizitätskonstante repräsentiert und so Cutoff-Artefakte unterdrückt. Der Einfluss dieses Kontinuums auf die explizit beschriebenen Ladungen kann anschaulich durch das Auftreten von Spiegelladungen oder -multipolen dargestellt und durch ein Reaktionsfeld (RF) quantitativ erfasst werden [96, 100, 109]. Mit Hilfe dieses Ansatzes und durch Vergleich mit Ewald-Simulationen konnten erstmals absolute Angaben zur Genauigkeit der Elektrostatikberechnung bei MD-Simulationen gemacht werden. So hat sich anhand von Simulationen an großen Wassersystemen gezeigt, dass die FAMUSAMM/RF- und Ewald-Methoden voneinander nur um wenige Promille abweichen.

Zusammenfassend ermöglicht EGO-MMII mit Hilfe des erweiterten FAMUSAMM-Algorithmus die genaue Beschreibung der gelösten Modellpeptide in ihrer natürlicheren Umgebung und unter Vermeidung jener Artefakte, die durch Selbstwechselwirkungen oder Abschneiden der langreichweitigen Elektrostatik entstehen können. Auf Details der Implementierung von FAMUSAMM in EGO-MMII werde ich in Kapitel 2 eingehen.

DFT/MM-Hybridrechnungen

Tabelle 1.2 führt neben dem oben vorgestellten akkuraten, aber in der Zahl der zu behandelten Atome beschränkten, DFT-Verfahren und den eben eingeführten klassischen MM-Modellen auch die Methode der *DFT/MM-Hybridrechnungen* [110, 111] auf. In der vorliegenden Arbeit dient diese Methode zur Ableitung eines Kraftfeldes für das Lösungsmittel Dimethylsulfoxid (DMSO). Dabei wird ein quantenmechanisch beschriebenes DMSO-Molekül in der Umgebung molekularmechanisch modellierter DMSO-Moleküle simuliert.

In einer DFT/MM-Hybridrechnung wird der Ansatz einer mehrskaligen Behandlung der Simulationssysteme fortgeführt, der schon für die Elektrostatik in den MM-Verfahren erläutert wurde. Die Verbindung eines DFT-Programms mit einer MM-Implementierung ermöglicht es, einen kleinen Teil der Atome eines Simulationssystems mit der DFT-Methode akkurat zu beschreiben, während die Umgebung dieses Teilsystems durch eine MM-Modellierung näherungsweise repräsentiert wird. Die Einwirkung des MM-Teils auf die quantenmechanisch beschriebenen Atome wird in einer solchen Hybridsimulation in Form des durch die Partialladungen im QM-Teil erzeugten elektrischen Potentials erfasst. Auf diese Weise können im DFT-Teil Polarisierungseffekte untersucht werden.

Algorithmen zur Kopplung des DFT- und des MM-Teils wurden in unserer Gruppe entwickelt [111], sind in EGO-MMII implementiert und wurden schon mehrfach erfolgreich zur Berechnung der Schwingungsspektren von Molekülen in Lösung eingesetzt [112–114].

1.4 Lichtschaltbare Modellpeptide als kombinierter Ansatz

Weiter oben wurde die Bedeutung zeitaufgelöster Untersuchungen an Biomolekülen für das Verständnis der Faltung und Funktionsweise von Proteinen betont. Viele Konzepte über die Vorgänge auf der Zeitskala von Pikosekunden bis Nanosekunden wurden bisher ausschließlich aus der Analyse molekularmechanischer Simulationen abgeleitet. Ein Vergleich mit experimentellen Beobachtungen konnte zumeist nicht durchgeführt werden, da experimentelle Methoden mit dieser Zeitauflösung bisher nicht verfügbar waren. Eaton et al. [20] formulierten daher den aus diesem Umstand erwachsenden Ansporn an experimentell tätige Wissenschaftler:

„A (...) motivation for developing fast kinetic methods has been to provide a much-needed reality check on computer simulations, which are flooding the protein-folding literature.“

In dieser Hinsicht vielversprechend waren und sind die schon im Vorangehenden beschriebenen optischen Pump-Probe-Techniken, die einen Einblick in molekulare Reaktionen im Zeitbereich von Femto- bis Nanosekunden ermöglichen. In Zusammenarbeit mehrerer Forschungsgruppen (Zinth, Moroder und Tavan) im SFB 533 wurde diese schnelle Spektroskopie mit NMR-Untersuchungen und der theoretischen Beschreibung durch MM-Simulationen kombiniert, um einen integrierten Ansatz zur Untersuchung der Peptiddynamik zu erhalten. Dieser Ansatz steht im Zentrum der vorliegenden Dissertation, in welcher die dazu entwickelten Simulationsmethoden und statistischen Analyseverfahren vorgestellt, sowie die dabei erzielten Simulationsergebnisse präsentiert werden sollen. Mit der angesprochenen Kombination von Methoden wurde angestrebt, nicht nur die Interpretation der experimentellen Beobachtungen zu erleichtern, sondern auch die Evaluation der Methode der MM-Simulationen und vor allem der MM-Kraftfelder voranzutreiben. Ein besonderes Augenmerk sollte dabei der Aussagefähigkeit der Simulationen in Bezug auf die Kinetik von Konformationsübergängen in Peptiden gelten.

Voraussetzung für diese Zusammenarbeit war das Design eines Systems, das mehreren Anforderungen gleichzeitig zu genügen hat. So musste es für die Pump-Probe-Experimente einen Schalter beinhalten, der nach optischer Anregung Konformationsübergänge oder Faltungsprozesse in einem Peptid initiiert. Zusätzlich war eine spektroskopische „Sonde“ notwendig, über die mit einem Abtastpuls Aussagen über die Kinetik des Biomoleküls getroffen werden können. Ferner sollte das System nicht zu komplex sein, um eindeutige Interpretationen zu ermöglichen, und auch nicht zu viele Atome umfassen, um die theoretische Beschreibung zu erleichtern.

Die letzten beiden Forderungen schlossen beispielsweise das früher untersuchte [115, 116] Protein Bakteriorhodopsin (BR) als Modellsystem aus. BR und der verwandte Sehfärbstoff Rhodopsin weisen zwar einen Photozyklus auf, in dem nach Lichtanregung die

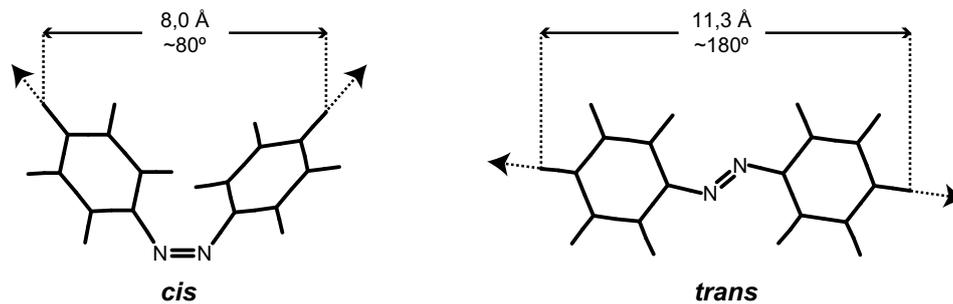


Abbildung 1.12: Die Geometrieänderung von Azobenzol durch Photoisomerisierung. *cis*- und *trans*-Isomere des Farbstoffs Azobenzol, die durch Photoanregung ineinander überführt werden können. Die Geometrien der beiden Isomere unterscheiden sich deutlich. Gemessen an den Positionen der beiden äußersten Wasserstoffe führt die gestreckte *trans*-Form zu einer Elongation um etwa 3 Å und besonders zu einer großen Änderung des „Öffnungswinkels“.

Isomerisierung eines Farbstoffs Konformationsänderungen des umgebenden Proteins bewirkt, sind für die hier angestrebte Behandlung aber zu komplex. Aus diesem Grund wurde in unserem Verbundprojekt aus dem SFB 533 ein mit acht Residuen recht kurzes Peptid ausgewählt, bei dem davon ausgegangen werden kann, dass sein Konformationsraum noch überschaubar ist. Dabei wurde die für das Peptid gewählte Aminosäuresequenz (Ala-Cys-Ala-Thr-Cys-Asp-Gly-Phe) der Bindungstasche der Thioredoxin Reduktase, des schon in Abbildung 1.6 vorgestellten Proteins, entnommen. Dort ist die aufgeführte Peptidsequenz in den Residuen 134 bis 141 wiederzufinden und beinhaltet die zwei für die Funktion des Enzyms unverzichtbaren Cystein-Residuen.

Als optischer Schalter wurde das *Azobenzol* gewählt. Dieser organische Farbstoff und aus ihm abgeleitete Verbindungen sind dafür bekannt, nach Photoanregung mit einer hohen Quantenausbeute eine *cis/trans*- oder – bei Anregung mit einer anderen Wellenlänge – eine *trans/cis*-Isomerisierung durchzuführen [117, 118]. Wie Abbildung 1.12 zeigt, führt die Photoisomerisierung zu einer deutlichen Änderung der Geometrie des Azobenzols. Von Vorteil ist auch, dass *cis*- und *trans*-Azobenzol stark unterschiedliche Absorptionsspektren aufweisen, die zudem empfindlich auf Verspannungen des Chromophors reagieren. Auf diese Weise erfüllt das Azobenzol nicht nur die Forderung nach einem schnellen Schalter, sondern auch die nach einer spektroskopisch auswertbaren Sonde. Eine Integration des Azobenzols in Peptide wurde durch die Synthese von zwei Derivaten des Farbstoffs ermöglicht, die als künstliche Aminosäuren in ein Peptidrückgrat integriert werden können [14, 119].

Die chemische Struktur dieser als APB und AMPB abgekürzten Azoschalter wird zusammen mit der des freien Azobenzols und der vollständigen zyklischen Modellpeptide in Abbildung 1.13 vorgestellt. Durch die direkte Integration des Lichtschalters in das Peptidrückgrat ist zu erwarten, dass sich dessen Geometrieänderung als eine Änderung der „Randbedingungen“ stark auf die möglichen Konformationen des Peptids auswirkt. Die

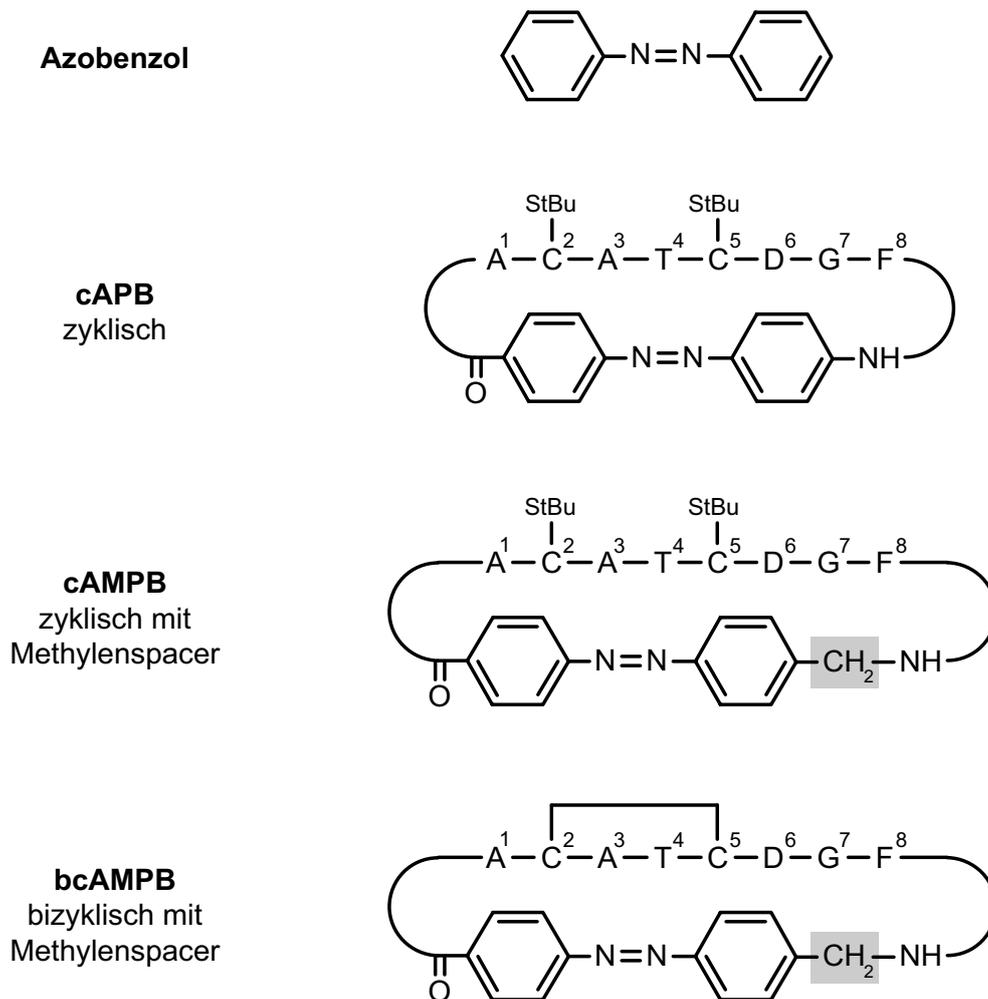


Abbildung 1.13: Die chemische Struktur von Azobenzol und den Modellpeptiden. Das freie Azobenzol umfasst eine Azo-Gruppe, d.h. zwei durch eine Doppelbindung verbundene Stickstoffe, sowie zwei Phenylringe. In obiger Darstellung sind die meisten Wasserstoffatome nicht gezeigt. Zu einem Farbstoff wird Azobenzol durch ein sich über das ganze Molekül ausdehnendes konjugiertes π -Elektronensystem. Wird Azobenzol um eine Carboxyl- und eine Aminogruppe erweitert, erhält man (4-Aminophenyl)azobenzoessäure (APB), die als künstliche Aminosäure in ein Peptid eingebunden werden kann. Durch Ringschluss entstehen die lichtschaltbaren Zylopeptide, deren zwei Cystein-Residuen eine Variation durch eine geschlossene oder offene Disulfidbrücke erlauben [37]. Zusätzlich wurde mit AMPB noch ein um eine Methylgruppe (grau hinterlegt) erweiterter Schalter synthetisiert, der sich durch eine größere Flexibilität auszeichnet [38].

Die in der Grafik aufgeführten Kurzbezeichnungen sind: Azo bzw. Azobenzol für den freien Farbstoff, c... für die monozyklischen Peptide, bc... für die bicyklischen Peptide, ...APB für Peptide mit dem einfachen Schalter und schließlich ...AMPB für Peptide mit Schalter und Methylenspacer.

lichtinduzierte Umkehrbarkeit des Vorgangs ermöglicht dabei ein unbeschränktes Hin- und Herschalten des Peptids und dadurch eine einfache experimentelle Handhabung. Zur Illustration der erzielbaren Konformationsänderungen zeigt Abbildung 1.14 beispielhaft je eine Konformation des *cis*- und *trans*-cAPBs. Die beiden Strukturen wurden den später präsentierten MD-Simulationen entnommen.

Die vorgestellten nützlichen Eigenschaften des Azobenzols als Lichtschalter sind schon lange von technischem Interesse. So wird dieser Farbstoff beispielsweise in der Forschung und Industrie verwendet, um optische Schalter [120–122] oder Datenspeicher [123–125] zu entwickeln. Auch selbstorganisierende [126] oder funktionalisierte [127, 128] Materialien wurden schon mit seiner Hilfe realisiert. Selbst die Anwendung von Azobenzol als Schalter für biologische Systeme wurde bereits diskutiert [129] und auch seine Einbindung in ein Peptidrückgrat [130] gezeigt. Ähnlich wie in dem hier vorgestellten Projekt arbeitet die Gruppe um Andrew Woolley [131, 132] mit einem α -helikalen Peptid, dessen Seitengruppen mit einem Azoschalter verbunden sind. Azobenzol wurde auch schon in eine DNA Dreifach-Helix integriert [133]. Schließlich untersuchten Hugel et al. [134] ein mehrere Azoschalter beinhaltendes Polymer mit Hilfe der Einzelmolekül-Kraftspektroskopie.

Experimentelle Charakterisierung der Modellpeptide

Erste fs-spektroskopische Messungen an Azobenzol, die im Rahmen des SFB 533 in den Arbeitsgruppen von Wolfgang Zinth (LMU München) und Josef Wachtveitl (jetzt Frankfurt) durchgeführt wurden [135], zeigten für den Zerfall des elektronisch angeregten Zustands nach $n\pi^*$ -Anregung^a eine aus zwei exponentiellen Termen zusammengesetzte Kinetik mit Zeitkonstanten von 170 fs und 2 ps. Anschließend befindet sich der Farbstoff in seinem – eventuell heißen – elektronischen Grundzustand, wobei der Großteil seiner mit der Photoisomerisierung verbundenen Geometrieänderung eingetreten ist. Relaxationszeiten um 15 ps wurden deshalb einem Abkühlen der vibronischen Anregungen durch Wärmedissipation in das umgebende Lösungsmittel zugeschrieben [136]. Entsprechende Messungen am APB, dem von Azobenzol abgeleiteten Schalter, ergaben vergleichbare Zeiten für diesen Prozess [137]. Selbst der Einbau des APBs in die Peptide beeinflusst die schnelle Dynamik des Schalters kaum, allerdings traten in Abhängigkeit von der Peptidstruktur (siehe Abbildung 1.13) zusätzliche Relaxationszeiten zwischen 30 ps und 1 ns auf [37, 138], welche deutliche Anzeichen für anhaltende Konformationsrelaxationen der geschalteten Peptide sind. Vor allem konnte somit gezeigt werden, dass der Chromophor sensitiv auf die Struktur des angeschlossenen Peptidrückgrats und deren Verspannung reagiert. Eine vertiefte Präsentation und Diskussion der spektroskopischen Ergebnisse findet sich in den zitierten Veröffentlichungen und besonders in der Dissertation von Sebastian Spörlein [139].

Wie schon die Synthese der Modellpeptide wurde auch ihre Strukturaufklärung in der Gruppe von Luis Moroder am MPI für Biochemie durchgeführt. Die bisher vorgestell-

^aÜbergang eines nichtbindenden Elektrons der Stickstoffatome in ein antibindendes Molekülorbital π^* .

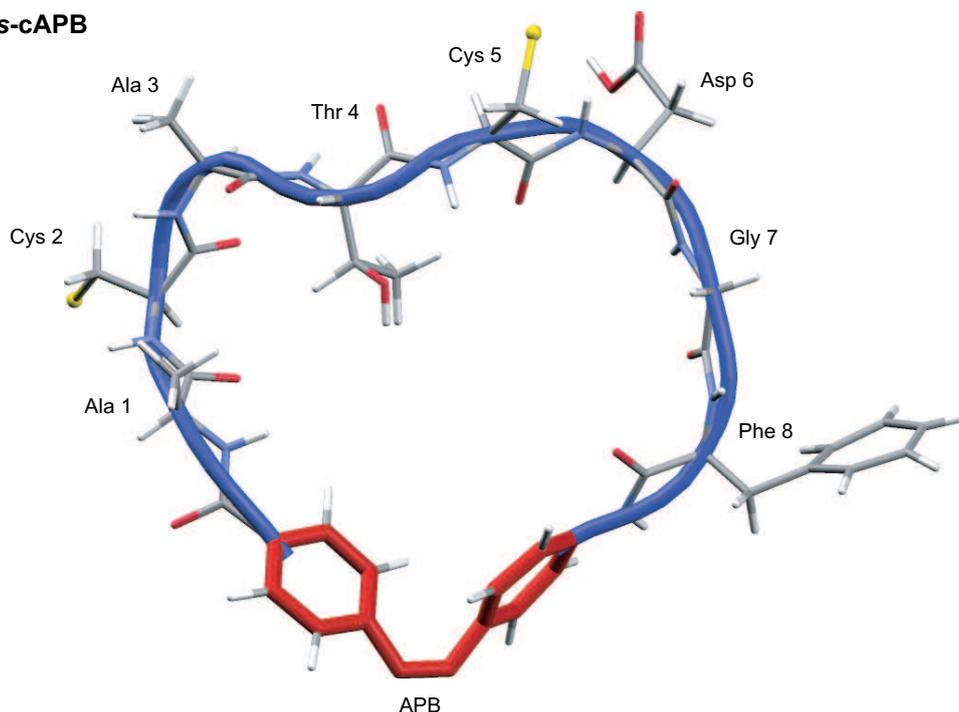
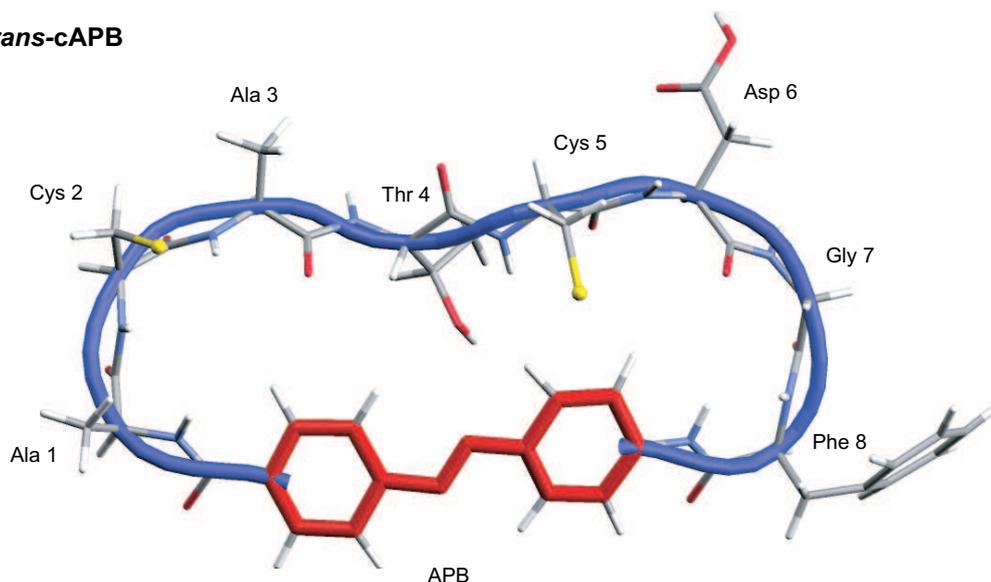
a *cis*-cAPBb *trans*-cAPB

Abbildung 1.14: Je eine Konformation des *cis*- und *trans*-cAPBs. Ergänzend zu Abbildung 1.13 wird für beide Isomerisierungszustände des Azobenzol-Farbstoffs exemplarisch jeweils eine der in den Simulationen beobachteten Konformationen des cAPBs gezeigt. Beide Strukturen werden als Stäbchenmodell dargestellt, bei dem Kohlenstoffe grau, Sauerstoffe rot, Stickstoffe blau und Wasserstoffe weiß eingefärbt sind. Die Schwefelatome der Cysteine und die an sie gebundenen Schutzgruppen wurden zu gelben Kügelchen vereinfacht. Dem Stäbchenmodell ist jeweils noch eine „Tube“-Darstellung überlagert, bei der die blau gefärbte Spline-Funktion die C_{α} -Atome der Residuen verbindet und damit den Verlauf des Peptidrückgrats hervorhebt.

te und in dieser Arbeit behandelte erste Generation von Peptiden ist nicht wasserlöslich, weswegen das Lösungsmittel Dimethylsulfoxid (DMSO) Verwendung fand.^a Die durch NMR-Messungen bestimmten Strukturen wurden von Renner et al. [37, 38] veröffentlicht. Sie zeigten in Abhängigkeit vom Isomerisierungszustand des Chromophores stark unterschiedliche Konformationsensembles der Modellpeptide. Besonders in der *trans*-Form schränken die Randbedingungen den Peptidteil stark ein, insbesondere wurde im Fall des cAPBs nur eine einzige Konfiguration gefunden. Auf diese Weise wurde die Vermutung belegt, dass mit der Photoisomerisierung des Azoschalters eine starke Konformationsänderung des Peptidrückgrats einhergeht.

1.5 Ziele und Gliederung der Arbeit

Wie wir gesehen haben, stellen die zyklischen Modellpeptide ein Modellsystem dar, an dem die Konformationsdynamik von Peptiden auf der Zeitskala von Femtosekunden bis Nanosekunden experimentellen Untersuchungen offen steht. Das Ziel dieser Arbeit ist nun eine komplementäre theoretische Beschreibung. Die folgenden zentralen Fragestellungen stehen dabei am Anfang:

- Kann die Struktur und die Relaxationsdynamik der lichtschtbaren Peptide mit molekularmechanischen Simulationen adäquat modelliert werden?
- Ist eine Vergleichbarkeit der Ergebnisse der Simulationen mit den Resultaten der NMR-Strukturanalyse und denen der fs-Spektroskopie gegeben?
- Ermöglicht die theoretische Beschreibung einen vertieften Einblick in die Konformationsräume der Peptide und in die mit der Relaxation verbundenen Konformationsübergänge?
- Welche Prozesse spielen sich nach Anregung des Chromophores auf den verschiedenen Zeitskalen ab?
- Können daraus allgemeine Schlussfolgerungen über die lichtinduzierte Relaxations- und Konformationsdynamik in Proteinen und Peptiden abgeleitet werden?

Eine erfolgreiche molekularmechanische Simulation der Modellpeptide hängt von einer Reihe von Voraussetzungen ab, die ich nach einer detaillierten Darstellung der Standardkraftfelder und Simulationsmethoden (Kapitel 2) in jeweils einzelnen Kapiteln behandeln werde.

Kapitel 3 beschäftigt sich mit dem DMSO als dem experimentell verwendeten Lösungsmittel. Eine möglichst realitätsgetreue Modellierung des DMSOs ist unerlässlich für die

^aIn der zweiten Generation der Modellpeptide wurde durch eine Änderung der Aminosäuresequenz die Wasserlöslichkeit der Modellpeptide erreicht [142].

Simulation der Relaxationsdynamik der Peptide, nicht nur weil der Dissipation von Wärme aus dem gelösten Molekül nach der Photoisomerisierung eine große Bedeutung zukommen sollte, sondern auch weil das moderat polare Lösungsmittel DMSO einen starken Einfluss auf die Struktur des Peptids ausüben kann. Neben einer Diskussion der bisher schon existierenden DMSO-Parametrisierungen wird im Rahmen dieser Arbeit ein neues Modell hauptsächlich aus „ersten Prinzipien“, d.h. quantenmechanischen Berechnungen, abgeleitet.

Das darauf folgende Kapitel 4 behandelt die Erstellung eines molekularmechanischen Kraftfeldes für die Azo-Schalter, die als in der Biologie unübliche Verbindungen nicht Teil der weit verbreiteten Kraftfelder für Peptide und Proteine sind. Die Bindungsparameter und Partialladungen der Chromophore und deren Schnittstellen zum Peptidrückgrat wurden aus DFT-Berechnungen bestimmt. Ferner wird gezeigt, wie ein mechanisches Modell der Photoisomerisierung des Azobenzols konstruiert und in die molekularmechanischen Simulationen integriert werden kann, um eine vollständige Beschreibung der Pump-Probe-Experimente zu erreichen.

Resultate von molekularmechanischen MD-Simulationen sind zunächst die *Trajektorien der Molekülbewegung*, d.h. die Menge aller Atomkoordinaten als Zeitreihe. Die dabei entstehenden Datenmengen können sehr umfangreich sein. So ergeben beispielsweise zehn Trajektorien eines Moleküls mit 100 Atomen, aufgezeichnet mit Zeitschritten von 100 fs über einen Zeitraum von 1 ns, einen Datensatz von $10 \cdot 3 \cdot 100 \cdot 10^4 = 3 \cdot 10^7$ Gleitkommazahlen oder etwa 110 MByte an Daten. Eine Analyse der in einem System ablaufenden Prozesse bedarf deshalb einer Datenreduktion, die auch auf die hohe Dimensionalität der Daten (3^N bei N Atomen) eingeht. Entsprechende Konzepte der Neuroinformatik wurden in den letzten Jahren in unserer Arbeitsgruppe entwickelt [144–147] und beispielsweise auf die Klassifikation von Sprachdaten angewandt [148–150]. Ein weiteres Ziel der vorliegenden Arbeit ist die Übertragung dieser statistischen Ansätze auf die Konformationsanalyse von Peptiden. Kapitel 5 beschreibt neuartige Methoden zur unüberwachten Klassifizierung von Molekülkonformationen sowie ihre Erweiterung im Fall von zirkulären Winkelvariablen.

Die darauf folgenden Kapitel 6, 7 und 8 widmen sich den durchgeführten MD-Simulationen und ihrer Auswertung. Zuerst werden in Kapitel 6 die eingesetzten Peptidkraftfelder untersucht. Dazu werden die Konformationen und die Energielandschaften eines einfachen Beispielmoleküls, des Alanin-Dipeptids, im Vakuum und in wässriger Lösung bestimmt und mit Literaturwerten verglichen.

An diese Verifikation der Methoden schließt sich in Kapitel 7 die Untersuchung der zyklischen Modellpeptide in ihren *cis*- und *trans*-Gleichgewichtsensembles an. Die Energielandschaften der verschiedenen Peptid-Spezies werden charakterisiert und ihre prototypischen Konformationen bestimmt. Ferner werden die in den Simulationen zu beobachtenden Peptidkonfigurationen detailliert mit den Ergebnissen der NMR-Strukturuntersuchungen verglichen.

In Kapitel 8 werden schließlich an den Beispielen des freien Azobenzols und des monozyklischen APB-Peptids (cAPB) die Auswirkungen der Wahl verschiedener Wärmeleitfähigkeiten des Lösungsmittels, verschieden steif parametrisierter Azobenzol-Farbstoffe und unterschiedlicher Peptidkraftfeldvarianten auf die simulierte Relaxationskinetik und die dabei ablaufenden Prozesse untersucht und mit den Beobachtungen aus der fs-Spektroskopie verglichen. Neben der Betrachtung von energetischen Observablen wird die strukturelle Konformationsdynamik des Peptidrückgrats analysiert.

2 Standardkraftfelder und Simulationsmethoden

Schon in der Einführung wurde die Methode der MD-Simulationen vorgestellt, mit der die lichtinduzierte Relaxationsdynamik der Modellpeptide untersucht werden soll. Das folgende Kapitel dient dazu, die eingesetzten Methoden detaillierter darzustellen. Eine ausführliche Formulierung der molekularmechanischen Kraftfelder ist notwendig, um später meine eigenen Modifikationen einordnen und notieren zu können. Ferner gehe ich hier vertieft darauf ein, wie diese Kraftfelder in unserem Simulationsprogramm EGO-MMII [80] implementiert sind und welche Weiterentwicklungen von meinen Kollegen und mir an diesem Programmpaket vorgenommen wurden.

2.1 Grundlagen der Molekulardynamik

Die Dynamik eines Ensembles von Molekülen wird durch die Schrödingergleichung

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, \mathbf{R}) = \hat{H}(\mathbf{r}, \mathbf{R}) \Psi(\mathbf{r}, \mathbf{R}) \quad (2.1)$$

beschrieben, wobei in der Notation der Wellenfunktion $\Psi(\mathbf{r}, \mathbf{R})$ die Elektronenspins der Einfachheit halber weggelassen wurden. Die Wellenfunktion hängt von den n Elektronenkoordinaten $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_n\}$ und den N Kernkoordinaten $\mathbf{R} = \{\mathbf{R}_1, \dots, \mathbf{R}_N\}$ ab. Der Hamiltonoperator $\hat{H}(\mathbf{r}, \mathbf{R})$ setzt sich, unter Verwendung atomarer Einheiten, d.h.

$$\hbar = e = m_e = (4\pi\epsilon_0)^{-1} = 1, \quad (2.2)$$

aus drei Termen zusammen

$$\hat{H}(\mathbf{r}, \mathbf{R}) = \hat{T}_K(\mathbf{R}) + \hat{T}_E(\mathbf{r}) + \hat{V}(\mathbf{r}, \mathbf{R}), \quad (2.3)$$

nämlich aus den Operatoren der kinetischen Energien der Kerne und der Elektronen,

$$\hat{T}_K(\mathbf{R}) = - \sum_{\mu=1}^N \frac{1}{2M_\mu} \nabla_\mu^2 \quad \text{und} \quad \hat{T}_E(\mathbf{r}) = - \frac{1}{2} \sum_{i=1}^n \nabla_i^2, \quad (2.4)$$

mit den Kernmassen M_μ , sowie dem Operator der potentiellen Energie

$$\hat{V}(\mathbf{r}, \mathbf{R}) = \hat{V}_{KK}(\mathbf{R}) + \hat{V}_{KE}(\mathbf{r}, \mathbf{R}) + \hat{V}_{EE}(\mathbf{r}),$$

der die Kern-Kern-, Kern-Elektron- und Elektron-Elektron-Wechselwirkungen beinhaltet und durch

$$\hat{V}(\mathbf{r}, \mathbf{R}) = \sum_{\mu < \nu} \frac{Z_\mu Z_\nu}{|\mathbf{R}_\mu - \mathbf{R}_\nu|} - \sum_{i, \mu} \frac{Z_\mu}{|\mathbf{r}_i - \mathbf{R}_\mu|} + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (2.5)$$

gegeben ist. Die Z_μ sind dabei die Kernladungszahlen.

In der bisherigen Betrachtung wurden die Atomkerne und Elektronen gleichwertig behandelt. Jedoch ist das Verhältnis der Elektronenmasse zur Kernmasse sehr klein (10^{-3} bis 10^{-5}), was dazu führt, dass die Kerne sich viel langsamer als die Elektronen bewegen und stärker lokalisiert sind. Aus diesem Grund kann mit der Born-Oppenheimer-Näherung [152, 153] angenommen werden, dass die Elektronen der Bewegung der Kerne unmittelbar folgen können. Als Konsequenz bewegen sich die Atomkerne in einem effektiven elektrostatischen Potential, das von den übrigen Kernen und allen Elektronen erzeugt wird [152].

Im stationären Fall ergibt sich dieses Potential aus der Lösung der elektronischen Schrödingergleichung gemäß

$$\hat{H}_E(\mathbf{r}|\mathbf{R}) \Psi_E(\mathbf{r}|\mathbf{R}) = E_E(\mathbf{R}) \Psi_E(\mathbf{r}|\mathbf{R}), \quad (2.6)$$

wobei $\Psi_E(\mathbf{r}|\mathbf{R})$ die elektronische Wellenfunktion zur vorgegebenen Kernkonfiguration \mathbf{R} und

$$\hat{H}_E(\mathbf{r}|\mathbf{R}) = \hat{T}_E(\mathbf{r}) + \hat{V}(\mathbf{r}|\mathbf{R}) \quad (2.7)$$

der elektronischer Hamiltonoperator ist, der üblicherweise auch die Kern-Kern-Wechselwirkung $\hat{V}_{KK}(\mathbf{R})$ als rein additiven Term beinhaltet. Für gebundene Zustände sind die Lösungen der elektronischen SG (2.6) quantisiert

$$\Psi_E(\mathbf{r}|\mathbf{R}) = \Psi_E^k(\mathbf{r}|\mathbf{R}), \quad k = 0, 1, 2, \dots, \quad (2.8)$$

mit der Quantenzahl $k = 0$ für den elektronischen Grundzustand und den Quantenzahlen $k = 1, 2, \dots$ für die elektronisch angeregten Zustände. Die zugehörigen Eigenwerte $E_E^k(\mathbf{R})$ definieren separat für jeden elektronischen Zustand das effektive Potential der Kernbewegung.

Statt nun im Rahmen der Born-Oppenheimer-Näherung für die Kernbewegung die Schrödingergleichung

$$i\hbar \frac{\partial}{\partial t} \Psi_K^k(\mathbf{r}, \mathbf{R}) = \left[\hat{T}_K(\mathbf{R}) + E_E^k(\mathbf{R}) \right] \Psi_K^k(\mathbf{r}, \mathbf{R}) \quad (2.9)$$

zu lösen, kann in vielen Anwendungen zum klassischen Limes, d.h. zu den Newtonschen Bewegungsgleichungen der Kerne

$$M_\mu \frac{d}{dt^2} \mathbf{R}_\mu = -\nabla_\mu E_E^k(\mathbf{R}) \quad (2.10)$$

mit Massen M_μ übergegangen werden. Um die Gültigkeit, aber auch die Grenzen dieser Näherung abzuschätzen, kann man die thermische de Broglie-Wellenlänge [154]

$$\Lambda = \frac{h}{\sqrt{2\pi m k_B T}} \quad (2.11)$$

eines Atoms der Masse m und Temperatur T betrachten. Bei $T = 300$ K erhält man für ein Wasserstoffatom etwa $\Lambda = 1,0$ Å, während die Wellenlänge eines schwereren Kohlenstoffatoms etwa $0,3$ Å beträgt. Vergleicht man diese Längen mit einer charakteristischen Größe des Systems, beispielsweise dem Abstand zweier Atome in einem Molekül von etwa 1 bis 3 Å, so erkennt man, dass die Bewegung des Kohlenstoffatoms in guter Näherung klassisch beschrieben werden kann, beim leichteren Wasserstoff jedoch Quanteneffekte eine große Rolle spielen können.

Für Vielteilchensysteme sind die Newtonschen Bewegungsgleichungen (2.10), die ein System gekoppelter Differenzialgleichungen zweiter Ordnung darstellen, nicht analytisch lösbar. Bei MD-Simulationen wird deshalb auf eine numerische Integration in diskreten Zeitschritten zurückgegriffen. Das bei den MD-Simulationen der Modellpeptide eingesetzte Programmpaket EGO-MMII verwendet hierzu eine Mehrschrittvariante [155, 156] des Verlet-Algorithmus [157], dessen Integrationsschritt sich als

$$\mathbf{R}_\mu(t + \Delta t) = 2\mathbf{R}_\mu(t) - \mathbf{R}_\mu(t - \Delta t) - \frac{\Delta t^2}{M_\mu} \nabla_\mu E_E^k(\mathbf{R}(t)), \quad \mu = 1, \dots, N \quad (2.12)$$

schreiben lässt. Sowohl im Verlet-Algorithmus als auch bei dem hier verwendeten Mehrschrittintegrator wird die neue Position eines Atoms zum Zeitpunkt $t + \Delta t$ aus den Positionen zu den vergangenen zwei Zeitschritten und der auf das Atom wirkenden Gesamtkraft berechnet, ohne dass die Geschwindigkeit des betrachteten Atoms explizit auftritt. Δt ist der Zeitschritt pro Integrationsschritt, der so zu wählen ist, dass eine ausreichend glatte Abtastung der schnellsten Freiheitsgrade gewährleistet wird. Da die Frequenz einer Normalmode durch das Verhältnis der auftretenden Kräfte zu den bewegten Massen bestimmt ist, treten die kleinsten Periodendauern von etwa 10 fs bei den Bindungsstreckschwingungen der leichten Wasserstoffe auf.

An dieser Stelle macht sich die beschränkte Gültigkeit der klassischen Näherung der Kernbewegung bemerkbar, denn in einer quantenstatistischen Betrachtungsweise sind derart hochfrequente Moden bei Raumtemperatur überhaupt nicht angeregt. Im Sinne einer realistischen Beschreibung der Moleküldynamik sollten daher in einer MD-Simulation alle solche hochfrequenten Moden durch Zwangsbedingungen unterdrückt werden. Rein praktisch können Zwangsbedingungen allerdings nur für interne Koordinaten und damit für stark auf diese Koordinaten lokalisierte Normalmoden eingeführt werden. Vorteilhaft ist, dass die große Zahl von Wasserstoff-Streckschwingungen mit Wellenzahlen von etwa 3.000 cm^{-1} nur sehr schwach mit anderen Normalmoden mischen, deren Wellenzahlen meist unterhalb 2.000 cm^{-1} liegen. Aus diesem Grund kann durch ein Festhalten der Bindungslängen der Wasserstoffe der getrennte Unterraum dieser Schwingungen ausgeschaltet werden, ohne die anderen Normalmoden zu beeinflussen. In MD-Simulationen werden

dazu Algorithmen wie SHAKE [158] oder MSHAKE [159] eingesetzt. Um die übrig bleibenden Moden glatt abzutasten, wird in dieser Arbeit durchgängig ein Integrationszeitschritt von $\Delta t = 1$ fs verwendet.

2.2 Das eingesetzte MM-Kraftfeld

Nun möchte ich detailliert auf das in dieser Arbeit eingesetzte Kraftfeld eingehen und erläutern, welche für die Simulation der Modellpeptide relevanten Weiterentwicklungen dazu von meinen Kollegen und mir an unserem Programmpaket EGO-MMII vorgenommen wurden. Das von EGO-MMII verwendete Kraftfeld basiert auf dem von CHARMM22 [81] und denjenigen Änderungen und Erweiterungen, die daran von den Entwicklern des Programmpaketes XPLOR [160] vorgenommen wurden. Daher lehnt sich meine Darstellung teils an jene von MacKerell et al. [81], teils an jene von Brünger [160] an.

Jedes molekularmechanische Kraftfeld ist eine empirische und analytische Parametrisierung für das oben hergeleitete effektive Potential der Kernbewegung $E_E^0(\mathbf{R})$ im elektronischen Grundzustand. Dabei wird das Modellpotential $E_{MM}(\mathbf{R})$ stets in E_{bonded} , die lokalen Wechselwirkungen der Atome mit ihren Bindungsnachbarn, und in $E_{\text{nonbonded}}$, die langreichweitigen Energiebeiträge zerlegt

$$E_E^0(\mathbf{R}) \approx E_{MM}(\mathbf{R}) = E_{\text{bonded}}(\mathbf{R}) + E_{\text{nonbonded}}(\mathbf{R}). \quad (2.13)$$

Diese Unterteilung entspricht dem Konzept von lokalisierten chemischen Bindungen, die von den Valenzelektronen der Atome ausgebildet werden. Auch der Hybridisierung der Atome mit den daraus resultierenden Bindungswinkeln wird in E_{bonded} Rechnung getragen. E_{bonded} beinhaltet nach Definition den Großteil der Wechselwirkungen von Atomen, die durch drei oder weniger kovalente Bindungen miteinander verbunden sind. Um die räumliche Anordnung von vier solcher Atome darzustellen, ist ein Übergang von den kartesischen Koordinaten der Atomkerne \mathbf{R} in einen Satz von so genannten *internen Koordinaten* hilfreich. Abbildung 2.1 zeigt Beispiele solcher interner Koordinaten: für die Bindungslängen $\mathbf{b} = (b_1, \dots, b_u)$ zwischen je zwei gebundenen Atomen, die Winkel $\boldsymbol{\theta} = (\theta_1, \dots, \theta_v)$ zwischen zwei solcher Bindungen und die schon erwähnten Dihedralwinkel $\boldsymbol{\phi} = (\phi_1, \dots, \phi_w)$. Die Hintransformation in interne Koordinaten, $\mathbf{R} \rightarrow (\mathbf{b}, \boldsymbol{\theta}, \boldsymbol{\phi})$, besteht in der Auswertung einfacher geometrischer Operationen. Dagegen ist die Rücktransformation häufig nichttrivial, beispielsweise wenn der Satz von internen Koordinaten nicht vollständig ist. Ein Beispiel hierfür ist die später verwendete Darstellung der Konfiguration eines Peptids durch die ϕ/ψ -Winkel der Aminosäuren. In anderen Fällen, wie bei der kartesischen Darstellung von Kräften, die mit Hilfe interner Koordinaten berechnet wurden, kann sich die Berechnung der Jacobi-Matrix der Koordinatentransformation als kompliziert erweisen.

Das CHARMM22-Kraftfeld und aus ihm abgeleitete Varianten vernachlässigen mögliche Kopplungen von internen Freiheitsgraden. Durch diese Annahme kann E_{bonded} als Summe

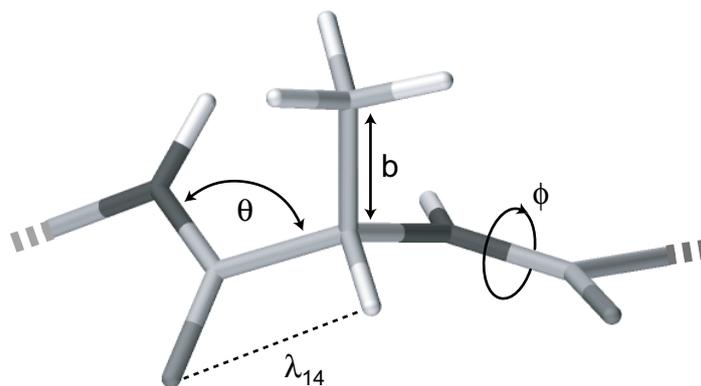


Abbildung 2.1: Skizze der internen Koordinaten. Ausschnitt einer Peptidkette mit einem Alanin-Residuum, in dem Kohlenstoffatome grau, Stickstoffe schwarz, Sauerstoffe dunkelgrau und Wasserstoffe weiß eingefärbt wurden. Als Beispiele für die internen Koordinaten sind eine Bindungslänge b , ein Bindungswinkel θ und ein Torsionswinkel ϕ eingezeichnet. Die gestrichelte Linie verbindet zwei Atome, zwischen denen die um den Faktor λ_{14} reduzierte elektrostatische Wechselwirkung berücksichtigt wird.

von Beiträgen geschrieben werden

$$E_{\text{bonded}}(\mathbf{b}, \boldsymbol{\theta}, \phi) = E_{\text{stretch}}(\mathbf{b}) + E_{\text{angle}}(\boldsymbol{\theta}) + E_{\text{dihedral}}(\phi), \quad (2.14)$$

die den verschiedenen internen Freiheitsgraden zugeordnet sind. Zur Vereinfachung habe ich an dieser Stelle die eng mit den Torsionbeiträgen verwandten so genannten *Improper*-Terme vernachlässigt, die von quadratischer Form sind und für die Erzwingung planarer Konfigurationen Verwendung finden. Dieser Effekt kann aber zumeist auch durch entsprechende Dihedralbeiträge erreicht werden.

Unter der Annahme kleiner Auslenkungen werden die Potentiale der kovalenten Bindungen zwischen zwei Atomen i und j als

$$E_{\text{stretch}}(\mathbf{b}) = \sum_{\text{bonds}} k_b (b - b_0)^2 \quad (2.15)$$

parametrisiert. Dabei ist b der aktuelle Abstand der zwei Atome, b_0 die Länge der Bindung in Ruhelage und k_b die harmonische Kraftkonstante. Die hier angegebene Summe soll über alle in der *Topologie* des Moleküls, d.h. der vorgegebenen chemischen Bindungsstruktur, festgelegten Bindungen laufen. Die Angabe von Indizes für b , b_0 und k_b und auch für die weiteren nun folgenden Kraftfeldparameter wird unterlassen, um eine unnötig komplexe Darstellung zu vermeiden.

Analog zu den Bindungslängen werden auch die Winkelpotentiale mit den Gleichgewichtswinkeln θ_0 und den Kraftkonstanten k_θ harmonisch genähert

$$E_{\text{angle}}(\boldsymbol{\theta}) = \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2. \quad (2.16)$$

Die für Bindungen und Winkel gültige Annahme kleiner Auslenkungen gilt nicht für die Torsionswinkel. Unter Ausnutzung der Periodizität eines solchen Winkels wird sein zugehöriges Potential durch eine Kosinus-Entwicklung

$$E_{\text{dihedral}}(\phi) = \sum_{\text{dihedrals}} \sum_n k_{\phi,n} [1 + \cos(n\phi + \delta_n)] \quad (2.17)$$

dargestellt. Ein Dihedral-Parametersatz besteht somit aus einer oder mehreren Periodizitäten n , für die jeweils eine Kraftkonstante $k_{\phi,n}$ und eine Phasenverschiebung δ_n angegeben wird. Für einen einfachen, d.h. nur mit einem Kosinus-Term beschriebenen, Dihedralwinkel ist demzufolge die Höhe der Rotationsbarriere durch $2 \cdot k_{\phi,n}$ gegeben.

Die während einer MM-Simulation einmal pro Zeitschritt notwendige Auswertung der bis hier aufgeführten *bonded* Energiebeiträge ist vom Rechenaufwand her unproblematisch, da die Anzahl der zu berechnenden Wechselwirkungen nur linear mit der Zahl der betrachteten Atome ansteigt. Wie in der Einführung (Abschnitt 1.3) ausführlich diskutiert wurde, gilt dies nicht für die langreichweitigen Beiträge zum Kraftfeld

$$E_{\text{nonbonded}}(\mathbf{R}) = E_{\text{VDW}}(\mathbf{R}) + E_{\text{elec}}(\mathbf{R}), \quad (2.18)$$

die sich aus Van der Waals (VDW) und elektrostatischen Beiträgen zusammensetzen. Die VDW-Wechselwirkung zwischen zwei Atomen i und j ist eine Funktion ihres Abstandes

$$r_{ij} = |\mathbf{R}_i - \mathbf{R}_j| \quad (2.19)$$

und hat die übliche Form eines 12-6 Lennard-Jones (LJ)-Potentials

$$E_{\text{VDW}}(\mathbf{R}) = \sum_{i < j} 4 \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \cdot sw_{\text{VDW}}(i, j). \quad (2.20)$$

Der r^{-12} -abhängige Teil der Funktion modelliert die Pauli-Abstoßung der sich überlappenden Elektronenverteilungen zweier chemisch nicht gebundener Atome, die sich bei nahem Kontakt fast wie harte Kugeln verhalten. Der mit einem r^{-6} -Verhalten langreichweitigere Anteil entspricht der attraktiven Dispersionswechselwirkung, die in erster Näherung durch die Wechselwirkung von durch Fluktuationen der Elektronendichte induzierten atomaren Dipolen erzeugt wird. Für ein Atompaar ist σ_{ij} der Radius der Paarwechselwirkung und ε_{ij} die Energie am Minimum der LJ-Funktion. Diese Parameter werden aus den atomaren Parametern der einzelnen Atome als arithmetisches bzw. geometrisches Mittel

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \quad \varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad (2.21)$$

berechnet. In Kraftfeldern wie CHARMM werden diese atomaren Parameter, genau wie die in (2.15-2.17) auftretenden Kraftkonstanten, jeweils für die so genannten *Atomtypen* angegeben, die anhand einer Klassifikation kovalenter Bindungsmuster definiert sind. Jedes in einem Molekül auftretende Atom wird nach seiner chemischen Umgebung einem

der etwa 50 Atomtypen zugeordnet und seine chemischen Bindungen durch den entsprechenden Satz von Parametern beschrieben. Die Summe in Gleichung (2.20) läuft über alle Atumpaare des Simulationssystems und umfasst damit $N^2/2$ Summanden bei N Atomen. Die Schaltfunktion

$$sw_{\text{VDW}}(i, j) = \begin{cases} 0 & \text{falls Atom } i \text{ von Atom } j \text{ weniger als drei Bindungen entfernt ist,} \\ 1 & \text{sonst.} \end{cases} \quad (2.22)$$

schaltet die VDW-Wechselwirkung für solche Atome ab, die über einen Winkel bzw. weniger als drei Bindungen miteinander verbunden sind. Dem liegt die Annahme zu Grunde, dass während der Parametrisierung diese nun abgeschalteten Anteile der VDW-Wechselwirkung schon in den Bindungs- oder Winkelkräften berücksichtigt wurden. Genaugenommen ist es nur schwer möglich, diese Beiträge beispielsweise in QM-Rechnungen voneinander zu trennen. Das CHARMM-Kraftfeld beinhaltet darüberhinaus noch einen eigenen Satz von VDW-Parametern, der Verwendung findet, sobald zwei Atome durch genau drei Bindungen getrennt sind. Diese so genannten 1-4 VDW-Parameter unterscheiden sich aber nur in einigen wenigen Fällen von den sonst verwendeten.

Im Fall der mit r^{-6} schnell abfallenden VDW-Wechselwirkung kann das quadratische Anwachsen des Rechenaufwands durch die Einführung eines Cutoffs verhindert werden, ohne die Genauigkeit der Berechnung von E_{VDW} zu beeinträchtigen. Die VDW-Wechselwirkung wird dazu gemäß

$$E_{\text{VDW}}(\mathbf{R}) = E_{\text{VDW,cut}}(\mathbf{R}) + E_{\text{VDW,corr}}(V) \quad (2.23)$$

aufgeteilt in einen direkt zu berechnenden Anteil $E_{\text{VDW,cut}}$ für Atomabstände kleiner r_{cut} und für größere Abstände in einen Korrekturterm $E_{\text{VDW,corr}}$ [74]. Während der Durchführung der ersten Simulationen an Systemen des reinen Lösungsmittels DMSO (siehe Kapitel 3) wurden Abweichungen der berechneten Enthalpien und Dichten von den Literaturwerten offenbar. Es stellte sich heraus, dass dies auf die fehlende VDW-Korrektur in EGO-MMII zurückzuführen war. In Anhang B beschreibe ich die von mir geschaffene Implementation einer VDW-Korrektur in EGO-MMII, welche die Besonderheiten des erweiterten FAMUSAMM-Algorithmus zur Berechnung der Elektrostatik berücksichtigt, die weiter unten erläutert werden.

In nicht-polarisierbaren MM-Kraftfeldern wird der elektrostatischen Wechselwirkung der Atome dadurch Rechnung getragen, dass ihnen integrale oder partielle Ladungen q_i zugewiesen werden. Diese Parameter sind unabhängig vom oben eingeführten Atomtyp und müssen somit für jedes Atom eines Moleküls spezifiziert werden. In atomaren Einheiten (2.2) geschrieben, wirkt zwischen zwei solchen Ladungen die Coulomb-Wechselwirkung

$$E_{\text{elec}}(\mathbf{R}) = \sum_{i < j} \frac{q_i q_j}{r_{ij}} \cdot sw_{\text{elec}}(i, j). \quad (2.24)$$

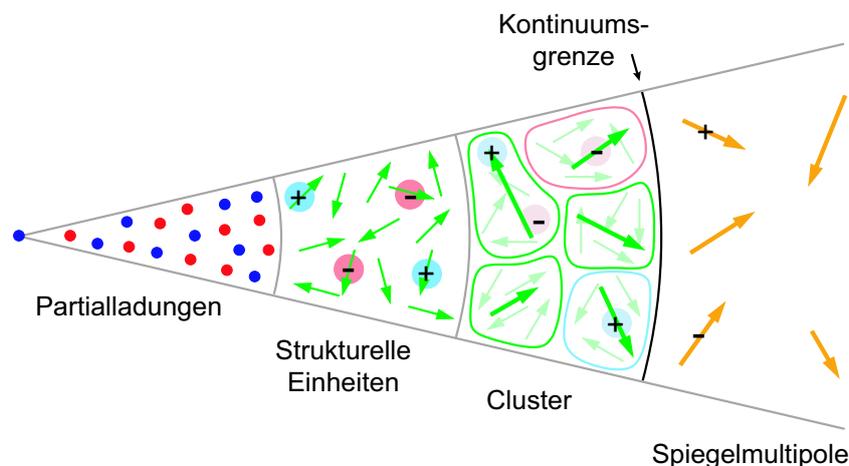


Abbildung 2.2: SAMM, die strukturadaptierte Multipolmethode. Ausgehend von einem betrachteten Atom werden die umgebenden Partiellladungen sukzessive zu Gruppen und Supergruppen zusammengefasst und deren Wechselwirkung durch eine Multipolentwicklung genähert. Ein Reaktionsfeld berücksichtigt zusätzlich die Beiträge des dielektrischen Kontinuums. Erklärung im Text.

Wie schon bei der VDW-Wechselwirkung werden auch bei der elektrostatischen Wechselwirkung jene Anteile in E_{bonded} mit aufgenommen, die zwischen benachbarten Atomen wirken. Um ein doppeltes Zählen zu vermeiden, wird auch in Gleichung (2.24) eine Schaltfunktion $sw(i, j)_{\text{elec}}$ eingeführt,

$$sw_{\text{elec}}(i, j) = \begin{cases} 0 & \text{falls Atom } i \text{ von Atom } j \text{ weniger als 3 Bindungen entfernt ist,} \\ \lambda_{14} & \text{falls Atom } i \text{ von Atom } j \text{ genau 3 Bindungen entfernt ist,} \\ 1 & \text{sonst.} \end{cases} \quad (2.25)$$

mit der die Elektrostatik für Atome abgeschaltet wird, zwischen denen eine oder zwei chemische Bindungen liegen. Wie Gleichung (2.25) aufführt, gilt ein Sonderfall für Atome, die genau drei Bindungen voneinander entfernt, und damit durch einen Torsionswinkel verbunden sind. Für diese wird E_{elec} mit einem Faktor λ_{14} skaliert, für den in verschiedenen Kraftfeldern unterschiedliche Werte vorgegeben sind. So wurde das Kraftfeld CHARMM19 unter der Annahme von $\lambda_{14} = 0,4$ parametrisiert [77], das von CHARMM22 dagegen mit $\lambda_{14} = 1,0$ [81].

2.3 Die strukturadaptierte Multipolmethode

In EGO-MMII wird zur näherungsweisen Berechnung der Coulomb-Summe (2.24) die schon in Abschnitt 1.3 skizzierte strukturadaptierte Multipolmethode SAMM [107, 161]

eingesetzt. Durch den SAMM-Algorithmus wird an der Position jedes Atoms i ein genähertes elektrostatisches Potential $\tilde{\Phi}_i$ bestimmt und daraus die elektrostatische Energie einer Partialladung E_{elec} berechnet. Die elektrostatische Gesamtenergie ist damit

$$E_{\text{elec,EGO}}(\mathbf{R}) = \sum_{i=1}^N q_i \tilde{\Phi}_i(\mathbf{R}). \quad (2.26)$$

Die Berechnung von $\tilde{\Phi}$ durch SAMM folgt dem in Abbildung 2.2 skizzierten mehrskaligen Schema. In einem Umkreis von etwa 10 \AA um die betrachtete Ladung wird deren Wechselwirkung mit den anderen Partialladungen unter Beachtung der Schaltfunktion (2.25) exakt berechnet. Weiter entfernte Atome werden zu so genannten *strukturellen Einheiten* zusammengefasst, deren Definition sich an der Topologie der Moleküle und den auftretenden Partialladungen orientiert, also strukturadaptiert ist. Dabei wird darauf geachtet, auf der einen Seite kompakte Einheiten zu bilden, auf der anderen Seite aber die vorliegenden Multipolmomente möglichst akkurat zu repräsentieren [107, 161, 162]. Beispielsweise sollten aus diesem Grund die Peptidgruppen, welche ein starkes Dipolmoment tragen, zu strukturellen Einheiten zusammengefasst werden. Die Beiträge jeder einzelnen strukturellen Einheit zu $\tilde{\Phi}$ wird anschließend durch eine Multipolentwicklung angenähert. EGO-MMII ermöglicht es, neben den Nettoladungen und den Dipolen der strukturellen Einheiten auch ihre Quadrupole miteinzubeziehen.

Abbildung 2.2 zeigt auch wie bei größeren Abständen von der betrachteten Punktladung die strukturellen Einheiten zu Clustern und diese sukzessive zu Superclustern zusammengefasst werden. Die in EGO-MMII verwendete Implementierung geht allerdings genau genommen den umgekehrten Weg, indem ausgehend von den größten Clustern Unterteilungen in Teilcluster vorgenommen werden. Diese Unterteilung wird durch den aus der Neuroinformatik entliehenen „Neural Gas“-Algorithmus [163] bewerkstelligt. Die Beiträge eines Clusters zum elektrostatischen Potential $\tilde{\Phi}$ werden analog zu jenen der strukturellen Einheiten (welche die Cluster auf Ebene 0 darstellen) durch ihre führenden Multipole genähert. Da, von Atom i aus betrachtet, weit entfernte Cluster zwar größer werden, die Entfernungsklassen aber auf eine Weise eingeteilt sind, dass diese Cluster unter einem konstanten Raumwinkel erscheinen, ist die Multipolnäherung stets gültig und führt zu einem Skalierungsverhalten des Rechenaufwandes mit $N \log(N)$. Eine weitere Reduktion des Aufwands ermöglicht die Ausnutzung der räumlichen Glattheit des von entfernten Clustern erzeugten Potentials. Vermittelt durch eine zweite Hierarchie und lokale Taylorentwicklungen kann nämlich das Potential $\tilde{\Phi}$ innerhalb eines Clusters an allen Orten von Subclustern oder Atomen effizient berechnet werden. Zusammen mit einem Mehrschrittverfahren zur Integration der Bewegungsgleichungen konnte ein insgesamt als FAMUSAMM bezeichneter Algorithmus formuliert werden [108, 164], der mit der Atomzahl N linear skaliert.

In der ursprünglichen Implementierung von FAMUSAMM wurde ein starkes algorithmisches Rauschen beobachtet, das eine Erwärmung des Simulationssystems zur Folge hatte. Das Rauschen konnte darauf zurückgeführt werden, dass in Dynamik-Simulationen

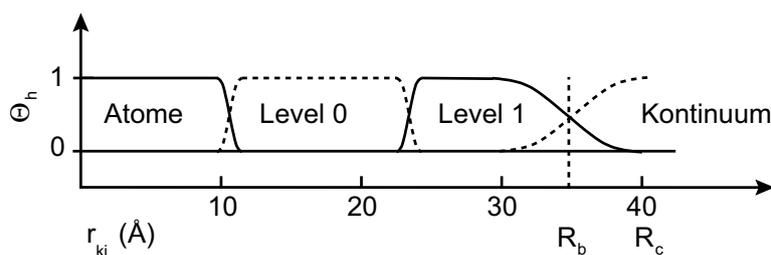


Abbildung 2.3: Der Überlapp der Hierarchieebenen in FAMUSAMM. Statt der eindeutigen Zuordnung einer Ladungsverteilung zu genau einer Entfernungsklasse, wie sie Abbildung 2.2 zeigt, wird in der weiterentwickelten Form von FAMUSAMM eine weiche Partitionierung durch die hier gezeigten Zuordnungsfunktionen Θ_h vorgenommen. Dies ermöglicht eine Interpolation der auf den beiden Darstellungstufen berechneten Beiträge der Ladungsverteilung zum elektrostatischen Potential und vermindert das algorithmische Rauschen. R_b bezeichnet den Radius der dielektrischen Grenzfläche und R_c den Cutoff-Radius an. Diese Abbildung wurde der Veröffentlichung von Mathias et al. [100] entnommen.

häufig Atomgruppen oder ganze Cluster aus einer Entfernungsklasse in eine höhere oder niedrigere Klasse übertreten, und damit Sprünge bei der Beschreibung ihrer Elektrostatik einhergehen. Abbildung 2.3 zeigt den von Gerald Mathias [96, 100] in der aktuellen Version von EGO-MMII eingeführten Überlapp der Entfernungsklassen und der damit verbundenen Beschreibungen. Interpoliert man die Beiträge eines Clusters zu $\tilde{\Phi}$ während seines Überganges zwischen zwei Entfernungsklassen, so kann eine deutliche Reduktion des Rauschens erreicht werden. Auch eine zweite nicht minder wichtige Erweiterung von FAMUSAMM wird in den Abbildungen 2.2 und 2.3 skizziert: die schon in der Einführung (Abschnitt 1.3) ausführlich diskutierte Einbeziehung eines Reaktionsfelds, welches für Abstände größer der äußersten Entfernungsklasse die Auswirkung eines dielektrischen Lösungsmittels auf das Potential $\tilde{\Phi}$ erfasst [100]. Erst durch diese Erweiterung ermöglicht FAMUSAMM eine vollständige und konsistente Beschreibung der Elektrostatik auf allen auftretenden Skalen, von den nahen, explizit berücksichtigten Partialladungen bis hin zum Lösungsmittelkontinuum.

Die Anwendung des FAMUSAMM-Verfahrens bei MD-Simulation der in DMSO gelösten Modellpeptide hat gewisse Abweichungen von den soeben geschilderten Vorgehen zur Folge. Dies liegt zum einen in der geringen Größe der Modellpeptide begründet, bei denen von einer maximalen räumlichen Ausdehnung von etwa 25 \AA ausgegangen werden kann (vgl. Abbildung 1.14). Um unter Berücksichtigung der Minimum-Image-Konvention keine intramolekularen Wechselwirkungen durch Anwendung des Cutoffs der langreichweitigen Elektrostatik zu vernachlässigen ist es notwendig, ein Simulationssystem zu verwenden, dessen Innenradius größer als die maximale Ausdehnung des gelösten Moleküls ist. Aus diesem Grund wurde in den meisten Simulationen der Modellpeptide ein Rhombendodekaeder als Simulationssystem gewählt, der einen Innenradius von nur etwas mehr als 27 \AA aufweist. Unter der Annahme der experimentellen Dichte von DMSO

wurde dieses relativ kleine (und damit in der Simulation mit einem geringen Rechenaufwand verbundene) Volumen mit etwa 1000 DMSO-Molekülen gefüllt. Nun ist ein DMSO-Molekül, verglichen mit einem Molekül des zumeist verwendeten Lösungsmittels Wasser, relativ groß. Sein Durchmesser, berechnet als der doppelte Gyrationradius, beträgt mehr als 3 Å, so dass in jeder Richtung des Simulationssystems stets weniger als 10 DMSO-Moleküle angeordnet sind.

Vergleicht man die soeben beschriebenen Abmessungen und Größenverhältnisse mit den üblichen radialen Ausdehnungen der FAMUSAMM-Hierarchieebenen (vgl. Abbildung 2.3), so wird deutlich, dass in den hier durchgeführten Simulationen der FAMUSAMM-Algorithmus nur bis zur Hierarchieebene der strukturellen Einheiten (Level 0), d.h. der einzelnen Lösungsmittelmoleküle angewendet werden kann. Auf Grund der kleinen Ausdehnung des hier untersuchten Simulationssystems und der geringen Anzahl von simulierten Atomen (etwa 10.000) hatte die Verwendung von nur zwei Hierarchieebenen im FAMUSAMM-Verfahren bei den vorgestellten Simulationen keine negativen Auswirkungen auf den Rechenaufwand.

Weitere Details der durchgeführten MD-Simulationen, wie die jeweils gewählten Parameter des FAMUSAMM-Algorithmus, finden sich in den Abschnitten 3.1.2, 6.2 und 7.1.

3 DMSO als explizites Lösungsmittel

Als erste Voraussetzung für eine erfolgreiche Simulation der Relaxationsdynamik der Modellpeptide wurde im Abschnitt 1.5 eine möglichst realitätsgetreue Modellierung des Lösungsmittels Dimethylsulfoxid (DMSO, C_2H_6SO , vgl. Abbildung 3.1) gefordert. Durch seine stark polare S=O Doppelbindung ist DMSO in der Lage, Wasserstoffbrücken auszubilden und die relativ große Dielektrizitätskonstante von 45,8. Es besitzt jedoch auch zwei apolare und damit hydrophobe Methylgruppen. Diese zwitterartige Struktur des DMSOs bedingt seine komplexen Wechselwirkungen mit Wasser, durch welche die physikalischen Eigenschaften von Mischungen dieser zwei Substanzen große Abweichungen vom Verhalten der beiden einzelnen Komponenten zeigen. So liegt der Gefrierpunkt einer Mischung von 1 mol DMSO auf 3 mol Wasser bei -70°C , ganz im Gegensatz zu reinem DMSO ($+19^\circ\text{C}$) oder reinem Wasser (0°C).

Zum Zeitpunkt des Beginns der vorliegenden Arbeit existierten eine ganze Reihe von MM-Modellen des DMSOs [165–168], die in MD-Simulationen eingesetzt wurden, um ein Verständnis für die Eigenschaften reinen DMSOs Rao und Singh [165], Luzar und Chandler [166], Vaisman und Berkowitz [169], Luzar et al. [170] und seiner Mischungen mit Wasser [171–174] zu erhalten. Auch wurden mit diesen Modellen MD-Simulation von in DMSO gelösten Zuckermolekülen [175], Salzen [176–178], Peptiden [179] und Proteinen [180]

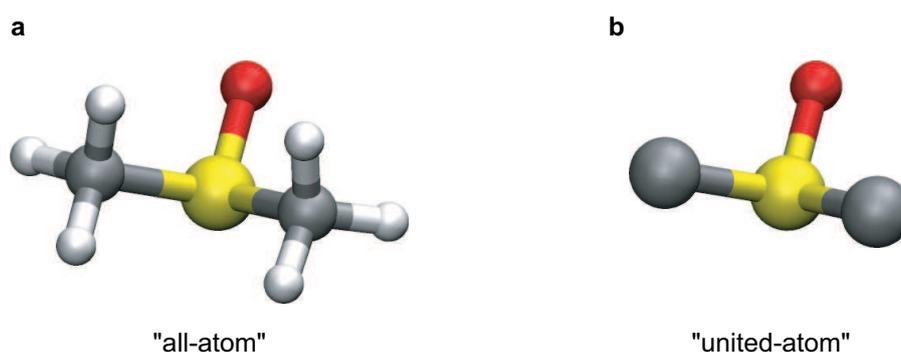


Abbildung 3.1: Dimethylsulfoxid. a Die Struktur des Dimethylsulfoxids (DMSO), optimiert mit der DFT-Methode B3LYP/6-31G*. Der Sauerstoff ist rot, der Schwefel gelb, die Kohlenstoffe grau und die Wasserstoffe sind weiß eingefärbt. In einem „all-atom“ Modell des DMSOs werden alle Atome explizit berücksichtigt. b In einem „united-atom“ Modell werden die Methylgruppen jeweils durch ein CH_3 -Pseudoatom repräsentiert, ein Vorgehen, welches zu einer deutlichen Rechenzeitersparnis in einer MD-Simulation führt.

durchgeführt. Die Wasserstoffatome des DMSOs wurden bei den diesen Modellen [165–168] aus Gründen der Rechenzeitersparnis nicht explizit berücksichtigt. Vielmehr wurden die Methylgruppen des DMSOs jeweils als Ganzes durch ein CH₃-Pseudoatom repräsentiert (vgl. Abbildung 3.1). Dieses mit „united-atom“ bezeichnete Vorgehen steht in Übereinstimmung mit älteren Protein-Kraftfeldern, wie dem von CHARMM19 [77].

Zu Beginn der vorliegenden Arbeit stellte sich die Frage, ob eine realistische Beschreibung der Modellpeptide unter Verwendung eines der „united-atom“ Modelle möglich ist oder ob nicht vielmehr eine explizite Berücksichtigung der Wasserstoffatome der DMSO-Moleküle und der mit ihnen verbundenen Freiheitsgrade notwendig ist, um die Wärmedissipation aus den gelösten Peptiden nach der Lichtanregung quantitativ richtig zu beschreiben. Es war auch zweifelhaft, ob die Zusammenfassung der acht Partialladungen der Kohlenstoff- und Wasserstoffatome eines DMSO-Moleküls zu den zwei Partialladungen der beiden CH₃-Pseudoatome eine realistische Darstellung der auftretenden Ladungsverteilung und insbesondere des Dipolmoments des DMSOs ergibt.

Aus diesen Gründen wurde die Schaffung eines expliziten (oder auch „all-atom“) Modells des DMSOs angestrebt. Im Gegensatz zum bisher üblichen Vorgehen zur Parametrisierung eines Lösungsmittelmodells sollten dabei möglichst wenige empirische Annahmen einfließen, um die Leistungsfähigkeit der theoretischen Beschreibung eines Moleküls aus ersten Prinzipien zu analysieren. Das entsprechende Vorgehen werde ich in Abschnitt 3.1 präsentieren. Anschließend wird das neue, im Folgenden mit „HC“ bezeichnete, DMSO-Modell nicht nur mit einigen ausgewählten „united-atom“-Parametrisierungen, sondern auch mit zwei weiteren „all-atom“ Modellen verglichen, die zeitgleich mit der vorliegenden Arbeit erstellt und veröffentlicht wurden [181, 182]. Bei diesen Vergleichen können die in der Literatur angegebenen Eigenschaften der bekannten Modelle nicht herangezogen werden, da zu erwarten ist, dass diese Eigenschaften durch die hier eingesetzte Methode zur Behandlung der langreichweitigen Wechselwirkungen (vgl. Abschnitt 2.3) stark modifiziert werden können. Vielmehr ist es angebracht, die Lösungsmitelegenschaften für alle Modelle unter den gleichen Voraussetzungen neu zu berechnen.

3.1 Parametrisierung des neuen DMSO-Modells

Der gesuchte neue Parametersatz für ein flexibles und explizites DMSO-Modell wurde durch drei aufeinander aufbauende Schritte erzeugt.

1. DFT-Berechnungen an einem einzelnen isolierten DMSO-Molekül im Vakuum dienten der Evaluierung und Optimierung des DFT-Zuganges durch Vergleich mit experimentellen Daten zu DMSO in der Gasphase. Vermittels der so bestimmten „optimalen“ DFT-Methode wurden die Geometrie, die Bindungsstärken und die Partialladungen eines vorläufigen Vakuum-Modells („VAK“) bestimmt.

2. Ein Simulationssystem aus VAK-DMSO wurde generiert und equilibriert. Durch den in der Einführung beschriebenen Ansatz der hybriden DFT/MM-Simulationen wurde anschließend eines der DMSO-Moleküle mit der in Schritt 1 bestimmten DFT-Methode in der Umgebung molekularmechanisch repräsentierter DMSO-Moleküle simuliert. Auf diese Weise konnten Polarisierungseffekte untersucht und die Ladungsverteilung des endgültigen „HC“-Modells festgelegt werden.
3. Im letzten Schritt wurden die VDW-Parameter des HC-Modells mit einem automatisierten Verfahren optimiert, um makroskopische Observable, wie die Dichte und die Verdampfungswärme, zu reproduzieren.

Auf die Details dieser Schritte möchte ich in den nächsten drei Unterabschnitten eingehen. Die auf diese Weise erzeugten Parametersätze der Modelle „VAK“ und „HC“ werden weiter unten in den Tabellen 3.6 und 3.7 zusammengefasst.

3.1.1 DFT-Berechnungen im Vakuum

Bisher wurden nur wenige quantenmechanische Berechnungen an DMSO veröffentlicht, darunter eine Arbeit von Müller-Plathe [183], in der die Aussagekraft verschiedener Austausch-Korrelationsfunktionale der DFT anhand von Strukturoptimierungen verglichen wurde. Eine weitere Arbeit beschäftigte sich mit Komplexen von DMSO an RU(II) [184] und kürzlich erschienen zwei Veröffentlichungen zu quantenmechanisch beschriebenen DMSO-Wasser-Mischungen [185, 186]. Dabei ist die letztgenannte Untersuchung von besonderem Interesse, da in ihr, wie auch in dieser Arbeit, das Programmpaket CPMD zur Berechnung herangezogen wurde. Um eine vertiefte Diskussion der Strukturberechnung von DMSO unter Verwendung unterschiedlicher DFT-Methoden zu ermöglichen, wurden in der vorliegenden Arbeit jedoch auch Vergleichsrechnungen mit GAUSSIAN durchgeführt.

Tabelle 3.1 zeigt die Ergebnisse der Geometrieoptimierung von DMSO im Vakuum, nämlich die Bindungslängen und -winkel zwischen den schweren Atomen und das Dipolmoment der Ladungsverteilung. Im Fall der Gaußschen Basisfunktionen wurde der 6-31G* Basissatz verwendet und fünf verschiedene Austausch-Korrelations-Funktionale miteinander verglichen.

Es zeigt sich, dass SVWN [190, 191] als Lokale-Dichte-Approximation die in der Gasphase gemessene Geometrie besser reproduziert als die gradientenkorrigierten Funktionale BP86 [64, 192] und BLYP [64, 65]. Letztere unterschätzen auch deutlich das Dipolmoment. Erkennbar ist das schon häufig beobachtete erfolgreiche Abschneiden des Drei-Parameter Hybrid-Funktional von Becke [66], das in Kombination mit den beiden Korrelations-Funktionalen von Perdew [192] (B3P86) bzw. Lee et al. [65] (B3LYP) gute Geometrien und auch eine genaue Schätzung des Dipolmoments ergibt.

| | Bindungslänge [Å] | | Winkel [°] | | Dipol [D] |
|---------------|-------------------|--------------|-------------|--------------|-------------|
| | S-O | C-S | C-S-C | C-S-O | |
| Exp. gas | 1,485 | 1,799 | 96,6 | 106,7 | 3,96 |
| Exp. liquid | 1,496 | 1,800 | 99,2 | 107,2 | - |
| GAUSSIAN [68] | | | | | |
| SVWN | 1,505 (1,3) | 1,812 (0,7) | 94,6 (-2,1) | 107,6 (0,8) | 3,79 (-4,3) |
| BP86 | 1,523 (2,6) | 1,855 (3,1) | 95,1 (-1,6) | 107,9 (1,1) | 3,71 (-6,3) |
| BLYP | 1,530 (3,0) | 1,874 (4,1) | 95,1 (-1,6) | 107,9 (1,1) | 3,67 (-7,3) |
| B3P86 | 1,505 (1,3) | 1,822 (1,3) | 95,7 (-0,9) | 107,5 (0,7) | 3,93 (-0,8) |
| B3LYP | 1,512 (1,8) | 1,838 (2,2) | 95,8 (-0,8) | 107,5 (0,7) | 3,93 (-0,8) |
| CPMD [69] | | | | | |
| LDA/MT | 1,564 (5,3) | 1,843 (2,4) | 95,4 (-1,2) | 105,3 (-1,3) | 4,32 (9,2) |
| BP/MT | 1,581 (6,5) | 1,884 (4,7) | 95,8 (-0,8) | 105,6 (-1,0) | 4,27 (7,8) |
| BLYP/MT | 1,591 (7,1) | 1,902 (5,7) | 95,8 (-0,8) | 105,7 (-0,9) | 4,32 (9,2) |
| LDA/SG | 1,479 (-0,4) | 1,784 (-0,8) | 95,8 (-0,8) | 106,7 (0,0) | 3,83 (-3,2) |
| BLYP/SG | 1,508 (1,5) | 1,842 (2,4) | 95,8 (-0,8) | 107,3 (0,6) | 3,81 (-3,8) |

Tabelle 3.1: Mit DFT optimierte Geometrien und Dipolmomente von DMSO *in vacuo*. Zum Vergleich werden zusätzlich die experimentell durch Mikrowellenspektroskopie in der Gasphase [187] bzw. durch Neutronenstreuung an flüssigem DMSO [188] bestimmten Geometrien aufgeführt. Das Dipolmoment kann dagegen nur in der Gasphase gemessen werden [189]. Die Strukturoptimierungen wurden unter Verwendung des 6-31G* Basissatzes bzw. eines Cutoffs der ebenen Wellen von 100 Ry und den angegebenen Austausch-Korrelationsfunktionalen durchgeführt. Im Fall der CPMD-Berechnungen ist zusätzlich die Art der eingesetzten Pseudopotentiale, Martins-Troullier (MT) [70] oder Goedecker (SG) [71], angegeben. Die Zahlen in Klammern geben die relative Abweichung der berechneten Größen von der in der Gasphase gemessenen Geometrie an.

Die beiden letztgenannten Funktionale stehen bedauerlicherweise in CPMD nicht zur Verfügung, da ihre Berechnung in der Basis der ebenen Wellen zu rechenaufwändig ist. Tabelle 3.1 präsentiert die ermittelten Geometrien für alle Kombinationen von Austausch-Korrelationsfunktionalen und Pseudopotentialen (PP), die in der Standardverteilung von CPMD für Schwefelatome verfügbar sind. Um eine Vergleichbarkeit bei konstantem Rechenaufwand zu erreichen, wurde in allen Rechnungen der Cutoff der Basis zu 100 Ry gewählt. Es zeigen sich deutliche Abweichungen bei Verwendung der PP von Martins und Troullier (MT) [70]. Dabei traten, unabhängig vom gewählten Funktional, zu lange Bindungen am Schwefelatom und ein zu großes Dipolmoment auf. Geeigneter erschienen die PP von Goedecker (SG) [71], die in beiden Fällen zu einer guten Geometrie und einem etwas unterschätzten Dipolmoment führten. Wie schon im Fall der GAUSSIAN-Rechnungen konnte dabei die Lokale-Dichte-Methode (LDA/SG) die Gasphasengeometrie sehr gut reproduzieren, wogegen BLYP/SG etwas zu lange Bindungen des Schwefelatoms ergab.

Bei der Auswahl des im Folgenden anzuwendenden CPMD-Funktionalen waren mehrere

Aspekte zu bedenken. Auf der einen Seite können im Fall der LDA üblicherweise kleinere Abstände zwischen den quantenmechanisch beschriebenen Atomen und dem Rand der von CPMD verwendeten QM-Box gewählt werden. Aus dem dadurch verminderten Volumen der QM-Box resultiert, verglichen mit BLYP, ein geringerer Rechenaufwand. Auf der anderen Seite sind gradientenkorigierte Funktionale, wie BLYP, dafür bekannt, bessere Energien und damit Bindungsstärken zu ergeben, d.h. die Form des von den Elektronen in ihrem Grundzustand auf die Atomkerne eines Moleküls ausgeübten Kraftfeldes $E_E(\mathbf{R})$ (siehe Abschnitt 2.1) genauer zu beschreiben. Eine sehr sensitive Vorgehensweise, um die Güte der Berechnung eines solchen Kraftfeldes zu beurteilen, ist die Betrachtung der durch dieses Kraftfeld bestimmten Schwingungsspektren eines Moleküls. Dabei ist nicht nur ein Vergleich bezüglich des eingesetzten Funktionals möglich, sondern auch ein weiterer Vergleich mit experimentellen Messungen in der Form von IR-Spektren.

Zur Berechnung eines Schwingungsspektrums beschreibt man $E_E(\mathbf{R})$ in der Nähe eines lokalen Minimums durch die harmonische Näherung

$$E_E(\mathbf{R}) \approx E_E(\mathbf{R}_0) + \frac{1}{2}(\mathbf{R} - \mathbf{R}_0)^T \mathbf{H} (\mathbf{R} - \mathbf{R}_0), \quad (3.1)$$

mit der Hessematrix \mathbf{H} der zweiten partiellen Ableitungen der Energie nach den Kernkoordinaten

$$\mathbf{H} = \left(\frac{\partial^2 E_E(\mathbf{R})}{\partial R_i \partial R_j} \right) \Big|_{\mathbf{R}=\mathbf{R}_0}. \quad (3.2)$$

Die Einträge von \mathbf{H} können dabei mittels der Bildung finiter Differenzen mit ausreichender Genauigkeit und vertretbarem Aufwand aus einer Reihe von Einzelpunktrechnungen bestimmt werden. Aus den Eigenwerten der mit den Atommassen gewichteten Hessematrix ergeben sich anschließend die Schwingungsfrequenzen der Normalmoden des Gesamtmoleküls. Hierbei verschwinden die ersten sechs Eigenwerte, da sie mit dessen Rotations- und Translationsfreiheitsgraden verbunden sind. Die zu den anderen $3N - 6$ Eigenwerten gehörigen Eigenvektoren beinhalten die Beiträge der einzelnen Atomkoordinaten zu der jeweiligen Normalmode und können als Auslenkung der Atome oder Verschiebungsvektoren visualisiert werden. Dies zeigt Abbildung 3.2 für die IR-aktiven Moden des DMSOs, die sich durch ein hohes Übergangsdipolmoment und damit große Absorptionsintensitäten auszeichnen.

Ergänzend dazu klassifiziert Tabelle 3.2 die Normalmoden nach den hauptsächlich auftretenden Deformationen und zeigt die unter Anwendung von LDA/SG bzw. BLYP/SG ermittelten Frequenzen und Intensitäten. Nimmt man an, dass die Kondensation des Lösungsmittels nicht zu einer sehr großen Änderung der Kraftkonstanten führt, so ermöglicht ein Vergleich mit den an flüssigem DMSO gemessenen IR-Absorptionsbanden eine Bewertung der beiden Funktionale.

Beginnt man die Diskussion bei kleinen Wellenzahlen, so zeigt sich, dass beide Methoden die niederenergetischen Verformungen der Bindungsgeometrie des Schwefels (Mode

3 DMSO als explizites Lösungsmittel

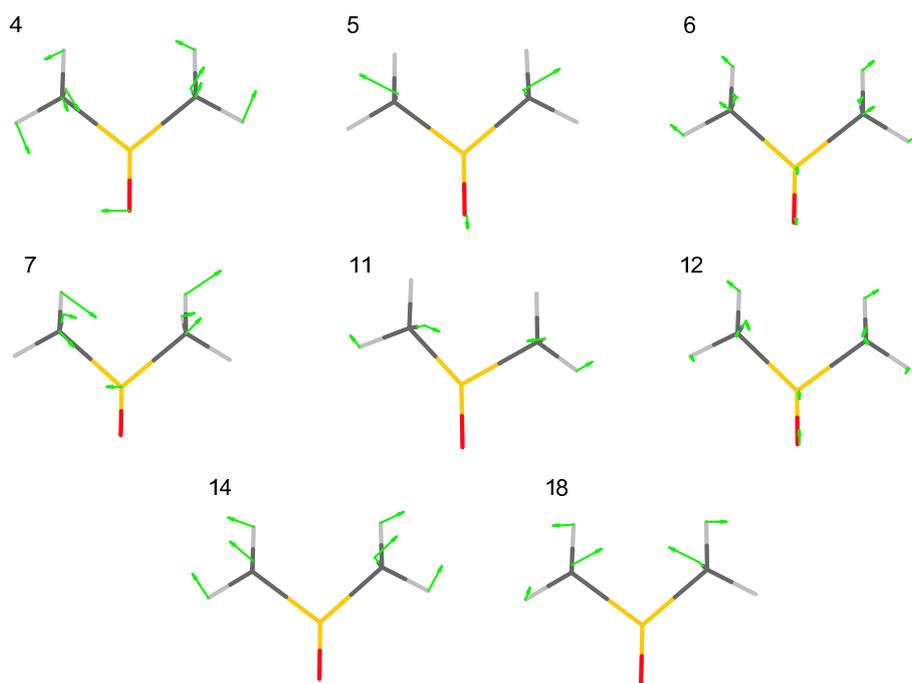


Abbildung 3.2: Die IR-aktiven Normalmoden des DMSO. Grüne Pfeile zeigen die mit der jeweiligen Normalmode verbundene Auslenkung der einzelnen Atome an, die Modennummern entsprechen denen von Tabelle 3.2.

| Nr. | Typ | Exp. k | BLYP/SG | | | LDA/SG | | |
|-----|--------------------------|-----------|---------|------|------|--------|------|------|
| | | | k | Int. | Abw. | k | Int. | Abw. |
| 4 | C-S-O Biegeschwingung | 332 | 334 | 7 | 2 | 318 | 8 | -14 |
| 5 | S pyramidiale Verformung | 381 | 360 | 7 | -21 | 367 | 5 | -14 |
| 6 | S-C gleichphasig | 668 | 586 | 10 | -82 | 657 | 9 | -11 |
| 7 | S-C gegenphasig | 698 | 605 | 18 | -93 | 677 | 16 | -21 |
| 11 | Schaukeln der Methylgr. | 954 | 990 | 12 | 36 | 975 | 13 | 21 |
| 12 | S-O Streckung | 1055 | 1061 | 128 | 6 | 1120 | 132 | 65 |
| 14 | Methylgr. symm. Deform. | 1311 | 1292 | 8 | -19 | 1247 | 28 | 64 |
| 18 | Methylgr. asymm. Deform. | 1437 | 1440 | 16 | 3 | 1385 | 24 | 52 |

Tabelle 3.2: Wellenzahlen der IR-aktiven Normalmoden des DMSOs. Unter Verwendung der beiden DFT-Funktionale BLYP bzw. LDA wurden die den Normalmoden zugeordneten Eigenfrequenzen aus der diagonalisierten Hessematrix bestimmt und in Wellenzahlen [cm^{-1}] umgerechnet. Die Amplituden wurden aus dem jeweiligen Übergangsdipolmoment ermittelt. Zum Vergleich wird die Lage der an flüssigem DMSO gemessenen Absorptionslinien angegeben, dabei wurden die Werte für kleine Wellenzahlen aus [193] und für große Wellenzahlen aus [194] entnommen. Auch die Abweichung der berechneten Frequenzen von den experimentellen Werten wird in Wellenzahlen angegeben.

4 und 5) relativ gut beschreiben. Auffällig ist für BLYP jedoch eine Unterschätzung der Stärke der S–C Bindung, die auch eine Entsprechung in einer zu großen Gleichgewichtslänge (Tabelle 3.1) findet. Auf der anderen Seite trifft BLYP sehr gut die Frequenz der sehr intensiven S–O Streckschwingung, die von LDA um 65 Wellenzahlen überschätzt wird. Es zeigt sich damit, dass durch den Vergleich der Spektren keine der beiden Methoden als untauglich ausgeschlossen werden kann. Weiterhin wird auch kein Skalierungsfaktor offenbar, der üblicherweise in vielen Untersuchungen pauschal angesetzt wird, um Über- oder Unterschätzungen in den Kraftkonstanten zu korrigieren. Unter Berücksichtigung der schon erwähnten Ersparnisse an Rechenzeit, die durch Verwendung des LDA-Funktional erwartet werden können, wurde deshalb den folgenden Rechnungen die LDA zu Grunde gelegt.

Parametrisierung des Vakuum-Modells

Die späteren Hybridrechnungen an einem in DMSO gelösten DMSO-Molekül bauen auf einem vorläufigen Lösungsmittelmodell auf. Da dieses Modell auf der Betrachtung eines einzelnen DMSO-Moleküls im Vakuum beruht, wird es im Folgenden als das Vakuum-Modell mit dem Akronym „VAK“ bezeichnet. In Tabelle 3.6 wird es zusammen mit dem HC-Modell aufgeführt, da beide Parametrisierungen die gleichen Bindungsparameter verwenden. Nach der im vorangehenden Abschnitt begründeten Wahl von LDA/SG als der zu verwendenden Kombination von Funktional und PP, wurde die Geometrie der beiden Modelle, d.h. deren Bindungslängen und -winkel, den entsprechenden Einträgen von Tabelle 3.1 entnommen.

Die Bestimmung von dem CHARMM-Kraftfeld angemessenen, lokalisierten Kraftkonstanten der Bindungen, Winkel und Torsionswinkel aus der DFT-Rechnung gestaltete sich allerdings etwas komplizierter und soll hier nur in der Form eines Überblicks dargestellt werden. Hierzu war es notwendig, die Einträge der kartesischen Hessematrix auf die entsprechenden, stark lokalisierten, internen Freiheitsgrade zu projizieren. Als Zwischenschritt bewährte sich die Verwendung der redundanten internen Koordinaten von Pulay et al. [195], die eine Basis aus kleinräumigen Bewegungen bilden. Beispiele solcher Koordinaten sind die symmetrischen oder die asymmetrischen Verformungen von Methylgruppen, die auch als Teile der Normalmoden des DMSO auftreten und in Abbildung 3.2 gezeigt werden.

Mit vorhandenen Hilfsprogrammen („hess2hess“, „ego2gva“ und schließlich „gva“) ist ein Übergang von der Hessematrix hin zu einer Darstellung der Kraftkonstanten in Pulay-Koordinaten möglich, aus welcher durch Mittelwertbildung bzw. die Berechnung von anderen Linearkombinationen die benötigten Kraftkonstanten zusammengestellt werden können. Diese werden ebenfalls in Tabelle 3.6 gezeigt. Anschließend wurde die Gültigkeit und Korrektheit dieses Verfahrens durch die Untersuchung seiner Selbstkonsistenz verifiziert. Dazu wurde aus dem auf diesem Wege parametrisierten molekülmechanischen DMSO-Modell wiederum die Hessematrix berechnet und das Projektionsverfahren auf deren Einträge angewendet. Es ergaben sich, mit nur kleinen Abweichungen, die in das MM-Modell

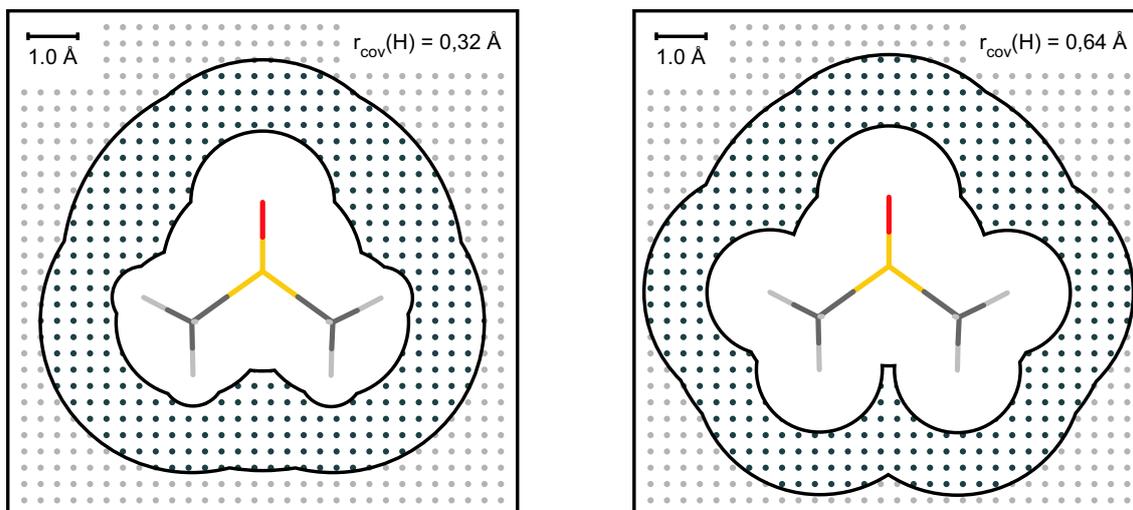


Abbildung 3.3: Die Wahl der Gitterpunkte zur Bestimmung der ESP-Ladungen des DMSOs. Schematische aber maßstabsgetreue Darstellung eines DMSO-Moleküls in Aufsicht. Weiß eingefärbt ist das Gebiet, das von mindestens einem Atomort weniger als $2r_{\text{cov}}$ entfernt ist. Die zweite schwarze Linie umfasst das Gebiet, in dem der Abstand höchstens $4r_{\text{cov}}$ beträgt. Zur Anpassung der ESP-Ladungen tragen alle dunkel hervorgehobenen Gitterpunkte bei. Die linke Darstellung geht von dem ursprünglichen Kovalenzradius der Wasserstoffe aus, in der rechten wurde dieser als doppelt so groß angenommen.

eingesetzten Kraftkonstanten, wogegen ein fehlerhaftes Vorgehen hier zu Diskrepanzen geführt hätte.

Um die elektrostatische Wechselwirkung eines DMSO-Moleküls mit seiner Umgebung zu beschreiben, wurde ein Satz von an den Atomorten lokalisierten Partialladungen mit Hilfe des Verfahrens der *ESP-Ladungen* [196] aus der quantenmechanisch berechneten Ladungsverteilung des Moleküls bestimmt. Dieses Verfahren geht vom gemeinsamen elektrostatischen Potential der Elektronendichte und der Kernladungen aus, das als Zwischenergebnis der DFT-Rechnung in dem durch Gitterpunkte diskretisierten Volumen der QM-Box vorliegt. Abbildung 3.3 zeigt anschaulich die QM-Box um ein zu berechnendes DMSO-Molekül und einige der Gitterpunkte im Ortsraum. Der ESP-Algorithmus berücksichtigt das quantenmechanisch berechnete Potential an denjenigen Gitterpunkten, die in einer gewissen Schale um das Gesamtmolekül liegen. Zur Bestimmung der zu berücksichtigenden Gitterpunkte wird jedem Atom nach seiner Ordnungszahl ein tabellierter Kovalenzradius r_{cov} zugewiesen und es werden alle jene Gitterpunkte ausgewählt, die von irgendeinem Atom weniger als vier solcher Radien entfernt liegen, aber zu keinem Atom einen geringeren Abstand als $2r_{\text{cov}}$ aufweisen. Das Optimierungsproblem besteht anschließend darin, an den ausgewählten Gitterpunkten die quadratischen Abweichungen des berechneten Potentials von demjenigen zu minimieren, das durch die fiktiven Punktladungen erzeugt werden würde. Diese Aufgabenstellung führt zu einem überbestimmten Gleichungssys-

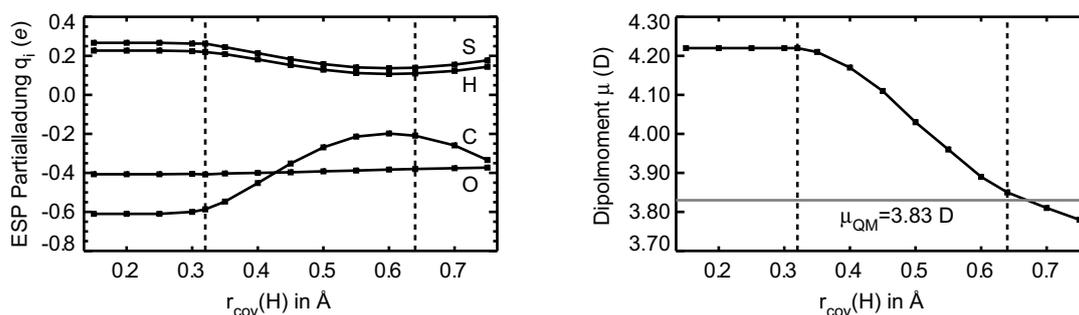


Abbildung 3.4: Die Abhängigkeit der ESP-Ladungen und des Dipolmoments von $r_{\text{cov}}(\text{H})$. Die linken Abbildung zeigt die über gleiche Atomtypen gemittelten Partiellladungen des DMSO in Einheiten der Elementarladung, die mit der Methode der ESP-Ladungen aus der Vakuum-Rechnung (LDA/SG) für verschiedene Kovalenzradien des Wasserstoffs $r_{\text{cov}}(\text{H})$ bestimmt wurden. Das rechte Diagramm gibt den Betrag des aus den einzelnen atomaren Partiellladungen und der Minimum-Geometrie berechneten Dipolmoments in Debye an.

tem, dessen Lösung durch die Bildung einer Pseudoinversen leicht bestimmt werden kann. Hierbei kann weiterhin durch das Hinzufügen einer zusätzlichen Gleichung die Ladungsneutralität des Gesamtmoleküls erzwungen werden. Eine detailliertere Formulierung des Vorgehens findet sich beispielsweise in der Diplomarbeit von Weiss [197]. Zur Generierung von Partiellladungen, die besonders geeignet sind, in einem MD-Modell das Dipolmoment eines DMSO-Moleküls als das führende Moment seiner Ladungsverteilung zu repräsentieren, wurde das beschriebene Verfahren basierend auf der in CPMD vorliegenden Implementierung in zwei Details abgeändert.

Die erste Änderung betrifft den Kovalenzradius der Wasserstoffatome $r_{\text{cov}}(\text{H})$. Tabelle 3.1 zeigt, dass für DMSO im Vakuum im Fall der Verwendung des LDA-Funktional aus der Verteilung der Kernladungen und der Elektronendichte ein Dipolmoment von $\mu_{\text{LDA}} = 3,83 \text{ D}$ berechnet wurde. Die mit dem ESP-Verfahren bei $r_{\text{cov}}(\text{H}) = 0,32 \text{ \AA}$ optimierten Partiellladungen ergaben jedoch, zusammen mit der Minimalgeometrie des Moleküls, ein Dipolmoment von $4,22 \text{ D}$ und überschätzen somit μ_{LDA} um etwa 10%. Als mögliche Begründung für diesen Effekt, wurde der sehr klein erscheinende Wert von $r_{\text{cov}}(\text{H})$ in Betracht gezogen. Das Resultat von Vergleichsrechnungen mit verschiedenen Werten von $r_{\text{cov}}(\text{H})$ zeigt Abbildung 3.4. Auffällig ist die Abhängigkeit der mittleren optimierten Partiellladungen der Wasserstoffe, Kohlenstoffe, des Schwefels und des zugehörigen MM-Dipolmoments von $r_{\text{cov}}(\text{H})$. Besonders die Ladung der Kohlenstoffe nimmt für wachsendes r_{cov} von $q_{\text{C}} = -0,587 e$ für $r_{\text{cov}}(\text{H}) = 0,32 \text{ \AA}$ ab bis hin zu einem deutlich geringeren Wert von $q_{\text{C}} = -0,209 e$ im Fall von $r_{\text{cov}}(\text{H}) = 0,64 \text{ \AA}$. Der Parameter $r_{\text{cov}}(\text{H})$ wurde daher so gewählt, dass das aus den ESP-Ladungen entstehende Dipolmoment den aus der Elektronendichte resultierenden Wert von $3,82 \text{ D}$ reproduziert.

Die angesprochene zweite Änderung des ESP-Verfahrens wurde durch die Einführung einer zweiten Nebenbedingung in das überbestimmte Gleichungssystem vorgenommen.

Neben der Ladungsneutralität des Gesamtmoleküls wurden zusätzlich gleiche Ladungen für alle sechs Wasserstoffatome eines DMSO-Moleküls gefordert. Diese Bedingung liegt darin begründet, dass in einem MM-Modell die drei Wasserstoffe einer Methylgruppe ununterscheidbar sind und deswegen mit identischen Ladungen beschrieben werden müssen. Der erste Ansatz, unterschiedliche ESP-Ladungen zu bestimmen und diese anschließend zu mitteln, führte erwartungsgemäß zu einer schlechten Darstellung des molekularen Dipolmoments durch das auf diese Weise generierte MM-Modell. Erst die erwähnte Nebenbedingung führte zu einer guten Übereinstimmung der Dipolmomente und damit zu sinnvollen Partialladungen.

Schließlich mussten als die letzten noch fehlenden Parameter des VAK-Modells, die VDW-Radien und -energien gewählt werden. Da dieses Modell nur eine erste Approximation darstellen sollte, wurden aus dem CHARMM22-Kraftfeld die entsprechenden Parameter der verwandten Atomtypen „O“, „S“, „CT3“ bzw. „HA“ übernommen. Damit wurde das VAK-Modell zusammengestellt, das in Tabelle 3.1 aufgeführt wird.

3.1.2 Hybride DFT/MM-Simulationen von DMSO in DMSO

Als zweiter Schritt zur Erzeugung eines Modells für DMSO in der kondensierten Phase wurden, aufbauend auf dem zuvor bestimmten vorläufigen Satz von Parametern, in einer (SG/LDA)/MM-Hybridsimulation die Auswirkung der elektrostatischen Umgebung auf das Dipolmoment und die Bindungsgeometrie eines solvatisierten, quantenmechanisch beschriebenen DMSO-Moleküls untersucht. Durch Polarisierungseffekte ist beim Übergang vom Vakuum in die Flüssigkeit eine Zunahme des Dipolmoments und eventuell eine Änderung seiner Richtung in Bezug auf die Geometrie des Moleküls zu erwarten. Experimentell konnte in der Gasphase ein Betrag von 3,96 D und ein Winkel zwischen dem Dipolmoment und der S=O-Bindung von $26,4^\circ$ bestimmt werden [189]. Entsprechende Messungen sind in kondensierter Phase jedoch nicht durchführbar. Eine Abschätzung der Polarisierungseffekte wurde zuvor mit DFT-Rechnungen unter Annahme eines dielektrischen Kontinuums und dessen Beschreibung durch ein selbstkonsistentes Reaktionsfeld getroffen. Dabei wurde eine Zunahme des Dipolmoments um 20% beobachtet [183].

Als Grundlage der (SG/LDA)/MM-Hybridrechnungen diente ein würfelförmiges Simulationenvolumen mit einer Kantenlänge von $58,1 \text{ \AA}$, das periodischen Randbedingungen unterworfen wurde und 1664 DMSO-Moleküle (und damit 16.640 Atome) umfasste, die klassisch mit Hilfe des VAK-Modells behandelt wurden. Die Größe des Systems wurde so gewählt, dass die experimentelle Dichte von flüssigem DMSO unter Normalbedingungen ($1,095 \text{ g/cm}^3$) herrschte. Zur Berechnung der langreichweitigen elektrostatischen Wechselwirkungen mittels des FAMUSAMM-Algorithmus' (siehe Abschnitt 2.2) wurden unter Berücksichtigung der relativ kleinen Ausmaße des Simulationssystems nur zwei Entfernungsklassen definiert. Demzufolge wurde für Atome mit weniger als etwa 10 \AA Abstand die Coulomb-Summe explizit berechnet und wurden die elektrostatischen Beiträge aller weiter entfernten DMSO-Moleküle, welche hier die strukturellen Einheiten darstellen,

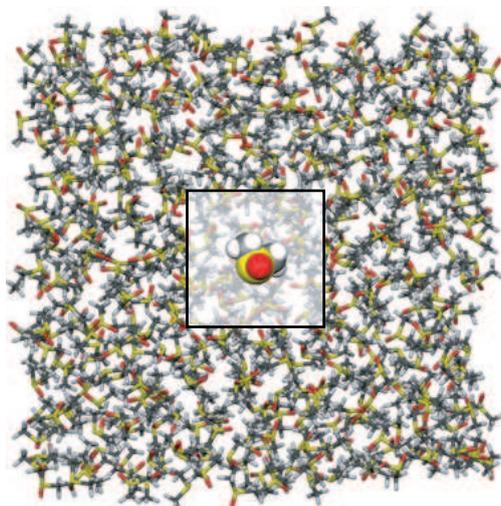


Abbildung 3.5: Hybride Simulation von DMSO in DMSO. Gezeigt wird ein Schnitt durch das würfelförmige Simulationssystem, das vollständig mit molekularmechanischen DMSO-Molekülen des VAK-Modells gefüllt ist. Ein einzelnes Molekül wurde ausgewählt und innerhalb der hell angedeuteten QM-Box im Rahmen der DFT quantenmechanisch beschrieben.

durch die ihrer Dipolmomente approximiert. Der Überlapp der beiden Entfernungsklassen wurde zu $1,25 \text{ \AA}$ gewählt. Konform zu der Minimum-Image Konvention lag der verwendete Reaktionsfeld-Cutoff bei der halben Kantenlänge des Systems. Das sich anschließende dielektrische Kontinuum wurde durch ein Reaktionsfeld mit der experimentellen Dielektrizitätskonstante von $\epsilon = 45,8$ berücksichtigt. Nach einer kurzen Relaxation des neu aufgesetzten Systems, während der sich die Geschwindigkeiten der Atome in jedem Zeitschritt dissipativ um den Faktor 0,9 verringerten, wurde zur Equilibrierung des Systems eine rein molekülmechanische Simulation mit Zeitschritten von 1 fs durchgeführt. Dabei hielt der M-SHAKE Algorithmus [159] die Länge der Wasserstoffbindungen starr auf ihrem Gleichgewichtswert. Während der ersten 15 ps der Equilibrierung des Systems wurde es mit einer Zeitkonstante von $\tau_T = 0,5 \text{ ps}$ an das Wärmebad angebunden, worauf 35 ps mit $\tau_T = 1 \text{ ps}$ folgten.

Abbildung 3.5 zeigt einen Schnitt durch das Simulationssystem, in dem ein in der Mitte des Würfels liegendes DMSO-Molekül ausgewählt wurde, um in der Hybridsimulation quantenmechanisch beschrieben zu werden. Als das dabei zu berücksichtigende Volumen im Ortsraum wurde eine QM-Box von 10 \AA Kantenlänge gewählt, deren Mittelpunkt während der Simulation dem gewählten Molekül nachgeführt, dabei aber nicht rotiert wurde. Auf diese Weise betrug der Abstand zwischen den Kernen des DMSO-Moleküls und den Begrenzungen der Box, unter Berücksichtigung der Rotation und eventueller Verformungen des Moleküls, immer mindestens $2,5 \text{ \AA}$. Mit dem auch in den Vakuumrechnungen verwendeten Cutoff der ebenen Wellen von 100 Ry benötigte die Berechnung der Grundzustands-Elektronendichte des DMSO-Moleküls in dieser Box etwa 6 CPU-Minuten auf

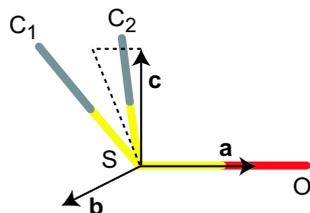


Abbildung 3.6: Das molekülfeste Koordinatensystem für DMSO. Farblich dargestellt sind die schweren Atome des DMSO-Moleküls mit ihren Bindungen. Die Einheitsvektoren $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ des gezeigten molekülfesten Koordinatensystems werden zu jedem Zeitschritt der Trajektorie aus den Ortsvektoren der Atome $\{\mathbf{r}_O, \mathbf{r}_S, \mathbf{r}_{C_1}, \mathbf{r}_{C_2}\}$ durch

$$\mathbf{a} = \mathbf{r}_O - \mathbf{r}_S, \quad \mathbf{b} = \mathbf{a} \times (\mathbf{r}_{C_1} + \mathbf{r}_{C_2} - 2\mathbf{r}_S), \quad \mathbf{c} = \mathbf{a} \times \mathbf{b}$$

und eine sich anschließende Normierung berechnet. Nach dieser Konstruktion spannen \mathbf{a} und \mathbf{c} die Ebene auf, in der sowohl die S=O-Bindung als auch die Winkelhalbierende der beiden C–S-Bindungen liegen. \mathbf{b} steht senkrecht auf dieser Ebene, zu der das Molekül in seiner Gleichgewichtsgeometrie symmetrisch ist.

einer mit 600 MHz getakteten Alpha-Workstation des Typs XP1000. Die Integrationszeit betrug 0,25 fs, um die Bewegung der Wasserstoffatome ausreichend häufig abzutasten. Die hier vorgestellte hybride Trajektorie verfolgte die Zeitentwicklung des einzelnen DMSO-Moleküls in seiner natürlichen Umgebung über eine Gesamtzeit von 5 ps oder 20.000 Zeitschritte und entsprach damit einem Rechenaufwand von ungefähr 80 CPU-Tagen. Das molekularmechanische System bildete ein NVT-Ensemble, wobei die quantenmechanisch beschriebenen Atome von der Geschwindigkeitsskalierung ausgenommen wurden.

Im Verlauf der hybriden Trajektorie wurde neben der Geometrie des quantenmechanisch beschriebenen DMSO-Moleküls das Dipolmoment seiner Ladungsverteilung in Zeitschritten von 1 fs gespeichert. Zur genaueren Analyse der auftretenden Polarisierungseffekte sollen hier nicht nur der Betrag des Dipolmoments, sondern auch seine einzelnen Komponenten diskutiert werden. Um die Rotationen und Translationen des DMSO-Moleküls zu berücksichtigen, wurde ein Übergang von den kartesischen Koordinaten des Simulationssystems hin zu einem Satz von molekülfesten Basisvektoren durchgeführt, die in Abbildung 3.6 definiert werden. Abbildung 3.7 zeigt die Zeitentwicklung und die Verteilung der Projektion des Dipolmoments auf diese Basisvektoren. In Tabelle 3.3 sind die Mittelwerte und Standardabweichungen der Einzelkomponenten zusammengefasst.

Während der ersten 0,5 ps der Trajektorie sind starke Veränderungen der Komponenten des Dipolmoments zu beobachten. Dies ist auf die Relaxation der lokalen Umgebung des Moleküls nach dem erfolgten Beschreibungswechsel zurückzuführen. Für μ_a , die erste und parallel zur S=O-Bindung liegende Komponente des Dipolmoments, beobachtet man eine periodische Zeitabhängigkeit mit großer Amplitude. Dieses Verhalten stammt offensichtlich von der harmonischen Schwingung der stark polaren S=O Bindung um ihre

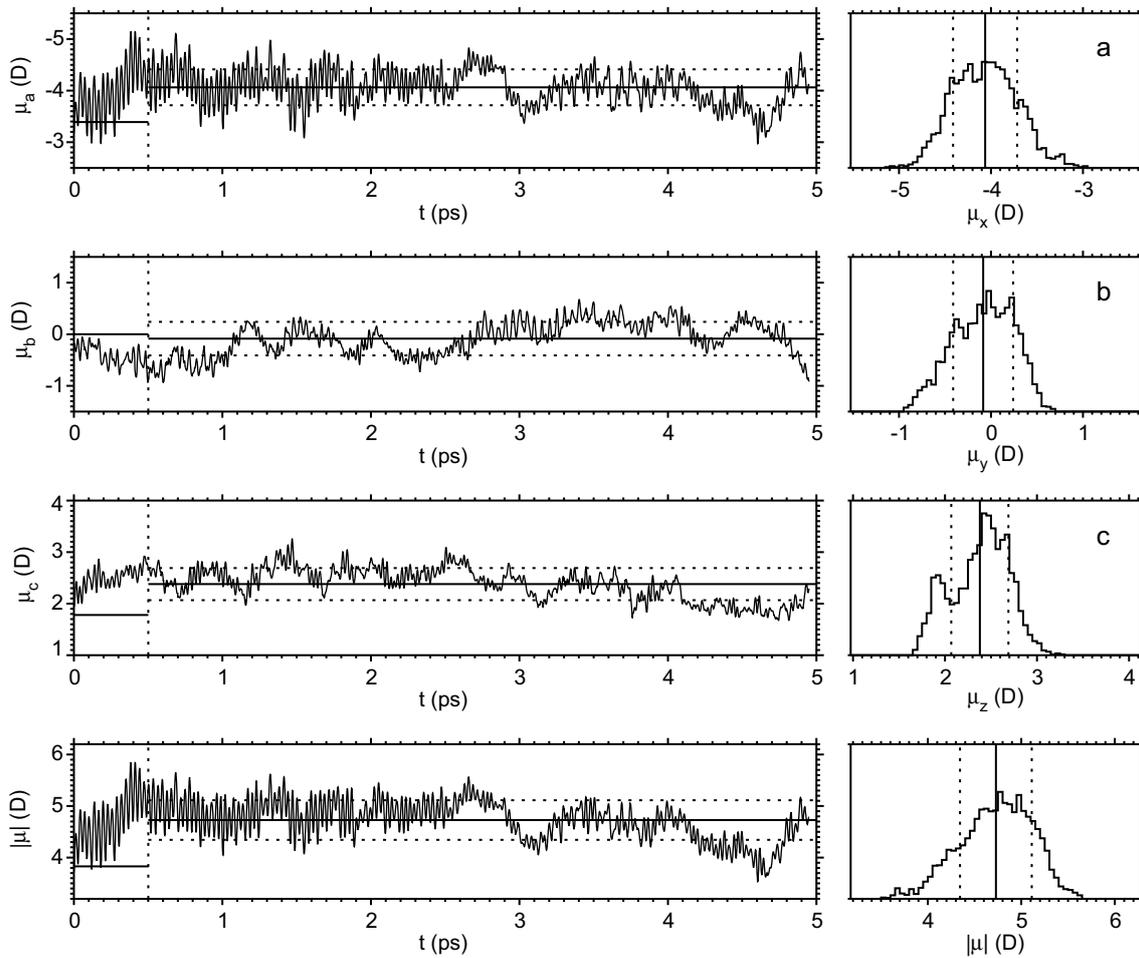


Abbildung 3.7: Das Dipolmoment des DMSO-Moleküls in der 5-ps Hybridsimulation. Aus der Ladungsverteilung des quantenmechanisch beschriebenen DMSO-Moleküls wurde das Dipolmoment berechnet und auf die molekülfesten Koordinaten projiziert, die in Abbildung 3.6 definiert sind. Die linke Spalte zeigt sowohl die Zeitentwicklung der drei Dipolkomponenten als auch die des Gesamtbetrages. In den ersten 500 fs wurde eine Relaxation des Systems angenommen, die durchgezogene Linie zeigt hier die Werte an, die dem Dipolmoment der Vakuumsrechnung entsprechen. Zu späteren Zeiten ist mit der durchgezogenen Linie der jeweilige Mittelwert und mit den gestrichelten Linien die Standardabweichung dargestellt. Letzteres gilt auch für die rechte Spalte von Diagrammen, in denen die aufsummierten Verteilungen der Dipolkomponenten vorgestellt werden. Zahlenwerte hierzu gibt Tabelle 3.3 an, Diskussion im Text.

Gleichgewichtslänge. Interessant ist die in der Mitte der Trajektorie auftretende Veränderung dieses Verhaltens. Eine genaue Betrachtung des Simulationssystems in einer dreidimensionalen Darstellung zeigte, dass sich von diesem Zeitpunkt an der Sauerstoff des DFT-Fragments stark dem klassisch beschriebenen Schwefelatom eines anderen DMSO-Moleküls annähert. Durch die entstehende elektrostatische Wechselwirkung nimmt effektiv die Masse des DFT-Sauerstoffs zu und es ändern sich die Frequenz und Amplitude der S=O-Schwingung. An dieser Stelle wurde daraufhin verifiziert, dass kein so genannter „spill-out“ auftrat, d.h. keine Elektronendichte auf den Ort des klassisch modellierten Schwefelatoms übergang, welches in der hybriden Simulation durch eine gaussförmige Ladungsverteilung in der QM-Box repräsentiert wird. Anders als die starke Komponente μ_a , sollte μ_b im Zeitmittel aus Symmetriegründen verschwinden. Der dazugehörige Mittelwert liegt entsprechend nahe bei Null. Allerdings zeigt die Zeitreihe von μ_b mit ihrer langsamen Veränderlichkeit auf Zeitskalen von Pikosekunden, dass die Gesamtlänge der hier vorgestellten Trajektorie bei weitem keine ausreichende Abtastung der Konfigurationen der Lösungsumgebung des quantenmechanisch beschriebenen DMSO-Moleküls gestattet. Deshalb sind die Ergebnisse der Hybridrechnung mit einer gewissen statistischen Unsicherheit verbunden. Der zur S=O-Bindung orthogonale Anteil des Dipolmoments, μ_c , zeigt schließlich eine eher zufällig wirkende Zeitabhängigkeit, wobei sein Beitrag zum Gesamtdipolmoment eher gering ist. Wie die letzte Reihe von Diagrammen in Abbildung 3.7 zeigt, ist dessen Betrag hauptsächlich durch den Wert von μ_a bestimmt.

Tabelle 3.3 zeigt die Mittelwerte der besprochenen Zeitreihen und ermöglicht eine Abschätzung der Zunahme des Dipolmoments beim Übergang eines DMSO-Moleküls vom Vakuum in die kondensierte Phase. Das Anwachsen der beiden Komponenten μ_a und μ_b führt zu einer Vergrößerung des Betrags von μ von 3,83 D auf 4,73 D und damit um 23%. Diese Zunahme stimmt gut mit den schon zitierten Reaktionsfeldrechnungen [183] überein, die eine Steigerung um 20% beobachteten. Tabelle 3.4 ermöglicht zur Diskussion einen Vergleich mit den Dipolmomenten der molekularmechanischen DMSO-Modelle. Während die beiden anderen expliziten Modelle (RESP[181], SF[182]) ein zu großes Dipolmoment aufweisen, wurde für das HC-Modell gefordert, den Betrag und die relative Ausrichtung des in der Hybridtrajektorie beobachteten Dipolmoments zu reproduzieren. Dazu wurden ausgehend von den für das VAK-Modell berechneten ESP-Ladungen q_i^{VAK} die Partialladungen des HC-Modells optimiert. Das molekularmechanische Dipolmoment μ wurde dazu aus \mathbf{r} , der als konstant angenommenen Gleichgewichtsgeometrie des mit den VAK-Parametern modellierten DMSO-Moleküls, und aus den zu bestimmenden Partialladungen q_i^{HC} nach

$$\mu = \sum_{i=1}^{10} \mathbf{r}_i \cdot q_i^{HC} \quad (3.3)$$

berechnet. μ soll die in Tabelle 3.3 gezeigten mittleren Werte für μ_a und μ_c annehmen. Aus den schon erwähnten Symmetrieüberlegungen heraus wurde μ_b dabei als ver-

| | DFT Vakuum | DFT/MM | | Veränderung | |
|----------------------------|-------------|-----------------|--------------------|---------------------|-----------------------------|
| | μ^V [D] | $\bar{\mu}$ [D] | err($\bar{\mu}$) | $\sigma(\bar{\mu})$ | $(\bar{\mu} - \mu^V)/\mu^V$ |
| μ_a (\parallel S=O) | -3,39 | -4,06 | (0,25) | 0,35 | +20% |
| μ_b (symm.) | 0,00 | -0,08 | (0,30) | 0,33 | - |
| μ_c (\perp S=O) | 1,78 | 2,38 | (0,29) | 0,31 | +33% |
| $ \mu $ | 3,83 | 4,73 | (0,32) | 0,38 | +23% |
| $\angle(\mu, \text{S=O})$ | 27,7° | 30,3° | - | - | - |

Tabelle 3.3: Zunahme des Dipolmoments beim Übergang in die kondensierte Phase. Das aus der Ladungsverteilung des DMSO-Moleküls berechnete Dipolmoment wird sowohl für die LDA/SG-optimierte Vakuum-Struktur (μ^V , siehe Tabelle 3.1) als auch als Mittelwert $\bar{\mu}$ der hybriden DFT/MM-Trajektorie (Abbildung 3.7) angegeben. Aufgeführt sind der Gesamtbetrag des Dipolmoments und seine Projektionen auf das in Abbildung 3.6 eingeführte molekülfeste Koordinatensystem. Für die Observablen der Hybridtrajektorie wird neben der einfachen Standardabweichung der Verteilung (σ) eine Fehlerabschätzung (*err*) angegeben. Diese Abschätzung basiert auf der Standardabweichung von Mittelwerten, die auf 50 verschiedenen Teilstücken der Zeitreihe ermittelt wurden, ein Vorgehen das als so genanntes „Block-averaging“ in der Monographie von Allen und Tildesley [74] beschrieben wird. Zusätzlich gibt die letzte Spalte der Tabelle die prozentuale Zunahme des Dipolmoments beim Übergang vom Vakuum in die Flüssigkeit. Außerdem wird in der letzten Zeile der Tabelle der gemessene und der in der Simulation beobachtete mittlere Winkel zwischen dem Gesamtdipolmoment und der S=O-Bindung aufgeführt.

| | OPLS | VG | MP | MI | RESP | SF | HC |
|-----------------|------|------|------|------|------|------|------|
| Dipolmoment [D] | 4,48 | 4,59 | 4,44 | 4,45 | 5,01 | 5,24 | 4,68 |

Tabelle 3.4: Dipolmomente der DMSO-Modelle. Für die starren Modelle (OPLS, VG [167], MP[168]) wird das Dipolmoment eines einzelnen Moleküls im Vakuum angegeben. Im Fall der flexiblen Parametrisierungen wurde das mittlere Dipolmoment pro Molekül aus der Nanosekunden-Trajektorie bestimmt.

schwindend angenommen. Als Nebenbedingungen dieser Optimierung wurden weiterhin die Neutralität des Gesamtmoleküls

$$\sum_{i=1}^{10} q_i^{HC} = 0 \quad (3.4)$$

und das Auftreten gleicher Ladungen an allen Wasserstoffen gefordert, wie auch eine möglichst geringe quadratische Abweichung der neuen von den alten Partialladungen

$$\sum_{i=1}^{10} (q_i^{HC} - q_i^{VAK})^2 \rightarrow \min. \quad (3.5)$$

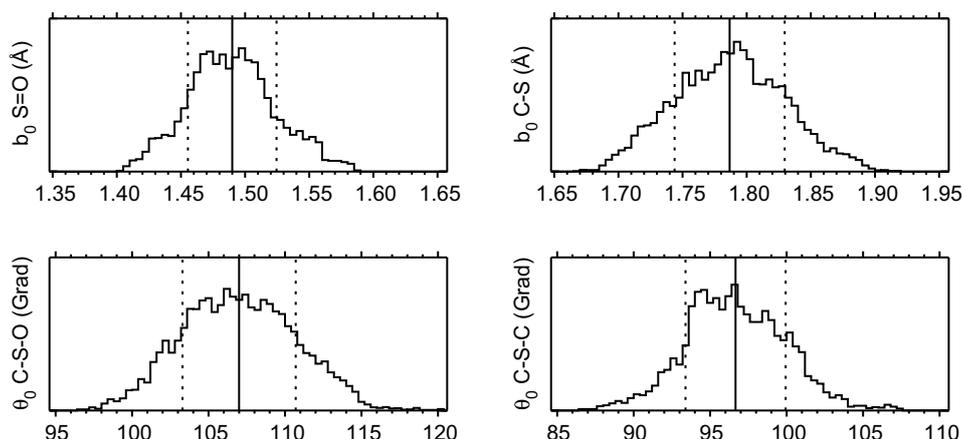


Abbildung 3.8: Die Verteilung der internen Koordinaten während der Hybridtrajektorie. Gezeigt sind die Histogramme der Bindungslängen und -winkel des quantenmechanisch beschriebenen DMSO-Moleküls. Wie zuvor geben die durchgezogenen Linien den Mittelwert und die gestrichelten Linien die Standardabweichungen der Verteilungen an, deren Zahlenwerte in Tabelle 3.5 aufgeführt werden.

Die letzte Forderung stabilisierte die Optimierung und konnte eingeführt werden, da das Gleichungssystem bis zu diesem Zeitpunkt noch unterbestimmt war. Das Ergebnis der Optimierung der q_i^{HC} unter den aufgeführten Nebenbedingungen wird in Form der Partialladungen des HC-DMSO Modells in Tabelle 3.7 aufgeführt. Ein Vergleich mit den Ladungen des VAK-Modells zeigt eine Zunahme der negativen Ladung des Sauerstoffatoms um etwa 13%, dagegen für den Schwefel eine Abnahme der auch schon vorher geringen Ladung um 26% und nur eine geringe Veränderung der Ladung der Kohlenstoffe. Deutlich ist dagegen wiederum die Zunahme der Partialladungen der Wasserstoffe um etwa 15%, deren Betrag von $0,140 e$ größer ist als der im CHARMM22-Protein-Kraftfeld für Wasserstoffe in Methylgruppen übliche Wert von $0,090 e$. Das resultierende mittlere Dipolmoment eines MM-DMSO-Moleküls in Lösung liegt mit $4,68 D$ sehr nahe am oben berechneten Mittelwert der Hybridtrajektorie von $4,73 D$ (vergleiche Tabelle 3.4).

Abbildung 3.8 und Tabelle 3.5 zeigen die Geometrie des DMSO-Moleküls während der Hybridtrajektorie. Die Bindungslängen der S=O Doppelbindung und die der C-S Einfachbindungen, als auch die Winkel zwischen den Atomen C-S-O und C-S-C wurden schon in Tabelle 3.1 zum Vergleich der verschiedenen DFT-Funktionale herangezogen und geben hier Einblick in die Änderung der Geometrie von DMSO in der kondensierten Phase im Vergleich zu seiner Vakuum-Struktur. Tabelle 3.5 zeigt die Mittelwerte dieser Koordinaten über alle Zeitpunkte der Hybridtrajektorie an. Berücksichtigt man den abgeschätzten Fehler der Mittelwerte, so kann nur für die S=O-Bindung eine signifikante Zunahme der Bindungslänge beobachtet werden, während die anderen drei Koordinaten nur Änderungen innerhalb des Fehlerbereichs zeigen. Da nur eine sehr geringe Verlängerung der S=O-Bindung um $0,011 \text{ \AA}$ auftritt, wurde zur Parametrisierung des endgültigen

| | DFT Vakuum | DFT/MM | | Veränderung | |
|-----------------------|------------|-----------|----------|-------------|-----------------------|
| | x_V | \bar{x} | $err(x)$ | $\sigma(x)$ | $(\bar{x} - x_V)/x_V$ |
| Bindungslänge S=O [Å] | 1,479 | 1,490 | (0,005) | 0,036 | +0,7% |
| Bindungslänge C-S [Å] | 1,784 | 1,787 | (0,006) | 0,043 | +0,2% |
| Winkel C-S-O [Grad] | 106,7 | 107,0 | (0,8) | 3,7 | +0,3% |
| Winkel C-S-C [Grad] | 95,8 | 96,7 | (1,2) | 3,3 | +0,9% |

Tabelle 3.5: Die Geometrie des DMSO-Moleküls in der Hybridtrajektorie. Aufgeführt werden die Mittelwerte, die durch Blockmittelung abgeschätzten Fehler und die Varianzen der Bindungslängen und -winkel des quantenmechanisch beschriebenen DMSO-Moleküls. Zum Vergleich mit den Vakuum-Rechnungen werden die entsprechenden Werte aus Tabelle 3.1 noch einmal wiederholt und die jeweilige Veränderung beim Übergang in die kondensierte Phase angegeben.

DMSO-Modells die Geometrie des Vakuum-Modells unverändert übernommen. Außerdem ist mit der gezeigten Beobachtung einer fast konstanten Geometrie auch keine starke Änderung der Kraftkonstanten beim Übergang in die Flüssigkeit zu erwarten, so dass auch diese unverändert übernommen wurden, wie Tabelle 3.6 zeigt.

3.1.3 Optimierung der Van der Waals Parameter

Im Vorangehenden konnte aufgezeigt werden, wie durch geeignete DFT-Berechnungen, und damit aus „ersten Prinzipien“ und unabhängig von Experimenten, die Geometrie und die Partialladungen eines Lösungsmittelmodells bestimmt werden können, die dessen Elektrostatik in der Flüssigkeit möglichst realitätsgetreu nachbilden. Für die noch fehlenden VDW-Parameter ist ein solches Vorgehen allerdings nicht möglich, da (i) DFT-Berechnungen die Dispersionswechselwirkungen nicht erfassen können [198], diese jedoch für den attraktiven Teil des VDW-Potentials verantwortlich sind und (ii) die makroskopischen thermodynamischen Eigenschaften des Lösungsmittelmodells sehr empfindlich von den gewählten VDW-Parametern abhängen, weshalb diese sehr genau festgelegt werden müssen.

Traditionell werden diese Parameter deshalb ausgehend von Werten in ähnlichen Molekülen in Handarbeit durch kleine Änderungen und sich anschließenden Testrechnungen an gemessene Werte von thermodynamischen Observablen angepasst. Meistens sind dies die Dichte des Lösungsmittels ρ und seine *Verdampfungswärme* H_{vap} (engl. heat of vaporization). In erster Näherung erwartet man dabei eine Abnahme der Dichte bei Vergrößerung der VDW-Radien σ (vgl. die Definition der VDW-Energie (2.20)) und eine Zunahme der Verdampfungswärme mit den Minimumsenergien ε [74]. Mehrere Versuche im Rahmen dieser Arbeit zeigten für DMSO jedoch eine weit kompliziertere Abhängigkeit, bei der diese Zusammenhänge sogar umgekehrt wurden. Das Zusammenspiel der VDW-Wechselwirkung mit der starken Elektrostatik des DMSOs führte beispielsweise zu einer Zu-

nahme der Dichte, sobald der VDW-Radius des Schwefels vergrößert wurde. Auch sind für DMSO acht verschiedene Parameter, vier Radien und vier Energien, zu optimieren. Es liegt also ein hochdimensionales Optimierungsproblem vor. Im Gegensatz zu den am Anfang des Kapitels vorgestellten DMSO-Modellen, bei denen nur einige der Parameter per Hand in großen Schritten variiert wurden, konnten in der vorliegenden Arbeit die Parameter des Sauerstoffs und Schwefels (σ_O , ε_O , σ_S und ε_S) frei optimiert werden. Anwendung fand ein automatisiertes Verfahren, das auf der Simplex-Methode (siehe Anhang A) der mehrdimensionalen Optimierung [199] beruht und zuerst von Faller et al. [200] zur Kraftfeldoptimierung eingesetzt wurde. Da das Simplex-Verfahren eine skalare Zielfunktion benötigt, wurde diese durch die Summe der relativen Abweichungen der Dichte und der Verdampfungswärme von ihren experimentellen Werten (siehe Tabelle 3.8) definiert.

$$f(x_i) = \left| 1 - \frac{\rho(x_i)}{\rho^{\text{exp}}} \right| + \left| 1 - \frac{H_{\text{vap}}(x_i)}{H_{\text{vap}}^{\text{exp}}} \right| \quad (3.6)$$

Eine hier mögliche Gewichtung der beiden Einzelbeiträge war nicht notwendig. Zur Evaluation der Zielfunktion (3.6) für einen bestimmten Satz von VDW-Parametern wurde jeweils eine eigenständige MD-Simulation durchgeführt. Dabei wurde der Startzustand des Systems, d.h. die Atomkoordinaten und -geschwindigkeiten, aus derjenigen schon durchgeführten Simulation entnommen, welche die zu dem aktuellen Satz ähnlichsten VDW-Parameter aufwies. Das Simulationsvolumen hatte die Form eines rhombischen Dodekaeders mit einem Innenradius von 28 Å, beinhaltete 956 DMSO-Moleküle und unterlag periodischen Randbedingungen. Abweichend von den Grundeinstellungen in EGOMMII wurde in allen folgenden Simulationen die SAMM-Abstandsklasse der Ebene 0 von 6 Å auf 9 Å vergrößert, d.h. die explizite Auswertung der Coulomb-Summe als auch die Berechnung der VDW-Wechselwirkung wurde in einem Umkreis von etwa 14 Å (anstatt 10 Å) um jede betrachtete Ladung vorgenommen. Wie zuvor wurde als höchste Clusterebene die der strukturellen Einheiten (Level 0) gewählt und ein Reaktionsfeld mit einer Dielektrizitätskonstante von $\varepsilon = 45,8$ berücksichtigt. Der Zeitschritt der Integration betrug wie üblich 1 fs. Die Equilibrierung des Systems wurde jeweils in zwei Schritten durchgeführt: zuerst 5 ps im NVT-Ensemble mit $T = 298$ K und einer Zeitkonstante der Anbindung des Thermostaten von $\tau_T = 0,25$ ps, anschließend im NPT-Ensemble unter Hinzunahme des Barostaten mit $P = 1$ atm, $\tau_P = 1$ ps und der experimentellen isothermen Kompressibilität κ_T aus Tabelle 3.8. Anschließend erfolgte die Berechnung der auszuwertenden Trajektorie von 100 ps Länge.

Unter Verwendung von zwei Knoten eines Linux-Clusters, die jeweils mit zwei Prozessoren des Typs AthlonMP 1900+ bestückt sind, betrug die Rechenzeit pro Integrations-schritt etwa 0,65 s und damit etwas mehr als einen Tag pro Satz von VDW-Parametern. Als Startsimplex wurden die Werte des Vakuum-Modells (aus CHARMM22 entnommen) zusammen mit vier Auslenkungen entlang der Koordinatenachsen des Parameterraumes um je 10% gewählt. Nach der Berechnung der Observablen für 37 Wertekombinationen erzielte der Simplex-Algorithmus eine Abweichung von weniger als 1% in der Dichte und in der Verdampfungswärme. Dabei ist zu beachten, dass diese Observablen, bedingt durch

die kurze 100-ps Trajektorie, mit einem Fehler dieser Größenordnung behaftet sind. Tabelle 3.7 zeigt das Endergebnis der Optimierung in Form der VDW-Parameter des HC-DMSO Modells und Tabelle 3.8 die auf einer längeren Trajektorie für dieses Modell bestimmten Werte von ρ und H_{vap} .

Ein weiteres neues Modell: MI-DMSO

Aus Beobachtungen, die ich später vorstellen möchte, entstand die Idee, ein weiteres Modell für DMSO zu parametrisieren, dessen Eigenschaften zwischen den starren „united-atom“ Modellen, und den neueren expliziten und flexiblen Modellen liegen. Dazu wurden die Geometrie und die Partialladungen des MP-Modells [168] verwendet, aber die Bindungslängen und Winkel als flexibel angenommen. Die dazu notwendigen Kraftkonstanten entstammen der in diesem Abschnitt vorgestellten Parametrisierung des HC-Modells. Die Tabellen 3.6 und 3.7 verdeutlichen dies, da sie diesen neuen Satz von Parametern unter der Bezeichnung MI-DMSO (wie „Mitte“) aufführen. Die angegebenen VDW-Parameter des Sauerstoffs und des Schwefels wurden, ausgehend von denen des MP-Modells, wiederum mit Hilfe der Simplex-Methode angepasst. Dabei wurde für die VDW-Radien eine leichte Zunahme beobachtet, die sich auch in der etwas geringeren Dichte des MI-Modells im Vergleich zum MP-Modell bemerkbar macht.

3.2 Evaluation der DMSO-Modelle

Die Tabellen 3.6 und 3.7 führen die Parametrisierungen der hier neu geschaffenen DMSO-Modelle zusammen mit den Parametern der zuvor veröffentlichten Modelle auf. Dabei werden die Modelle mit Akronymen bezeichnet, die in der Legende von Tabelle 3.6 erläutert werden. Tabelle 3.6 gibt für jedes der DMSO-Modelle an, ob dies ein kompaktes „united-atom“ oder ein explizites „all-atom“ Modell ist und ob die Bindungswinkel flexibel sind oder festgehalten werden. Aus diesen beiden Informationen ergibt sich jeweils die zusätzlich aufgeführte Anzahl der Freiheitsgrade pro Molekül. Weiterhin enthält Tabelle 3.6 die Bindungsparameter und die Kraftkonstanten der Modelle, wogegen Tabelle 3.7 die Parameter der langreichweitigen Wechselwirkungen beinhaltet.

Im Folgenden werden die beiden hier neu vorgestellten Modelle (HC, MI) hinsichtlich ihrer thermodynamischen, kinetischen und strukturellen Eigenschaften sowohl mit den rigiden „united-atom“ Modellen (OPLS, VG, MP) als auch mit den flexiblen „all-atom“ Modellen (RESP, SF) verglichen.

3.2.1 Thermodynamische und kinetische Observablen

Für jede der hier diskutierten DMSO-Parametrisierungen wurde ein 956 Moleküle umfassendes Simulationssystem generiert. Diese Systeme unterlagen den gleichen periodischen

3 DMSO als explizites Lösungsmittel

| | | OPLS ^a | VG ^b | MP ^c | MI ^d | RESP ^e | SF ^f | VAK/HC ^g |
|-----------------------|------------|-------------------|-----------------|-----------------|-----------------|-------------------|-----------------|---------------------|
| Wasserstoffe | | united | united | united | united | all | all | all |
| Starre Moleküle | | ja | ja | ja | nein | nein | nein | nein |
| Freiheitsgrade | | 6 | 6 | 6 | 12 | 24 | 24 | 24 |
| Bindungen | | | | | | | | |
| S-O | b_0 | 1,530 | 1,530 | 1,530 | 1,530 | 1,530 | 1,530 | 1,479 |
| | k_b | - | - | - | 573,8 | 570,0 | 540,0 | 573,8 |
| S-C | b_0 | 1,800 | 1,950 | 1,950 | 1,950 | 1,810 | 1,800 | 1,784 |
| | k_b | - | - | - | 206,9 | 227,0 | 240,0 | 206,9 |
| C-H | b_0 | - | - | - | - | 1,090 | 1,110 | 1,100 |
| | k_b | - | - | - | - | 340,0 | 322,0 | 357,2 |
| Winkel | | | | | | | | |
| O-S-C | θ_0 | 106,8 | 106,8 | 106,8 | 106,7 | 106,8 | 106,8 | 106,7 |
| | k_θ | - | - | - | 88,6 | 80,0 | 79,0 | 88,6 |
| C-S-C | θ_0 | 97,4 | 97,4 | 97,4 | 95,8 | 97,4 | 95,0 | 95,8 |
| | k_θ | - | - | - | 72,2 | 62,2 | 34,0 | 72,2 |
| S-C-H | θ_0 | - | - | - | - | 109,5 | 111,3 | 108,5 |
| | k_θ | - | - | - | - | 50,0 | 46,1 | 35,9 |
| H-C-H | θ_0 | - | - | - | - | 109,5 | 108,4 | 110,4 |
| | k_θ | - | - | - | - | 35,0 | 35,5 | 34,0 |
| Torsionswinkel | | | | | | | | |
| X-C-S-X | k_ϕ | - | - | - | - | 0,28 | 0,20 | 0,37 |

Tabelle 3.6: Bindungsparameter der DMSO-Modelle. Die Modelle werden durch Akronyme bezeichnet, welche unten werden. Die Gruppierung der Spalten bzw. der Modelle spiegelt die Behandlung der Wasserstoffatome wieder: „united-atom“, heißt dass Methylgruppen zu Pseudoatomen zusammengefasst werden. Alternativ dazu werden in den „all-atom“ Modellen alle Atome explizit berücksichtigt. Weiterhin sind die Modelle entweder starr oder flexibel, wobei auch bei den flexiblen Parametrisierungen die Bindungslängen der Wasserstoffatome konstant gehalten werden. Die resultierende Anzahl von Freiheitsgraden pro Molekül wird aufgeführt. Die Bindungslängen b_0 werden in Å, die Bindungsstärken k_b in kcal/(mol · Å²), die Gleichgewichtswinkel θ_0 in Grad und die Winkelstärken k_θ, k_ϕ in kcal/(mol · rad²) angegeben. Die Multiplizität der Torsionswinkel der Methylgruppen ist hier immer $n = 3$ und die Phase $\delta = 0$ (vergleiche mit Gleichungen (2.15) bis (2.17)).

Erklärung der Akronyme und Referenzen:

^a(OPLS) Aus dem OPLS-Kraftfeld. Bisher unpubliziert. Parameter aus [201].

^b(VG) Das Modell der Gruppe van Gunsteren [167].

^c(MP) Eine Neuparametrisierung des VG-Modells [168].

^d(MI) Eine Mischung aus den Modellen MP und HC. Hier neu vorgestellt.

^e(RESP) Dieses Modell wurde für das AMBER-Kraftfeld erstellt[181].

^f(SF) Die Parametrisierung von Strader und Feller [182].

^g(HC) Das hier neu vorgestellte flexible und explizite Modell, zu dessen Erstellung das Vakuum-Modell (VAK) als Zwischenstufe dient. Beide Modelle weisen den gleichen Satz von Bindungsparametern auf.

| | | OPLS | VG | MP | MI | RESP | SF | VAK ^a | HC |
|--------------------------------|------------|---------|---------|---------|---------|---------|---------|------------------|---------|
| Partiellladungen | | | | | | | | | |
| O | q | -0,4590 | -0,4590 | -0,4367 | -0,4367 | -0,5205 | -0,5560 | -0,3720 | -0,4210 |
| S | q | 0,1390 | 0,1390 | 0,1167 | 0,1167 | 0,3155 | 0,3120 | 0,1640 | 0,1210 |
| C/CH ₃ | q | 0,1600 | 0,1600 | 0,1600 | 0,1600 | -0,3244 | -0,1480 | -0,2620 | -0,2710 |
| H | q | - | - | - | - | 0,1423 | 0,0900 | 0,1220 | 0,1400 |
| Van der Waals-Parameter | | | | | | | | | |
| O | ϵ | 0,2797 | 0,4097 | 0,4079 | 0,4165 | 0,2100 | 0,1200 | 0,1200 | 0,1354 |
| | σ | 2,9300 | 2,6300 | 2,6300 | 2,6616 | 2,9600 | 3,0290 | 3,0291 | 3,2754 |
| S | ϵ | 0,3946 | 0,3098 | 0,3098 | 0,3159 | 0,2500 | 0,3500 | 0,4500 | 0,4398 |
| | σ | 3,5600 | 3,5600 | 3,5600 | 3,8061 | 3,5640 | 3,5640 | 3,5636 | 3,4863 |
| C/CH ₃ | ϵ | 0,1598 | 0,2248 | 0,2248 | 0,2248 | 0,1094 | 0,0780 | 0,0800 | 0,0800 |
| | σ | 3,8100 | 3,6600 | 3,7390 | 3,7390 | 3,4000 | 3,6350 | 3,6710 | 3,6710 |
| H | ϵ | - | - | - | - | 0,0157 | 0,0240 | 0,0220 | 0,0220 |
| | σ | - | - | - | - | 2,4710 | 2,3880 | 2,3520 | 2,3520 |

Tabelle 3.7: Partiellladungen und VDW-Parameter der DMSO-Modelle. Für die in Tabelle 3.6 aufgeführten Modelle sind die Partiellladungen der Atome in Einheiten der Elementarladung e , die VDW-Energien ϵ in kcal/mol und die Radien σ in Å angegeben. ^aZusätzlich zu den Ladungen und VDW-Parametern des hier neu vorgestellten HC-Modells werden auch die Werte des, als Zwischenstufe bei der Erstellung des HC-Modells verwendeten, Vakuum-Modells (VAK) aufgeführt. Referenzen siehe Tabelle 3.6.

Randbedingungen, wie das in Abschnitt 3.1.3 zur Optimierung der VDW-Parameter des HC-DMSOs eingesetzte System. Weiterhin wurden auch die gleichen Einstellungen verwendet, beispielsweise die des FAMUSAMM-Algorithmus', eine Temperatur von 298 K oder ein Druck von 1 atm. Abweichend von den Optimierungsläufen wurde jedoch die Dauer der Equilibrierungen auf 100 ps verlängert und für die Auswertung jeweils Trajektorien von einer Nanosekunde Länge generiert. Die verschiedenen Energiebeiträge und die Atomkoordinaten wurden in Zeitabständen von 100 fs abgespeichert. Als Endergebnis zeigt Tabelle 3.8 die aus diesen Simulationen extrahierten thermodynamischen und dynamischen Observablen, sowie experimentelle Vergleichswerte, die an flüssigem DMSO unter Normalbedingungen gemessen wurden.

Betrachtet man als erstes die Dichten der verschiedenen Modelle, so wird deutlich, dass die neu eingeführte VDW-Korrektur im Zusammenspiel mit der von Martina Stork implementierten Druckmessung eine korrekte Beschreibung des Lösungsmittels im NPT-Ensemble ermöglicht. Besonders die schon von anderen Arbeitsgruppen gemachte Beobachtung, dass ältere DMSO-Modelle (VG, RESP) eine zu hohe Dichte aufweisen, sobald ihre langreichweitige Elektrostatik vollständig mit Hilfe eines RF- oder Ewald-Ansatzes Berücksichtigung findet [168], konnte reproduziert werden. Es zeigt sich aber auch, dass das aus diesem Grund nachparametrisierte MP-Modell bei Verwendung der FAMUSAMM-Methode und des hier eingesetzten, recht großen Simulationssystems immer noch eine leicht

zu hohe Dichte aufweist. Dagegen reproduzieren das OPLS-, das kürzlich vorgestellte SF-Modell und die hier neu eingeführten MI- und HC-Parametrisierungen die experimentelle Dichte. Im Fall der neuen Modelle, bei denen offensichtlich identische Simulationsbedingungen während der Optimierung und der hier diskutierten Evaluation vorlagen, sind die kleinen Abweichungen zur gemessenen Dichte durch die in der Optimierung durchgeführten, recht kurzen Simulationen zu erklären, bei denen der Fehler in der Bestimmung von $\rho = N/\langle V \rangle$ in der Größenordnung der jetzt beobachteten Abweichungen lag.

Die in Tabelle 3.8 gezeigte Fehlerabschätzung basiert im Fall der thermodynamischen Observablen auf der Bestimmung der Anzahl der statistisch unabhängigen Messwerte einer Größe \mathcal{A} [74]. Die Varianz des Mittelwerts von \mathcal{A} hängt von der eigentlichen Varianz von \mathcal{A} , der Anzahl der Messwerte N_t und der Korrelationszeit dieser Messwerte $\tau_{\mathcal{A}}$ ab.

$$\sigma^2(\langle \mathcal{A} \rangle) = \frac{2 \tau_{\mathcal{A}}}{N_t} \sigma^2(\mathcal{A}) \quad (3.7)$$

$\tau_{\mathcal{A}}$ wird dabei in Einheiten der Zeitabstände der Messwerte angegeben und bezeichnet den Zeitpunkt, zu dem die Autokorrelationsfunktion von $\mathcal{A}(t)$ „ausreichend“ weit abgefallen ist. Die hier angewendete Bindung war, dass der natürliche Logarithmus der Autokorrelationsfunktion auf einen willkürlichen Wert von -3 abgefallen sein muss. Auf diese Weise können vor allem die Fehler untereinander gut verglichen werden. So zeigt sich beispielsweise für die Dichte ρ ein relativ kleiner Fehler, der in der starken Anbindung des Barostaten an die DMSO-Systeme begründet ist.

Die zweite Observable ist die Verdampfungswärme H_{vap} des Lösungsmittels. Sie ist als Differenz der Enthalpien pro Molekül zwischen der kondensierten und der gasförmigen Phase definiert.

$$H_{vap} = H_{liq} - H_{gas} \quad (3.8)$$

Berücksichtigt man zur Berechnung der inneren Energie des kondensierten Systems $U = E_{VDW} + E_{elec}$ nur die intermolekularen Beiträge, so kann unter Annahme des Verhaltens eines idealen Gases H_{gas} zu $RT = 0,596$ kcal/mol abgeschätzt werden und man erhält die Berechnungsvorschrift für H_{vap}

$$H_{vap} = -\langle U \rangle \frac{N_A}{N} - RT. \quad (3.9)$$

Für die Verdampfungswärme gelten in Tabelle 3.8 ähnliche Beobachtungen wie für die zuvor diskutierte Dichte: als eine der Observablen, die üblicherweise bei der Generierung der VDW-Parameter optimiert wird, sollte bei allen Modellen für die Verdampfungswärme eine gute Übereinstimmung zu erwarten sein. Allerdings kann auch U , und damit die Verdampfungswärme, sehr sensitiv auf Veränderungen in der Methode der Berechnung der langreichweitigen Wechselwirkungen reagieren. Zu beobachten ist ein deutlich zu geringer Wert von H_{vap} für das OPLS-Modell, aber keine großen Abweichungen für VG

| | ρ | H_{vap} | α | κ_T | C_V | D | τ_1 | τ_2 |
|------|--------------------------------|----------------------------------|------------------------------|--------------------------------|--|---------------------------------------|----------|----------|
| | $\frac{\text{g}}{\text{cm}^3}$ | $\frac{\text{kcal}}{\text{mol}}$ | $10^{-4} \frac{1}{\text{K}}$ | $10^{-5} \frac{1}{\text{atm}}$ | $\frac{\text{cal}}{\text{mol}\cdot\text{K}}$ | $10^{-9} \frac{\text{m}^2}{\text{s}}$ | ps | ps |
| Exp | 1,0950 | 12,63 | 9,3 | 5,3 | 28,3 | 0,80 | - | 5 |
| OPLS | 1,0925 (2) | 11,07 (0) | 8,4 (4) | 3,1 (1) | 18,0 (4) | 1,04 (6) | 9 (1) | 4 (0) |
| VG | 1,1480 (3) | 12,56 (0) | 8,1 (3) | 4,0 (1) | 18,5 (6) | 0,70 (7) | 15 (1) | 6 (1) |
| MP | 1,1006 (2) | 11,94 (0) | 8,2 (2) | 2,9 (1) | 19,2 (4) | 0,85 (4) | 13 (1) | 6 (1) |
| MI | 1,0911 (7) | 12,68 (1) | 7,8 (5) | 2,4 (1) | 36,2 (6) | 0,64 (5) | 13 (1) | 6 (1) |
| RESP | 1,1228 (6) | 12,23 (1) | 9,7 (9) | 3,6 (1) | 72,6 (8) | 0,59 (5) | 16 (1) | 7 (1) |
| SF | 1,0960 (7) | 12,93 (2) | 9,2 (6) | 4,6 (5) | 72,5 (6) | 0,58 (4) | 18 (2) | 8 (1) |
| HC | 1,0991 (7) | 12,84 (2) | 8,9 (3) | 4,2 (3) | 73,9 (9) | 0,47 (4) | 43 (6) | 16 (3) |

Tabelle 3.8: Thermodynamische und dynamische Observablen der DMSO-Modelle. Gezeigt sind die Dichte ρ , die Verdampfungswärme H_{vap} , der thermische Expansionskoeffizient α , die isotherme Kompressibilität κ_T und die spezifische Wärme bei konstantem Volumen C_V . Weiterhin werden D , der Diffusionskoeffizient der Translation, und die $\tau_{1,2}$, die Zeiten der Rotations-Diffusion bezüglich der S=O-Bindung, tabelliert. Zahlen in Klammern geben jeweils eine Fehlerabschätzung für die letzte Stelle des davor gezeigten Wertes an. Die Messwerte für flüssiges DMSO unter Normalbedingungen wurden den Veröffentlichungen von Liu et al. [167], Vaisman und Berkowitz [169] bzw. [202] entnommen. Das Vorgehen zur Berechnung der Observablen aus den MD-Trajektorien und der Abschätzung der damit verbundenen Unsicherheiten wird im Text gegeben.

oder RESP, alles drei Modelle, die unter Verwendung eines kleinen Cutoffs erzeugt wurden. Das MP-Modell zeigt eine etwas zu kleine Verdampfungswärme, wogegen das SF-Modell zu stark gebunden zu sein scheint. Für die beiden neuen Parametrisierungen gilt wie schon zuvor, dass die beobachtete Abweichung durch die Endlichkeit der Simulationsläufe während der Optimierung zu erklären sind.

Die nächsten drei thermodynamischen Observablen, der thermische Expansionskoeffizient α , die isotherme Kompressibilität κ_T und die spezifische Wärme bei konstantem Volumen C_V , sind jeweils als die partiellen Ableitungen einer Zustandsgröße nach einer zweiten definiert

$$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P \quad \kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T \quad C_V = \left(\frac{\partial U}{\partial T} \right)_V. \quad (3.10)$$

Diese Größen wurden durch die Bildung finiter Differenzen bestimmt, wozu jeweils zwei zusätzliche Simulationen von je 200 ps Dauer notwendig waren. Für den Expansionskoeffizienten wurden diese Simulationen im NPT -Ensemble bei 283 und 313 K durchgeführt, für die Kompressibilität im NVT -Ensemble und den Dichten von 1,102 und 1,111 g/cm^3 und für C_V in NVT bei wiederum 283 und 313 K. Die Fehler in den einzelnen Observablen, die in die Gleichungen (3.10) eingehen, wurden jeweils mit der zuvor beschriebenen Methode bestimmt und die Fehlerfortpflanzung verfolgt. Wie Tabelle 3.8 zu entnehmen

ist, treten als Folge der Differenzbildung besonders für α und κ_T recht große Fehler auf. Der thermische Expansionskoeffizient hat deshalb keine große Aussagekraft zur Evaluation der DMSO-Modelle, vielmehr beobachtet man lediglich etwas zu geringe Werte für die „united-atom“ Modelle. Erstaunlicherweise zeigt das leicht flexible MI-Modell unter diesen den kleinsten Expansionskoeffizienten. Die expliziten Modelle weisen alle eine gute Übereinstimmung mit dem gemessenen α auf. Auch die Ergebnisse bezüglich der isothermen Kompressibilität κ_T sind ähnlich unspezifisch, alle hier besprochenen DMSO-Modelle unterschätzen diese mehr oder weniger stark.

Für die Kinetik der Wärmedissipation von einem durch die Photoisomerisierung thermisch angeregten Modellpeptid in das DMSO ist die mikroskopische Wärmeleitfähigkeit des Lösungsmittels bestimmend. Um diese Größe für ein gegebenes DMSO-Modell zu bestimmen, müsste ein Nichtgleichgewichts-Ensemble geschaffen werden, in dem durch einen Temperaturgradienten ein zu messender Wärmestrom entsteht. Im Zusammenspiel mit periodischen Randbedingungen ist dies nur schwer zu realisieren, weshalb Müller-Plathe [203] vorschlug, das Problem umgekehrt anzugehen, also einen Wärmefluss in einem System zu erzeugen und den dadurch entstehenden Temperaturgradienten zu bestimmen. Die Realisierung und Kalibrierung eines solchen Verfahrens ist leider sehr aufwändig und konnte deshalb im Rahmen der vorliegenden Arbeit nicht durchgeführt werden.

Aus diesem Grund möchte ich an dieser Stelle die Wärmekapazität C_V der zur Verfügung stehenden DMSO-Modelle diskutieren, denn es kann erwartet werden, dass Unterschiede in C_V für verschiedene Parametrisierungen auch qualitative Schlüsse auf das Verhältnis der jeweiligen Wärmeleitfähigkeiten zulassen. Tabelle 3.8 zeigt die für die DMSO-Modelle ermittelten Werte von C_V . Man erkennt, dass das C_V eines DMSO-Modells anscheinend stets proportional zur Anzahl der Freiheitsgrade f ist (vgl. Tabelle 3.6). Der Proportionalitätsfaktor zwischen diesen beiden Größen beträgt etwa 3,06 cal/(mol·K) (oder 1,54 in Einheiten der Gaskonstante R). Dementsprechend zeigen die rigiden „united-atom“ Modelle ($f = 6$) eine deutliche Unterschätzung des experimentellen Wertes von C_V um mehr als 30%, während die voll flexiblen „all-atom“ Modelle ($f = 24$) die spezifische Wärme konsistent um mehr als 160% überschätzen. Auch das in seiner Anzahl der Freiheitsgrade intermediäre flexible „united-atom“ Modell MI ($f = 12$) überschätzt C_V , jedoch lediglich um etwa 30%.

Die im letzten Fall der „all-atom“ Modelle auftretende Überschätzung von Wärmekapazitäten in MD-Simulationen ist ein bekanntes Phänomen, da in der klassischen Näherung alle vibronischen Freiheitsgrade unabhängig von der Temperatur angeregt werden können. In der quantenmechanischen Realität tragen hingegen nur diejenigen Normalmoden eines Moleküls zur Wärmekapazität bei, die bei einer gegebenen Temperatur eine endliche Besetzungszahl aufweisen. Für $T = 300$ K gilt $k_b T \approx 1/40$ eV. Dies ist die Energie, welche einer harmonischen Schwingung mit der Wellenzahl von etwa 200 1/cm entspricht. Betrachtet man das Spektrum der Normalmoden von DMSO, das ausschnittsweise in Tabelle 3.2 gezeigt wird, so erkennt man eine große Lücke in den Moden zwischen 381 1/cm und 668 1/cm. Geht man von einem „Auftauen“ der fünf Moden mit Wellenzahlen un-

terhalb dieser Lücke aus, so erhält man $f = 11$ beitragende Freiheitsgrade und mit der oben beschriebenen Proportionalität ein theoretisches C_V von ungefähr 33 cal/(mol·K), das sehr nahe am gemessenen Wert liegt. Um eine konsistente Beschreibung im Rahmen der MD-Simulationen zu erreichen, wurden trotzdem, in den expliziten DMSO-Modellen nur die Länge der die höchsten Eigenfrequenzen aufweisenden C–H Bindungen mittels des MSHAKE-Algorithmus festgehalten.

Die letzten drei betrachteten Observablen sind die Diffusionskonstanten der Translation und der Rotation, welche Aussagen über die Beweglichkeit der Lösungsmittelmoleküle machen. Die Berechnung der Diffusionskoeffizienten lehnt sich an die Darstellung in [103] an. Für die Translationsbewegung wurden die Orte der Schwefelatome als Aufpunkte gewählt, dabei soll $\mathbf{r}_i(t)$ den Ort des i ten Schwefels zum Zeitpunkt t bezeichnen. Als Maß für die in der Zeit τ zurückgelegte Distanz wurde die Funktion

$$d(\tau) = \frac{1}{6} \left\langle [\mathbf{r}_i(t + \tau) - \mathbf{r}_i(t)]^2 \right\rangle_{i, \Delta t} \quad (3.11)$$

betrachtet. $\langle \dots \rangle_{i, \Delta t}$ bezeichnet dabei die Mittelung über alle Moleküle i und über alle gegeneinander verschobenen Zeitintervalle der Länge Δt . Zu jedem τ wird somit das Mittel des Abstands der Moleküle von ihren Start- oder auch Referenzorten $\mathbf{r}_i(t)$ bestimmt. Die periodischen Randbedingungen und das damit verbundene Überwechseln einzelner Lösungsmittelmoleküle von einem der Spiegelbilder der Simulationssysteme in ein anderes fand Berücksichtigung, indem beim Auftreten eines jeden solchen Übergangs der Referenzort des betroffenen Moleküls mitversetzt wurde. In der vorliegenden Auswertung wurden, wie schon erwähnt, die Koordinaten während einer Trajektorie von 1 ns Länge alle 100 fs abgespeichert. Die Intervall-Länge Δt betrug 100 ps und die einzelnen Intervalle wurden jeweils um 50 ps gegeneinander verschoben. Auf diese Weise basierte die Auswertung auf 20 Teiltrajektorien. Der Diffusionskoeffizient D wurde anschließend durch einen linearen Fit an $d(\tau)$ im Zeitbereich von 10 bis 100 ps ermittelt, wobei die Auslassung der ersten 10 ps dabei notwendig ist, da die in diesem Zeitbereich auftretende ballistische Bewegung der Moleküle nicht zur Diffusion beitragen darf. Eine Diskussion dieses Vorgehens, besonders zur notwendigen Wahl langer Zeitintervalle und etwaiger Fehlerquellen, findet sich in einer Veröffentlichung von Chitra und Yashonath [204]. Als Fehlerabschätzung wurde für alle Diffusionskonstanten die Standardabweichung der auf den Teiltrajektorien bestimmten Einzelwerte angegeben.

Die in Tabelle 3.8 als gemessen angegebene Diffusionskonstante von $0,8 \cdot 10^{-9} \text{ m}^2/\text{s}$ ist mit einer gewissen Unsicherheit behaftet, da in anderen Quellen auch Beträge von 0,6 oder 0,9 angegeben wurden. Vergleicht man die für die verschiedenen Modelle gemessenen Werte, so zeigen die rigiden und kompakten Parametrisierungen eine mit der Messung übereinstimmende Beweglichkeit, nur das OPLS-Modell scheint eventuell zu beweglich zu sein. Dies kann im Zusammenhang mit seiner geringen Verdampfungswärme und den damit nur schwächer aneinander gebundenen Molekülen erklärt werden. Die expliziten Modelle und das flexible MI-DMSO zeigen dagegen eine zu kleine Diffusionskonstante, besonders das HC-Modell, das am unbeweglichsten ist.

Diese Beobachtung gilt auch und verstärkt für die Diffusionszeiten der Rotation τ_1 und τ_2 . Diese wurden durch Betrachtung der Funktionen

$$p_l(\tau) = -\ln \langle P_l(\mathbf{e}_i(t + \tau) \cdot \mathbf{e}_i(t)) \rangle_{i, \Delta t}, \quad l = 1, 2 \quad (3.12)$$

bestimmt. Hier bezeichnet $\mathbf{e}_i(t)$ den Einheitsvektor entlang der S=O-Bindung des i ten Moleküls zum Zeitpunkt t und P_l die beiden ersten Legendre-Polynome

$$P_1(x) = x \quad \text{und} \quad P_2(x) = \frac{1}{2}(3x^2 - 1). \quad (3.13)$$

Die gleiche Mittelung über die Moleküle und die Teiltrajektorien wie bei Translation wurde angewendet. Anschließend ergaben sich die τ_l als die Kehrwerte der Steigung eines linearen Fits an die p_l im Zeitbereich zwischen 1 ps und dem Zeitpunkt des ersten Auftretens von $p_l(\tau) > 3$. Experimentelle Messergebnisse sind nur für τ_2 verfügbar. Die kompakten Modelle reproduzieren die Rotations-Diffusionszeit relativ gut, die expliziten Modelle sind dagegen etwas zu unbeweglich. Besonders das neu geschaffene HC-DMSO weist ein τ_2 von 16 ps auf und zeigt damit eine Rotations-Diffusion, die etwa dreimal langsamer als die gemessene und zweimal langsamer als die der beiden anderen expliziten Modelle ist.

3.2.2 Radiale Paarverteilungsfunktionen

Abbildung 3.9 zeigt die Paarverteilungsfunktionen $g_{ij}(r)$ der schweren Atome für einen Teil der betrachteten DMSO-Modelle. Die $g_{ij}(r)$ geben die Wahrscheinlichkeit an, und zwar relativ zu einem vollkommen ungeordneten System der gleichen Dichte, zwei Atome der Atomsorten i und j in einem Abstand r voneinander zu finden. Damit ermöglichen die Paarverteilungsfunktionen einen Einblick in die Ordnung eines Lösungsmittels. Sie sollen hier dazu dienen, die strukturellen Eigenschaften der zur Verfügung stehenden DMSO-Parametrisierungen untereinander und mit Ergebnissen von Neutronenstreuungsexperimenten [170] zu vergleichen. Die Paarverteilungsfunktionen wurden für die sechs möglichen Kombinationen der drei schweren Atome (O, S, C) mit Hilfe des Programmpakets XPLOR als Histogramme der Atomabstände berechnet. Als Datenquelle dienten wiederum die schon im vorangehenden Abschnitt beschriebenen Nanosekunden-Trajektorien. Bei der Berechnung der Histogramme fanden die Abstände der Atome eines Moleküls untereinander keine Berücksichtigung. Der Übersichtlichkeit halber werden in Abbildung 3.9 die $g_{ij}(r)$ nur für zwei der kompakten DMSO-Modelle gezeigt (MP und MI), da die Kurven für das VG und das OPLS-Modell sich nur geringfügig von diesen unterscheiden.

Betrachtet man zuerst die Paarverteilungsfunktionen g_{OO} und g_{SS} , so erkennt man in beiden die Schalenstruktur des DMSOs. Neben einem ersten Maximum bei $r \approx 5 \text{ \AA}$ ist eine gut ausgeprägte zweite Lösungsmittelschale im Abstand von etwa 9 \AA und sogar eine Dritte am Rand des dargestellten Bereichs zu beobachten. Deutlich zeigen diese Darstellungen, dass die expliziten Modelle (RESP, SF und HC) strukturierter sind, als diejenigen

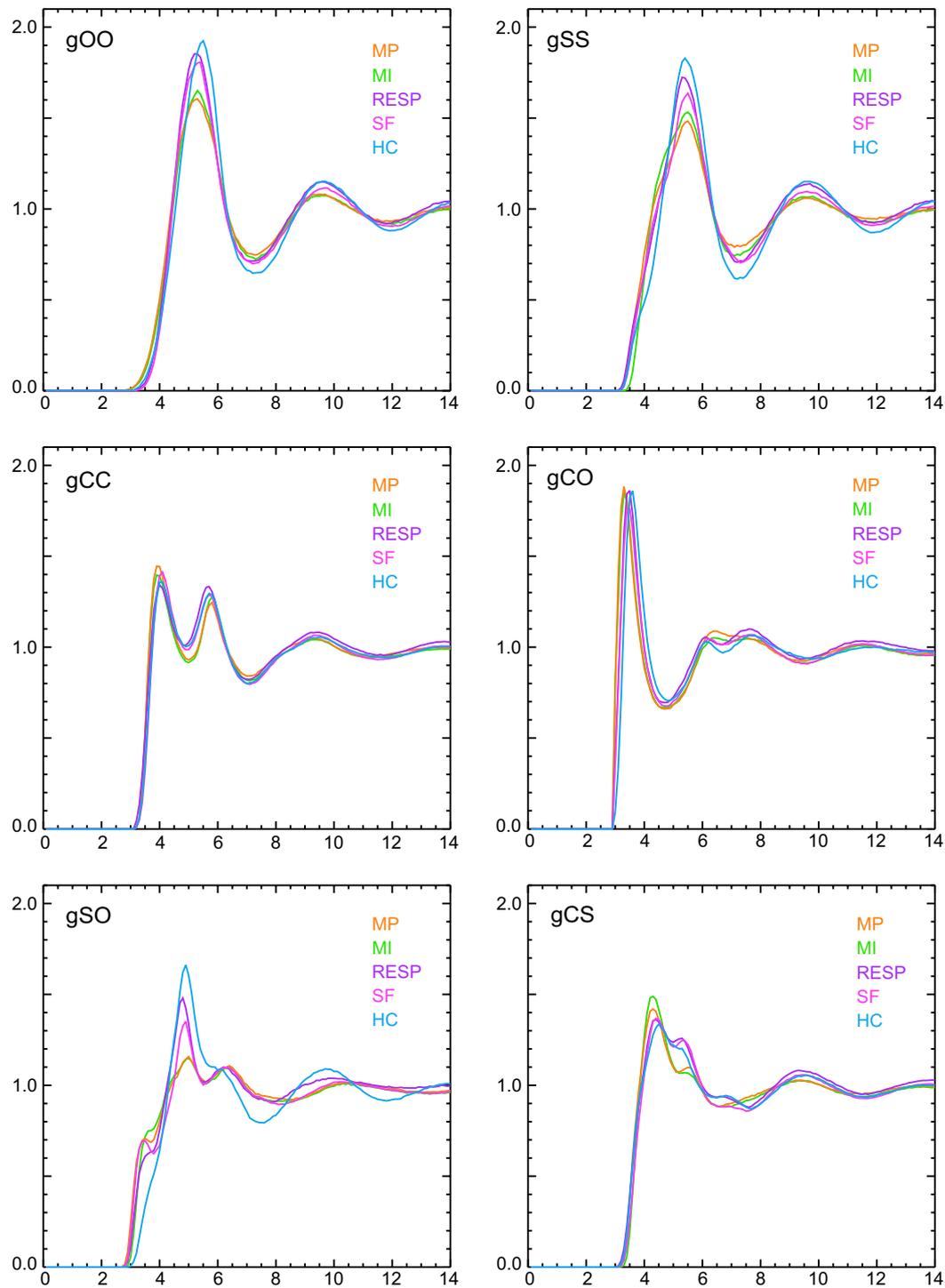


Abbildung 3.9: Die Paarverteilungsfunktionen der schweren Atome des DMSO. Für jede der sechs möglichen Kombinationen der schweren Atome wird g_{ij} über dem radialen Abstand r (in [Å]) dargestellt. Um die Übersichtlichkeit zu wahren, werden die Kurven nur für fünf der sieben betrachteten Modelle gezeigt.

| | S | O | C |
|---|--------|--------|--------|
| S | 0,0167 | | |
| O | 0,0684 | 0,0700 | |
| C | 0,1566 | 0,3208 | 0,3675 |

Tabelle 3.9: Experimentelle Gewichtungsfaktoren der Paarverteilungsfunktionen. Mit diesen sind die Paarverteilungsfunktionen der schweren Atome zu gewichten, um den Strukturfaktor der Neutronenstreuung zu erhalten. Diese experimentell ermittelten Gewichtungen wurden in der Veröffentlichung von Luzar et al. [170] in Tabelle I aufgeführt.

(MP, MI), in denen die Methylgruppen als „united-atoms“ beschrieben werden. Dagegen weist die CC-Verteilung durch das Vorkommen von zwei Kohlenstoffatomen pro Molekül eine etwas kompliziertere Struktur auf. Ihr Verlauf ist allerdings fast vollständig unabhängig von Details der betrachteten Modelle. Ähnliches gilt auch für g_{CO} und bis auf kleine Details auch für g_{CS} . In der letztgenannten Verteilungsfunktion ist allenfalls das Fehlen der Wasserstoffatome an den beiden kompakten Modellen im Bereich von 5 bis 6 Å zu bemerken.

Sehr interessant hingegen ist das Verhalten der Schwefel-Sauerstoff-Verteilungsfunktion g_{SO} , in der wieder die expliziten Modelle eine ausgeprägtere erste Schale zeigen. Besonders aber unterscheidet sich hier das HC-DMSO auffallend von den anderen Modellen. So fehlt ihm das Nebenmaximum bzw. die Schulter bei $r < 5$ Å. Weiterhin scheint dessen erste Schale etwas nach außen verschoben zu sein, so dass ein starker Überlapp mit der zweiten Schale auftritt. Für größere Radien ist für das HC-Modell auch eine stärkere Strukturierung zu erkennen. Um dies Verhalten zu verstehen, muss man die Details der in den Tabellen 3.6 und 3.7 aufgelisteten Modelle vergleichen. Unterschiede zwischen dem HC-DMSO und den beiden anderen Modellen finden sich hauptsächlich in einigen Parametern. Beispielsweise wurde für das HC-DMSO eine um etwa 3% kürzere S=O-Bindung und tendenziell schwächere Partialladungen angenommen. Anzunehmen ist jedoch, dass die Abweichungen in g_{SO} hauptsächlich in den unterschiedlichen Verhältnissen zwischen den VDW-Radien des Schwefels und des Sauerstoffs begründet liegen, die in den Parametrisierungen auftreten. So gilt das RESP-Modell ein Verhältnis von $\sigma_S/\sigma_O = 1,20$ und für das SF-Modell ein vergleichbarer Wert von 1,18. Dagegen verwendet das HC-Modell ein Verhältnis von 1,06, d.h. einen relativ sehr viel größeren Sauerstoff. Diese Beobachtung kann auch eine Erklärung für die langsame Rotations-Diffusionszeit des HC-DMSOs liefern, denn sie bedeutet auch eine geringere „Kugelförmigkeit“ dieses Modells.

Wie schon am Anfang dieses Abschnitts angedeutet wurde, sollte es durch einen Vergleich mit Ergebnissen der Neutronenstreuung möglich sein, die Paarverteilungsfunktionen und damit die DMSO-Modelle selbst zu bewerten. Dazu ist es ausreichend, die g_{ij} gewichtet aufzusummieren, um den partiellen Strukturfaktor g_{XX} zu erhalten, welcher in den Streuexperimenten bestimmt wird. Die Gewichtungsfaktoren sind als Teilergebnis der Auswertung der Experimente als bekannt vorauszusetzen und werden in Tabelle 3.9 aufgeführt.

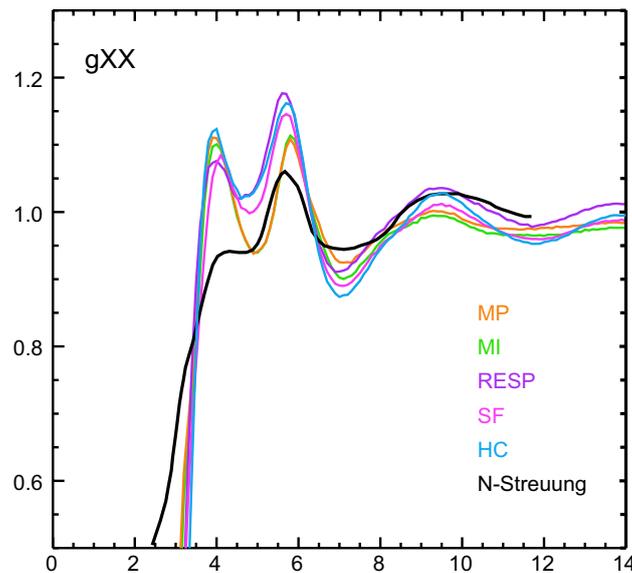


Abbildung 3.10: Partielle Strukturfaktoren aus der Simulation und der Neutronenstreuung. Die Strukturfaktoren der DMSO-Modelle sind die mit den Multiplikatoren aus Tabelle 3.9 gewichteten Paarverteilungsfunktionen, die in Abbildung 3.9 gezeigt wurden. Der Strukturfaktor der Neutronenstreuung wurde der Arbeit von Luzar et al. [170] entnommen.

Leider zeigt diese Tabelle auch, dass die Hauptbeiträge zu g_{XX} von den Verteilungen g_{CC} und g_{CO} stammen, die wie g_{CS} , dem nächst-wichtigen Beitrag, kaum sensitiv auf die Parametrisierung der Modelle reagieren. Dementsprechend zeigt Abbildung 3.10 auch sehr ähnliche Verläufe von g_{XX} für die unterschiedlichen Modelle. Insbesondere kann keine Unterscheidung zwischen den drei expliziten Modellen gemacht werden. Auffällig ist jedoch das Fehlen des ersten Maximums in der experimentellen Kurve. Dieses Maximum in den Daten der Simulationen wurde in den bisher veröffentlichten Arbeiten stets mit den dabei verwendeten „united-atom“ Modellen erklärt. Hier handelt es sich aber offenbar um eine bei allen diskutierten Modellen auftretende Eigenschaft.

3.2.3 Diskussion

Zusammenfassend stehen nach den Untersuchungen dieses Kapitels eine Reihe von verschiedenen Parametrisierungen des DMSOs als Lösungsmittel zur Verfügung. Die rigiden „united-atom“ Modelle (OPLS, VG, MP) zeigen eine zu geringe spezifische Wärme, während die flexiblen Modelle diese mehr (RESP, SF, HC) oder minder (MI) überschätzen. Der oben vermutete Zusammenhang zwischen der spezifischen Wärme eines DMSO-Modells, bzw. seiner unbekanntes mikroskopischen Wärmeleitfähigkeit, und der in den Simulationen der Dynamik des Azobenzols bzw. der Modellpeptide nach Lichtanregung beobachteten Geschwindigkeit der Wärmedissipation wird in Abschnitts 8.3 untersucht

werden.

Für Simulationen, bei denen möglichst lange Trajektorien berechnet werden sollen und bei denen zu vermuten ist, dass die Wärmedissipation des DMSO-Modells keine Rolle spielt, bietet sich das „united-atom“ Modell MP-DMSO zur Verwendung an, da es unter der vollständigen Behandlung der elektrostatischen Wechselwirkung eine relativ gute Dichte und Verdampfungswärme aufweist und einen geringeren Rechenaufwand benötigt, die deutlich „teureren“ flexiblen und expliziten Modelle.

Im Fall der flexiblen und expliziten Modellen ist das RESP-Modell durch seine zu hohe Dichte und zu geringe Verdampfungswärme auszuschließen. Das SF-Modell stand erst ganz zu Ende der Forschungsarbeit zur Verfügung, so dass alleine die neu bestimmte HC-Parametrisierung diese Klasse von Modellen vertreten kann. Wie im vorangehenden Abschnitt anhand der Paarverteilungsfunktionen gezeigt wurde, scheint das HC-Modell allerdings noch verbesserungswürdig zu sein. Um eine Veröffentlichung der Ergebnisse zu ermöglichen, wird deshalb eine erneute Optimierung der VDW-Parameter angestrebt, bei der das Verhältnis der VDW-Radien des Sauerstoffs und des Schwefels konstant gehalten werden soll. Es kann erwartet werden, dass ein solches verbessertes Modell nicht nur die Leistungsfähigkeit einer Parametrisierung aus ersten Prinzipien aufweist, sondern auch im Vergleich der thermodynamischen Observablen bestehen kann.

4 Modellierung der lichtschtaltbaren Peptide

Schon bei der Beschreibung der Ziele der vorliegenden Arbeit (Abschnitt 1.5) war angesprochen worden, dass neben einer neu zu schaffenden Modellierung des Lösungsmittels DMSO auch mehrere Ergänzungen am Kraftfeld der Modellpeptide notwendig sind, um deren lichtinduzierte Dynamik mit computergestützten Simulationen beschreiben zu können. Für den Peptidteil der ringförmigen Moleküle, deren chemische Struktur in Abbildung 1.13 gezeigt wird, fällt die Wahl des anzuwendenden Kraftfeldes leicht, da mit der in EGO-MMII implementierten Variante des CHARMM22-Kraftfeldes [81] schon eine passende Parametrisierung zur Verfügung steht. Die Form dieses Peptidkraftfeldes wurde in Kapitel 2.2 vorgestellt.

Anders verhält es sich mit den Azobenzolfarbstoffen, die als nicht-biologische Substanzen nicht von den existierenden Kraftfeldern abgedeckt werden. Auch eine Übertragung von existierenden Parametern, wie etwa derer für die in Proteinen auftretenden Phenylringe, auf Teilgruppen des Azobenzols ist nicht zulässig, da sie das ausgedehnte konjugierte π -Elektronensystem dieses Moleküls vernachlässigen würde. Daher wird im ersten Abschnitt dieses Kapitels aus ersten Prinzipien eine zum CHARMM22-Kraftfeld konforme Parametrisierung des Azobenzols, seiner Derivate (APB und AMPB) und deren Bindung an das Peptidrückgrat vorgenommen. Dabei kommen erneut diejenigen QM-Methoden zum Einsatz, die im letzten Kapitel schon erfolgreich zur Parametrisierung des DMSOs eingesetzt wurden.

Abschnitt 4.2 stellt schließlich ein molekularmechanisches Modell der Photoisomerisierung des Azobenzols vor, welches nicht nur die dabei auftretenden Geometrieänderungen der Farbstoffe, sondern auch die eingebrachten Anregungsenergien und die Isomerisierungskinetik insoweit realistisch beschreibt, als dies im Rahmen der klassischen Mechanik möglich ist. Erst diese vereinfachende klassische Beschreibung ermöglicht die Untersuchung der Pikosekunden-Dynamik der Modellpeptide nach Lichtanregung, da es bislang keine hinreichend exakten und recheneffizienten QM-Methoden zur Beschreibung von Photoisomerisierungsreaktionen in großen organischen Farbstoffen, wie den Azobenzolen, gibt.

4.1 Kraftfeldparameter für die Azobenzolfarbstoffe

Zur Festlegung der Geometrie, der Partiaalladungen und der Kraftkonstanten der Azobenzolfarbstoffe wurden DFT-Berechnungen am freien Azobenzol und an zwei Modellmolekülen durchgeführt, deren Strukturen in Abbildung 4.1 skizziert sind. Diese beiden Analoga sollen die durch Peptidbindungen mit dem Peptidrückgrat verbundenen Azobenzolderivate APB und AMPB repräsentieren. Aus einer Reihe von Veröffentlichungen ist bekannt, dass die Dichtefunktional-Theorie die beobachtete Struktur und die Schwingungsspektren des freien Azobenzols und seiner Derivate quantitativ beschreiben kann [205–207]. In der Literatur wird intensiv über die Planarität des *trans*-Azobenzols diskutiert, welche von DFT-Studien vorhergesagt wird [206], während MP2-Berechnungen eine nicht-planare Struktur zeigten [208, 209]. Da experimentelle Untersuchungen mit Elektronenstreuung [208] und pikosekunden-zeitaufgelöste Raman-Spektren [210] ebenfalls auf die Planarität hinweisen und ein MM-Kraftfeld des Azobenzols aus Symmetriegründen nur eine planare Struktur parametrisieren kann, wurde diese im Folgenden angenommen.

Alle hier an Azobenzol und seinen Derivaten vorgenommenen Rechnungen basierten auf dem B3LYP-Funktional und dem Basissatz 6-31G*. Diese Kombination wurde von Biswas und Umaphy [205] als diejenige bestimmt, welche die beste Übereinstimmung mit den experimentell bestimmten Strukturen zeigte. Alle Rechnungen wurden an einzelnen Molekülen im Vakuum durchgeführt. Das Programmpaket GAUSSIAN [68] erlaubte auf einfache Weise nicht nur die Bestimmung der Geometrie der Moleküle, sondern auch die Berechnung von Potentialflächen als Funktion der internen Koordinaten (Bindungslängen, Winkel, Torsionswinkel). In einem ersten Schritt wurden die Geometrien der drei in Abbildung 4.1 gezeigten Moleküle ausgehend von beliebigen Startstrukturen optimiert, jeweils getrennt für ihr *cis*- und ihr *trans*-Isomer.

Für Azobenzol zeigt Abbildung 1.12 die berechneten Geometrien, deren Bindungslängen und -winkel die zuvor veröffentlichten Werte reproduzieren [205]. Wie Tabelle 4.1 auführt, wurden die in den optimierten Geometrien des *trans*-Azobenzols, des *trans*-APBs und des *trans*-AMPBs beobachteten Abstände und Winkel direkt als die Gleichgewichtslängen und -winkel des Kraftfeldes übernommen. Die Wahl der *trans*-Strukturen als Referenz wurde hier durch das Auftreten sterischer Hinderungen in den *cis*-Strukturen und durch die Beobachtung nur kleiner Änderungen der Bindungslängen beim Übergang von *trans* nach *cis* motiviert. So verkürzt sich die berechnete N=N Bindungslänge beim Übergang vom *trans*-Azobenzol zu seinem *cis*-Isomer um 0,012 Å von 1,261 Å auf 1,249 Å, während die C–C Bindungen sich um 0,017 Å von 1,419 Å auf 1,436 Å verlängern. Dies entspricht in beiden Fällen relativen Änderungen von etwa 1%, die im Rahmen der groben Approximation, welche ein molekularmechanisches Kraftfeld darstellt, vernachlässigbar sind. Aus den gleichen Überlegungen und weil sich die Geometrien der Azo-Gruppe und der Phenylringe unter Austausch der in Abbildung 4.1 gezeigten Seitengruppen nur sehr schwach änderten, wurde die Geometrie des freien Azobenzols gleichermaßen auch für die dem Azobenzol äquivalenten Teile seiner Derivate APB und AMPB angenommen.

4.1 Kraftfeldparameter für die Azobenzolfarbstoffe

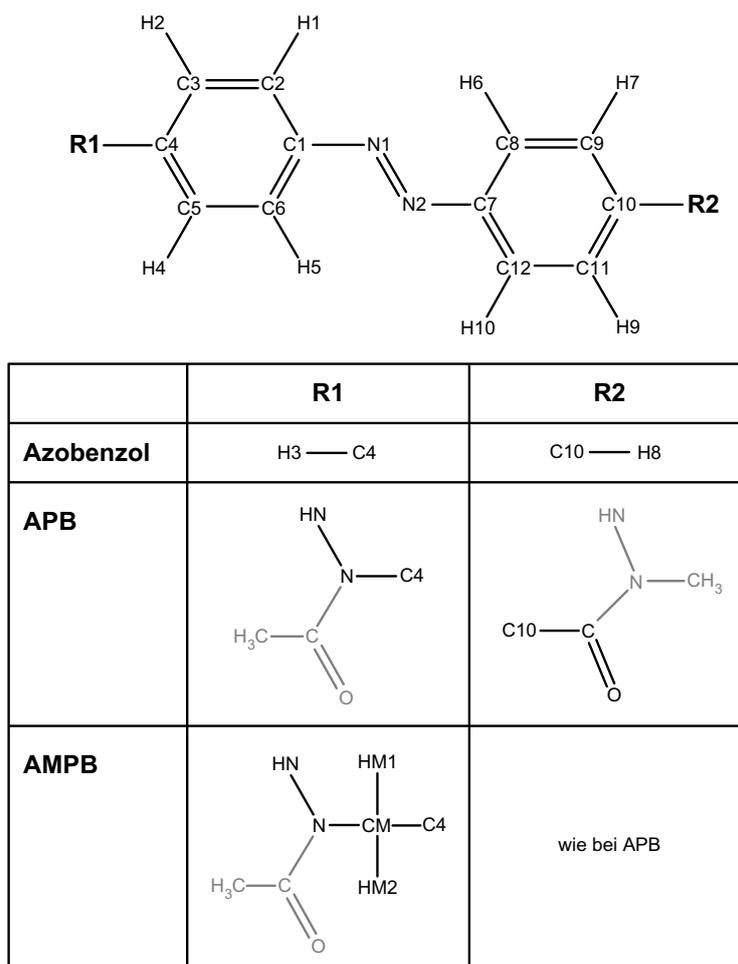


Abbildung 4.1: Die Struktur und die Atomnamen der Azobenzolfarbstoffe. Die Topologie des Azobenzols und seiner Derivate APB und AMPB unterscheidet sich durch die beiden Atomgruppen R1 und R2. Im Fall des freien Azobenzols werden diese lediglich durch zwei Wasserstoffe gebildet. Ein APB-Residuum umfasst, analog zu der Struktur der herkömmlichen Aminosäuren, jeweils eine „halbe“ Peptidgruppe, die dunkel hervorgehoben ist. Die grau dargestellten Atome sind deshalb kein Teil des APB-Residuums, sondern wurden in den DFT-Rechnungen mitberücksichtigt, um den Einfluss der vollständigen Peptidbindungen auf die Struktur und die Energetik des in das Peptidrückgrat eingebauten Farbstoffes zu erfassen. Offene Bindungen wurden mit Methylgruppen abgesättigt. Wie schon in Abbildung 1.13 gezeigt wurde, weist das AMPB im Vergleich zum APB eine zusätzliche Methylengruppe auf, welche die N-terminale Peptidbindung vom Farbstoff trennt. Der untere Teil der Abbildung benennt die zusätzlichen Atome der Methylengruppe.

| Bindung | | | | Winkel | | |
|-----------|--------------------|--------------------|-------------|--------------------|--------------------|--|
| | k_b | b_0 | | k_θ | θ_0 | |
| Azobenzol | | | | Azobenzol | | |
| CAZ CAZ | 419,4 ^c | 1,398 ^a | CAZ CAZ CAZ | 40,0 ^d | 120,0 ^e | |
| NAZ CAZ | 341,7 ^b | 1,419 ^a | HAZ CAZ CAZ | 34,4 ^b | 120,0 ^e | |
| HAZ CAZ | 404,7 ^b | 1,086 ^a | NAZ CAZ CAZ | 60,2 ^b | 120,0 ^e | |
| NAZ NAZ | 716,8 ^b | 1,261 ^a | NAZ NAZ CAZ | 121,1 ^b | 114,8 ^a | |
| APB/AMPB | | | | APB/AMPB | | |
| C CAZ | 250,0 ^d | 1,504 ^a | C CAZ CAZ | 45,8 ^d | 120,0 ^d | |
| NH1 CAZ | 320,0 ^d | 1,405 ^a | O C CAZ | 80,0 ^d | 121,0 ^d | |
| CT2 CAZ | 230,0 ^d | 1,526 ^a | NH1 C CAZ | 80,0 ^d | 116,5 ^d | |
| | | | NH1 CAZ CAZ | 70,0 ^d | 120,0 ^e | |
| | | | H NH1 CAZ | 35,0 ^d | 114,8 ^a | |
| | | | C NH1 CAZ | 50,0 ^d | 120,0 ^d | |
| | | | HA CT2 CAZ | 49,3 ^d | 107,5 ^d | |
| | | | CT2 CAZ CAZ | 45,8 ^d | 120,0 ^d | |
| | | | NH1 CT2 CAZ | 50,0 ^d | 116,3 ^d | |

Tabelle 4.1: Kraftfeldparameter der Bindungslängen und -winkel. Für das freie Azobenzol und seine Derivate APB bzw. AMPB werden in der linken Tabelle für alle zwischen den Atomtypen auftretenden Bindungen die Kraftkonstanten k_b [$\text{kcal mol}^{-1} \text{\AA}^{-2}$] und die zugehörigen Gleichgewichtslängen b_0 [\AA] angegeben. Die Zuordnung der Atomnamen aus Abbildung 4.1 zu den Atomtypen erfolgt in Tabelle 4.2. Analog zeigt die rechte Tabelle die Kraftkonstanten der Winkel k_θ [$\text{kcal mol}^{-1} \text{Rad}^{-2}$] und die Gleichgewichtswinkel θ_0 [Grad]. Die Fußnoten geben jeweils die Vorgehensweise an, die zur Bestimmung des betreffenden Wertes angewendet wurde:

^a Abstand oder Winkel wurden den DFT-optimierten *trans*-Strukturen des Azobenzols bzw. der APB- und AMPB-Analoga entnommen. ^b Die Kraftkonstante wurde durch einen Fit an die DFT-berechnete Potentialfläche festgelegt. ^c Die Kraftkonstante der Bindung wurde aus der Projektion der kartesischen Hesse-Matrix auf die internen Koordinaten der Bindungslängen bestimmt. ^d Entnommen aus ähnlichen Parametern des CHARMM22-Kraftfeldes. ^e Die Winkel von 120° wurden gewählt, um den Symmetrien des Moleküls zu genügen.

Unter Zugrundelegung der optimierten Geometrien wurden anschließend die Partialladungen der Atome mit dem Verfahren der ESP-Ladungen (siehe Abschnitt 3.1.1) ermittelt. Die Auswahl der Raumpunkte, an denen das elektrostatische Potential der Moleküle betrachtet wurde, geschah durch den als stabil bekannten CHelpG-Algorithmus [211]. Für jedes der drei betrachteten Moleküle wurden auf diese Weise die Partialladungen getrennt in der *cis*- und in der *trans*-Konfiguration bestimmt, anschließend jedoch bezüglich des Isomerisierungszustandes gemittelt, um eine Beschreibung zu finden, die gleichermaßen für beide Isomere gültig ist.

Hier bleibt anzumerken, dass eine realistischere Beschreibung der Elektrostatik von polarisierbaren Gruppen, wie den Azobenzolfarbstoffen oder den Peptidplättchen der Proteine, nur mit Hilfe von Kraftfeldern erreicht werden kann, welche elektronische Polarisierbarkeiten explizit berücksichtigen. Solche Kraftfelder werden seit einigen Jahren von ver-

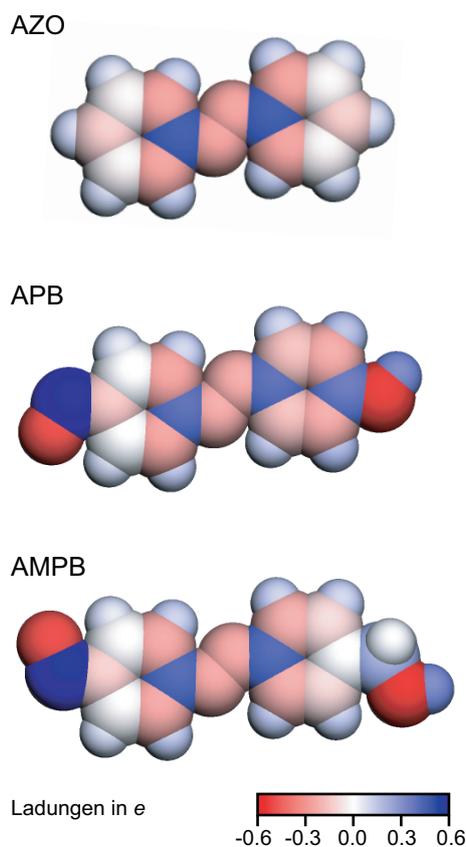


Abbildung 4.2: Eine farbkodierte Darstellung der MM-Partialladungen. Die in Tabelle 4.2 aufgeführten Partialladungen der Azobenzolfarbstoffe werden in der Farbskala der Atome des Kollontenmodells visualisiert.

schiedenen Gruppen entwickelt [212–214]. Weiterhin konnte nachgewiesen werden, dass solche Parametrisierungen grundsätzlich besser die Ladungsverschiebungen repräsentieren, welche als Folge einer strukturellen Umordnung der Atome eines Moleküls auftreten [215]. Die in dieser Arbeit betrachteten Isomerisierungsreaktionen oder auch die zu untersuchenden Konformationsübergänge des Rückgrats der Modellpeptide stellen genau solche strukturellen Umordnungen dar. Demzufolge wäre die Verwendung eines polarisierbaren Kraftfeldes anzustreben. Leider befinden sich diese Kraftfelder in einem frühen Stadium ihrer Entwicklung und konnten bisher nicht ausreichend getestet werden, was besonders problematisch ist, da sie auf einer, verglichen mit den schon kompliziert zu definierenden herkömmlichen Kraftfeldern, noch größeren Zahl von Parametern beruhen. Somit stellt das hier vorgestellte Vorgehen der mittleren Partialladungen die zur Zeit bestmögliche Lösung dar.

Tabelle 4.2 führt die Partialladungen für jedes Atom der drei Azobenzolderivate auf. Diese Daten sind in Abbildung 4.2 anschaulich dargestellt. In dieser Abbildung wird deutlich, dass die Partialladungen neben der besprochenen Mittelung über die beiden Isomere auch

| Atomname | Atomtyp | Partialladungen [e] | | |
|------------|---------|---------------------|---------|---------|
| | | Azobenzol | APB | AMPB |
| C1 | CAZ | 0,4543 | 0,4028 | 0,4204 |
| C2 | CAZ | -0,2565 | -0,2561 | -0,2477 |
| C3 | CAZ | -0,0419 | -0,0145 | -0,0221 |
| C4 | CAZ | -0,1352 | -0,1647 | -0,1439 |
| C5 | CAZ | -0,0419 | -0,0145 | -0,0221 |
| C6 | CAZ | -0,2565 | -0,2561 | -0,2477 |
| N1 | NAZ | -0,2564 | -0,2439 | -0,2462 |
| N2 | NAZ | -0,2564 | -0,2439 | -0,2462 |
| C7 | CAZ | 0,4543 | 0,4028 | 0,4204 |
| C8 | CAZ | -0,2565 | -0,1803 | -0,2250 |
| C9 | CAZ | -0,0419 | -0,2389 | -0,1113 |
| C10 | CAZ | -0,1352 | 0,3665 | 0,0004 |
| C11 | CAZ | -0,0419 | -0,2389 | -0,1113 |
| C12 | CAZ | -0,2565 | -0,1803 | -0,2250 |
| H1 | HAZ | 0,1229 | 0,1202 | 0,1251 |
| H2 | HAZ | 0,0946 | 0,0836 | 0,0823 |
| H3 | HAZ | 0,0991 | | |
| H4 | HAZ | 0,0946 | 0,0836 | 0,0823 |
| H5 | HAZ | 0,1229 | 0,1202 | 0,1251 |
| H6 | HAZ | 0,1229 | 0,1202 | 0,1251 |
| H7 | HAZ | 0,0946 | 0,1654 | 0,1191 |
| H8 | HAZ | 0,0991 | | |
| H9 | HAZ | 0,0946 | 0,1654 | 0,1191 |
| H10 | HAZ | 0,1229 | 0,1202 | 0,1251 |
| N-terminus | | | | |
| N | NH1 | | -0,5293 | -0,5351 |
| HN | H | | 0,3192 | 0,3182 |
| CM | CT2 | | | 0,1969 |
| HM1 | HA | | | 0,0263 |
| HM2 | HA | | | 0,0263 |
| C-terminus | | | | |
| C | C | | 0,5930 | 0,5723 |
| O | O | | -0,5017 | -0,5008 |

Tabelle 4.2: Atomtypen und Partialladungen der Azobenzolfarbstoffe. Den in Abbildung 4.1 benannten Atomen werden in dieser Tabelle die jeweiligen Atomtypen zugeordnet. Die Atomtypen NAZ, CAZ und HAZ werden zusätzlich zu denen aus CHARMM22 eingeführt. Weiterhin werden die Partialladungen für den Fall der Parametrisierung des freien Azobenzols und seiner zwei Derivate aufgeführt. Diese Ladungen gelten sowohl für die *cis*- als auch die *trans*-Isomere.

4.1 Kraftfeldparameter für die Azobenzolfarbstoffe

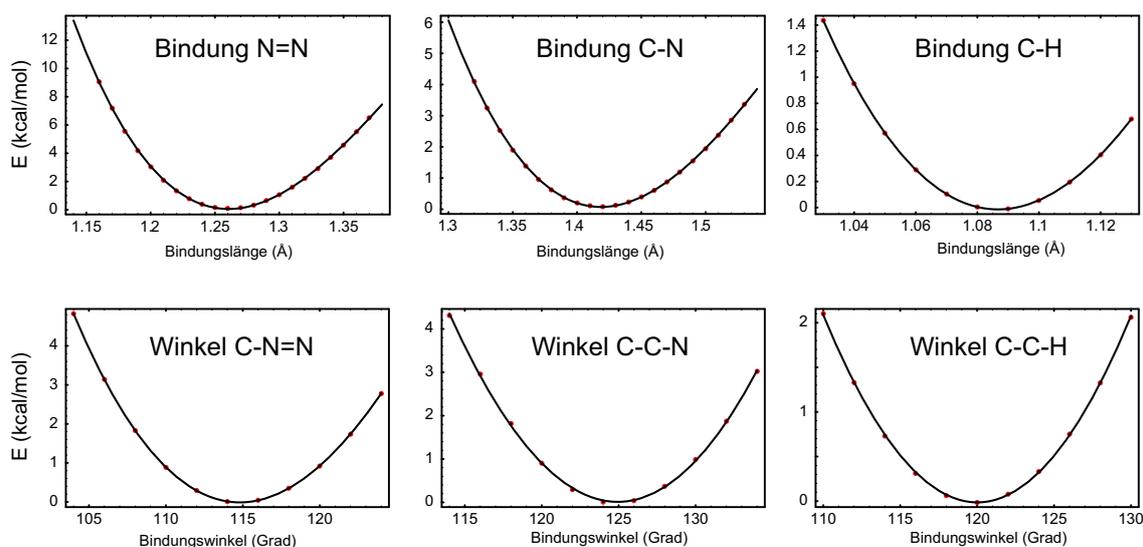


Abbildung 4.3: Die Potentialflächen der Bindungen und Winkel des Azobenzols. Die roten Punkte stellen die Gesamtenergien des Azobenzols dar, die unter Vorgabe einer der gezeigten internen Koordinaten und Relaxation der übrigen Freiheitsgrade mit Hilfe der DFT berechnet wurden. Die durchgezogenen Linien sind Fits von Polynomen dritter Ordnung an diese Datenpunkte.

hinsichtlich der in den Phenylringen auftretenden Symmetrien gemittelt wurden. So ist in Abbildung 4.1 beispielsweise offensichtlich, dass die Wasserstoffatome H2 und H4 zueinander äquivalent sind und deshalb mit den gleichen Partialladungen parametrisiert werden müssen. Abbildung 4.2 verdeutlicht die starke Elektronegativität der beiden Stickstoffatome der Azogruppe, die zu positiven Partialladungen an den beiden angebondenen Kohlenstoffatomen führt. Aus der entstehenden Ladungsverteilung folgt für das freie Azobenzol ein Dipolmoment in der *cis*-Form, aber keines im Fall des *trans*-Isomers. Betrachtet man die Derivate des Azobenzols, so sieht man, dass die Azogruppe durch die Substitution der beiden äußeren Wasserstoffe kaum beeinflusst wird, die Anbindung der rechten Peptidgruppe (sozusagen am N-Terminus des Farbstoffs) aber beim APB zu einer starken Ladungsverschiebung im rechten Phenylring führt. Die Anbindung der C-terminalen Peptidgruppe scheint dagegen nur einen geringen Einfluss auszuüben. Die Form der Molekülorbitale des APBs, wie sie aus den DFT-Rechnungen bestimmt werden können (hier nicht gezeigt), deutet an, dass sich das ausgedehnte π -Elektronensystem, welches das Azobenzol und seine Derivate erst zu Farbstoffen macht, sich ununterbrochen auch über die beiden angeschlossenen Peptidgruppen erstreckt. Dieser Effekt wird in dem molekularmechanischen Kraftfeld aus Konsistenzgründen vernachlässigt. Folglich werden die beiden angebondenen Peptidgruppen mit den gleichen Parametern beschrieben, wie alle übrigen Peptidgruppen.

4 Modellierung der lichtschtbaren Peptide

| Torsionswinkel | | | | k_ϕ^A | k_ϕ^B | n | ϕ_0 |
|----------------|-----|-----|-----|--------------------|--------------------|-----|----------|
| Azobenzol | | | | | | | |
| CAZ | NAZ | NAZ | CAZ | 12,58 ^a | 21,60 ^b | 2 | 180,0 |
| | | | | - | 4,36 ^b | 4 | 0,0 |
| NAZ | NAZ | CAZ | CAZ | 0,86 ^a | 1,67 ^b | 2 | 180,0 |
| | | | | - | 0,23 ^b | 4 | 0,0 |
| CAZ | CAZ | CAZ | CAZ | 3,10 ^d | 3,10 ^d | 2 | 180,0 |
| HAZ | CAZ | CAZ | X | 3,50 ^d | 3,50 ^d | 2 | 180,0 |
| NAZ | CAZ | CAZ | CAZ | 3,10 ^d | 3,10 ^d | 2 | 180,0 |
| APB/AMPB | | | | | | | |
| CAZ | NAZ | NAZ | CAZ | 12,47 ^a | 20,62 ^b | 2 | 180,0 |
| | | | | - | 3,19 ^b | 4 | 0,0 |
| NAZ | NAZ | CAZ | CAZ | 1,29 ^a | 2,03 ^b | 2 | 180,0 |
| | | | | - | 0,18 ^b | 4 | 0,0 |
| CAZ | CAZ | C | X | 0,85 ^d | 0,55 ^b | 2 | 180,0 |
| | | | | - | 0,13 ^b | 4 | 0,0 |
| C | CAZ | CAZ | CAZ | 3,10 ^d | 3,10 ^d | 2 | 180,0 |
| CAZ | C | NH1 | CT1 | 1,60 ^d | 1,60 ^d | 1 | 0,0 |
| | | | | 2,50 ^d | 2,50 ^d | 2 | 180,0 |
| CAZ | C | NH1 | H | 2,50 ^d | 2,50 ^d | 2 | 180,0 |
| CAZ | CAZ | NH1 | X | 0,42 ^a | 0,71 ^b | 2 | 180,0 |
| | | | | - | 0,13 ^b | 4 | 0,0 |
| NH1 | CAZ | CAZ | CAZ | 3,10 ^d | 3,10 ^d | 2 | 180,0 |
| CT1 | C | NH1 | CAZ | 1,60 ^d | 1,60 ^d | 1 | 0,0 |
| | | | | 2,50 ^d | 2,50 ^d | 2 | 180,0 |
| O | C | NH1 | CAZ | 2,50 ^d | 2,50 ^d | 2 | 180,0 |

Tabelle 4.3: Kraftfeldparameter der Torsionswinkel. Die Parameter werden für Azobenzol und seine beiden Derivate angegeben. Es wird sowohl die ältere (*A*) als auch die neuere (*B*) Parametrisierung aufgeführt. Eine Diskussion dazu findet sich im Text. In Notation der Gleichung (2.17) werden die Kraftkonstanten k_ϕ [kcal mol⁻¹], die Periodizitäten n und die Phasenverschiebungen δ [Grad] angegeben. Die Parametrisierung *B* benutzt teilweise mehrere Terme in der Kosinus-Entwicklung, diese werden als zusätzliche Zeilen in der Tabelle gezeigt. Die Herkunft der einzelnen Werte wird durch die folgenden Fußnoten erklärt: ^a Fit eines einzelnen Kosinus an die Potentialfläche in der nahen Umgebung des Minimums. ^b Fit einer Kosinus-Entwicklung an die gesamte Potentialfläche, unter Berücksichtigung der Torsionsbarrieren. ^d Entnommen aus ähnlichen Parametern des CHARMM22-Kraftfeldes.

Kraftkonstanten

Die in den Tabellen 4.1 und 4.3 aufgeführten Kraftkonstanten wurden, wie die Fußnoten angeben, auf unterschiedlichen Wegen festgelegt. So wurden einige der Werte aus dem CHARMM22-Kraftfeld der Peptide übernommen, sofern dort ähnliche chemische Umgebungen behandelt werden. Dies trifft beispielsweise auf die Methylengruppe des AMPBs zu, bei der aufgrund der Einfachbindungen zu den Nachbargruppen davon ausgegangen werden kann, dass ihre Struktur und Energetik kaum von der chemischen Natur dieser Nachbarn abhängt. Dagegen wurden die Kraftkonstanten des Azobenzol-Residuums explizit aus der Betrachtung der potentiellen Energiefläche abgeleitet. Aufwändige DFT-Berechnungen wurden durchgeführt, in denen für eine Reihe von festgelegten Werten der entsprechenden internen Koordinate eine Relaxation der übrigen Freiheitsgrade zugelassen wurde.

Die Ergebnisse dieser Optimierungen der Molekülgeometrie unter der Zwangsbedingung jeweils einer festgehaltenen Koordinate zeigt Abbildung 4.3 für einen Teil der Bindungslängen und -winkel. Durch die Relaxation des restlichen Moleküls entsprechen die dargestellten Potentialkurven denjenigen Einzeltermen, welche in einem MM-Kraftfeld den Beitrag der zugehörigen internen Koordinaten repräsentieren. Bei der Abtastung der Potentialkurven wurden auch größere Auslenkungen erzeugt, so dass Anharmonizitäten auftraten. Aus diesem Grund wurde an die berechneten Punkte jeweils ein Polynom dritten Grades angepasst, anschließend aber in Übereinstimmung mit der Form des Kraftfeldes (siehe Gleichungen (2.15) und (2.16) nur die quadratischen Terme berücksichtigt. Ein alternatives Vorgehen wäre an dieser Stelle die Selektion derjenigen Datenpunkte, deren Energien unter einer Grenze von der Größenordnung von $k_B T$ liegen, und somit desjenigen Bereichs der Energiefläche, an dem das System eine nicht-verschwindende Aufenthaltswahrscheinlichkeit aufweist. Ein quadratischer Fit an diese ausgewählten Punkte liefert beispielsweise für die N=N Doppelbindung eine leicht schwächere Kraftkonstante von 700 statt 717 kcal/(mol·Å²) und für die C–C Bindung 332 statt 342 kcal/(mol·Å²). Da beide Vorgehensweisen gleichermaßen berechtigt sind und erstere keine willkürliche Wahl eines Energiebereichs benötigt, zeigt Tabelle 4.1 Werte, die aus dem Fit des Polynoms dritter Ordnung entstanden sind.

Für die Dehnung der C=C Bindungen innerhalb der Phenylringe konnte dieses Verfahren nicht angewendet werden, vielmehr wurde für diesen Parameter die in kartesischen Koordinaten berechnete Hessematrix mittels des Programms GAMESS [216] auf den Koordinatensatz der Bindungslängen projiziert. Die dabei resultierenden 410 kcal/(mol·Å²) liegen über dem Wert von 305 kcal/(mol·Å²), den CHARMM22 für Phenylgruppen annimmt. Ein ähnliches Verhalten ist auch bei den Winkelstärken zu beobachten, die beispielsweise für die Wasserstoffe der Phenylringe um etwa 10% im Vergleich zu CHARMM22 erhöht ausfallen. Vergleichswerte für die Eigenschaften der Azo-Gruppe existieren in dem Peptid-Kraftfeld natürlich nicht.

Auch im Fall der Torsionswinkel möchte ich zwei unterschiedliche Ansätze zur Bestimmung der Kraftkonstanten diskutieren, die chronologisch aufeinander folgten. Beide basieren, wie schon zuvor, auf der Betrachtung der Potentialfläche als Funktion einer internen Koordinate, hier der Dihedralwinkel. Wiederrum wurden die Potentialkurven mit Hilfe der DFT unter Relaxation der übrigen Freiheitsgrade der Moleküle berechnet. Die Berechnungen wurden nicht nur am freien Azobenzol durchgeführt, sondern auch am APB, um den etwaigen Einfluss der angebundenen Peptidgruppen auf die Steifigkeit der Azogruppe zu untersuchen. Das erste Vorgehen und die mit ihm bestimmten Kraftkonstanten bezeichne ich im folgenden mit dem Buchstaben „A“. In diesem ersten Ansatz wurden die Potentialkurven in der unmittelbaren Umgebung der Gleichgewichtslage des Dihedralwinkels in der *trans*-Struktur der Farbstoffe betrachtet. In Abbildung 4.4 sind die betreffenden Datenpunkte blau hervorgehoben. An diese Punkte wurde anschließend ein einzelner Kosinusterm angepasst (vergleiche Gleichung (2.17)), wobei, wie Tabelle 4.3 aufführt, stets eine zweifache Periodizität der Winkel angenommen wurde. Wie Abbildung 4.4 weiterhin zeigt, konnte das Verfahren A am C-Terminus des APBs nicht angewendet werden, da dort in der planaren Struktur des Moleküls eine leichte sterische Hinderung auftritt. Stattdessen wurden für diesen Winkel Parameter aus dem CHARMM22-Kraftfeld übernommen.

Nachdem erste Simulationen der zyklischen Peptide unter Verwendung der Kraftfeldparameter A durchgeführt und veröffentlicht wurden [138], stellte sich die Frage, ob diese Parameter in der Lage sind, die Torsionsbarrieren der Dihedralwinkel zu reproduzieren. Zur Beantwortung dieser Frage wurden in zusätzlichen Berechnungen die Potentialkurven über größere Winkelbereiche abgetastet, welche ausgehend vom Minimum auch die Winkel maximaler Energie abdecken. So zeigt erwartungsgemäß die Doppelbindung C=C=N-C in Abbildung 4.4 eine hohe Barriere zwischen dem *trans*- (180°) und dem *cis*-Isomer (0° bzw. 360°). In der *cis*-Form tritt erneut eine sterische Hinderung auf, hier verursacht durch die Kollision der beiden Phenylringe. Diese Hinderung führt beim Azobenzol dazu, dass das *cis*-Isomer energetisch $15,7$ kcal/mol über dem *trans*-Isomer liegt. Die Rechnungen am APB ergaben eine ähnliche Energiedifferenz von $16,1$ kcal/mol. Aufbauend auf den umfassenderen Berechnungen an den Potentialflächen wurden im Vorgehen B alle in Abbildung 4.4 blau oder rot gefärbten Punkte berücksichtigt und an diese eine Kosinus-Entwicklung mit zwei Termen angepasst. Tabelle 4.3 zeigt die Kraftkonstanten der dabei verwendeten doppelt- und vierfach-periodischen Teilbeiträge. Abbildung 4.4 verdeutlicht, dass die so parametrisierten Funktionen die Torsionsbarrieren besser beschreiben, aber auch etwas flachere Verläufe in der Nähe der Minima aufweisen. Diese geringen Unterschiede werden sich später noch als sehr wichtig erweisen.

4.2 Ein MM-Modell der Isomerisierung

Eine theoretische Beschreibung der durch Lichtanregung ausgelösten Isomerisierungsreaktion von Azobenzol muss die elektronisch angeregten Zustände des Farbstoffs berück-

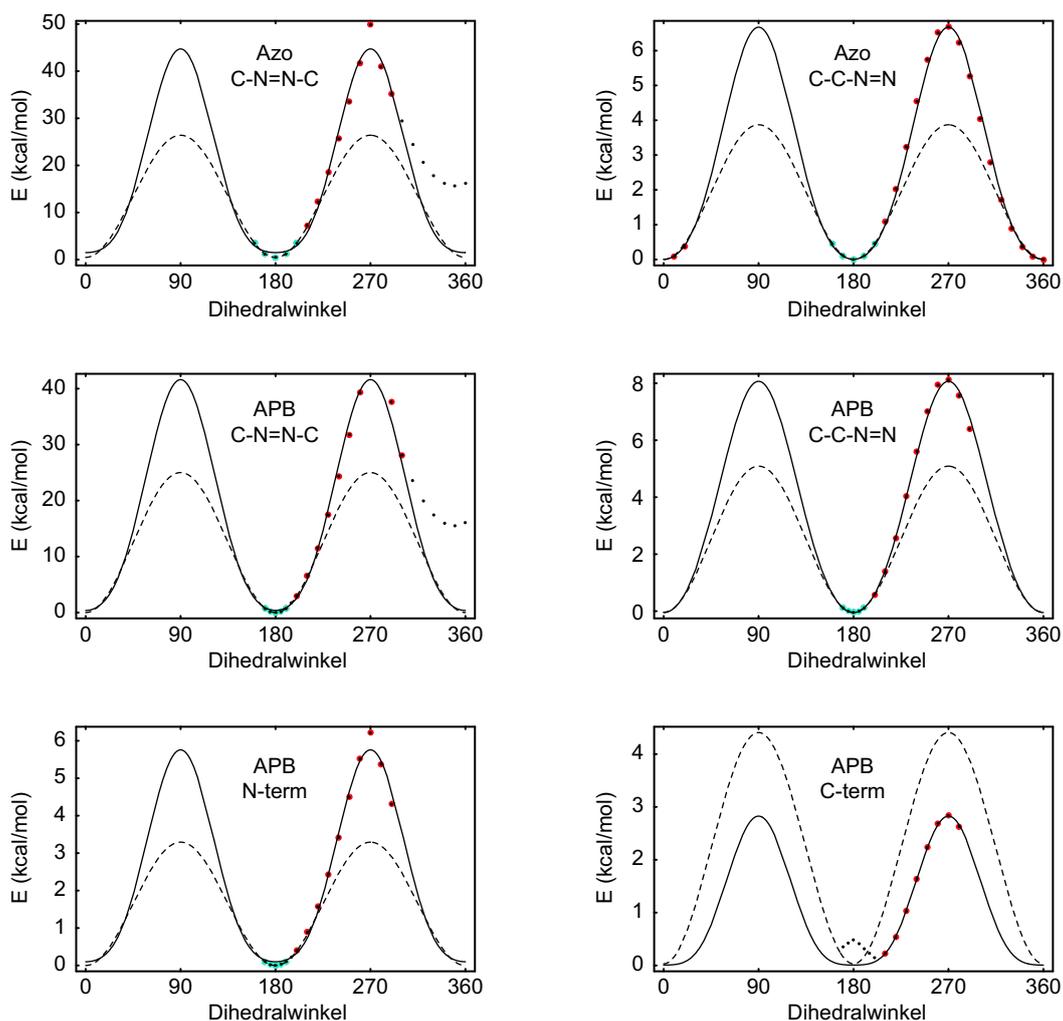


Abbildung 4.4: Die Potentialkurven der Dihedralwinkel. Für das freie Azobenzol und das APB wurde die Potentialfläche im elektronischen Grundzustand durch DFT-Rechnungen abgetastet. Schwarze Punkte stellen die einzelnen Datenpunkte dar. Die blau eingefärbten Punkte fanden Berücksichtigung beim Fit der gestrichelt dargestellten einfachen Kosinus-Funktionen (Verfahren A, siehe Text). Das Verfahren B resultierte in den durchgezogen gezeichneten Kurven und basiert neben den blauen auch auf den roten Datenpunkten. Von den sechs Winkeln, bezeichnet C–C=N–C jeweils die Torsion um die Doppelbindung. Die beiden Minima entsprechen den *cis*- bzw. *trans*-Isomeren der Moleküle. Zur Untersuchung der anderen Winkel wurde stets die *trans*-Struktur zu Grunde gelegt. Mit C–C–C=N ist die Rotation der Phenylringe in Bezug auf die Azo-Gruppe bezeichnet und N-term bzw. C-term betrifft die von den Atomtypen CAZ–CAZ–NH1–X bzw. CAZ–CAZ–C–X gebildeten Torsionswinkel (siehe Abbildung 4.1).

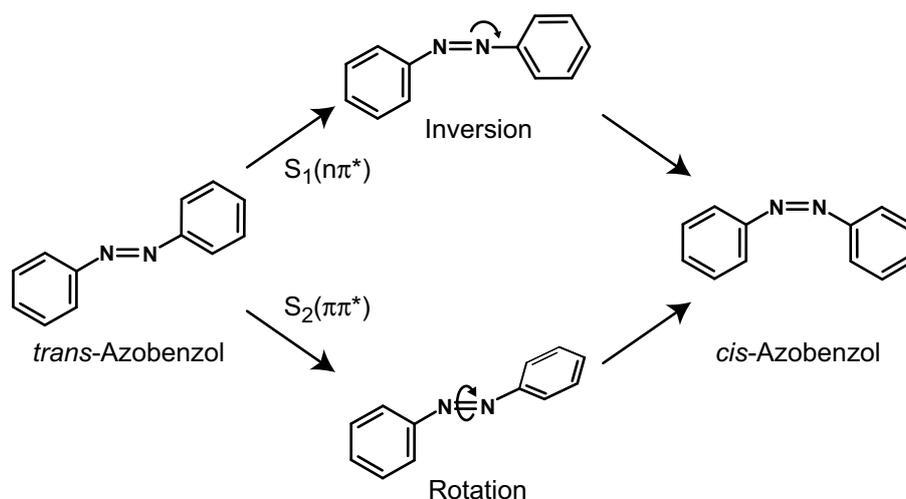


Abbildung 4.5: Die diskutierten Reaktionspfade der Isomerisierung des Azobenzols. Je nach Wellenlänge des Anregungspulses und des damit verbundenen Übergangs des Azobenzols in den elektronisch angeregten S_1 oder S_2 Zustand läuft die darauf folgende Isomerisierungsreaktion in Form einer Inversion an einem der beiden Stickstoffatome oder einer Rotation um die N=N-Bindung ab [217]. Diskussion im Text.

sichtigen. Ein entsprechender Ansatz wurde schon in der Einführung (Abschnitt 1.3) erwähnt. Die angeregten Zustände des Azobenzols waren bis heute Gegenstand einer Reihe von Untersuchungen. So berechneten Astrand et al. [218] für das Azobenzol und mehrere seiner Derivate durch Einzelpunktrechnungen die Anregungsenergien zwischen dem Grundzustand S_0 und den beiden ersten angeregten Singulett-Zuständen S_1 und S_2 . In anderen Arbeiten wurden zusätzlich die Potentialflächen des S_1 , des S_2 und anderer angeregter Zustände entlang gewisser eindimensionaler [217, 219] bzw. zweidimensionaler [220] Reaktionskoordinaten abgetastet. Dabei konnten die experimentell bestimmten vertikalen Anregungsenergien reproduziert werden, es wurden aber unterschiedliche Verläufe der Potentialflächen gefunden. Die genaue Form dieser Flächen ist jedoch wichtig für die Bestimmung der Reaktionskoordinate der Photoisomerisierung des Azobenzols.

Abbildung 4.5 zeigt die beiden in diesem Zusammenhang diskutierten Reaktionspfade: eine Rotation um die N=N-Bindung und eine Inversion an einem der beiden Stickstoffe. Im Folgenden bezeichne ich mit n das energetisch am höchsten liegende nichtbindende Molekülorbital, mit π das höchste besetzte π -Orbital und mit π^* das niedrigste unbesetzte π -Orbital. In dieser Schreibweise besteht eine Anregung in den S_2 hauptsächlich aus einem $\pi \rightarrow \pi^*$ Übergang. Die zuvor zitierten theoretischen Arbeiten stimmen darin überein, dass im Fall der Anregung in den S_2 die Isomerisierungsreaktion als Rotation abläuft, wie schon Monti et al. [217] vorschlugen. In den Experimenten der Femtosekunden-Spektroskopie wird jedoch der S_1 -Zustand durch einen $n \rightarrow \pi^*$ Übergang angeregt. Für diese Reaktion konnte bis zum jetzigen Zeitpunkt durch QM-Berechnungen keiner der beiden Reaktionspfade ausgeschlossen werden. Experimente an Azobenzolderivaten,

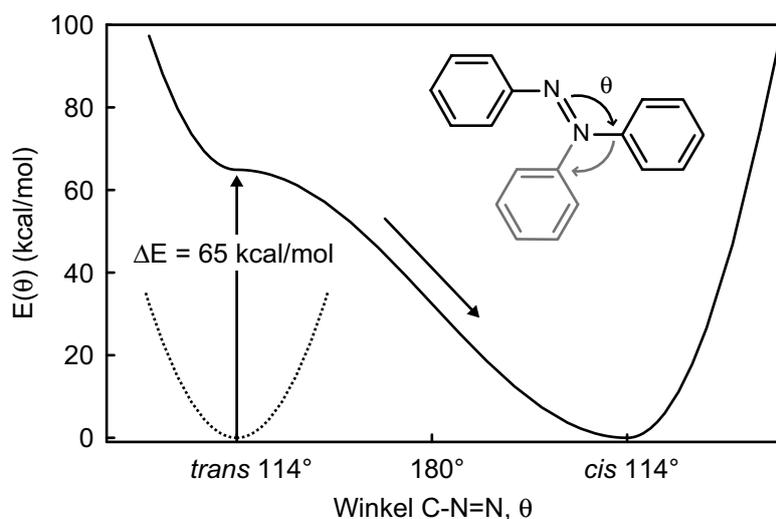


Abbildung 4.6: Das Zusatzpotential der erzwungenen Inversion. Die Inversion und ihre Reaktionskoordinate, der Bindungswinkel θ , werden in der kleinen Darstellung skizziert. Das Einschalten des gegen θ aufgetragenen Zusatzpotentials (hier in der *trans*-Form des Chromophors) führt instantan zum Eintrag der Anregungsenergie ΔE . Während der sich anschließenden Relaxation wird das Azobenzol hinüber in den *trans*-Zustand gezwungen.

in denen die Rotation der Phenylgruppen durch geeignete chemische Modifikationen verhindert wurde, lieferten zu dieser Frage jedoch eine eindeutige Aussage: sie zeigten, dass die Isomerisierung im S_2 durch die chemische Modifikation stark beeinflusst wird, nicht aber die Reaktion aus dem S_1 [221, 222]. Dieses Verhalten weist auf eine Rotation im S_2 und eine Inversion im S_1 hin, eine Schlussfolgerung, die später durch Femtosekunden-Fluoreszenz-Spektroskopie an Azobenzol von Fujino et al. [223] bestätigt wurde.

In Übereinstimmung mit diesen Resultaten wurde die Photoisomerisierung in der Form einer Inversion modelliert. Für das molekularmechanische Modell der *trans-cis*-Isomerisierung (bzw. vice versa) wurde dabei angenommen, dass der Azobenzolfarbstoff unmittelbar nach optischer Anregung in die Franck-Condon-Region des *trans*-Isomers auf der S_1 -Potentialfläche einen Reaktionspfad durchläuft, der ihn direkt durch eine konische Durchschneidung [224] auf die S_0 -Potentialfläche des *cis*-Isomers führt. Abbildung 4.6 verdeutlicht, wie dabei der Bindungswinkel θ an einem der beiden Stickstoffatome als Reaktionskoordinate dient. Um die in der Zeichnungsebene liegende Bewegung des einen Phenylrings hervorzurufen, wurde an θ das gezeigte Zusatzpotential angelegt. Die analytische Form dieses Winkelpotentials

$$E(\theta) = \begin{cases} (\theta - \theta_1) + \Delta E, & \theta \leq \theta_1 \\ \frac{\Delta E}{2} \left[1 + \cos\left(\pi \frac{\theta - \theta_1}{\theta_2 - \theta_1}\right) \right], & \theta > \theta_1 \wedge \theta \leq \theta_2 \\ k_\theta(\theta - \theta_2), & \theta > \theta_2 \end{cases} \quad (4.1)$$

ist für θ im ganzen Winkelbereich von 0° bis 360° definiert. Die Potentialfunktion setzt

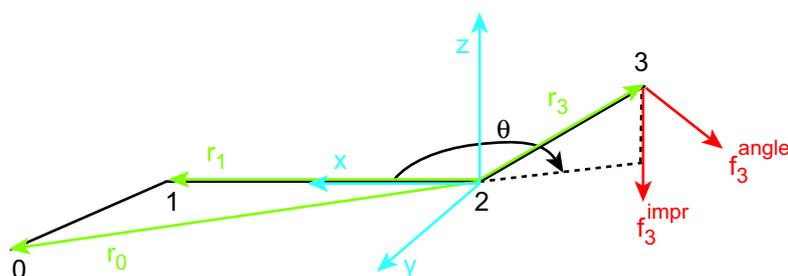


Abbildung 4.7: Das lokale Koordinatensystem der Isomerisierung. Die Orte der vier an der Isomerisierung beteiligten Atome (0, 1, 2 und 3, d.h. C–N=N–C) sind gezeigt. Das lokale Koordinatensystem (blau) hat seinen Ursprung am Ort von Atom 2, seine Definition ist im Text gegeben. Die Ortsvektoren der Atome 0, 1 und 3 im lokalen Koordinatensystem sind grün und die Kräfte auf Atom 3 rot hervorgehoben. Der Bindungswinkel θ ist als der Winkel zwischen der x-Achse und der Projektion von \mathbf{r}_3 auf die xy-Ebene definiert.

sich stückweise zusammen aus den üblichen quadratischen Termen der Winkel (2.16) und einem Kosinus, der die Isomerisierungsreaktion treibt. Winkel kleiner als 180° gehören zu dem Ausgangszustand der Isomerisierung, während größere Werte dem Endzustand entsprechen. Somit entsprechen die zwei Winkel θ_1 und θ_2 dem Gleichgewichtswinkel θ_0 , jeweils in der *cis*- oder der *trans*-Form des Chromophors. Die Kraftkonstante k_θ ist durch den für die N=N–C Gruppe angenommenen Wert aus Tabelle 4.1 gegeben, während die Energiedifferenz $\Delta E = 64,9$ kcal/mol der Energie eines Photons der Wellenlänge 480 nm entspricht. Letzteres ist die im Experiment eingesetzte Anregungswellenlänge. Mit diesen Festlegungen enthält das Zusatzpotential (4.1) und damit die Modellierung der Isomerisierungsreaktion keinen freien Parameter.

Während die Einführung eines zusätzlichen Winkelpotentials in ein MM-Kraftfeld frei von Schwierigkeiten ist, gilt dies leider nicht für das Isomerisierungspotential (4.1). Betrachtet man den in Abbildung 4.6 skizzierten Verlauf der Inversionsreaktion, so erkennt man, dass im Verlauf der Reaktion eine lineare Konfiguration der drei Atome N=N–C auftreten kann. Eine solche lineare Anordnung ist jedoch in MM-Kraftfeldern nicht vorgesehen: tritt sie auf, so ist nicht nur der Bindungswinkel zwischen den drei Atomen schlecht definiert, sondern es entstehen auch numerische Instabilitäten in der Berechnung der Dihedralwinkel, an denen das zur Isomerisierung gewählte Stickstoffatom beteiligt ist. Um dieses Problem zu beheben, wurden die betroffenen Dihedalkräfte durch äquivalente „Improper“-Winkel ersetzt und deren Kraftbeiträge mit der neu eingeführte Winkelkraft kombiniert.

Abbildung 4.7 zeigt die vier beteiligten Atome und ein lokales Koordinatensystem, welches die weitere Behandlung vereinfacht. Dabei wird davon ausgegangen, dass die Isomerisierungsreaktion an Atom 2, d.h. einem der beiden Stickstoffatome, stattfindet. Die x-Achse des neuen Koordinatensystem verbindet Atom 2 mit Atom 1, die z-Achse steht senkrecht auf der durch die Atome 0, 1 und 2 aufgespannten Ebene und die y-Achse soll

orthogonal zu den beiden anderen Achsen sein. Setzt man den Koordinatenursprung auf Atom 2 so lassen sich die Orte der anderen Atome als

$$\begin{aligned}\mathbf{r}_0 &= (r_{0x}, r_{0y}, 0) \\ \mathbf{r}_1 &= (r_{1x}, 0, 0) \\ \mathbf{r}_3 &= (r_{3x}, r_{3y}, r_{3z})\end{aligned}\quad (4.2)$$

schreiben. Die auf Atom 3 wirkende Kraft setzt sich aus der Winkelkraft und den Improper-Kräften zusammen

$$\mathbf{f}_3 = \mathbf{f}_3^{\text{angle}} + \mathbf{f}_3^{\text{impr}} = (f_{3x}, f_{3y}, f_{3z}). \quad (4.3)$$

Die Winkelkraft ist der Gradient des Zusatzpotentials (4.1), der in den Koordinaten des lokalen Koordinatensystems berechnet wird

$$\mathbf{f}_3^{\text{angle}} = \nabla E(\theta) \quad \text{mit} \quad \theta = \arccos\left(r_{3x} \left(r_{3x}^2 + r_{3y}^2\right)^{-1/2}\right). \quad (4.4)$$

Nach der in Abbildung 4.7 gezeigten Definition des Winkels θ beinhaltet $\mathbf{f}_3^{\text{angle}}$ nur Kraftkomponenten in x - und y -Richtung. Die Improperkraft wirkt senkrecht dazu und zwingt das Atom 3 in die Ebene der anderen drei Atome zurück. Aus einer Reihenentwicklung der Dihedralenergie (2.17) bezüglich r_{3z}

$$E^{\text{impr}} = 2k_{\phi,2} + (2k_{\phi,1} - 8k_{\phi,2})r_{3z}^2 + 8k_{\phi,2}r_{3z}^4 \quad (4.5)$$

kann die z -Komponente der Improperkraft zu

$$\mathbf{f}_{3z}^{\text{impr}} = (4k_{\phi,1} - 16k_{\phi,2})r_{3z} + 32k_{\phi,2}r_{3z}^3 \quad (4.6)$$

bestimmt werden. Die hier auftretenden Kraftkonstanten $k_{\phi,1}$ und $k_{\phi,2}$ sind die der Torsionskräfte, wie sie in Tabelle 4.3 aufgeführt wurden. Mit diesen zwei Beiträgen ist die auf Atom 3 wirkende Kraft in der Darstellung des internen Koordinatensystems festgelegt (4.3). Die Kräfte der anderen drei Atome können nun als Funktionen dieser Kraft (\mathbf{f}_3) und der Atomkoordinaten \mathbf{r}_i geschrieben werden. Als Nebenbedingungen müssen die Gesamtkraft und das Gesamtdrehmoment verschwinden

$$\mathbf{F} = \sum_{i=0}^3 \mathbf{f}_i \equiv 0, \quad \mathbf{M} = \sum_{i=0}^3 \mathbf{r}_i \times \mathbf{f}_i \equiv 0. \quad (4.7)$$

Weiterhin wurde festgelegt, dass keine Kräfte auf den Winkel 0-1-2 und auf die Bindung 1-2 wirken sollen, d.h. $f_{0x} = f_{0y} = f_{1x} = 0$. Mit diesen Bedingungen entstand ein eindeutig bestimmtes Gleichungssystem. Kürzt man die Projektionen der Komponenten von \mathbf{f}_3 auf diejenigen von \mathbf{r}_3 folgendermaßen ab

$$m_{a,b} = f_{3a} \cdot r_{3b}, \quad (4.8)$$

so ergab die Auflösung des Gleichungssystems die fehlenden drei nicht-trivialen Kraftkomponenten zu

$$\begin{aligned} f_{0z} &= (m_{y,z} - m_{z,y})/r_{0y} \\ f_{1y} &= (m_{x,y} - m_{y,x})/r_{1x} \\ f_{1z} &= [(m_{z,y} - m_{y,z}) r_{0x} + (m_{x,z} - m_{z,x}) r_{0y}] / (r_{0y} r_{1x}). \end{aligned} \tag{4.9}$$

Die hier hergeleiteten Zusatzkräfte und die mit ihnen verbundene potentielle Energie wurden in EGO-MMII implementiert und können in der Befehlsdatei mit dem Kommando „ISOMER“ genutzt werden. Ein Anschalten der Zusatzkraft führt automatisch zu einer Abschaltung des „normalen“ für den Winkel θ geltenden Winkelpotentials und aller betroffenen Dihedral-Potentiale. Sobald in einer Simulation das System vom Ausgangszustand bezüglich der Isomerisierung hinüber in den Zielzustand wechselt und anschließend eine gewählte Höchstauslenkung von θ aus der neuen Ruhelage unterschritten wird, erfolgt eine graduelle Abschaltung der Zusatzkraft innerhalb eines vorgegebenen Zeitraums und ein gleichzeitiges graduelles Anschalten des Standardpotentials. Diese lineare Interpolation zwischen den beiden Potentialen verhindert Sprünge in der Darstellung der Energiefläche des Chromophors, welche in Folge der näherungsweisen Ersetzung der Dihedralwinkel durch Improperwinkel entstehen können.

5 Datenanalyse und Modellbildung

In den letzten Kapiteln wurden diejenigen Vorarbeiten präsentiert, die für eine theoretische Beschreibung der Dynamik der lichtsichtbaren Modellpeptide mittels MD-Simulationen unabdingbar sind. Bevor jedoch die Simulationen und ihre vielfältigen Ergebnisse vorgestellt werden können, ist es notwendig, Methoden der statistischen Datenanalyse an das vorliegende Problem anzupassen oder neu zu entwickeln. Im Folgenden werde ich daher zunächst die bei der Analyse von Simulationsdaten anfallenden Aufgaben skizzieren, bevor ich genauer auf die von mir verwendeten Ansätze eingehe.

Das Ergebnis einer MD-Simulation ist stets eine Trajektorie $\mathbf{R}(t)$ der Ortskoordinaten aller Atome des Systems [75]. Als Folge der diskreten Integration (vgl. Abschnitt 2.1) der Bewegungsgleichungen werden die Atomorte $\mathbf{R}_i(t)$ an äquidistanten Zeitpunkten bestimmt

$$\mathbf{R}(t) = \{\mathbf{R}_1(t), \dots, \mathbf{R}_N(t)\}, \quad \text{mit } t = m \cdot \Delta t, \quad m = 1, \dots, M. \quad (5.1)$$

Die Anzahl N der Atome des simulierten Systems und die Gesamtdauer der Simulation $M \cdot \Delta t$ sind dabei durch die zu untersuchenden Prozesse oder Reaktionen vorgegeben. Durch die Verfügbarkeit immer leistungsfähigerer Computer können heutzutage große Proteine in expliziten Lösungsmitteln behandelt und dafür Trajektorien bis Mikrosekunden Länge aufgezeichnet werden [85]. Die dabei anfallenden Datensätze sind hochdimensional, denn es sind die Koordinaten von mehreren tausend Atomen zu verarbeiten. Sie sind zusätzlich auch sehr umfangreich, da bei Integrationsschritten Δt von Femtosekunden und typischen Simulationsdauern im Nanosekundenbereich mehrere Millionen Konfigurationen des Moleküls berechnet werden.

Die erste Aufgabe bei der Analyse solcher Daten ist somit eine Reduktion des Umfangs der anfallenden Datensätze, wobei darauf zu achten ist, dass die relevanten Informationen zum Systemverhalten erhalten bleiben. Abschnitt 5.1 ist dieser Fragestellung gewidmet. Dort werden mögliche Strategien zur Auswahl einer Teilmenge der in der Trajektorie (5.1) beinhalteten Daten, sowie zur Wahl passender Koordinaten diskutiert. Ferner wird in Abschnitt 5.1 die Methode der Hauptkomponentenanalyse [225–227] erläutert, welche es unter Ausnutzung linearer Korrelationen der Daten ermöglicht, die Datensätze auf essentielle Untermannigfaltigkeiten zu projizieren.

Weitere Ziele der statistischen Analyse von Peptidtrajektorien, auf die in den nachfolgenden Abschnitten eingegangen werden soll, lassen sich gut an einem einfachen Beispiel illustrieren. Abbildung 5.1 zeigt dazu eine MD-Trajektorie der ϕ , ψ -Dihedralwinkel des Alanin-Dipeptids, eines kleinen Moleküls, das häufig zu Testzwecken betrachtet wird.

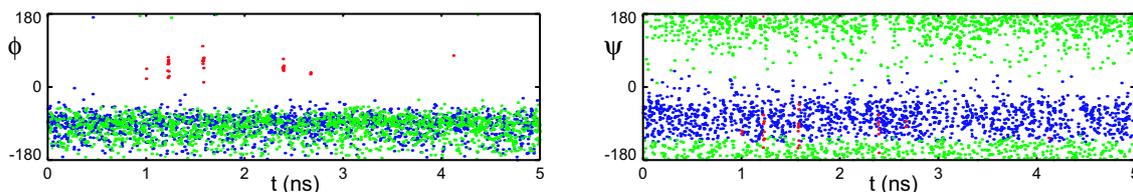


Abbildung 5.1: Beispiel einer MD-Trajektorie. Das Alanin-Dipeptid ist ein kleines Peptid-Analogon, das in Kapitel 6 noch detailliert untersucht wird. Über einen Zeitraum von 5 ns wird hier die Dynamik des Dipeptids in wässriger Lösung und bei einer Temperatur von 500 K mittels der Zeitserien der beiden Dihedralwinkel ϕ und ψ dargestellt. Mit der weiter unten vorgestellten UNIVAR-Methode wurde die Dichte der Datenpunkte im zweidimensionalen Winkelraum geschätzt, anschließend mit einem weiteren zu schildernden Verfahren die Konformationen des Moleküls identifiziert und in den Zeitserien farbig hervorgehoben. Weitere Erklärungen finden sich im Text.

Vorgreifend auf die Diskussion in Abschnitt 5.1 sei festgestellt, dass sich die Konfiguration, d.h. die räumliche Struktur, von Peptiden gut durch diese Winkel beschreiben lässt. Die in Abbildung 5.1 gezeigten Zeitverläufe der beiden Winkel machen deutlich, dass das Dipeptid mehrere geometrisch unterschiedliche Zustände einnehmen kann. Das Molekül zeigt thermische Fluktuationen innerhalb dieser Zustände und plötzliche Übergänge zwischen diesen metastabilen Zuständen, den so genannten Konformationen des Moleküls [22], die auch schon in der Einführung erwähnt wurden (vgl. Abschnitt 1.1). Die hier beschriebene Dynamik eines Moleküls wird als *Konformationsdynamik* bezeichnet und ist ein zentraler Gegenstand der vorliegenden Arbeit. Abbildung 5.1 zeigt, wieder im Vorgriff auf die Ergebnisse dieses Kapitels, drei farblich hervorgehobene Konformationen des Dipeptids, die durch ein noch zu spezifizierendes Verfahren identifiziert und markiert wurden. Man beobachtet häufige Übergänge zwischen der grün markierten und der blau markierten Konformation, während das System nur selten in die rot markierte Konformation wechselt.

Wie in Abschnitt 1.1 schon allgemein angedeutet und in Referenz [22] anhand eines vereinfachten Proteinmodells gezeigt wurde, gehören Peptid- oder Proteinkonformationen zu lokalen Minima einer freien Energiefunktion $F(\mathbf{R}_S)$. Diese Funktion, welche auch *Potential der mittleren Kraft* (engl. Potential of Mean Force, PMF) genannt wird, hat als Argument einen Satz von Koordinaten \mathbf{R}_S , welche die Bewegung einiger weniger ausgewählter Freiheitsgrade im „Wärmebad“ der vielen restlichen Freiheitsgrade des Simulationssystems charakterisieren [22]. Im Fall des Beispiels aus Abbildung 5.1 umfasst \mathbf{R}_S lediglich die beiden Dihedralwinkel des Dipeptids, während die übrigen Freiheitsgrade dieses Moleküls, sowie die des umgebenden Wassers das Wärmebad bilden.

In Referenz [22] wird genauer begründet, dass $F(\mathbf{R}_S)$ gemäß

$$F(\mathbf{R}_S) = -k_B T \ln \varrho(\mathbf{R}_S), \quad (5.2)$$

aus der invarianten Dichte $\varrho(\mathbf{R}_S)$ des Subsystems S der ausgewählten Freiheitsgrade entsteht. Hierbei ist k_B der Boltzmannfaktor und T die Temperatur des Systems S und damit

die des Gesamtsystems. Die invariante Dichte $\varrho(\mathbf{R}_S)$ kennzeichnet dabei das thermische Gleichgewicht von S [75] und kann bei hinreichend langen Simulationszeiten und im Fall einer ergodischen Dynamik von S durch die zur Trajektorie \mathbf{R}_S gehörige Punktdichte

$$\tilde{\varrho}(\mathbf{R}'_S) = \frac{1}{M} \sum_{t=1}^M \delta(\mathbf{R}'_S - \mathbf{R}_S(t)) \quad (5.3)$$

genähert werden. Um aus dieser Punktdichte die Orte der Dichtemaxima ableiten zu können, welche nach (5.2) den Minima von $F(\mathbf{R}_S)$ entsprechen und somit Molekülkonformationen kennzeichnen, muss zunächst zur Punktdichte $\tilde{\varrho}(\mathbf{R}_S)$ (5.3) ein glattes, differenzierbares Modell $\hat{\varrho}(\mathbf{R}_S)$ assoziiert werden. Diese Aufgabe ist ein Spezialfall eines allgemeinen Problems der statistischen Datenanalyse, nämlich des Problems der *Dichteschätzung* eines Datensatzes \mathcal{X} aus Datenpunkten \mathbf{x}_i [228]. Die Schätzung eines glatten Modells $\hat{\varrho}(\mathbf{R}_S)$ für die aus den Trajektorien extrahierten Datensätze \mathcal{X} basiert auf theoretischen Konzepten und Algorithmen, die in den letzten Jahren in der Gruppe von Paul Tavan erarbeitet wurden [145, 147–149]. In Abschnitt 5.2 fasse ich diese Konzepte zusammen und erläutere die zugehörigen Algorithmen anhand eines einfachen Beispiels.

Ist ein glattes Modell $\hat{\varrho}(\mathbf{R}_S)$ der invarianten Dichte $\varrho(\mathbf{R}_S)$ vorhanden, so können die Orte der Maxima dieser Dichte durch einen Gradientenaufstieg näherungsweise bestimmt werden. Bei diesem Zugang gibt es jedoch aufgrund der Endlichkeit der Trajektorie eine gewisse Unsicherheit, die darin besteht, dass auch statistische Schwankungen zu Dichtemaxima der Punktdichte $\tilde{\varrho}(\mathbf{R}_S)$ und im zugeordneten Modell führen können. Um die Peptidkonformationen von solchen zufälligen, aus unzureichender Statistik stammenden Dichtefluktuationen unterscheiden zu können, ist, wie in Abschnitt 5.3 erläutert wird, eine Analyse der Modelldichte auf multiplen Glättungsskalen notwendig [229–231]. Insbesondere werde ich im angeführten Abschnitt 5.3 den zugehörigen Stand der Forschung einschließlich meiner eigenen Beiträge [231] darstellen.

Schon in der Einleitung (Abschnitt 1.1) wurde gezeigt, dass die Konfiguration eines Peptids gut durch die Angabe seiner Dihedralwinkel beschrieben werden kann, ein Umstand, der schon bei dem Beispiel aus Abbildung 5.1 berücksichtigt wurde und der in Abschnitt 5.1 näher erläutert wird. Solche Dihedralwinkel sind jedoch, wie alle Winkel, zirkuläre Größen, bei deren statistischer Analyse Vorsicht angebracht ist. Wie Jammalamadaka und SenGupta [232] in ihrer Monographie zeigen, ist schon die Mittelwertbildung zweier Winkel nicht trivial zu bewerkstelligen, da in diesem Fall die naive Berechnung des arithmetischen Mittels zu unsinnigen Ergebnissen führen kann. Deshalb ist die letzte Aufgabe dieses Kapitels eine Erweiterung der bisher angesprochenen statistischen Verfahren in Bezug auf die Betrachtung zirkulärer Variablen. In Abschnitt 5.4 wird deshalb kurz die deskriptive Statistik zirkulärer Größen diskutiert, hauptsächlich werden dort aber die nötigen Erweiterungen am eingesetzten Verfahren zur Dichteschätzung vorgestellt.

5.1 Dimensionsreduktion

Als erste Aufgabe der Datenanalyse war die Reduktion des Umfangs des zur MD-Trajektorie (5.1) gehörigen Datensatzes formuliert worden, welcher die Ortskoordinaten der N Atome des betrachteten Systems an den M betrachteten Zeitpunkten umfasst. Zumeist findet der erste Schritt zur Datenreduktion schon während der Berechnung der MD-Trajektorie statt, da die Atomkoordinaten häufig nicht an jedem Integrationsschritt, sondern beispielsweise nur alle 500 fs oder nur jede Pikosekunde abgespeichert werden. Dies ist gerechtfertigt, denn zwar sind sehr kurze Integrations-Zeitschritte notwendig, um die Dynamik der schnellsten Freiheitsgrade des Systems ausreichend glatt abzutasten, jedoch laufen die meisten Prozesse von Interesse, wie Konformationsübergänge in Proteinen, auf viel längeren Zeitskalen ab (siehe Kapitel 1). Für die Untersuchung dieser Prozesse ist deshalb ein gröberes Abtasten der abgespeicherten Trajektorie vertretbar. In der vorliegenden Arbeit wurden die Nanosekunden-Simulationen auf diese Weise zu Datensätzen mit $M = 10^3 \dots 10^4$ Zeitschritten reduziert.

Die Dimension $D = 3N$ der Daten kann dagegen während der Simulation, abgesehen von einer eventuellen Vernachlässigung der Freiheitsgrade des Lösungsmittels, nicht ohne weiteres verringert werden, da beispielsweise im Rahmen einer späteren Visualisierung der Bewegung des betrachteten Moleküls, d.h. für die Erzeugung eines Films, alle seine Atomkoordinaten benötigt werden. Diese Einschränkung gilt allerdings nicht für die sich an den Simulationslauf anschließende statistische Analyse des abgespeicherten Datensatzes, in der eine dimensionsreduzierte Darstellung der Molekülkonfiguration sehr hilfreich ist. Im einfachsten Fall kann eine solche reduzierte Darstellung durch die Auswahl einer Teilmenge aller Atomkoordinaten erzeugt werden. So ist es beispielsweise möglich, die dreidimensionale Struktur eines Proteins in guter Näherung durch die alleinige Angabe der Orte seiner C_α -Atome (siehe Abbildung 1.2) wiederzugegeben. Ein Hinweis auf die Gültigkeit dieses Vorgehens stellt Abbildung 1.6 dar. Die dort gezeigte vereinfachte Darstellung des Peptidrückgrats eines Proteins basiert nur auf den Koordinaten der C_α -Atome und vermittelt dennoch einen guten Einblick in die Struktur dieses komplexen Moleküls.

Die Wahl der Koordinaten zur Beschreibung der Peptidkonfigurationen

Die Orte der zuvor ausgewählten Atome werden üblicherweise in kartesischen Koordinaten abgespeichert, nicht nur weil dies der internen Darstellung der meisten Simulationsprogramme entspricht, sondern auch weil kartesische Koordinaten in Berechnungen einfach zu handhaben sind. Bei einem Vergleich verschiedener Konfigurationen eines sich frei in Lösung bewegenden Proteins oder Peptids müssen jedoch stets die drei Translations- und die drei Rotationsfreiheitsgrade des Moleküls eliminiert werden, was beispielsweise durch ein „Matching“ der betrachteten Strukturen auf eine zuvor berechnete mittlere Referenzstruktur \mathbf{R}^R geschehen kann. Dazu ist es üblich, die mittlere quadratische Abweichung (engl. Root Mean Square Deviation, RMSD) der Atomkoordinaten \mathbf{R}_i der

anzupassenden Konfiguration von den Koordinaten \mathbf{R}_i^R der Referenzstruktur

$$RMSD = \sqrt{\frac{1}{D} \sum_{i=1}^D (\mathbf{R}_i - \mathbf{R}_i^R)^2} \quad (5.4)$$

durch eine Rotation und eine Verschiebung des Molekülschwerpunkts zu minimieren. Problematisch bei diesem Vorgehen ist, dass die RMSD im allgemeinen viele lokale Minima aufweist und die Resultate des Fits deshalb von der anfänglichen Ausrichtung der anzupassenden Struktur abhängig sind. Besonders stark ist diese Abhängigkeit bei einer Anwendung des RMSD-Fits auf die hier betrachteten sehr flexiblen [37, 38] Modellpeptide, da dort häufig sehr stark unterschiedliche Konfigurationen überlagert werden.

Um die hier skizzierten Probleme bei der Strukturanalyse von Peptiden zu vermeiden, ist es daher ratsam, statt kartesischer Koordinaten so genannte „interne“ Koordinaten zu verwenden. Hier bieten sich besonders die in Abschnitt 1.1 eingeführten Dihedralwinkel ϕ und ψ an, mit deren Angabe die Struktur des Peptidrückgrats eindeutig definiert ist. In dieser Darstellung tritt das Problem der Elimination der globalen Rotations- und Translationsfreiheitsgrade nicht auf, eine eventuell schlecht definierte mittlere Struktur muss nicht angegeben werden und ein Matching ist nicht notwendig. Auch unter dem Gesichtspunkt der Dimensionsreduktion ist die Beschreibung des Proteinrückgrats der Modellpeptide durch seine $2 \cdot 8 = 16$ Dihedralwinkel mindestens genauso günstig wie beispielweise die Angabe der $3 \cdot 8 = 24$ kartesischen Koordinaten der C_α -Atome.

Bei der hier getroffenen Festlegung der zur Beschreibung der Peptidkonfiguration verwendeten Koordinaten ist anzumerken, dass hier und auch im Folgenden die Konfiguration der eingebauten Lichtschalter getrennt von der des zu schaltenden Peptids behandelt wird. Dieses Vorgehen begründet sich durch die Betrachtung des Azobenzol-Schalters als eine extern auf das Peptid wirkende Randbedingung, unter der die Konfiguration und Dynamik des Peptidrückgrats untersucht werden soll.

Mit den bisherigen Vorüberlegungen können die Voraussetzungen für die in diesem Kapitel vorgestellten statistischen Analysemethoden folgendermaßen zusammengefasst werden: die Konfiguration des Rückgrats der Modellpeptide wird in einen $D = 16$ dimensionalen *Merkmalsvektor* zusammengefasst. Während ein kartesischer Merkmalsvektor als

$$\mathbf{x} = (x_1, \dots, x_D)^T \in \mathbb{R}^D \quad (5.5)$$

geschrieben wird, gelten griechische Buchstaben für zirkuläre Variablen

$$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_D)^T \in \mathbb{S}^D, \quad (5.6)$$

wobei \mathbb{S}^D den D -dimensionalen Torus bezeichnet, welcher durch die D Winkelvariablen aufgespannt wird. Der in einer typischen Simulation entstehende Datensatz

$$\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_M\} \subset \mathbb{R}^D \quad \text{bzw.} \quad \mathcal{A} = \{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_M\} \subset \mathbb{S}^D \quad (5.7)$$

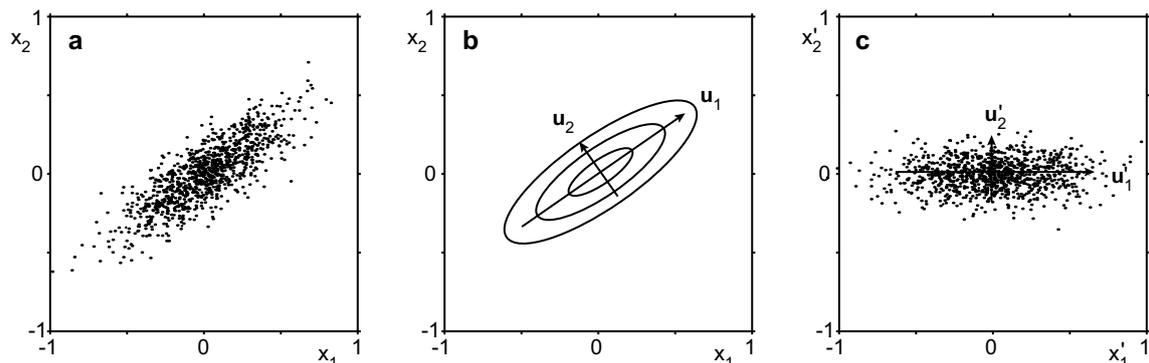


Abbildung 5.2: Hauptkomponentenanalyse. **a** Ein Beispiel linear korrelierter Daten. **b** Dargestellt sind die dem Datensatz zu Grunde liegende bivariate Normalverteilung und die zwei orthogonalen Eigenvektoren der Kovarianzmatrix. Letztere wurden mit den Varianzen σ_1^2 und σ_2^2 skaliert. **c** Das Ergebnis der Hauptkomponentenanalyse ist ein unkorrelierter Datensatz.

umfasst $M = 10^3 \dots 10^4$ Datenpunkte $\mathbf{x}_m \in \mathbb{R}^D$ (bzw. $\alpha_m \in \mathbb{S}^D$), deren statistische Unabhängigkeit durch ein hinreichend grobkörniges Abtasten der Trajektorien garantiert wird.

Die Hauptkomponentenanalyse

Die Hauptkomponentenanalyse (engl. principal component analysis, PCA) ist ein altbekanntes Verfahren der multivariaten Statistik. Pearson [233] formulierte sie schon 1901, jedoch wurde sie mehrfach, teils unabhängig von vorangehenden Arbeiten, neu entdeckt [234]. Die PCA basiert darauf, lineare Abhängigkeiten in Datensätzen auszunutzen, um die Daten auf essentielle Unterräume zu projizieren. Auf diese Weise kann eine Dimensionreduktion erreicht werden, die beispielsweise zu einer Verringerung des Rechenaufwandes in späteren Verarbeitungsschritten führt. Sie wird auch häufig dazu eingesetzt, hochdimensionale Datensätze auf ein- oder zweidimensionale Abbildungen zu reduzieren. Die Theorie der PCA und das praktische Vorgehen bei der Untersuchung eines Datensatzes findet sich in vielen Lehrbüchern, es existieren jedoch auch spezielle Monographien [225, 226]. In diesem Abschnitt soll die Theorie der PCA kurz zusammengefasst werden und ein Überblick über die zuvor veröffentlichten Anwendungen dieser Methode bei der Analyse von MD-Trajektorien gegeben werden. Bis auf Weiteres wird dabei von einem Datensatz nicht-zirkulärer Variablen ausgegangen, bevor in Abschnitt 5.4 die entsprechende Erweiterung der PCA auf zirkuläre Variablen beschrieben wird.

Die PCA basiert auf der Beschreibung des Datensatzes \mathcal{X} durch seine ersten beiden statistischen Kumulanten. Als Beispiel soll der in Abbildung 5.2 a gezeigte Datensatz verarbeitet werden. Die erste Kumulante eines Datensatzes ist der Mittelwert

$$\bar{\mathbf{x}} = \langle \mathbf{x} \rangle_{\mathcal{X}}, \quad (5.8)$$

die zweiten sind die Kovarianzen, welche in der *Kovarianzmatrix*

$$C(\mathcal{X}) = \langle (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T \rangle_{\mathcal{X}} \quad (5.9)$$

zusammengefasst werden. Die Schreibweise $\langle \dots \rangle_{\mathcal{X}}$ zeigt hier die Mittelung über alle M Datenpunkte an. Verwendet man zur Untersuchung eines Datensatzes \mathcal{X} eine solche Statistik zweiter Ordnung, so nimmt man implizit an, dass die dem Datensatz zu Grunde liegende Wahrscheinlichkeitsdichte eine einzige *multivariate Normalverteilung*

$$p(\mathbf{x}) = \frac{\exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{c})^T \Sigma^{-1} (\mathbf{x} - \mathbf{c})\right]}{(2\pi)^{D/2} (\det \Sigma)^{1/2}} \quad (5.10)$$

ist. Für die erste und zweite Kumulante einer solchen Verteilung gilt

$$\bar{\mathbf{x}} = \mathbf{c} \quad \text{und} \quad C = \Sigma. \quad (5.11)$$

Die Kovarianzmatrix C ist nach Definition symmetrisch und reell und weist deshalb D positive Eigenwerte σ_i^2

$$C\mathbf{u}_i = \sigma_i^2 \mathbf{u}_i, \quad i = 1, \dots, D \quad (5.12)$$

auf. Eine Entartung der Eigenwerte tritt bei aus realen Zufallsprozessen erzeugten Daten praktisch nicht auf, stellt aber auch keine Schwierigkeit dar. Ohne Einschränkung der Allgemeinheit sollen die Eigenwerte nach ihrer Größe sortiert sein

$$\sigma_i^2 \geq \sigma_j^2 \quad \text{für} \quad i < j.$$

Aus der Symmetrie von C folgt weiterhin, dass die zugehörigen Eigenvektoren stets orthogonal zueinander sind. Normiert man sie, so gilt

$$\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij}.$$

Abbildung 5.2 b skizziert die beiden Eigenvektoren der Kovarianzmatrix des Beispieldatensatzes, welche die Richtung der größten und der zweitgrößten (und hier damit der kleinsten) Varianz der Daten angeben. Das Verfahren der Hauptkomponentenanalyse besteht aus der Mittelwertbefreiung des Datensatzes und der Projektion der Differenzen $\mathbf{x} - \bar{\mathbf{x}}$ auf die Eigenvektoren von C

$$\mathbf{x}'_i = \mathbf{u}_i \cdot (\mathbf{x} - \bar{\mathbf{x}}), \quad (5.13)$$

zur Erzeugung neuer Koordinaten \mathbf{x}'_i . Das Ergebnis ist ein Datensatz \mathcal{X}' mit verschwindendem Mittelwert, der keine Korrelationen aufweist

$$C(\mathcal{X}') = \text{diag}(\sigma_1^2, \dots, \sigma_D^2). \quad (5.14)$$

Abbildung 5.2 c zeigt das Ergebnis dieser Transformation. Eine Dimensionsreduktion kann nun durch eine Projektion von \mathcal{X} , nicht auf alle, sondern nur auf die ersten q Eigenvektoren erreicht werden. Die Wahl der Anzahl der vernachlässigten Hauptachsen $q + 1, \dots, D$ bestimmt den Teil

$$\sigma_V^2 = \sum_{i=q+1}^D \sigma_i^2 \quad (5.15)$$

der Gesamtvarianz des Datensatzes \mathcal{X} , welcher im projizierten Datensatz \mathcal{X}' vernachlässigt wird.

Bei Zufallsverteilungen, die durch Messungen oder Simulationen komplexer Systeme bestimmt wurden, ist zumeist ein exponentielles Abfallen von σ_i^2 mit i zu beobachten, so dass häufig schon sehr wenige der *Hauptkomponenten*, d.h. der Hauptachsen mit großem σ_i^2 , einen großen Teil der Variabilität einer Datenverteilung erfassen. Dementsprechend geht bei der Projektion auf die Hauptkomponenten kaum Information verloren. Zieht man ferner in Betracht, dass die mit den vernachlässigten Dimensionen assoziierten Datenvariabilitäten häufig das unkorrelierte Rauschen darstellen, so führt eine PCA zusätzlich zur Rauschunterdrückung [225]. Abbildung 5.2 zeigt die PCA für einen linear korrelierten zweidimensionalen Datensatz. In diesem Beispiel könnte durch eine Projektion auf die größere der beiden Hauptkomponenten ein gutes Modell der Datenverteilung erzeugt werden.

Vergleichbar mit dem am Anfang dieses Abschnitts erwähnten mehrfachen historischen „Wiederentdecken“ der PCA als statistisches Analyseverfahren wurde die Idee, eine PCA zur Auswertung von MD-Simulationen zu nutzen, mehrfach von verschiedenen Gruppen eingeführt und deshalb auch mit unterschiedlichen Bezeichnungen versehen. Einen guten Überblick zu dieser Entwicklung gibt der Artikel von Kitao und Go [235]. In der Groninger Forschungsgruppe wurde die PCA treffend als ein Verfahren zur Untersuchung der „Essential Dynamics“ von Proteinen bezeichnet [236] und untersucht, welchen Einfluss die Dauer einer MD-Trajektorie auf das Ergebnis dieses Verfahrens hat [237].

Anwendungen der Methode der „Essential Dynamics“ beschäftigten sich unter anderem mit dem reversiblen Faltungsverhalten eines Heptapeptids [238], den von einem Protein in freien Simulationen abgetasteten Subräumen seiner Konformationen [239] und einem Vergleich dieser Räume mit NMR-Ensembles [240]. Abseher und Nilges [241] verwendeten schließlich das Wissen über die essentiellen Freiheitsgrade eines Proteins, um eine gesteuerte Dynamik zu formulieren, die auf effizientere Weise das Absuchen von Konformationsräumen ermöglicht. Andere Anwendungen, in denen die PCA genutzt wurde, um die Konfigurationen von Biomolekülen zu untersuchen, wurden in der letzten Zeit für Peptide [88, 90, 242] und Proteine [7, 243, 244] präsentiert. Probleme treten bei der Anwendung der PCA besonders durch die immer beschränkte Simulationszeit und das damit verbundene unvollständige Abtasten des Konformationsraumes der Moleküle auf, wie Balsera et al. [245] und Hess [246] aufgezeigt haben.

Eng verwandt mit der PCA – und unter gewissen Bedingungen dual zu ihr – ist die so genannte „Principal Co-ordinate Analysis“ (PCoordA), die nicht auf der Betrachtung der Korrelationen der einzelnen Variablen basiert, sondern die Ähnlichkeiten der einzelnen Datenpunkte in Form ihrer Distanzmatrix auswertet. Dabei bietet die PCoordA durch die freie Wahl der anzuwendenden Distanz-Metrik eine gewisse größere Flexibilität als die PCA. Der Hauptunterschied in der Anwendung der PCA und der PCoordA liegt jedoch darin, dass im ersten Fall eine $D \times D$ -Matrix zu diagonalisieren ist (5.12), während im zweiten Fall eine Matrix der Größe $M \times M$ vorliegt. Da in den meisten Fällen die Anzahl der Datenpunkte deutlich größer als die der Dimensionen ist ($M \gg D$), führt die PCA üblicherweise zu einem geringeren Rechenaufwand. Eingeführt in das Gebiet der MD-Simulationen wurde die PCoordA von Becker [247, 248] und dort mehrfach genutzt [249, 250], um beispielsweise den Einfluss von strukturellen Einschränkungen auf die Energielandschaft von Peptiden zu untersuchen.

5.2 Dichteschätzung

Wie schon in der Einleitung dieses Kapitels dargestellt wurde, ist es die Aufgabe der Dichteschätzung, aus einem Datensatz \mathcal{X} (5.7) ein glattes und evtl. differenzierbares Modell $\hat{p}(\mathbf{x})$ der ihm zu Grunde liegenden unbekanntenen Wahrscheinlichkeitsverteilung $p(\mathbf{x})$ zu erzeugen [228].

Eine direkte Methode zu einem Datenmodell (vgl. (5.3)) eine geglättete Dichte zu assoziieren, wurde von Parzen [251] vorgeschlagen und ist auch als *Kernel-Dichteschätzung* bekannt. Die Kernel-Dichteschätzung ist ein Dichtemodell

$$\hat{p}(\mathbf{x}|\mathcal{X}, \sigma) = \frac{1}{M} \sum_{i=1}^M g(\mathbf{x} - \mathbf{x}_i|\sigma) \quad (5.16)$$

in Form einer Mischung normierter Gaußfunktionen

$$g(\mathbf{x} - \mathbf{x}_i|\sigma) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp\left[-\frac{(\mathbf{x} - \mathbf{x}_i)^2}{2\sigma^2}\right] \quad (5.17)$$

identischer Breiten, welche an den Orten der Datenpunkte zentriert sind. Im Grenzfall beliebig kleiner Breiten σ entspricht die Kernelschätzung dem Datenmodell. Die Wahl eines endlichen Wertes für σ ermöglicht hingegen eine Glättung und Generalisierung der geschätzten Dichte. Jedoch ist es nicht möglich, eine optimale Wahl der Breite σ anzugeben, da für diesen frei wählbaren Parameter kein Optimierungskriterium existiert. Außerdem beschränkt die mit der Anzahl der Datenpunkte steigende Anzahl von Gaußfunktionen und der damit verbundene Berechnungsaufwand die Anwendung der Kernel-Dichteschätzung auf kleine Datensätze.

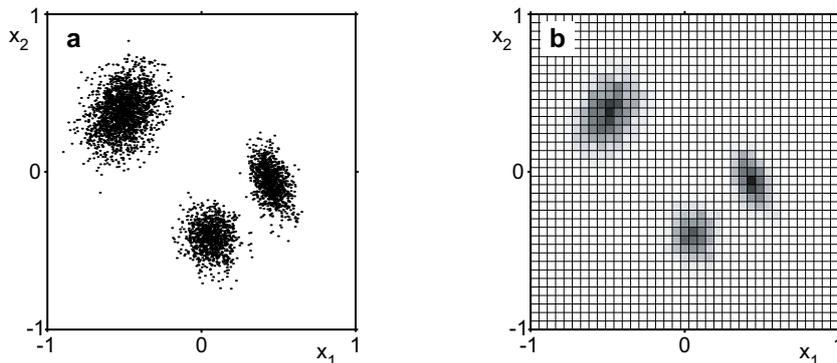


Abbildung 5.3: Die Histogrammschätzung. **a** Ein Beispiel multimodal verteilter Daten. **b** Dichteschätzung durch ein Histogramm.

Im Gegensatz zu der nicht-parametrischen Kernel-Dichteschätzung wird bei einer *parametrischen Dichteschätzung* eine bestimmte funktionale Form $\hat{p}(\mathbf{x}|\Theta)$ der Wahrscheinlichkeitsverteilung angenommen. Die Parameter Θ des geschätzten Modells sind dann durch geeignete Mittelwertbildungen aus dem Datensatz zu bestimmen. Offensichtlich stellt in diesem Sinne auch die im vorangehenden Abschnitt vorgestellte Hauptkomponentenanalyse eine Dichteschätzung dar: als Modelldichte nimmt sie eine multivariate Normalverteilung der Form (5.10) an und erlaubt es nach (5.11), deren Parameter aus den ersten beiden Kumulanten von \mathcal{X} zu schätzen. Allerdings können mit einer Normalverteilung nur Datenverteilungen in akzeptabler Weise repräsentiert werden, die lediglich ein Maximum aufweisen.

Abbildung 5.3 a zeigt dagegen das Beispiel eines multimodalen Datensatzes, der nur schlecht durch eine einzige Normalverteilung beschrieben werden kann. Ein erster und naheliegender Weg, die Dichte dieses Datensatzes durch ein parametrisches Modell zu schätzen, ist die Berechnung eines *Histogramms*, wie es in Abbildung 5.3 b gezeigt wird. Bei der Histogramm-Schätzung wird die zu Grunde liegende Wahrscheinlichkeitsdichte durch eine Mischung von lokalen Gleichverteilungen angenähert, deren Zentren auf einem regelmäßigen Gitter liegen und die alle die gleiche Breite b aufweisen. Die statistischen Gewichte dieser Gleichverteilungen werden dann durch einfaches Abzählen der in die entsprechenden „Töpfe“ fallenden Datenpunkte geschätzt.

Eine Histogrammschätzung ist leicht zu implementieren und liefert schnell anschauliche Ergebnisse. Jedoch ist diese Art der Dichteschätzung gleich mehreren Begrenzungen unterworfen. Erstens ist es nicht möglich, eine optimale Wahl für die Topfbreite b und damit die Auflösung des Histogramms anzugeben. Wird b als zu groß gewählt, können wichtige Details einer Datenverteilung übersehen werden, bei zu kleinem b sind die statistischen Gewichte mit sehr großen Fehlern versehen. Fatal ist jedoch die hier auftretende Verkörperung des so genannten „Fluchs der Dimensionalität“. Denn mit der Dimension D des Problems steigt – bei einer fest gewählten Kantenlänge b – die Anzahl der zur Untertei-

lung Raumes anzusetzenden Gleichverteilungen exponentiell an, so dass in den meisten der Volumina keine oder nur wenige Datenpunkte gezählt werden. Aus diesem Grund ist die Anwendung der Histogrammschätzung effektiv auf zwei- oder dreidimensionale Datensätze beschränkt.

Dichteschätzung mit Mischungen von Normalverteilungen

Im Gegensatz zu den bis hierhin vorgestellten Verfahren ist die Dichteschätzung mit Mischungen von Normalverteilungen frei von Beschränkungen bezüglich der Dimensionalität der zu schätzenden Datenverteilung und hat den Vorteil, mit einer frei wählbaren Datenreduktion verbunden zu sein. Hier möchte ich die in der Gruppe von Paul Tavan zu diesem Thema geleisteten Vorarbeiten [145–147, 149, 252] zusammenfassen.

Während Kloppenburg und Tavan [145] den verallgemeinerten Fall einer Dichteschätzung mit einer Mischung aus multivariaten Gaußverteilungen verschiedener statistischer Gewichte formulierten, werde ich mich auf den für die vorliegende Arbeit wichtigen Spezialfall univariater Normalverteilungen gleicher Gewichte $\hat{P}_r = 1/R$ beschränken. Die dabei angenommene parametrische Form der Wahrscheinlichkeitsdichte

$$\hat{p}(\mathbf{x}|\Theta) = \frac{1}{R} \sum_{r=1}^R g(\mathbf{x} - \mathbf{c}_r|\sigma) \quad (5.18)$$

ist eine Mischung aus R normierten Gaußkugeln $g(\mathbf{x} - \mathbf{c}_r|\sigma)$ (5.17). Im Gegensatz zur Kernel-Dichteschätzung (5.16) repräsentiert hier jede Normalverteilung nicht nur einen einzelnen Datenpunkt, sondern eine Teilmenge des Datensatzes, so dass eine *Datenkompression* um den Faktor $\kappa = M/R$ erzielt wird. Die Menge der Parameter des Mischungsmodells (5.18) wird mit Θ bezeichnet und setzt sich aus den Zentren \mathbf{c}_r der Mischungskomponenten sowie der globalen Varianz σ^2 zusammen. Kloppenburg und Tavan [145] gaben mit UNIVAR einen robusten Lernalgorithmus zur Optimierung des Parametersatzes Θ an, dessen Eigenschaften später von Albrecht et al. [147] noch detaillierter untersucht wurden. Bevor der Lernalgorithmus vorgestellt werden kann, ist es notwendig, die zu Grunde liegende Partitionierung des Merkmalsraumes zu diskutieren und das Kriterium für eine optimale Dichteschätzung festzulegen.

Bayessche Klassifikation und die Partitionsfunktion

In der Mustererkennung stellt sich häufig das Problem der Klassifikation, d.h. der Zuordnung einer durch einen Merkmalsvektor \mathbf{x} (5.5) beschriebenen Beobachtung zu einem von mehreren diskreten und grobkörnig definierten Zuständen oder Klassen $k = 1, \dots, K$ des betrachteten Systems. Bei der so genannten *überwachten Klassifikation* geht man davon aus, dass sowohl die Anzahl der möglichen Zustände K als auch die *a priori* Wahrscheinlichkeiten $P(k)$ bekannt sind, einen Zustand k unter Vernachlässigung der Beobachtung vorzufinden. Zusätzlich seien auch die klassenlokalen bedingten Wahrscheinlichkeitsdichten $p(\mathbf{x}|k)$ verfügbar, welche die Verteilungen der Zufallsvariablen \mathbf{x} für jede der Klassen angeben.

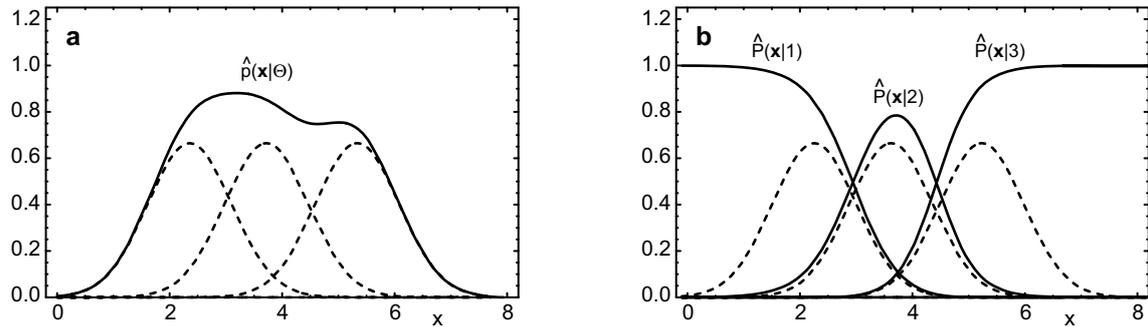


Abbildung 5.4: Die Partitionierung des Merkmalsraumes. **a** Ein eindimensionales Beispiel einer Mischungsdichte $\hat{p}(\mathbf{x}|\Theta)$ (5.18) aus univariaten Normalverteilungen identischer statistischer Gewichte. Jede der drei gestrichelt dargestellten Einzelkomponenten bildet eine eigene Klasse. **b** Die Partitionsfunktionen $\hat{P}(r|\mathbf{x})$ (5.22) unterteilen den Raum und ordnen die einzelnen Raumbereiche unscharf, d.h. nach Wahrscheinlichkeit, jeweils einer der drei Klassen zu.

Der Bayessche Satz [228, 253, 254] verbindet diese Größen mit den *a posteriori* Wahrscheinlichkeiten

$$P(k|\mathbf{x}) = \frac{P_k p(\mathbf{x}|k)}{p(\mathbf{x})}, \quad (5.19)$$

mit

$$p(\mathbf{x}) = \sum_{k'=1}^K P_{k'} p(\mathbf{x}|k'), \quad (5.20)$$

des Vorliegens eines Zustands k unter der Bedingung der Beobachtung von \mathbf{x} . Für jeden Merkmalsvektor \mathbf{x} gilt

$$\sum_{k=1}^K P(k|\mathbf{x}) = 1. \quad (5.21)$$

Betrachtet man die *a posteriori* Wahrscheinlichkeiten als Funktionen von \mathbf{x} , so erhält man eine unscharfe Aufteilung des Merkmalsraumes in Gebiete, die den einzelnen Klassen nach den Wahrscheinlichkeiten $P(k|\mathbf{x})$ zuzuordnen sind. Aus diesem Grund werden die Wahrscheinlichkeiten $P(k|\mathbf{x})$ auch die *Partitionsfunktionen* bezeichnet.

Identifiziert man die R Komponenten der geschätzten Mischung (5.18) als Schätzungen von klassenlokalen Dichten und fasst man die statistischen Gewichte $\hat{P}_r = 1/R$ als die *a priori* Wahrscheinlichkeiten der Zustände r auf, so kann der Bayessche Satz (5.19) dazu genutzt werden, die Zuordnungswahrscheinlichkeiten der Datenpunkte \mathbf{x} zu den einzelnen Normalverteilungen anzugeben. Durch Einsetzen in (5.19) erhält man die Partitionsfunktionen des univariaten Dichtemodells

$$\hat{P}(r|\mathbf{x}, \Theta) = \frac{g(\mathbf{x} - \mathbf{c}_r|\sigma)}{R \hat{p}(\mathbf{x}|\Theta)}. \quad (5.22)$$

Abbildung 5.4 zeigt die hier diskutierten Zusammenhänge an einem eindimensionalen Beispiel.

Aufbauend auf der unscharfen Partitionierung des Merkmalsraumes kann auch eine eindeutige und diskrete Klassifikation der Merkmalsvektoren vorgenommen werden. Das Bayessche Entscheidungskriterium sagt hierzu aus, dass die optimale Entscheidungsstrategie diejenige ist, \mathbf{x} einfach der Klasse k mit der größten Zuordnungswahrscheinlichkeit zuzuordnen [228]

$$\hat{P}(k|\mathbf{x}, \Theta) = \max_{k'} \left\{ \hat{P}(k'|\mathbf{x}, \Theta) \right\} \Rightarrow \mathbf{x} \rightarrow k. \quad (5.23)$$

Fasst man wie zuvor jede Komponente einer geschätzten Mischungsichte (5.18) als eine eigene Klasse auf, so wird jeder Datenpunkt durch (5.23) dem Zentrum \mathbf{c}_k einer der Komponenten zugeordnet^a und somit durch diesen kodiert. Aus diesem Grund wird die Menge dieser Zentren auch als *Kodebuch* der Dichteschätzung bezeichnet und werden die \mathbf{c}_k manchmal auch *Kodebuchzentren* genannt.

Das Prinzip der Maximum Likelihood

Wie lässt sich für einen gegebenen Datensatz \mathcal{X} statistisch unabhängiger Messwerte eine möglichst gute Dichteschätzung $\hat{p}(\mathbf{x}|\Theta)$ in Form einer Mischung aus R univariaten Normalverteilungen bestimmen, d.h. welches sind die optimalen Parameter Θ der Modelldichte? Zur Bewertung der Güte einer solchen Schätzung erweist sich die *Likelihood*

$$\mathcal{L}(\Theta|\mathcal{X}) := \prod_{\mathbf{x} \in \mathcal{X}} \hat{p}(\mathbf{x}|\Theta) \quad (5.24)$$

als zweckmäßig, welche als Funktion des Parametersatzes Θ ein Maß für die Wahrscheinlichkeit angibt, den Datensatz \mathcal{X} durch Ziehen aus der geschätzten Wahrscheinlichkeitsdichte $\hat{p}(\mathbf{x}|\Theta)$ zu erzeugen. Somit kann das Optimierungsproblem der Dichteschätzung in das Problem der Maximierung der Likelihood bezüglich der Parameter Θ überführt werden. Alternativ wird zumeist die mittlere *Log-Likelihood*

$$l(\Theta|\mathcal{X}) = \frac{1}{M} \sum_{\mathbf{x} \in \mathcal{X}} \ln [\hat{p}(\mathbf{x}|\Theta)] \quad (5.25)$$

maximiert. Dieses Vorgehen ist durch die Monotonität des Logarithmus gerechtfertigt und meist vorteilhaft, da die Log-Likelihood analytisch einfacher zu handhaben ist.

Kloppenburg und Tavan [145] haben gezeigt, dass die Likelihood von denjenigen Parametern Θ einer univariaten Dichteschätzung maximiert wird, die den Selbstkonsistenzbedin-

^a Im hier diskutierten univariaten Fall entspricht (5.23) einfach einer Zuordnung des Datenpunktes zu demjenigen Zentrum \mathbf{c}_k mit dem geringsten euklidischen Abstand $|\mathbf{x} - \mathbf{c}_k|$.

gungen der Mittelwerte

$$\mathbf{c}_r = \langle \mathbf{x} \rangle_{r, \Theta} \quad (5.26)$$

$$\hat{P}_r = \frac{1}{R} = \left\langle \hat{P}(r|\mathbf{x}, \Theta) \right\rangle_{\mathcal{X}} \quad (5.27)$$

genügen. Hier tritt neben dem üblichen globalen Mittelwert einer Funktion $f(\mathbf{x})$ in Bezug auf den gesamten Datensatz \mathcal{X}

$$\langle f(\mathbf{x}) \rangle_{\mathcal{X}} = \frac{1}{M} \sum_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad (5.28)$$

auch ein *lokaler Mittelwert*

$$\langle f(\mathbf{x}) \rangle_{r, \Theta} = \frac{\sum_{\mathbf{x} \in \mathcal{X}} \hat{P}(r|\mathbf{x}, \Theta) f(\mathbf{x})}{\sum_{\mathbf{x} \in \mathcal{X}} \hat{P}(r|\mathbf{x}, \Theta)} \quad (5.29)$$

auf, welcher auf der zuvor eingeführten Partitionierung (5.22) des Merkmalsraumes basiert. Man erkennt, dass die Selbstkonsistenzbedingung (5.26) bedeutet, dass bei einer optimalen Dichteschätzung die Zentren der Mischungskomponenten jeweils einen lokalen Mittelwert der Datenverteilung bilden. Die zweite Bedingung (5.27) fordert dazu die so genannte *Load-Balance*, d.h. dass jede Normalverteilung gleich viele Datenpunkte repräsentiert.

Eine ähnliche Formulierung der Selbstkonsistenzbedingungen für den allgemeinen Fall multivariater Normalverteilungen variabler statistischer Gewichte ergibt nicht nur ein Kriterium für eine optimale Schätzung der Parameter, sondern definiert auch einen iterativen Algorithmus zu deren Optimierung, der als *Expectation-Maximization* (EM) Verfahren [255] bezeichnet wird. Das EM-Verfahren weist allerdings schwerwiegende algorithmische Probleme auf [147, 228], so zum Beispiel eine starke Abhängigkeit der sich ergebenden Parameter von der Wahl ihrer Startwerte und eine Instabilität der statistischen Gewichte \hat{P}_r [252]. Kloppenburg und Tavan [145] erklärten dieses Verhalten, das häufig zu singulären Lösungen der Selbstkonsistenzbedingungen führt, durch die komplexe Struktur der Likelihood im hochdimensionalen Parameterraum, auf welcher der EM-Algorithmus einen naiven Gradientenaufstieg durchführt.

Der Lernalgorithmus

Das von Kloppenburg und Tavan [145] (vgl. auch [147] sowie [252]) geschaffene UNIVAR-Verfahren zur Optimierung der Parameter Θ einer Dichteschätzung mit Mischungen univariater Normalverteilungen erfüllt die Selbstkonsistenzbedingungen (5.27) und (5.26) und beseitigt gleichzeitig die beschriebenen Probleme des EM-Verfahrens. UNIVAR verbindet das aus dem EM-Verfahren übernommene Kriterium der maximalen Likelihood mit Konzepten des statistischen Clusterings, wie sie von Rose, Gurewitz und Fox [256] (RGF) eingeführt wurden. Die Autoren schlugen zur Optimierung eines Codebuchs unter Minimierung eines Fehlermaßes ein Verfahren des deterministischen *simulated annealing* vor.

1. Initialisierung der Kodebuchzentren \mathbf{c}_r mit dem Schwerpunkt der Daten, der Gaußbreite σ mit der Standardabweichung der Daten und Festlegung eines Startwerts des Lernparameters ε .
2. Zufällige Ziehung eines Datenpunktes \mathbf{x} aus \mathcal{X} .
3. Berechnung der neuen Kodebuchzentren $\mathbf{c}_r^{\text{neu}}$ zu

$$\mathbf{c}_r^{\text{neu}} = \mathbf{c}_r^{\text{alt}} + \varepsilon \hat{P}(r|\mathbf{x}, \Theta) (\mathbf{x} - \mathbf{c}_r^{\text{alt}}) + \boldsymbol{\nu}, \quad (5.30)$$

wobei $\boldsymbol{\nu}$ ein kleiner Rauschterm ist, der das Aufspalten der Komponenten erleichtert.

4. Verkleinerung von σ und ε nach den Abkühlregeln. Weiter mit 2., bis das Maximum der Likelihood erreicht ist.

Abbildung 5.5: Der UNIVAR-Algorithmus. Das Verfahren zur Optimierung der Parameter einer Dichteschätzung mit univariaten Normalverteilungen unter Maximierung der Likelihood.

Das RGF-Verfahren basiert auf der Variation eines Skalenparameters σ , der die Unschärfe einer Partitionsfunktion (5.22) angibt und dem eine Temperatur zugeordnet werden kann [256]. Eine hohe Temperatur führt zu einer glatten und konvexen Energielandschaft mit einem leicht aufzufindenden Minimum. Ein sich anschließendes Abkühlen resultiert dann in einer hierarchischen Serie von Kodebüchern, die jeweils eine näherungsweise Lösung des Problems auf feineren Skalen darstellen. Das UNIVAR-Verfahren übernimmt die Idee des Abkühlverfahrens, sein Optimierungskriterium, die Log-Likelihood, weist jedoch, im Gegensatz zum Fehlermaß des RGF-Verfahrens, ein Extremum beim optimalen Wert des Parameters σ auf und erlaubt damit die Festlegung der optimalen minimalen Temperatur im Abkühlprozess.

Abbildung 5.5 führt die Schritte des UNIVAR-Algorithmus auf. In dem gezeigten Lernschritt werden die Kodebuchzentren auf einen beliebig gezogenen Datenpunkt zu bewegt, wobei die Schrittweite jeweils mit der Zuordnungswahrscheinlichkeit $\hat{P}(r|\mathbf{x}, \Theta)$ des Datenpunktes zu dem betreffenden Kodebuchzentrum und einem Lernparameter ε gewichtet wird. Nach Albrecht et al. [147] konvergiert diese sequentielle stochastische Lernregel auf Parameterwerte, welche der Selbstkonsistenzbedingung (5.26) genügen, d.h. die Likelihood maximieren. Im Verlauf des Trainings der Parameter wird die Gaußbreite σ und der Lernparameter ε nach einem vorgegebenen Schema abgekühlt. Um das während des Abkühlprozesses stattfindende Aufspalten der Kodebuchzentren zu erleichtern, wird in Gleichung (5.30) zusätzlich ein kleiner normalverteilter Rauschterm $\boldsymbol{\nu}$ addiert. Dersch und Tavan [252] haben schließlich zum skizzierten Verfahren noch einige weitere algorithmische Schritte hinzugefügt, die dafür sorgen, dass auch die zweite Selbstkonsistenzbedingung (5.27) erfüllt wird und somit suboptimale Lösungen vermieden werden.

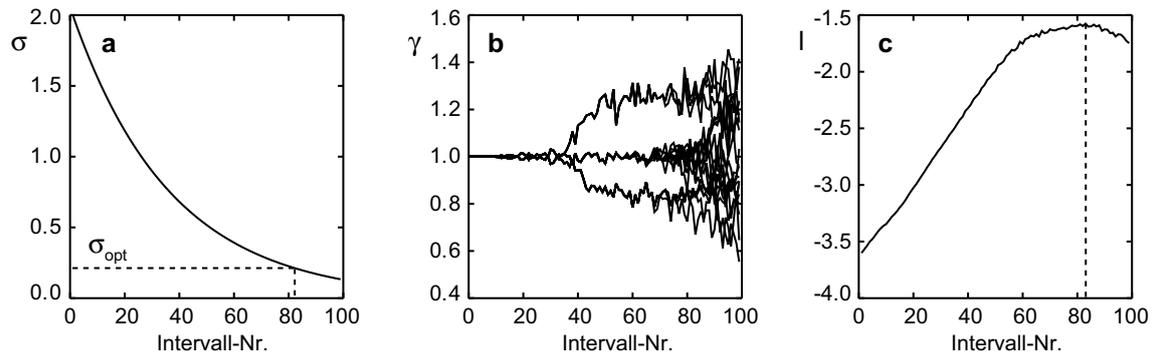


Abbildung 5.6: Der Verlauf eines UNIVAR-Trainings. Als Funktion des Trainingsintervalls wird für das in Abbildung 5.8 gezeigte Beispiel einer Dichteschätzung **a** die Gaußbreite σ , **b** die beobachtete Load γ_r (5.31) der 20 Kodebücher und **c** die resultierende Log-Likelihood l aufgetragen.

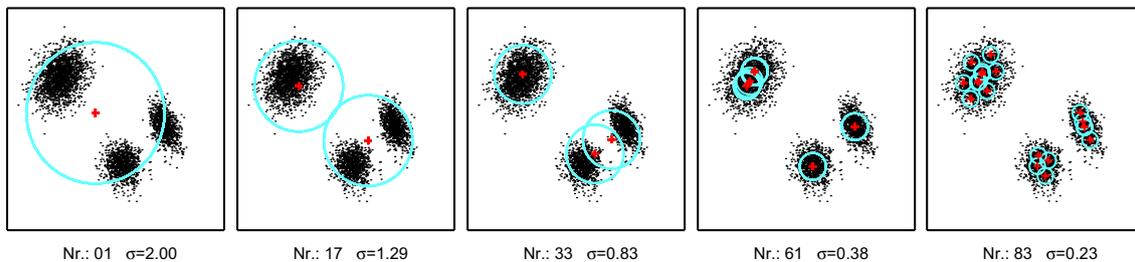


Abbildung 5.7: Die Entwicklung der Orte der Kodebuchzentren. Zu dem in Abbildung 5.6 gezeigten Verlauf eines UNIVAR-Trainings wird hier die entsprechende Entwicklung der Aufpunkte c_r verdeutlicht. Zu einer Auswahl der Lernintervalle, mit deren Nummer die Bilder jeweils bezeichnet sind, werden der Beispieldatensatz als schwarze Punkte und die Orte der Kodebuchzentren als rote Kreuze dargestellt. Die Kreise verdeutlichen die jeweilige Größe σ der Gaußfunktionen.

Anhand des in Abbildung 5.3 a gezeigten Beispieldatensatzes soll der Ablauf des UNIVAR-Lernprozesses vorgeführt werden. Zur Schätzung der Dichte dieses 4.000 Datenpunkte umfassenden Datensatzes wurde eine Mischung von $R = 20$ univariaten Normalverteilungen angesetzt. Abbildung 5.6 a zeigt das exponentielle Abkühlen der Breite σ im Verlauf des Trainings. Etwas abweichend von dem in Abbildung 5.5 aufgeführten Algorithmus und motiviert durch Effizienzbetrachtungen geschah das Abkühlen von σ und ε nicht kontinuierlich, sondern in diskreten Intervallen. In jedem dieser Trainingsintervalle wurde jeweils eine Anzahl von Datenpunkten zufällig aus dem Datensatz gezogen und trainiert, welche von der Größenordnung des Gesamtumfangs des Datensatzes ist, so dass während jedes Intervalls ein quasistationärer Zustand der Dichteschätzung angestrebt wurde.

Der Verlauf des Trainings kann anhand mehrerer Größen verfolgt werden. Abbildung 5.6

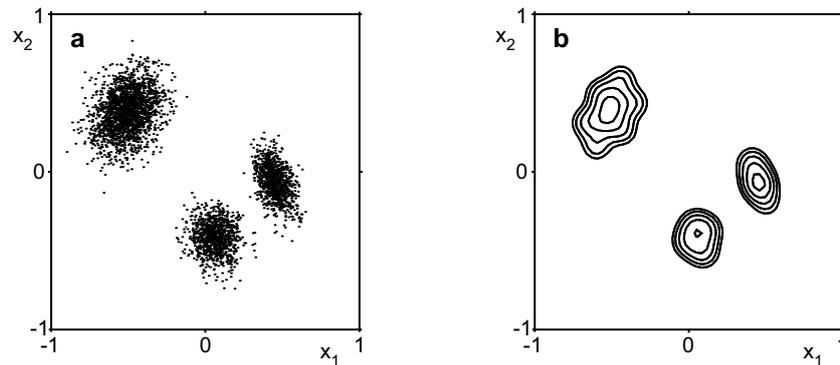


Abbildung 5.8: Das Ergebnis einer UNIVAR Dichteschätzung. **a** Das schon in Abbildung 5.3 gezeigte Beispiel multimodal verteilter Daten. **b** Dichteschätzung durch eine Mischung von $R = 20$ univariaten Normalverteilungen. Die Konturlinien verbinden Punkte gleicher Wahrscheinlichkeitsdichte $\hat{p}(\mathbf{x}|\Theta)$. Der Abstand der Konturen ist logarithmisch gewählt.

b zeigt die so genannte *Load*

$$\gamma_r = R \cdot \left\langle \hat{P}(r|\mathbf{x}, \Theta) \right\rangle_{\mathcal{X}} \quad (5.31)$$

der R Kodebücher, d.h. der beobachtete Anteil der von einem Kodebuch repräsentierten Datenpunkte relativ zu dem Anteil von $1/R$, der durch die nach (5.27) geforderten Load-Balance vorgeschrieben wird [252]. Die Load-Balance gilt im Fall einer endlichen Anzahl R von Kodebüchern jedoch immer nur näherungsweise, da bei den im Abkühlprozess auftretenden Aufspaltungen (siehe Abbildung 5.7) immer eine diskrete Zahl von Kodebuchzentren auf mehrere Datenverteilungen unterschiedlichen Umfangs aufgeteilt werden müssen. Die in Abbildung 5.6 b nach der Aufspaltung in drei Verteilungen zu beobachtende Abweichung der Loads vom Idealwert $\gamma_r = 1$ ist ein Beispiel dafür. Um eine zu breite Verteilung der Loads zu verhindern, wird, sobald für zwei Kodebuchzentren i und j ein Verhältnis der Loads von $\gamma_i > 2\gamma_j$ auftritt, ein Umsortieren vorgenommen, d.h. der Ort des Kodebuchzentrums j wird gleich dem des Zentrums i gesetzt [252]. Laut Abbildung 5.6 b ist dies im gezeigten Beispiel aber nicht zu beobachten.

Abbildung 5.6 c zeigt mit der Log-Likelihood (5.25) die entscheidende Observable des Trainings. Wie schon erläutert wurde, nimmt sie in einem bestimmten Trainingsintervall ein Maximum l_{\max} an und bestimmt damit die optimale Gaußbreite σ_{opt} . Im Beispieltraining wird das Maximum der Log-Likelihood im 83ten Trainingsintervall beobachtet, was einer optimalen Gaußbreite σ_{opt} von 0,23 entspricht. Diese Breite bildet zusammen mit den Orten der Kodebuchzentren am Ende des ausgezeichneten Trainingsintervalls den Parametersatz Θ_{opt} der optimalen Dichteschätzung. Der optimale Parametersatz wird im letzten Bild der Serie 5.7 gezeigt, während Abbildung 5.8 noch einmal den Beispieldatensatz und dessen optimal geschätzte Dichte im Detail vorführt.

5.3 Unüberwachte Klassifikation und Prototypen

Während man bei der überwachten Klassifikation, die weiter oben dargestellt wurde, davon ausgeht, die Komponenten $p(\mathbf{x}|k)$ und Gewichte $P(k)$ der Mischung (5.19) seien bekannt, wird beim Problem der unüberwachten Klassifikation lediglich die Mischung $p(\mathbf{x})$ als bekannt vorausgesetzt, so dass die zu lösende Aufgabe darin besteht, $p(\mathbf{x})$ in eine gewisse Zahl von Komponenten $p(\mathbf{x}|k)$ mit Gewichten $P(k)$ zu zerlegen. Praktisch stellt sich das Problem dabei so, dass $p(\mathbf{x})$ lediglich in Form einer endlichen Stichprobe \mathcal{X} bekannt ist.

Als Ansatz zur Lösung dieser Aufgabenstellung schlug Schnell [229] 1964 vor, Klassen anhand der Maxima \mathbf{y}^k von $p(\mathbf{x})$ zu definieren und alle Punkte $\mathbf{x} \in \mathbb{R}^D$, die durch Gradientenaufstieg auf $p(\mathbf{x})$ mit \mathbf{y}^k verbunden sind, in eine Klasse k zusammenzufassen. Zur Erzeugung einer glatten Schätzung $\hat{p}(\mathbf{x})$ von $p(\mathbf{x})$ wählte Schnell die zuvor vorgestellte Methode der Kernel-Schätzung (5.16). Er schlug weiterhin vor, die geschätzte Dichte $\hat{p}(\mathbf{x})$ einer variablen Glättung der Breite σ zu unterziehen und den Gradientenaufstieg auf den unterschiedlich geglätteten Dichtemodellen durchzuführen. Schnell erkannte ferner, dass die Betrachtung der Anzahl N_σ der bei einer Auflösungsstufe σ auftretenden Maxima es ermöglicht, diejenigen Klassenaufteilungen zu identifizieren, die über weite Bereiche von σ stabil sind und die damit eine besondere Bedeutung für die Struktur des Datensatzes aufweisen. Insbesondere können so statistische Dichteschwankungen von signifikanten Dichtemaxima unterschieden werden.

Ein verwandtes Verfahren haben Tavan et al. [230] im Jahr 1990 formuliert. Sie legten jedoch der unüberwachten Klassifikation ein dimensionsreduziertes Dichtemodell [148, 257] des Datensatzes zu Grunde, welches auf den von Kohonen [258] vorgeschlagenen selbstorganisierenden Merkmalskarten basiert. Erst durch diese Dimensionreduktion wurde die Schnellsche Methode [229] für größere Datensätze anwendbar. Aufgrund der Verwendung neurophysiologisch motivierter Merkmalskarten wies das Verfahren von Tavan et al. [230] jedoch eine Reihe unschöner Eigenschaften auf, die mit einer statistisch wohlbegründeten und konsistenten Modellbildung nicht vereinbar sind.

Hirschberger, Carstens, Stork und Tavan [231] formulierten daher eine Kombination der im vorangehenden Kapitel vorgestellten Methode der Dichteschätzung mit Mischungen von Normalverteilungen [145, 147] mit dem oben skizzierten Gradientenaufstieg und erweiterten auf diese Weise die statistisch fundierte und sehr leistungsfähige Methode der Dichteschätzung um die Fähigkeit zur unüberwachten Klassifikation. Die Theorie zu dieser Kombination von Methoden wurde von Hirschberger [259] in seiner Diplomarbeit, welche von Paul Tavan vorgeschlagen und von mir betreut wurde, ausgearbeitet und soll hier kurz zusammengefasst werden.

Die benötigte parametrische Glättung der aus einem Datensatz bestimmten Modelldichte ist bei Verwendung der Dichteschätzung mit univariaten Normalverteilungen (5.18) einfach zu erzielen. Dazu konstruiert man eine ein-parametrische Mannigfaltigkeit von Dich-

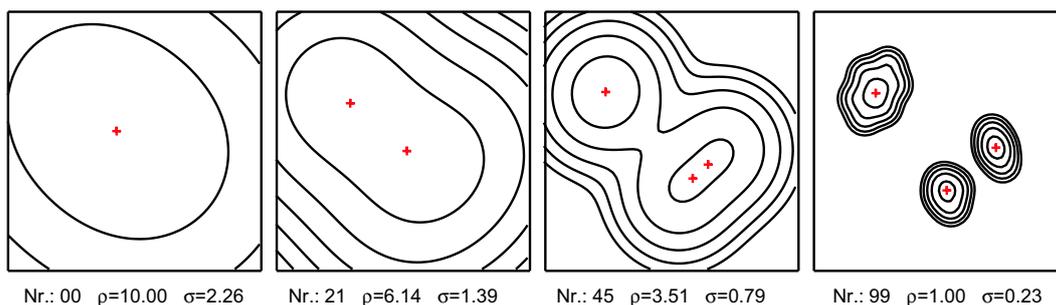


Abbildung 5.9: Die skalierte Modelldichte und ihre Maxima. Beispielhaft sind einige Elemente der Schar von skalierten Modelldichten $\hat{p}(\mathbf{x}|\Theta)$ des Beispieldatensatzes aus Abbildung 5.8 abgebildet. Der Skalenparameter ρ und die resultierende Gaußbreite σ sind für jede der Dichten angegeben. Die Abstände der Konturlinien wurden wie zuvor logarithmisch gewählt. Die durch den Gradientenaufstieg bestimmten Prototypen werden durch Kreuze markiert. Die beiden mittleren Diagramme zeigen die Modelldichten, bei denen gerade eine Aufspaltung der Minima aufgetreten ist. Die beiden anderen Diagramme repräsentieren die maximale Vergrößerung des Modells bzw. die optimale Dichteschätzung.

teschätzungen $\hat{p}(\mathbf{x}|\Theta_{\text{opt}}, \sigma)$ deren Gaußbreiten σ durch eine einfache Skalierung

$$\sigma = \rho \cdot \sigma_{\text{opt}}, \quad \text{mit } \rho \geq 1 \quad (5.32)$$

aus der im vorangehenden Abschnitt festgelegten optimalen Gaußbreite σ_{opt} entstehen. Abbildung 5.9 zeigt einige geglättete Dichten des schon oben betrachteten Datensatzes.

Dadurch, dass die Faltung einer Gaußfunktion mit einer zweiten wieder eine Gaußfunktion ergibt, ist die Skalierung der Breiten σ (5.32) äquivalent zu einer Gaußglättung der optimal geschätzten Dichte

$$\hat{p}(\mathbf{x}|\Theta_{\text{opt}}, \sigma) = \int_{\mathbb{R}^D} d^D x' g(\mathbf{x} - \mathbf{x}', \sigma_G) \hat{p}(\mathbf{x}'|\Theta_{\text{opt}}). \quad (5.33)$$

Aus diesem Grund steht das hier gewählte Vorgehen in engem Zusammenhang mit Methoden der *Skalenraumanalyse*, die schon in den 50er Jahren in Japan entwickelt wurden [260]. Skalenraummethoden sind besonders in der Bildverarbeitung von großer Bedeutung, da sie dort die vorurteilsfreie Analyse eines aufgenommenen Bildes ermöglichen, ohne dass zuvor eine Auflösungskala festgelegt werden muss [261, 262].

Folgt man dem Vorschlag von Tavan et al. [230] und formuliert den Gradientenaufstieg aus Gründen der numerischen Stabilität auf dem Logarithmus der skalierten Modelldichte $\hat{p}(\mathbf{x}|\Theta_{\text{opt}}, \sigma)$, so lässt sich zeigen [231], dass er durch die zeitliche Dynamik eines Vektors $\mathbf{x}(t)$

$$\mathbf{x}(t+1) = \sum_{r=1}^R \mathbf{c}_r \hat{P}(r|\mathbf{x}(t), \Theta_{\text{opt}}, \sigma) \quad (5.34)$$

als Fixpunktiteration formuliert werden kann. Dabei ist der Startwert $\mathbf{x}(0)$ ein Datenpunkt aus \mathcal{X} und die Partitionsfunktionen $\hat{P}(r|\mathbf{x}(t), \Theta_{\text{opt}}, \sigma)$ sind diejenigen, welche nach (5.22) zu den skalierten Komponenten (5.33) gehören. Es sei angemerkt, dass Kowalewski [263] aufbauend auf der Arbeit von Tavan et al. [230] ein verwandtes Verfahren vorgeschlagen hat.

Die durch (5.34) definierte Dynamik, welche nach Tavan et al. [230] auch als rekurrente autoassoziative Dynamik in einem neuronalen Zweischichten-Netzwerk aufgefasst werden kann, wird an einem der Extrema \mathbf{y}_σ^k der skalierten Modelldichte stationär

$$\mathbf{y}_\sigma^k = \lim_{t \rightarrow \infty} \mathbf{x}(t). \quad (5.35)$$

In der Implementierung des Gradientenaufstiegs, dem so genannten ASSO-Verfahren^a, wurde unter Ausnutzung der Tatsache, dass die stationären Punkte (5.35) der Dynamik nur an den Maxima stabil sind, durch Erweiterungen des Algorithmus dafür Sorge getragen, dass die Dynamik auch tatsächlich nur an einem solchen Maximum abbricht. Im Sinne der unüberwachten Klassifikation wird anschließend der Datenpunkt, der als Startpunkt des Gradientenaufstiegs diente, dem erreichten Maximum und der mit diesem verbundenen Klasse (5.35) zugeordnet

$$\mathbf{x}(0) \rightarrow \mathbf{y}_\sigma^k. \quad (5.36)$$

Führt man den Gradientenaufstieg auf einer der geglätteten Dichten für alle Datenpunkte des Datensatzes \mathcal{X} – oder alternativ auch nur für alle Kodebuchzentren \mathbf{c}_r – durch, so erhält man mit

$$\mathcal{Y}_\sigma = \{\mathbf{y}_\sigma^k | k = 1, \dots, N_\sigma\}, \quad (5.37)$$

die Menge aller stationären Punkte \mathbf{y}_σ^k der Auflösungsstufe σ , deren Anzahl wie zuvor mit N_σ (oder nach (5.32) äquivalent mit N_ρ) bezeichnet werden soll.

Abbildung 5.10 präsentiert das Ergebnis der Anwendung des ASSO-Verfahrens auf die skalierte Dichte des Beispieldatensatzes, die schon in Abbildung 5.9 dargestellt wurde. Abbildung 5.10 a zeigt erwartungsgemäß, dass für große σ (große ρ) alle Normalverteilungen verschmelzen und nur ein einziges Maximum aufgefunden wird. Auf feineren Skalen treten erst zwei und schließlich drei Maxima \mathbf{y}_σ^k , so genannte *Prototypen* zu Tage. Prinzipiell können selbst bei $\sigma = \sigma_{\text{opt}}$ ($\rho = 1$) nie mehr Maxima existieren, als Normalverteilungen zur Dichteschätzung herangezogen wurden. Damit gilt

$$1 \leq N_\sigma \leq R \quad \forall \sigma. \quad (5.38)$$

Abbildung 5.10 a zeigt, dass in der Beispieldichte selbst bei $\rho = 1$ nur drei Maxima existieren, was auf die Glattheit der Schätzung und die Kompaktheit der drei Einzelverteilungen zurückzuführen ist. Abbildung 5.10 b verdeutlicht schließlich das Aufspalten der Prototypen bei immer feineren Darstellungen der Modelldichte.

^a Die Bezeichnung „ASSO-Verfahren“ für die Implementierung des Gradientenaufstiegs auf der Modelldichte bezieht sich auf die Tatsache, dass dieses Modell als Autoassoziativspeicher [264] aufgefasst werden kann.

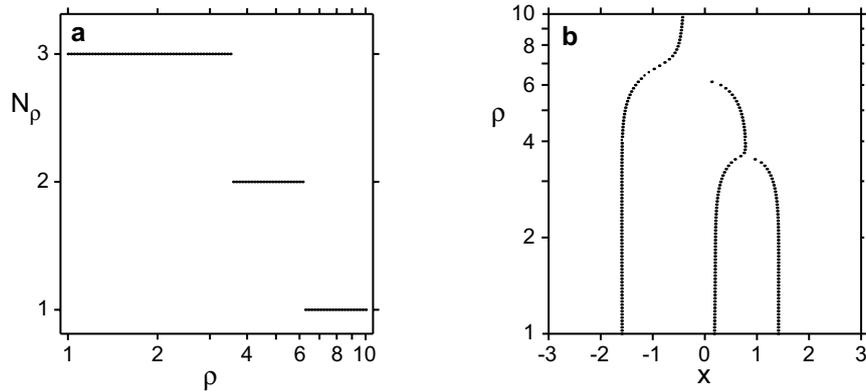


Abbildung 5.10: Das Ergebnis des ASSO-Verfahrens. Beispielhaft wurde das ASSO-Verfahren auf die in Abbildung 5.8 gezeigte Datenverteilung angewendet. **a** Die Anzahl der detektierten Prototypen N_ρ als Funktion des Skalensparameters ρ in logarithmischer Auftragung. **b** Die x_1 -Komponenten der gefundenen Maxima y_σ^k geben einen Einblick in die Entwicklung der Aufspaltung.

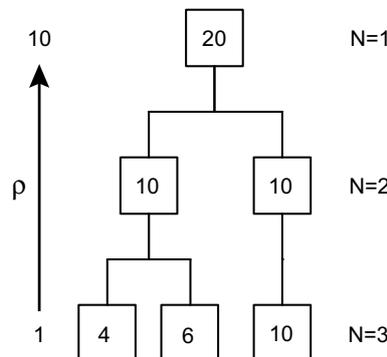


Abbildung 5.11: Die Hierarchie der Prototypen als Baumstruktur. Die minimale Hierarchie der vom ASSO-Verfahren (Abbildung 5.10) in dem zweidimensionalen Beispieldatensatz aufgefundenen Prototypen wird vermittels einer Baumstruktur gezeigt. Diese illustriert die Vater-Kind-Beziehungen der auf verschiedenen Auflösungsstufen ρ existierenden Prototypen y_σ^k . Jeder Prototyp wird durch ein Quadrat repräsentiert, in dem die Anzahl der zugeordneten Kodebuchzentren angegeben ist (zur Dichteschätzung wurden 20 Zentren verwendet). Berücksichtigt man, dass die Datenpunkte im Verhältnis 2:1:1 aus drei Normalverteilungen gezogen wurden, so fasst die gezeigte Darstellung die hierarchische Struktur des Datensatzes in guter Näherung und auf sehr übersichtliche Weise zusammen.

Bis hierhin wurde in der beispielhaften Anwendung des ASSO-Verfahrens der Skalenparameter ρ_s sehr fein abgetastet, so dass in Abbildung 5.10 der Eindruck einer kontinuierlichen Skala entsteht. Versieht man den Skalenparameter mit einem Index: ρ_s mit $s = 1, \dots, S$, dann wurde bisher $S = 100$ gewählt. Die hierarchische Struktur der im Skalenraum der Modelldichte zu beobachtenden Maxima bzw. der Prototypen lässt sich jedoch noch weiter in eine „minimale Hierarchie“ kondensieren. Dazu werden aus der Schar der Mengen der Prototypen \mathcal{Y}_{ρ_s} diejenigen ausgewählt, bei denen eine weitere Verfeinerung der Skala zu einer neueren Aufspaltung führen würde. Auf diese Weise wird für jede auftretende Anzahl N_σ nur genau ein repräsentativer Satz von Prototypen ausgewählt. Nach der Auswahlvorschrift wird dabei genau derjenige Satz von Prototypen selektiert, bei dem kurz vor dem Aufspalten die Orte der Prototypen am besten definiert sind.

Wird die rekurrente Dynamik (5.34) nicht nur innerhalb einer Skalenstufe durchgeführt, sondern werden die Orte der Prototypen einer der selektierten Stufen in der nächsthöheren und damit größeren Stufe zu den dort bestimmten Prototypen zugeordnet, so kann die hierarchische Struktur des der Dichteschätzung zu Grunde liegenden Datensatzes in Form eines abstrakten Baumes dargestellt werden. Abbildung 5.11 zeigt die Hierarchie der Prototypen wiederum an dem bisher betrachteten Beispieldatensatz und verdeutlicht die Aussagekraft dieser Darstellung.

Effiziente zweistufige Klassifizierung

Im Vorangehenden wurde dargestellt, wie durch das ASSO-Verfahren eine unüberwachte Klassifikation eines Datensatzes \mathcal{X} vorgenommen werden kann. Als Ergebnis ist dann, abhängig von der gewählten Auflösungsstufe, die Anzahl der Klassen N_σ bekannt, in die der Datensatz \mathcal{X} einzuteilen ist. Jede Klasse k wird dabei durch einen prototypischen Merkmalsvektor \mathbf{y}_σ^k charakterisiert. Um in einem zweiten Schritt überwacht jeden Datenpunkt \mathbf{x} des Datensatzes einer der gefundenen Klassen zuzuordnen, kann prinzipiell wieder das ASSO-Verfahren herangezogen werden. Die Zuordnungsregel (5.36) definiert dazu eine diskrete und scharfe Klassifikation. Ein solches Vorgehen ist aber mit zwei Problemen verbunden: (i) Eine scharfe Klassifikation ist nicht immer erwünscht, vielmehr kann die Angabe einer fraktionellen Zugehörigkeit eines Datenpunktes zu mehreren Klassen eine wichtige Information darstellen. (ii) Die Klassifikation mit dem ASSO-Verfahren ist sehr rechenaufwändig, da für jeden Datenpunkt \mathbf{x} die rekurrente Dynamik (5.34) vollständig auszuwerten ist.

Aus diesen Gründen soll hier ein zweistufiges Verfahren zur Klassifikation diskutiert werden. Es beruht auf einer Unterteilung der geschätzten Modelldichte $\hat{p}(\mathbf{x}|\Theta)$ (5.18) in klassenlokale Teildichten. Dazu wird das Kodebuch $C = \{\mathbf{c}_1, \dots, \mathbf{c}_R\}$ in N_σ disjunkte Teilmengen C_k zerlegt

$$C = \bigcup_{k=1}^{N_\sigma} C_k \quad \text{mit} \quad C_k \cap C_l = \emptyset \quad \forall k \neq l, \quad (5.39)$$

und zwar derart, dass C_k genau diejenigen Kodebuchzentren enthält, die nach (5.36) dem Prototypen \mathbf{y}_σ^k zugeordnet werden. Die klassenlokale Wahrscheinlichkeitsdichte $\hat{p}(\mathbf{x}|k)$ ist dann ein einfaches Mischungsmodell

$$\hat{p}(\mathbf{x}|k) = \frac{1}{|C_k|} \sum_{r \in C_k} \hat{p}(\mathbf{x}|r). \quad (5.40)$$

Hier bezeichnet $|C_k|$ die Mächtigkeit der Menge C_k . Um unter dieser Annahme die Zuordnungswahrscheinlichkeit $\hat{P}(k|\mathbf{x})$ eines Datenpunkts \mathbf{x} zu einer Konformation \mathbf{y}_σ^k zu berechnen, wendet man den Bayessche Satz (5.19) auf das Mischungsmodell (5.40) an und setzt die bekannten *a priori* Wahrscheinlichkeiten der Klassen

$$\hat{P}_k = \frac{|C_k|}{R}, \quad (5.41)$$

ein. Es zeigt sich dann, dass die $\hat{P}(k|\mathbf{x})$ durch eine einfache Summenbildung aus den Zuordnungswahrscheinlichkeiten zu den einzelnen Kodebuchzentren $\hat{P}(r|\mathbf{x})$ (5.22) berechnet werden können

$$\hat{P}(k|\mathbf{x}) = \sum_{r \in C_k} \hat{P}(r|\mathbf{x}). \quad (5.42)$$

Damit konnte gezeigt werden, dass die Zuordnung der Datenpunkte zu den gefundenen Prototypen eines Datensatzes einfach in einem zweistufigen Verfahren durchgeführt werden kann: zuerst werden die Kodebuchzentren aufwändig durch das ASSO-Verfahren zu den Prototypen assoziiert. Für jeden Datenpunkt sind dann nur die Zuordnungswahrscheinlichkeiten zu den Kodebuchzentren $\hat{P}(r|\mathbf{x})$ zu berechnen und klassenlokal aufzaddieren.

5.4 Zirkuläre Statistik

Am Beginn dieses Kapitels wurde erwähnt, dass die statistische Analyse so genannter *zirkulärer Variablen* mit besonderen Problemen verbunden ist. Eine zirkuläre Variable zeichnet sich dadurch aus, dass eine durch sie festgelegte Beobachtung invariant unter der Addition des Vielfachen einer bestimmten Periodenlänge ist. Bei Winkelvariablen im Bogenmaß beträgt diese Länge 2π , so dass die Transformation

$$\alpha' = \alpha + p 2\pi \quad \text{mit} \quad p \in \mathbb{Z} \quad (5.43)$$

keine Auswirkung auf die Beobachtung hat. In später folgenden Abbildungen und Auswertungen der Simulationsdaten werde ich den Wertebereich $[-\pi, \pi]$ und die entsprechende Angabe in Grad ($[-180^\circ, 180^\circ]$) austauschbar verwenden.

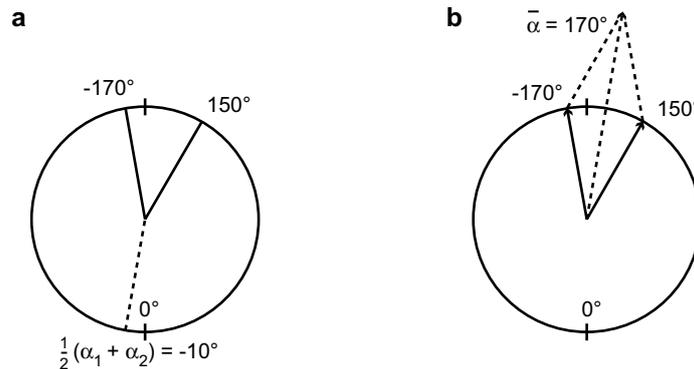


Abbildung 5.12: Der Mittelwert zweier Winkelvariablen. Beispielhaft werden zwei Winkel, -170° und 150° , auf dem Einheitskreis dargestellt. **a** Werden diese beiden Winkel wie einfache lineare Größen behandelt und ihr Mittelwert berechnet, so sieht man, dass der berechnete Winkel genau die falsche Richtung angibt. **b** Der richtige Mittelwert zirkulärer Variablen $\bar{\alpha}$ kann durch die Addition der zugehörigen Einheitsvektoren bestimmt werden. Details im Text.

Wendet man die bisher vorgestellten statistischen Verfahren in naiver Weise auf zirkuläre Variablen an, so werden schnell gravierende Probleme offenbar [232]. Abbildung 5.12 a erklärt, dass schon die Berechnung des Mittelwertes zweier oder mehrerer Winkelvariablen nicht einfach durch die Berechnung des arithmetischen Mittels (5.28) bewerkstelligt werden kann. Vielmehr ist es notwendig, den in Abbildung 5.12 b gezeigten Weg zu gehen: statt als skalare Größen werden die Winkel als Ortsvektoren der durch sie bestimmten Punkte auf dem Einheitskreis betrachtet und diese Vektoren komponentenweise aufaddiert. Die Richtung des resultierenden Vektors gibt dann den mittleren Winkel an, während seine Länge eine Aussage über die Definiertheit dieses Mittelwertes macht [232]. Die Komponenten der Vektoren sind dabei offensichtlich durch den Sinus und den Kosinus der Winkel gegeben, so dass der *zirkuläre Mittelwert* als

$$\bar{\alpha} = \langle \alpha \rangle_{\mathcal{A}} = \arctan^* \left(\frac{\sum_{\alpha \in \mathcal{A}} \sin \alpha}{\sum_{\alpha \in \mathcal{A}} \cos \alpha} \right) \quad (5.44)$$

definiert werden kann. Weil die übliche Arkustangens-Funktion nur einen Wertebereich von $-\frac{\pi}{2}, \frac{\pi}{2}$ aufweist, muss hier die mit \arctan^* bezeichnete Erweiterung dieser Funktion auf alle vier Quadranten des Vollkreises eingesetzt werden. Die in den meisten Programmiersprachen (beispielsweise C und auch IDL) vorhandene Arkustangens-Implementierung zeigt genau die gewünschten Eigenschaften.

Auch die Differenz zweier zirkulärer Variablen bedarf einer gesonderten Definition, wofür auch wieder mehrere mögliche Vorgehensweisen existieren. Hier wurde die *zirkuläre Differenz*

$$\alpha \ominus \beta = [(\alpha - \beta + \pi) \bmod 2\pi] - \pi \quad (5.45)$$

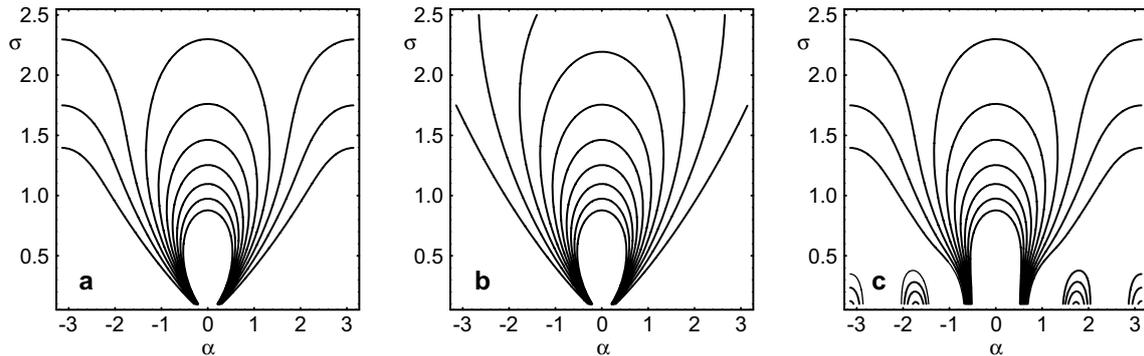


Abbildung 5.13: Die Zirkuläre Normalverteilung und ihre Darstellungen. **a** Die um den Ursprung zentrierte zirkuläre Normalverteilung $g_w(\alpha|0, \sigma)$ wird als zweidimensionale Funktion des Winkels α und der Breite σ dargestellt. Die in dieser Teilabbildung gezeigte Referenz der „wahren“ zirkulären Verteilung wurde unter Berücksichtigung einer großen Zahl ($P = 100$) von Termen aus (5.47) berechnet. Die Konturlinien sind äquidistant gewählt, jedoch ist der dargestellte Wertebereich nach oben hin beschränkt. Man sieht, dass für Breiten die sich π annähern, die zirkuläre Normalverteilung zusehends flacher wird. **b** Zum Vergleich wird die nicht-zirkuläre Normalverteilung $g(\alpha|0, \sigma)$ dargestellt, die in der UNIVAR-Implementierung als Darstellung der zirkulären Verteilung für kleine Gaußbreiten σ gewählt wurde. Für $\sigma < 0,7$ ist keine Abweichung zwischen den Abbildungen a und b sichtbar. **c** Für größere σ wird die Darstellung $g_c(\alpha|\mu, \sigma)$ (5.48) der zirkulären Verteilung verwendet, die auf einer Entwicklung nach Kosinusfunktionen beruht. Hier wurden $P = 4$ Terme der Entwicklung berücksichtigt. Für kleine σ sind periodische Artefakte zu erkennen, für $\sigma > 0,7$ jedoch wieder eine gute Übereinstimmung mit der Abbildung a.

so gewählt, dass ihr Betrag dem Abstand von α zu der am nächsten liegenden periodischen Kopie von β entspricht, das Vorzeichen die entsprechende Drehrichtung angibt und der Wertebereich der Differenz wieder $[-\pi, \pi]$ beträgt.

Jammalamadaka und SenGupta [232] schlagen weiterhin auch mehrere alternative Maße für eine *zirkuläre Kovarianz* zwischen zwei Variablen α_i und α_j vor. Die Definition

$$C(\mathcal{A})_{i,j} = \langle \sin(\alpha_i - \bar{\alpha}_i) \sin(\alpha_j - \bar{\alpha}_j) \rangle_{\mathcal{A}} \quad (5.46)$$

hat neben einer guten Anschaulichkeit und einfachen Implementierung auch den Vorteil, für den Fall recht kompakter Datenverteilungen näherungsweise in die normale nicht-zirkuläre Kovarianzmatrix (5.9) des Datensatzes überzugehen.

Eine direkte Übertragung der Normalverteilung (5.17) auf den Fall zirkulärer Daten, kann erreicht werden, indem diese um den Einheitskreis „gewickelt“ wird (engl. wrapped normal distribution, WND)

$$g_w(\alpha|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{p=-\infty}^{\infty} \exp\left[-\frac{(\alpha - \mu - 2\pi p)^2}{2\sigma^2}\right]. \quad (5.47)$$

Es existieren zwar auch hier mehrere alternative zirkuläre Verteilungen, die Eigenschaften

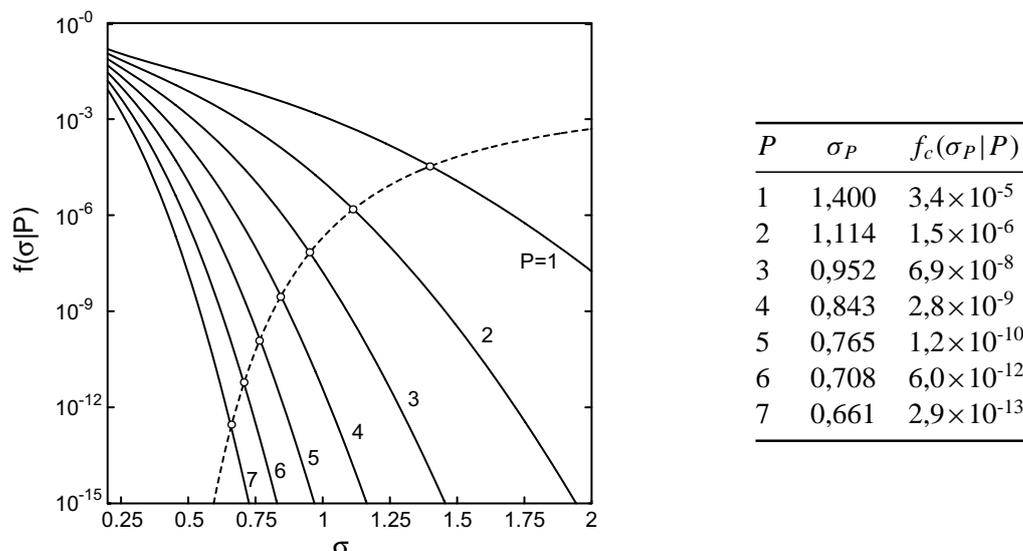


Abbildung 5.14: Fehler in den Darstellungen der zirkulären Normalverteilung. In der linken Abbildung wird $f_n(\sigma)$ (5.49), die integrale quadratische Abweichung der nicht-zirkulären von der zirkulären Normalverteilung, als Funktion der Gaußbreite σ gestrichelt dargestellt. Die durchgezogenen Linien zeigen die $f_c(\sigma|P)$ (5.50), die Fehler der Darstellung (5.48) unter der Berücksichtigung von P von Kosinus-Termen. Die hervorgehobenen Schnittpunkte der Graphen bestimmen nach Gleichung (5.51) die optimalen Schwellwerte σ_P . Die Ergebnisse werden zusätzlich in der Tabelle zusammengefasst.

einer Normalverteilung zeigen (besonders häufig wird die von-Mises-Verteilung betrachtet [232]), doch hat die hier verwendete Definition (5.47) den entscheidenden Vorteil, dass die Faltung einer solchen Verteilung mit einer Gaußfunktion (5.33) wieder eine zirkuläre Normalverteilung mit passender Varianz ergibt. Nur mit dieser Wahl ist eine Übertragung der Theorie des ASSO-Verfahrens auf zirkuläre Variablen möglich.

Wird anstelle der einfachen univariaten Normalverteilung (5.17) deren zirkuläre Entsprechung (5.47) in den beschriebenen Verfahren der Dichteschätzung eingesetzt, so muss die Berechnung der Reihenentwicklung in (5.47) bei einem festzulegenden Term P abgebrochen werden. Die Summation über p (5.47) läuft dann nur noch von $-P$ bis P . Während die zirkuläre Verteilung für kleine Breiten σ in die nicht-zirkuläre Form übergeht und auch für $\sigma \ll \pi$ eine Berücksichtigung nur der ersten Ordnung der Überträge ($P = 1$) notwendig ist, steigt das zu wählende P für $\sigma > \pi$ rapide an. In diesem Fall ist eine alternative Darstellung der zirkulären Normalverteilung sehr hilfreich [232, 269]

$$g_c(\alpha|\mu, \sigma) = \frac{1}{2\pi} \left[1 + 2 \sum_{p=1}^{\infty} \varrho^{2p} \cos p(\alpha - \mu) \right] \quad \text{mit} \quad \varrho = e^{-\frac{\sigma^2}{2}}, \quad (5.48)$$

da diese eine Entwicklung nach periodischen Kosinus-Funktionen darstellt. Damit konvergiert g_c (5.48) für große σ deutlich schneller als g_w (5.47). Umgekehrtes gilt jedoch im

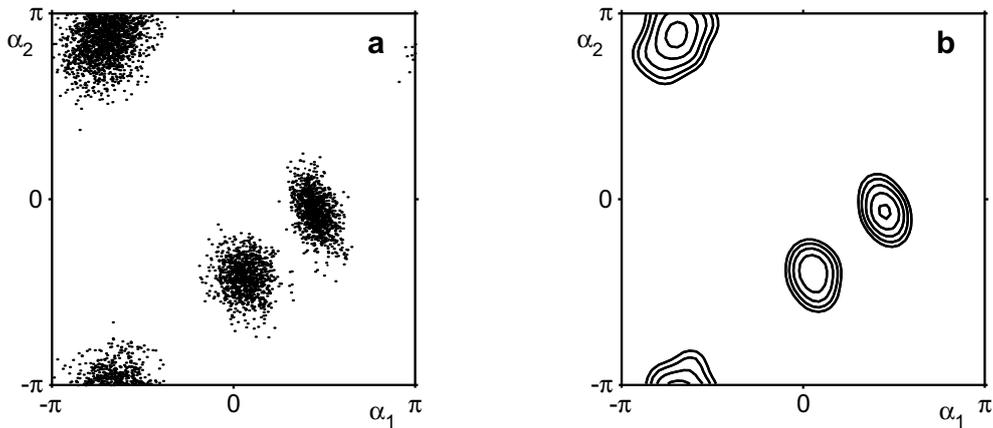


Abbildung 5.15: Dichteschätzung im Fall zirkulärer Variablen. **a** Eine leichte Abänderung des bisher betrachteten Datensatzes (siehe Abbildung 5.3) führt zu einem zirkulären Datensatz, bei dem eine der Teilverteilungen über die Grenze des Wertebereichs hinausragt und zurückgefaltet wird. **b** Wie zuvor wurde eine Dichteschätzung mit einer Mischung von $R = 20$ univariaten, hier aber zirkulären Normalverteilungen, vorgenommen.

Fall kleiner Gaußbreiten. Abbildung 5.13 verdeutlicht dies auf anschauliche Weise.

Bei der Erweiterung der Implementierung des UNIVAR-Algorithmus zur Berücksichtigung der zirkulären Statistik wird für kleine σ die nicht-zirkuläre Normalverteilung (5.17) verwendet und ab einem gewissen Schwellwert σ_P auf die Kosinus-Darstellung g_c umgeschaltet. Der Index P deutet dabei an, dass σ_P natürlich in Abhängigkeit von der vorgegebenen Anzahl der zu berücksichtigenden Kosinus-Terme bestimmt werden muss. Als Fehlermaß wurde für einige Werte von P numerisch das Integral der quadratischen Abweichung der beiden Darstellungen von der „wahren“ Verteilung bestimmt, welche unter Aufsummierung vieler (100) Terme aus (5.47) berechnet wurde

$$f_n(\sigma) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [g_w(\alpha'|0, \sigma) - g(\alpha'|0, \sigma)]^2 d\alpha' \quad (5.49)$$

$$f_c(\sigma|P) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [g_w(\alpha'|0, \sigma) - g_c^P(\alpha'|0, \sigma)]^2 d\alpha'. \quad (5.50)$$

Der optimale Schwellwert σ_P ist dadurch bestimmt, dass die Fehler der beiden möglichen Darstellungen gerade gleich groß sind

$$f_n(\sigma_P) \stackrel{!}{=} f_c(\sigma_P|P). \quad (5.51)$$

Abbildung 5.14 zeigt den Verlauf der Fehlerfunktionen (5.49) und (5.50) in Abhängigkeit von der Gaußbreite σ und die numerisch bestimmten Punkte, an denen sich diese Graphen

schneiden. Die so bestimmten Schwellwerte σ_P werden erwartungsgemäß kleiner, je mehr Terme in der Kosinus-Entwicklung verwendet werden. Die Tabelle in Abbildung 5.14 gibt mithin auch eine Abschätzung des maximal auftretenden, über α gemittelten Fehlers in der Berechnung der zirkulären Normalverteilung an.

Abbildung 5.15 illustriert den Erfolg der Erweiterung des UNIVAR-Algorithmus um die Behandlung zirkulärer Merkmalsvektoren. Eine hier nicht gezeigte Anwendung des ASSO-Verfahrens auf der Modelldichte würde unweigerlich mögliche Fehler oder Ungenauigkeiten im Übergang zwischen den diskutierten Darstellungen der zirkulären Normalverteilung aufzeigen, da diese Methode auf der Annahme einer glatten Vergrößerung der Modelldichte basiert. Testrechnungen zeigten jedoch, dass beim Übergang von einer Darstellung zur anderen keine willkürlichen Aufspaltungen von Prototypen oder Diskontinuitäten auftraten.

Mit der Erweiterung der statistischen Verfahren um die Möglichkeit, Winkelvariablen angemessen behandeln zu können, sind nun alle Voraussetzungen dafür geschaffen worden, um die MD-Simulationen der Modellpeptide strukturell auswerten zu können. Die Erweiterungen wurden in den entsprechenden Programmen implementiert und der Arbeitsgruppe als Werkzeuge zur Verfügung gestellt.

6 Validierung des Peptid-Kraftfeldes: das Alanin-Dipeptid

Nachdem in Kapitel 3 Modelle für das Lösungsmittel DMSO vorgestellt, in Kapitel 4 das eingesetzte Kraftfeld diskutiert und schließlich im letzten Kapitel statistische Methoden zur Analyse der Simulationsdaten präsentiert wurden, sind nun alle notwendigen „Zutaten“ für die MD-Simulationen und deren Auswertung vorhanden. Bevor in den nächsten beiden Kapiteln die Ergebnisse der Simulationen der lichtschtbaren Modellpeptide vorgestellt werden können, ist es jedoch angebracht, die verwendeten Methoden und vor allem das eingesetzte Peptid-Kraftfeld anhand eines gut charakterisierten einfachen Beispielmoleküls zu validieren.

Eine solche Validierung erschien insbesondere deshalb angebracht, da im Laufe der vorgestellten Arbeit deutlich wurde, dass eventuell Unstimmigkeiten zwischen dem publizierten Peptidkkräftfeld [81] und seiner in unserer Arbeitsgruppe eingesetzten Implementierung bestanden. Um diese Problematik zu erläutern sei noch einmal auf Abschnitt 2.2 zurückverwiesen, in dem die CHARMM-Energiefunktion (2.13) eingeführt wurde. Ein Detail dieses Kraftfeldes ist die Schaltfunktion (2.25), welche die Coulomb-Wechselwirkung (2.24) für solche Atome abschaltet oder vermindert, die chemisch über höchstens drei kovalente Bindungen miteinander verbunden sind. Wie Gleichung (2.25) ausweist, wird der Wert dieser Abschaltfunktion bei einem Abstand der Atome von genau drei Bindungen durch einen Parameter λ_{14} gegeben. Es wurde auch schon erwähnt, dass der Wert von λ_{14} beim Übergang von Version 19 auf Version 22 des CHARMM-Kraftfelds von zuerst 0,4 auf den Wert 1,0 geändert wurde [77, 81].

Die ersten Versionen unseres MD-Programms EGO verwendeten die ältere der beiden Versionen des Kraftfeldes (CHARMM19) und zwar in einer leicht veränderten Form, welche mit dem Programmpaket XPLOR [160] ausgeliefert wurde. Dies ist der Grund dafür, dass ($\lambda_{14} = 0,4$) als Standardwert Aufnahme in die Befehlsdateien von EGO fand [270]. Allerdings wurde einige Jahre später in der Arbeitsgruppe zur Verwendung des neueren, explizite Wasserstoffatome umfassenden Kraftfelds CHARMM22 übergegangen, zu dem jedoch zu dem damaligen Zeitpunkt noch keine Veröffentlichung vorlag. Die entsprechende Publikation erschien erst 1998 [81], weshalb bei der Einführung des CHARMM22-Kraftfeldes in EGO die Information übersehen wurde, dass diese Version des Kraftfeldes auf einem λ_{14} von 1,0 basiert.

Nachdem ich die Abweichung des in der Arbeitsgruppe bisher verwendeten Kraftfeldes [270] von seiner publizierten Form [81] bemerkt hatte, stellte sich die Frage nach den

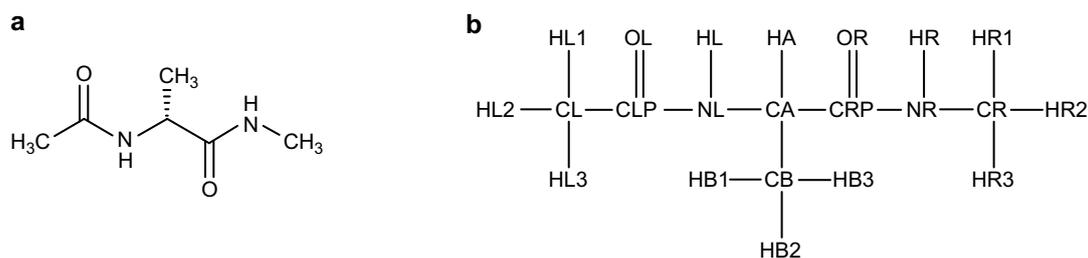


Abbildung 6.1: Das Alanin-Dipeptid. **a** Die chemische Struktur des Alanin-Dipeptids. **b** Die dazugehörige Topologie, in der die Atomnamen und chemischen Bindungen festgelegt werden. Diese Topologie wurde nicht, wie für Peptide üblich, aus der des Alanin-Residuums und zweier Termini gebildet, sondern existiert in der zum CHARMM22-Kraftfeld gehörigen Datei „topallh22x.mod“ als eigenständige Definition.

Auswirkungen der Wahl des Nahfeld-Skalierungsfaktors λ_{14} auf die Struktur und Dynamik von Peptiden.

Anstatt entsprechende Untersuchungen an den relativ komplexen Modellpeptiden durchzuführen, entschied ich mich, Vergleichsrechnungen an einem kleineren und besser untersuchten Molekül vorzunehmen. Dazu bot sich das so genannte *Alanin-Dipeptid*^a an, dessen chemische Struktur und Topologie in Abbildung 6.1 gezeigt werden. Von besonderem Vorteil ist, dass für dieses Molekül sowohl die Geometrien seiner Minimalstrukturen bekannt sind, als auch seine durch die Referenzimplementierung des CHARMM22-Kraftfelds gegebene Energielandschaft vorliegt [81].

Das Alanin-Dipeptid umfasst ein einzelnes Alanin-Residuum und zwei vollständige Peptidgruppen, welche durch Methylgruppen terminiert sind. Es kann somit als einfachstes Modell einer in ein Proteinerückgrat eingebundenen Aminosäure angesehen werden. Seine Konfiguration lässt sich in guter Näherung durch die Angabe der beiden Dihedralwinkel ϕ und ψ des Alanin-Residuums beschreiben (vgl. Abschnitt 5.1) und damit direkt in einem zweidimensionalen Ramachandran-Plot visualisieren. Wegen dieser einfachen Darstellbarkeit und seiner geringen Zahl von Atomen, stellt das Alanin-Dipeptid so etwas wie das „Wasserstoffatom“ oder die „weiße Labormaus“ der computergestützten molekularen Biophysik dar.

6.1 Minimalstrukturen im Vakuum

In einem ersten Schritt wurden mit EGO-MMII die Minimalstrukturen des Alanin-Dipeptids im Vakuum ermittelt, um deren Geometrien und potentielle Energien vergleichen zu können. Zu diesem Zweck wurden MD-Simulationen durchgeführt, denen das MM-

^aN-Acetyl Alanin N'-Methyl Amid

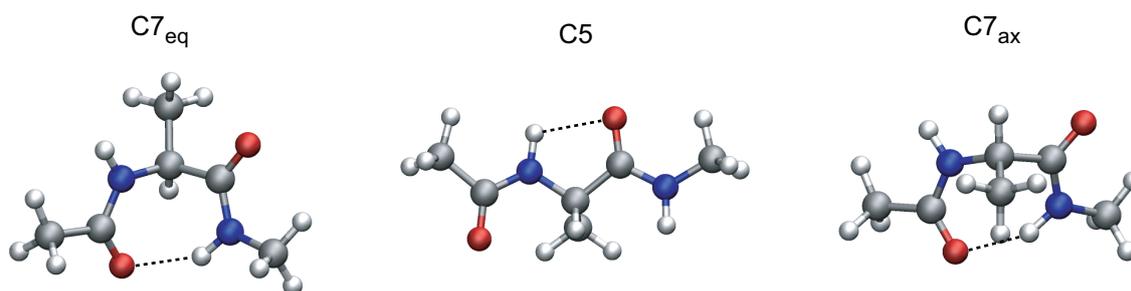


Abbildung 6.2: Drei stabile Strukturen des Alanin-Dipeptids im Vakuum. Unter Verwendung des CHARMM22-Kraftfeldes ($\lambda_{14} = 1,0$) weist das Alanin-Dipeptid im Vakuum drei Minimalstrukturen auf, die üblicherweise als $C7_{eq}$, $C5$ und $C7_{ax}$ bezeichnet werden. Die zugehörigen potentiellen Energien und Geometrien werden in Tabelle 6.1 aufgeführt.

| Struktur | | EGO | | CHARMM ^a | HF ^b | DFT ^c | MP2 ^d |
|-----------|-----------------|----------------------|----------------------|---------------------|-----------------|------------------|------------------|
| | | $\lambda_{14} = 0,4$ | $\lambda_{14} = 1,0$ | | | | |
| $C7_{eq}$ | ΔE_{MM} | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| $C5$ | ΔE_{MM} | -3,01 | 0,90 | 0,92 | 0,40 | 1,41 | 1,39 |
| $C7_{ax}$ | ΔE_{MM} | 3,71 | 2,04 | 2,05 | 2,82 | 2,60 | 2,66 |
| $C7_{eq}$ | ϕ | -88,7 | -81,3 | -81,3 | -85,8 | -82,8 | -82,6 |
| | ψ | 76,9 | 70,6 | 70,6 | 79,0 | 73,3 | 75,8 |
| $C5$ | ϕ | -157,5 | -151,6 | -151,4 | -157,2 | -158,0 | -161,1 |
| | ψ | 176,6 | 171,0 | 170,6 | 159,8 | 163,9 | 155,5 |
| $C7_{ax}$ | ϕ | 75,1 | 69,8 | 69,7 | 76,0 | 73,6 | 73,7 |
| | ψ | -74,2 | -67,8 | -67,6 | -55,4 | -57,8 | -53,7 |

Tabelle 6.1: Die Energien und Minimalstrukturen des Alanin-Dipeptids im Vakuum. Für die drei in Abbildung 6.2 gezeigten Minimalstrukturen des Alanin-Dipeptids im Vakuum werden sowohl die inneren Energien ΔE [kcal/mol] relativ zu $C7_{eq}$ als auch die jeweiligen Geometrien angegeben. Neben den mit EGO-MMII optimierten Konfigurationen (unter Verwendung der beiden diskutierten Werte für λ_{14}) werden auch verschiedene Referenzwerte angegeben: ^a Literaturwert des CHARMM22-Kraftfeldes [81], ^b HF/6-31G* [271], ^c B3LYP/6-31G* (vorliegende Arbeit), ^d MP2/CBS limit [272].

Kraftfeld CHARMM22 mit den beiden alternativen Werten für λ_{14} zu Grunde gelegt wurde. Ausgehend von unterschiedlichen Startkonfigurationen des Dipeptids, d.h. mehreren unterschiedlichen ϕ/ψ Winkelkombinationen, wurde jeweils durch eine dissipative Geschwindigkeitsskalierung eine Strukturminimierung erreicht. Zusätzlich wurden ähnliche Optimierungen auch mit dem Programmpaket GAUSSIAN unter Verwendung der DFT-Methode B3LYP/6-31G* durchgeführt.

Abbildung 6.2 zeigt jene drei Minimalstrukturen, die für das Alanin-Dipeptid im Vakuum durch das CHARMM22-Kraftfeld festgelegt werden. Ergänzend hierzu führt Tabelle 6.1 die Geometrien dieser Minimalstrukturen und ihre Energiedifferenzen auf. Dabei werden nicht nur die Ergebnisse für die zwei sich im Skalierungsfaktor λ_{14} unterscheidenden Implementation des CHARMM22 Kraftfeldes präsentiert, sondern auch die Referenzwerte, die zusammen mit der Parametrisierung des CHARMM22-Kraftfeldes veröffentlicht wurden [81]. Weiterhin beinhaltet Tabelle 6.1 auch die Ergebnisse der DFT-Rechnung und Literaturwerte die mit HF- [271, 273] oder aufwändigen MP2-Rechnungen [272] ermittelt wurden. Zu allen bisher veröffentlichten QM-Rechnungen am Alanin-Dipeptid im Vakuum bietet der Artikel von Vargus et al. [272] einen umfassenden Überblick.

Die in Abbildung 6.2 verwendeten Bezeichnungen der drei Minimalstrukturen des Dipeptids im Vakuum beziehen sich auf die Anzahl der Atome, die der durch eine Wasserstoffbrückenbindung stabilisierte Ring von Atomen beinhaltet: sieben Atome im Fall von $C7_{eq}$ bzw. $C7_{ax}$ und fünf bei der $C5$ -Konfiguration. Das Kürzel „eq“ bezeichnet die äquatoriale Lage des C_{β} außerhalb des Rings, „ax“ seine axiale Lage innerhalb des Rings der anderen Atome. Betrachtet man die gestreckte Form der Struktur $C5$ und vergleicht man die in Tabelle 6.1 aufgeführten dazugehörigen Dihedralwinkel mit den Ramachandran-Plots aus Abbildung 1.4, so wird deutlich, dass diese Struktur der eines Ausschnitts einer Peptidkette in einem β -Faltblatt entspricht. Dagegen können die beiden anderen Strukturen nicht einfach einer der vorgestellten Protein-Sekundärstrukturen zugeordnet werden. Beide Strukturen machen einen eher gefalteten Eindruck, jedoch liegen die Dihedralwinkel von $C7_{eq}$ in einem Bereich des Ramachandran-Plots, der bei β -Faltblättern auftritt und die Winkel der Struktur $C7_{ax}$ fallen in einen Bereich, der in Abbildung 1.4 als sterisch verboten klassifiziert wurde. Erst eine genauere Betrachtung zeigt, dass diese beiden Minimalstrukturen so genannten γ -Turns [274, 275] entsprechen, die bei Faltblättern drei Residuen umfassen und in denen das Peptidrückgrat um fast 180° auf sich zurückgefaltet wird. Zwar fehlen dem Alanin-Dipeptid natürlich Teile des ersten und des dritten Residuums, die Wasserstoffbrücke, die beiden Dihedralwinkel des Alanins und die Atomanordnung stimmen für die Minimalstruktur $C7_{ax}$ jedoch sehr gut mit der Definition des „klassischen“ γ -Turns überein [274]. Gleichermäßen entspricht $C7_{eq}$ einem „inversen“ γ -Turn [274].

Ein Vergleich der beiden in Tabelle 6.1 aufgeführten EGO-Minimalstrukturen, die aus der Verwendung der beiden alternativen Werte für λ_{14} resultieren, zeigt bei den Dihedralwinkeln nur relativ geringe Unterschiede ($< 10^\circ$). Eine gute Übereinstimmung mit den CHARMM22-Referenzwerten kann allerdings nur für $\lambda_{14} = 1,0$ beobachtet werden. Ver-

gleicht man weiterhin diese MM-Strukturen mit den entsprechenden Ergebnissen aus den QM-Rechnungen so ist zu beobachten, dass, obwohl das CHARMM22-Kraftfeld teilweise anhand von HF-Rechnungen am Alanin-Dipeptid optimiert wurde, die CHARMM-Geometrien nicht den HF-Geometrien entsprechen. Die mit den akkurateren DFT und MP2-Methoden bestimmten Geometrien sind untereinander sehr gut vergleichbar, gegenüber den CHARMM-Geometrien sind hier bei den C5- und C7_{ax}-Strukturen jedoch um etwa 15° kleinere Werte für ψ zu beobachten.

Die zusätzlich zu den Geometrien in Tabelle 6.1 aufgeführten potentiellen Energien der EGO-Minimalstrukturen relativ zu C7_{eq} weisen einen deutlichen Einfluss des Skalierungsfaktors λ_{14} auf. Das Ergebnis für den Literaturwert $\lambda_{14} = 1,0$ stimmt gut mit den CHARMM-Referenzwerten überein, insbesondere darin, dass C7_{eq} der energetisch günstigere Zustand ist, während C5 um etwa 1 kcal/mol und C7_{ax} um ungefähr 2 kcal/mol darüber liegen. Bei einer Wahl von $\lambda_{14} = 0,4$ wird die energetische Reihenfolge der beiden geometrisch relativ ähnlichen Strukturen C7_{eq} und C5 jedoch vertauscht, C5 liegt hier etwa 3 kcal/mol unterhalb von C7_{eq}. Die aus den QM-Berechnungen ermittelten Energien bestätigen das CHARMM-Kraftfeld, wobei mit den akkurateren Methoden etwas größere Energiedifferenzen gefunden werden.

Zusammenfassend konnte in diesem Unterabschnitt aufgezeigt werden, dass die Geometrien und Energetik der mit EGO-MMII berechneten Minimalstrukturen des Alanin-Dipeptids gut mit den publizierten Resultaten des CHARMM22-Kraftfeldes übereinstimmen, sobald der Literaturwert von $\lambda_{14} = 1,0$ für den Skalierungsfaktor der elektrostatischen 1-4-Wechselwirkung verwendet wird. In diesem Fall ist auch eine zufriedenstellende Übereinstimmung mit den Resultaten der DFT- und MP2-Rechnungen zu beobachten. Im Fall der Wahl von $\lambda_{14} = 0,4$ sind zwar nur geringe Abweichungen bei den Geometrien der drei Strukturen des Dipeptids zu beobachten, es tritt jedoch eine auffällige Überstabilisierung der gestreckten β -Faltblatt-ähnlichen C5-Struktur gegenüber den γ -Turn-ähnlichen Strukturen C7_{eq} und C7_{ax} auf.

6.2 Energieflächen im Vakuum und in Lösung

Zwar hat die oben betrachtete Lage der Minima in der durch die verschiedenen Kraftfelder definierten potentiellen Energie des Alanin-Dipeptids einen ersten Einblick in die Auswirkungen der jeweiligen Wahl des Skalierungsfaktors λ_{14} ermöglicht, jedoch wird die bei einer MD-Simulation beobachtete Dynamik eines Moleküls durch die Form seiner vollständigen Energielandschaft bestimmt. Wie zuvor sollen in diesem Abschnitt die durch den Skalierungsfaktor λ_{14} induzierten Veränderungen der Energiefläche eines beliebigen Peptids durch die stellvertretende Betrachtung des Alanin-Dipeptids untersucht werden. Wie seine Konfiguration kann die potentielle Energie $U(\mathbf{R}) \equiv E_{MM}(\mathbf{R})$ (2.13) des Dipeptids in guter Näherung als Funktion der beiden Dihedralwinkel ϕ und ψ angegeben werden, was wiederum eine einfache Darstellung und Diskussion ermöglicht. Noch

wichtiger als ein Verständnis der Veränderungen an der potentiellen Energie $U(\phi, \psi)$ ist für die späteren Simulationen der in DMSO gelösten Modellpeptide allerdings eine Untersuchung der freien Energie $F(\phi, \psi)$ des in Wasser gelösten Alanin-Dipeptids, da die Dynamik der Modellpeptide auf dieser durch Solvatationseffekte und entropische Beiträge modifizierten Energiefläche stattfindet.

Die zu einer solchen Untersuchung der freien Energie notwendige MD-Simulation des Alanin-Dipeptids in explizitem Wasser wurde erstmals 1979 von der Gruppe um Martin Karplus vorgestellt [276, 277]. Mit der Zunahme der Leistungsfähigkeit der verfügbaren Computer konnten später längere Trajektorien berechnet werden, was es Tobias und Brooks [278] 1992 ermöglichte, den Einfluss des Wassers auf die Konformationen des Dipeptids detailliert zu untersuchen. Seitdem wurden viele der im Bereich der MD-Simulationen neu entwickelten Methoden zuerst am Alanin-Dipeptid getestet. Dies gilt besonders für solche, die dazu dienen, auf effektive Weise den Konfigurationsraum eines Moleküls abzutasten [279–283], wie z.B. für das „Adaptive Umbrella-Sampling“ [284], das von Bartels und Karplus [285] und auch Smith [286] eingesetzt wurde, um die zweidimensionale Fläche der freien Energie des Dipeptids möglichst akkurat zu bestimmen. Auch die Reaktionspfade zwischen den durch hohe Barrieren getrennten Konformeren des Alanin-Dipeptids in Lösung wurden mehrfach untersucht [287, 288].

Erwähnenswert im Zusammenhang mit der Abschätzung der Qualität von MM-Kraftfeldern ist schließlich die vor kurzem erschienene Arbeit von Hu et al. [289], in welcher die Dynamik des quantenmechanisch beschriebenen und in Wasser gelösten Alanin-Dipeptids mittels einer hybriden QM/MM-Methode über einen Zeitraum von 5 ns verfolgt wurde. Hu et al. [289] haben die aus dieser Hybridtrajektorie berechnete Landschaft der freien Energie mit den entsprechenden Ergebnissen von klassischen MD-Simulationen verglichen und dabei deutliche Abweichungen, nicht nur zwischen den verschiedenen verfügbaren MM-Kraftfeldern, sondern insbesondere im Vergleich zu den Ergebnissen der verwendeten QM/MM-Methode gefunden. Angesichts der bei der verwendeten QM-Methode (SCCDFTB) [290] eingesetzten Näherungen muss dies jedoch nicht bedeuten, dass die QM/MM-Fläche der freien Energie verlässlicher ist als jene, die von den klassischen Simulationen vorhergesagt werden. Die in dieser Publikation und in anderen MD-Untersuchungen beobachtete beträchtliche Bandbreite der Ergebnisse zeigt lediglich die Unsicherheiten auf, mit denen gegenwärtig noch alle Versuche theoretischer Beschreibungen der Konformationsdynamik von Peptiden behaftet sind.

Zur Bestimmung der potentiellen Energie des Alanin-Dipeptids im Vakuum $U(\phi, \psi)$ wurden die Dihedralwinkel ϕ und ψ in 6° -Schritten abgetastet. Für jedes der vorgegebenen Winkelpaare wurde in einem mehrstufigen Prozess der Energieminimierung zunehmend stärkere Zusatzpotentiale an diese beiden Dihedralwinkel angelegt, während die restlichen Freiheitsgrade des Moleküls relaxieren konnten. Auf diese Weise konnte $U(\phi, \psi)$ glatt abgetastet werden. Das entsprechende Ergebnis für die Wahl von $\lambda_{14} = 0,4$ wird im Folgenden mit $U_{0,4}$ bezeichnet, während unter Verwendung des Literaturwertes $\lambda_{14} = 1,0$ die Energiefläche $U_{1,0}$ berechnet wurde.

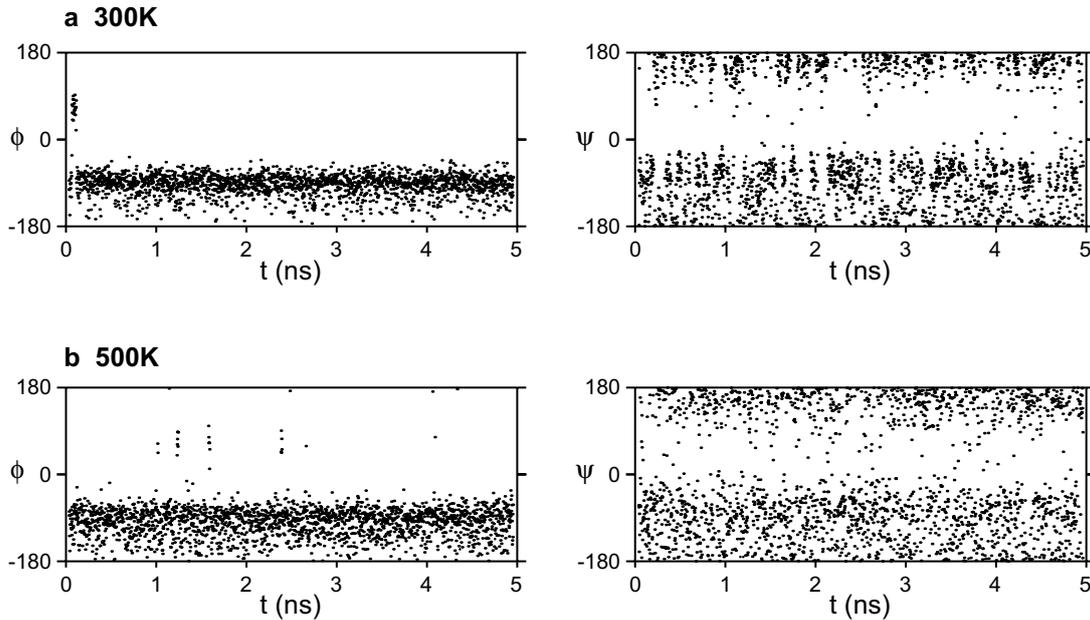


Abbildung 6.3: Zeitserien der Dihedralwinkel des Alanin-Dipeptids in Lösung. **a** zeigt die Zeitentwicklung der Dihedralwinkel während der Trajektorie, die bei 300 K unter der Wahl von $\lambda_{14} = 1,0$ berechnet wurde. **b** die entsprechende Zeitserie bei 500 K.

Die Bestimmung der freien Energie $F(\phi, \psi)$ basierte auf MD-Simulationen von jeweils 5 ns Dauer. Dazu wurde das Dipeptid in einen mit etwa 900 Wassermolekülen gefüllten Rhombendodekaeder eingesetzt. Für das Wasser wurde eine an das CHARMM22-Kraftfeld angepasste [81] und flexible Form des TIP3P Wassermodells [292] eingesetzt. Im NVT-Ensemble wurden für jeden der beiden diskutierten Werte von λ_{14} Simulationen bei zwei verschiedenen Temperaturen, 300 K und 500 K, durchgeführt. Der Thermostat [101] wurde mit einer Relaxationszeit von $\tau_T = 0,5$ ps an das System angekoppelt, das Volumen wurde so gewählt, dass eine Dichte von $1,0 \text{ g/cm}^3$ vorlag. Die Länge der Zeitschritte betrug 1 fs und der MSHAKE-Algorithmus [159] hielt die Bindungslängen aller Wasserstoffe konstant. Die Elektrostatik wurde mit FAMUSAMM (siehe Abschnitt 2.2) bis zur Ebene der strukturellen Einheiten unter Berücksichtigung ihrer Quadrupolmomente ausgewertet, wobei ein Reaktionfeld [100] das weiter entfernte Wasser beschrieb, dessen Dielektrizitätskonstante (DK) $\varepsilon = 80$ betrug. Die in Anhang B vorgestellte Korrektur der VDW-Wechselwirkung wurde angewendet. Schließlich wurden die berechneten Konfigurationen des Dipeptids während der Simulationsdauer von 5 ns alle 500 fs abgespeichert. Das Vorgehen, um aus diesen Trajektorien die freie Energie $F(\phi, \psi)$ zu berechnen, wurde in Kapitel 5 erläutert. Ergänzend bleibt hier nur anzumerken, dass zur Schätzung der invarianten Dichte $\varrho(\phi, \psi)$ aus der Trajektorie der ϕ - und ψ -Winkel eine Mischung von $R = 100$ univariaten Normalverteilungen verwendet wurde. Analog zur potentiellen Energie wird auch für die freie Energie die Notation $F_{0,4}$ bzw. $F_{1,0}$ verwendet, um die beiden alternativen Werte von λ_{14} zu verdeutlichen.

Zur Beantwortung der Frage, ob die Dauer der vorgestellten Trajektorien des gelösten Dipeptids ausreichend ist, um eine ausreichende Abtastung seines Konformationsraumes zu erreichen, zeigt Abbildung 6.3 die Zeitentwicklungen der Dihedralwinkel ϕ und ψ während der beiden Simulationsläufe (300 K und 500 K), bei denen $\lambda_{14} = 1,0$ gewählt wurde. Die dargestellte Trajektorie bei 500 K war schon zuvor (Kapitel 5) als Beispiel für eine Konformationsdynamik herangezogen worden. Vergleicht man die in Abbildung 6.3 vorgestellten Zeitreihen mit den Geometrien der Vakuum-Minimalstrukturen aus Tabelle 6.1, so erkennt man, dass erst bei 500 K einige wenige Übergänge von den Konformeren mit $\phi < 0^\circ$ zu dem Konformer $C7_{ax}$ ($\phi \approx 70^\circ$) auftreten, nicht aber bei 300 K. Aus diesem Grund wurden die Trajektorien bei 500 K verwendet, um die Konfigurationsraumdichte des Alanin-Dipeptids zu bestimmen. Offensichtlich kann aber auch in diesem Fall nur von einer relativ schlechten Approximation der invarianten Dichte des Systems und auch nicht von einer ergodischen Dynamik ausgegangen werden, wie sie in Kapitel 5 angenommen wurde.

Abbildung 6.4 zeigt zusammenfassend die zweidimensionalen Landschaften sowohl der potentiellen Energien $U_{0,4}$ und $U_{1,0}$ des Alanin-Dipeptids im Vakuum, als auch der freien Energien $F_{0,4}$ und $F_{1,0}$ in wässriger Lösung. Beginnt man die Betrachtung mit Teilbild a, so ist festzustellen (Vergleich hier nicht gezeigt), dass $U_{1,0}$ sehr gut mit der Energielandschaft übereinstimmt, die im Rahmen der CHARMM-Kraftfeldoptimierung von MacKerell et al. [81] veröffentlicht wurde. In Teilbild a sind auch die drei in Abbildung 6.2 vorgestellten und in Tabelle 6.1 quantitativ charakterisierten Minimalstrukturen in Form lokaler Minima der potentiellen Energie zu erkennen. Vergleicht man die Höhen der Barrieren zwischen diesen Minima, so sieht man, dass zwischen den geometrisch benachbarten Strukturen $C7_{eq}$ und $C5$ nur 1-2 kcal/mol zu überwinden sind, sofern der Übergang über den Sattelpunkt S_2 erfolgt. Der Übergang über S_1 ist dagegen mit einer Barriere von 4-5 kcal/mol verbunden. Zwar ist auch der geometrisch weiter entfernte Zustand $C7_{ax}$ über zwei Sattelpunkte zu erreichen, die mit ihnen verbundenen Barrieren sind mit etwa 8 kcal/mol aber noch einmal deutlich höher.

Behält man nun den Literaturwert $\lambda_{14} = 1,0$ für den Skalierungsfaktor bei und geht von der potentiellen Energie $U_{1,0}$ (Teilbild a) zur freien Energie $F_{1,0}$ (Teilbild c) über, so wird zuerst die bei der freien Simulation beschränkte Abtastung der Energielandschaft deutlich. Auf Grund dieser Beschränkung kann nur derjenige Teil der Energielandschaft mittels einer freien Simulation untersucht werden, der höchstens ein kleines (drei- bis vierfaches) Vielfaches von $k_B T$ über dem absoluten Minimum von $F_{1,0}$ liegt ($k_B \cdot 500 \text{ K} \approx 1 \text{ kcal/mol}$). Eine weitere Begrenzung der hier verwendeten einfachen Methode zur Bestimmung der Landschaft der freien Energie hatten schon die in Abbildung 6.3 gezeigten Zeitreihen aufgezeigt: dadurch, dass bei den MD-Simulationen bei weitem nicht genug Übergänge zum Konformer $C7_{ax}$ beobachtet wurden, ist eine Angabe der Energie dieses Zustands relativ zur Energie der beiden anderen Konformere mit einem sehr großen Fehler behaftet. Gleiches gilt auch für die Form und Höhe der Barriere am Sattelpunkt S_3 und besonders auch für den überhaupt nicht aufgelösten Bereich der Energiefläche in der Nähe des (hier fiktiv eingezeichneten) Sattelpunkts S_4 .

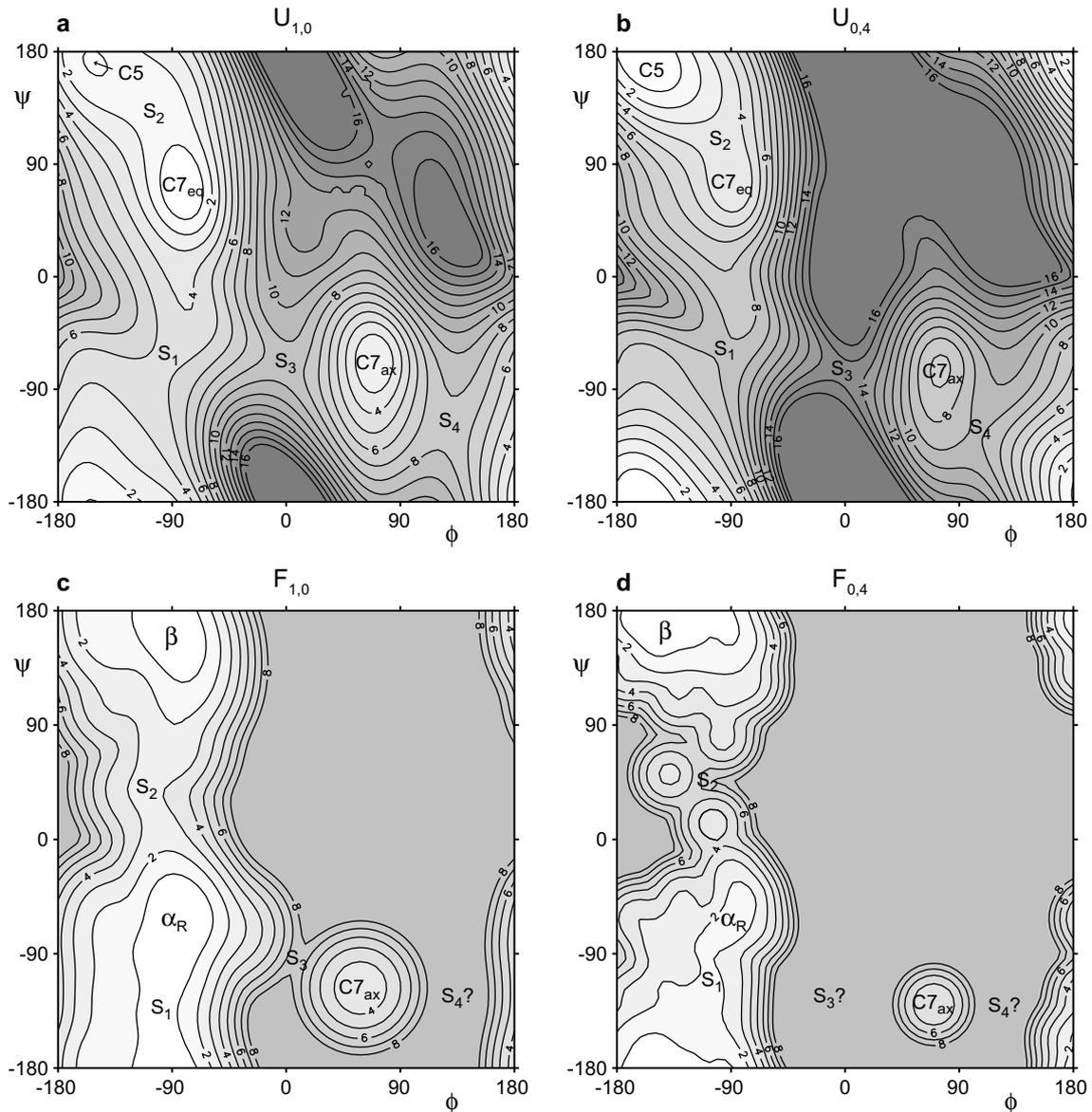


Abbildung 6.4: Die Energieflächen des Alanin-Dipeptids. **a** Die potentielle Energie $U_{1,0}(\phi, \psi)$ des Dipeptids im Vakuum, berechnet unter der Annahme von $\lambda_{14} = 1,0$, dem Literaturwert des Skalierungsfaktors λ_{14} der elektrostatischen 1-4-Wechselwirkung. Wie auch in den anderen Teilbildern beträgt der Abstand der Konturlinien jeweils 1 kcal/mol. Die Minima von $U(\phi, \psi)$ werden durch die in Abbildung 6.2 eingeführten Bezeichnungen gekennzeichnet. **b** Analog zu Teilbild a wird die potentielle Energie $U_{0,4}(\phi, \psi)$ dargestellt, hier jedoch unter der Wahl von $\lambda_{14} = 0,4$. **c** und **d** stellen die freien Energien $F(\phi, \psi)$ des Alanin-Dipeptids in wässriger Lösung dar, die nach Gleichung (5.2) aus Trajektorien bei 500 K bestimmt wurden. Auch hier werden die Ergebnisse für die beiden Werte von λ_{14} miteinander verglichen. Die Bezeichnungen der Minima der freien Energie finden ihre Erklärung im Text. Zusätzlich sind in allen Teilbildern die Sattelpunkte (S_1 bis S_4) zwischen den stabilen Konformationen hervorgehoben.

Eine ausreichende Statistik und damit eine gut definierte Energielandschaft kann jedoch für denjenigen Teil der Abbildung 6.4 c angenommen werden, in dem $\phi < -40^\circ$ gilt. Dort beobachtet man bei $F_{1,0}$ im Vergleich zu $U_{1,0}$ eine durch den Einfluss des Lösungsmittels und evtl. intramolekularer entropischer Beiträge verursachte Verschmelzung der zwei vorher unterscheidbaren Strukturen C5 und $C7_{eq}$ zu einer einzigen Konformation. Nach der weiter oben angegebenen Interpretation dieser beiden Strukturen als β -faltblattartig bzw. einem γ -Turn entsprechend und im Vergleich mit dem Ramachandran-Plot aus Abbildung 1.4 liegt es nahe, die entstehende Konformation mit einem β zu bezeichnen. Auffällig ist das Auftreten eines Minimums in $F_{1,0}$, das keine Entsprechung in $U_{1,0}$ findet. Seine Geometrie entspricht der einer rechtshändigen α -helikalen Peptidstruktur, weshalb diese Konformation in Abbildung 6.4 c die Bezeichnung „ α_R “ trägt. Die β - und die α_R -Konformation sind energetisch gleichwertig und bei $F_{1,0}$ über zwei Sattelpunkte miteinander verbunden, wobei bei einem Übergang über S_2 eine Barriere von 2-3 kcal/mol zu überwinden ist, während der Sattelpunkt S_1 nicht einmal 1 kcal/mol über den beiden Minima liegt. Die hier beobachtete Höhe der Barriere bei S_1 ist der einzige auffällige Unterschied zu den von anderen Gruppen präsentierten Landschaften der freien Energie des Alanin-Dipeptids [286, 287, 289], in denen für diese Barriere größere Werte im Bereich von 4 kcal/mol gefunden wurden (Vergleich nicht gezeigt).

Die Auswirkung des Wechsels des Skalierungsfaktors λ_{14} vom Literaturwert 1,0 auf den Wert von 0,4 auf die Minima in U wurde schon zuvor anhand der Tabelle 6.1 diskutiert. Ein Vergleich der Teilbilder a und b der Abbildung 6.4 verdeutlicht die dabei gemachten Beobachtungen noch einmal: Der Wechsel von λ_{14} führt zwar zu keiner deutlichen Verschiebung der Lage der Minima, aber zu deutlichen Veränderungen ihrer Energien, insbesondere ihrer energetischen Reihenfolge. In Teilbild b und damit bei $U_{0,4}$ ist nicht mehr $C7_{eq}$ der energetisch niedrigste Zustand, sondern C5. Dies hat auch zur Folge, dass der Zustand $C7_{eq}$ bei $U_{0,4}$ kaum noch als lokales Minimum in Erscheinung tritt und der Sattelpunkt S_2 mit einer sehr kleinen Barriere verbunden ist. Auf der anderen Seite führt der Wechsel von λ_{14} dazu, dass die anderen Sattelpunkte angehoben werden und dies zum Teil sehr deutlich: S_1 um etwa 4 kcal/mol auf 8-9 kcal/mol, S_3 um etwa 6 kcal/mol auf etwa 15 kcal/mol und S_4 immerhin um etwa 2 kcal/mol auf etwa 9 kcal/mol. Insgesamt ist zu beobachten, dass $U_{0,4}$ deutlich „schroffer“ erscheint als $U_{1,0}$, da der Anstieg der Energiefläche hin zu den verbotenen Bereichen steiler und diese verbotenen Bereiche selbst ausgedehnter werden.

Beim Vergleich von $F_{0,4}$ (Teilbild d) mit $F_{1,0}$ (Teilbild c) werden ähnliche Veränderungen sichtbar. So ist auch hier eine zunehmende Rauheit der Energielandschaft zu beobachten. Während die Lage der Konformationen α_R und $C7_{ax}$ beim Wechsel von λ_{14} unverändert bleibt, tritt bei der β -Konformation bei $F_{0,4}$ eine deutliche Zunahme des Winkels ϕ auf. Auch liegt die Konformation nicht wie zuvor hauptsächlich in einem Bereich von $\psi < 180^\circ$, sondern weist eine fast gleichmässige Verteilung von ψ um den Wert von $180^\circ / -180^\circ$ auf. Diese Verschiebung der β -Konformation liegt offensichtlich in der zuvor diskutierten Absenkung der potentiellen Energie der gestreckten C5-Struktur begründet, welche nun einen stärkeren Einfluß auf die β -Konformation ausübt. Weiterhin

kann an Teilbild d beobachtet werden, dass, während die beiden Konformere β und α_R ungefähr gleiche Energien $F_{1,0}$ aufweisen, α_R bei $F_{0,4}$ um etwa 1-2 kcal/mol über der β -Konformation liegt. Dagegen kann hier über die freie Energie von $C7_{ax}$ relativ zu den beiden anderen Konformationen keine Aussage getroffen werden, da die entsprechenden Bereiche der Energielandschaft nicht verbunden sind. Wegen der unzureichenden Statistik können bei $F_{0,4}$ auch nicht die Formen und Höhen der Sattelpunkte S_3 und S_4 bestimmt werden. Dies gilt prinzipiell auch für S_2 , in dessen Umgebung die berechnete freie Energie $F_{0,4}$ eindeutig artifizielle Strukturen aufweist. Die Höhe der Barriere am Sattelpunkt S_1 beträgt weiterhin weniger als 1 kcal/mol.

Zusammenfassend zeichnet Abbildung 6.4 ein klares Bild der Veränderungen an den Landschaften der potentiellen und der freien Energie des Alanin-Dipeptids, die durch einen Wechsel des Skalierungsfaktors λ_{14} vom Literaturwert 1,0 auf den alternativ diskutierten Wert von 0,4 ausgelöst werden. Im Vergleich mit den in der Literatur veröffentlichten Energieflächen des Dipeptids [81, 286, 287, 289] wird deutlich, dass der Literaturwert ($\lambda_{14} = 1,0$) zu einer guten Reproduktion dieser Ergebnisse führt. Die alternative Wahl von $\lambda_{14} = 0,4$ führt zu einer rauerer Energielandschaft und einer Überstabilisierung der β -Faltblatt-Konformation im Vergleich zur α_R -Helix. Eine direkte Übertragung dieser Beobachtungen auf längere Peptide basiert auf der Annahme, dass der Wechsel von λ_{14} primär die Wechselwirkungen innerhalb des betrachteten Alanin-Residuums betrifft. Eine Betrachtung der in Abbildung 6.2 gezeigten Strukturen und der sich darin räumlich nahe kommenden und durch drei Bindungen verbundene Atome legt die Vermutung nahe, dass diese Annahme gültig ist und somit näherungsweise von einer Additivität des λ_{14} -Effekts bezüglich der Residuen eines Peptids ausgegangen werden kann. Somit ist auch in längeren Peptiden für $\lambda_{14} = 0,4$ eine Überstabilisierung der β -Faltblatt-Konformation gegenüber einer α -Helix um etwa 1-2 kcal/mol pro Residuum zu erwarten. Bedenkt man, dass beim Übergang eines Polyalaninpeptids von einer ungeordneten (Random Coil) Konformation in eine α -Helix eine Differenz der freien Energie von etwa dem gleichen Betrag pro Residuum beobachtet wurde [293], so ist zu erwarten, dass die Wahl $\lambda_{14} = 0,4$ die Stabilität einer α -Helix in MD-Simulationen entscheidend vermindert.

7 Die Modellpeptide im Gleichgewicht

Mit der Verifikation des Kraftfeldes im vorangehenden Kapitel sind nun alle notwendigen Vorarbeiten für die MD-Simulationen der Modellpeptide geleistet worden. Die Aufgabe besteht nun darin, die Konformationsräume dieser Peptide zu bestimmen und auf geeignete Weise zu analysieren. Dazu werden Langzeittrajektorien der verschiedenen Modellpeptide in Lösung vorgestellt und mit den in Kapitel 5 entwickelten Verfahren analysiert. Ziel dieser Analyse ist es, die Landschaften der freien Energie der Modellsysteme in Abhängigkeit vom Isomerisierungszustand des Farbstoffs zu bestimmen. Diese Energielandschaften bilden anschließend die Grundlage für eine detaillierte Untersuchung der Konformationsräume der Modellpeptide und der dort ablaufenden Konformationsdynamik.

Die Betrachtung der Resultate der MD-Simulationen wird anschließend um einen Vergleich mit den Ergebnissen der NMR-Strukturuntersuchungen [37, 38] ergänzt. Das einfachste Vorgehen bei einem solchen Vergleich ist es, die Geometrien der bei den Simulationen beobachteten Konformationen der Peptide direkt den aus dem NMR-Refinement hervorgegangenen Strukturen gegenüberzustellen und daraus Gemeinsamkeiten und Unterschiede zu erkennen. Dagegen kann eingewandt werden, dass das beim NMR-Refinement eingesetzte Kraftfeld und das verwendete Verfahren zur Strukturoptimierung den Vergleich beeinflussen können. Daher wurde in der Literatur vorgeschlagen, die Ergebnisse einer MD-Simulation direkt mit den NOE-Abstandsbedingungen zwischen den Wasserstoffatomen oder den ^3J -Kopplungskonstanten der Dihedralwinkel des Peptidrückgrats in Beziehung zu setzen. Beispielsweise wurden in der Gruppe van Gunsteren schon mehrfach Simulationen von Peptiden auf diese Weise mit NMR-Messungen verglichen [87, 294]. Die gleichen Autoren haben auch untersucht, mit welchen Problemen die NMR-Strukturaufklärung im Fall kleiner in Lösung auf NMR-Zeitskalen schnell rotierender Peptide verbunden ist [295] und welchen Effekt das Vorliegen einer Mischung verschiedener Konformationen auf die gemessenen NMR-Spektren hat [179]. Andere erwähnenswerte Beispiele für Vergleiche zwischen NMR-Restraints und MD-Simulationen sind die von Colombo et al. [296] an dem künstlichen Peptid „Betanova“ durchgeführten Untersuchungen und die ausführliche Analyse einer MD-Trajektorie des Lysozyms der Hühner [297–299].

In den letzten Jahren wurde auch mit Erfolg versucht, nicht die Abstandsbedingungen, sondern die „rohen“ Daten der NMR-Messungen, etwa die NOESY-Spektren, zur Beurteilung von MD-Simulationen heranzuziehen [300–302]. Hierzu ist es notwendig, ausgehend von den tabellierten chemischen Verschiebungen der einzelnen Atome, die Spektren aus der MD-Trajektorie zu berechnen. Diese „back-calculation“ ist in dem von mir verwendeten MD-Programm nicht vorgesehen. Ihre Implementierung hätte einen außerordentlichen

Programmieraufwand erfordert und konnte deshalb hier nicht eingesetzt werden.

Zwei weitere Arbeiten sind für die in diesem Abschnitt untersuchte Frage nach der konformellen Flexibilität von Peptiden und deren Bestimmung durch MD-Simulationen und NMR-Messungen von Interesse: Lei und Smith [303] konnten zeigen, dass in Simulationen eine weit größere Flexibilität einer ungeordneten Loop-Region eines Proteins zu beobachten ist als sie von den NMR-Strukturen gezeigt wird. Ferner beobachteten Stote et al. [304] in der MD-Trajektorie eines kurzen Peptids stärker entfaltete Konformationen als sie durch NMR-Messungen vorausgesagt wurden. Schließlich ist noch die Arbeit von Abseher et al. [240] zu erwähnen, in welcher die Autoren eine Hauptkomponentenanalyse verwendet haben, um gleichermaßen in MD-Trajektorien und NMR-Strukturensembles die „weichen Moden“, d.h. kollektive Koordinaten großer Beweglichkeit, zu bestimmen. Hier zeigte sich, dass diese Moden in der Theorie und im Experiment einen vergleichbar ausgedehnten Raum aufspannen.

Das vorliegende Kapitel ist wie folgt gegliedert: zuerst werden die Details der durchgeführten Simulationen vorgestellt. Insbesondere wird dabei auf die behutsame Equilibrierung der Simulationssysteme eingegangen. Mit dem geschilderten Vorgehen kann die Stabilität der NMR-Strukturen bezüglich des verwendeten MD-Kraftfeldes beurteilt werden. Hieran schließt sich in Abschnitt 7.2 die Betrachtung der aus den MD-Trajektorien berechneten Landschaften der freien Energie der Modellpeptide an. Die Landschaften werden sowohl in Form von Ramachandran-Plots als auch in der Projektion auf ihre Hauptkomponenten dargestellt. Anhand dieser Darstellungen kann ein erster Einblick in die Konformationsräume der drei Peptide und ihrer beiden Isomere gewonnen werden. In Abschnitt 7.3 folgt dann die quantitative Analyse der Konformations-Subzustände, bei der die in Abschnitt 5.3 vorgestellten Methoden der unüberwachten Klassifikation zum Einsatz kommen. Die unterschiedlichen Ensembles von Konformationen werden anschließend visualisiert und den veröffentlichten NMR-Strukturen gegenübergestellt, bevor sie in Abschnitt 7.4 direkt mit den NMR-Restraints beurteilt werden. Schließlich werden in Abschnitt 7.5 die Ergebnisse dieses Kapitels diskutiert.

7.1 Details der Simulationen

Die Konformationsräume aller drei in Abbildung 1.13 gezeigten Mitglieder der Familie von zyklischen Peptiden wurden untersucht: des cAPBs, des cAMPBs mit Methylenspacer und des bcAMPBs, bei welchem zusätzlich zu dem Spacer die Disulfidbrücke geschlossen ist. Die in Kapitel 4 bestimmte Energiebarriere von mehr als 40 kcal/mol für eine Drehung um die N=N-Bindung in den Azobenzolfarbstoffen verhindert innerhalb der erreichbaren Simulationszeiten einen thermisch angeregten Übergang zwischen den *cis*- und den *trans*-Isomeren der Modellpeptide, die deshalb jeweils in getrennten Simulationen untersucht werden. Statt von einer einzigen Energielandschaft wird deshalb in Zukunft von getrennten *cis*- bzw. *trans*-Landschaften gesprochen.

Um die Konformationsräume der Modellpeptide abzutasten, wurden lange MD-Simulationen durchgeführt. Ähnlich wie im Fall des Alanin-Dipeptids zeigten Voruntersuchungen in Form einer Reihe von Simulationen bei unterschiedlichen Temperaturen, dass erst ab 450 K einige wenige Konformationsübergänge pro Nanosekunde Simulationszeit auftreten. Um im Rahmen der hier vorgestellten Trajektorien von jeweils 10 ns Dauer wenigstens eine grobe Abtastung der Konformationsräume zu erreichen, wurde $T = 500$ K gewählt.

Als Ausgangsstrukturen für die Simulationen dienten jeweils die ersten der veröffentlichten NMR-Strukturen [37, 38], die nach Auffassung der Autoren die entsprechenden NMR-Restraints am besten erfüllen. Jede dieser sechs Startkonfigurationen wurde in einen mit etwa 960 DMSO-Molekülen gefüllten Rhombendodekaeder eingesetzt. Dies entspricht einem Innenradius des Simulationssystems von etwa 27 Å. In Hinblick auf den hohen Rechenaufwand, der mit der Berechnung einer Langzeittrajektorie verbunden ist, und da angenommen werden kann, dass bei einer Gleichgewichtssimulation die Wärmeleitfähigkeit des Lösungsmittelmodells eine eher untergeordnete Rolle spielt, wurde die als MP bezeichnete Parametrisierung für die DMSO-Moleküle gewählt (siehe Kapitel 3). Das Kraftfeld der Modellpeptide wurde schon ausführlich in Kapitel 4 besprochen. Hier wurde stets die Variante B der Dihedralstärken der Azobenzolfarbstoffe und der CHARMM22-Standardwert von 1,0 für den Skalierungsfaktor λ_{14} gewählt.

Bei dem Einsetzen der Peptide in die DMSO-Dodekaeder wurden alle DMSO-Moleküle die von einem der Peptidatome weniger als 2 Å Entfernung aufwiesen, ausgeschnitten. Es galten periodische Randbedingungen. Die langreichweitigen elektrostatischen Wechselwirkungen wurden mit dem FAMUSAMM-Algorithmus (siehe Abschnitt 2.2) berechnet, wobei wegen der relativ kleinen Ausmaße des Simulationssystems nur zwei Entfernungsklassen definiert wurden: die explizite Berechnung der Coulomb-Summe galt für alle Ladungen mit einem Abstand von weniger als etwa 13 Å, wogegen die elektrostatischen Beiträge aller weiter entfernten strukturellen Einheiten durch ihre Dipol- und Quadrupolmomente approximiert wurden. Der Überlapp der beiden Entfernungsklassen betrug 1,25 Å. Außerhalb des durch die Minimum-Image Konvention vorgegebenen Abschneideradius von etwas weniger als 27 Å wurde das sich anschließende dielektrische Kontinuum durch ein Reaktionsfeld [100] mit einer DK von 45,8 berücksichtigt. Schließlich fand wieder die in Anhang B vorgestellte Korrektur der langreichweitigen VDW-Wechselwirkungen Anwendung, wobei der effektive Gyrationradius 1,2 Å betrug. In allen vorgestellten Simulationen der Modellpeptide wurde ein Integrationszeitschritt von 1 fs Länge gewählt. Die Geometrie der starren DMSO-Modelle und auch die Längen der Bindungen aller Wasserstoffatome des Peptids wurden mit Hilfe des MSHAKE Algorithmus [159] konstant gehalten.

Equilibrierung der Simulationssysteme

Durch das Einsetzen der NMR-Strukturen in das Lösungsmittel und das darauf folgende Entfernen überlappender Atome entstanden stark verspannte Konfigurationen des Simulationssystems. Die so genannte *Equilibrierung* eines solchen Systems, die vor der eigentli-

chen MD-Simulation durchzuführen ist, hat zum Ziel, aus einer solchen Startkonfiguration ein Simulationssystem zu erzeugen, in dem die Konfiguration der Lösungsmittelmoleküle an das gelöste Molekül angepasst ist, die Geschwindigkeitsverteilung der Atome einer Maxwell-Verteilung bei der gewählten Temperatur entspricht und schließlich die Struktur des simulierten Biomoleküls in Übereinstimmung mit dem MD-Kraftfeld steht [74]. Hier und auch bei den später folgenden Simulationen wurde eine sehr vorsichtige Equilibrierung vorgenommen, die abrupte Veränderungen in der Beschreibung der Modellpeptide und somit Kraftstöße auf Teile der Moleküle vermeidet. Sie bestand jeweils aus den folgenden sechs Schritten bzw. Simulationsläufen:

1. Zuerst wurde das gelöste Peptid vollständig immobilisiert und die Energie des Lösungsmittels 200 Zeitschritte lang minimiert, indem die Geschwindigkeiten der Atome in jedem Schritt mit einem Faktor von 0,95 skaliert wurde. Auf diese Weise wurden diejenigen Energiebeiträge aus dem System entfernt, die an einzelnen Atomen durch einen evtl. noch vorhandenen Überlapp mit den VDW-Kugeln anderer Atome auftreten können.
2. Bei weiterhin festgehaltenem Peptid wurde die freie Dynamik des Lösungsmittels im NVE-Ensemble 5.000 Schritte lang verfolgt. Die Bewegung der Lösungsmittelmoleküle entstand dabei durch die Relaxation noch vorhandener Verspannungen. Während dieser Simulation heizten sich die Systeme, auch durch die immer auftretenden Rauschkkräfte, auf eine Temperatur von etwa 100 K auf. Auch konnten so die DMSO-Moleküle mögliche Leerräume in der Nähe des gelösten Peptids ausfüllen.
3. In den darauf folgenden 15 ps der Simulation wurde der Thermostat [101] für das Lösungsmittel eingeschaltet und dieses im NVT-Ensemble auf eine Temperatur von 298 K erwärmt. Die Peptide blieben weiterhin immobilisiert, die Zeitskala der Anbindung an das Wärmebad betrug $\tau_T = 0,5$ ps.
4. Bei einem immer noch festgehaltenen Peptid wurde nun vom NVT- in das NPT-Ensemble übergegangen und der Druck durch den Barostaten [101] auf 1 atm festgelegt. Die Zeitkonstante der Ankopplung betrug $\tau_P = 2,5$ ps und die experimentelle isotherme Kompressibilität des DMSOs ($\kappa_T = 5,32 \times 10^{-5}$ 1/atm) wurde angenommen. Dieser Schritt der Equilibrierung dauerte weitere 80 ps.
5. Das Peptid wurde nun mit Hilfe des vollständig equilibrierten Lösungsmittels vorsichtig „aufgetaut“. Dazu wurden in den nächsten 50 ps der Simulation die Atome der Modellpeptide freigegeben, allerdings nicht vollständig: für die Positionen der C_α -Atome und der Stickstoffatome der Azo-Gruppe galten harmonische Zwangsbedingungen mit einer Stärke von $0,4 \text{ kcal}/(\text{mol} \cdot \text{\AA}^2)$. Auf diese Weise wurde die Konfiguration des Peptidrückgrats weiterhin festgehalten, während die Seitengruppen relaxieren konnten. Von diesem Equilibrierungsschritt an wurden der Thermo- und der Barostat nur halb so stark an das System angebunden ($\tau_T = 1$ ps und $\tau_P = 5$ ps).

6. Jede Equilibrierung wurde durch eine Simulation von 100 ps Dauer bei einer Temperatur von 298 K abgeschlossen, bei der die Dynamik der Modellpeptide frei von allen Einschränkungen verfolgt wurde.

Stabilität der NMR-Strukturen während der Equilibrierung

Ein Zweck der vorgestellten aufwändigen Equilibrierung ist es, eine Aussage über die Kompatibilität der veröffentlichten NMR-Strukturen mit dem hier eingesetzten Kraftfeld treffen zu können. Dazu ist zuerst anzumerken, dass bei der Optimierung der NMR-Strukturen das Programmpaket INSIGHT II mit dem dort implementierten Kraftfeld CVFF Verwendung fand. Dabei wurde der Einfluss des DMSOs als Lösungsmittel implizit durch die Annahme einer DK von 46,7 abgeschätzt [37]. Unterschiede zwischen dem im NMR-Refinement verwendeten und dem hier eingesetzten Kraftfeld bestehen daher unter anderem in unterschiedlichen Parametern der Bonded- und Nonbonded-Wechselwirkungen (siehe Abschnitt 2.2) und in der ganz anderen Modellierung des Lösungsmittels. Die Stabilität der NMR-Strukturen während der Equilibrierung ermöglicht daher eine Aussage über das Ausmaß dieser Unterschiede.

Abbildung 7.1 zeigt am Beispiel der Equilibrierung einer der NMR-Strukturen des *cis*-cAPBs die während der Schritte 5 und 6 im Peptidrückgrat auftretenden geometrischen Änderungen. Wie zu Beginn des Kapitels 5 erörtert wurde, dienen zur Beschreibung der Geometrie des Peptidrückgrats die 16 Dihedralwinkel ϕ und ψ der acht Aminosäuren. Die naheliegenste und in Abbildung 7.1 verwendete Darstellung dieser Daten ist die in Form von acht Ramachandran-Plots.

Vergleicht man zuerst die in Abbildung 7.1 gezeigte NMR-Struktur (Lage der roten Kreuze) mit dem schon mehrfach diskutierten abstrakten Ramachandran-Plot (Abbildung 1.4) bzw. den für das Alanin-Dipeptid berechneten Energieflächen (Abbildung 6.4), so ist zu erkennen, dass durch die NMR-Strukturoptimierung teilweise Winkelkombinationen für das Peptid festgelegt wurden, die (nicht nur) im CHARMM22-Kraftfeld als energetisch ungünstig angesehen werden. In dem gezeigten Beispiel liegen beispielsweise die Geometrien der Residuen Ala 3 und Cys 5 in einem Bereich des Ramachandran-Plots, welcher der in Peptiden selten beobachteten α_L -Sekundärstruktur entspricht. Im Verlauf der gezeigten Equilibrierung ist auch sehr schnell (nach etwa 35 ps, Zeitachse hier nicht gezeigt) eine Relaxation des Peptids hin zu einer Konfiguration geringerer innerer Energie zu beobachten. Diese Relaxation bestand bei vielen der untersuchten NMR-Strukturen in einem oder mehreren, sehr schnell ablaufenden so genannten *Peptid-Flips*, bei denen sich eines der starren Peptidplättchen (vgl. Abbildung 1.2) um etwa 180° dreht. Ein solcher Flip führt nicht unmittelbar zu großen Änderungen an der räumlichen Struktur des Peptidrückgrats, jedoch sind Auswirkungen des umgedrehten Dipols der Peptidgruppe zu erwarten. Ein Flip der Peptidgruppe zwischen einem Residuum i und dem darauf folgenden Residuum $i + 1$ bewirkt eine gleichzeitige Änderung der Winkel ψ_i und ϕ_{i+1} um jeweils etwa 180° . Somit ist in Abbildung 7.1 ein solcher Flip bei den Residuen 2 und 3 zu beobachten. Er wird durch einen blauen Pfeil hervorgehoben. Gleichzeitig ist ein ähnlicher Übergang

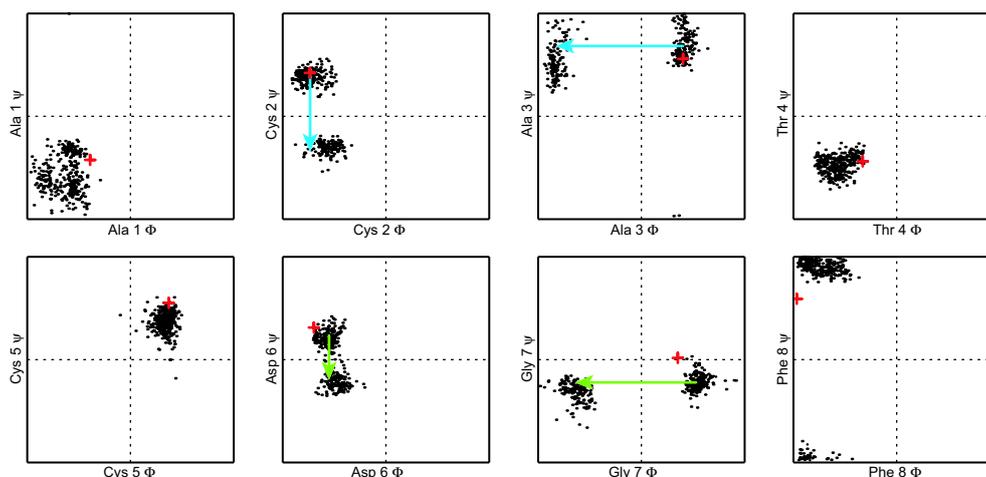


Abbildung 7.1: Relaxation einer NMR-Struktur während der MD-Simulation. Ramachandran-Plots der 16 Dihedralwinkel ϕ und ψ des Peptidrückgrats im *cis*-cAPB. Die Achsenbeschriftungen (-180° , 0° , 180°) wurden aus Gründen der Übersichtlichkeit weggelassen. Rote Kreuze kennzeichnen die Geometrie der NMR-Startstruktur (hier die fünfte der zehn veröffentlichten Strukturen). Schwarze Punkte stellen die während der Schritte 5 und 6 der Equilibrierung aufgezeichnete Trajektorie dar. Farbige Pfeile verdeutlichen den Konformationsübergang in Form zweier fast simultaner Peptid-Flips, der etwa bei $t = 35$ ps in Schritt 6, der freien Simulation, auftrat.

auch an den Residuen 6 und 7 zu beobachten (grüner Pfeil). Durch die hohe Flexibilität des Glyzins ist der zweite Flip im Dihedralraum jedoch nicht im gleichen Maße ausgeprägt. Schließlich sind in Abbildung 7.1 an fast allen Dihedralwinkeln, die nicht von einem der Flips betroffen sind, Verschiebungen der bei der Simulation beobachteten Winkelverteilung gegenüber der NMR-Struktur zu erkennen.

Zusammenfassend kann also festgestellt werden, dass viele der NMR-Strukturen bei Verwendung des CHARMM22-Kraftfeldes und einer expliziten Modellierung des Lösungsmittels keine Strukturen minimaler freier Energie darstellen. Vielmehr scheint das beim NMR-Strukturrefinement eingesetzte Verfahren des „Simulated Annealings“ in vielen Fällen Verspannungen im Peptidrückgrat erzeugt zu haben. Insbesondere ist darauf hinzuweisen, dass die in den NMR-Strukturen festgelegten Ausrichtungen der Peptidplättchen und damit die ihrer Dipole, von dem bei den Simulationen verwendeten Kraftfeld häufig als energetisch äußerst ungünstig beurteilt werden.

7.2 Die Energielandschaften der *cis*- und *trans*-Isomere

Eine Untersuchung der Landschaften der freien Energie der Modellpeptide ermöglicht es, Aussagen über die Ausdehnung ihrer Konformationsräume zu treffen. Von Interesse ist dabei ein Vergleich der Energielandschaften der unterschiedlichen Spezies der zyklischen Peptide (Abbildung 1.13). Hier soll unter anderem der Einfluß des Methylenspacers im cAMPB auf die Flexibilität des Peptidrückgrats untersucht werden. Gleiches gilt natürlich auch für die Auswirkungen der Disulfidbrücke im bcAMPB. Weiterhin ist es von größter Wichtigkeit, für jedes der drei Modellpeptide die Veränderung der Energielandschaft bei einem Wechsel des Isomerisierungszustandes des eingebauten Azobenzolfarbstoffs zu bestimmen. Die sich hier stellenden Fragen sind: Wie schränken die durch die Geometrie des Farbstoffs gebildeten Randbedingungen den Peptidteil der Modellsysteme ein? Welche Unterschiede sind dabei zwischen den *cis*- und *trans*-Isomeren des Farbstoffs zu beobachten? Und wie stark überlappen die Energielandschaften der beiden Isomere jeweils eines der Modellpeptide?

Wie schon am Anfang dieses Abschnitts erläutert wurde, dienten zur Bestimmung der Energielandschaften freie Simulationen von jeweils 10 ns Dauer, die bei 500 K durchgeführt wurden. Diese Erhöhung der Temperatur auf einen Wert, der leicht über dem experimentellen Siedepunkt von DMSO liegt, führte im NPT-Ensemble zu einer Abnahme der Dichte auf etwa $0,9 \text{ g/cm}^3$. Jedoch konnten bei einer visuellen Inspektion der Lösungskonfigurationen keine besonderen Dichteschwankungen oder entstehenden Hohlräume beobachtet werden. Während der MD-Simulationen wurden die Koordinaten der Modellpeptide alle 500 fs abgespeichert, so dass sich Trajektorien aus jeweils 20.000 Konfigurationen ergaben.

7.2.1 Ramachandran-Plots

Abbildung 7.2 zeigt die Energielandschaften der drei betrachteten Modellpeptide, wobei die Isomerisierungszustände jeweils zusammen in einem Diagramm dargestellt werden. Zur Beschreibung der Konfiguration des Peptidrückgrats wurden wiederum die 16 Dihedralwinkel verwendet und die Daten in Form von jeweils acht Ramachandran-Plots dargestellt. Die freie Energie wurde mit den in Kapitel 5 erläuterten Methoden aus den Gleichgewichtstrajektorien bestimmt. Im Gegensatz zu den im NVT-Ensemble durchgeführten Simulationen am Alanin-Dipeptid und der daraus berechneten Helmholtz'schen freien Energie F ist bei dem hier erzeugten NPT-Ensemble präziserweise von der Gibbs'schen freien Energie G zu sprechen. Da bei MD-Simulationen der PV -Term allerdings nur eine sehr untergeordnete Rolle spielt [309], werden im Folgenden beide Energien vereinfacht als die „freie Energie“ bezeichnet.

Aufgrund der endlichen Länge der aufgezeichneten Trajektorien kann natürlich nicht von

7 Die Modellpeptide im Gleichgewicht

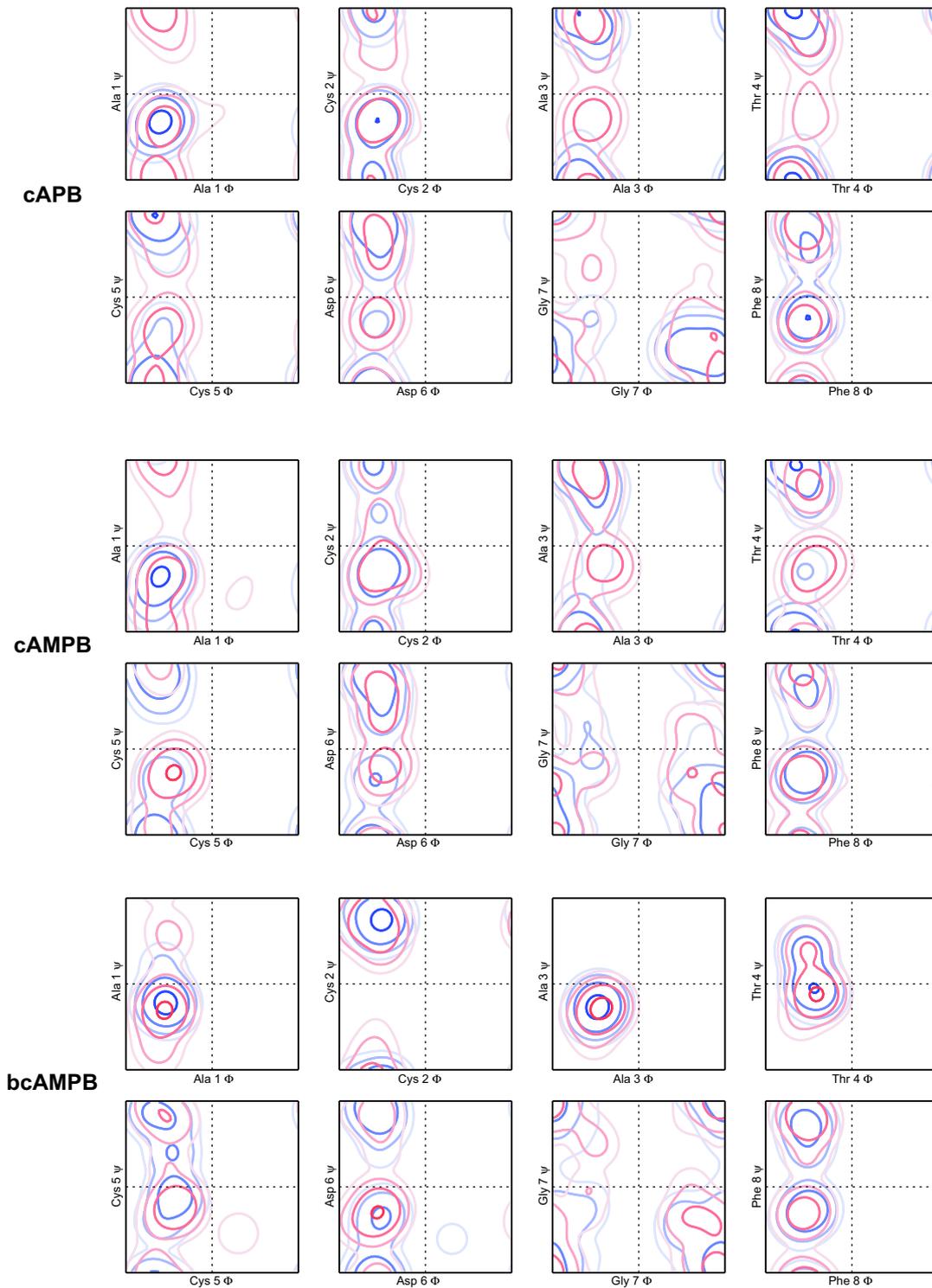


Abbildung 7.2: Ramachandran-Plots der freien Energie der Modellpeptide. Für jedes der drei zyklischen Peptide wird die aus den Gleichgewichtstrajektorien berechnete Schätzung der freien Energie (5.2) als Funktion der acht Dihedralwinkelpaare dargestellt. Rote Konturlinien repräsentieren die Energielandschaft der *cis*-Isomere, blaue Konturlinien die der *trans*-Isomere. Der Abstand der Konturlinien beträgt jeweils 1 kcal/mol.

einer wirklich ergodischen Dynamik des Systems ausgegangen werden, so dass die beobachteten Konfigurationsraumdichten wahrscheinlich mit hohen Fehlern behaftet sind. Dennoch wurden die in Kapitel 5 vorgestellten Methoden unverändert angewendet, es muss aber jeweils die Signifikanz der gewonnenen Resultate diskutiert werden. Die Schätzung der Konfigurationsraumdichten aus den Trajektorien geschah jeweils durch eine Mischung von 100 univariaten Normalverteilungen, wobei der UNIVAR-Algorithmus in seiner Erweiterung auf zirkuläre Variablen Einsatz fand (siehe Abschnitt 5.4). Für die sechs betrachteten Datensätze ergaben sich sehr ähnliche optimale Gaußbreiten σ_{opt} , so dass zur besseren Vergleichbarkeit der Schätzungen in allen Fällen $\sigma = 0,43$ gewählt wurde.

Beginnt man die Diskussion der Abbildung 7.2 mit der Energielandschaft bzw. Konfigurationsraumdichte des *cis*-Isomers (rot) des monozyklischen APB-Peptids (cAPB), so zeigt sich zuerst, dass alle Residuen bei der MD-Simulation nur solche Winkelkombinationen aufweisen, die den erlaubten Bereichen der α -helikalen oder β -Faltblatt-ähnlichen Sekundärstrukturen entsprechen. Eine Ausnahme bildet natürlich immer das Glyzin mit der Residuenummer 7 (Gly 7), dessen Konfiguration relativ unspezifisch ist und deshalb hier nicht diskutiert werden soll. Die beiden an jedem Residuum auftretenden Minima der freien Energie sind beim *cis*-cAPB ungefähr gleich tief, nur beim Thr 4 tritt eine leichte Bevorzugung der β -Faltblatt-Konfiguration auf. Uneinheitlich ist die Aussage über die Höhe der Sattelpunkte zwischen den beiden Minima. Während bei den Residuen 1, 2, 3 und 5 die Barriere bei negativen Werten von ψ geringer ist (vgl. auch Abbildung 6.4), scheint im Fall der Residuen 6 und 8 ein Übergang über denjenigen Sattelpunkt günstiger zu sein, der bei positiven Werten von ψ zu finden ist.

Ein anderes Bild bietet sich bei Betrachtung der blau eingezeichneten Energielandschaft des *trans*-cAPBs. Hier ist der Konfigurationsraum der Mehrzahl der Residuen merklich durch den gestreckten Azobenzolfarbstoff eingeschränkt: deutlich ist dies am Ala 1 zu beobachten, bei dem überhaupt nur die α -helikale Konfiguration auftritt. Auch beim Cys 2 ist jetzt diese Konfiguration leicht energetisch gegenüber der weiterhin beobachtbaren β -Strang-Konfiguration bevorzugt. Dagegen erzwingt der *trans*-Chromophor bei den Residuen 3 und 4 die β -Strang-Struktur, die nun auch bei den Residuen 5 und 6 bevorzugt wird. Am Phe 8 ist dagegen keine nennenswerte Veränderung durch das Umschalten des Farbstoffs festzustellen. Im Vergleich zum *cis*-cAPB drängt sich somit für das *trans*-cAPB das Bild einer durch den Farbstoff konformell stark eingeeengte und gestreckte Konfiguration des Peptidrückgrats auf.

Die Gegenüberstellung der Energieflächen des cAPBs und des cAMPBs ermöglicht es nun, den Einfluss des zusätzlichen Methylenspacers im cAMPB abzuschätzen, der zwischen dem Azobenzolfarbstoff und dem Phenylalanin (Phe 8) eingebaut ist (siehe Abbildung 1.13). Vergleicht man in Abbildung 7.2 die Energielandschaften des cAPBs und des cAMPBs, so sind dort für die *cis*-Isomere keinerlei signifikante Unterschiede festzustellen. Dagegen scheint beim *trans*-cAMPB der Einfluss des Farbstoffs nicht so ausgeprägt zu sein wie beim *trans*-cAPB. Beispielsweise können im *trans*-cAMPB die Residuen 3 bis 5 weiterhin eine α -helikale Konfiguration annehmen.

Die Beobachtungen an den *trans*-Isomeren zeigen demnach, dass dort der Einbau des Spacers zu einer erhöhten Flexibilität des Modellpeptids bzw. zu schwächeren Ankopplung der durch den Farbstoff ausgeübten Zwangsbedingungen führt. Warum ist dies aber nicht im Vergleich der beiden *cis*-Isomere zu beobachten? Ein mögliche Erklärung hierfür ist, dass das Peptidrückgrat schon im *cis*-cAPB nur so wenig durch den Farbstoff beeinflusst wird, dass es dort problemlos seine „natürlichen“ Konformationen einnehmen kann. Eine weiteren Abschwächung der Anbindung des Farbstoffs durch den Spacer ist dann nicht mehr möglich und bleibt ohne Folge. Alternativ könnte man vermuten, dass der Einbau der Methylengruppe in das *cis*-Isomer prinzipiell keine Abschwächung der Anbindung des Farbstoffes zur Folge hat. Die letztere Hypothese ist aber unter Berücksichtigung der in Abbildung 1.12 gezeigten Geometrien der beiden Azobenzol-Isomere zu verwerfen. Dort wurde insbesondere der sich mit der Isomerisierung verändernde Öffnungswinkel des Farbstoffs als eine an das Peptid angelegte Zwangsbedingung hervorgehoben. Der Methylenspacer erlaubt jedoch eine frei rotierbare Anbindung des Peptidrückgrats an den Farbstoff und führt damit zwangsläufig zu einer Abschwächung der Einschränkung des Peptid-Konformationsraumes.

Einen weitaus stärkeren Einfluss auf die Energielandschaft des Peptids hat das Schließen der Schwefelbrücke zwischen den beiden Cystein-Residuen, d.h. der Übergang vom monozyklischen cAMPB zum bityklischen bcAMPB. Abbildung 7.2 zeigt, dass die Schwefelbrücke die Konfiguration der „eingeklammerten“ Residuen Cys 2, Ala 3 und Thr 4 extrem einschränkt, so dass dort in der Energielandschaft nur ein einzelnes und gut definiertes Minimum existiert. Diese Beobachtung ist unabhängig vom Isomerisierungszustand des Farbstoffs. Während an den Residuen 5, 6 und 8 des bcAMPBs teilweise noch zwei Minima der freien Energie zu beobachten sind, also noch eine gewisse konformelle Flexibilität erhalten bleibt, sind auch dort die Energieflächen der beiden Isomere sehr ähnlich. Nur am Ala 1 weist das *cis*-bcAMPB zwei Minima auf, das *trans*-bcAMPB jedoch nur eines. Verglichen mit den cAMPB zeigt das bcAMPB somit eine viel eingeschränktere Energielandschaft, bei der nur noch ein geringer Effekt des Umschaltens des Farbstoffs zu beobachten ist.

Zusammenfassend zeigen die Ramachandran-Plots der Energielandschaften, dass sich beim Umschalten des Azobenzolfarbstoffs im cAPB der Konformationsraum des Peptidrückgrats deutlich verändert. Während die *cis*-Energielandschaft eine hohe Komplexität aufweist, scheint das Peptid in der *trans*-Form auf eine oder wenige Konformationen festgelegt zu sein. Durch den Einbau des Methylenspacers in das cAMPB wird der Einfluss des Farbstoffs auf das Peptidrückgrat merklich abgeschwächt. Schließlich führt die Disulfidbrücke im bcAMPB zu einer extremen Versteifung des Peptids zwischen den beiden Cysteinen. Die Energielandschaft lässt sich dort durch den Farbstoff nur noch wenig beeinflussen.

Bei der Diskussion der in den Ramachandran-Plots dargestellten Energielandschaften ist jedoch stets zu berücksichtigen, dass diese acht zweidimensionalen Projektionen der freien Energie keinen Einblick in die Korrelationen zwischen der Konfiguration eines der Re-

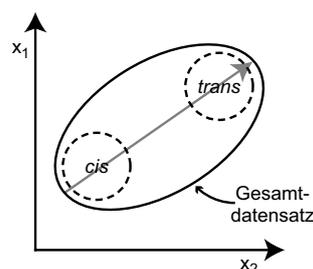


Abbildung 7.3: Die Motivation der Anwendung der Hauptkomponentenanalyse. Sind die Geometrien der *cis*- und der *trans*-Isomere ausreichend verschieden voneinander, so kann angenommen werden, dass die Hauptkomponenten des aus den beiden Trajektorien zusammengesetzten Gesamtdatensatzes die Richtung bezeichnen bzw. denjenigen Unterraum aufspannen, in dem eine gute Trennung der beiden Verteilungen beobachtet werden kann.

siduen mit derjenigen der anderen Aminosäuren erlauben. So ist nicht ersichtlich, welche Kombinationen solcher Konfigurationen überhaupt auftreten. Nimmt man am Beispiel des *cis*-cAPBs für die sieben nicht-Glyzin Residuen je zwei mögliche Konfigurationen an (α -helikal oder β -Strang-ähnlich), so lassen sich daraus $2^7 = 128$ mögliche Konfigurationen des Gesamtpeptids konstruieren. Die Fragen, welche dieser Kombinationen als Konformationen während der Simulation aufgetreten sind, wie diese Konformationen energetisch zueinander stehen und welche Übergänge vorgekommen sind, können in der bisher gezeigten Darstellung nicht beantwortet werden.

7.2.2 Projektion der Energieflächen auf ihre Hauptkomponenten

In Abschnitt 5.1 wurde die Hauptkomponentenanalyse (PCA) als Methode zur Auswertung der Korrelationen in hochdimensionalen Datensätzen vorgestellt und gezeigt, wie mit ihrer Hilfe niedrigdimensionale Darstellungen dieser Daten erzeugt werden können. Betrachtet man, wie oben, die Peptidkonfigurationen bzw. Energielandschaften nur in ihrer Projektion auf die einzelnen Residuen und deren Winkelverteilungen, so ist es unmöglich, beispielsweise Aussagen über die auftretenden Sekundärstrukturmotive zu machen, nach deren Definition die Geometrien mehrerer aufeinanderfolgender Residuen gleichzeitig in bestimmten Gebieten des Ramachandran-Plots liegen müssen. Deshalb spielen Korrelationen zwischen den einzelnen Winkelkoordinaten eine große Rolle. Es liegt also nahe, die PCA auf die Energielandschaften anzuwenden.

Würde man die Hauptkomponenten einer der Trajektorien, beispielsweise der des *cis*-cAPBs, untersuchen, so könnte man einen Einblick in die „weichen“ Freiheitsgrade des dabei beobachteten Gemisches von Konformationen gewinnen. Im Hinblick auf die später durchzuführenden Nichtgleichgewichtssimulationen ist aber vielmehr eine gemeinsame Auswertung der jeweiligen *cis*- und *trans*-Energielandschaften der Modellpeptide von In-

teresse. Abbildung 5.2 skizziert die hiermit verbundene Idee: nimmt man an, dass sich die Geometrien der *cis*-Isomere eines der Peptide grundlegend von denen des dazugehörigen *trans*-Isomers unterscheiden, so sollte der aus beiden Simulationen zusammengesetzte Gesamtdatensatz eine hohe Varianz in diejenige Richtung aufweisen, welche die beiden Isomere im Winkelraum voneinander trennt. Eine Projektion der Daten auf diese ausgezeichnete Richtung erzeugt im gezeigten Beispiel zwei deutlich getrennte Verteilungen. Dagegen hätte eine Projektion auf eine der eingezeichneten kartesischen Koordinatenachsen ein mehr oder minder starker Überlapp der beiden Verteilungen zur Folge.

Ein solcher Überlapp konnte auch zuvor in den Ramachandran-Plots der Abbildung 7.2 beobachtet werden. Kann nun durch die PCA der drei Gesamtdatensätze aus den Simulationen des cAPBs, des cAMPBs und des bcAMPBs jeweils eine optimale Richtung zur Trennung der *cis*- und *trans*-Konformationsräume bestimmt werden? Solch ein einfach zu berechnendes und lineares Maß zur Klassifikation einer Peptidstruktur als *cis*- oder *trans*-artig wäre zur Auswertung der Isomerisierungssimulationen sehr hilfreich. Und welche Einsichten erlauben die Projektionen der hochdimensionalen Energielandschaften auf die ersten Hauptkomponenten?

Der Hauptkomponentenanalyse wurde nicht direkt auf den MD-Trajektorien ausgeführt. Vielmehr stützt sie sich auf die schon im vorangehenden Unterabschnitt verwendeten Schätzungen der Konfigurationsraumdichten durch jeweils 100 univariate Normalverteilungen. Auf diese Weise konnte der hohe Rechenaufwand vermieden werden, der mit der Verarbeitung der zweimal 20.000 Merkmalsvektoren der Trajektorien verbunden ist. Zusätzlich umfasst dieses Vorgehen schon vor der PCA eine Verminderung des unkorrelierten Rauschens der Daten, denn die Dichteschätzung mit univariaten Normalverteilungen erzeugt ein Datenmodell, in dem Rauschbeiträge senkrecht zu der durch die Daten aufgespannten Mannigfaltigkeit im Merkmalsraum unterdrückt sind [146].

Führt man die PCA auf den Orten der Kodebuchzentren aus, so erhält man das in Abbildung 7.4 gezeigte Eigenwertspektrum der Kovarianzmatrizen der drei Gesamtdatensätze. Bei allen drei Modellpeptiden tritt der erwartete exponentielle Abfall der Varianzen mit der Nummer der Hauptkomponente auf, besonders interessant ist jedoch der Umstand, dass sich gerade die ersten zwei Hauptkomponenten stark von den anderen Komponenten absetzen. Diese Beobachtung motiviert eine Projektion der Energielandschaften auf genau diese ersten zwei Komponenten, wie sie in Abbildung 7.5 dargestellt wird. Es werden zwei alternative Darstellungen gezeigt: einmal unter Verwendung von Höhenlinien, in denen Details der Energielandschaften besser erkennbar sind und eine dreidimensionale Berg- und Tallandschaft, die eine bessere Übersicht und Anschaulichkeit bieten. Schon ein erster Blick auf die abgebildeten Energielandschaften macht deutlich, dass durch die PCA bei keinem der drei Peptide eine vollständige Trennung der *cis*- und *trans*-Konformationsräume erreicht wird.

Am Beispiel der beiden cAPB Isomere ist zu sehen, dass die *cis*-Energielandschaft (rot) weiter ausgedehnt ist als die dazugehörige *trans*-Landschaft (blau). Während der Schwerpunkt der *cis*-Konfigurationen des cAPBs eher im oberen linken Quadranten der Abbil-

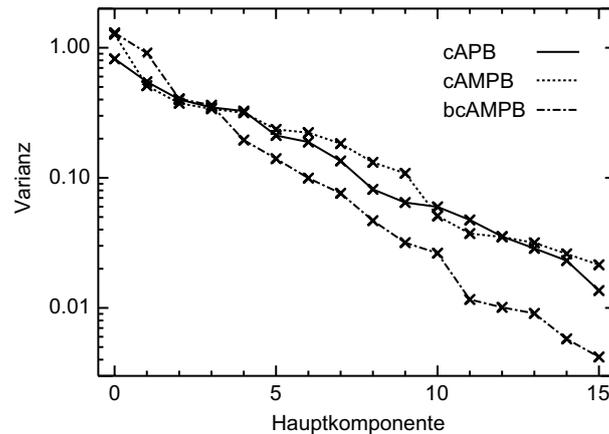


Abbildung 7.4: Die mit den Hauptkomponenten verbundenen Varianzen. Die nach ihrer Größe sortierten Eigenwerte der Kovarianzmatrix (5.12) bzw. die Varianzen der 16 Hauptkomponenten für die drei Gesamtdatensätze. Diese Datensätze wurden jeweils aus den Trajektorien der *cis*- und *trans*-Isomere der drei Modellpeptide zusammengesetzt.

Abbildung 7.5 a zu finden ist, liegt der Schwerpunkt der *trans*-Konfigurationen im unteren rechten Quadranten. Eine gewisse Trennung der beiden Konformationsräume kann auf diese Weise beobachtet werden, allerdings zeigen die Verteilungen einen starken Überlapp. Da dieser Überlapp auch in Projektion auf die weiteren Hauptkomponenten auftritt und dort erwartungsgemäß zunimmt (Daten hier nicht gezeigt), ist festzustellen, dass bei den Simulationen *cis*-Strukturen auftreten, welche in der Winkeldarstellung den *trans*-Strukturen nicht unähnlich sind. Innerhalb der *cis*-Energielandschaft des cAPBs erkennt man ein größeres und in sich strukturiertes Minimum und vier bis fünf Nebenminima, die durch Barrieren von etwa 0,5 bis 1 kcal/mol voneinander getrennt sind. Die dazugehörige *trans*-Energielandschaft weist zwei sehr ausgeprägte Minima auf, die über eine ähnlich hohe Barriere verbunden sind, während schwach ein drittes Minimum zu entdecken ist.

Die projizierten Energielandschaften des cAMPBs zeigen, verglichen mit denen des cAPBs, keine gravierenden Unterschiede. Im Fall des cAMPBs erscheint die *cis*-Landschaft etwas kompakter zu sein als zuvor. Auch treten ihre Nebenminima weniger deutlich hervor. Die *trans*-Landschaft des cAPBs hat ungefähr die gleiche Ausdehnung wie die des cAPBs und zeigt neben einem ausgeprägten Maximum noch ein einzelnes Nebenminimum. Interessanterweise scheint der in Abbildung 7.3 a skizzierte Ansatz, durch eine PCA die *cis*- und *trans*-Isomere zu trennen, im Fall des cAMPBs besser realisiert zu sein: zwar ist auch hier ein deutlicher Überlapp der beiden Konformationsräume festzustellen, jedoch gibt die erste Hauptkomponente eine gute Richtung zur deren Trennung an, wie man an der Lage der beiden Schwerpunkte erkennen kann.

Von einer Trennung der beiden Konformationsräume kann dagegen beim bcAMPB nicht gesprochen werden. Wie schon die Ramachandran-Plots der bcAMPB-Energielandschaften (Abbildung 7.2) vermuten ließen, tritt hier ein fast vollständiger Überlapp der beiden

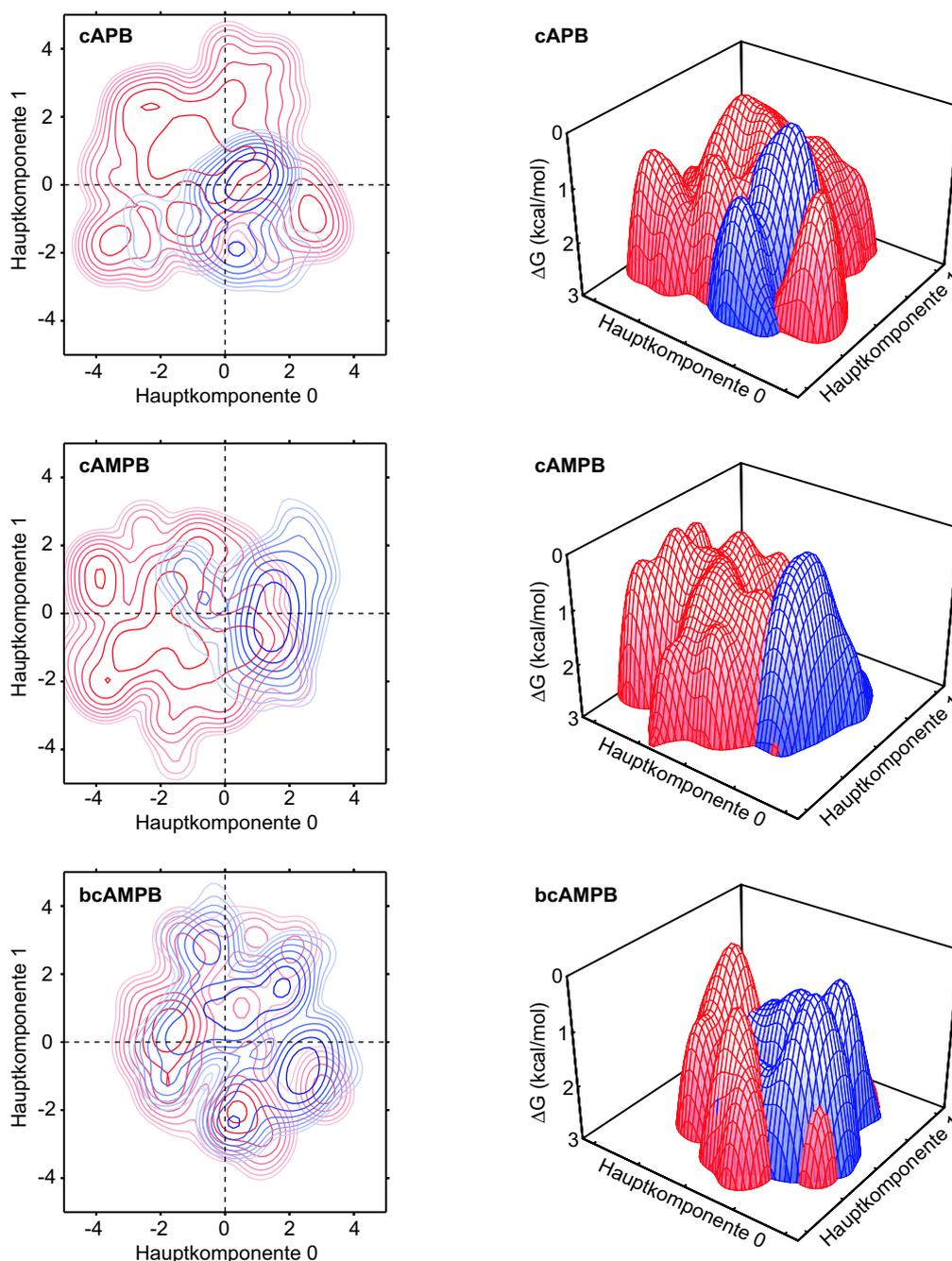


Abbildung 7.5: Die freie Energie in Projektion auf die ersten beiden Hauptkomponenten. Für die drei Modellpeptide cAPB, cAMPB und bcAMPB wird die aus den Gleichgewichtssimulationen bestimmte freie Energie ΔG relativ zum Energieminimum in ihrer Projektion auf die ersten beiden Hauptkomponenten der Gesamtdatenverteilung als Kontur- und als 3D-Plot dargestellt. Die Energieflächen der *cis*-Isomere sind rot, die der *trans*-Isomere blau gezeichnet. Die Hauptkomponenten sind in Radians angegeben und die Konturlinien haben einen Abstand von jeweils 0,5 kcal/mol. Zur besseren Übersicht wurde in den 3D-Plots die z-Achse umgedreht, so dass die Minima der freien Energie als Maxima erscheinen.

Verteilungen auf. Auch in ihrer Ausdehnung, Strukturierung und in der Anzahl der zu beobachtenden Minima sind sich beim bcAMPB beide Konformationsräume sehr ähnlich.

Die in Abbildung 7.5 vorgestellten zweidimensionalen Projektionen der Energielandschaften auf die ersten zwei Hauptkomponenten der Peptidkonfigurationen ermöglichen somit einen ersten anschaulichen Einblick in die Konformationsräume der Modellpeptide. Der Vergleich der Ergebnisse der *cis*- und der *trans*-Isomere zeigt für das cAPB und das cAMPB, dass dort ein Umschalten des Azobenzolfarbstoffs, wie erwünscht, zu deutlichen Veränderungen an dem für das Peptidrückgrat zugänglichen Konformationsraum führt. Im Fall des bcAMPBs scheint dies nicht im gleichen Maße zuzutreffen. Die angestrebte Definition einer linearen Reaktionskoordinate zur Bewertung einer beliebigen Peptidkonfiguration als mehr *cis*- oder *trans*-artig konnte nicht erreicht werden, was zu dem Schluss führt, dass sich die beiden Konformationsräume nicht linear trennen lassen. Auf die Diskussion der in Abbildung 7.5 beobachteten Minima der auf die Hauptkomponenten projizierten freien Energie möchte ich verzichten, da die Konformationen im folgenden Unterabschnitt der Peptide zuverlässiger im vollen 16-dimensionalen Raum aller Dihedralwinkel bestimmt werden können.

7.3 Konformationsanalyse durch Klassifikation der MD-Trajektorien

Mit dem in Abschnitt 5.3 vorgestellten ASSO-Verfahren steht eine Methode der unüberwachten Klassifikation zur Verfügung, die es erlaubt, die Hierarchie von Molekülkonformationen aus den Gleichgewichtssimulationen zu bestimmen. Als Ausgangspunkt dieses Verfahrens dienen wiederum die parametrischen Schätzungen der Konfigurationsraumdichten, die schon in den vorangehenden Unterabschnitten in Form der Energielandschaften diskutiert wurden. Das Resultat des ASSO-Verfahrens ist nach Abschnitt 5.3 in einer Menge von prototypischen Merkmalsvektoren \mathcal{Y}_σ (5.37) in Abhängigkeit von einer zu wählenden Auflösungsstufe σ , die alternativ auch durch den Skalenparameter ρ (5.32) angegeben werden kann. Nach Kapitel 5 bezeichnet jeder dieser Prototypen den Ort eines lokalen Maximums der skalierten Datendichte und damit eine Konformation des simulierten Moleküls.

Das ASSO-Verfahren wurde jeweils auf die Trajektorien der drei Modellpeptide in ihren beiden Isomerisierungszuständen angewendet. Die zu Grunde liegende Schätzung der Konfigurationsraumdichten wurde schon in den vorangehenden Unterabschnitten vorgestellt. Wichtig für die nun folgende Diskussion ist nur, dass bei den Dichteschätzungen stets $R = 100$ Kodebuchzentren Verwendung fanden und dass, zur besseren Vergleichbarkeit, einheitlich für alle Peptide und Isomerisierungszustände die selbe Gaußbreite $\sigma = 0,43$ angenommen wurde.

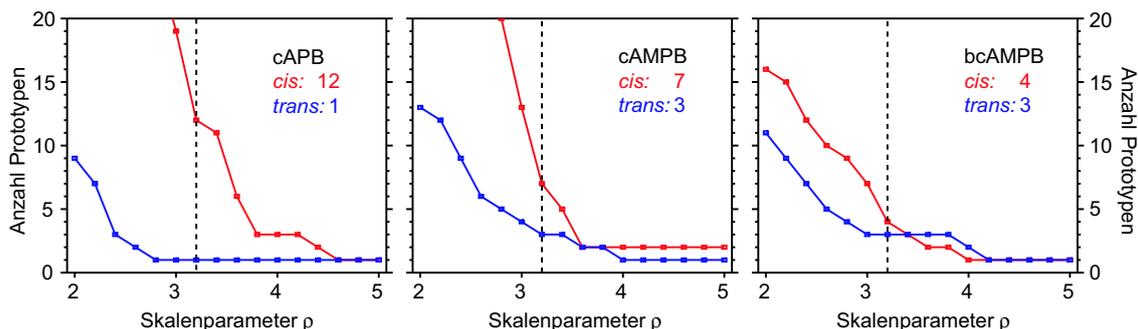


Abbildung 7.6: Die Anzahlen der Konformationen. Für die drei Modellpeptide und deren Isomere wurden mit Hilfe des ASSO-Verfahrens die Zahlen N_ρ der Konformationen als Funktionen des Skalenparameters ρ bestimmt. Eine gestrichelte Linie hebt die zur weiteren Diskussion ausgewählte Skalenstufe $\rho = 3,2$ ($\sigma = 1,39$) hervor. Die auf dieser Stufe gefundenen Anzahlen von Konformationen werden zusätzlich jeweils für die *cis*- (rot) und die *trans*-Isomere (blau) in die Diagramme eingetragen.

7.3.1 Die Anzahlen der bestimmten Konformationen

Die erste Observable bei der unüberwachten Klassifikation der Trajektorien ist die Zahl N_σ der gefundenen Konformationen als Funktion der Auflösungsstufe σ (bzw. ρ). Abbildung 7.6 zeigt diese Funktionen für die drei Modellpeptide und ermöglicht jeweils einen Vergleich der *cis*- und der *trans*-Isomere. Beginnt man die Diskussion mit dem *cis*-Isoomer des cAPBs, so wird dort für große Glättungen ($\rho \geq 4,6$) nur ein Prototyp detektiert, welcher den globalen Mittelwert aller bei der Simulation durchlaufenen Peptidkonfigurationen darstellt. Mit abnehmendem ρ beobachtet man das nach den Ausführungen von Abschnitt 5.3 erwartete monotone Ansteigen von N_ρ , d.h., dass bei verfeinerter Darstellung immer kleinräumigere Strukturen sichtbar werden. Beim *cis*-cAPB fällt auf, dass die Unterteilung in drei Konformationen über einen größeren Bereich von ρ stabil ist. Danach folgt eine stetige Zunahme von N_ρ . Die hochdimensionale Energielandschaft dieses Peptids weist demnach drei ausgeprägte Minima oder Konformationen auf, die bei näherer Betrachtung eine starke Strukturierung zeigen, also in Sub-Konformationen zerfallen.

Dagegen tritt das in Abbildung 7.5 beobachtete zweite Minimum der freien Energie beim *trans*-cAPB erst bei sehr viel kleineren Werten von ρ auf. Dies geschieht in einem Bereich, in dem das *cis*-Isoomer schon fast 20 Konformationen aufweist. Man erkennt an diesem Beispiel, dass der Vergleich der zu den beiden Isomeren gehörigen Graphen und besonders die Größe der dazwischen eingeschlossenen Fläche es ermöglicht, die Ausdehnung der beiden Konformationsräume in Beziehung zu setzen. Damit kann die Fähigkeit des Azobenzolfarbstoffes beurteilt werden, den Konformationsraum des Peptidteils durch die mit der Isomerisierung verbundene Geometrieänderung zu modifizieren. Im Fall des cAPBs bestätigt sich die schon zuvor gemachte Beobachtung, dass der Konformationsraum der *cis*-Isomere viel weiter ausgedehnt und strukturierter ist, als derjenige des *trans*-

Isomers. Es kann hier somit von einer hohen „Schaltfähigkeit“ des Modellsystems gesprochen werden kann.

Geht man zur Betrachtung des cAMPBs über, so fällt auf, dass im *cis*-Isomer auch bei großen Glättungen zwei Prototypen auftreten. Diese Beobachtung verweist auf zwei geometrisch stark unterschiedliche Gruppen von Konformationen. Ansonsten ist die *cis*-Kurve ihrem Gegenstück beim cAPB sehr ähnlich und zeigt einen vergleichbaren Anstieg im selben Bereich von ρ . Dagegen ist an der Kurve des *trans*-Isomers eine im Vergleich zum cAPB viel früher einsetzende Zunahme von N_σ zu beobachten, die auf eine Zunahme der konformellen Flexibilität des *trans*-Peptids durch den Einbau des Methylen-Spacers hinweist. Wie schon in den vorangehenden Abschnitten vermutet wurde, beeinflusst der Spacer dagegen den *cis*-Konformationsraum nur in geringem Maße. Betrachtet man wieder die Fläche zwischen den beiden Kurven als Maß für die Schaltfähigkeit des cAMPBs, so nimmt diese durch den Einbau des Spacers deutlich ab.

Das Schließen der Disulfid-Brücke führt nur zu geringen Veränderungen an der Kurve der N_ρ des *trans*-bcAMPBs, wie ein Vergleich mit dem Graphen des *trans*-cAMPBs zeigt. Tendenziell werden beim *trans*-bcAMPB etwas weniger Konformationen detektiert. Auffällig ist hier auch die Stabilität der Unterteilung in genau drei Konformationen. Wie auch schon in den vorangehenden Diskussionen vermutet wurde, schränkt die Disulfid-Brücke jedoch den *cis*-Konformationsraum des Peptids sehr stark ein. Während beim cAPB und beim cAMPB für Skalenstufen von $\rho \leq 3,0$ eine schier unüberschaubare Anzahl von Konformationen beobachtet werden konnte, ist der Anstieg beim *cis*-bcAMPB weit weniger ausgeprägt. Durch die Ähnlichkeit der Kurven der beiden Isomere des bcAMPBs in Abbildung 7.6 ist für dieses Modellpeptid nur noch von einer sehr geringen Schaltfähigkeit auszugehen.

Bevor eine detailliertere Untersuchung der beobachteten Konformationen stattfinden kann, ist es notwendig, eine Auflösungsstufe ρ zur Betrachtung auszuwählen, da sich sonst der Umfang der zu diskutierenden Ergebnisse als zu groß darstellt. Diese Auswahl ist mit einer gewissen Willkür verbunden. Zwar legt Abbildung 7.6 nahe, dass sich in der Umgebung von $\rho = 3,8$ die Zahl der klassifizierten Prototypen bei kleinen Variationen der Auflösung nicht stark ändert. Jedoch wird für das *cis*-bcAMPB bei dieser Auflösung eine geringere Zahl von Konformationen identifiziert als für den *trans*-Zustand, obwohl der Verlauf der Kurven für kleine Skalenstufen eine gegenläufige Tendenz andeutet. Wie Abbildung 7.6 ferner zeigt, steigt für $\rho < 3,0$ die Zahl der Konformationen des *cis*-cAPBs und des *cis*-cAMPBs so stark an, dass eine Darstellung und Diskussion aller dieser Strukturen sehr unübersichtlich wäre. Als Kompromiss wurde deshalb $\rho = 3,2$ gewählt. Diese Skalenstufe und die auf ihr beobachtete Anzahl von Konformationen ist in Abbildung 7.6 hervorgehoben. Somit wird die Gleichgewichtstrajektorie des *cis*-cAPBs im Folgenden in 12 verschiedene Konformationen unterteilt, während für das entsprechende *trans*-Isomer noch keine Unterteilung vorgenommen wird (12:1). Für das cAMPB wird ein Verhältnis von sieben *cis*-Konformationen zu drei *trans*-Konformationen beobachtet (7:3), während das bcAMPB vier zu drei Konformationen zeigt (4:3). Anhang C gibt einen tabellarischen

Überblick über die zu allen diesen Konformationen gehörenden (ϕ/ψ) -Winkel, die dort zusätzlich in einer einfachen binären Klassifikation S entweder als „ α -helikal“ oder „ β -faltblattartig“ bezeichnet werden.

7.3.2 Repräsentative Strukturen

In der bisherigen Diskussion der unüberwachten Klassifikation der Gleichgewichtstrajektorien wurde ausschließlich auf die Anzahl der gefundenen Konformationen N_ρ als Funktion der Auflösungsstufe eingegangen. Die mit einer der Konformationen verbundene Molekülgeometrie könnte zur Betrachtung wie zuvor in einen Ramachandran-Plot eingezeichnet werden. Da aber schon gezeigt wurde, dass diese Art der Darstellung nur einen beschränkten Einblick in eine Molekülkonfiguration erlaubt, und um einen weiteren Vergleich mit den veröffentlichten NMR-Strukturen zu ermöglichen, soll hier der Schritt zurück aus dem Raum der Dihedralwinkel in den der kartesischen Koordinaten aller Atome der Modellpeptide gemacht werden. Wie in Abschnitt 5.1 diskutiert wurde, fehlen in der dimensionsreduzierten Winkeldarstellung alle Informationen über die Orte der Atome der Peptid-Seitengruppen oder der Azobenzolfarbstoffe. Um dennoch für einen gegebenen Satz von Dihedralwinkeln dazu passende kartesische Atomkoordinaten des Gesamtmoleküls zu erhalten, ist es naheliegend, in der Gleichgewichtstrajektorie nach einer Molekülgeometrie zu suchen, deren Dihedralwinkel eine möglichst kleine quadratische Abweichung von den gegebenen Winkeln aufweisen. Die so gefundene repräsentative Struktur kann dann als Grundlage der Visualisierung der Konformation dienen.

Abbildung 7.7 zeigt die auf der im vorhergehenden Abschnitt ausgewählten Skalenstufe von $\rho = 3,2$ beobachteten Konformationen der drei Modellpeptide in ihren jeweils zwei Isomerisierungszuständen. Um diese Abbildung zu erzeugen, wurden die jeweiligen repräsentativen Strukturen der Konformere durch einen RMSD-Fit (siehe Abschnitt 5.1) auf eine dieser Strukturen abgebildet, danach eine mittlere Struktur berechnet und anschließend alle Strukturen auf diese mittlere gefittet. Bei den Fits wurde jeweils die Koordinaten aller C_α -Atome berücksichtigt. Um eine gute Anpassung der Farbstoffe aneinander zu erzielen, wurden von diesen jeweils sechs Atome (die beiden Stickstoffe und die Kohlenstoffe auf den Symmetrieachsen der Phenylringe) hinzugezogen. Tabelle 7.1 zeigt die beobachteten mittleren RMSDs. Für die dreidimensionale Zeichnung des Peptidrückgrats der Strukturen wurde die schon in Abbildung 1.14 verwendete Darstellungsform verwendet, bei der eine Spline-Funktion die C_α -Atome verbindet. Ferner werden die schweren Atome des Farbstoffs durch eine weitere zylinderförmige Darstellung hervorgehoben. Die Seitenketten und leichten Atome der Moleküle werden in dieser vereinfachten Abbildung nicht gezeigt.

Zusätzlich wird in Abbildung 7.7 auch noch eine Abschätzung der relativen freien Energien der Konformere vorgenommen. Diese Abschätzung beruht auf den bei den Gleichgewichtssimulationen beobachteten Besetzungen der Konformationssubzustände. Am Ende von Abschnitt 5.3 wurde erläutert, wie die Zuordnungswahrscheinlichkeit $\hat{P}(k|\mathbf{x})$ (5.42)

| Peptid | RMSD (Å) | |
|--------|----------|-------|
| | cis | trans |
| cAPB | 1,69 | 0,00 |
| cAMPB | 1,91 | 0,63 |
| bcAMPB | 1,75 | 0,61 |

Tabelle 7.1: Mittlere RMSDs der Peptid-Konformationen. Für die drei Peptide in ihren jeweils zwei Isomerisierungszuständen wurden die mittleren RMSDs (5.4) der in Abbildung 7.7 gezeigten repräsentativen Strukturen aller Konformationen in Bezug auf eine gemittelte Struktur berechnet. Da für das *trans*-cAPB auf der gewählten Auflösungsstufe nur eine Konformation gefunden wurde, verschwindet die RMSD für dieses Peptid.

einer durch den Merkmalsvektor \mathbf{x} beschriebenen Molekülgeometrie zur Konformation k aus dem Dichtemodell und dem Ergebnis der unüberwachten Klassifikation der Trajektorie zu berechnen ist. Die in einer Trajektorie beobachtete Besetzungswahrscheinlichkeit P_k einer Konformation k ist dann der zeitliche Mittelwert dieser Größe

$$P_k = \frac{1}{M} \sum_{t=1}^M \hat{P}(k|\mathbf{x}(t)), \quad (7.1)$$

aus der analog zu (5.2) die Differenz der freien Energie zweier Konformationen k und l zu

$$\Delta F = -k_B T \ln \frac{P_k}{P_l} \quad (7.2)$$

berechnet werden kann. In Abbildung 7.7 wird diese Differenz immer in Bezug auf die Konformation kleinster freier Energie in Einheiten von $k_B T$ angegeben, was bei $T = 500$ K fast exakt 1 kcal/mol entspricht. Die gezeigte Energieskala verdeutlicht die in den Graustufen der Moleküldarstellung kodierte Energiedifferenzen: dunkle Schattierungen für energetisch niedrige Konformationen, helle Töne für nur schwach besetzte, energetisch höherliegende Konformationen.

Schließlich stellt Abbildung 7.8 den aus den MD-Simulationen bestimmten Konformationen die entsprechenden von Renner et al. [37, 38] veröffentlichten NMR-Strukturen gegenüber. Es wurden jeweils die 10 energetisch niedrigsten Strukturen des NMR-Refinements ausgewählt und deren C_α -Atome durch einen RMSD-Fit aufeinander gelegt. Es ist zu betonen, dass im Fall des bcAMPBs nicht die Strukturen der Originalveröffentlichung [38] gezeigt werden, sondern die bisher unveröffentlichten Ergebnisse neuer Messungen.

Um den Vergleich zwischen den MD-Konformationen und den NMR-Strukturen besser herausarbeiten zu können, möchte ich die Abbildungen 7.7 und 7.8 gleichzeitig diskutieren. Beginnt man die Betrachtung mit dem *cis*-cAPB, so weisen die zwölf dort gefundenen

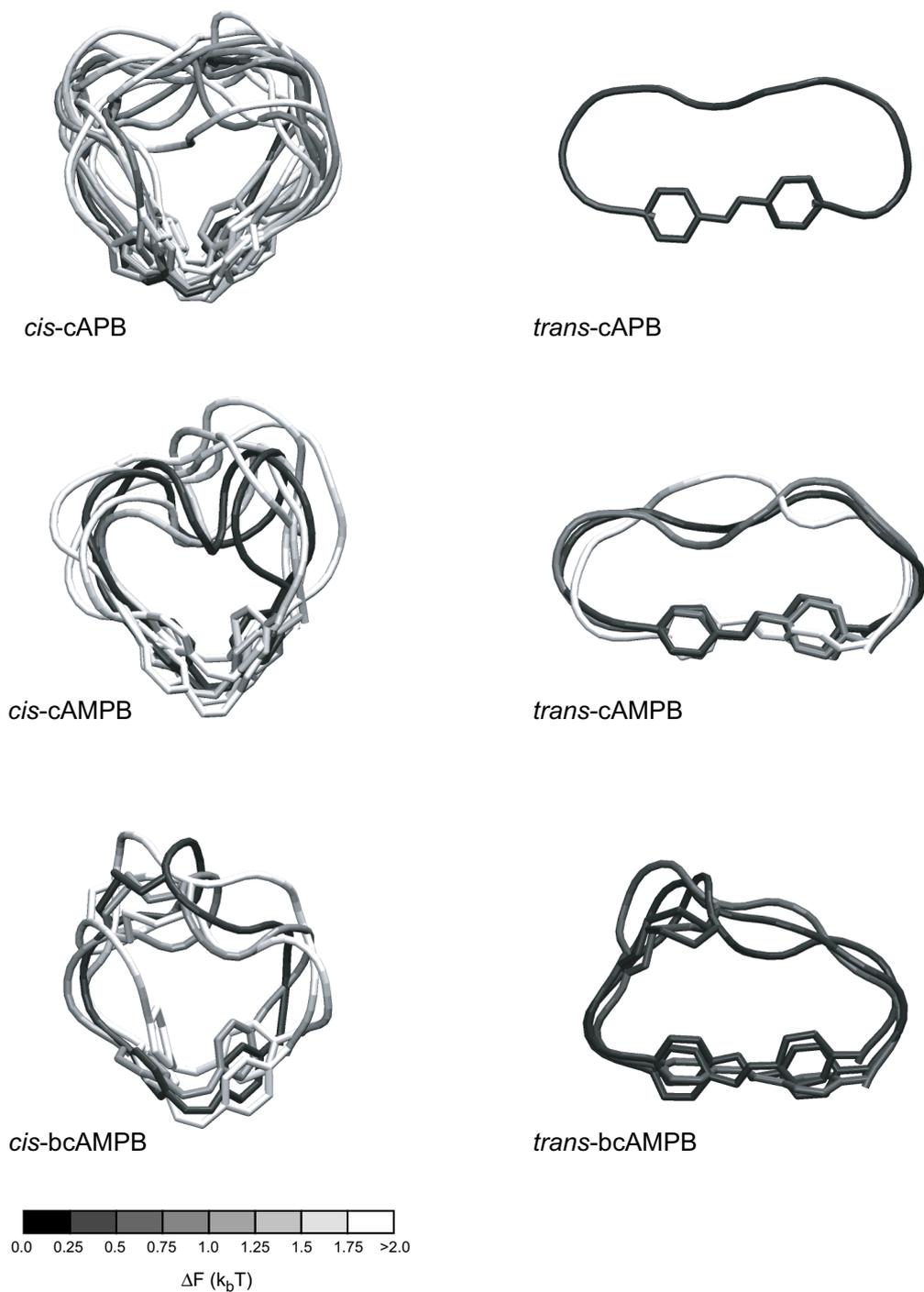


Abbildung 7.7: Die beobachteten Konformationen der Modellpeptide. Für jedes der drei Modellpeptide werden die in den MD-Simulationen bei 500 K beobachteten *cis*- und *trans*-Konformationen dargestellt. Die jeweiligen Anzahlen und Geometrien der Konformationen wurden durch eine unüberwachte Klassifikation aus den Trajektorien bei ermittelt. Die Legende gibt die Skala der Graustufen an, mit denen die freie Energie der Konformationen relativ zum energetisch niedrigsten Zustand angegeben wird.

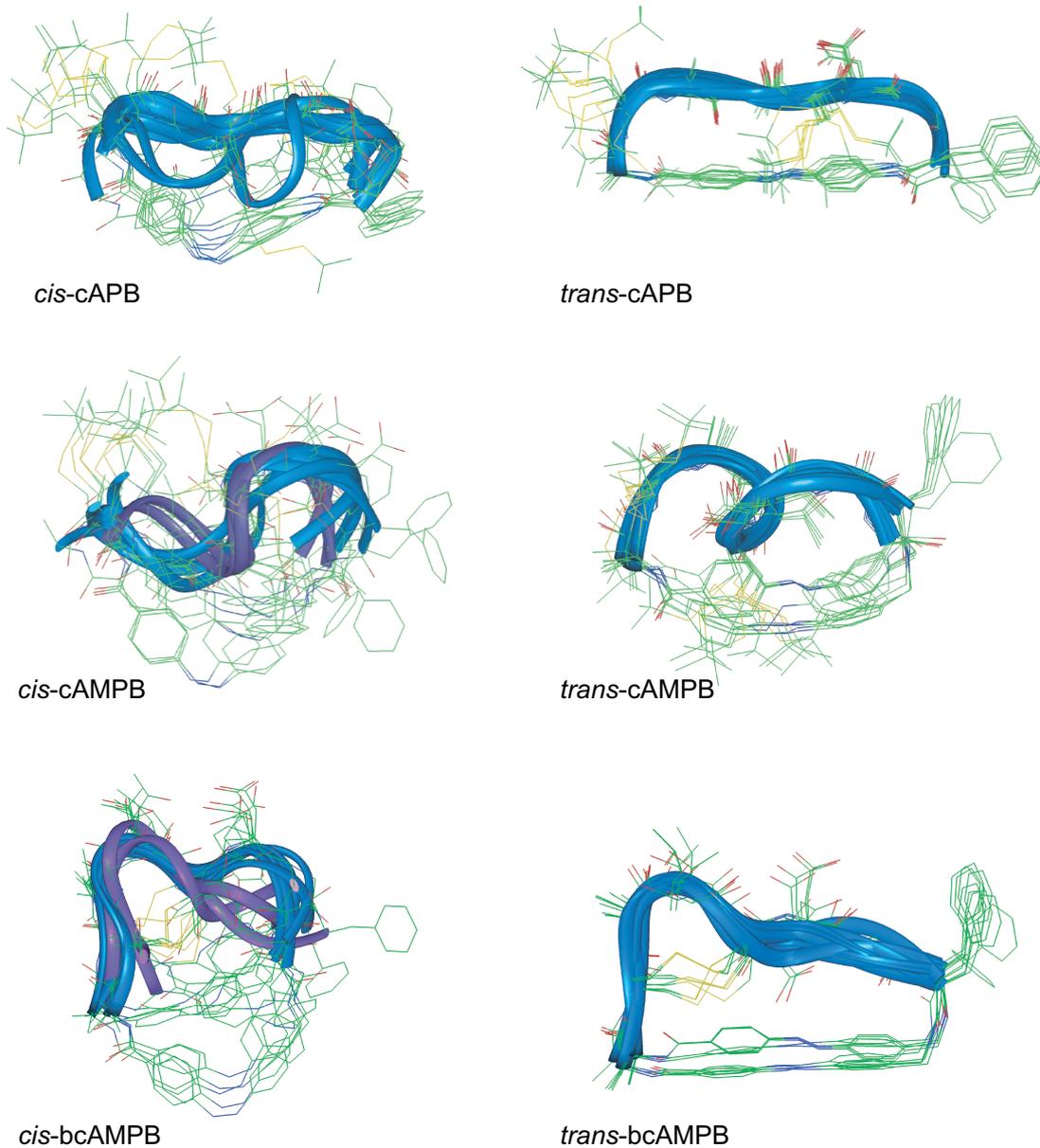


Abbildung 7.8: Die NMR-Strukturen. Die hier gezeigten Abbildungen für das cAPB und das cAMPB wurden den Veröffentlichungen von Renner et al. [37, 38] entnommen. Die Strukturen des bcAMPBs entstammen erneuten, bisher unveröffentlichten Messungen.

MD-Konformationen recht unterschiedliche Geometrien auf. Sie sind teils helikaler Natur, einige Konformationen weisen aber auch weite, ungeordnete Geometrien auf, die üblicherweise als „Loops“ bezeichnet werden. Die Einfärbung der dargestellten Strukturen zeigt, dass mehrere sehr unterschiedliche Konformationen eine vergleichbar niedrige freie Energie besitzen, also davon ausgegangen werden kann, dass diese Konformere auch bei 300 K zu beobachten sind. Dies gilt nicht unbedingt für diejenigen Konformationen, deren Energiedifferenz mehr als etwa 1 kcal/mol beträgt. Dieser Umstand besitzt eine gewisse Relevanz für den Vergleich mit den entsprechenden NMR-Strukturen, die als Minimalstrukturen aus dem Refinement hervorgegangen sind und damit einem Ensemble bei 0 K entsprechen. Hier, im Fall des *cis*-cAPBs, wirken die NMR-Strukturen viel kompakter als diejenigen der MD-Konformationen. Neben zwei helikaler Strukturen beobachtet man dort eine Reihe von gestreckten Geometrien, die einander sehr ähnlich sind. Bei allen der zehn gezeigten NMR-Strukturen tritt eine sehr kleine Distanz zwischen den mittleren Teilen des Peptidrückgrats und dem Farbstoff auf. Dies steht im Gegensatz zu den MD-Strukturen, bei denen häufig eine fast kreisförmige Geometrie des Moleküls auftritt.

Eine große Übereinstimmung zwischen den NMR- und den MD-Untersuchungen herrscht bei der Frage der Geometrie des Peptidrückgrats des *trans*-cAPBs: sowohl die einzige MD-Konformation als auch die sehr uniformen NMR-Strukturen weisen eine gestreckte Form des Peptidteils auf. Unterschiede der beiden Abbildungen sind auf verschiedene Definitionen der Tube-Darstellung zurückzuführen. Eine deutliche Differenz zeigt sich jedoch in der Orientierung der Phenylringe des Azobenzolfarbstoffs, die bei allen dreien der *trans*-Modellpeptide bei den MD-Simulationen in der Ebene des ringförmigen Moleküls liegen, während sie bei den NMR-Strukturen senkrecht dazu stehen.

Geht man zum *cis*-cAMPB über, ist bei den MD-Simulationen im Vergleich zum *cis*-cAPB noch einmal die schon diskutierte Abnahme der Anzahl von Konformationen zu beobachten. Während bei den bisherigen Untersuchungen zwischen diesen beiden Modellsystemen keine gravierenden Veränderungen durch den Einbau des Methylenspacers festgestellt werden konnten, fällt in Abbildung 7.7 die stärker ausgeprägte Helizität der energieärmsten Konformationen des *cis*-cAMPBs auf. Diese Beobachtung deutet darauf hin, dass erst durch die Kombination des Spacers mit dem *cis*-Farbstoff die konformelle Einschränkung des Peptidteils so weit aufgehoben wird, dass dem Peptidrückgrat die Ausbildung der „natürlichen“ helikalen Struktur gelingt. Die NMR-Messungen am *cis*-cAMPB bestätigen diese Beobachtung: für alle zehn Strukturen wird hier eine Windung mit zwei Umdrehungen angenommen. Allerdings wird diese helikale Struktur aus den NMR-Messungen auch für das *trans*-cAMPB vorhergesagt. Hier lässt das Refinement nur eine einzige Konformation zu. Die dabei bestimmte Helix des Peptidrückgrats ist offensichtlich nicht ganz mit den vom Farbstoff erzeugten Randbedingungen vereinbar, was zu teilweise stark gebogenen und unter Spannung stehenden Konfigurationen der Azobenzolgruppe führt. Diesen gewundenen Strukturen des *trans*-cAMPBs widersprechen die MD-Simulationen, bei denen die Konformere des *trans*-cAMPBs zwar eine etwas größere Variabilität als beim *trans*-cAPB, aber immer noch eine gestreckte Gestalt aufweisen. Wenn man die energetisch ungünstigste der drei Konformationen vernachlässigt, lassen die Si-

| <i>cis</i>-cAPB | | | | | | | | | | | | |
|--------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Konformation | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| RMSD (Å) | 2,00 | 0,47 | 1,27 | 1,44 | 0,00 | 2,27 | 1,08 | 1,97 | 1,72 | 1,47 | 1,56 | 1,15 |
| <i>trans</i>-cAPB | | | | | | | | | | | | |
| Konformation | 1 | | | | | | | | | | | |
| RMSD (Å) | 1,15 | | | | | | | | | | | |

Tabelle 7.2: Mittlere RMSDs der cAPB-Konformationsensembles. Für jeder der in Abbildung 7.9 gezeigten Konformationen wurden die RMSDs der zugeordneten Strukturen gemittelt.

mulationsergebnisse nur im rechten Teil des Peptids Auswirkungen des Spacers erkennen.

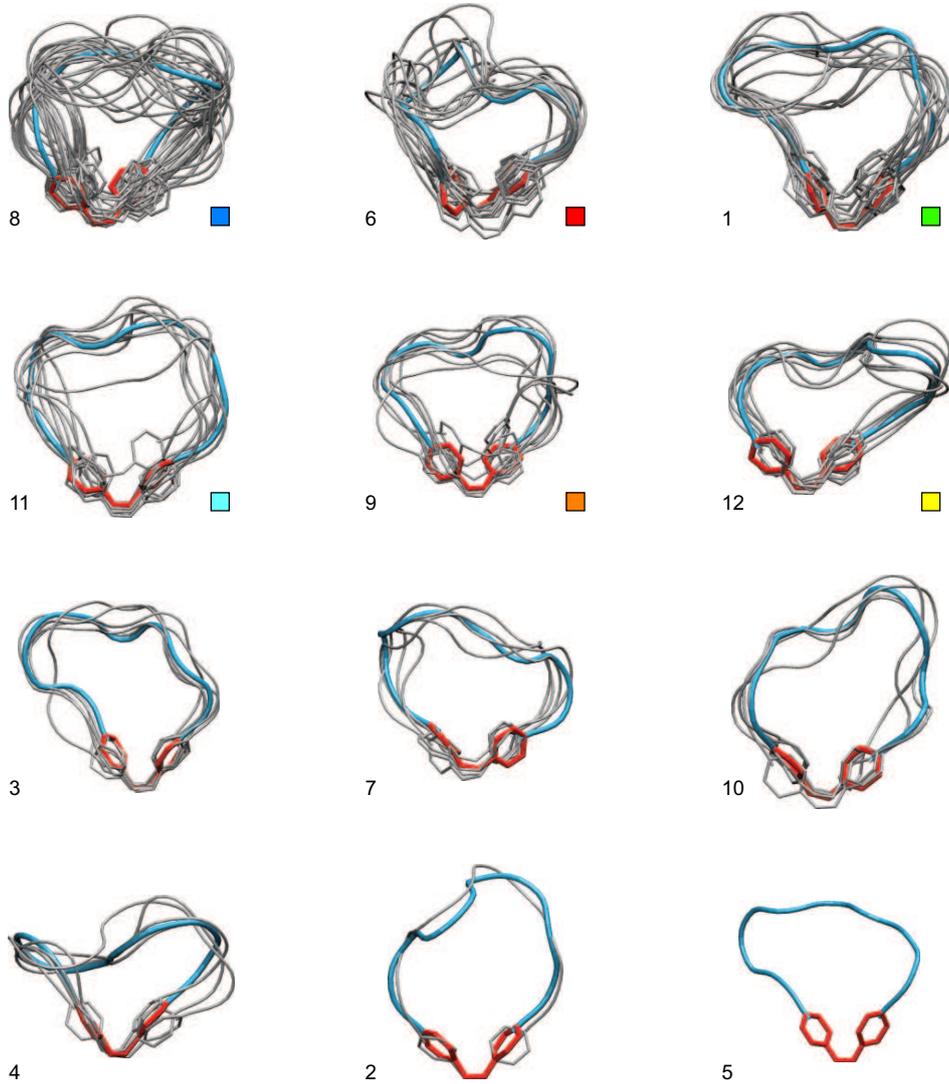
Wie schon zuvor festgestellt wurde, beobachtet man bei den Simulationen des bcAMPBs eine vergleichbare Flexibilität in den *cis*- und den *trans*-Konformationsräumen. Nach Abbildung 7.7 zeigen die Geometrien des Peptidrückgrats bei beiden Isomeren größere Ähnlichkeiten als dies bei den anderen Modellpeptiden der Fall ist. Während allerdings im *cis*-Zustand neben einer energetisch favorisierten Konformation auch einige abweichende, energetisch höherliegende Konformere existieren, sind sich die drei Konformationen des *trans*-bcAMPBs geometrisch sehr ähnlich und wurden auch in der Trajektorie zu fast gleichen Anteilen beobachtet. Der Vergleich der MD-Konformationen mit den aus den neuen Messungen bestimmten NMR-Strukturen fördert beim bcAMPB eine fast überraschende Übereinstimmung zu Tage. Nur über das Ausmaß der Variabilität der Konformationen besteht keine Einigkeit. Auch im Fall des bcAMPBs erzeugt das NMR-Refinement deutlich kompaktere Verteilungen.

Zusammenfassend kann folgendes festgehalten werden: beschreibt man wie in den Abbildungen 7.7 und 7.8 die Geometrie des Peptidrückgrats alleine durch die räumliche Lage der C_{α} -Atome, so ist für die Geometrien der Peptidteile einiger der Modellsysteme, nämlich das *trans*-cAPB, das *cis*-cAMPB und die beiden bcAMPBs, eine erstaunlich gute Übereinstimmung zwischen Simulation und NMR-Messung festzustellen. Bei den anderen Peptiden treten hingegen teilweise deutliche Unterschiede auf. Ferner beobachtet man tendenziell bei den MD-Simulationen eine größere Variabilität der möglichen Peptidkonformationen, wogegen die Ensembles der NMR-Strukturen zumeist kompakter erscheinen.

Die Konformationen des cAPBs im Detail

Für die im nächsten Abschnitt anstehende Diskussion der NMR-Restraints ist es unabdingbar, auf die in den Simulationen beobachteten Konformationen der Peptide detaillierter als zuvor geschehen einzugehen. An Stelle der in Abbildung 7.7 gezeigten Überlagerungen aller Konformationen, werden in Abbildung 7.9 am Beispiel des cAPBs die gefundenen Konformere einzeln dargestellt. Zusätzlich zu der farbig hervorgehobenen Mo-

***cis*-cAPB**



***trans*-cAPB**

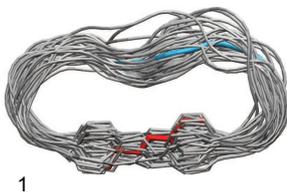


Abbildung 7.9: Die Konformationen des cAPBs. Die in Abbildung 7.7 gezeigten Konformationen des *cis*- und *trans*-cAPBs werden einzeln dargestellt und sind farblich hervorgehoben. Zusätzlich werden in grau die zugeordneten Strukturen gezeigt, die jeweils ein Kodebuchzentrum der Dichteschätzung repräsentieren. Die Ziffern benennen die Konformationen, die absteigend nach ihrer Besetzung (7.1) sortiert sind. Ferner werden den ersten sechs Konformationen eindeutige Farben zugeordnet.

lekülgeometrie mit welcher der Prototyp der entsprechenden Konformation repräsentiert wird, gibt Abbildung 7.9 aber auch einen Einblick in die Ausdehnung des entsprechenden Konformationssubraumes. Dazu werden jeder Struktur minimaler freier Energie eine Auswahl von Molekülkonfigurationen überlagert, die als zu dieser Konformation zugehörig klassifiziert wurden. In Übereinstimmung mit dem am Ende von Abschnitt 5.3 formulierten zweistufigen Klassifikationsverfahren wurden hierzu diejenigen Molekülgeometrien aus der Trajektorie verwendet, die den 100 Kodebuchzentren der Dichteschätzung der Gesamtrajektorie entsprechen. Die jeweilige Anzahl der zu einer Konformation zugeordneten Strukturen spiegelt vergrößert die schon zuvor betrachtete Besetzung (7.1) wieder. Tabelle 7.2 führt die gemittelten RMSDs auf, die bei der Anpassung der repräsentativen Strukturen an die Struktur des Minimums beobachtet wurden. Die in Abbildung 7.9 und im Folgenden für die Konformationen verwendeten Bezeichnungen („1“-„12“) entstammen der Reihenfolge, in der diese Zustände während der Klassifikation der Trajektorien aufgefunden wurden. Da diese Reihenfolge keinerlei physikalische Bedeutung hat, wurden die Konformationen in Abbildung 7.9 absteigend nach ihrer Besetzung (7.1) bzw. nach zunehmender freier Energie sortiert. Schließlich wird den ersten sechs Konformationen jeweils zusätzlich eine Farbe zugeordnet, was zur besseren Übersichtlichkeit späterer Abbildungen beitragen soll.

In ihrer Gesamtheit betrachtet, verstärken die in Abbildung 7.9 gezeigten Konformationen des cAPBs noch einmal die Beobachtungen, die schon anhand von Abbildung 7.7 gemacht worden: das *trans*-Isomer des cAPBs weist genau eine gut definierte Konformation auf. Tabelle 7.2 weist für die der *trans*-Konformation zugeordneten Strukturen eine sehr kleine mittlere RMSD von 1,15 Å aus. Diese Strukturen sind sich demnach nicht nur im Winkelraum sondern auch in ihren kartesischen Koordinaten sehr ähnlich. Allenfalls im Bereich der Residuen^a 5 bis 7 ist eine gewisse Variabilität zu beobachten. Dagegen weist das *cis*-cAPB eine Vielzahl von Konformationen auf, deren Geometrien sich stark voneinander unterscheiden. Bei vielen der MD-Konformationen des *cis*-cAPBs tritt, im Gegensatz zu den Strukturen des NMR-Refinements (Abbildung 7.8), ein großer Abstand zwischen dem Azobenzol-Farbstoff und dem ihm gegenüberliegenden Teil des Peptidrückgrats auf. Gute Beispiele hierfür sind die Konformationen 11, 2 und besonders die 10. Betrachtet man nur die farblich hervorgehobenen „Mittelpunkte“ der *cis*-Konformationen, so kann man in deren Geometrien eine Vielzahl Sekundärstruktur motive entdecken: eine helikale Anordnung beispielsweise beim Konformer 12, gestreckte Formen bei den Konformeren 11 und 7 und vor allem eine Vielzahl von „Turns“, in denen das Peptidrückgrat eine scharfe Wendung vollführt. Bei genauerer Betrachtung dieser β -Turns (hier nicht gezeigt) ist häufig eine Stabilisierung durch Wasserstoff-Brücken zu beobachten.

Ich möchte die ersten sechs Konformationen im Detail diskutieren, soweit dies Abbildung 7.9 erlaubt, in der die Strukturen nur von einem Blickpunkt aus dargestellt werden. Die Konformation Nummer 8 weist die größte Besetzung (7.1) auf. Ihre Geometrie ist nicht allzu gut definiert, wie auch der RMSD von knapp 2 Å zeigt. Bestimmend für die

^aZur Reihenfolge der Residuen siehe Abbildung 1.14 oder auch Abbildung 7.10.

Konformation 8 scheint jedoch eine helikale Anordnung der Residuen 1 bis 4 zu sein, während die Residuen 5 bis 7 einen Turn bilden. Dieser Turn ist in der gewählten Darstellung auf den Beobachter hin ausgerichtet.

Im Vergleich dazu tritt bei der Konformation 6 ein Turn im linken Teil der Darstellung auf, etwa bei den Residuen 1 bis 4. Auffällig ist, dass die Konformation 6 aus zwei Sub-Konformationen zusammengesetzt ist: bei der einen weist der Turn in Richtung des Betrachters, bei der anderen weg von ihm. Dies führt auch hier zu einem relativ großen RMSD des Ensembles von Strukturen. Betrachtet man die Öffnung des durch das Molekül gebildeten Rings, erscheint die Konformation 6 jedoch etwas kompakter zu sein, als Nummer 8.

Konformation 1 zeigt wieder einen etwas weiter geöffneten Ring und eine gestreckte Konfiguration zwischen den Residuen 3 und 6. In diesem Bereich ist die räumliche Lage des Rückgrats nicht festgelegt, während dies bei den anderen Residuen der Fall ist. Bei genauerer Betrachtung stellt man fest, dass auch die Konformation 1 aus mindestens zwei Sub-Konformationen aufgebaut ist.

Ignoriert man diejenige Struktur bei der Konformation 11, die nicht richtig angepasst werden konnte, so sind die übrigen Strukturen als weit ausgedehnte „Loops“ zu klassifizieren, die im Bereich der Residuen 3 und 4 eine leichte Helizität aufweisen. Insgesamt stellt Konformation 11 wieder einen weit geöffneten Ring dar.

In Konformation 9 ist ein Konformer relativ niedriger freier Energie mit zwei weiteren Konfigurationen zusammengefasst worden. Während ersteres fast symmetrisch bezüglich einer Spiegelung an der Vertikalen erscheint, weisen die anderen Strukturen ein geschwungenes Peptidrückgrat im rechten Teil der Abbildung auf.

Schließlich kann mit Konformation 12 ein gut definierter Zustand beobachtet werden. Die zugeordneten Strukturen weisen untereinander eine große Ähnlichkeit auf, was zu einem RMSD von nur etwas mehr als 1 Å führt. Die Geometrie dieser Konformation zeigt einen gut ausgeprägten Turn bei den Residuen 5 bis 7, der bezüglich des Farbstoffs stark nach rechts verschoben ist.

7.4 Erfüllung der NMR-Restraints

Schon zu Beginn des vorliegenden Abschnitts wurde erläutert, dass die oben präsentierte Gegenüberstellung der bei den MD-Simulationen auftretenden Molekülkonformationen mit den im Zuge des NMR-Refinements bestimmten Strukturen nicht der optimale Weg ist, um die Ergebnisse von Theorie und Experiment miteinander zu vergleichen. Von größerer Aussagekraft ist ein Rückgriff auf die im NMR-Experiment gemessenen Abstände zwischen Paaren von Wasserstoffkernen, die als Zwangsbedingungen oder „Restraints“ in das NMR-Refinement eingehen. Zu diesem Thema wurde in Abschnitt 1.2 kurz erläutert, wie aus dem bei der 2D-NMR-Spektroskopie beobachteten Nuklearen Overhauser Effekt

(NOE) jeweils eine Obergrenze für solche Abstände ermittelt werden kann.

Die zum Vergleich mit NOE-Constraints benötigten Abstände gegebener Wasserstoffatome liegen bei den MD-Trajektorien direkt vor. Für den Vergleich ist es vorteilhaft, die Abhängigkeit des NOEs von der inversen sechsten Potenz des Protonenabstandes zu berücksichtigen, indem der bei einer MD-Trajektorie beobachtete Abstand r nach

$$\bar{r} = \left\langle r^{-6} \right\rangle^{-1/6} \quad (7.3)$$

gewichtet gemittelt wird [87, 294]. Die Implementierung dieser Abstandsberechnung in einem MD-Programmpaket wird etwas durch die Tatsache erschwert, dass manche Protonen in den NMR-Spektren aus Symmetriegründen ununterscheidbar sind. Die Orte solcher Atome werden deshalb im NMR-Refinement geometrisch gemittelt und die auf diese Weise bestimmte räumliche Position als „Pseudoatom“ behandelt. Solche Pseudoatome sind also bei der Auswertung der MD-Trajektorien zu berücksichtigen.

In Anhang D zeigen die zu den drei verschiedenen Modellpeptiden und ihren Isomerisierungszuständen gehörigen Tabellen die Ergebnisse der Abstandsmittlung (7.3) für diejenigen Protonenpaare, für die experimentell ein NOE beobachtet wurde. Die entsprechenden Daten wurden freundlicherweise von Christian Renner^a zur Verfügung gestellt. Die bei den Simulationen auftretenden Protonenabstände wurden jeweils sowohl in Bezug auf die gesamte Trajektorie, als auch unter Berücksichtigung der zuvor diskutierten Konformationen gemittelt. In Anhang D werden anstatt der konformationslokalen Mittelwerte deren Abweichung von den experimentellen NOEs tabelliert.

7.4.1 Die beispielhafte Diskussion einiger NOEs

Anstatt sofort mit der Diskussion der Vielzahl von NOEs zu beginnen, die an den drei unterschiedlichen Modellpeptiden beobachtet wurden, möchte ich zuerst beispielhaft und vertieft auf fünf das *cis*-cAPB betreffende Abstände eingehen. Diese ausgewählten NOEs werden in Abbildung 7.10 vorgestellt und in Tabelle D.1 unter den Nummern 2, 8, 11, 20 und 33 geführt.

Um nicht ausschließlich darauf beschränkt zu sein, die experimentellen NOEs mit den gemittelten Protonenabständen aus Anhang D zu vergleichen, zeigt Abbildung 7.11 ergänzend die Zeitentwicklungen dieser Abstände im Verlauf der aufgezeichneten Trajektorien und ihre Verteilungen in Form von Histogrammen, jeweils unter Berücksichtigung der Klassifikation der Molekülgeometrien zu einer der sechs wichtigsten Konformationen des Peptidrückgrats aus Abbildung 7.9. Um die Verbindung der Abstandsverteilungen mit den dort abgebildeten Konformationen hervorzuheben, wird die gleiche Farbkodierung verwendet. Zeitpunkte der Trajektorien, bei denen die Peptidkonfiguration nicht einer der sechs wichtigen, sondern einer der anderen, energetisch höherliegenden Konformationen

^aMPI für Biochemie, Martinsried

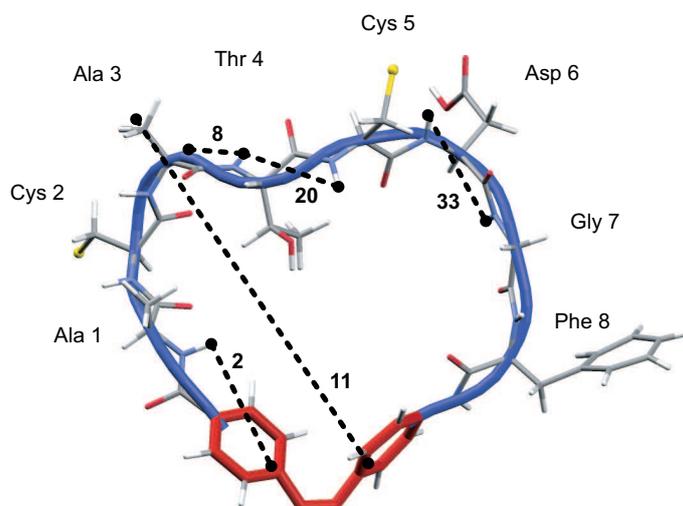


Abbildung 7.10: Einige NOEs des *cis*-cAPBs. In eine Darstellung der Konformation „11“ des *cis*-cAPBs wurden die im Text detailliert diskutierten NOEs eingezeichnet. Die gestrichelten Linien verbinden die jeweils betroffenen Wasserstoffatome oder Pseudoatome. Zusätzlich werden die NOEs durch ihre Nummern gekennzeichnet (siehe Anhang D).

zugeordnet wurden, werden in Abbildung 7.11 aus Gründen der Übersichtlichkeit weggelassen. Abbildung 7.11 zeigt zusätzlich die experimentellen NOEs als gestrichelte Linien, während die durchgezogenen schwarzen Linien das Ergebnis der nach (7.3) bezüglich der Gesamtrajektorie durchgeführten Mittelung darstellen.

Beginnt man die Diskussion der Abbildung 7.11 mit dem NOE 2, so erkennt man anhand der Molekülstruktur aus Abbildung 7.10, dass dieser NOE den Abstand zwischen dem HN des Ala 1 und einem Pseudoatom misst, das in der Ebene eines der Phenylringe des Azobenzols liegt. Die Starrheit dieses Rings und die des angehängten Peptidplättchens hat zur Folge, dass zwischen den beiden betroffenen Atomen nur eine einzige C–C-Bindung frei drehbar ist. Da das Pseudoatom aber stets nahe der Achse dieser Drehung liegt, ändert sie den betrachteten Abstand nicht. Diese Überlegung wird durch Abbildung 7.11 bestätigt, die für den NOE 2 nur geringe zeitliche Variationen ausweist. Diese Abstandsvariationen werden durch thermische Fluktuationen von Bindungslängen und -winkeln verursacht. Die konformationsspezifischen Abstandsverteilungen sind daher alle unimodal und miteinander identisch. Die Simulationen reproduzieren also den experimentellen NOE 2 trivialerweise für alle Konformationen. Damit ist anzunehmen, dass der NOE 2 keinen Einfluß auf das Ergebnis eines NMR-Refinement haben kann.

Eine andere Beobachtung ist am Beispiel des NOEs 8 zu machen, welcher sich auf den Abstand zwischen dem HA des Ala 3 und dem HN des darauf folgenden Residuums bezieht. Wie im vorangehenden Beispiel ist auch hier zwischen den beiden betroffenen Atomen nur ein einziger Dihedralwinkel frei drehbar: der Winkel ψ des Ala 3. In der Monographie von Wüthrich [39] wird ein solcher Protonenabstand mit $d_{\alpha N}$ bezeichnet. Aus der

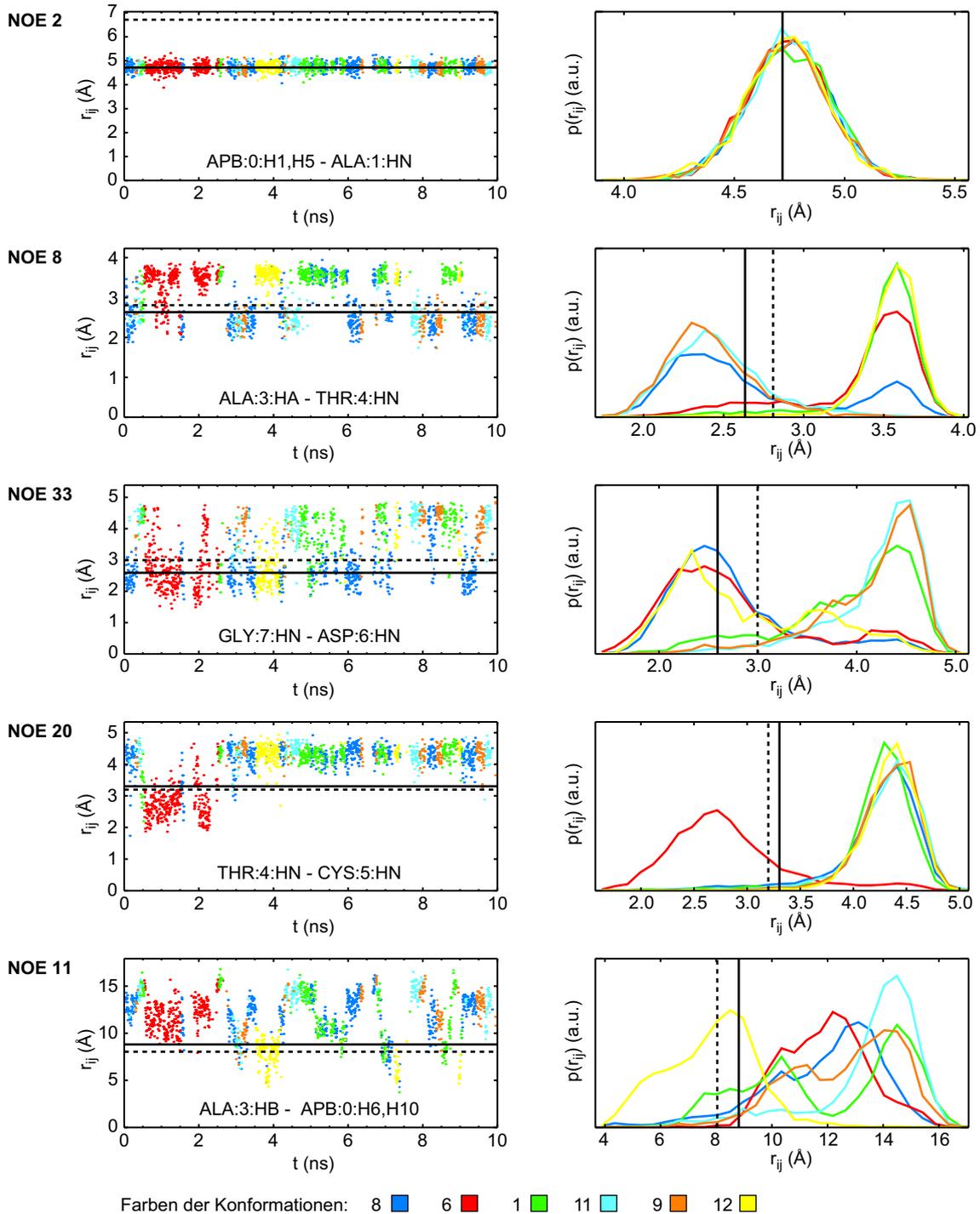


Abbildung 7.11: Detaillierte Ergebnisse der Simulation zu einigen der NOEs. Für fünf der NOEs des *cis*-cAPBs werden die bei der Simulation beobachteten Inter-Proton-Abstände gezeigt: links die Zeitreihe und rechts die Verteilung der Abstände als Histogramm. Farblich werden die sechs wichtigen Konformationen aus Abbildung 7.9 kodiert. Die gestrichelte Linie zeigt jeweils die experimentelle Obergrenze des NOEs und die durchgezogene Linie den nach (7.3) über die Gesamtrajektorie gemittelten Abstand.

Geometrie der Bindungen zwischen den beiden Wasserstoffatomen folgt, dass $d_{\alpha N}$ mit dem Kosinus von ψ variiert: für $\psi \approx 130^\circ$, also im Bereich der β -Faltblatt-Sekundärstruktur, ergibt sich ein Wert von weniger als $2,5 \text{ \AA}$, während $\psi \approx 90^\circ$ (α -Helix) einen weit größeren Wert von etwa $3,5 \text{ \AA}$ zur Folge hat. Die für den Dihedralwinkel ψ (Ala 3) in Abbildung 7.2 auftretenden beiden Minima der freien Energie erzeugen dementsprechend in Abbildung 7.11 eine bimodale Abstandsverteilung. In einer einzelnen Konformation des *cis*-cAPBs bestimmt ψ (Ala 3) damit, ob sie den NOE 8 erfüllt oder nicht. Dies ist gut an den zugehörigen Histogrammen zu erkennen, die für die Konformationen 11 und 9 hauptsächlich kleine Protonenabstände ausweisen, während die Abstände bei den Konformationen 6, 1 und 12 die experimentelle Obergrenze von $2,8 \text{ \AA}$ überschreiten (diese Unterteilung stimmt mit der einfachen Klassifikation der Geometrien der Residuen aus Tabelle C.1 überein). Auffällig ist dagegen, dass die Konformation 8 schon für sich alleine betrachtet eine bimodale Abstandsverteilung besitzt. Diese Beobachtung weist darauf hin, dass die Konformation 8 mehrere Sub-Konformationen umfasst. Betrachtet man den gewichteten Mittelwert des mit dem NOE 8 verbundenen Protonenabstandes (schwarze Linie) so erkennt man, dass das Auftreten von Konformationen mit großen Abständen $d_{\alpha N}$ nur eine geringe Auswirkung auf den gewichteten Mittelwert hat. Vielmehr führt wegen (7.3) das zeitweilige Auftreten von Konformationen geringer Protonenabstände umgekehrt dazu, dass die Gesamtrajektorie den NOE 8 effektiv erfüllt.

Nach Abbildung 7.10 misst NOE 33 den Abstand des Amid-Wasserstoffs HN des Asp 6 vom HN des Gly 7. Dieser von Wüthrich [39] als d_{NN} bezeichnete Abstand hängt sowohl von ϕ als auch von ψ am Asp 6 ab. In jenem erlaubten Bereich des Ramachandran-Plots, der den α -helikalen Strukturen entspricht, ist d_{NN} stets kürzer als $3,0 \text{ \AA}$, während in β -Strängen Abstände zwischen $2,0$ und $4,8 \text{ \AA}$ auftreten können. Da im *cis*-cAPB am Asp 6 Fluktuationen zwischen diesen beiden Strukturen auftreten, zeigt Abbildung 7.11 auch für NOE 33 eine bimodale Verteilung. Drei der sechs Konformationen (8, 6 und 12) erfüllen die Abstandsbedingung, während die drei anderen (1, 11 und 9) diesen NOE verletzen. Auch hier entspricht diese Unterteilung der Konformationen in zwei Gruppen genau der Klassifikation des Residuums Asp 6 als entweder helikal oder faltblattartig.

Auch der NOE 20 gehört zu einem Abstand des Typs d_{NN} , welcher durch die Dihedralwinkel des Residuums Thr 4 bestimmt wird. Tabelle C.1 weist hierfür fast ausschließlich gestreckte Geometrien aus. Nur Konformation 6 zeigt am Thr 4 ein deutlich abweichendes ψ . Blickt man zurück auf Abbildung 7.9, so wurde für diese Konformation im Bereich der Residuen 3 bis 4 ein stark zurückgefalteter β -Turn beobachtet. Nur diese Konfiguration des Peptidrückgrats führt zu Abständen der beiden HNs, die unterhalb von $3,0 \text{ \AA}$ liegen, wie man Tabelle D.1 entnehmen kann. Abbildung 7.11 bestätigt dieses Ergebnis und zeigt außerdem, dass Übergänge zwischen der Konformation 6 (rot) und den anderen Konformation sehr selten sind. So wird dieser Zustand überhaupt nur innerhalb der ersten 3 ns der Trajektorie angenommen.

Alle drei der zuletzt diskutierten NOEs waren lokaler Natur und die durch sie bestimmten Protonenabstände wurden jeweils ausschließlich durch einen oder zwei der Dihedralwin-

kel des Peptidrückgrats bestimmt. Aus diesem Grund ist davon auszugehen, dass diese NOEs kaum zur Festlegung der großräumigen dreidimensionalen Struktur der Modellpeptide beitragen. Von großer Bedeutung für das NMR-Strukturrefinement und auch für die Bewertung der MD-Konformationen sind deshalb NOEs zwischen Wasserstoffatomen, die in der chemischen Struktur der Peptidkette weit voneinander entfernt liegen. Beispielfür hierfür ist NOE 11. Nach Abbildung 7.10 wird der zugehörige Abstand quer durch das Molekül gemessen. Dabei beeinflusst die Orientierung der Seitengruppe des Ala 3 in Bezug auf den Azobenzolfarbstoff diesen Abstand in starker Weise.

Tabelle D.1 weist für NOE 11 einen experimentellen Maximalabstand von etwa 8,0 Å aus, während der gewichtete Mittelwert der Trajektorie mit 8,8 Å etwas größer ist. Aus Abbildung 7.11 wird deutlich, dass auch hier nur während kurzer Teilabschnitte der Trajektorie kleine Abstände auftreten. Beispielsweise nähern sich die beiden Wasserstoffe in der gelb eingezeichneten Konformation 12, welche schon in der Diskussion der Abbildung 7.9 als sehr kompakt beschrieben wurde, teilweise bis auf 4,0 Å aneinander an. Auch eine der Sub-Konformationen des Zustands 1 weist teilweise geringe Abstände des Ala 3 vom rechten Teil des Farbstoffs auf, wie das Histogramm in Abbildung 7.11 nahe legt.

7.4.2 Die Erfüllung der NOEs durch die Konformationen des *cis*-cAPBs

Die obige Diskussion der fünf ausgewählten NOEs hat gezeigt, dass jede dieser Abstandsbedingungen von jeweils mindestens einer der sechs betrachteten Konformationen des *cis*-cAPBs erfüllt wird. Doch gilt diese Beobachtung für alle 41 NOEs? Von Interesse ist hier ferner die Frage, welche dieser Abstandsbedingungen von welchen Konformationen erfüllt bzw. verletzt werden.

Um einen Überblick zu erhalten sind die in Tabelle D.1 aufgeführten konformationslokalen Abweichungen in Abbildung 7.12 grafisch dargestellt. Dort entsprechen die Zeilen den verschiedenen NOEs und die Spalten den 12 Konformationen aus Abbildung 7.9. Die in der Matrix verwendeten Graustufen repräsentieren auf einer logarithmischen Skala die festgestellten Abweichungen der MD-Konformationen von den experimentellen NOEs. Zur Erleichterung der Diskussion wurde in Abbildung 7.12 zusätzlich eine Tabelle eingefügt, die für jede Konformation die Gesamtzahl der aufgetretenen Verletzungen, die mittlere Abweichung in Ångström und die freie Energie ΔG aufführt.

Ich möchte Abbildung 7.12 zuerst entlang der unterschiedlichen NOEs diskutieren: eine erste Betrachtung zeigt, dass 14 der 41 NOEs ausnahmslos von allen Konformationen erfüllt werden. Die betroffenen NOEs wurden stets zwischen Protonen zweier aufeinanderfolgender Residuen beobachtet, die nur durch wenige chemische Bindungen voneinander getrennt sind. Wie oben am Beispiel des NOE 2 diskutiert wurde, bestimmt in diesen Fällen die lokale Geometrie des Moleküls vollständig die Protonenabstände. Man erkennt deshalb an der Erfüllung dieser NOEs allenfalls die Qualität der Parametrisierung der Bin-

7 Die Modellpeptide im Gleichgewicht

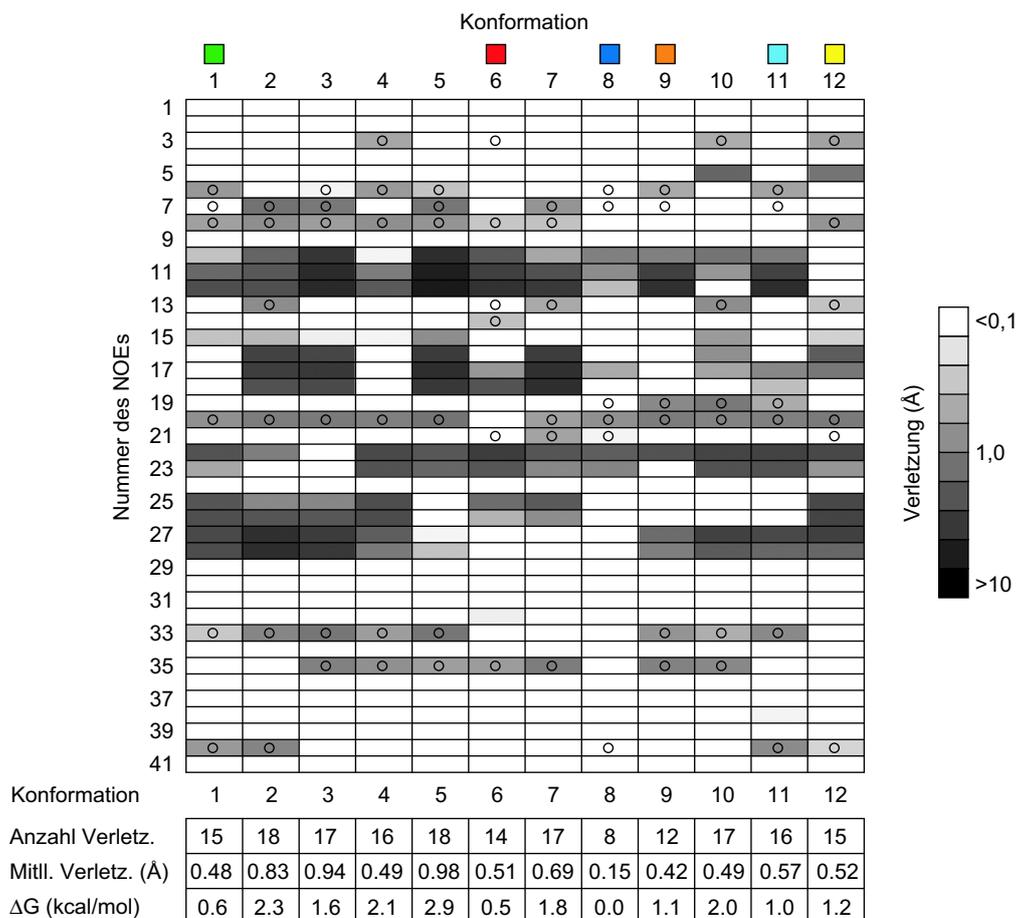


Abbildung 7.12: Abweichung der *cis*-cAPB Konformationen von den NOEs. Die Abweichungen der konformationslokal nach(7.3) gemittelten und in Tabelle D.1 aufgeführten Protonenabstände von den experimentellen Abstandsbedingungen werden auf einer logarithmischen Skala in Graustufen kodiert. Die Farbfelder über den Spalten markieren die in den Abbildungen 7.9 und 7.11 diskutierten Konformationen. Kleine Kreise zeigen auf NOEs der Form $d_{\alpha N}$ oder d_{NN} diejenigen Konformationen an, bei denen durch eine einfache Betrachtung der Geometrie des entsprechenden Residuums eine Verletzung erwartet werden kann. Erläuterung hierzu im Text.

dungslängen und -winkel im verwendeten Kraftfeld. Deshalb ist zu erwarten, dass auch das im NMR-Refinement verwendete Kraftfeld diese NOEs von sich aus erfüllt und sie somit keinen Einfluß auf die experimentell ermittelten Strukturen haben.

Abbildung 7.12 zeigt weiterhin, dass eine andere Gruppe von NOEs jeweils von etwa der Hälfte der Konformationen erfüllt bzw. verletzt wird. Zu dieser Gruppe von Abstandsbedingungen gehören vor allem solche des Typs $d_{\alpha N}$ oder d_{NN} , bei denen der Protonenabstand direkt von der Winkel-Konfiguration des entsprechenden Residuums abhängt. Anhand der Beispiele der NOEs 8, 33 und 20 wurde oben diskutiert, dass – vereinfacht betrachtet – eine im Ramachandran-Plot (Abbildung 1.4) als α -helikal bezeichnete Konfiguration eines Residuums stets mit einem größeren Abstand $d_{\alpha N}$ verbunden ist, als eine gestreckte faltblattartige Geometrie, während für d_{NN} genau die Umkehrung gilt. Damit ist es leicht möglich, basierend auf der in Anhang C aufgeführten Zuordnung der einzelnen Residuen zu den beiden möglichen Geometrien (α oder β), die bei den verschiedenen Konformationen des *cis*-cAPBs auftretenden Verletzungen der NOEs vorherzusagen. Um dies zu verdeutlichen, wurden in Abbildung 7.12 diejenigen NOE/Konformations-Kombinationen mit einem kleinen Kreis versehen, bei denen eine Verletzung zu erwarten ist: bei den NOEs des Typs $d_{\alpha N}$ (Nummern 3, 6, 8, 14, 21 und 35) wurden diejenigen Konformationen markiert, bei denen die betroffenen Residuen im Gegensatz zum Wert des NOEs als helikal eingestuft wurden, während bei den NOEs des Typs d_{NN} (7, 13, 19, 20, 33, 40) diejenigen Zustände hervorgehoben sind, bei denen das entsprechende Residuum faltblattartig erscheint. Eine Betrachtung der Markierungen zeigt, dass alle zugehörigen Abstandsüberschreitungen von dieser einfachen Betrachtung erklärt werden.

Abzüglich der Menge der trivial erfüllten NOEs und der soeben diskutierten kurzreichweitigen Abstandsbedingungen, die nur von einem Teil der Konformationen erfüllt werden, bleiben von den 41 am *cis*-cAPB beobachteten NOEs noch 14 übrig. Davon ist die Mehrzahl wirklich langreichweitiger Natur, d.h. es wird eine Abstandsbedingung zwischen Wasserstoffatomen angegeben, deren Residuen in der Peptidkette mindestens durch eine weitere Aminosäure voneinander getrennt sind. Mit dem NOE 11 wurde oben schon einer dieser Abstände genauer untersucht und festgestellt, dass dieser genau von einer Konformation (Nummer 12) erfüllt wird, bei welcher das Ala 3 dem Farbstoff sehr nahe kommt. Da die NOEs 10 und 12 mit dem NOE 11 stark verwandt sind, gilt für sie die gleiche Aussage. Dagegen werden die NOEs 16 bis 18, welche die Distanz der Seitengruppe des Thr 4 vom Farbstoff betreffen, sogar von drei der Konformationen eingehalten (1, 4 und 9), bei denen diese Seitengruppe in Richtung der Mitte des zyklischen Moleküls gedreht ist. Die Abstandsbedingungen 25 bis 28 werden schließlich von der zentralen Konformation 8 erfüllt, da dort anscheinend durch den in der Umgebung des Asp 6 beobachteten Loop dieses Residuum näher an den Farbstoff und an das Phe 8 gelangen kann.

Zusammenfassend zeigt diese Betrachtung von Abbildung 7.12, dass ausnahmslos jede Abstandsbedingung von mindestens einer der 12 Konformationen erfüllt wird. Für die über die Gesamttrajektorie gemittelten Protonenabstände tritt in Tabelle D.1 jedoch trotzdem nicht immer eine vollständige Übereinstimmung mit den experimentellen NOEs auf.

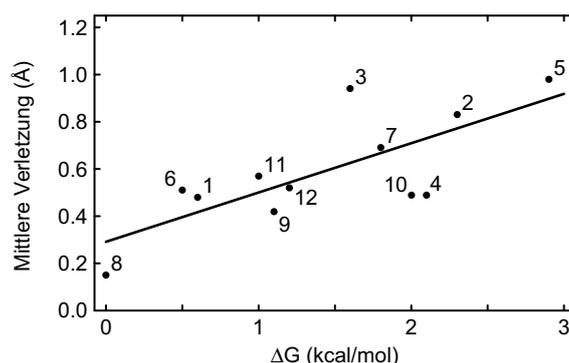


Abbildung 7.13: Das Ansteigen der mittleren NOE-Verletzungen mit ΔG . Für jede der durchnummerierten Konformationen des *cis*-cAPBs wird das zuvor in Abbildung 7.12 tabellierte arithmetische Mittel der NOE-Verletzungen über die freie Energie ΔG aufgetragen. Die durchgezogene Linie zeigt einen linearen Fit an diese Daten nach der Methode der kleinsten Quadrate.

Dies kann beispielsweise daran liegen, dass die relativen Besetzungen der unterschiedlichen Konformationen in der relativ kurzen MD-Trajektorie noch nicht ausreichend gut bestimmt werden oder dass sich die Peptidstruktur bei der Simulationstemperatur von 500 K leicht von jener bei der experimentellen Temperatur von etwa 300 K unterscheidet. Außerdem führt die erhöhte Simulationstemperatur dazu, dass Konformationen höherer freier Energie ΔG verstärkt angenommen werden. Nimmt man an, dass die Konformationen mit steigendem ΔG einen zunehmend „entfalteteren“ Charakter aufweisen, so erwartet man eine damit verbundene Zunahme der Anzahl der NOE-Verletzungen oder alternativ ein Ansteigen des arithmetischen Mittels der konformationslokalen Abweichungen.

Um diese Hypothese zu prüfen, wurden in Abbildung 7.13 die mittleren NOE-Verletzungen der Konformationen über ΔG aufgetragen. Die vermutete Korrelation ist deutlich zu erkennen. Die gezeigte Gerade stellt einen Fit an die Datenpunkte dar und weist eine Steigung von etwa $0,2 \text{ (Å} \cdot \text{mol)/kcal}$ auf. Damit wird die Annahme bestätigt, dass Konformationen höherer freier Energie weiter ausgedehnt, weniger kompakt oder stärker entfaltet sind. Demzufolge ist davon auszugehen, dass eine MD-Simulation bei Raumtemperatur verglichen mit den hier diskutierten Simulationen bei 500 K zu tendenziell kleineren Protonenabständen und damit zu einer noch besseren Erfüllung der experimentellen NOEs führen würde.

7.4.3 Die Protonenabstände der anderen Modellpeptide

Im Vorangehenden wurden am Beispiel des *cis*-cAPBs die bei den MD-Simulationen auftretenden konformationslokalen Protonenabstände ausführlich und detailliert mit den NOE-Abstandsbedingungen der NMR-Untersuchungen verglichen. Für das *trans*-cAPB und die übrigen beiden Modellpeptide cAMPB und bcAMPB in ihren beiden Isomeren

werden die entsprechenden Daten gleichfalls in Anhang D aufgeführt, sie sollen hier jedoch nur kurz diskutiert werden.

Da für das *trans*-cAPB bei der gewählten Auflösungsstufe nur eine einzige Konformation gefunden wurde, müssen in diesem Fall nur die globalen Mittelwerte der Protonenabstände betrachtet werden. Die Tabelle D.2 zeigt, dass ein Großteil der NOEs erfüllt wird. Neben einigen kleinen und insignifikanten Abweichungen fallen jedoch die langreichweitigen NOEs 15, 16, 24 und 25 auf, bei denen die MD-Simulationen deutlich größere Protonenabstände zeigen. Interessanterweise sind die an diesen NOEs beteiligten Residuen genau jene beiden, deren Seitengruppen chemisch labile Protonen aufweisen: das HG1 der OH-Gruppe des Thr 4 und das HD2 der Carboxylgruppe des ASP 6. Diese Protonen können mit Protonen aus der Umgebung des Peptids austauschen. Zwar stehen hierfür in reinem DMSO keine Protonen des Lösungsmittels zur Verfügung, jedoch ist davon auszugehen, dass sich das in der Realität stets im DMSO gelöste Wasser in der Umgebung der hydrophilen Seitengruppen ansammelt. Ein solcher Austausch kann zur Verstärkung der betroffenen NOEs führen. Die MD-Simulationen weisen also darauf hin, dass die erwähnten Abstandsbedingungen kritisch zu hinterfragen sind.

Geht man zum *cis*-cAMPB über, so zeigt sich, dass dort 17 der insgesamt 34 Abstandsbedingungen trivial von allen Konformationen erfüllt werden. Wie schon zuvor beim *cis*-cAPB, liegt bei den übrigen NOEs entweder schon der über die Trajektorie gemittelte Protonenabstand unter der gegebenen Obergrenze oder es findet sich mindestens eine Konformation, bei der ein entsprechend kleiner Abstand auftritt. Die einzige Ausnahme hierbei bilden der NOE 12 und in geringerem Maße auch der diesem stark verwandte NOE 11. Jedoch wurde hier nur ein sehr schwacher NOE beobachtet, und diesem ein sehr großer Protonenabstand von 4,8 Å zugeordnet, so dass die absoluten Abweichungen nicht signifikant zu sein scheinen.

Für das *trans*-Isomer des cAMPBs gelten ähnliche Aussagen, wie sie oben für das *trans*-cAPB getroffen wurden. Auch hier treten einige deutliche Verletzungen auf, wobei wie zuvor auffällig oft das Residuum Thr 4 beteiligt ist. Bei einem Teil dieser Abweichungen, beispielsweise bezüglich der NOEs 15 und 16, lässt eine Betrachtung der zugehörigen Zeitreihen der Protonenabstände (hier nicht gezeigt) vermuten, dass bei der Analyse längerer MD-Trajektorien eine Übereinstimmung erreicht werden könnte. Dies gilt jedoch nicht für den NOE 23, bei dem sich die betroffenen Wasserstoffatome in der Trajektorie nicht auch nur annähernd nahe kommen, weshalb dessen Gültigkeit in Frage zu stellen ist. Auch zeigt eine Betrachtung der im NMR-Refinement verwendeten Eingabedatei der NOE-Constraints, dass diese Abstandsbedingungen teilweise nicht wie zu erwarten zwischen Wasserstoffatomen, sondern auch zwischen Kohlenstoffatomen angegeben wurden. Bei der hier vorgestellten Auswertung wurde die Position dieser Kohlenstoffatome jeweils durch die an sie gebundenen Wasserstoffe ersetzt.

Schließlich ist der Vergleich zwischen den MD-Trajektorien und den NMR-Constraints noch für das bazyklische bcAMPB zu diskutieren. Hierzu wurden im Verlauf der Erstellung der vorliegenden Arbeit zuerst diejenigen NOEs herangezogen, die bei einer ersten

Messung an den beiden Isomeren des bcAMPBs beobachtet wurden und die der Strukturaufklärung der veröffentlichten NMR-Strukturen [38] zu Grunde lagen. Dieser erste Vergleich offenbarte eine Vielzahl von deutlichen Abweichungen zwischen den MD-Trajektorien und den NOE-Constraints (Daten nicht gezeigt), was einer der Gründe dafür war, diese Messungen zu wiederholen. Die Tabellen D.5 und D.6 in Anhang D zeigen die NOEs aus der erneuten Messung. Für das *cis*-bcAMPB wurden bei der zweiten Messung nur 28 NOEs beobachtet, die, bis auf eine Ausnahme, von jeweils mindestens einer MD-Konformation erfüllt werden. Einzig beim NOE 8 treten stets geringfügige Abweichungen von etwa 0,2 Å auf, wobei hierbei ein weiteres Mal das Thr 4 beteiligt ist.

Wie schon beim *trans*-cAPB sind auch beim *trans*-bcAMPB auffallend mehr Verletzungen zu beobachten, als bei den dazugehörigen *cis*-Isomeren. Signifikant größere Protonenabstände erkennt man im Fall des *trans*-bcAMPBs beispielsweise bei den NOEs 3, 13 und 16, die alle Abstände des Typs $d_{\alpha N}$ mit Beteiligung eines Alanin-Residuums betreffen. Bei den meisten der restlichen Verletzungen weist zumindest eine der drei Konformationen einen näherungsweise passenden Abstand auf, so dass auch hier eine Simulation bei Raumtemperatur oder eine längere Trajektorie zu passenderen Ergebnissen führen könnte. Zuletzt ist noch der NOE Nummer 23 von Interesse, bei dem experimentell ein Abstand von 4 Å festgestellt wurde. In der MD-Trajektorie liegen das HA des Cysteins 5 und das HN des Glyzins jedoch meist mindestens 6 Å voneinander entfernt. Nur wenn kurzzeitig das Peptidplättchen zwischen dem Asp 6 und dem Gly 7 einen Flip (siehe Abschnitt 7.1) in eine nach dem CHARMM22-Kraftfeld als ungünstig angesehene Ausrichtung ausführt, kann sich dieser Abstand verkleinern. Dies tritt während der Simulation aber nur selten und immer nur kurzzeitig auf, so dass der mittlere Protonenabstand deutlich zu groß ist.

7.4.4 Vergleich der ^3J -Kopplungskonstanten

Neben den oben diskutierten NOEs wurden bei den NMR-Messungen an den Modellpeptiden auch jeweils einige der durch chemische Bindungen vermittelten Spin-Spin-Kopplungen vermessen (siehe Abschnitt 1.2). Im Fall zweier, durch drei intermediäre chemische Bindungen getrennten Protonen hängt die ^3J -Kopplungskonstante von dem durch die Bindungen gebildeten Dihedralwinkel θ ab. Die funktionale Form dieser Abhängigkeit, die Karplus-Beziehung, wurde schon in der Einführung vorgestellt (1.1)

$$^3J = A \cos^2 \theta + B \cos \theta + C. \quad (7.4)$$

In Anhang D werden die experimentell ermittelten Kopplungskonstanten $^3J_{exp}$ den aus den MD-Trajektorien berechneten Werten $^3J_{MD}$ gegenübergestellt. Um $^3J_{MD}$ zu bestimmen, wurden die Winkel θ aus den Gleichgewichtstrajektorien extrahiert, nach (7.4) die Zeitreihen der Kopplungen berechnet und hieraus der arithmetische Mittelwert gebildet.

Da nach Aussage von Christian Renner die ^3J -Kopplungskonstanten, verglichen mit den NOEs, keinen übermäßigen Einfluß auf das Ergebnis des NMR-Refinements hatten, möchte ich nicht im Detail auf die tabellierten Werte eingehen. Man erkennt bei vielen Winkeln

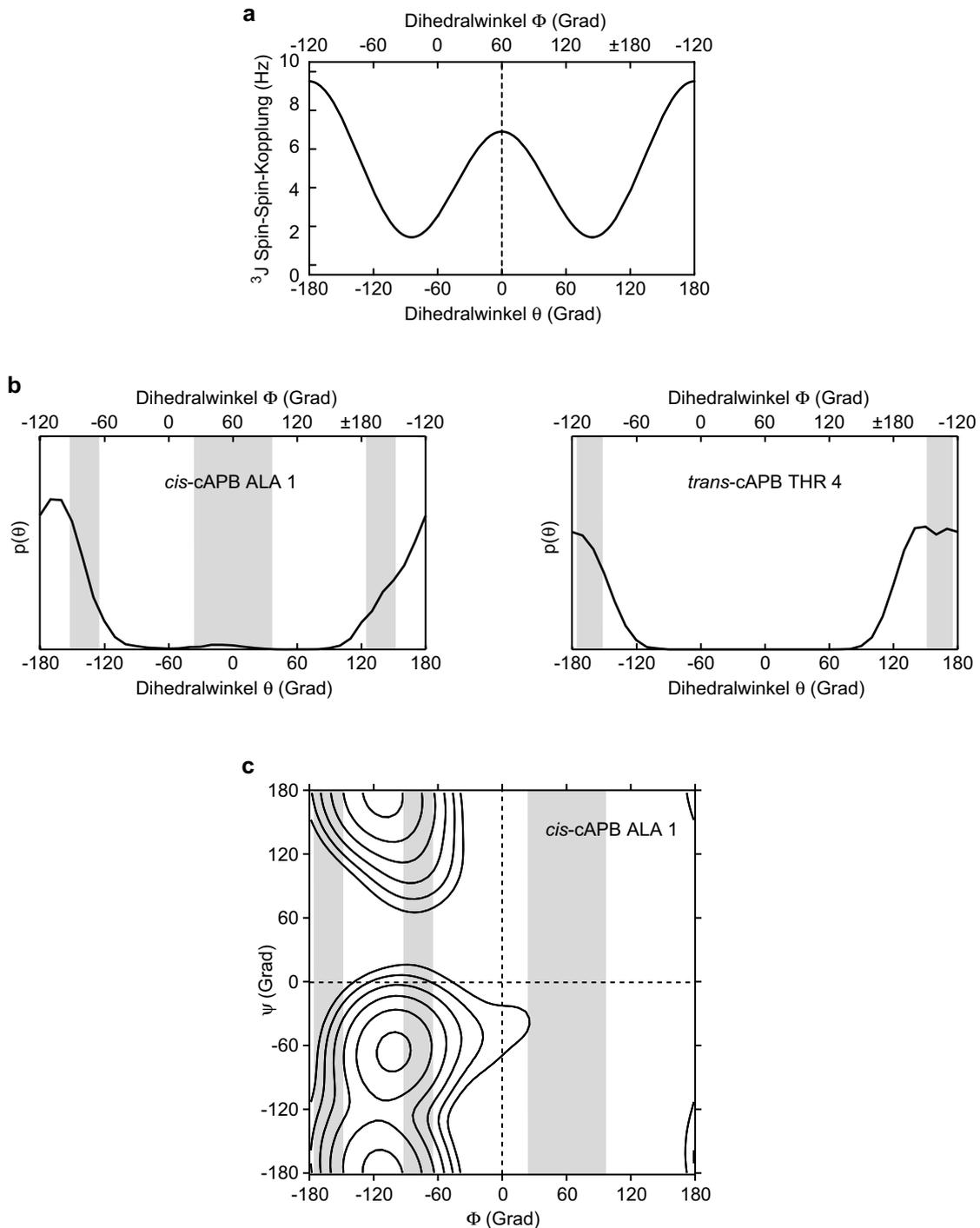


Abbildung 7.14: 3J -Kopplungskonstanten und die Dihedralwinkel. **a** Die Abhängigkeit (7.4) der 3J -Kopplungskonstanten von den Dihedralwinkeln θ bzw. $\phi = |\theta + 60^\circ|$. **b** Zwei Beispiele für die bei den Simulationen des cAPBs beobachteten Winkelverteilungen. Links: am Ala 1 des *cis*-cAPBs, rechts: am Thr 4 des *trans*-cAPBs. Die im NMR-Refinement auf Grundlage der Kopplungskonstanten als erlaubt angesehenen Winkelbereiche sind grau hinterlegt. **c** Im Ramachandran-Plot wird, ergänzend zum ersten Diagramm in Teilbild **b**, die ϕ/ψ -Winkelverteilung am Ala 1 des *cis*-cAPBs gezeigt.

eine zufriedenstellende Übereinstimmung, d.h. dort unterscheiden sich die experimentellen und theoretischen Kopplungskonstanten um weniger als 1 Hz voneinander. Zwei Trends fallen jedoch auf: bei vielen der Alanin-Residuen findet sich ein deutlich größerer Wert von ${}^3J_{MD}$ als ${}^3J_{exp}$, während umgekehrt bei einigen der Threoninen eine Abweichung zu geringeren Kopplungen zu beobachten ist.

Um diese Beobachtung zu interpretieren, ist eine Betrachtung der geometrischen Zusammenhänge hilfreich, die Abbildung 7.14 ermöglicht. Die Tabellen in Anhang D machen deutlich, dass experimentell fast ausschließlich Kopplungskonstanten an Dihedralwinkeln θ des Typs $\text{HN-N-C}_\alpha\text{-H}_\alpha$ gemessen wurden. Ein Rückblick auf die Struktur der Aminosäureresiduen (Abbildung 1.2) zeigt, dass unter Annahme eines perfekt tetraedisch koordinierten C_α , der zuvor schon mehrfach diskutierte Dihedralwinkel ϕ des Peptidrückgrats nach $\phi = |\theta + 60^\circ|$ auf einfache Weise mit dem Winkel θ zusammenhängt. Aus diesem Grund wird in Abbildung 7.14 a die Karplus-Beziehung (7.4) gleichermaßen als Funktion von θ und von ϕ skizziert. Die Parameter A , B , C wurden dabei in Übereinstimmung mit dem NMR-Refinement [37] zu $A = 6,7 \text{ Hz}$, $B = -1,3 \text{ Hz}$ und $C = 1,5 \text{ Hz}$ gewählt.

Abbildung 7.14 a zeigt, dass die Kopplungskonstante über einen Bereich von etwa 1 bis 9 Hz variieren kann und dass – aufgrund der quadratischen Abhängigkeit dieser Größe vom Kosinus des Dihedralwinkels – einer gemessenen Kopplung zwei bis vier erlaubte Winkelbereiche zugeordnet werden können. Diese erlaubten Bereiche, deren Breite sich jeweils aus dem Messfehler der Kopplungskonstanten ergibt, werden in Abbildung 7.14 b am Beispiel von zwei Residuen des *cis*- bzw. *trans*-cAPBs durch grau hinterlegte Flächen angedeutet, während die dort gezeigten Graphen die jeweils bei den Simulationen beobachteten Verteilungen der Dihedralwinkel verdeutlichen.

Die Winkelverteilung am Ala 1 des *cis*-cAPBs weist ihr Maximum etwas links (oder gedreht entgegen dem mathematischen Drehsinn) von einem der erlaubten Winkelbereiche auf. Da die Karplus-Beziehung in diesem Bereich eine negative Ableitung aufweist (vgl. Abbildung 7.14 a), erklärt diese Beobachtung den erhöhten Wert der Kopplungskonstanten. Dieser Zusammenhang ist interessant, da schon zuvor, beispielsweise von Hu et al. [289], festgestellt wurde, dass bei Verwendung des CHARMM22-Kraftfeldes im Vergleich zu anderen Kraftfeldern, wie GROMOS oder OPLS, die Verteilung der Dihedralwinkel des Alanin-Dipeptids sich als nach links verschoben präsentiert. Dieser Eindruck wird durch die hier gemachten Beobachtungen verstärkt, wie Abbildung 7.14 c zeigt.

Schließlich erkennt man in Abbildung 7.14 b eine ähnliche Verschiebung auch am Thr 4 des *trans*-cAPBs, die hier sicherlich einen Teil zu den etwas geringen Kopplungskonstanten ${}^3J_{MD}$ beiträgt. Relevanter scheint jedoch die durch die erhöhte Temperatur in der Simulation ausgelöste Verbreiterung der Winkelverteilung zu sein, durch die Winkelbereiche erreicht werden, die nach Abbildung 7.14 a mit sehr kleinen Werten für die Kopplung verbunden sind. Dieser Effekt ist zu erwarten und stellt damit keine systematische Abweichung zwischen den Messungen und den Simulationen dar.

7.5 Diskussion

Die in diesem Kapitel vorgestellten Untersuchungen haben bewiesen, dass es möglich ist, mit dem Stand der Technik entsprechenden molekularmechanischen MD-Simulationen die Gleichgewichtskonformationen der zyklischen Modellpeptide quantitativ korrekt vorherzusagen. Dies zeigt sich vor allem daran, dass die mit den Simulationen berechneten Konformationsensembles die aus den NMR-Messungen abgeleiteten NOE-Abstandsbedingungen bis auf einige wenige Ausnahmen erfüllen.

Jedoch ist hervorzugeben, dass die hier festgestellte Vorhersagefähigkeit der MD-Simulationen und die damit verbundene Validierung des eingesetzten MM-Kraftfeldes, eventuell nicht ohne weiteres auf die Untersuchung anderer Biomoleküle, wie beispielsweise nicht-zyklischer Peptide oder ganzer Proteine, übertragen werden darf, denn diese Systeme weisen verglichen mit den hier untersuchten zyklischen Peptiden weitaus komplexere Konformationsräume auf. Solche komplexen Konformationsräume umfassen üblicherweise eine Vielzahl von entfalteten Strukturen und einige wenige native bzw. gefaltete Konformationen, wobei letztere zumeist nur um wenige kcal/mol energetisch gegenüber den ungefalteten Zuständen begünstigt sind. Während bei der Simulation eines solchen, in seinem Konformationen unbeschränkten Moleküls das verwendete Kraftfeld in der Lage sein muss, die erwähnten sehr kleinen Energieunterschiede akkurat zu reproduzieren, ist eine derartige Genauigkeit zur erfolgreichen Beschreibung der zyklischen Peptide nicht notwendig, denn diese durch den Ringschluss konformell stark eingeschränkten Moleküle können sich nicht entfalten oder konformell denaturieren. So haben die Gleichgewichtssimulationen gezeigt, dass die Modellpeptide selbst bei einer erhöhten Temperatur von 500 K jeweils nur ein vergleichsweise überschaubares Ensemble von möglichen Konformationen aufweisen. Welche Konformationen einem der Modellpeptide zugänglich sind und welche nicht, scheint dabei hauptsächlich durch den Ringschluss, die vom Farbstoff auf das Peptid ausgeübten Zwänge und durch sterische Hinderungen im Peptidteil selbst bestimmt zu werden. Diese Zwänge werden in erster Linie durch die im Kraftfeld verwendeten Bindungslängen beeinflusst und weniger durch schwieriger festzulegende Parameter wie den Kraftkonstanten oder den Partialladungen. Aus diesem Grund ist zu erwarten, dass auch andere Kraftfelder zu sehr ähnlichen Gleichgewichtskonformationen der Modellpeptide führen würden.

Das vorliegende Kapitel diene jedoch nicht nur dazu, nachzuweisen, dass die Gleichgewichtsensembles der Peptide mit den verwendeten MD-Methoden quantitativ richtig beschrieben werden können. Vielmehr wurde hier auch die Leistungsfähigkeit der in Kapitel 5 vorgestellten Methoden zur datengetriebenen Modellbildung und unüberwachten Klassifikation von MD-Trajektorien untersucht. Ein erstes Ergebnis ist hier die Feststellung, dass mit dem „state of the art“ der Auswertung von MD-Simulationen, d.h. mit einer einfachen Hauptkomponentenanalyse, keine besonders tiefe Einsicht in die Topologie des Konformationsraums der Modellpeptide gewonnen werden kann. Weiterhin konnte jedoch die Ausgereiftheit des um eine Behandlung von Winkelräumen erweiterten UNIVAR-Verfah-

rens zur Schätzung einer analytischen Dichtemodells und des darauf aufbauenden ASSO-Verfahren zur unüberwachten Identifikation von Molekülkonformationen aufgezeigt werden.

Basierend auf den durch das ASSO-Verfahren identifizierten Konformationen der Modellpeptide war es anschließend möglich, die Simulationsdaten in Bezug auf die NOE-Abstandsbedingungen konformationslokal auszuwerten. Dabei zeigte sich besonders am Beispiel der Strukturen des *cis*-CAPBs, dass nicht alle NOEs von jeder Konformation erfüllt wird, zu den meisten NOEs jedoch mindestens eine Konformation existiert, die mit dieser Abstandsbedingung konform ist. Dies erklärt sich aus der Tatsache, dass bei der Aufnahme des NMR-Spektrums bekannterweise ein Zeit- und Scharmittel gebildet wird. Im Gegensatz zu den hier vorgestellten virtuellen „Einzelmolekülexperimenten“ wird dabei über lange Zeiten (etwa 0,01 s) und über alle etwa 10^{17} in der Probe vorhandenen Moleküle gemittelt. Treten, wie hier, während der Messzeit Konformationsübergänge in den untersuchten Molekülen auf, bzw. liegen diese äquivalent als Konformerengemisch vor, so stellen die gemessenen NOEs gleichermassen Mittelwerte der einzelnen Spin-Spin-Wechselwirkungen dar. In einem einfachen Modell geschieht dies analog zu der mit r^{-6} gewichteten Mittelung (7.3), so dass das Signal von Konformationen, die an einem Paar von Protonen einen kleinen Abstand aufweisen, das Signal der anderen Konformationen „überstrahlt“. Es werden damit in der Messung somit gleichzeitig alle *möglichen* Kontakte beobachtet, jedoch gilt nicht, dass diese *notwendigerweise* jeweils von jeder Konformation erfüllt werden.

Problematisch ist in diesem Zusammenhang das allgemein übliche und auch bei der Strukturaufklärung der Modellpeptide [37] gewählte Vorgehen des NMR-Strukturrefinements (siehe Abschnitt 1.2), bei dem stets alle NOEs in der Form von Zusatzpotentialen berücksichtigt werden. Auf diese Weise wird eine strikte und gleichzeitige Einhaltung aller Abstandsbedingungen erzwungen. Damit erklärt sich nicht nur die in diesem Kapitel festgestellte weitaus größere Kompaktheit der NMR-Strukturen im Vergleich zu den simulierten Konformationen, sondern auch die Verspannung der NMR-Strukturen, die sich beispielsweise in der unnatürlichen und willkürlichen Ausrichtung der Peptidplättchen bemerkbar macht. Es erscheint so, als ob die durch die vielen Abstandsbedingungen stark eingeschränkte Beweglichkeit der Peptide im Zusammenspiel mit einem zu schnellen Abkühlen im Annealing-Prozess zu einem regelrechten Einfrieren der Ausrichtung der Peptidplättchen geführt hat. Dies hat natürlich auch eine willkürliche Ausrichtung der mit den Peptidplättchen verbundenen Dipole zur Folge. Basierend auf den veröffentlichten NMR-Strukturen war die Ausrichtung dieser Dipole jedoch auch der Ausgangspunkt für eine Berechnung oder Abschätzung der Lage der Amidbanden der Modellpeptide Bredenbeck et al. [310, 311]. Im Hinblick auf die hier gezeigten Ergebnisse erscheint diese Berechnung der Amidbanden zumindest mit einem hohen Fehler verbunden zu sein.

8 Die lichtinduzierte Relaxationsdynamik der Modellpeptide

Im vorangehenden Abschnitt wurde gezeigt, dass die vorgestellten Langzeitsimulationen der Modellpeptide, trotz der bekannten [93, 94] Fehler und Unzulänglichkeiten der verwendeten MM-Kraftfelder, in der Lage sind, die Konformationen dieser Moleküle in Lösung verlässlich vorherzusagen. Denn es wurde eine gute quantitative Übereinstimmung der berechneten Strukturen mit den aus den NMR-Messungen ermittelten Abstandsbedingungen beobachtet. Ähnliche Validierungen der Methode der MD-Simulation in Hinsicht auf strukturelle Observablen wurden schon zuvor mehrfach durchgeführt und publiziert (siehe Abschnitt 1.3). Über solche strukturellen Fragestellungen hinausgehend werden MD-Simulationen jedoch häufig auch zur Aufklärung dynamischer Prozesse eingesetzt. Dieses Vorgehen stützt sich auf die Tatsache, dass aus MD-Simulationen – im Gegensatz beispielsweise zu Monte-Carlo-Simulationen – unmittelbar Aussagen über die in einem System auftretenden Dynamiken gewonnen werden können. In der Einführung (Kapitel 1) wurde berichtet, dass die meisten Ideen und Einsichten zur Konformationsdynamik von Proteinen auf kurzen Zeitskalen der Auswertung von MD-Trajektorien entstammen.

Jedoch war bisher keine Verifikation der simulierten Dynamiken durch einen Vergleich mit experimentellen Daten möglich, denn es fehlten geeignete Modellsysteme und ausreichend schnelle Messmethoden, um die Dynamik solcher Moleküle in dem den Simulationen zugänglichen Zeitbereich der Piko- bis Nanosekunden experimentell anzuregen und im Detail zu verfolgen. Erst durch die Synthese der lichtschtbaren Modellpeptide und die Anwendung der Femtosekunden-zeitaufgelösten Spektroskopie auf diese hinreichend einfachen Systeme konnte ein Überlapp zwischen den in Simulationen und den in Experimenten zugänglichen Zeitskalen erzielt werden.

Aufbauend auf der in Kapitel 3 und 4 vorgestellten Parametrisierung eines Kraftfeldes für die in DMSO gelösten Modellpeptide soll nun untersucht werden, ob und unter welchen Bedingungen MD-Simulationen in der Lage sind, die lichtinduzierte Relaxationsdynamik der Peptide quantitativ korrekt zu beschreiben. Ein erster Schritt zur Beantwortung dieser Fragen wurde im Rahmen der hier vorgestellten Arbeit durch die Auswertung einer kleinen Anzahl von Simulationsläufen gemacht, in denen die Lichtanregung des *cis*-cAPBs und die des freien *cis*-Azobenzols sowie die sich anschließende Relaxation der Moleküle hin in ihren jeweiligen *trans*-Zustand verfolgt wurde. Die dabei festgestellte, erstaunlich gute Übereinstimmung zwischen den in den Simulationen beobachteten Dynamiken und den aus der fs-Spektroskopie ermittelten Zeitkonstanten wurde anschließend veröffentlicht [138] und als Hinweis darauf interpretiert, dass MD-Simulationen prinzipiell dazu

geeignet sind, die Konformationsdynamik von Proteinen und Peptiden in Lösung korrekt vorherzusagen.

So erfreulich diese fast perfekte Übereinstimmung zuerst erschien, so unbefriedigend ist sie, wenn man bedenkt, dass die hier präsentierte Kombination von fs-Experimenten und MD-Simulationen eigentlich dazu auserkoren war, einen Test für die Beschreibung von Proteindynamik durch MD-Simulationen und insbesondere für die dabei eingesetzten Kraftfelder zu bilden. Wenn es, wie bei diesen ersten Simulationen, ausreichend ist, *ein* beliebiges Peptidkraftfeld (CHARMM) mit *einem* möglichen Lösungsmittelmodell und *einer* denkbaren Parametrisierung der Farbstoffe zu kombinieren und dabei *eine* vorgegebene Implementierung (EGO-MMII) dieses Gesamtkraftfeldes einzusetzen, um eine fast perfekte Übereinstimmung zu erreichen, so kann auf diese Weise kein Beitrag zur Verbesserung der Methodik oder der Kraftfelder geleistet werden. Aber dieser Eindruck täuscht, denn es beruht, wie sich zeigen wird, die bei den ersten Simulationen beobachtete gute Übereinstimmung mit den Experimenten tatsächlich gerade auf einer glücklichen Wahl der möglichen Kraftfeldparameter, auf deren Veränderung die simulierten Dynamiken empfindlich reagieren.

Natürlich ist es unmöglich, den Einfluss der Vielzahl der benötigten Kraftfeldparameter auf die Moleküldynamik vollständig und umfassend zu untersuchen. Schon zur Abschätzung der Auswirkungen einer Variation eines einzelnen Details des Kraftfeldes ist die Berechnung einer gewissen Anzahl (mindestens 10) von Nanosekunden-Trajektorien der Modellsysteme nach Lichtanregung notwendig. Über diese Einzeltrajektorien ist zu mitteln, um ein akzeptables Signal-zu-Rausch-Verhältnis für die auszuwertenden Observablen zu erreichen. Aus den experimentellen Voruntersuchungen ist bekannt, dass sich die Konformationsdynamik der zyklischen Peptide nach Umschalten des Azobenzols mindestens in den Zeitbereich einer Nanosekunde hinein erstreckt [139], so dass der mit jeder einzelnen Trajektorie verbundene Rechenaufwand recht erheblich ist. Deshalb kann in endlicher Zeit und mit den zur Verfügung stehenden Ressourcen nur exemplarisch der Einfluß einer sehr geringen Anzahl von Kraftfeldparametern auf die Peptiddynamik untersucht werden. Aus den gleichen Gründen möchte ich mich neben dem Azobenzol auf die alleinige Betrachtung des cAPBs beschränken und nur eine mögliche Richtung der Isomerisierungsreaktion untersuchen. Während mit dem in Kapitel 4 vorgestellten Zusatzpotential sowohl die *cis*→*trans* als auch die *trans*→*cis*-Isomerisierung modelliert werden kann, soll hier nur der erste Fall betrachtet werden, da dieser im Experiment besser charakterisiert werden konnte [139].

Mit diesen Einschränkungen kann der Einfluss der folgenden Kraftfelddetails auf die simulierten Dynamiken untersucht werden:

(i) Das Modell des Lösungsmittels DMSO

In Kapitel 3 wurden mehrere mögliche Modelle für das Lösungsmittel DMSO vorgestellt, darunter solche, bei denen die Wasserstoffatome der Methylgruppen explizit berücksichtigt werden, und andere, bei denen diese Gruppen jeweils ein Pseudoatom bilden. Ferner unterscheiden sich die DMSO-Modelle auch hinsichtlich ihrer Flexibilität und weiterer

Kraftfeldparameter. Aus der Vielzahl der bekannten Modelle wurden das neue und flexible Modell mit der Bezeichnung HC (bzw. sein Vorläufer VAK-DMSO), das rigide „united-atom“ Modell MP und das semi-flexible Modell MI ausgewählt, um den Einfluss der Wahl des Lösungsmittelmodells zu ermitteln.

(ii) Der Skalierungsfaktor λ_{14} der Elektrostatik

In Kapitel 6 wurde detailliert auf den Faktor λ_{14} eingegangen, mit dem im CHARMM-Kraftfeld die elektrostatische Wechselwirkung zweier durch drei Bindungen voneinander getrennter Atome skaliert wird (siehe (2.25)). Im Gegensatz zum CHARMM-Standard von $\lambda_{14} = 1,0$ basierten die zuerst durchgeführten und publizierten [138] Simulationen auf der Annahme des Wertes $\lambda_{14} = 0,4$, der gestreckte Peptidkonformationen begünstigt. Es ist zu untersuchen, welche Auswirkungen die mit diesem Parameter verbundenen Änderungen der Energielandschaft des Peptids (vgl. Abschnitt 6.2) auf dessen Konformationsdynamik haben.

(iii) Die Dihedralpotentiale in den Azobenzolfarbstoffen

Als drittes Detail des Kraftfeldes werden die beiden alternativen Parametrisierungen A und B der Dihedralpotentiale der Azobenzolfarbstoffe verglichen, die in Kapitel 4 (Tabelle 4.3) vorgestellt wurden. Zurückblickend auf Abbildung 4.3 sei daran erinnert, dass die Parametrisierung A einen steileren Verlauf der Potentiale in der Nähe der Minima zeigt als die Variante B. Dagegen gibt B die Höhen der Potentialbarrieren exakt wieder, während A sie unterschätzt.

Selbst mit der Beschränkung auf nur diese drei zu variierenden Aspekte des Kraftfeldes war es unmöglich, Simulationen für alle möglichen Parameterkombinationen durchzuführen, denn hierzu wären unter Berücksichtigung dreier DMSO-Modelle, zweier möglicher Werte für λ_{14} und zweier alternativer Dihedralpotentiale immerhin noch 12 Kombinationen zu evaluieren gewesen. Die Auswahl der tatsächlich untersuchten Kombinationen und die Details der Simulationen werden in Abschnitt 8.1 erläutert.

Um die Relaxation der Modellsysteme in den Simulationen verfolgen zu können, bedarf es noch geeigneter Observablen. In Abschnitt 8.2 werden zwei hierzu geeignete Größen besprochen: als erste skalare Observable wird aus den Trajektorien jeweils die Gesamtenergie der gelösten Moleküle berechnet. Durch die Anpassung eines Ratenmodells von multiexponentiellen Zerfällen an die so entstehenden Zeitreihen der Gesamtenergie können dann jeweils die Zeitkonstanten der Relaxationsdynamiken bestimmt und mit den aus der fs-Spektroskopie bekannten Werten verglichen werden. Daneben beinhalten die MD-Trajektorien in Form aller Atomkoordinaten jedoch noch eine Vielzahl weiterer Informationen über die stattfindenden Relaxationsprozesse. Dabei ist es offensichtlich von besonderem Interesse, einen Einblick in die im Peptidteil des cAPBs stattfindenden Konformationsübergänge zu gewinnen. Um die auftretende Konformationsdynamik quantitativ auswerten zu können, ist ein Maß dafür notwendig, wie weit sich die Geometrie eines einzelnen Moleküls zu einem gegebenen Zeitpunkt vom Raum der *cis*-Konformere entfernt und dem der *trans*-Konformere angenähert hat. Basierend auf den in Kapitel 5 erarbeiteten Methoden zur statistischen Datenanalyse wird in Abschnitt 8.2 ein solches Maß definiert.

| Bez. | Mol. | DMSO | λ_{14} | Dihe-Par. | Anz. | t_{500K} (ps) | t_{298K} (ps) | t_{isom} (ps) | CPU-Tage |
|------|------|------|----------------|-----------|------|-----------------|-----------------|------------------------|----------|
| R1 | Azo | VAK | 0,4 | A | 6 | 0 | 100 | 100 | 44 |
| R2 | Azo | HC | 1,0 | B | 10 | 0 | 500 | 1000 | 323 |
| R3 | Azo | MI | 1,0 | B | 10 | 0 | 500 | 1000 | 56 |
| R4 | Azo | MP | 1,0 | B | 10 | 0 | 500 | 1000 | 56 |
| R5 | cAPB | VAK | 0,4 | A | 6 | 0 | 0 | 1000 | 239 |
| R6 | cAPB | HC | 0,4 | A | 10 | 50 | 200 | 1000 | 312 |
| R7 | cAPB | HC | 1,0 | A | 10 | 50 | 200 | 1000 | 312 |
| R8 | cAPB | HC | 0,4 | B | 10 | 50 | 200 | 1000 | 312 |
| R9 | cAPB | HC | 1,0 | B | 10 | 50 | 200 | 1000 | 312 |
| R10 | cAPB | MI | 0,4 | B | 10 | 500 | 500 | 1000 | 63 |
| R11 | cAPB | MI | 1,0 | B | 10 | 500 | 500 | 1000 | 63 |
| R12 | cAPB | MP | 1,0 | B | 10 | 500 | 500 | 1000 | 63 |

Tabelle 8.1: Übersicht über die Relaxations-Simulationen. Es werden die 12 durchgeführten Simulationen der lichtinduzierten Relaxationsdynamik benannt und die jeweils gewählten Kraftfelddetails aufgeführt. Neben der Bezeichnung der Simulation werden das dabei untersuchte Molekül, entweder „Azo“ für das freie Azobenzol oder „cAPB“ für das erste Modellpeptid (siehe Abbildung 1.13), und das verwendete DMSO-Modell (siehe Kapitel 3) angegeben. Die beiden folgenden Spalten beinhalten den in der Simulation angenommenen Skalierungsfaktor λ_{14} der elektrostatischen 1-4-Wechselwirkung (siehe (2.25)) und den ausgewählten Satz *A* oder *B* der Dihedralpotentiale im Farbstoff (siehe Tabelle 4.3). Es folgt die Anzahl der Einzeltrajektorien, aus der sich die Simulation zusammensetzt, die Zeitdauer der zusätzlich zur Equilibrierung durchgeführten Gleichgewichtssimulation bei 500 K und 298 K und die Dauer t_{isom} der jeweils berechneten Trajektorien der Relaxation. In der letzten Spalte wird der Rechenaufwand der Simulationen abgeschätzt. Angegeben sind CPU-Tage auf Prozessoren des Typs AMD AthlonMP 1900+. In der Tabelle werden mit R1 und R5 diejenigen Simulationen bezeichnet, deren Ergebnisse schon zuvor veröffentlicht wurden [138].

Anschließend werden in Abschnitt 8.3 die Ergebnisse einer Vielzahl von Simulationen präsentiert und diskutiert. Ein Vergleich der einzelnen Simulationsergebnisse untereinander ermöglicht dabei eine Beantwortung der Fragen (i) bis (iii). Ferner werden exemplarisch einige Trajektorien der Relaxation des cAPBs nach Lichtanregung strukturell untersucht, um Einsichten in die ablaufende Konformationsdynamik zu gewinnen und damit gleichzeitig das Potential der in dieser Arbeit vorgestellten datengetriebenen Modellbildung zur Analyse von MD-Trajektorien aufzuzeigen.

8.1 Details der Simulationen

Tabelle 8.1 listet alle zur Untersuchung der Relaxationsdynamik durchgeführten Simulationen auf und zeigt neben den dabei vorgenommenen Variationen der oben diskutierten

Kraftfeldaspekte auch die Anzahl der jeweils berechneten Einzeltrajektorien. Hinsichtlich ihrer Durchführung unterscheiden sich die beiden älteren und zuvor veröffentlichten [138] Simulationen R1 und R5^a teilweise deutlich von den übrigen, neueren Simulationen, für welche die Details der Durchführung schon in Abschnitt 7.1 ausführlich dargestellt wurden und hier deshalb nicht wiederholt werden müssen.

Die Simulationen R1 und R5 entstanden unter Verwendung einer älteren Version des Programmpaketes EGO, der einige wichtige Funktionen fehlten. Aus diesem Grund konnte beispielsweise anstatt des mit den neueren Simulationen beschriebenen Rhombendodekaeders (etwa 960 DMSO-Moleküle) nur ein würfelförmiges Simulationsvolumen angesetzt werden, welches mit etwa 29 Å Kantenlänge mehr als 1.600 DMSO-Moleküle umfasst. Weitere Unterschiede zu den neueren Simulationen bestehen darin, dass die Coulomb-Summe nur bis zu einem Abstand von 11 Å (anstatt 13 Å) explizit berechnet wurde, dass die Quadrupolmomente weiter entfernter struktureller Einheiten keine Berücksichtigung fanden und auch darin, dass ein kleinerer Überlapp der SAMM-Distanzklassen von 1 Å verwendet wurde (siehe [100]).

Als Ausgangspunkte für die Berechnung von Relaxationstrajektorien nach Lichtanregung dienten equilibrierte Schnappschüsse der Simulationssysteme, die getrennt jeweils für alle ausgewählten Kombinationen von Kraftfelddetails erzeugt wurden. Solche Schnappschüsse könnten prinzipiell aus einer Gleichgewichtstrajektorie entnommen werden, wie sie in Kapitel 7 vorgestellt wurde. Da diese Langzeitsimulationen aufgrund des mit ihnen verbundenen hohen Aufwands jedoch nur für eine der Parameterkombinationen (HC-DMSO; $\lambda_{14} = 1,0$; Dihedralparametersatz B) berechnet wurden, musste ein anderer Weg gewählt werden, um equilibrierte Simulationssysteme zu erhalten und gleichzeitig einen zumindest eingeschränkten Überblick über die Konformationsräume der *cis*- und *trans*-Isomere zu erhalten. Es wurden, entsprechend der in Tabelle 8.1 gezeigten Anzahl von Einzeltrajektorien, die besten 6 oder 10 der veröffentlichten NMR-Strukturen des cAPBs [37] einzeln in Lösungsmittelboxen eingesetzt und anschließend die in Abschnitt 7.1 erläuterten Schritte der Equilibrierung durchgeführt. An den Schritt 6 dieses Verfahrens schlossen sich dann jeweils zwei weitere Gleichgewichtssimulationen im NPT-Ensemble an, deren Zeitdauern Tabelle 8.1 als t_{500K} bzw. t_{298K} angibt. Die erste Simulation bei 500 K diente einer vollständigen Relaxation der verspannten NMR-Strukturen, während die zweite bei 298 K eine Erforschung des Peptid-Konformationsraumes ermöglichte. Als Startpunkt der Relaxationssimulationen wurde anschließend jeweils die Systemkonfiguration am Ende der zweiten Gleichgewichtstrajektorie gewählt.

Abweichungen von dem beschriebenen Vorgehen ergaben sich erstens bei den Simulationen des freien Azobenzols, welches offensichtlich – abgesehen von der möglichen *cis-trans*-Isomerisierung – keine weitere konformelle Variabilität aufweist. Hier wurden die Schnappschüsse jeweils einer einzelnen Gleichgewichtstrajektorie von 500 ps Dauer entnommen. Ferner wichen bei der Simulation R5 die Equilibrierungen von dem in Ab-

^aZur besseren Unterscheidung werden im Folgenden die beiden älteren Simulationen R1 und R5 auch teilweise zusätzlich mit dem Kürzel „PNAS“ gekennzeichnet.

schnitt 7.1 aufgeführten Schema wie folgt ab: die Schritte 2, 4 und 5 wurden ausgelassen, dagegen betrug die Simulationsdauer 30 ps bei Schritt 3 und 50 ps bei Schritt 6. Da zu dem damaligen Zeitpunkt kein Barostat in EGO implementiert war, konnten die Simulationen R1 und R5 nur im NVT-Ensemble durchgeführt werden. Dabei entsprach die Dichte der Lösungsmittelboxen der experimentellen Dichte von reinem DMSO bei Normalbedingungen.

Details der Relaxationssimulationen

Ausgehend von den oben beschriebenen Startkonfigurationen wurden die *cis*-Isomere der Modellsysteme jeweils eine weitere Pikosekunde lang im thermischen Gleichgewicht verfolgt, um deren Ausgangszustand zu charakterisieren. Hierbei und während der gesamten Relaxationssimulation, wurde der Barostat abgeschaltet und somit in das NVT-Ensemble gewechselt. Alle anderen Details der Simulationen, wie die Auswertung der langreichweitigen Wechselwirkungen, die Zeitkonstanten des Thermostaten, etc. entsprachen denen der oben bzw. in Abschnitt 7.1 beschriebenen Gleichgewichtssimulationen. Insbesondere wurde auch hier der Thermostat ausschließlich an die Freiheitsgrade der Lösungsmittelmoleküle angebunden, um Artefakte zu vermeiden.

Die Modellierung der lichtinduzierten Isomerisierungsreaktion der Azobenzolfarbstoffe wurde in Abschnitt 4.2 geschildert. Dementsprechend wurde zum Zeitpunkt $t = 0$ der Relaxationssimulationen das Zusatzpotential (4.1) angeschaltet, um die Lichtanregung zu beschreiben. Die auf diese Weise instantan in den Farbstoff eingetragene Energie $\Delta E = 64,9$ kcal/mol entsprach derjenigen eines Photons der experimentell verwendeten Anregungswellenlänge (480 nm). Während der nächsten 10 ps der Simulation wurde die Geometrie der Azogruppe überwacht. Sobald der *trans*-Zustand erreicht und der Winkel N=N-C (die Reaktionskoordinate θ) nur noch um weniger als 20° aus seiner Gleichgewichtslage ausgelenkt war, wurde damit begonnen, über einen Zeitraum von 1 ps linear zwischen dem Zusatzpotential und den normalen Kraftfeldbeiträgen zu interpolieren. Anschließend war die Standardbeschreibung des *trans*-Isomers der Modellsysteme erreicht, deren Relaxationsdynamiken bis zu den in Tabelle 8.1 aufgeführten Zeitpunkten t_{som} verfolgt wurden. Während der Simulationen wurden die unterschiedlichen Beiträge zur MM-Gesamtenergie (2.13) aufgezeichnet und dabei zwischen dem gelösten Molekül und dem Lösungsmittel unterschieden. Auch die Geometrien des cAPBs bzw. des Azobenzols wurden abgespeichert. In beiden Fällen geschah dies in Zeitschritten von 5 fs ($t = -1$ ps ... 10 ps), 50 fs ($t = 10$ ps ... 100 ps) und 0,5 ps ($t = 100$ ps ... 1 ns). Die aus diesen Aufzeichnungen extrahierten Observablen werden in den nächsten Abschnitten vorgestellt.

An dieser Stelle sei noch auf den mit den vorgestellten Simulationen verbundenen erheblichen Rechenaufwand eingegangen, der in Tabelle 8.1 in Einheiten von CPU-Tagen aufgeführt wird. Die Abschätzung des Aufwands bezieht sich dabei auf Workstations, die mit Prozessoren des Typs AMD AthlonMP 1900+ bestückt sind, da ein Cluster solcher Doppelprozessor-Maschinen zur Berechnung der neueren Trajektorien eingesetzt wurde.

Der Gesamtaufwand betrug mehr als 2000 CPU-Tage. Besonders auffällig ist, dass die Verwendung eines „all-atom“ DMSO-Modells (VAK, HC) einen deutlich höheren Aufwand zur Folge hat, als die eines „united-atom“ Modells. Es wird sich jedoch zeigen, dass dieser Aufwand in Kauf genommen werden muss.

8.2 Energetische und geometrische Observablen

Bei einer MD-Simulation stehen die verschiedenen Einzelbeiträge zur Gesamtenergie des untersuchten Systems zumeist direkt in Form von Zeitreihen zur Verfügung und können dazu dienen, einen solchen Simulationslauf auszuwerten. Es liegt deshalb nahe, die licht-induzierte Relaxationsdynamik des Azobenzols und des cAPBs durch eine Betrachtung der zu einem Zeitpunkt t nach Anschalten des Zusatzpotentials noch im gelösten Molekül verbleibenden Anregungsenergie quantitativ zu analysieren. Dazu wurde für jede der Trajektorien, aus denen sich die Simulationen zusammensetzen, jeweils die als

$$E_{\text{ges}}(t) = E_{\text{MM,sol}}(t) + E_{\text{kin,sol}}(t) \quad (8.1)$$

definierte Gesamtenergie des Moleküls berechnet. E_{ges} setzt sich aus $E_{\text{MM,sol}}$, der MM-Energie (2.13) aller Atome des gelösten Moleküls, und deren kinetischen Energien $E_{\text{kin,sol}}$ zusammen. Zwar wäre es im Fall des cAPBs günstiger, anstelle des Gesamtmoleküls nur den Azobenzol-Farbstoff alleine zu betrachten, da dies einen direkteren Bezug zu den experimentellen Beobachtungen erlauben würde, jedoch führt die relativ starre Anbindung des Farbstoffes an den Peptidteil dazu, dass zwischen diesen beiden Teilen der zyklischen Peptide starke Energiefluktuationen stattfinden, deren Amplituden deutlich größer sein können, als die während der Energierelaxation zu untersuchenden Energiedifferenzen.

Bei der Auswertung der Messergebnisse der fs-Spektroskopie werden die beobachteten Dynamiken zumeist durch ein Ratenmodell beschrieben, in dem das Ensemble der untersuchten Moleküle nach ihrer Anregung sequentiell eine Reihe von getrennten Zwischenzuständen $i = 1, \dots, n$ mit jeweils unterschiedlichen Absorptionseigenschaften durchläuft, bevor es wieder in seinen Grundzustand übergeht [312]. In einem solchen linearen Ratenmodell kann der Verlauf der Abgabe der Anregungsenergie durch einen multi-exponentiellen Zerfall der Form

$$E_{\text{Rate}}(t) = \begin{cases} 0 & , t < 0 \\ \sum_{i=1}^n a_i \exp(-t/\tau_i) & , t \geq 0 \end{cases} \quad (8.2)$$

modelliert werden. Die Parameter dieses Modells sind die Zerfallskonstanten τ_i der einzelnen Zustände und die Amplituden a_i . Ein möglicher energetischer Offset zwischen dem *cis*-Isomer als Ausgangszustand vor der Anregung und dem vollständig relaxierten *trans*-Isomer wird im Folgenden durch die Angabe einer „unendlich“ großen Zeitkonstante mit zugehöriger Amplitude angegeben.

Zur Anpassung der Modellfunktion (8.2) an die mittels der Simulationen beobachtete Abnahme der Gesamtenergie E_{ges} wurde letztere zuerst auf ihren mittleren Wert während der einen Pikosekunde vor Anschalten des Zusatzpotentials normiert. Weiterhin wurden aus der Zeitreihe der E_{ges} eine Teilmenge von Zeitpunkten ausgewählt, die auf der in der fs-Spektroskopie üblichen linear-logarithmischen Zeitskala möglichst äquidistant sind. Diese spezielle Zeitskala, die auch in vielen der nachfolgenden Abbildungen verwendet wird, ist linear im Bereich von -1 ps bis 1 ps und wächst anschließend logarithmisch bis zu $t = 1$ ns an.

Als Optimierungskriterium der Anpassung des Ratenmodells diente die mittlere quadratische Abweichung

$$\chi^2(\tau_i, a_i) = \frac{1}{2n} \sum_{t=1}^M [E_{\text{ges}}(t) - E_{\text{Rate}}(t)]^2 \quad (8.3)$$

der Modellfunktion E_{Rate} (8.2) von den ausgewählten Datenpunkten E_{ges} , wobei die τ_i und die a_i als freie Parameter betrachtet, die Anzahl der exponentiellen Zerfälle hingegen fest vorgegeben wurde. Für die freien Parameter wurden interaktiv Startwerte vorgegeben und diese anschließend mit Hilfe des von den Mitarbeitern des Labors Z20 an unserem Lehrstuhl entwickelten IDL-Programms [313] auf eine minimale Abweichung (8.3) hin optimiert.

Die Anpassung der Modellfunktion (8.2) erweist sich nicht immer als unproblematisch, da beispielsweise die optimierten Parameter deutlich von den gewählten Startbedingungen abhängen können. Ferner ist festzustellen, dass sobald zwei Zerfallsprozesse nicht mindestens um eine Größenordnung voneinander differierende Zeitkonstanten aufweisen, deren Beträge nur sehr schlecht definiert sind. Dies führt zu der Beobachtung, dass eine Datenreihe, die durch zwei Zerfallszeiten von beispielsweise 3 ps und 10 ps modelliert werden kann, ebenso gut durch einen einzelnen Zerfall mit $\tau = 6$ ps zu beschreiben ist.

Um dieses Verhalten besser verstehen zu können und um die Aussagekraft der ermittelten Zeitkonstanten zu stärken, wurde von angestrebt, eine Fehlerabschätzung für die Anpassung der Zeitkonstanten zu treffen. Hierzu fand ein von Helmut Satzger [314] implementiertes Verfahren Anwendung, bei dem nach der erfolgten Optimierung des Modells untersucht wird, welche Auswirkung die Variation eines der Parameter auf die Güte des gefundenen Fits hatte. Dazu wurde jeweils iterativ getestet, um wieviel Prozent dieser Parameter, unter Konstanthalten der übrigen Parameter, nach unten und oben verändert werden kann, ohne dass sich das Maß χ^2 (8.3) um mehr als – willkürliche gewählte – 3% erhöht. Die auf diese Weise bestimmten Unsicherheiten für die Zeitkonstanten τ_i werden im Folgenden mit $e(\tau_i)$ bezeichnet.

Klassifizierung der Peptidkonfigurationen

Um einen Einblick in die während der Relaxation im Peptidteil des cAPBs stattfindenden Konformationsübergänge zu erhalten, ist es notwendig, ein Maß dafür festzulegen, wie

weit sich die Geometrie des Peptidrückgrats zu einem gegebenen Zeitpunkt von der eines *cis*-Isomers entfernt und sich derjenigen eines *trans*-Isomers angenähert hat. Die in Kapitel 5 vorgestellten Überlegungen lassen vermuten, dass es schwierig ist, im Raum der kartesischen Atomkoordinaten eine Metrik für den Abstand einer beliebigen Peptidkonfiguration von den als bekannt angenommenen Konformationen der *cis*- und *trans*-Isomere zu definieren. Stattdessen sind die Dihedralwinkel des Peptidrückgrats die zu diesem Zweck geeigneteren Koordinaten. Nun hat aber die in Abschnitt 7.2.1 präsentierte Untersuchung der Peptid-Konformationsräume gezeigt, dass eine Projektion dieser Räume auf die Ramachandranplots der einzelnen Aminosäuren häufig einen deutlichen Überlapp der *cis*- und *trans*-Geometrien zur Folge hat. In Abschnitt 7.2.2 wurde deshalb versucht, durch eine Hauptkomponentenanalyse diejenige Linearkombination der Winkelvariablen zu finden, die eine optimale Trennung zwischen den beiden Konformationsräumen ermöglicht. Laut Abbildung 7.5 gelingt dies im Fall des cAPBs allerdings nur mit beschränktem Erfolg. Dementsprechend ist festzustellen, dass man einen stark mit Rauschen und Fehlern behafteten Eindruck des *cis-trans*-Übergangs erhält, wenn man die während der Relaxationstrajektorien durchlaufenen Peptidgeometrien auf die erste der Hauptkomponenten – oder auch die optimale Kombination der ersten zwei oder drei Komponenten – projiziert (Daten hier nicht gezeigt).

Über eine einfache PCA hinausgehend wurde bei der Auswertung der Gleichgewichtstrajektorien jedoch auch schon die Leistungsfähigkeit der in Kapitel 5 entwickelten Methoden zur Klassifikation von Datensätzen nachgewiesen. So konnten mit diesen Methoden nicht nur die Peptidkonformationen aufgefunden, sondern anschließend auch alle in den Trajektorien auftretenden Peptidgeometrien einer dieser Konformationen zugeordnet werden.

Angewendet auf die nun vorliegende Problemstellung liegt es nahe, die *cis*- und die *trans*-Konformationen als Klassen zu betrachten, denen eine gegebene Peptidkonfiguration zuzuordnen ist. Als Modelle dieser beiden Klassen können zwei klassenlokale Dichteschätzungen (5.40) dienen, die gemeinsam zu einem Kodebuch zusammengefasst werden (für eine genauere Erklärung dieses Ansatzes zum Bau eines Klassifikators siehe die Veröffentlichung von Albrecht et al. [147]). Bezeichnet man die beiden Klassen mit C_c für *cis* und C_t für *trans*, so erhält man für einen Daten- oder Winkelvektor x nach (5.42) die Zuordnungswahrscheinlichkeiten $\hat{P}(C_c|\mathbf{x})$ und $\hat{P}(C_t|\mathbf{x})$, die sich stets zu Eins ergänzen. Als skalares Maß für den Fortschritt des Übergangs des Peptide von einer *cis*- zu einer *trans*-artigen Geometrie bietet sich somit $\hat{P}(C_t|\mathbf{x})$ an, da diese Größe am Anfang einer Relaxationstrajektorie einen Wert nahe bei Null aufweist, Konformationsübergänge hin auf den *trans*-Zustand mit einem Anstieg dieser Größe verbunden sind und deren Wert von 1 einen geometrisch vollständig relaxierten *trans*-Zustand anzeigt.

Es war notwendig, für jede der in Tabelle 8.1 aufgeführten Simulationen des cAPBs ein eigenständiges Modell der *cis* und *trans* Konformationslandschaften zu generieren, da davon auszugehen ist, dass deren Gestalt durch die gewählten Kraftfeldparameter beeinflusst wird. Die zur Erzeugung der Modelle notwendigen Gleichgewichtssimulationen wurden

in Abschnitt 8.1 erläutert. Im Gegensatz zu den in Kapitel 7 vorgestellten Langzeittrajektorien wurden hier die Konformationsräume jeweils mit Hilfe mehrerer kurzer Simulationen charakterisiert, die sich hinsichtlich der Startgeometrien der Modellpeptide unterschieden. Aus den in diesen Simulationen bei Raumtemperatur beobachteten Peptidkonfigurationen wurden die durchlaufenen Dihedralwinkelkombinationen extrahiert und, jeweils getrennt für die beiden Isomere, zu einem Gesamtdatensatz vereinigt, aus dem anschließend die Konfigurationsraumdichte mit Hilfe des UNIVAR-Verfahrens geschätzt wurde (siehe Abschnitt 5.2). Die dabei angesetzten Mischungsmodelle umfassen, wie zuvor, jeweils 100 univariate zirkuläre Normalverteilungen. Während die optimale Gaußbreite σ_{opt} im Fall der Langzeitsimulationen etwa 0,43 betrug (siehe Unterabschnitt 7.2.1), beobachtet man hier die maximale Log-likelihood im Mittel schon bei $\sigma_{\text{opt}} = 0,22$, was auf die geringere Temperatur und das, durch die kürzeren Simulationen bedingte, weniger umfassende Abtasten der Konformationsräume zurückzuführen ist.

Verwendet man die auf diese Weise konstruierten klassenlokalen Dichteschätzungen, um aus den einzelnen Relaxationstrajektorien die Zeitreihen der geometrischen Observablen $\hat{P}(C_t|\mathbf{x}(t))$ zu berechnen, so wird schnell deutlich, dass die relativ kleinen Gaußbreiten σ_{opt} eine sehr scharfe Partitionsfunktion (siehe Abbildung 5.4) erzeugen. Dies hat zur Folge, dass die geometrische Zuordnung zu einem der beiden Konformationsräume auf fast diskrete Weise geschieht, d.h. dass $\hat{P}(C_t|\mathbf{x}(t))$ im Zeitverlauf plötzlich von 0 auf 1 springt. Dieses – hier nicht gezeigte – Verhalten weist auf den Mangel an Informationen über die Gestalt der Energielandschaft der Peptide im Bereich zwischen den Konformationsräumen der *cis*- und der *trans*-Isomere hin, welcher nicht durch die Gleichgewichtstrajektorien abgetastet wurde. Um dennoch ein graduelles Maß für die ablaufenden geometrischen Änderungen zu erhalten, wurde das „Wissen“ über die Energielandschaften durch ein Hochskalieren der Gaußbreiten in den Mischungsmodellen spekulativ ausgedehnt. Dieses Vorgehen entspricht der im Rahmen des ASSO-Verfahrens in Abschnitt 7.3 vorgenommenen Vergrößerung des Konformationsraumes der Peptide. Dort wurde zur detaillierten Analyse der Peptidkonformationen ein Skalierungsfaktor von $\rho = 3,2$ gewählt, was einer Gaußbreite von 1,29 entspricht, und zeigt, dass auf dieser Auflösungsstufe die wichtigen Merkmale der Konformationsräume in Form von deutlich unterschiedlichen Konformationen sichtbar sind, während überflüssige Details durch die Glättung entfallen. Aus diesem Grund und um die Konsistenz zu wahren, wurden die Gaußbreiten der hier verwendeten Dichtemodelle auf den gleichen Wert von 1,29 gesetzt, was einer Skalierung um etwa $\rho = 6,0$ entspricht.

8.3 Ergebnisse und Diskussion

8.3.1 Übereinstimmung mit den Experimenten: Azobenzol und cAPB

Bei der Vorstellung und Diskussion der Ergebnisse der diversen Relaxations-Simulationen möchte ich mit den beiden Simulationsläufen beginnen, die in der Chronologie zuerst durchgeführt wurden und deren Ergebnisse schon zuvor Grundlage einer Veröffentlichung waren [138]. Dies sind diejenige Simulation des Azobenzols, die in Tabelle 8.1 als R1 bezeichnet wird und die Simulation R5 des cAPBs.

Abbildung 8.1 zeigt die aus diesen Rechnungen folgenden Zeitverläufe der Gesamtenergie E_{ges} (8.1) auf der erwähnten linear/logarithmischen Zeitskala. An beide Zeitreihen wurde sowohl ein bi- als auch ein tri-exponentieller Zerfall der Form (8.2) angepasst. Alle auf diese Weise ermittelten Zeitkonstanten, ihre Unsicherheiten und die zugehörigen Amplituden werden in Tabelle 8.2 aufgeführt, während in Abbildung 8.1 jeweils nur der tri-exponentielle Zerfall als durchgezogener Graph dargestellt ist. Zum Verständnis dieser Art von Abbildungen ist zu beachten, dass durch die lineare Auftragung von E_{ges} über der logarithmischen Zeitskala die Zerfallskonstanten des Ratenmodells als Wendepunkte der Modellfunktion in Erscheinung treten.

Im Fall des Azobenzols (R1, grün) weisen die Datenpunkte der Gesamtenergie nur eine geringe Streuung auf. Tabelle 8.2 zeigt dementsprechend, dass die Anpassung der Modellfunktion an diesen Datensatz gut definiert ist und dass ein bi-exponentieller Zerfall ausreicht, um die Dynamik dieser Größe zu beschreiben. Eine dritte Zeitkonstante ist nicht notwendig, wird sie dennoch angesetzt, so ist sie mit einer sehr hohen Unsicherheit verbunden. Verfolgt man bei R1 den Verlauf von E_{ges} bzw. den der angepassten Modellfunktion, so erkennt man bei $t = 0$ die durch das Anschalten des Zusatzpotentials (4.1) instantan in das Molekül eingebrachte Anregungsenergie von etwa 65 kcal/mol. Direkt im Anschluss an die Anregung wird in einem kurzen Zeitbereich ein großer Teil der Anregungsenergie vom Farbstoffmolekül an das Lösungsmittel abgegeben. Die diesem Prozess entsprechende Zeitkonstante beträgt 144 fs und ist lediglich mit einer Unsicherheit von $\pm 10\%$ verbunden. Visualisiert man die kartesischen Trajektorien des Azobenzols nach Lichtanregung als Film, so beobachtet man in diesem Zeitbereich die durch das Zusatzpotential getriebene ballistische Isomerisierungsreaktion der Azo-Gruppe. Die mit dieser Reaktion verbundene starke Änderung der Geometrie des Azobenzols (siehe Abbildung 1.12) ist mit einer raschen Bewegung der beiden Phenylringe durch das sie umgebende Lösungsmittel verbunden. Als Folge der dabei stattfindenden Beschleunigung und Verlagerung von DMSO-Molekülen wird ein großer Teil der Anregungsenergie direkt vom Azobenzol auf das umgebende Lösungsmittel übertragen.

Die spektroskopischen Messungen am *cis*→*trans*-Übergang des Azobenzols ergaben sehr gut vergleichbare schnellste Zeitkonstanten: Nägele et al. [135] bestimmten τ_1 zu 170 fs,

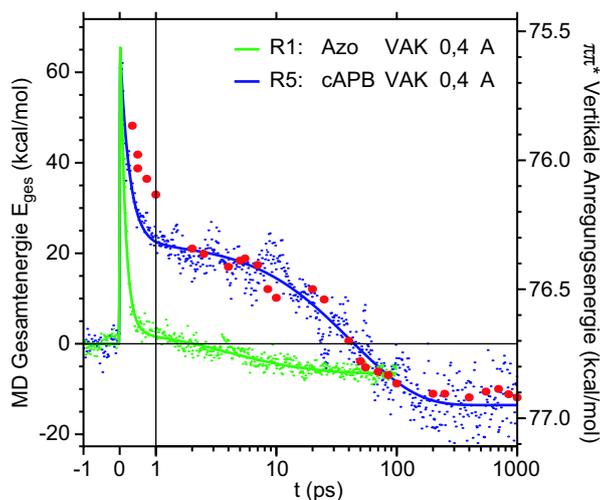


Abbildung 8.1: Energierelaxation: Azobenzol (R1) und cAPB (R5). Der Zeitverlauf der über die Einzeltrajektorien gemittelten Gesamtenergie E_{ges} des Azobenzols (R1, kleine grüne Punkte) und des cAPBs (R5, kleine blaue Punkte) wird über einer Zeitskala aufgetragen, die linear im Bereich von -1 ps bis 1 ps und logarithmisch für größere Zeiten ist. Die simulierte Lichtanregung findet zum Zeitpunkt $t = 0$ statt. Durchgezogenen Linien stellen die an diese Daten angepassten Modellfunktion (8.2) eines tri-exponentiellen Zerfalls dar, deren Parameter in Tabelle 8.2 aufgeführt werden. Zum Vergleich werden zusätzlich experimentelle Daten (rote Punkte) in Form der vertikalen Anregungsenergien des $\pi\pi^*$ -Übergangs im Farbstoff des cAPBs gezeigt. Erläuterungen im Text.

| R1: Azo VAK 0,4 A PNAS | | | | | | R5: cAPB VAK 0,4 A PNAS | | | | | |
|------------------------|-------------|-------|----------|-------------|-------|-------------------------|-------------|-------|----------|-------------|-------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 144 | -10 +10 | 82 | 141 | -15 +15 | 82 | 280 | -30 +40 | 50 | 258 | -35 +40 | 50 |
| 5,9 | -25 +35 | 10 | 3,1 | -100 +100 | 7 | 45,7 | -15 +20 | 39 | 19,3 | -100 +100 | 12 |
| | | | 12,7 | -55 +10 | 5 | | | | 61,7 | -40 +10 | 30 |
| ∞ | | -7 | ∞ | | -7 | ∞ | | -15 | ∞ | | -15 |

Tabelle 8.2: Zeitkonstanten: Azobenzol (R1) und cAPB (R5). An die in den ersten Simulationen des Azobenzols und des cAPBs beobachtete Relaxationsdynamik der Gesamtenergie E_{ges} (8.1) wurde jeweils ein Ratenmodell (8.2) mit zwei (jeweils links) bzw. drei (jeweils rechts) exponentiellen Zerfällen angepasst (siehe Abbildung 8.1). Die Zeitkonstanten τ_i dieser Zerfälle sind untereinander angeordnet, wobei die schnellste Zerfallszeit in Femtosekunden und alle weiteren in Pikosekunden angegeben werden. Die Unsicherheiten $e(\tau_i)$ der Zeitkonstanten sind in Prozent und die Amplituden a_i in kcal/mol aufgeführt. Ein „ ∞ “ bezeichnet den Offset für große Zeiten.

während neuere Messungen von Spörlein [139] auf einen etwas schnelleren Verlauf von 130 fs deuteten. Die schon in Abschnitt 4.2 erläuterte Vorstellung bezüglich des in diesem Zeitbereich ablaufenden Prozesses ist, dass die Azo-Gruppe unmittelbar nach der Lichtanregung des $n\pi^*$ -Übergangs im ersten elektronisch angeregten Zustand (S_1) von *cis* nach *trans* isomerisiert und dabei durch eine konische Durchschneidung in den vibronisch angeregten elektronischen Grundzustand (S_0) übergeht. Die Übereinstimmung der in den Simulationen auftretenden Zeitkonstante mit den Messergebnissen zeigt, dass das hier angesetzte Zusatzpotential den Isomerisierungsprozess im Rahmen der MM-Näherung nicht nur qualitativ, sondern auch quantitativ richtig beschreibt. Sie ist damit ein weiterer Hinweis auf die Korrektheit der Annahme einer Inversionsreaktion an einem der Stickstoffatome der Azo-Gruppe.

Im Hinblick auf die Beschreibung der Photoisomerisierung ist das verwendete Modell jedoch limitiert. Im Gegensatz zu Abbildung 8.1 wird in den Experimenten eine zweite Zeitkonstante von etwa 1 bis 2 ps beobachtet [139]. Diese Zerfallszeit wurde als diffusive Dynamik des Moleküls auf der S_1 -Fläche identifiziert [138], d.h. als „Suche“ des Moleküls nach der konischen Durchschneidung [138]. Bei der Entwicklung des Modellpotentials wurde jedoch angenommen, dass das Molekül einen Reaktionspfad durchläuft, der es direkt in den Grundzustand führt. Das Fehlen der angesprochenen Zerfallszeit in den Simulationen sollte demnach konstruktionsbedingt sein.

Laut Abbildung 8.1 verbleiben eine Pikosekunde nach Anschalten des Zusatzpotentials noch etwa 10 kcal/mol der anfangs eingebrachten Energie im Azobenzol. Nimmt man an, dass sich diese Energie gleichmäßig auf die 62 Freiheitsgrade des Azobenzols^a verteilt, so entspricht dies einer Erwärmung des gelösten Moleküls um etwa 100 K. Ferner ist davon auszugehen, dass sich zu diesem Zeitpunkt auch das umgebende DMSO in einem Nichtgleichgewichtszustand befindet, nicht nur durch die schon übertragene Anregungsenergie, sondern auch durch die Störung der das Azobenzol umgebenden Solvathülle aufgrund der Geometrieänderung des Farbstoffs. Basierend auf diesen Überlegungen ist die in Tabelle 8.2 aufgeführte zweite Zeitkonstante von 5,9 ps mit der Relaxation des Gesamtsystems in das thermische Gleichgewicht zu identifizieren. Diese Schlussfolgerung wurde auch schon bei der Interpretation der zeitaufgelösten spektroskopischen Messungen am Azobenzol gezogen [135], in denen diesem Prozess Zeitkonstanten zwischen 10 und 20 ps zugeordnet wurden. In späteren Messungen am Azobenzol ergaben sich Kühlzeiten von 15 ps [138], 11 ps [139] und 5,6 ps [315]. Der Vergleich mit den Experimenten zeigt, dass in der hier betrachteten Simulation R1 die Kühlzeit des Azobenzols in DMSO höchstwahrscheinlich unterschätzt wird. Zu erklären ist dies mit dem verwendeten flexiblen „all-atom“ Modell des DMSOs (VAK-DMSO). Bei dieser Art von DMSO-Modellen wurde in Kapitel 3 stets eine Überschätzung der spezifischen Wärme festgestellt, durch die ein schneller Wärmeübergang vom gelösten Molekül in das Lösungsmittel begünstigt wird. Der Einfluß des DMSO-Modells auf die Kühlzeiten des Azobenzols soll weiter unten

^aEin Azobenzol-Molekül umfasst 24 Atome, von deren 72 Freiheitsgraden jedoch 10 durch das Konstanthalten der Bindungslängen der Wasserstoff-Atome wegfallen.

noch detaillierter untersucht werden.

Zuvor möchte ich jedoch die Besprechung von Abbildung 8.1 fortsetzen. Als letztes Merkmal bleibt dort in der Gesamtenergie des Azobenzols noch ein Offset von 7 kcal/mol (siehe Tabelle 8.2) zwischen Ausgangs- und Endzustand ($t = 100$ ps) zu erklären. Dieser Offset spiegelt die unterschiedlichen freien Energien der beiden Azobenzol-Isomere wider. Zieht man in Betracht, dass die in Abschnitt 4.1 vorgestellten DFT-Berechnungen am freien Azobenzol im Vakuum für das *cis*-Isomer eine um etwa 16 kcal/mol höhere innere Energie als für das *trans*-Isomer ergaben, so ist der hier beobachtete geringere Offset wohl auf Fehler des MM-Kraftfeldes oder auf Solvatationseffekte zurückzuführen.

Geht man nun zur Simulation R5 des Modellpeptids cAPB über, dessen Gesamtenergie in Abbildung 8.1 blau dargestellt ist, so zeigt Tabelle 8.2, dass auch in diesem Fall zwei exponentielle Zerfallszeiten ausreichend sind, um die beobachtete Relaxationsdynamik zu modellieren. Werden drei Zeitkonstanten angesetzt, so hat dies zur Folge, dass zwei von ihnen, τ_2 und τ_3 , sich nur um einen Faktor 3 voneinander unterscheiden und dementsprechend schlecht definiert sind. Die schnellere Zeitkonstante von 280 fs beschreibt wie zuvor die ballistische Isomerisierung des Farbstoffs, welcher hier in Form des APBs in den Peptidring eingebunden ist. Offensichtlich führt die Anbindung des Farbstoffs an den Peptidteil zu einer Verlangsamung des Umschaltprozesses. Die fs-Messungen stimmen hiermit überein. So finden sich für das freie APB, welches nicht durch Simulationen untersucht wurde, Zeitkonstanten, die denen des unmodifizierten Azobenzols sehr nahe kommen [137]. Die Einbindung des APBs in das zyklische Molekül führt experimentell zu verlangsamten Isomerisierungszeiten von 270 fs [138] bzw. 230 bis 300 fs [139].

Abbildung 8.1 zeigt, dass im Fall des cAPBs nach der Isomerisierung des Farbstoffs ein deutlich größerer Teil der Anregungsenergie im Molekül verbleibt. Während beim Azobenzol direkt etwa 65 kcal/mol in das DMSO übergehen, sind es hier nur etwa 45 kcal/mol. Es ist davon auszugehen, dass sich die verbleibende Energie teils auf die vibronischen Freiheitsgrade des Gesamtmoleküls verteilt, aber anschließend auch teils in Form einer Verspannung sowohl des Farbstoffes als auch des Peptids vorliegt. Eine solche Verspannung muss notwendigerweise auftreten, da durch die Geometrieänderung des Azobenzolfarbstoffs die Randbedingungen des Peptidrückgrats nicht zu seiner aktuellen Konformation passen. Während gezeigt wurde, dass der Farbstoff seine Geometrieänderung in weniger als einer Pikosekunde durchläuft, kann auf dieser kurzen Zeitskala keine signifikante Änderung der Konformation des Peptidrückgrats stattfinden. Eine visuelle Inspektion der kartesischen Trajektorien zeigt in den ersten Pikosekunden nach der Lichtanregung allenfalls Veränderungen bei den direkt am Farbstoff angebondenen Peptid-Residuen Ala 1 und Phe 9. Erst die zweite der in Tabelle 8.2 aufgeführten Zeitkonstanten von 45,7 ps ist mit der Konformationsdynamik des Peptidrückgrats zu identifizieren, durch welche die Relaxation des Peptidteils von einer *cis*- zu einer *trans*-artigen Konformation abläuft.

Um beim Vergleich zwischen den Simulationen und den Experimenten nicht nur auf diejenigen Zeitkonstanten zurückgreifen zu können, die aus der Anpassung einer globalen Fitfunktion an das gesamte zeitaufgelöste Spektrum ermittelt wurden [139], wurde in Abbil-

dung 8.1 zusätzlich die Zeitabhängigkeit der vertikalen Anregungsenergie ΔE des $\pi\pi^*$ -Übergangs im *trans*-APB aufgenommen. Um ΔE zu berechnen, wurde an ausgewählten Zeitpunkten die spektrale Lage des Maximums der entsprechenden Absorptionsbande ermittelt. Die $\pi\pi^*$ -Absorptionsbanden von Farbstoffen mit konjugierten Doppelbindungen sind dafür bekannt, mit deutlichen spektralen Verschiebungen auf Torsionen im π -Elektronen-System zu reagieren. Im Fall des Azobenzols zeigten Monti et al. [217], dass der elektronische Grundzustand S_0 durch solche Verdrehungen weit weniger beeinflusst wird als der angeregte Zustand S_2 . Aus diesen Gründen ist davon auszugehen, dass der Zeitverlauf von ΔE die vibronische und konformelle Relaxation im Grundzustand widerspiegelt. Eine äquivalente Eigenschaft wurde weiter oben auch der in den MD-Simulationen auftretenden Gesamtenergie E_{ges} attestiert. Zusammen motiviert dies den in Abbildung 8.1 angestellten Vergleich der Zeitreihen von ΔE und E_{ges} , bei dem zur Anpassung der beiden Energien ein Offset und eine Skalierung frei gewählt wurden.

Die abweichenden Skalen der beiden Energien haben mehrere Ursachen. Erstens wird in den Simulationen die Energie des ganzen Moleküls betrachtet, während ΔE die Energiedifferenz des Farbstoffs misst. Zweitens liegt immer nur ein Bruchteil der Gesamtenergie in denjenigen Freiheitsgraden vor, die zur Änderung des π -Elektronensystems beitragen und drittens führt das Verhältnis der Steigungen der beiden Energieflächen S_0 und S_1 als Funktion der Kernkoordinaten zu einer weiteren Skalierung. Da diese drei Effekte nicht quantifiziert werden können, wurde die Skala der Anregungsenergie ΔE in Abbildung 8.1 bestmöglich an den Verlauf des Ratenmodells angepasst. Dabei wurde nur geringer Wert auf eine Übereinstimmung im Zeitbereich bis zu 1 ps gelegt, da dort, wie oben erläutert wurde, mit Abweichungen der Modellierung von der realen Dynamik des Farbstoffes gerechnet werden muss.

Abbildung 8.1 lässt für die Relaxationsdynamik nach der ersten Pikosekunde eine schon zuvor als erstaunlich gute bezeichnete Übereinstimmung zwischen den Simulationen und den Experimenten erkennen. In beiden Fällen tritt die gleiche Kinetik mit einer Zeitkonstanten von etwa 45 ps und einer Amplitude von etwa 45 kcal/mol auf. Neuere Messungen am cAPB ergaben durch einen globalen Fit der spektralen Daten alternativ eine Zerfallszeit von 64 ps [139]. Dieser Wert ist mit den hier vorgestellten Daten vereinbar, wenn man neben seinem möglichen Fehler auch die starke Streuung der Gesamtenergie E_{ges} berücksichtigt, durch die laut Tabelle 8.2 die im Ratenmodell gefundene Relaxationszeit mit einem Fehler mindestens der Größenordnung $\pm 15\%$, d.h. ± 7 ps behaftet ist. Ferner konnte mit Hilfe der Spektroskopie noch eine zusätzliche Zeitkonstante von 1,4 ns bestimmt werden [139], die in den auf 1 ns Dauer beschränkten und mit starkem Rauschen behafteten MD-Trajektorien nicht beobachtbar ist. Anscheinend ist diese große Zeitkonstante auch nicht aus der 1 ns-Zeitreihe der ΔE ablesbar.

Auch hier bleibt der Offset zwischen den Energien der *cis*- und der *trans*-Konformere des cAPBs zu diskutieren. Dieser wird von den Simulationen und den Experimenten übereinstimmend mit etwa 10 bis 15 kcal/mol angegeben. Er ist größer als derjenige im Azobenzol (R1), weshalb davon auszugehen ist, dass beim deutlich größeren und polareren cAPB die

Solvatationseffekte zu einer Absenkung der freien Energie des *trans*-Isomers führen.

8.3.2 Der Einfluss des DMSO-Modells auf die Kühlzeit des Azobenzols

Weiter oben wurde vermutet, dass der in der Simulation R1 am Azobenzol beobachtete schnelle Abkühlprozess (6 ps) auf die typische Überschätzung der spezifischen Wärme und damit vermutlich auch der mikroskopischen Wärmeleitfähigkeit des Lösungsmittels durch das verwendete flexible „all-atom“ Modell des DMSOs zurückzuführen ist. Diese Vermutung lässt sich leicht verifizieren, indem man gleichartige Simulationen der Relaxationsdynamik des Azobenzols mit mehreren der in Kapitel 3 vorgestellten DMSO-Modelle durchführt und dabei Modelle mit unterschiedlichen spezifischen Wärmen C_V auswählt.

Ein Vergleich dieser Art wurde mittels der Simulationen R2 bis R4 angestellt. Bei der Simulation R2 wurde das flexible „all-atom“ Modell HC-DMSO eingesetzt, das die endgültige Parametrisierung des in der Simulation R1 verwendeten VAK-DMSOs darstellt und dessen C_V von etwa 74 cal/(mol·K) etwa 160% über dem experimentellen Wert von 28,3 cal/(mol·K) liegt (siehe Tabelle 3.8). Als Gegenstück hierzu basiert die Simulation R4 auf dem rigiden „united-atom“ Modell MP-DMSO, das mit einem C_V von etwa 19 cal/(mol·K) eine um etwa 30% zu geringe spezifische Wärme aufweist. Als ein Kompromiss zwischen diesen beiden in ihren Eigenschaften extremen DMSO-Modellen wurde in Kapitel 3 mit dem MI-DMSO noch ein flexibles „united-atom“ Modell vorgestellt, dessen Parametrisierung eine Mischung der beiden Extrema ist. Durch die Annahme flexibler Bindungswinkel weist das MI-DMSO ein C_V von etwa 36 cal/(mol·K) auf, ein Wert der erwartungsgemäß zwischen denen der beiden anderen Modelle liegt, jedoch die experimentelle spezifische Wärme immer noch etwa 30% überschätzt.

Um die Ergebnisse der Simulationen R2, R3 und R4 einordnen zu können, ist es zunächst angebracht, R2 mit der zuvor publizierten [138] Simulation R1 zu vergleichen, weil bei diesen beiden Simulationen sehr ähnliche DMSO-Modelle verwendet wurden, so dass ähnliche Kühlzeiten erwartet werden können. Andererseits beinhaltet der Übergang von R1 zu R2, wie Tabelle 8.1 ausweist, eine Reihe weiterer Änderungen des MM-Kraftfelds. So wurden zusätzlich der Skalierungsfaktor λ_{14} und die Dihedralpotentiale des Farbstoffs verändert. Dementsprechend ist nicht auszuschließen, dass diese Änderungen die ballistische Isomerisierung des Azobenzols modifizieren.

Eine Antwort auf diese Frage liefert der Vergleich der Daten zur Energierelaxation in Tabelle 8.2 für R1 und Tabelle 8.3 für R2. Die in diesen beiden Tabellen fett gedruckten Zeitkonstanten für die ballistische Isomerisierung von 144 fs (R1) und 160 fs (R2), sowie für die Energiedissipation von 5,9 ps (R1) und 8,4 ps (R2) sind nahezu gleich. Daraus folgt, dass die Simulationen R1 und R2 die Photoisomerisierung von Azobenzol in DMSO trotz der Kraftfeldunterschiede in äquivalenter Weise beschreiben.

Es ist lediglich festzustellen, dass die Relaxationsdynamik bei R2 durch einen tri-exponentiellen Zerfall modelliert werden muss, während, wie oben schon festgestellt wurde, bei R1 ein bi-exponentielles Modell ausreicht. Mithin liefert R2 eine zusätzliche Zeitkonstante von 0,7 ps, deren Ursache unklar ist. Es könnte sein, dass es sich hierbei um ein Artefakt der Modellbildung handelt, welches auf den Umstand zurückzuführen ist, dass nach Annäherung an die *trans*-Geometrie für eine Zeitdauer von 1 ps zwischen dem die Isomerisierung treibenden Modellpotential und dem harmonischen Zielpotential interpoliert wird. In R1 wurde eine solche Interpolation nicht vorgenommen.

Andererseits könnte es auch sein, dass aufgrund der bei R2 verbesserten Statistik unterschiedliche Azobenzol-Populationen sichtbar werden. So könnten beispielsweise bei einem Teil der Farbstoffmoleküle die Geschwindigkeitsvektoren der Atome im Moment des Anschaltens des Modellpotentials in Richtung der Isomerisierungsordinate weisen und auf diese Weise unmittelbar die ballistische Isomerisierungsreaktion einleiten. Dagegen könnten bei einem anderen Teil der Moleküle die Geschwindigkeiten in andere Richtungen deuten, so dass die Reaktion erst durch die Wirkung der am Franck-Condon Punkt noch schwachen Kräfte des Modellpotentials (vgl. Abbildung 4.6) induziert wird und daher langsamer abläuft. Falls diese Erklärung zutreffend wäre, so würde das hier präsentierte klassische Modell sogar eine Interpretation der experimentell beobachteten Zeitkonstante von etwa 3 ps erlauben.

Ein auffälliger Unterschied zwischen den Simulationen R1 und R2 existiert jedoch: durch die in R2 stärker berücksichtigte elektrostatische 1-4-Wechselwirkung ($\lambda_{14} = 1,0$) nimmt der Offset zwischen dem *cis*- und dem *trans*-Azobenzol von 7 kcal/mol auf 17 kcal/mol zu. Dies ist zu erwarten, wenn man sich die in Abbildung 4.2 gezeigten Partialladungen im Azobenzol und dessen Geometrieänderung während der Isomerisierung (siehe Abbildung 1.12) in Erinnerung ruft. Beispielsweise kommen sich die beiden stark negativ geladenen, durch drei Bindungen voneinander getrennten Kohlenstoffatome, die an die Stickstoffe gebunden sind, im *cis*-Azobenzol sehr nahe. Eine Heraufskalierung ihrer Coulomb-Wechselwirkung führt damit zu einer Anhebung der inneren Energie dieses Isomers. Auf die Geschwindigkeit der modellierten Inversionsreaktion hat dieser Energieunterschied aber anscheinend keine große Auswirkung.

Wendet man sich nun dem Vergleich der drei alternativen DMSO-Modelle zu, so sind die in Abbildung 8.2 sowie in Tabelle 8.3 aufgeführten Daten zur Energierelaxation zu analysieren. Tabelle 8.3 zeigt, dass auch die in den Simulationen R3 und R4 auftretenden Dynamiken am besten mit einem tri-exponentiellen Zerfall beschrieben werden können. Insbesondere weist Tabelle 8.3 nach, dass die Wahl des Lösungsmittelmodells weder einen signifikanten Einfluss auf die schnellen Zeitkonstanten der Isomerisierungsreaktion noch auf den energetischen Offset zwischen den Isomeren hat. Im Gegensatz hierzu ist jedoch eine deutliche Abhängigkeit der beobachteten Kühlzeit τ_3 vom verwendeten DMSO-Modell zu beobachten: von 8,4 ps beim HC-DMSO über 33,3 ps beim MI-DMSO zu 39,3 ps beim MP-DMSO, d.h. es zeigt sich eine Zunahme der Kühlzeit mit Abnahme der Anzahl der Freiheitsgrade des eingesetzten DMSO-Modells bzw. mit dessen spezifischer Wärme.

8 Die lichtinduzierte Relaxationsdynamik der Modellpeptide

| R2: Azo HC 1,0 B | | | | | | R3: Azo MI 1,0 B | | | | | |
|------------------|-------------|-------|------------|-------------|-------|------------------|-------------|-------|-------------|-------------|-------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 215 | -10 +10 | 76 | 160 | -20 +20 | 66 | 317 | -10 +10 | 70 | 180 | -25 +25 | 62 |
| 7,2 | -10 +15 | 25 | 0,7 | -40 +80 | 17 | 23,2 | -15 +20 | 16 | 1,1 | -35 +45 | 20 |
| | | | 8,4 | -15 +15 | 22 | | | | 33,3 | -20 +20 | 13 |
| ∞ | | -17 | ∞ | | -17 | ∞ | | -14 | ∞ | | -14 |

| R4: Azo MP 1,0 B | | | | | |
|------------------|-------------|-------|-------------|-------------|-------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 293 | -15 +15 | 70 | 169 | -20 +20 | 63 |
| 30,3 | -15 +15 | 21 | 1,1 | -30 +50 | 19 |
| | | | 39,3 | -15 +15 | 19 |
| ∞ | | -16 | ∞ | | -17 |

Tabelle 8.3: Zeitkonstanten: Azobenzol in verschiedenen DMSO-Modellen (R2-R4). Analog zu Tabelle 8.2 werden die Parameter des Ratenmodells für die Simulationen des Azobenzols in unterschiedlichen DMSO-Modellen aufgeführt.

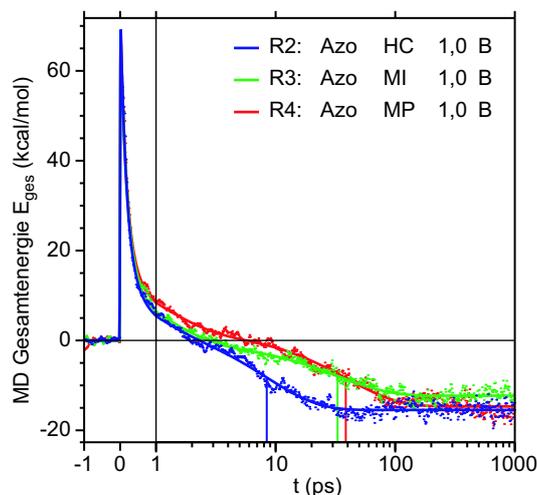


Abbildung 8.2: Energierelaxation: Azobenzol in verschiedenen DMSO-Modellen (R2-R4). Zur Beurteilung der Auswirkung des gewählten DMSO-Modells auf die Kühlzeit des Azobenzols werden hier drei Simulationen verglichen: R2 (blau), bei der das flexible „all-atom“ Modell HC-DMSO eingesetzt wurde, R3 (grün), in der das semi-flexiblen „united-atom“ Modell MI-DMSO Verwendung fand und schließlich die Simulation R4 (rot), die auf dem rigiden „united-atom“ Modell MP-DMSO basiert. Alle drei DMSO-Modelle wurden in Kapitel 3 vorgestellt und diskutiert. Die den Kühlzeiten des Azobenzols (die in Tabelle 8.3 aufgeführten Zeitkonstanten τ_3) entsprechenden Wendepunkte der Modellfunktionen werden durch senkrechte Linien hervorgehoben. Zur Erläuterung der gewählten Darstellung siehe die Legende von Abbildung 8.1.

Auch die Energierelaxationskurven aus Abbildung 8.2 bestätigen diesen Trend, wie aus einer visuellen Inspektion unmittelbar hervorgeht.

Damit ist die Vermutung bestätigt, dass die Kühlzeit des gelösten Azobenzols stark vom DMSO-Modell bestimmt wird. Es zeigt sich jedoch auch, dass die beobachteten Kühlzeiten nicht alleine durch die unterschiedlichen spezifischen Wärmen der DMSO-Modelle erklärt werden können, da entgegen den Erwartungen beim MI-DMSO, welches die experimentelle spezifische Wärme um etwa 30% überschätzt, keine zu kleine sondern eine zu große Kühlzeit auftritt. Vielmehr deuten die relativ ähnlicher Kühlzeiten bei den beiden „united-atom“ Modellen (MI, MP) darauf hin, dass die Frage, ob in einem DMSO-Modell die Wasserstoffatome explizit berücksichtigt werden oder nicht, einen großen Einfluss auf dessen mikroskopische Wärmeleitfähigkeit und damit auf die Kühlzeit hat.

Da in keinem der Experimente eine Kühlzeit von mehr als 20 ps gemessen wurde, erscheinen die beiden Modelle MI und MP für die Simulation der Relaxationsdynamik des Azobenzols und auch des cAPBs nicht geeignet zu sein. Besonders schwer wiegt hier, dass bei Verwendung eines dieser beiden unflexiblen Modelle die am Azobenzol beobachtete Kühlzeit in den Bereich der Zeitkonstanten der zu untersuchenden Konformationsdynamik der Modellpeptide verschoben wird, was deren Untersuchung – zumindest mittels der Observablen der Gesamtenergie – unmöglich macht. Hierauf möchte ich weiter unten noch einmal zurückkommen. Festzustellen bleibt, dass das flexible „all-atom“ Modell (HC) des DMSOs aufgrund seiner sehr schnellen Wärmedissipation offensichtlich die optimale Wahl zur Untersuchung der Relaxationsdynamik der Modellpeptide darstellt. Insbesondere wurde also in diesem Fall in der Originalarbeit [138] „intuitiv“ die richtige Wahl getroffen.

8.3.3 Die Reproduzierbarkeit der Simulationen des cAPBs

Als nächster Schritt bei der Analyse der simulierten Relaxationsdynamiken ist es an der Zeit, von den „alten“ (R1, R5) zu den „neuen“ Simulationen überzugehen. Die Simulation R5 des cAPBs wurde schon oben diskutiert. Tabelle 8.1 zeigt, dass die später durchgeführte Simulation R6 auf einem ähnlichen Kraftfeld basiert wie R5, da hier der gleiche Wert für λ_{14} und die gleichen Dihedralpotentiale gewählt wurden. Ein geringer Unterschied ergibt sich durch das Lösungsmittelmodell: VAK-DMSO in R5 und dessen mit geringfügig veränderten VDW-Parametern behafteter Nachfolger HC-DMSO in der Simulation R6. Die übrigen Unterschiede der beiden Simulationssysteme, wie die Systemgröße und das statistische Ensemble, wurden schon in Abschnitt 8.1 beschrieben.

Unter der begründeten Annahme, dass die Auswirkungen der aufgezählten Unterschiede auf die Relaxationsdynamik des cAPBs vernachlässigbar klein sind, kann die Simulation R6 als eine Wiederholung des „Experiments“ R5 angesehen werden und dazu dienen, die Reproduzierbarkeit der eingesetzten Methodik zu untersuchen. Wie schon zuvor, werden die Ergebnisse der Simulationen R5 und R6 in einer Abbildung (8.3) zur Energierelaxation

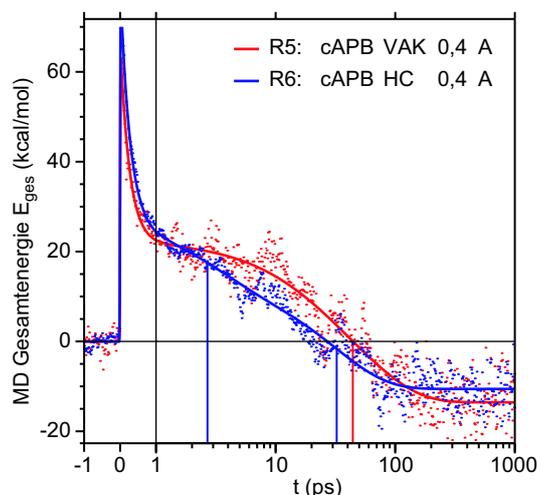


Abbildung 8.3: Energierelaxation: alte (R5) und neue (R6) Simulation des cAPBs. Analog zu Abbildung 8.1 wird die in den Simulationen R5 (rot) und R6 (blau) des cAPBs beobachtete Relaxation der Gesamtenergie dargestellt. Beide Simulationen verwendeten die gleichen Kraftfeldparameter für das Modellpeptid, und sehr ähnliche DMSO-Modelle. Die senkrechten Linien markieren die in Tabelle 8.4 hervorgehobenen Zeitkonstanten.

| R5: cAPB VAK 0,4 A PNAS | | | | | | R6: cAPB HC 0,4 A | | | | | |
|-------------------------|-------------|-------|----------|-------------|-------|-------------------|-------------|-------|-------------|-------------|-------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 280 | -30 +40 | 50 | 258 | -35 +40 | 50 | 395 | -25 +30 | 61 | 263 | -55 +45 | 59 |
| 45,7 | -15 +20 | 39 | 19,3 | -100 +100 | 12 | 22,8 | -20 +20 | 34 | 2,8 | -75 +100 | 16 |
| | | | 61,7 | -40 +10 | 30 | | | | 31,8 | -25 +50 | 28 |
| ∞ | | -15 | ∞ | | -15 | ∞ | | -11 | ∞ | | -11 |

Tabelle 8.4: Zeitkonstanten: alte (R5) und neue (R6) Simulation des cAPBs. Analog zu Tabelle 8.2 werden die Parameter des Ratenmodells aufgeführt.

und in einer Tabelle (8.4) mit Zeitkonstanten und Amplituden (8.4) zusammengefasst.

Man erkennt in Abbildung 8.3, dass die Zeitreihen der Gesamtenergie aus den beiden Simulationen sehr ähnlich verlaufen. Berücksichtigt man ihre jeweiligen statistischen Schwankungen, die im Fall der nur aus sechs Trajektorien zusammengesetzten Simulation R5 größer sind als bei R6, so ist in weiten Teilen der Zeitachse eine gute Übereinstimmung festzustellen. Allenfalls im Zeitbereich zwischen etwa 4 ps und 20 ps könnte davon gesprochen werden, dass die Energie bei R6 etwas niedriger verläuft als bei R5. Da der Zeitverlauf der Energie bei R6 in der logarithmischen Auftragung fast linear erscheint, kann er nicht mit einem bi-exponentiellen Zerfall repräsentiert werden. Übereinstimmend damit, dass die beiden Modellfunktionen im Zeitbereich unter einer Pikosekunde fast deckungsgleich verlaufen, werden für die beiden Simulationen mit 280 fs bzw. 263 fs fast

identische Zeitkonstanten für die Isomerisierung des Farbstoffs gefunden. Nennenswerte Unterschiede treten jedoch bei der Zeitkonstante der Konformationsdynamik auf, die bei R6 mit etwa 32 ps anzugeben ist, während für R5 ein Wert von etwa 46 ps bestimmt wurde. Diese beiden Zeiten sind aber im Rahmen ihrer angegebenen Ungenauigkeiten miteinander vereinbar. Speziell bei der Simulation R6 ist die Zeitkonstante τ_3 mit einer Unsicherheit von etwa +50% in Richtung längerer Zeiten behaftet, was einer absoluten Unsicherheit von etwa 16 ps entspricht.

Mit dem Vergleich der Simulationen R5 und R6 konnte somit gezeigt werden, dass die Computer-Experimente zur Relaxationsdynamik der Modellpeptide zwar wiederholbar sind, die berechneten Zeitkonstanten jedoch aufgrund unzureichender Statistik mit einer relativ hohen Unsicherheit verbunden sind. Um diese Unsicherheit zu vermindern, müsste über eine deutlich größere Anzahl von Einzeltrajektorien gemittelt werden.

8.3.4 Veränderungen an λ_{14} und den Dihedralpotentialen

Nachdem oben vermittels der Simulation R6 des cAPBs eine Verbindung zu den zuvor veröffentlichten Ergebnissen hergestellt wurde, kann diese Simulation nun als Ausgangspunkt für die Beantwortung der weiter oben formulierten Fragestellungen (ii) und (iii) dienen, welche sich mit den Auswirkungen einer Veränderung des Skalierungsfaktors λ_{14} bzw. der Dihedralpotentiale des Farbstoffs auf die simulierte Relaxationsdynamik des cAPBs beschäftigen. Als Datenquellen stehen dafür die folgenden Simulationen zur Verfügung: R7, bei der – im Gegensatz zu R6 – der Literaturwert 1,0 für λ_{14} angenommen wurde, R8, die sich von R6 darin unterscheidet, dass hier der „weichere“ Satz *B* der Dihedralpotentiale angenommen wurde und schließlich R9, die von beiden Veränderungen des Kraftfelds gleichzeitig betroffen ist (siehe hierzu auch Tabelle 8.1).

Ich möchte die Diskussion der Kraftfeldmodifikationen mit einem Vergleich der Simulationen R6 und R8 beginnen und mithin zuerst lediglich den Übergang von den steifen Dihedralpotentialen *A* zu den weichen Potentialen *B* untersuchen (siehe Abschnitt 4.1). In Abbildung 8.4 werden die zugehörigen Energielaxationsdaten und tri-exponentiellen Fitkurven miteinander verglichen. Die dort abgebildete Kurve R6 (blau) ist mit jener aus Abbildung 8.3 identisch und wurde schon dort diskutiert. Bei R8 (grüne Kurve) verläuft die Energielaxation offensichtlich langsamer. Sie wird im Pikosekundenbereich, wie aus Tabelle 8.5 hervorgeht, durch Relaxationszeiten von 11,2 ps und 188 ps charakterisiert, während dort bei R6 lediglich eine Zeitkonstante von 31,8 ps auftrat.

Anscheinend verzögert der Übergang zu den weicheren Dihedralpotentialen *B* den Konformationsübergang des Peptids von *cis* nach *trans* so sehr, dass die Prozesse der Energiedissipation (11,2 ps) und die der Konformationsrelaxation (188 ps) in den Daten getrennt sichtbar werden, während sie bei den Potentialen *A* (R6) noch untrennbar überlagert waren und dort in Form der 31,8 ps-Kinetik gemeinsam in Erscheinung traten. Eine anschauliche Deutung der Funktion der Dihedralpotentiale als mehr (*A*) oder minder (*B*) steife Fe-

8 Die lichtinduzierte Relaxationsdynamik der Modellpeptide

| R6: cAPB HC 0,4 A | | | | | | R7: cAPB HC 1,0 A | | | | | |
|-------------------|-------------|-------|-------------|-------------|------------|-------------------|-------------|-------|-------------|-------------|-----------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 395 | -25 +30 | 61 | 263 | -55 +45 | 59 | 320 | -25 +35 | 52 | 260 | -35 +40 | 49 |
| 22,8 | -20 +20 | 34 | 2,8 | -75 +100 | 16 | 10,1 | -20 +20 | 34 | 3,6 | -60 +100 | 18 |
| | | | 31,8 | -25 +50 | 28 | | | | 16,6 | -35 +100 | 21 |
| ∞ | | -11 | ∞ | | -11 | ∞ | | -3 | ∞ | | -4 |

| R8: cAPB HC 0,4 B | | | | | | R9: cAPB HC 1,0 B | | | | | |
|-------------------|-------------|-------|-------------|-------------|------------|-------------------|-------------|-------|------------|-------------|------------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 248 | -30 +50 | 43 | 192 | -30 +35 | 43 | 253 | -35 +55 | 45 | 182 | -30 +40 | 44 |
| 21,9 | -15 +20 | 40 | 11,2 | -20 +25 | 32 | 16,4 | -25 +30 | 33 | 7,3 | -25 +30 | 30 |
| | | | 188 | -40 +80 | 16 | | | | 293 | -50 +100 | 15 |
| ∞ | | -9 | ∞ | | -14 | ∞ | | -5 | ∞ | | -12 |

Tabelle 8.5: Zeitkonstanten: cAPB, R6 bis R9. In zu Tabelle 8.2 analoger Weise werden die Parameter des Ratenmodells für die Simulationen R6 bis R9 des cAPBs zusammengefasst. Die Anordnung der vier Simulationen soll den Vergleich der Ergebnisse bezüglich der beiden im Text diskutierten Kraftfelddetails vereinfachen: in der linken Spalte gilt $\lambda_{14} = 0,4$, während rechts die beiden Simulationen mit $\lambda_{14} = 1,0$ angeordnet sind. In der oberen Zeile wurde der Satz A der Dihedralpotentiale des Farbstoffs verwendet, in der unteren Zeile entsprechend der Satz B.

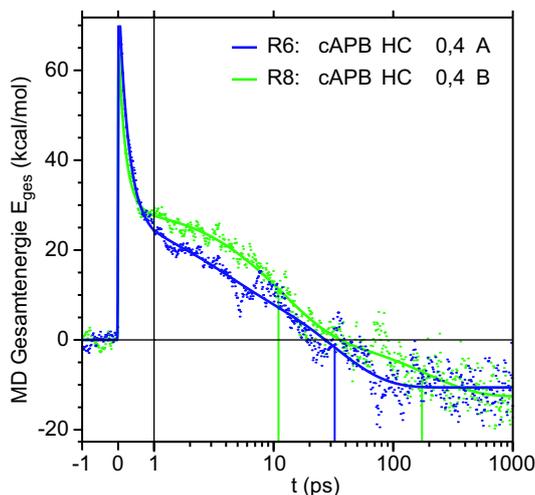


Abbildung 8.4: Energierelaxation: Veränderung der Dihedralpotentiale (R6 und R8). Analog zu Abbildung 8.1 wird die in den Simulationen R6 (blau, Dihedralpotentiale A) und R8 (grün, Dihedralpotentiale B) des cAPBs beobachtete Relaxation der Gesamtenergie dargestellt. Die senkrechten Linien markieren die Zeitkonstanten aus Tabelle 8.5.

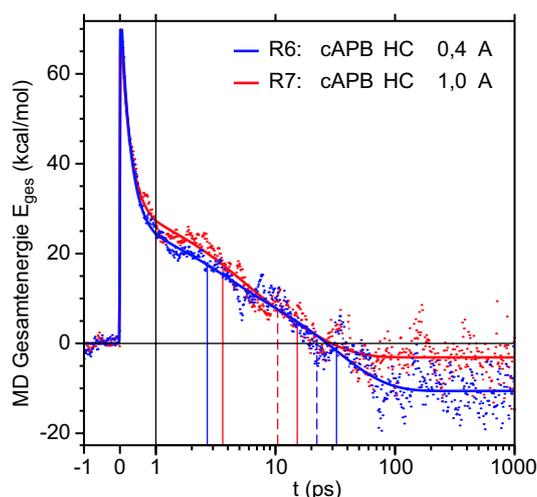


Abbildung 8.5: Energierelaxation: Änderung des Skalierungsfaktors λ_{14} (R6 und R7). Analog zu Abbildung 8.1 werden die Zeitreihen der Gesamtenergie für die Simulationen R6 (blau, $\lambda_{14} = 0,4$) und R7 (rot, $\lambda_{14} = 1,0$) gezeigt.

dern, welche das Peptid nach der *cis-trans*-Isomerisierung des Farbstoffs in Richtung der *trans*-Konformation ziehen, wird weiter unten (Abschnitt 8.3.6) anhand der Diskussion einzelner Trajektorien nachgeliefert werden. Schon hier sei jedoch festgehalten, dass die Identifikation der Dihedralpotentiale B mit einer weicheren Zugfeder die Verzögerung des Peptid-Konformationsübergangs im Fall R8 erklären kann. Erstaunlich ist lediglich, wie stark sich eine relativ geringfügige Änderung dieses Kraftfelddetails (vgl. Tabelle 4.3) auf die Konformationsdynamik auswirkt.

Wie Tabelle 8.5 ausweist, verkürzt der Übergang von den Dihedralpotentialen A nach B beim cAPB zusätzlich (und im Gegensatz zu den Befunden am Azobenzol in DMSO, vgl. Abschnitt 8.3.2) auch noch die Zeitkonstante der ballistischen Isomerisierung von 263 fs auf 192 fs. Der letztere Wert ist dabei deutlich kleiner als das experimentelle Ergebnis von etwa 270 fs.

Die hier festgestellte Beschleunigung der Isomerisierung des Farbstoffs ist vermutlich darauf zurückzuführen, dass bei dem in das zyklische Peptid eingebundenen APB-Farbstoff im Verlauf der Isomerisierung große Auslenkungen der Dihedralwinkel in der Azo-Gruppe auftreten. Dabei wird der bei größeren Auslenkungen steilere Verlauf der Dihedralpotentiale B abgetastet, was zu größeren Kräften und damit zur Beschleunigung der Reaktion führt. Mithin wurde in der Originalpublikation [138] mit den Dihedralpotentialen A sowohl in Hinblick auf die Modellierung der ballistischen Isomerisierungsreaktion als auch auf die Stärke der „Zugfeder“ die richtige Wahl getroffen.

Im nächsten Schritt kann nun der Übergang von $\lambda_{14} = 0,4$ zu $\lambda_{14} = 1,0$ untersucht werden (vgl. Frage (ii) weiter oben), wozu es nützlich ist, zunächst den Fall der steifen Feder (A) d.h. den Übergang von Simulation R6 zu R7 zu betrachten. Man erkennt in Abbildung 8.5,

dass die Energierelaxation während der ersten etwa 20 ps bei R6 und R7 ähnlich verläuft. Dementsprechend weist Tabelle 8.5 für diese beiden Simulationen sehr ähnliche schnellste Zeitkonstanten aus, ein Befund der mit den Ergebnissen zur ballistischen Isomerisierung des freien Azobenzols übereinstimmt (vgl. Abschnitt 8.3.2). Während anschließend die Gesamtenergie bei R6 bis auf 11 kcal/mol absinkt, scheint die Energierelaxation bei R7 jedoch abgeschlossen zu sein und endet in einem *cis-trans*-Offset von lediglich 4 kcal/mol (siehe Tabelle 8.5). Der bei R7 relativ kleine Offset ist überraschend, ruft man sich in Erinnerung, dass im Fall des Azobenzols beim äquivalenten Übergang von R1 nach R2 die Energiedifferenz von 7 kcal/mol auf 17 kcal/mol anwuchs, was dort anhand einer Analyse der Partialladungsverteilung und der daraus zu erwartenden Wirkung der λ_{14} -Änderung erklärt werden konnte. Zur Interpretation des überraschenden Ergebnisses von R7 muss daher angenommen werden, dass dort die Konformationsrelaxation auch nach einer Nanosekunde noch nicht weit fortgeschritten und ein beträchtlicher Teil der Anregungsenergie im Peptid als Verspannung verblieben ist.

Die soeben skizzierte Interpretation wird durch den Vergleich der Relaxationsdaten in Tabelle 8.5 bestätigt. Für R7 wird dort als langsamste Kinetik lediglich eine Zeitkonstante von 16,6 ps ausgewiesen, welche im typischen Zeitbereich der Energiedissipations-Kinetiken liegt (vgl. dazu die ähnliche 11,2 ps-Kinetik von R8, die oben solchen Prozessen zugeordnet wurde). Im Hinblick auf die Konformationsdynamik führt also ein Kraftfeld, welches die gestreckte β -Faltblatt-Konformation des *trans*-Zielzustands weniger begünstigt ($\lambda_{14} = 1,0$; R7) als die bisher diskutierten Kraftfelder ($\lambda_{14} = 0,4$; R6 und R8), zu einer so stark verzögerten Relaxationskinetik, dass sie im abgetasteten Zeitbereich von einer Nanosekunde noch nicht in Form einer Energieabgabe erkennbar ist. Zumindest für die steife Feder (A) ist diese Schlussfolgerung unabweisbar.

Wie die Untersuchung des Alanin-Dipeptids gezeigt hat (siehe Kapitel 6), führt eine Wahl von $\lambda_{14} = 1,0$ im Vergleich mit $\lambda_{14} = 0,4$ nicht nur zu einer energetischen Bevorzugung gestreckter Strukturen, sondern auch zu einer deutlich glatteren Landschaft der freien Energie sowie zu geringeren Barrieren zwischen den α -helikalen (*cis*) und den β -Faltblattartigen (*trans*) Konformationen. Im diskutierten Fall von R7 wird diese glattere Energielandschaft mit einem Farbstoff kombiniert, der eine harte Feder darstellt. Es erhebt sich die Frage, welche Auswirkungen diese Kombination auf die strukturellen Aspekte der Peptiddynamik hat.

Zur Beantwortung dieser Frage kann der in Abbildung 8.6 gezeigte Zeitverlauf der geometrischen Observablen (siehe Abschnitt 8.2) dienen. Dort beobachtet man bei R6 schon nach etwa 10 ps ein erstes Ansteigen der Zuordnungswahrscheinlichkeit der Peptidkonfigurationen hin zum Raum der *trans*-Konformationen, ganz im Gegensatz zur Simulation R7, für die über einen deutlich längeren Zeitraum von etwa 80 ps eine unveränderte Zugehörigkeit zu den *cis*-Konformationen sichtbar ist. Diese Beobachtung untermauert die Annahme einer bei $\lambda_{14} = 0,4$ (R6) inhärenten Präferenz des Peptidrückgrats für eine gestreckte Konfiguration. Ferner bestätigt Abbildung 8.6 noch einmal die Annahme einer bei R7 im Vergleich zu R6 verzögerten Relaxationskinetik, da sie im Zeitbereich zwi-

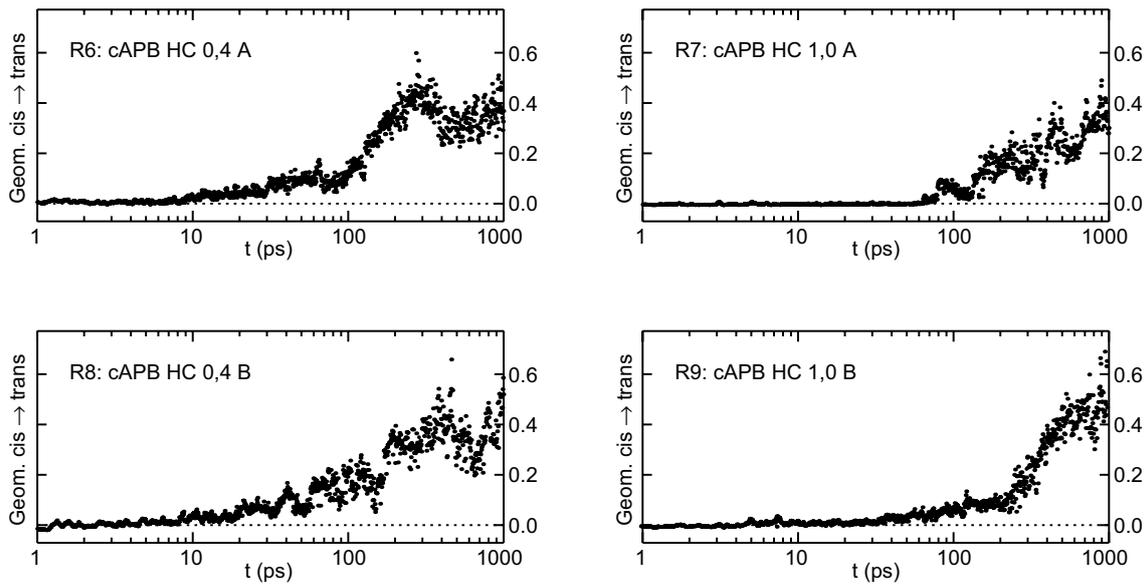


Abbildung 8.6: Klassifikation der Peptidgeometrien: cAPB, R6 bis R9. Auf einer logarithmischen Skala werden für die Simulationen R6 bis R9 die Zeitreihen der in Abschnitt 8.2 eingeführten geometrischen Observablen gezeigt, welche die Zuordnungswahrscheinlichkeit der Konfiguration des Peptidrückgrats entweder zum Raum der *cis*- oder dem der *trans*-Konformationen angibt. Der Wert von 0 dieser Observablen entspricht dem Ausgangszustand einer reinen *cis*-Geometrie, während der Wert von 1 zu einer vollständig relaxierten *trans*-Struktur gehört.

schon 100 ps und 1 ns bei R7 einen geringeren Anstieg der *trans*-Artigkeit ausweist, als bei R6.

Zur Erklärung der verzögerten Kinetik bei R7 ist es jedoch nicht ausreichend, lediglich den Kraftfeldparameter λ_{14} heranzuziehen. Auch dies zeigt Abbildung 8.6, wenn man in ihr die strukturelle Dynamik von R7 mit der von R9 vergleicht. In diesen beiden Simulationen galt $\lambda_{14} = 1,0$, im Gegensatz zu R7 (*A*) wurden in R9 jedoch die weicheren Dihedralwinkelpotentiale *B* angenommen. Die oben diskutierte Gegenüberstellung der Simulationen R6 und R8 hatte gezeigt, dass bei R8 die schwächere Federkraft der Dihedralpotentiale *B* die Konformationsdynamik des Peptidteils verlangsamt.

Analog ist beim Übergang von R7 zu R9 eine weitere Verzögerung der strukturellen Dynamik zu erwarten. Abbildung 8.6 macht jedoch deutlich, dass diese Erwartung nicht erfüllt wird. Vielmehr führt unter der Annahme von $\lambda_{14} = 1,0$ ein Aufweichen der Feder (*B*) zu einer Beschleunigung der strukturellen Relaxation. Dies ist daran zu erkennen, dass bei R9 eine erste Abweichung von den reinen *cis*-Strukturen schon etwa 40 ps nach der Lichtanregung auftritt und dort auch ein, verglichen mit R7, deutlich steilerer Anstieg der geometrischen Observablen im Zeitbereich nach 200 ps zu beobachten ist.

Wie kommt es dazu, dass ein weicherer die Konformationsübergänge treibender Farbstoff zu einer Beschleunigung der strukturellen Relaxation des Peptidrückgrats führt? Eine

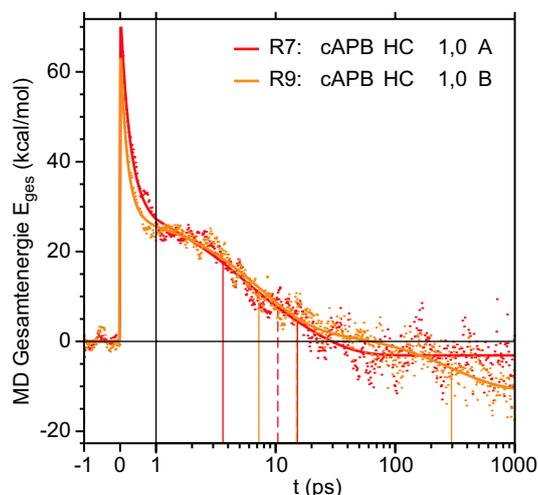


Abbildung 8.7: Energierelaxation: Veränderung der Dihedralpotentiale II (R7 und R9). Analog zu Abbildung 8.1 wird die in den Simulationen R7 (rot, Dihedralpotentiale A) und R9 (orange, Dihedralpotentiale B) des cAPBs beobachtete Relaxation der Gesamtenergie dargestellt. Im Gegensatz zu den in Abbildung 8.4 verglichenen Simulationen R6 und R8 galt hier $\lambda_{14} = 1,0$.

mögliche Erklärung hierfür ist, dass bei R7 die harte Feder (A) den Peptidteil auf dessen glatter *cis*-Energielandschaft ($\lambda_{14} = 1,0$) schnell und sehr direkt in Richtung auf die *trans*-Konformation treibt, das Peptid auf diesem direkten Weg jedoch häufig in einem *cis*-artigen Zustand gefangen bleibt, aus dem es nur über eine sehr hohe Energiebarriere in eine *trans*-artige Konformation entkommen kann.

Diese Interpretation ist konsistent mit den an den anderen Simulationen gemachten Beobachtungen. So erlaubt bei R9 die weichere Feder (B) größere Fluktuationen der Peptidkonfiguration, da sie weniger starke Randbedingungen anlegt. Durch den geringeren Zug des Farbstoffs stehen dem Peptid bei R9 mehr gangbare Wege in Richtung auf die *trans*-Konformation zur Verfügung, so dass es seltener in einen für den Ablauf der Relaxation ungünstigen Zustand getrieben wird. Ferner ist nach diesem Modell davon auszugehen, dass bei den Simulationen R6 und R9 ($\lambda_{14} = 0,4$) die Rauheit der Energielandschaft eine weitaus diffusivere Dynamik des Peptids bedingt als bei den Simulationen R7 und R9 ($\lambda_{14} = 1,0$). Solch eine diffusive Dynamik wird weniger stark vom Zug des Farbstoffs bestimmt und mündet deshalb auch nur seltener in einen der „getrappten“ Zustände, was die beobachtete Beschleunigung der Relaxationskinetik zur Folge hat.

Eine Rückkehr zur Observablen der Gesamtenergie und der diesbezüglich in Abbildung 8.7 gezeigte Vergleich von R7 und R9 bestätigen die oben ausgesprochene Vermutung, dass bei R9 die Konformationsdynamik des Peptids schneller und vollständiger abläuft als bei R7. Dies wird besonders an dem bei R9 (orange) größeren Offset von 12 kcal/mol deutlich und daran, dass Tabelle 8.5 für R9 eine Zeitkonstante von 293 ps aufführt. Das Auftreten dieser großen Zeitkonstante weist darauf hin, dass bei R9 die Konformationsdynamik,

verglichen mit R7, beschleunigt wird und sich damit in den von den Computer-Experimenten abgedeckten Zeitbereich unter einer Nanosekunde verschiebt. Auch die übrigen der in Tabelle 8.5 für R9 aufgeführten Zeitkonstanten sind mit den zuvor diskutierten Beobachtungen kompatibel. So ist, verglichen mit R7 und in guter Übereinstimmung mit dem Übergang von R6 zu R8, auch hier eine Beschleunigung der ballistischen Isomerisierung von 260 fs auf 182 fs zu beobachten, die weiter oben durch die mit diesem Übergang verbundenen Veränderungen der Dihedralpotentiale erklärt wurde. Ferner werden in R9, wie auch in R8, die Prozesse der Energiedissipation (7,3 ps) von denen der Konformationsdynamik getrennt.

Schließt man die Untersuchung der Kraftfeldmodifikationen mit einem Vergleich der in Tabelle 8.5 aufgeführten Dynamiken der Simulationen R8 und R9 ab, so zeigt es sich, dass mit den obigen Annahmen ein konsistentes Bild der Auswirkungen dieser Modifikationen auf die Relaxationsdynamik des cAPBs gewonnen werden konnte. Interpretiert man in diesem Bild die bei R8 und R9 verwendeten Kraftfelddetails, so findet sowohl bei R8 als auch bei R9 die weiche Variante des Farbstoffs Verwendung. Es ist somit bei beiden Simulationen eine sehr ähnliche Geschwindigkeit der Isomerisierungsreaktion des Azobenzols zu erwarten. Tabelle 8.5 bestätigt diese Erwartung.

Weiterhin wurde der Wechsel des Skalierungsfaktors λ_{14} zwischen R8 ($\lambda_{14} = 0,4$) und R9 ($\lambda_{14} = 1,0$) mit einem Übergang identifiziert, der von einer rauen Energielandschaft des Peptids, welche gestreckte Konformationen begünstigt, zu einer glatteren Energielandschaft führt, die keine solche Bevorzugung aufweist. Durch das Fehlen einer starken treibenden Feder ist bei den Simulationen R8 und R9 kein Auftreten des oben beschriebenen „Trappings“ zu erwarten, so dass bei R9 die fehlende Präferenz der Zielstruktur zu einer Verlangsamung der Konformationsdynamik führen sollte. Dies bestätigt Tabelle 8.5, die für diesen Prozess ein Anwachsen der Zeitkonstante von 188 ps auf 293 ps ausweist.

Auch die geometrische Observable aus Abbildung 8.6 zeigt eine Übereinstimmung mit den bisher formulierten Erwartungen bezüglich des Übergangs von R8 zu R9. Bei R8 ist hier ein initiales Ansteigen der *trans*-Artigkeit des Peptids nach etwa 20 ps zu beobachten, während dies bei R9 erst nach etwa 40 ps geschieht. Auch hier wird also die bei R8 dem Peptidrückgrat inhärente Präferenz für eine gestreckte Konfiguration sichtbar. Interessant ist jedoch auch, dass in Abbildung 8.6 zu späteren Zeiten der Anstieg der geometrischen Observablen bei R9 schneller zu verlaufen scheint als bei R8. Nach den zuvor vorgebrachten Erläuterungen weist dies darauf hin, dass die späteren Konformationsübergänge auf der bei R9 glatteren Energielandschaft etwas schneller ablaufen können als auf der rauen Landschaft von R8.

Zusammenfassend ist festzustellen, dass durch die Simulationen der Relaxation des cAPBs vier dynamische Prozesse identifiziert werden konnten, die jeweils mit einer *intrinsischen* Zeitkonstante verbunden sind:

- (i) die ballistische Isomerisierung des Azobenzol-Farbstoffs mit einer intrinsischen Zeitkonstante von etwa 260 fs,
- (ii) eine weitere unspezifische Dynamik im Bereich 1 bis 3 ps, die mit der Isomerisierung

des Farbstoffs verbunden ist,

(iii) der Prozess der Energiedissipation in das umgebende DMSO (etwa 7 ps) und schließlich

(iv) die Konformationsdynamik des Peptidteils des Modellsystems, für die Zeitkonstanten beginnend bei 30 ps bis hin in den Bereich der Nanosekunden gefunden wurden.

Der teilweise bei der Dynamik der Energierelaxation auftretende Überlapp dieser vier Prozesse führt jedoch dazu, dass die beobachteten *apparenten* Zeitkonstanten zumeist Mischungen der intrinsischen Zeitkonstanten sind. Als Beispiel für das Auftreten solcher Mischungen wurde weiter oben schon die Simulation R6 diskutiert. Aber auch in dem zuletzt angestellten Vergleich zwischen R8 und R9 ist in Tabelle 8.5 zu beobachten, dass mit dem Herausschieben der Konformationsdynamik von 188 ps auf 293 ps auch eine Abnahme der Kühlzeit von 11,2 ps auf 7,3 ps verbunden ist. Diese Beobachtung kann gleichfalls mit einer Abnahme der Überlagerung der Prozesse (iii) und (iv) zwischen R8 und R9 gedeutet werden.

8.3.5 Das cAPB in unterschiedlichen DMSO-Modellen

Mit dem im vorangehenden Abschnitt erzielten Verständnis für die während der Relaxation des cAPBs ablaufenden Prozesse und die damit verbundenen Zeitskalen ist es nun an der Zeit, noch einmal auf die Wichtigkeit der Wahl eines geeigneten Lösungsmittelmodells einzugehen. Schon in Abschnitt 8.3.2 wurde anhand der Simulationen R2 bis R4 gezeigt, dass die beiden „united-atom“ Modelle MI und MP die Zeitkonstante des Kühlprozesses (iii) (vgl. oben) beim Azobenzol von etwa 7 ps auf mehr als 30 ps vergrößern. Daraus wurde geschlossen, dass eine Beschreibung der Relaxation des cAPBs unter Verwendung dieser beiden DMSO-Modelle zu falschen *apparenten* Relaxationszeiten führen kann. Der Vergleich der Simulationen R9, R11 und R12, bei denen die selben DMSO-Modelle eingesetzt wurden wie bei R2 bis R4 (siehe Tabelle 8.1), ermöglicht es nun, diese Vermutung zu überprüfen.

Abbildung 8.8 zeigt den schon oben diskutierten Zeitverlauf der Gesamtenergie bei R9 und stellt ihm die entsprechenden Daten aus den Simulationen R11 und R12 gegenüber. Man erkennt, dass die Energieabgabe bei allen drei Simulationen bis etwa zum Zeitpunkt $t = 1$ ps sehr ähnlich verläuft. Dementsprechend stimmen die in Tabelle 8.6 für R11 und R12 aufgeführten Zeitkonstanten des ballistischen Isomerisierungsprozesses (i) (vgl. oben) innerhalb ihrer Unsicherheiten mit der bei R9 festgestellten Zerfallszeit von 182 fs überein. Diese Zeitkonstanten sind bei R11 und R12 relativ schlecht definiert, da in diesen Simulationen auch der im Pikosekundenbereich angesiedelte Prozess (ii) erscheint und in die Zerfallszeit des Isomerisierungsprozesses (i) einmischt.

Während Tabelle 8.6 für R9 neben einer Kühlzeit von etwa 7 ps noch eine Zeitkonstante der Konformationsdynamik von 293 ps ausweist, wird bei R11 lediglich eine weitere Zerfallszeit von etwa 39 ps und bei R12 von etwa 31 ps detektiert. In beiden Fällen liegt

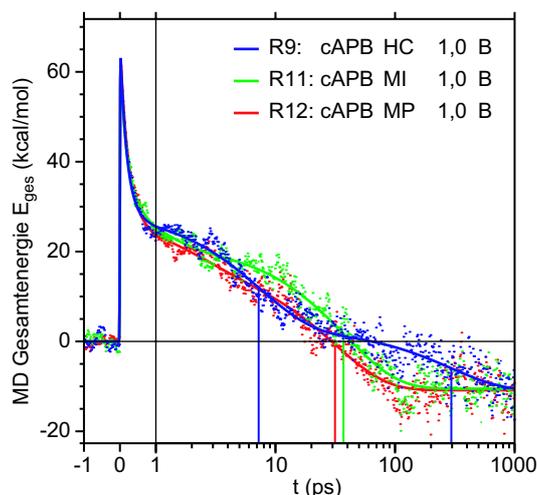


Abbildung 8.8: Energierelaxation: cAPB in diversen DMSO-Modellen (R9, R11 und R12). Analog zu Abbildung 8.1 wird die in den Simulationen R9 (blau, HC-DMSO), R11 (grün, MI-DMSO) und R12 (rot, MP-DMSO) des cAPBs beobachtete Relaxation der Gesamtenergie dargestellt. Die senkrechten Linien markieren die Zeitkonstanten aus Tabelle 8.5.

| R9: cAPB HC 1,0 B | | | | | | R11: cAPB MI 1,0 B | | | | | |
|-------------------|-------------|-------|------------|-------------|-------|--------------------|-------------|-------|-------------|-------------|-------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 253 | -35 +55 | 45 | 182 | -30 +40 | 44 | 385 | -25 +30 | 41 | 135 | -100 +100 | 29 |
| 16,4 | -25 +30 | 33 | 7,3 | -25 +30 | 30 | 37,0 | -10 +15 | 37 | 0,7 | -55 +100 | 22 |
| | | | 293 | -50 +100 | 15 | | | | 38,6 | -10 +15 | 36 |
| ∞ | | -5 | ∞ | | -12 | ∞ | | -12 | ∞ | | -12 |

| R12: cAPB MP 1,0 B | | | | | |
|--------------------|-------------|-------|-------------|-------------|-------|
| τ_i | $e(\tau_i)$ | a_i | τ_i | $e(\tau_i)$ | a_i |
| 395 | -25 +25 | 45 | 222 | -65 +50 | 43 |
| 26,1 | -15 +15 | 35 | 2,1 | -85 +100 | 13 |
| | | | 31,4 | -15 +25 | 31 |
| ∞ | | -12 | ∞ | | -12 |

Tabelle 8.6: Zeitkonstanten: cAPB, R9, R11 und R12. In zu Tabelle 8.2 analoger Weise werden die Parameter des Ratenmodells für die Simulationen R9, R11 und R12 des cAPBs zusammengefasst, in denen unterschiedliche Modelle des Lösungsmittels verwendet wurden.

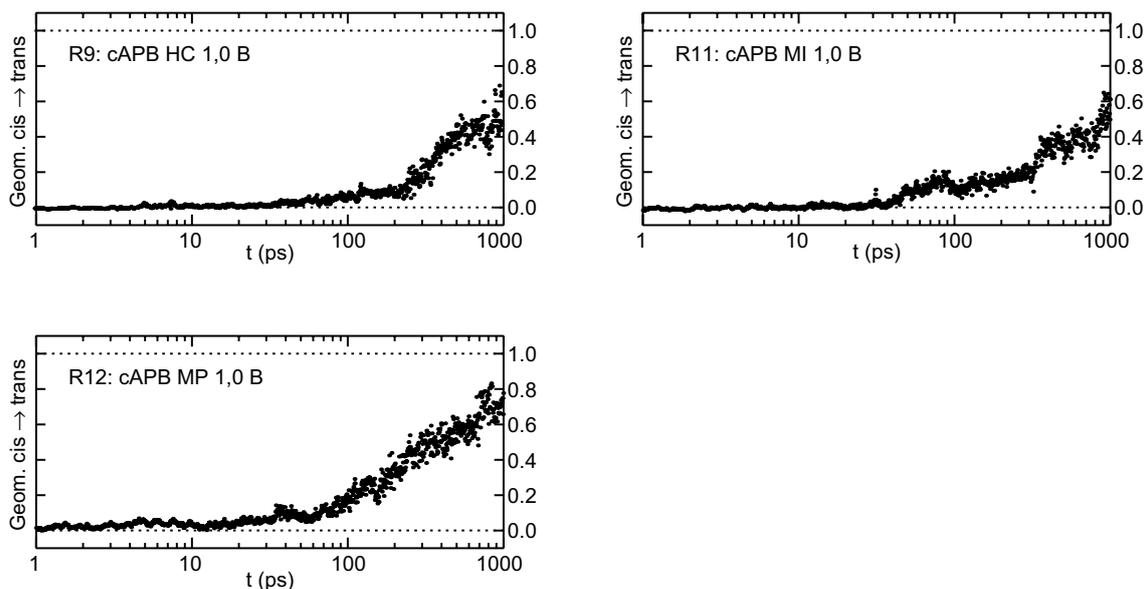


Abbildung 8.9: Klassifikation der Peptidgeometrien: cAPB, R9, R11 und R12. Analog zu Abbildung 8.6 wird die Klassifikation der in den Trajektorien durchlaufenen Peptidgeometrien als *cis*- oder *trans*-artig gezeigt, hier für diejenigen Simulationen, die sich im Lösungsmittelmodell unterscheiden.

diese Zeitkonstante im Bereich der am Azobenzol unter Verwendung der entsprechenden DMSO-Modelle (MI, MP) beobachteten Kühlzeiten.

Hier stellt sich die Frage, weshalb die Energierelaxation bei den Simulationen R11 und R12 keine Signatur der Konformationsdynamik aufweist. Als Erklärung ist eine durch die kompakten DMSO-Modelle bewirkte Verschiebung der Konformationsdynamik zu Zeiten länger einer Nanosekunde auszuschließen, denn Tabelle 8.6 weist für alle drei Simulationen stets den gleich großen Energieoffset von 12 kcal/mol aus. Dies beweist, dass die Energierelaxation am Ende aller drei Simulationen ähnlich weit vorangeschritten ist.

Eine zweite mögliche Erklärung für die Unbeobachtbarkeit der Konformationsdynamik bei R11 und R12 ist, dass zwar deren intrinsische Zeitkonstante nicht durch das DMSO-Modell beeinflusst wird und dementsprechend etwa 300 ps beträgt, jedoch ein starker Überlapp mit dem auf etwa 30 ps verlangsamten Kühlprozess zu nur einer einzelnen apparenten Zerfallszeit führt. Gegen diese Erklärung spricht jedoch sowohl der fast eine ganze Größenordnung betragende Unterschied zwischen diesen beiden intrinsischen Zeitkonstanten, als auch der Umstand, dass das Ergebnis ihrer Mischung deutlich von den reinen Kühlzeiten abweichen müsste. Das ist hier aber nicht der Fall.

Demzufolge bleibt nur eine mögliche Erklärung für die in Abbildung 8.8 bei R11 und R12 beobachteten Energierelaxationen übrig, nämlich diejenige, nach der die „united-atom“ Modelle des DMSOs (MI, MP) die Konformationsdynamik des Peptidteils derart beschleunigen, dass die intrinsische Zeitkonstante dieses Prozesses in einen Zeitbereich

von 30 bis 60 ps verschoben wird, in dem bei diesen DMSO-Modellen auch die Kühlzeiten liegen.

Bestätigt wird die soeben formulierte Interpretation durch Abbildung 8.9, die den Zeitverlauf der geometrischen Observablen für die drei Simulationen zeigt. Dabei ist zu beachten, dass (im Gegensatz zu der oben diskutierten Abbildung 8.6) in Abbildung 8.9 der gesamte Wertebereich $[0, 1]$ der Observablen abgebildet wird. Man beobachtet dort bei R11 (MI-DMSO), verglichen mit R9 (HC-DMSO), sowohl eine leichte Zunahme der strukturellen Umordnungen, als auch ein Ansteigen der *trans*-Artigkeit am Ende der Trajektorien von etwa 0,5 auf etwa 0,6, eine Beobachtung, die auf eine Beschleunigung der Konformationsdynamik hindeutet. Ein noch stärkeren Hinweis liefert jedoch die Simulation R12 mit dem steifsten DMSO-Modell (MP-DMSO), bei der ein kräftiger, schon bei $t = 60$ ps einsetzender Anstieg erkennbar ist, der in einem Wert der *trans*-Artigkeit von fast 0,8 endet. Damit haben bei R12 etwa 80% der Trajektorien nach einer Nanosekunde Simulationszeit den strukturell vollständig relaxierten *trans*-Zustand erreicht.

Zusammenfassend wurde hier gezeigt, dass die Verwendung der „united-atom“ Modelle MI-DMSO und MP-DMSO zumindest bei solchen Simulationen des cAPBs zu einer deutlichen Beschleunigung der strukturellen Umordnungen im Peptidteil führt, bei denen der Standardwert $\lambda_{14} = 1,0$ und die weiche Farbstofffeder (B) eingesetzt werden. Die im Abschnitt 8.3.2 für diese beiden DMSO-Modelle beobachteten verlängerten Kühlzeiten zeigen, dass dort länger ein „heißeres“ Modellpeptid vorliegt, welchem es leichter fällt, energetische Barrieren auf dem Reaktionspfad hin zu der *trans*-Konformation zu überqueren. Andererseits verdecken hier die überlagerten Kühlzeiten die mit der Konformationsdynamik verbundenen Kinetiken der Energierelaxation.

8.3.6 Ein Einblick in die lichtinduzierten Konformationsübergänge

Bislang wurden ausschließlich Mittelwerte von Observablen betrachtet. Wendet man diese Observablen stattdessen exemplarisch auf einzelne ausgewählte Relaxationstrajektorien an, so kann man einen detaillierten Einblick in die nach der Lichtanregung in den Modellsystemen ablaufenden Konformationsübergänge zu gewinnen. Um diese Konformationsübergänge auch auf atomarer Ebene diskutieren zu können, wird im Folgenden zusätzlich auch auf die „Rohdaten“ der Trajektorien in Form der kartesischen Atomkoordinaten und der Zeitreihen der Dihedralwinkel zurückgegriffen.

Bei der Untersuchung einzelner Trajektorien möchte ich mich auf die Simulation R6 (siehe Tabelle 8.1) beschränken, da die dort eingesetzte Parameterkombination die beobachtete Relaxationskinetik am besten reproduziert. Außerdem ist es von Vorteil, dass bei R6 die Konformationsdynamik des Peptids in dem von den Simulationen erfassten Zeitbereich von 1 ns weit voranschreitet.

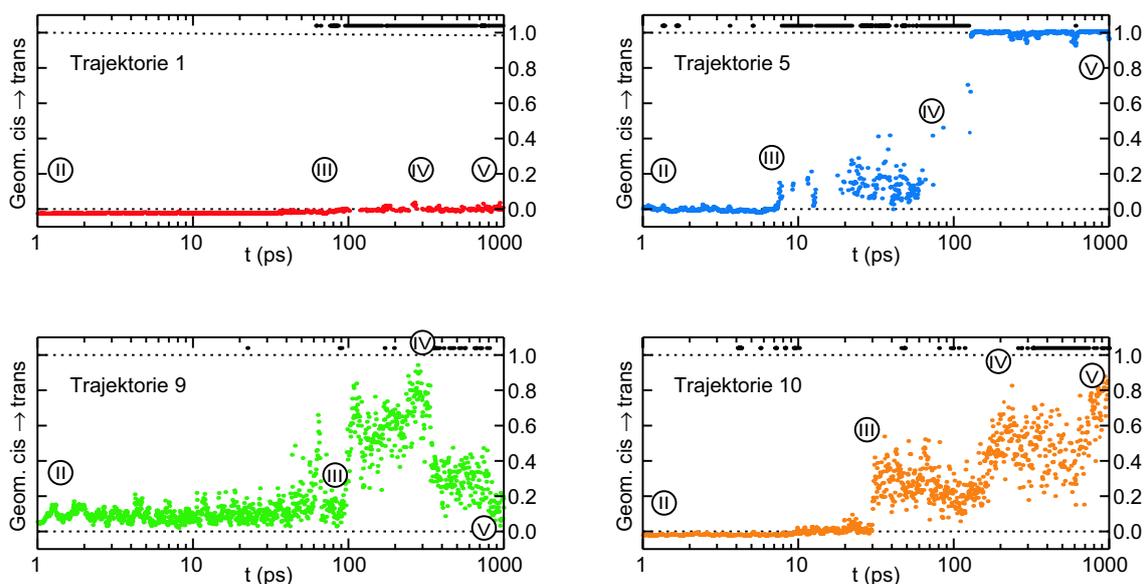


Abbildung 8.10: Geometrische Klassifikation einiger Einzeltrajektorien. Analog zu Abbildung 8.6 werden für vier aus der Simulation R6 ausgewählte Trajektorien (Nr. 1, 5, 9, 10) die Zeitreihen der geometrischen Observablen gezeigt. Wie zuvor, bezeichnet ein Wert nahe Null eine reine *cis*-Geometrie des Peptids, während ein Wert von 1 einer vollständig relaxierten *trans*-Geometrie entspricht. Schwarz eingezeichnete Datenpunkte oberhalb von 1 kennzeichnen Konfiguration des Peptids, die keinem der beiden Konformationsräume zugeordnet werden konnten. Römische Ziffern markieren diejenigen Zeitpunkte, an denen die Molekülstrukturen in Abbildung 8.12 visualisiert werden.

Abbildung 8.10 zeigt den Zeitverlauf der geometrischen Observablen für vier Trajektorien aus R6. Diese aus den zehn zur Verfügung stehenden Trajektorien ausgewählten Beispiele geben jede für sich einen typischen Verlauf der geometrischen Observablen wieder. Zu beachten ist, dass Abbildung 8.10 im Gegensatz den weiter oben gezeigten Abbildungen der strukturellen Observablen noch eine zusätzliche Angabe enthält, nämlich welche der durchlaufenen Peptidkonfigurationen weder dem *cis*- noch dem *trans*-Konformationsraum zugeordnet werden können. Um diese Information zu erhalten, wurde während der Klassifikation der Relaxationstrajektorien (vgl. Abschnitt 8.2) zuerst für jeden Datenpunkt des aus allen Trajektorien zusammengesetzten Datensatzes der zugehörige Wert der Gesamt-Modellichte (5.18) berechnet und nach einer Sortierung jene 10% aller Peptidkonfigurationen ermittelt, welche am unwahrscheinlichsten aus der Modellichte gezogen werden. Dieses Vorgehen definiert eine so genannte *Restklasse* [147]. Da über Datenpunkte aus der Restklasse im Dichtemodell nur wenige Informationen vorhanden sind, werden diese in Abbildung 8.10 gesondert hervorgehoben.

Beginnt man die Betrachtung von Abbildung 8.10 mit der rot eingefärbten Trajektorie 1 (im folgenden auch kurz mit T1 bezeichnet), so erkennt man im ganzen Zeitbereich bis zu 1 ns nach der Lichtanregung eine ausschließliche Zuordnung des Peptids zu den *cis*-Struk-

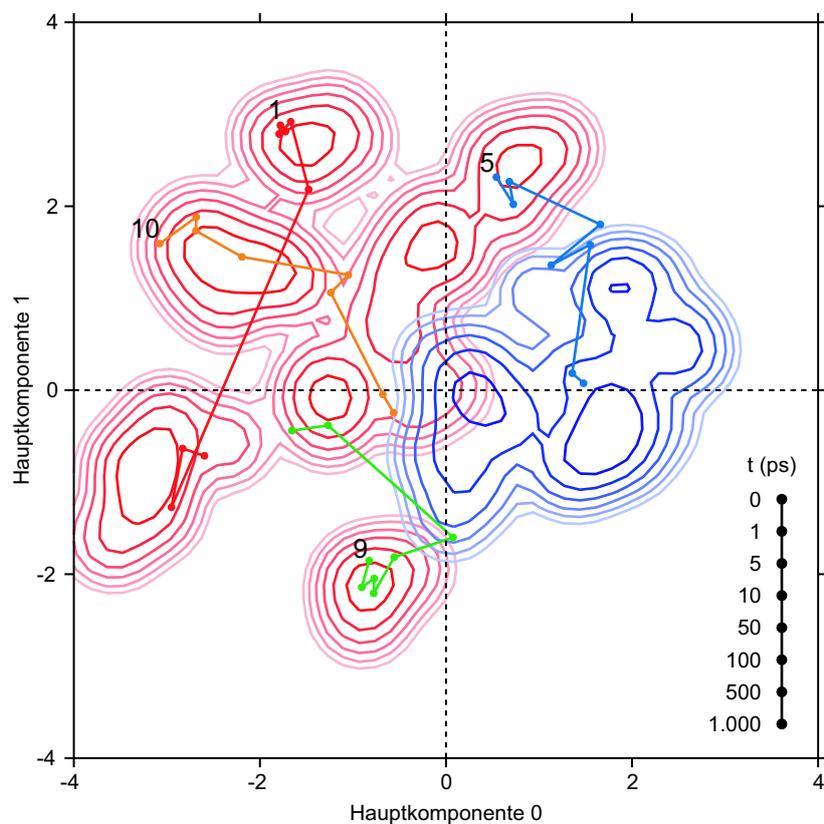


Abbildung 8.11: Hauptkomponentenprojektion einiger Relaxations-Trajektorien. Als Konturlinien werden die in der Simulation R6 im thermischen Gleichgewicht beobachteten Landschaften der freien Energie des cAPBs in ihrer Projektion auf die Hauptkomponenten der Konformationsräume dargestellt. Die Konturlinien weisen einen Abstand von 0,5 kcal/mol auf und wurden rot für die *cis*-Isomere und blau für *trans*-Isomere eingefärbt. In der gleichen Projektion werden zusätzlich vier der Relaxationstrajektorien (1, rot; 5, blau; 9, grün; 10 orange) aus R6 mit in die Abbildung aufgenommen. Beginnend bei der Nummer der Trajektorie verdeutlichen die Punkte die eingezeichnete logarithmische Zeitskala.

turen. Ausgenommen sind hiervon lediglich diejenigen Zeitpunkte in T1, an denen sich die Peptidkonfiguration so weit vom bekannten *cis*-Konformationsraum hin in die Restklasse entfernt hat, dass keine sichere Zuordnung zu den bekannten Strukturen möglich ist.

Der ständige Aufenthalt von T1 im Raum der *cis*-Strukturen ist auch gut in Abbildung 8.11 zu beobachten. Dort werden die untersuchten Trajektorien jeweils in Form des Zeitverlaufs der Peptid-Dihedralwinkel auf die Hauptkomponenten der *cis*- und *trans*-Energielandschaften projiziert, welche zuvor durch die Gleichgewichtssimulationen (vgl. Tabelle 8.1) bestimmt wurden. Man erkennt in Abbildung 8.11, dass T1 in der Nähe einer der *cis*-Konformationen startet, anschließend zwischen $t = 50$ ps und $t = 100$ ps den *cis*-Raum durchquert und sich schließlich nach einer Nanosekunde immer noch in diesem Raum befindet. Die Frage, welche Konformationsübergänge hinter diesem Verlauf verborgen sind und weshalb das Peptid den *trans*-Konformationsraum innerhalb der gegebenen Simulationszeit nicht erreicht, werde ich weiter unten anhand von Schnappschüssen aus der kartesischen Molekültrajektorie diskutieren.

Zuvor möchte ich jedoch noch die drei anderen Trajektorien anhand der Abbildungen 8.10 und 8.11 kurz charakterisieren. Die Trajektorie 5 (blau) zeigt laut Abbildung 8.10 ein zu T1 in gewisser Weise komplementäres Verhalten, da sie schon nach weniger als 10 ps den Raum der *cis*-Strukturen zügig verlässt und bereits nach 100 ps den *trans*-Zustand erreicht. Auch hier ermöglicht Abbildung 8.11 einen ergänzenden Einblick in die ablaufenden Strukturübergänge. Man stellt dort fest, dass T5 bei einer *cis*-Konformation beginnt, welche nahe an der Grenze zum *trans*-Konformationsraum liegt. Nach etwa 10 ps überschreitet T5 diese Grenze um anschließend nach 500 ps im Schwerpunkt der *trans*-Strukturen anzukommen.

Auch die Trajektorie 9 (grün) beginnt in Abbildung 8.11 bei einer *cis*-Konformation, die in der PCA nahe am *trans*-Raum liegt. Nach 100 ps streift T9 den *trans*-Raum, verlässt ihn jedoch anschließend wieder, um nach 500 ps zu einer, sich vom Ausgangszustand unterscheidenden, *cis*-Konformation zurückzukehren. Dieser strukturelle Verlauf von T9 lässt sich im Detail anhand von Abbildung 8.10 verfolgen, in der schon von Beginn der Trajektorie an keine ausschließliche Zuordnung zu den *cis*-Strukturen gezeigt wird. Dies deutet auf einen Überlapp der beiden Konformationsräume in der Nähe des Startpunkts von T9 hin (vgl. Abbildung 8.11). Im weiteren Verlauf von T9 steigt die *trans*-Artigkeit zuerst steil an, wobei nur sehr wenige nicht zuordnungsfähige Konfigurationen auftreten, um dann nach etwa 200 ps ein Maximum zu erreichen. Anschließend fällt die geometrische Observable für T9 steil ab und endet, in Übereinstimmung mit den an der PCA gemachten Beobachtungen, bei einer erneuten Zuordnung zum *cis*-Konformationsraum.

Schließlich ist die Trajektorie 10 (orange) ein gutes Beispiel für eine Trajektorie, bei der sich die Annäherung an den *trans*-Zustand monoton und in mehreren diskreten Schritten vollzieht, wie man Abbildung 8.10 entnehmen kann. In Abbildung 8.11 äußert sich dieses Verhalten in einer relativ geradlinigen Bewegung, die von T10 ausgehend von einer *cis*-Konformation in Richtung auf den *trans*-Raum ausgeführt wird. In Übereinstimmung mit Abbildung 8.10 zeigt Abbildung 8.11 ferner, dass T10 nach 1 ns den *trans*-Konforma-

tionsraum noch nicht ganz erreicht hat.

Die soeben diskutierte Charakterisierung der vier Trajektorien hat gezeigt, wie unterschiedlich und vielfältig die während der Relaxation des cAPBs ablaufenden strukturellen Veränderungen sein können. Um zusätzlich die Geometrie des Azobenzol-Farbstoffs mit in die Betrachtung einbeziehen zu können und um die Konformationsübergänge im Peptidrückgrat auf atomarer Ebene zu untersuchen, zeigt Abbildung 8.12 Schnappschüsse der Molekülstrukturen zu verschiedenen aus den Trajektorien herausgegriffenen Zeitpunkten.

Versucht man zuerst, die Gemeinsamkeiten der vier Trajektorien in Abbildung 8.12 zu untersuchen, so sieht man den Azobenzol-Farbstoff APB am Beginn der Trajektorien und damit direkt vor der Lichtanregung ($t = 0$ ps) erwartungsgemäß jeweils in seiner gewinkelten *cis*-Form vorliegen. Dagegen weist der Peptidteil in jeder der Trajektorien eine andere Konformation auf, die entweder mehr oder minder kompakt erscheint, jedoch stets vereinbar ist mit den in Abbildung 7.9 gezeigten *cis*-Konformationen.

Zeile II aus Abbildung 8.12 wurde den Trajektorien eine Pikosekunde nach der Lichtanregung entnommen. Man erkennt, wie in der ersten Pikosekunde das in Abschnitt 4.2 vorgestellte Modellpotential die als Inversion ablaufende Isomerisierungsreaktion der Azo-Gruppe bewirkt hat. Der Farbstoff liegt nun in allen Fällen als *trans*-Isomer vor, weist jedoch eine stark nicht-planare Geometrie und damit eine große Verspannung auf. Insbesondere bei den Trajektorien 1, 5 und 10 ist an den Strukturen II deutlich ersichtlich, dass – wie weiter oben schon vermutet wurde – das verspannte APB die Rolle einer Feder spielt, deren Federstärke durch die Dihedralpotentiale in der Azo-Gruppe bestimmt wird. Dabei sei in Erinnerung gerufen, dass in dem hier betrachteten Fall der Simulation R6 die steifere Parametrisierung A der Dihedralpotentiale Verwendung fand. Abbildung 8.12 zeigt ferner erwartungsgemäß, dass der Zug der Farbstoff-Feder erwartungsgemäß innerhalb der kurzen Zeitspanne der ersten Pikosekunde zu keinen sichtbaren Veränderungen an der Konformation des Peptidrückgrats führt. Allenfalls bei den beiden direkt an den Farbstoff angebondenen Residuen sind kleine Verschiebungen der Atompositionen zu beobachten.

Sucht man nach weiteren Gemeinsamkeiten der Trajektorien, so lässt ihr anschließender Verlauf in Abbildung 8.12 eine fortschreitende Relaxation des Azobenzol-Farbstoffs in Richtung auf seine planare *trans*-Struktur erkennen, die bei allen vier Beispielen nach einer Nanosekunde nahezu erreicht ist. Mit etwas Abstand betrachtet, zeigt Abbildung 8.12 ferner, wie das mit der Relaxation des Farbstoffs verbundene Öffnen des von ihm eingeschlossenen Winkels (vgl. Abbildung 1.12) dazu führt, dass das Peptidrückgrat zunehmend von seiner Startkonformation in Richtung auf eine gestreckte Konfiguration gezwungen wird.

Eine genauere Betrachtung der einzelnen Struktur-Schnappschüsse und der Vergleich zwischen zeitlich aufeinanderfolgenden Strukturen ermöglicht es nicht nur, zu beurteilen, welche Arten von Konformationsübergängen im Peptidrückgrat auftreten, sondern auch herauszufinden, an welchen Stellen sich die Peptidstruktur nach einer Nanosekunde noch

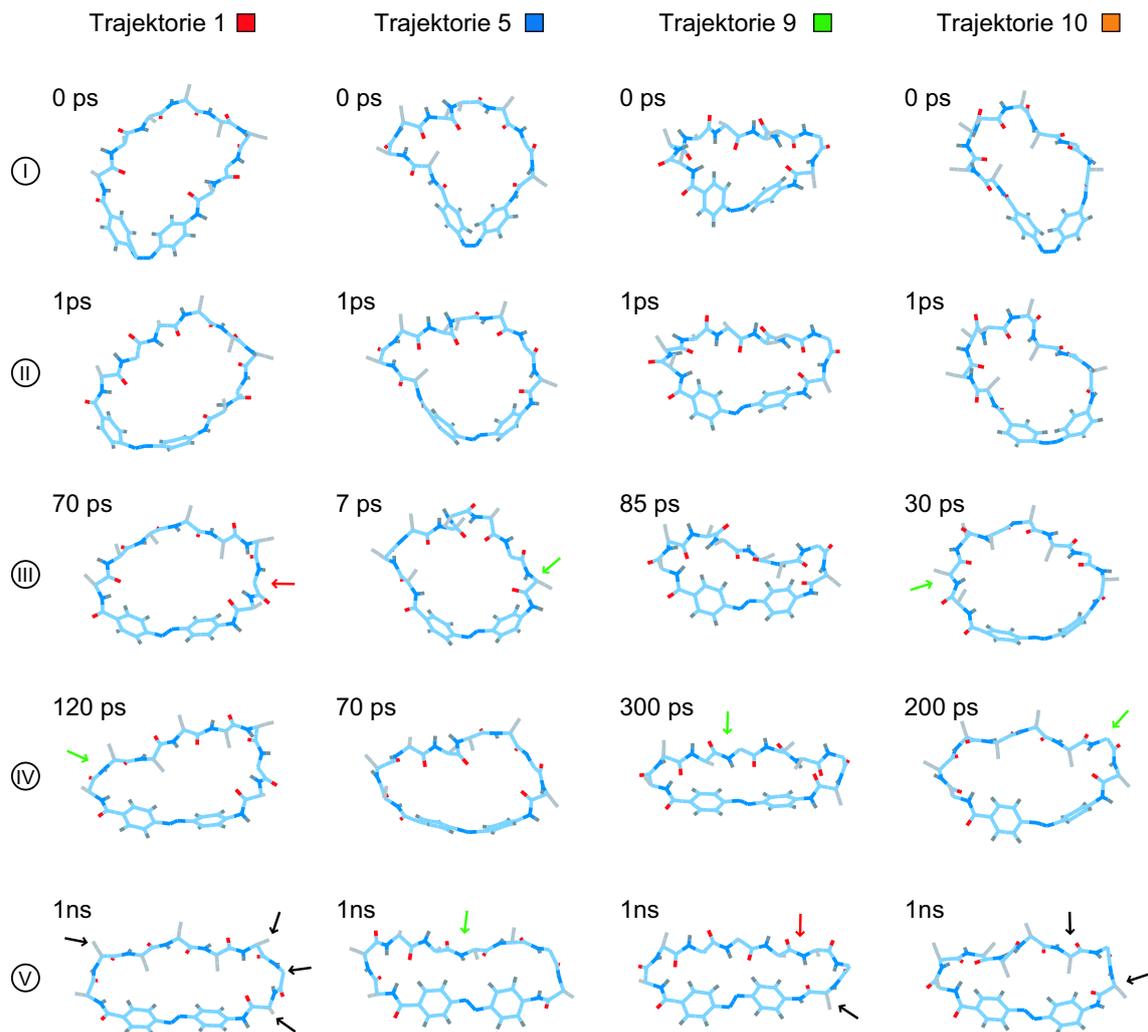


Abbildung 8.12: Schnappschüsse der Molekülstrukturen. Die gezeigten Strukturen wurden den vier näher untersuchten Relaxationstrajektorien der Simulation R6 entnommen und mit Hilfe des Programms VMD [17] visualisiert. Kohlenstoffe sind hellblau, Stickstoffe dunkelblau, Sauerstoffe rot und Wasserstoffe dunkelgrau dargestellt. Die Ausrichtung der Seitengruppen der Aminosäuren wird durch hellgraue Linien verdeutlicht, welche jeweils die Positionen des C_{α} und des C_{β} miteinander verbinden. Die Zeitachse verläuft senkrecht aber nicht gleichförmig. Bei jeder Trajektorie wird die Molekülstruktur im Moment der Lichtanregung (0 ps), nach 1 ps und am Ende der berechneten Trajektorie gezeigt. Die mit den römischen Ziffern III und IV markierten Strukturen wurden den Trajektorien zu solchen Zeitpunkten entnommen, bei denen laut Abbildung 8.10 deutliche Änderungen in der geometrischen Observablen und damit in der Konformation der Moleküle auftreten.

Grüne Pfeile deuten auf Residuen, an denen im letzten Schritt zwischen zwei Schnappschüssen eine Veränderung in Richtung auf die *trans*-Konformation eintrat. Rote Pfeile zeigen eine Veränderung weg vom Raum der *trans*-Strukturen an. Schwarze Pfeile verweisen auf diejenigen Residuen, bei denen am Ende der Trajektorie eine andauernde Abweichung von der strukturell vollständig relaxierten *trans*-Konformation existiert.

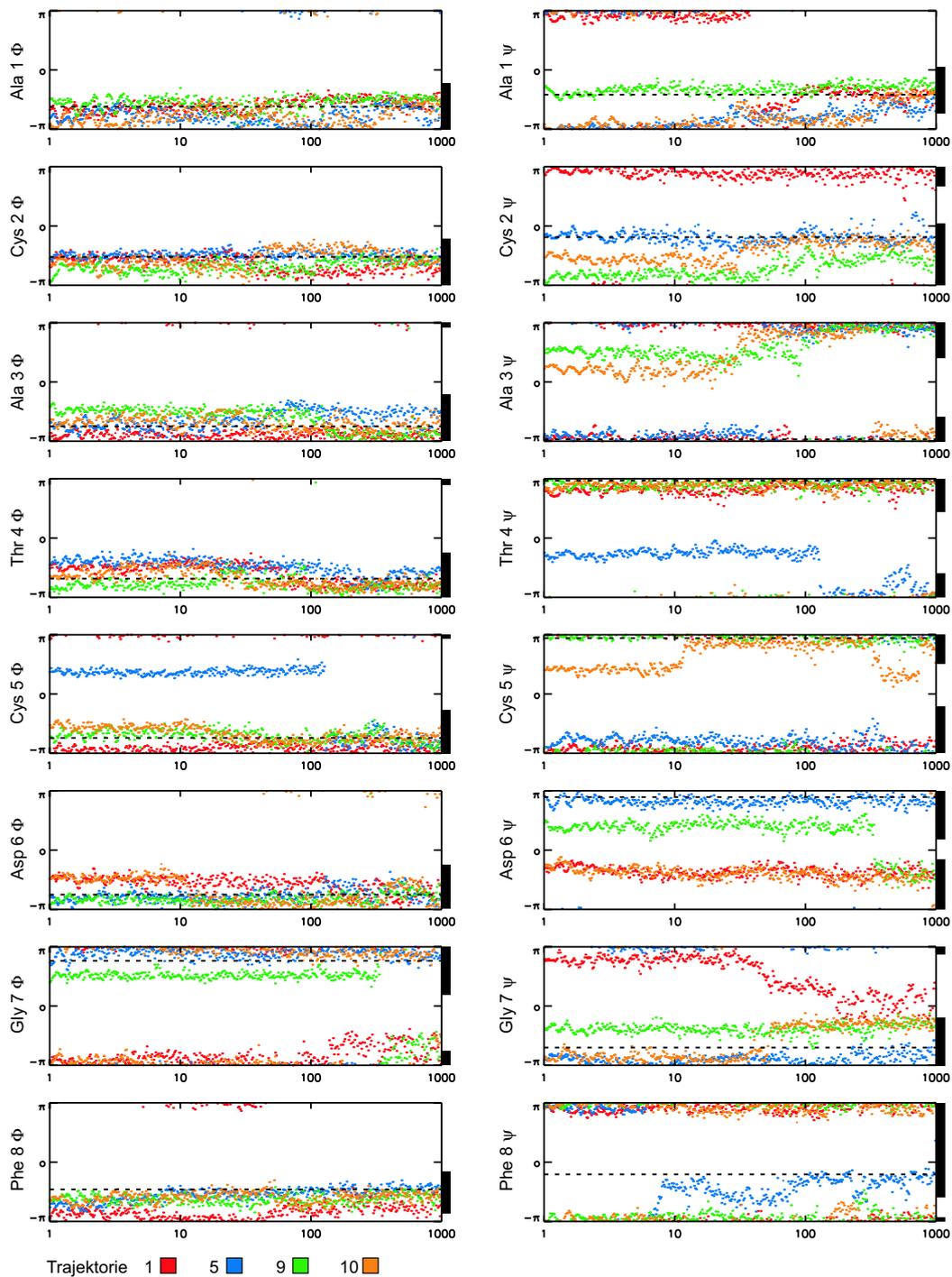


Abbildung 8.13: Die Einzeltrajektorien der Dihedralwinkel. Für die vier ausgewählten Relaxationstrajektorien der Simulation R6 werden die Zeitverläufe aller Dihedralwinkel ϕ und auf einer logarithmischen Zeitskala gezeigt. Die Farbkodierung der Trajektorien entspricht dabei derjenigen aus den Abbildungen 8.10 und 8.11. Die gestrichelten horizontalen Linien geben die Geometrie der mittleren *trans*-Konformation an, die mit Hilfe des ASSO-Verfahrens aus der entsprechenden Gleichgewichtstrajektorie der Simulation R6 ermittelt wurde. Die schwarzen Balken an der rechten Seite der Teilbilder markieren diejenigen Winkelbereiche, in denen die freie Energie der *trans*-Konformationen laut Abbildung 7.2 kleiner als 3 kcal/mol ist.

immer von einer vollständig relaxierten *trans*-Struktur unterscheidet. Für eine vertiefte Diskussion und eine Zuordnung der konformellen Strukturänderungen zu den Aminosäureresiduen, zeigt Abbildung 8.13 zusätzlich für alle vier Trajektorien die Zeitreihen der Dihedralwinkel ϕ und ψ .

Für die Trajektorie 1 weist Abbildung 8.13 im Zeitbereich zwischen etwa 40 ps und etwa 120 ps einen langsamen und kontinuierlichen Übergang des Ala 1 ψ von einer gestreckten zu einer turn-artigen Geometrie aus, wie sie in der *trans*-Konformation des Peptids stets anzutreffen ist. Abbildung 8.12 macht deutlich, dass eine solche Krümmung beim Ala 1 und auch beim Phe 9 (für die Lage der Residuen siehe auch Abbildung 1.14) notwendig ist, um die durch den Farbstoff vorgegebenen Randbedingungen zu erfüllen. Gleichzeitig mit dieser für den Übergang des Peptids in Richtung auf die *trans*-Konformation günstigen Änderung der Geometrie des Ala 1 ist jedoch auch eine Umordnung am Gly 7 festzustellen. Dieses Residuum ist zwar nicht sehr strukturbestimmend, jedoch zeigt der in Schnappschuss bei 120 ps aus Abbildung 8.12, dass die Peptidkonformation zu diesem Zeitpunkt, übereinstimmend mit der obigen Charakterisierung, immer noch sehr *cis*-artig erscheint. Die gleichzeitigen Änderungen an den Winkeln Ala 1 ψ (grüner Pfeil in Abbildung 8.12) und Gly 7 ψ (roter Pfeil) führen somit zu dem oben bei T1 beobachteten Übergang des Peptids von einer *cis*-Konformation zu einer anderen.

Vergleicht man in Abbildung 8.13 die am Ende der Trajektorie 1 vorliegenden Winkelwerte mit der dort eingezeichneten Vergleichswerten der mittleren *trans*-Geometrie, so kann man mehrere Residuen identifizieren, bei denen auch nach einer Nanosekunde der Dihedralwinkel ψ noch nicht relaxiert ist: das Cys 2, das Asp 6, das schon erwähnte Gly 7 und schließlich das an den Farbstoff angebundene Phe 8. In Abbildung 8.12 werden diese Residuen bei der Struktur V mit schwarzen Pfeilen markiert. Der Umstand, dass hier vier von 16 Dihedralwinkeln am Ende der Trajektorie noch von der *trans*-Struktur abweichen, erklärt auch, weshalb Abbildung 8.10 für T1 nur eine verschwindende *trans*-Artigkeit ausweist.

Wendet man sich der Trajektorie 5 (blau) zu, so wurde in deren Zeitverlauf schon nach etwa 7 ps ein Verlassen des *cis*-Konformationsraumes beobachtet (vgl. Abbildung 8.10). Abbildung 8.13 lässt erkennen, dass zu diesem Zeitpunkt der Winkel Phe 8 ψ eine relativ abrupte Veränderung erfährt, nach welcher er einen *trans*-konformen Wert annimmt. sehr interessant ist auch der zweite sprunghafte Anstieg der *trans*-Artigkeit zwischen $t = 70$ ps (Struktur IV) und der Struktur V. Laut Abbildung 8.13 ist dieser Anstieg auf einen plötzlichen Peptid-Flip (vgl. Abbildung 7.1) zurückzuführen, der nach etwa 110 ps an den Winkeln Thr 4 ψ und Cys 5 ϕ stattfindet. Nach Abbildung 8.12 führt dieser Flip direkt von einer recht helikalen Konformation zur gestreckten *trans*-Konformation des Peptids. Abbildung 8.13 zeigt dementsprechend, dass bei T5 nach dem Flip im Peptidrückgrat keinerlei Abweichungen der Dihedralwinkel von der mittleren *trans*-Struktur mehr existieren.

Für die Trajektorie 9 (grün) können keine gleichermaßen klare Aussagen gemacht werden, da durch den oben beobachteten Überlapp der Konformationsräume in den von dieser Trajektorie durchlaufenen Winkelbereichen ein sehr empfindlicher funktionaler Zu-

sammenhang zwischen den Dihedralwinkeln und der durch die Klassifikation angegebenen *trans*-Artigkeit besteht. Man erkennt jedoch, dass der in Abbildung 8.10 beobachtete Anstieg der geometrischen Observablen im Zeitbereich vor 300 ps vor allem auf eine graduelle Änderung an den Winkeln Cys 2 ψ und Ala 3 ψ zurückzuführen ist. In Abbildung 8.12 entspricht dieser Vorgang einer Auflösung der helikalen Peptidstruktur zwischen den Schnappschüssen III und IV (grüner Pfeil). Das spätere steile Abfallen der *trans*-Artigkeit ist schließlich auf einen Peptid-Flip an den Winkeln Asp 6 ψ und Gly 7 ϕ zurückzuführen. Es bleibt anzumerken, dass in T9 die Struktur V des cAPBs nach einer Nanosekunde etwas ungewöhnlich erscheint, da in ihr der Peptidteil stark nach „rechts“ bezüglich des Farbstoffs verschoben ist.

Die Schnappschüsse (Abbildung 8.12) aus der Trajektorie 10 zeigen schließlich auf eingängliche Weise, wie die eine Loop-Struktur aufweisende *cis*-Konformation durch den vom Farbstoff ausgeübten Zug aufgelöst und in eine gestreckte Konformation überführt wird. Bei den Dihedralwinkeln sind diese Übergänge laut Abbildung 8.13 nicht mit Peptid-Flips sondern vielmehr mit einer größeren Zahl von relativ unabhängigen schnellen Veränderungen an einzelnen Winkeln verbunden. So wird der erste Anstieg der *trans*-Artigkeit bei $t = 30$ ps durch die gleichzeitige Relaxation der Winkel Cys 2 ψ und Ala 3 ψ bewirkt, während der zweite Anstieg bei etwa $t = 110$ ps auf eine Geometrieänderung des Gly 7 zurückzuführen ist. Am Ende der Trajektorie 10 bleiben Abweichungen von der *trans*-Konformation beim dann noch immer helikal erscheinenden Asp 6 und wiederum beim Phe 8 bestehen.

Zusammenfassend zeigt die detaillierte Untersuchung der Einzeltrajektorien, dass die Konformationsübergänge, welche die Relaxation des Peptids hin auf eine *trans*-Konformation ausmachen, vielfältiger Natur sind. Man beobachtet sowohl graduelle und langsame Veränderungen an einzelnen Residuen, als auch plötzliche Sprünge an einzelnen Dihedralwinkeln und sogar korrelierte Sprünge mehrerer Winkel, letztere teilweise in der Form ganzer Peptid-Flips. Solche Flips führen in der hier verwendeten Darstellung der Peptidkonfiguration durch die Dihedralwinkel des Peptidrückgrats natürlich zur Feststellung großer konformeller Änderungen.

Weiterhin hat dieser Abschnitt deutlich gemacht, in welcher Weise diese Konformationsübergänge vom Azobenzol-Farbstoff getrieben werden. Es bestätigt sich, dass der Farbstoff die Rolle einer Feder spielt. Zuletzt bleibt die Beobachtung hervorzuheben, dass sich die nach einer Nanosekunde verbleibenden Abweichungen der Peptidstrukturen von der vollständig relaxierten *trans*-Struktur zumeist auf die Residuen Phe 8 bzw. Gly 7 beziehen. Es ist anzunehmen, dass dies einerseits auf die Rolle des Phe 8 als ein „Gelenk“ zurückzuführen ist, in dem das Peptidrückgrat in der *trans*-Konformation um fast 180° zurückgefaltet wird, andererseits aber auch eventuell die große Flexibilität des Gly 7 eine Rolle spielt. Auf der anderen Seite des Peptids scheint der stets graduell erfolgende Übergang des mit einer kleineren Seitenkette behafteten Ala 1 dagegen im Hinblick auf das Erreichen der *trans*-Konformation unproblematischer zu sein.

9 Zusammenfassung und Ausblick

Wir haben in dieser Arbeit einen Ansatz zur Validierung von molekularmechanischen (MM) Kraftfeldern kennengelernt, der darin besteht, die Molekulardynamik (MD) der Relaxation kleiner lichtschtbarer Modellpeptide zu simulieren und die dabei auftretenden Kinetiken mit Ergebnissen der Femtosekunden-Spektroskopie zu vergleichen.

Erste Untersuchungen dieser Art zeigten überraschenderweise, dass die Relaxationsdynamik der Modellpeptide nach der *cis-trans*-Isomerisierung des eingebauten Azobenzol-Farbstoffs durch MD-Simulationen mit der ersten Wahl eines MM-Kraftfelds quantitativ richtig wiedergegeben wurde [138]. Dieses in Bezug auf die Validierung der Kraftfelder enttäuschende Ergebnis ließ keinen Rückschluss auf mögliche Mängel der Parametrisierung zu. Nachfolgende Untersuchungen zeigten jedoch, dass der erste Eindruck täuschte. Vielmehr trat in den ersten Simulationen eine glückliche Kompensation von Fehlern auf. Exakt die dort verwendete Kombination eines expliziten und flexiblen DMSO-Modells, eines Peptidkraftfelds, welches eine Präferenz für gestreckte Peptidstrukturen aufweist, mit einer Farbstoffparametrisierung, die eine starke Feder bildet, hatte eine Kinetik zur Folge, die zufälligerweise mit der experimentell beobachteten übereinstimmte.

Ein sehr überraschendes Ergebnis der nachfolgenden Parametrisierungsstudien war die Erkenntnis, dass die Modellierung des Lösungsmittels DMSO einen entscheidenden Einfluss auf die beobachteten Kinetiken hat. So konnte an der Photoisomerisierung des Azobenzols in Lösung beobachtet werden, dass das flexible und explizite Modell HC-DMSO zu schnell kühlt, während die kompakteren „united-atom“ Modelle (das flexible MI-DMSO und das rigide MP-DMSO) den Farbstoff zu lange heiß bleiben lassen. Überraschend war nun, dass die Kühleigenschaften des eingesetzten DMSO-Modells die Kinetik der Konformationsrelaxation des vollständigen Modellpeptids noch zu Zeiten beeinflussen, bei denen das Peptid längst auf Umgebungstemperatur abgekühlt ist. Dies lässt den Schluss zu, dass während der kurzen, etwa 15 ps dauernden Abkühlperiode, welche der Photoisomerisierung nachfolgt, die Weichen für die sich anschließende Konformationsrelaxation des Peptidrückgrats in unterschiedlicher Weise gestellt werden können: Eine verlängerte Abkühlperiode ermöglicht offensichtlich eine Anfangsrelaxation in Zustände, von denen aus der vollständig relaxierte *trans*-Zustand durch Zimmertemperatur-Fluktuationen um Größenordnungen schneller erreicht werden kann, als von jenen Zuständen, die nach einem rapiden Abkühlprozess besetzt sind.

Damit könnte die Ausdehnung von Kühlzeiten ein mögliches Optimierungskriterium und Designprinzip photobiologischer Prozesse in Proteinen sein, wenn unterstellt wird, dass die dabei ablaufenden Konformationsdynamiken möglichst schnell zum jeweiligen Ziel-

zustand führen sollen.

Auch Veränderungen an zwei zusätzlich untersuchten Kraftfelddetails, der Stärke der durch den Farbstoff gebildeten Feder und der Präferenz des Peptids für eine gestreckte Struktur, haben jeweils deutliche Auswirkungen auf die Kinetiken der Peptidrelaxation. Dabei sind die Auswirkungen einzelner Parameteränderungen voneinander abhängig. Daher führt jede Kombination von Kraftfelddetails und einem DMSO-Modell zu spezifischen Veränderungen all jener Prozesse, aus denen sich die gesamte Relaxationsdynamik der Modellpeptide zusammensetzt. Hier hat sich gezeigt, dass die steifere Feder (A) die bessere Wahl darstellt, da die Photoisomerisierung andernfalls zu schnell, die Konformationsdynamik dagegen zu langsam abläuft. Die Kombination der steifen Feder mit dem schnell kühlenden HC-DMSO führt jedoch zu einem „Trapping“ der Konformationsdynamik des Modellpeptids, sofern das Standardkraftfeld für den Peptidteil angenommen wird. Erst durch den Einsatz des veränderten Peptidkraftfelds mit seiner Präferenz für die Zielstruktur der Konformationsübergänge können, wie in den ersten Simulationen [138], die experimentellen Relaxationszeiten reproduziert werden.

Daher kann vermutet werden, dass ein Gesamtkraftfeld aus einem langsamer kühlenden DMSO-Modell, der steifen Feder und der Standardparametrisierung des Peptids zu akzeptablen Ergebnissen in der Relaxationsdynamik führen könnte. Dabei sollte die Kühldynamik jedoch schneller sein als bei den steifen DMSO-Modellen, damit sich ihre Zeitkonstante von jener der Konformationsdynamik, wie im Experiment, noch hinreichend unterscheidet. Andernfalls wäre eine Trennung dieser Prozesse durch Beobachtung der Energierelaxation nicht möglich.

Hervorzuheben ist, dass die hier und in Kapitel 8 präsentierten Einsichten in die komplexen Auswirkungen der Kraftfelddetails nur gewonnen werden konnten, nachdem die in den Kapiteln 3 bis 6 vorgestellten umfangreichen Vorarbeiten gewissenhaft geschaffen worden waren. Zu diesen Vorarbeiten gehört sowohl die Parametrisierung des flexiblen „all-atom“ Modells des DMSOs als auch die Charakterisierung einer Vielzahl von DMSO-Modellen. Beide waren notwendig, um die überragende Rolle des DMSO Lösungsmittelmodells für die Steuerung der Kinetik der Konformationsrelaxation erkennen zu können. Ferner ermöglichte es erst die aufwändige Modellierung des Farbstoffs und seiner lichtinduzierten Isomerisierungsreaktion, die schnellsten Prozesse der Relaxationsdynamik der Modellpeptide in den MD-Simulationen quantitativ zu beschreiben. Schließlich konnte erst durch die erarbeiteten Methoden der Datenanalyse der Zusammenhang zwischen den Prozessen der strukturellen Relaxation der Peptide und den bei der Energierelaxation auftretenden Zeitkonstanten aufgeklärt werden.

Neben den in Kapitel 7 diskutierten Untersuchungen der Gleichgewichtskonformationen der Modellpeptide, mit denen Probleme der NMR-Strukturaufklärung flexibler Moleküle analysiert werden konnten, stellen auch die diversen Vorarbeiten, jede für sich, eigenständige Ergebnisse dar. So hat die hybride DFT/MM-Beschreibung des DMSOs die Leistungsfähigkeit dieser Methode zur Untersuchung von Polarisierungseffekten gezeigt. Ferner führte der Versuch, die thermodynamischen Eigenschaften dieses Lösungsmittels zu

reproduzieren, zu einer Validierung unseres Simulationsprogramms EGO-MMII und zur Einführung der Korrektur der langreichweitigen VDW-Energiebeiträge. Schließlich wurde im Rahmen der vorliegenden Arbeit ein automatisiertes Verfahren zu Optimierung der VDW-Parameter eines Lösungsmittels implementiert. Dieses Verfahren könnte beispielsweise in Zukunft dazu genutzt werden, um ein weiteres, im oben angesprochenen Sinne verbessertes DMSO-Modell zu parametrisieren. Basierend auf den festgestellten Zusammenhängen zwischen der Bauart („all-atom“ bzw. „united-atom“) und der Anzahl der Freiheitsgrade eines DMSO-Modells auf der einen und dessen Kühlverhalten auf der anderen Seite ist zu vermuten, dass ein „all-atom“ DMSO-Modell, bei dem zusätzlich zu den Wasserstoff-Streckschwingungen weitere Freiheitsgrade eingefroren werden (beispielsweise durch die Annahme rigider Methylgruppen), die experimentellen Kühlzeiten reproduzieren könnte.

Auch stellt die Anwendung der neuen Methoden zur statistischen Auswertung von MD-Trajektorien, einschließlich ihrer in Kapitel 5 vorgestellten Erweiterungen, ein eigenständiges Ergebnis dar. Es ist zu erwarten, dass diese Auswertungsmethoden in Zukunft besonders bei der Analyse von MD-Simulationen größerer und noch komplexer Biomoleküle eine wichtige Rolle spielen werden.

A Optimierung von VDW-Parametern mit der Simplex-Methode

Die Simplex-Methode [199] ist ein einfach zu implementierender Algorithmus zur Minimierung einer skalaren Funktion $f(x_1, \dots, x_d)$ von d Variablen x_i . Diese Methode ist sehr robust in der Lage, ein lokales Minimum von f zu finden, benötigt aber dazu üblicherweise mehr Schritte als andere Methoden. Ein Simplex ist Menge von $d + 1$ Punkten im d -dimensionalen Raum, die einen konvexen Körper aufspannen. Für drei Dimensionen zeigt Abbildung A.1 diesen Körper als einen „verformten Tetraeder“. In jedem Schritt des Algorithmus wird f an allen Punkten des Simplexes ausgewertet und derjenige mit dem höchsten Funktionswert bestimmt. Dieser Punkt wird anschließend durch die Hyperebene der anderen Punkte reflektiert. Ist diesem neuen Punkt ein Funktionswert zugeordnet, der niedriger als der bisher kleinste ist, so führt der Algorithmus zusätzlich zu der Spiegelung eine Ausdehnung in die neue Richtung durch. Ist jedoch der neue Funktionswert größer als alle anderen Werte, so wird statt der Spiegelung das Volumen des Simplex durch eine Kontraktion verkleinert. Visualisiert man die Abfolge der vom Simplex-Algorithmus evaluierten Funktionswerte, so ähnelt das Verhalten dem einer Amöbe, die sich durch kleine Öffnungen zwängt aber auch breite Täler ausfüllen kann, um flache Bereiche von f effizient abzudecken.

Die Anbindung der Simplex-Methode an EGOMMII basiert auf einer Sammlung von Programmskripten, die von der Gruppe von Florian Müller-Plathe (jetzt International Uni-

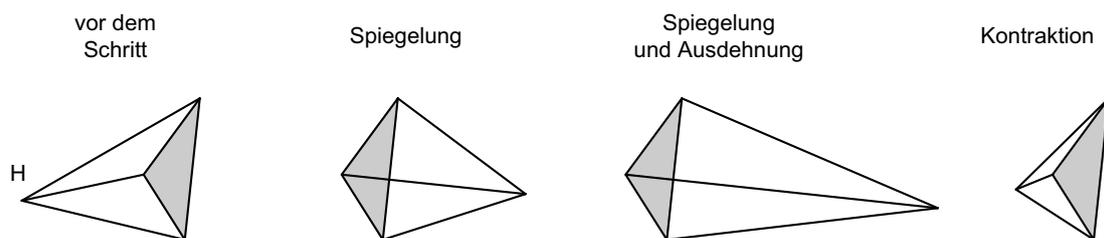


Abbildung A.1: Der Simplex-Optimierungsalgorithmus in zwei Dimensionen. Der aus vier Punkten bestehende Simplex und die möglichen Schritte der Simplex-Methode werden skizziert. Vor dem Schritt wird der Eckpunkt mit dem höchsten Funktionswert mit H markiert und die durch die anderen Punkte definierte Hyperebene grau hervorgehoben. Bei der Spiegelung wird der Punkt H durch die Hyperebene reflektiert, das Volumen des Simplex bleibt konstant. Ausdehnung und Kontraktion sind lineare Transformationen, die zu keiner Entartung des Simplex führen können, jedoch sein Volumen vergrößern bzw. verkleinern.

versity of Bremen) zur Verfügung gestellt wurde. Neben einer Anpassung der Skripte an die Formate der Dateien, die im Fall von EGOMMII die Parametrisierung des Simulationssystems definieren, mussten vor allem ein automatisierte Verfahren zur Auswertung der Simulationen erstellt werden. Das dabei gewählte Vorgehen zur Berechnung der thermodynamischen Observablen entspricht jenem, welches in Abschnitt 3.2 beschrieben wird. Zur Automatisierung des Verfahrens wurde der Optimierungsalgorithmus schließlich an das so genannte „Batch queueing“ System unseres Rechenclusters angebunden, welches eine automatische Verwaltung von Rechenjobs und ihrer Abhängigkeiten untereinander ermöglicht.

B Korrektur der langreichweitigen VDW-Beiträge

Gleichung (2.23) zeigte die vorgenommene Aufspaltung der VDW-Wechselwirkung in einen explizit zu berechnenden Anteil für Abstände kleiner als r_{cut} , dem Cutoff-Radius, und in den zu korrigierenden langreichweitigen Anteil $E_{\text{VDW,corr}}$. Um den Korrekturterm zu berechnen, geht man davon aus, dass die Paarverteilungsfunktion $g(r)$ aller Atomsorten im Bereich außerhalb des Cutoffs keine Strukturen mehr aufweisen, also $g(r) \approx 1$ für $r \geq r_{\text{cut}}$ gilt. Die Korrektur kann dann für ein, nur aus einer Teilchensorte bestehendes, System als Integral über die Kugelschalen

$$E_{\text{VDW,corr}} = 2\pi N\rho \int_{r_{\text{cut}}}^{\infty} r^2 E_{\text{VDW}}(r) dr \quad (\text{B.1})$$

geschrieben werden [74]. N ist dabei die Anzahl der Teilchen, ρ die Teilchendichte und $E_{\text{VDW}}(r)$ das Lennard-Jones Potential der Paarwechselwirkung. Mit der Definition dieses Potentials (Gleichung (2.20)) und unter der gerechtfertigten Vernachlässigung der mit r^{-12} schnell abfallenden repulsiven Anteile erhält man

$$E_{\text{VDW,corr}} = \frac{8\pi N\rho \varepsilon \sigma^6}{3 r_{\text{cut}}^3}, \quad (\text{B.2})$$

mit den VDW-Parametern ε und σ . In einem Simulationssystem, das N_{types} verschiedenen Atomsorten enthält, ist über alle Sortenpaare i, j zu summieren

$$E_{\text{VDW,corr}} = \frac{8\pi}{3 r_{\text{cut}}^3 V(t)} \sum_{i \leq j}^{N_{\text{types}}} \varepsilon_{ij} \sigma_{ij}^6 N_i N_j. \quad (\text{B.3})$$

Im letzten Schritt wurde zusätzlich ρ durch N/V ersetzt, was eine in i und j symmetrischen Darstellung zur Folge hat. Von Vorteil ist, dass in dieser Gleichung (B.3) alleine das Volumen V eine Zeitabhängigkeit aufweist. Deshalb muss die Berechnung der Summe nur einmal zu Beginn einer Simulation durchgeführt werden und die Berechnung der Korrektur-Energie beschränkt sich auf eine einfache Division.

Wie in Abschnitt 2.2 dargestellt wurde, tritt als Besonderheit bei Verwendung der FAMU-SAMM-Methode kein einfach definierter atombasierter Cutoff-Radius auf. Vielmehr wird auf der Ebene der strukturellen Einheiten ein Abstandskriterium auf die Distanz ihrer Massenschwerpunkte angewendet, das auch ihre jeweiligen Gyrationradien berücksichtigt.

B Korrektur der langreichweitigen VDW-Beiträge

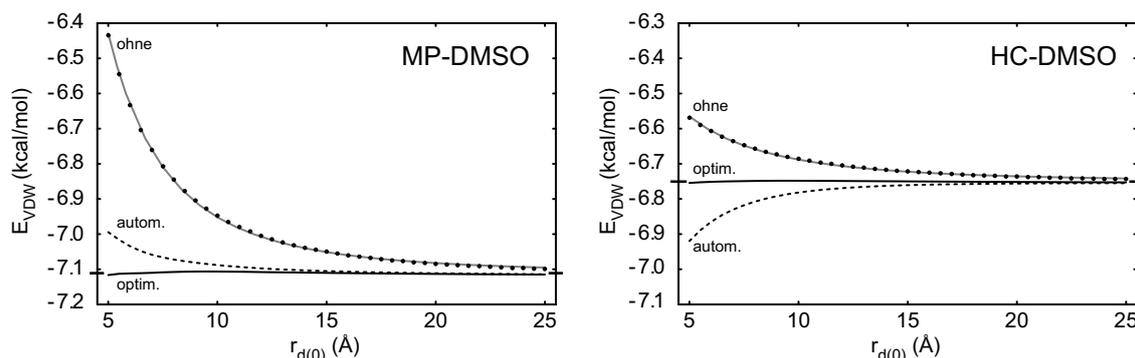


Abbildung B.1: Die optimale Korrektur der VDW-Wechselwirkung. Für zwei der in Kapitel 3 vorgestellten DMSO-Modelle, dem Pseudoatome verwendenden MP-DMSO (links) und dem expliziten HC-DMSO (rechts), wird das Vorgehen zur optimalen Korrektur der langreichweitigen Anteile der VDW-Wechselwirkung illustriert. Die x-Achse zeigt $r_{d(0)}$, den gewählten Abstand der strukturellen Einheiten (Cluster-Level 0), der in den Standardeinstellungen von EGOMMII 6 Å beträgt. Schwarze Punkte stellen den aus Einzelpunktrechnungen eines Testsystems bei abgeschalteter Korrektur bestimmten VDW-Beitrag pro Molekül dar. Die daran angepasste Funktion der Form (B.3) ist grau dargestellt. Die VDW-Beiträge mit Korrektur sind für den Fall der automatischer Wahl des mittleren Gyrationradiuses (gestrichelt) und den der Verwendung des optimalen $r_{gyr,eff}$ (durchgezogen) gezeigt. An den y-Achsen ist der Wert von E_{VDW} markiert, der asymptotisch für $r_{d(0)} \rightarrow \infty$ angenommen wird. Details und Diskussion im Text.

Auch wird durch den neu eingeführten Überlapp der Distanzklassen die VDW-Wechselwirkung [100] nicht ab einem bestimmten Abstand hart abgeschaltet, sondern existiert ein Übergangsbereich der Breite $r_{overlap}$, in dem zwischen der vollen Wechselwirkung und der Nicht-Berücksichtigung mittels einer kubischen Funktion interpoliert wird. Da die Korrektur der VDW-Wechselwirkung unabhängig von der räumlichen Anordnung der Atome oder der strukturellen Einheiten im System sein soll, wurde ein effektiver mittlerer Abschneideradius \tilde{r}_{cut} abgeschätzt. Dieser hängt von $r_{d(0)}$, dem vom Anwender vorgegebenen Abstand der strukturellen Einheiten auf Clusterebene 0, von der halben Breite des Überlappbereichs $r_{overlap}$ und auch von dem mittleren Gyrationradius der strukturellen Einheiten $\langle r_{gyr(0)} \rangle$ ab.

$$\tilde{r}_{cut} = r_{d(0)} + \frac{1}{2}r_{overlap} + 2 \begin{cases} \langle r_{gyr(0)} \rangle & \text{automatisch} \\ r_{gyr,eff} & \text{optimiert} \end{cases} \quad (\text{B.4})$$

Der mittleren Gyrationradius kann alternativ vom Anwender durch einen effektiven Gyrationradius $r_{gyr,eff}$ ersetzt werden, der in einer externen Optimierung zu bestimmen ist. Abbildung B.1 illustriert das Vorgehen, für das zuerst Einzelpunktberechnungen des VDW-Beitrags ohne jegliche Korrektur durchgeführt werden müssen, wobei der Abstand der strukturellen Einheiten $r_{d(0)}$, und damit der effektive Abschneideradius \tilde{r}_{cut} variiert wird. Die, in Abbildung B.1 gezeigten, großen Werte für $r_{d(0)}$ wurden durch Missachtung der

| | DMSO | | | | | Wasser |
|---|------|----------|------|------|------|--------|
| | OPLS | VG/MP/MI | RESP | SF | HC | TIP3 |
| $\langle r_{\text{gyr}(0)} \rangle$ [Å] | 1,38 | 1,46 | 1,79 | 1,83 | 1,81 | 0,67 |
| $r_{\text{gyr,eff}}$ [Å] | 1,21 | 1,19 | 2,07 | 2,21 | 2,91 | 2,00 |

Tabelle B.1: Effektive Gyrationradien zur VDW-Korrektur. Die mittleren $\langle r_{\text{gyr}(0)} \rangle$ und die optimierten effektiven Gyrationradien $r_{\text{gyr,eff}}$ der verschiedenen, in Kapitel 3 diskutierten, DMSO-Modelle werden gezeigt. Zur Vollständigkeit sind auch die Werte für das, an das CHARMM22-FF [81] angepasste, TIP3-Modell [292] des Wassers aufgeführt.

Minimum-Image Konvention erreicht, deren Erfüllung hier nicht notwendig ist. Die auf diese Weise entstehenden Kurven zeigen die erwartete invers-kubische Abhängigkeit von $r_{\text{d}(0)}$. Somit ist es möglich, die durch die Gleichungen (B.3) und (B.4) gegebene Funktion an diese Werte anzupassen. Für die zwei freien Parameter, den Grenzwert von E_{VDW} für $r_{\text{d}(0)} \rightarrow \infty$ und den effektive Radius $r_{\text{gyr,eff}}$, wurden mit Hilfe des Programmpakets MATHEMATICA jeweils optimale Werte bestimmt. Tabelle B.1 zeigt die Ergebnisse dieses Vorgehens für die untersuchten DMSO-Modelle und das TIP3-Wassermodell. Die verwendeten Konfigurationen des Simulationssystems wurden aus Trajektorien mit $T = 289$ K und Normaldruck entnommen und ein Distanzklassen-Überlapp von $r_{\text{overlap}} = 1,25$ Å verwendet.

Die kompakten vieratomigen DMSO-Modelle (OPLS, VG, MP und MI, siehe Kapitel 3) weisen eine große geometrische Ähnlichkeit auf und besitzen deshalb auch die gleichen mittleren $r_{\text{gyr}(0)}$ von etwa $1,4$ Å. Dieser Wert ist gut vergleichbar mit dem der angepassten optimalen Radien, der für diese Modelle einheitlich $1,2$ Å beträgt. Abbildung B.1 zeigt dazu passend, dass E_{VDW} mit der automatischen Korrektur im Fall des MP-DMSOs für $r_{\text{d}(0)} > 8$ Å einen ausreichend flachen Verlauf aufweist. Mit den optimierten Radien wird eine fast perfekte Korrektur ermöglicht, so dass die Abhängigkeit der VDW-Wechselwirkung von $r_{\text{d}(0)}$ praktisch aufgehoben wird. Der Fall der expliziten DMSO-Modelle weist allerdings darauf hin, dass die automatische Korrektur nicht immer anwendbar ist. Trotz ähnlicher $r_{\text{gyr}(0)}$, führt Tabelle B.1 für die drei Parametrisierungen (RESP, SF und HC) stark unterschiedliche optimale Radien auf. Diese hängen offensichtlich in einer komplizierten Weise von der Geometrie der Moleküle, den VDW-Parametern und Details der Bestimmung der Wechselwirkungsklassen in EGOMMII ab. Das rechte Diagramm in Abbildung B.1 zeigt eine Überkorrektur bei Verwendung des geometrisch bestimmten Radiuses, die von der gleichen Größenordnung wie der zu behebende Fehler ohne Korrektur ist. Auch hier führt der optimierte Radius zu einem sehr flachen Kurvenverlauf.

C Geometrien der Konformationen

C Geometrien der Konformationen

| K. | Ala 1 | | | Cys 2 | | | Ala 3 | | | Thr 4 | | | Cys 5 | | | Asp 6 | | | Gly 7 | | | Phe 8 | | |
|---------------------|--------|--------|----------|--------|--------|----------|--------|--------|----------|--------|--------|----------|--------|--------|----------|--------|--------|----------|--------|--------|---|--------|--------|----------|
| | ϕ | ψ | S | ϕ | ψ | S | ϕ | ψ | S |
| cis-cAPB | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | -105 | -171 | β | -100 | -54 | α | -99 | -75 | α | -130 | 164 | β | -128 | -170 | β | -98 | 83 | β | 172 | -152 | / | -100 | 146 | β |
| 2 | -103 | 168 | β | -100 | 104 | β | -106 | -50 | α | -121 | 174 | β | -92 | -63 | α | -101 | 111 | β | 180 | -175 | / | -107 | 165 | β |
| 3 | -93 | 138 | β | -114 | -66 | α | -99 | -42 | α | -114 | 159 | β | -96 | -35 | α | -89 | 120 | β | 176 | -168 | / | -102 | -58 | α |
| 4 | -101 | -90 | α | -107 | -66 | α | -100 | -57 | α | -123 | -177 | β | -121 | -150 | β | -98 | 118 | β | 119 | -113 | / | -102 | -48 | α |
| 5 | -118 | 169 | β | -109 | -47 | α | -101 | -43 | α | -88 | 124 | β | -95 | 113 | β | -129 | 171 | β | -133 | 39 | / | -106 | -83 | α |
| 6 | -106 | -75 | α | -121 | -154 | β | -87 | -64 | α | -89 | -48 | α | -99 | 94 | β | -110 | -60 | α | 136 | -72 | / | -105 | -58 | α |
| 7 | -110 | 178 | β | -86 | 67 | β | -121 | -96 | α | -95 | 104 | β | -101 | 142 | β | -101 | -41 | α | 149 | -58 | / | -102 | -61 | α |
| 8 | -116 | -158 | β | -95 | -52 | α | -105 | 154 | β | -93 | 123 | β | -107 | -67 | α | -100 | -37 | α | -174 | -149 | / | -95 | 156 | β |
| 9 | -113 | -171 | β | -96 | -56 | α | -102 | 139 | β | -109 | 149 | β | -109 | -115 | α | -99 | 100 | β | 177 | -176 | / | -99 | -60 | α |
| 10 | -92 | -54 | α | -108 | -174 | β | -91 | 155 | β | -90 | 113 | β | -95 | -65 | α | -107 | 157 | β | -105 | 90 | / | -96 | -52 | α |
| 11 | -107 | 174 | β | -96 | -52 | α | -96 | 101 | β | -119 | 170 | β | -117 | -144 | α | -96 | 120 | β | 172 | -171 | / | -99 | 141 | β |
| 12 | -98 | -55 | α | -108 | 162 | β | -97 | -63 | α | -97 | 129 | β | -112 | -92 | α | -94 | -26 | α | -145 | -155 | / | -95 | 140 | β |
| trans-cAPB | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | -113 | -65 | α | -101 | -61 | α | -123 | 171 | β | -120 | 175 | β | -124 | 179 | β | -108 | 158 | β | 114 | -112 | / | -93 | -36 | α |
| cis-cAMPB | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | -113 | -115 | α | -91 | -49 | α | -95 | 136 | β | -94 | -73 | α | -83 | -58 | α | -88 | 93 | β | 142 | -59 | / | -106 | -68 | α |
| 2 | -104 | -80 | α | -106 | -53 | α | -104 | 138 | β | -100 | -65 | α | -101 | -56 | α | -95 | 77 | β | 176 | -174 | / | -99 | 132 | β |
| 3 | -117 | -159 | β | -104 | -58 | α | -96 | 150 | β | -90 | 127 | β | -110 | -169 | β | -97 | 96 | β | 172 | -174 | / | -99 | 156 | β |
| 4 | -113 | -151 | β | -95 | -54 | α | -102 | 141 | β | -94 | 133 | β | -108 | -76 | α | -96 | -16 | α | -170 | -171 | / | -95 | 150 | β |
| 5 | 73 | -74 | α | -104 | -160 | β | -98 | -170 | β | -64 | 83 | β | -107 | -65 | α | -95 | -19 | α | -152 | -160 | / | -101 | -61 | α |
| 6 | -101 | -162 | β | -67 | -58 | α | -75 | -27 | α | -77 | -42 | α | -74 | -54 | α | -83 | 45 | β | -173 | -54 | / | -103 | -61 | α |
| 7 | -87 | -63 | α | -79 | -25 | α | -60 | -44 | α | -68 | -36 | α | -73 | -37 | α | -84 | -39 | α | 128 | 31 | / | -111 | 172 | β |
| trans-cAMPB | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | -114 | -74 | α | -99 | -54 | α | -120 | 174 | β | -108 | 164 | β | -123 | 172 | β | -106 | 146 | β | 160 | -129 | / | -96 | -44 | α |
| 2 | -102 | -57 | α | -108 | 170 | β | -89 | 97 | β | -132 | 173 | β | -103 | 139 | β | -108 | -66 | α | 155 | -95 | / | -98 | -52 | α |
| 3 | -114 | -72 | α | -100 | -59 | α | -120 | 172 | β | -115 | 173 | β | -116 | 168 | β | -98 | 113 | β | 167 | -161 | / | -95 | 114 | β |
| cis-bcAMPB | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | -104 | -75 | α | -109 | 111 | β | -74 | -47 | α | -74 | -20 | α | -80 | -48 | α | -107 | -60 | α | 154 | -163 | / | -92 | 139 | β |
| 2 | -98 | -57 | α | -109 | 122 | β | -76 | -52 | α | -70 | -20 | α | -73 | -36 | α | -94 | -52 | α | 151 | -87 | / | -102 | -55 | α |
| 3 | -101 | -53 | α | -109 | -165 | β | -83 | -53 | α | -87 | -30 | α | -85 | 136 | β | -129 | -69 | α | 134 | -57 | / | -99 | -61 | α |
| 4 | -90 | 86 | β | -128 | 150 | β | -91 | -59 | α | -88 | 37 | β | -110 | 165 | β | -96 | 172 | β | -96 | 93 | / | -101 | -63 | α |
| trans-bcAMPB | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | -97 | -41 | α | -95 | 133 | β | -84 | -46 | α | -80 | -5 | α | -83 | 11 | α | -101 | 174 | β | -179 | -90 | / | -94 | -47 | α |
| 2 | -101 | -35 | α | -89 | 140 | β | -88 | -51 | α | -89 | 28 | α | -103 | 133 | β | -103 | 166 | β | -180 | -171 | / | -95 | -51 | α |
| 3 | -99 | -36 | α | -88 | 135 | β | -87 | -46 | α | -89 | 14 | α | -95 | 124 | β | -103 | 160 | β | 175 | 179 | / | -94 | 115 | β |

Tabelle C.1: Die Geometrien der Peptidkonformationen. Für die in den Gleichgewichtssimulationen beobachteten und in Abbildung 7.7 gezeigten Konformationen werden die ϕ/ψ -Winkel des Peptidrückgrats in Grad aufgeführt. Zusätzlich gibt eine weitere Spalte „S“ die jeweilige Zuordnung eines Residuums zu einem der beiden Sekundärstruktur motive „ α -helikal“ oder „ β -faltblattartig“ an. Diese Zuordnung geschah analog zu dem von Hu et al. [289] gewählten Vorgehen durch die einfache Betrachtung des Winkels ψ . Dabei bezeichnet ein α diejenigen Residuen, die in einer Konformation ein $\alpha \in [-150^\circ, 30^\circ]$ aufweisen und ein β die übrigen Geometrien.

D Tabellen des MD-NMR Vergleichs

Dieser Anhang beinhaltet in tabellarischer Form diejenigen Daten, die in Abschnitt 7.4 zum Vergleich der in den MD-Simulationen beobachteten Konformationen mit den NMR-Messungen herangezogen wurden. In der jeweils oberen Tabelle werden als r_{exp} die aus den NOE-Intensitäten ermittelten maximalen Proton-Proton Abstände aufgeführt. Die NOEs sind nummeriert und es werden die beteiligten Wasserstoffatome angegeben. Diese Angabe erfolgt in der Form „Residuum:Residuums-Nummer:Wasserstoffe“. Ist mehr als ein Wasserstoff aufgeführt, entweder in der Form „Hx,Hy“ oder durch die Angabe eines Wildcards „H*“, so können die betroffenen Protonen aus Symmetriegründen nicht unterschieden werden und bilden ein Pseudoatom, dessen Ort dem geometrischen Mittel der Orte der beteiligten Protonen entspricht. Die Benennung der Wasserstoffatome folgt beim APB der Abbildung 4.1 und im Rest des Peptides dem Standard von CHARMM22 mit Ausnahme der Bezeichnung „HPHE*“, die alle Wasserstoffe des Phenylrings umfasst.

Den NMR-Restraints r_{exp} werden die mittleren Protonen-Abstände gegenübergestellt, wie sie durch gewichtete Mittelung (7.3) aus den Gleichgewichtstrajektorien (Kapitel 7) berechnet wurden. Neben den Ergebnissen der Mittelung über die gesamte Trajektorie (r_{MD}), wurden die Mittelwerte auch bezüglich der in Abschnitt 7.3 diskutierten Konformationen gebildet. Die mit „Abw.“ bezeichnete Spalte gibt die Abweichung $r_{\text{MD}} - r_{\text{exp}}$ an. Hierbei werden negative Abweichungen nicht aufgeführt und auch nicht als Verletzung des NMR-Restraints gewertet, da die experimentellen Abstände nur eine obere Grenzen angeben können. Alle Abstände und Abweichungen werden in Ångström angegeben.

Für jedes der Modellpeptide gibt die zweite Tabelle die ^3J -Kopplungskonstanten an. Analog zu den NOEs sind diese durchnummeriert und werden durch die Angabe der vier beteiligten Wasserstoffatome bezeichnet. Neben der gemessenen Kopplungskonstanten $^3\text{J}_{\text{exp}}$ wird jeweils der in der Simulation beobachtete Wert $^3\text{J}_{\text{MD}}$ und dessen Abweichung vom experimentellen Wert angegeben. Auch hier wird die Einheit Hertz nicht explizit in den Tabellen aufgeführt.

Die experimentellen Daten wurden für das cAPB und das cAMPB den entsprechenden Publikationen entnommen [37, 38]. Im Fall des bcAMPBs werden unpublizierte Werte der erneut durchgeführten NMR-Messungen verwendet.

D Tabellen des MD-NMR Vergleichs

| NOE | Atom 1 - Atom2 | r_{exp} | r_{MD} | Abw. | Abweichung bez. Konformation | | | | | | | | | | | | |
|-----|---------------------------|-----------|----------|------|------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
| 1 | APB:0:H2,H4 - ALA:1:HN | 4,52 | 2,68 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 2 | APB:0:H1,H5 - ALA:1:HN | 6,71 | 4,72 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 3 | ALA:1:HA - CYS:2:HN | 3,01 | 2,68 | 0,00 | 0,00 | 0,00 | 0,00 | 0,48 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,43 | 0,00 | 0,53 | 0,00 |
| 4 | ALA:1:HB* - CYS:2:HN | 4,07 | 2,86 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 5 | ALA:1:HB* - ALA:3:HN | 5,08 | 4,40 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 1,57 | 0,00 | 1,21 | 0,00 |
| 6 | CYS:2:HA - ALA:3:HN | 2,88 | 2,80 | 0,00 | 0,62 | 0,00 | 0,12 | 0,60 | 0,29 | 0,00 | 0,00 | 0,00 | 0,46 | 0,00 | 0,46 | 0,51 | 0,00 |
| 7 | CYS:2:HN - ALA:1:HN | 3,11 | 2,76 | 0,00 | 0,00 | 1,25 | 1,12 | 0,00 | 1,16 | 0,00 | 0,67 | 0,00 | 0,00 | 0,00 | 0,04 | 0,00 | 0,00 |
| 8 | ALA:3:HA - THR:4:HN | 2,81 | 2,63 | 0,00 | 0,54 | 0,74 | 0,57 | 0,73 | 0,67 | 0,27 | 0,29 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,67 |
| 9 | ALA:3:HB* - THR:4:HN | 4,32 | 3,03 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 10 | ALA:3:HB* - APB:0:H1,H5 | 7,86 | 8,19 | 0,33 | 0,29 | 1,62 | 4,00 | 0,12 | 4,71 | 2,06 | 0,47 | 0,97 | 0,98 | 1,21 | 1,04 | 0,00 | 0,00 |
| 11 | ALA:3:HB* - APB:0:H6,H10 | 8,03 | 8,81 | 0,78 | 1,51 | 2,03 | 4,90 | 1,04 | 6,42 | 3,24 | 2,37 | 0,77 | 3,22 | 0,64 | 3,16 | 0,00 | 0,00 |
| 12 | ALA:3:HB* - APB:0:H7,H9 | 7,21 | 7,82 | 0,61 | 2,48 | 2,39 | 5,10 | 1,96 | 7,06 | 4,29 | 3,68 | 0,32 | 4,44 | 0,00 | 4,77 | 0,00 | 0,00 |
| 13 | ALA:3:HN - CYS:2:HN | 3,09 | 2,54 | 0,00 | 0,00 | 0,77 | 0,00 | 0,00 | 0,00 | 0,00 | 0,46 | 0,00 | 0,00 | 0,76 | 0,00 | 0,29 | 0,00 |
| 14 | THR:4:HA - CYS:5:HN | 2,77 | 2,39 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,30 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 15 | THR:4:HB - CYS:5:HN | 3,30 | 3,09 | 0,00 | 0,29 | 0,35 | 0,13 | 0,12 | 0,75 | 0,00 | 0,00 | 0,00 | 0,00 | 0,60 | 0,00 | 0,22 | 0,00 |
| 16 | THR:4:HG2* - APB:0:H1,H5 | 8,02 | 7,38 | 0,00 | 0,00 | 3,12 | 2,88 | 0,00 | 3,55 | 0,00 | 3,42 | 0,00 | 0,00 | 0,73 | 0,00 | 1,90 | 0,00 |
| 17 | THR:4:HG2* - APB:0:H6,H10 | 7,33 | 7,24 | 0,00 | 0,00 | 3,38 | 3,78 | 0,00 | 4,64 | 0,63 | 4,41 | 0,45 | 0,00 | 0,49 | 0,84 | 1,14 | 0,00 |
| 18 | THR:4:HG2* - APB:0:H7,H9 | 6,79 | 6,27 | 0,00 | 0,00 | 2,35 | 2,65 | 0,00 | 3,85 | 2,21 | 4,64 | 0,00 | 0,00 | 0,00 | 0,31 | 0,00 | 0,00 |
| 19 | THR:4:HN - ALA:3:HN | 3,28 | 2,65 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,81 | 1,11 | 0,44 | 0,00 | 0,00 |
| 20 | THR:4:HN - CYS:5:HN | 3,20 | 3,30 | 0,10 | 0,71 | 1,03 | 1,09 | 0,98 | 1,20 | 0,00 | 0,55 | 0,71 | 1,04 | 0,93 | 0,98 | 1,09 | 0,00 |
| 21 | ASP:6:HA - GLY:7:HN | 2,94 | 2,54 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,04 | 0,51 | 0,12 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 22 | ASP:6:HB1 - THR:4:HG1 | 4,80 | 6,01 | 1,21 | 2,26 | 1,00 | 0,00 | 2,71 | 2,07 | 3,40 | 2,16 | 1,83 | 2,25 | 2,97 | 3,05 | 2,77 | 0,00 |
| 23 | ASP:6:HB2 - THR:4:HG1 | 4,80 | 5,23 | 0,43 | 0,47 | 0,00 | 0,00 | 2,27 | 1,55 | 2,11 | 0,84 | 0,89 | 0,03 | 2,31 | 2,30 | 0,66 | 0,00 |
| 24 | ASP:6:HB1 - GLY:7:HN | 5,30 | 2,86 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 25 | ASP:6:HB1 - PHE:8:HPHE* | 7,10 | 5,94 | 0,00 | 2,14 | 0,83 | 0,85 | 2,48 | 0,00 | 1,37 | 1,95 | 0,00 | 0,00 | 0,00 | 0,00 | 2,75 | 0,00 |
| 26 | ASP:6:HB2 - PHE:8:HPHE* | 7,10 | 5,75 | 0,00 | 2,45 | 2,14 | 2,06 | 2,59 | 0,00 | 0,38 | 0,76 | 0,00 | 0,00 | 0,00 | 0,00 | 3,04 | 0,00 |
| 27 | ASP:6:HB1 - APB:0:H7,H9 | 6,20 | 5,92 | 0,00 | 2,72 | 4,53 | 3,49 | 1,81 | 0,12 | 0,00 | 0,00 | 0,00 | 1,39 | 3,17 | 2,70 | 3,26 | 0,00 |
| 28 | ASP:6:HB2 - APB:0:H7,H9 | 6,78 | 5,99 | 0,00 | 2,53 | 4,76 | 3,82 | 1,09 | 0,29 | 0,00 | 0,00 | 0,00 | 0,99 | 1,93 | 1,52 | 1,55 | 0,00 |
| 29 | GLY:7:HA1 - ASP:6:HA | 4,90 | 4,55 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 30 | GLY:7:HA2 - ASP:6:HA | 5,02 | 4,51 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 31 | GLY:7:HA1 - PHE:8:HN | 3,24 | 2,43 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,10 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 32 | GLY:7:HA2 - PHE:8:HN | 3,24 | 2,72 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,13 | 0,03 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 33 | GLY:7:HN - ASP:6:HN | 3,00 | 2,60 | 0,00 | 0,26 | 0,87 | 1,16 | 0,57 | 1,16 | 0,00 | 0,00 | 0,00 | 0,67 | 0,42 | 0,81 | 0,00 | 0,00 |
| 34 | PHE:8:HA - APB:0:H7,H9 | 6,01 | 4,26 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 35 | PHE:8:HA - APB:0:HN | 2,53 | 2,51 | 0,00 | 0,00 | 0,00 | 0,96 | 0,72 | 0,56 | 0,58 | 1,02 | 0,00 | 0,94 | 0,83 | 0,00 | 0,00 | 0,00 |
| 36 | PHE:8:HB1 - APB:0:H7,H9 | 6,92 | 4,64 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 37 | PHE:8:HB2 - APB:0:H7,H9 | 6,78 | 4,36 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 38 | PHE:8:HB1 - APB:0:HN | 3,47 | 2,76 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,12 | 0,00 | 0,00 |
| 39 | PHE:8:HB2 - APB:0:HN | 3,80 | 2,61 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 40 | PHE:8:HN - APB:0:HN | 3,35 | 2,65 | 0,00 | 0,61 | 0,89 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,79 | 0,21 | 0,00 |
| 41 | PHE:8:HPHE* - APB:0:HN | 6,54 | 4,83 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |

| ³ J Nr. | Dihedral | ³ J _{exp} | ³ J _{MD} | Abw. |
|--------------------|-------------------|-------------------------------|------------------------------|-------|
| 1 | ALA:1:HN-N-CA-HA | 6.20 | 7.56 | 1.36 |
| 2 | CYS:2:HN-N-CA-HA | 7.00 | 7.43 | 0.43 |
| 3 | ALA:3:HN-N-CA-HA | 6.70 | 7.47 | 0.77 |
| 4 | THR:4:HN-N-CA-HA | 7.50 | 7.12 | -0.38 |
| 5 | CYS:5:HN-N-CA-HA | 7.00 | 7.48 | 0.48 |
| 6 | ASP:6:HN-N-CA-HA | 6.50 | 7.69 | 1.19 |
| 7 | GLY:7:HN-N-CA-HA1 | 6.00 | 5.21 | -0.79 |
| 8 | GLY:7:HN-N-CA-HA2 | 5.20 | 4.73 | -0.47 |
| 9 | PHE:8:HN-N-CA-HA | 7.70 | 7.82 | 0.12 |

Tabelle D.1: *cis*-cAPB: Protonenabstände und ³J-Kopplungen.

| NOE | Atom 1 - Atom2 | r_{exp} | r_{MD} | Abw. | Konf. | |
|-----|---------------------------|------------------|-----------------|------|-------|------|
| | | | | | 1 | |
| 1 | APB:0:H2,H4 - ALA:1:HN | 4,70 | 2,65 | 0,00 | 0,00 | 0,00 |
| 2 | ALA:1:HB* - CYS:2:HN | 4,02 | 2,92 | 0,00 | 0,00 | 0,00 |
| 3 | ALA:1:HB* - ALA:3:HN | 5,80 | 5,32 | 0,00 | 0,00 | 0,00 |
| 4 | CYS:2:HA - ALA:3:HN | 2,64 | 3,00 | 0,36 | 0,36 | 0,36 |
| 5 | CYS:2:HB1 - ALA:3:HN | 3,19 | 2,61 | 0,00 | 0,00 | 0,00 |
| 6 | CYS:2:HB2 - ALA:3:HN | 3,42 | 2,70 | 0,00 | 0,00 | 0,00 |
| 7 | CYS:2:HN - ALA:1:HN | 2,63 | 2,29 | 0,00 | 0,00 | 0,00 |
| 8 | ALA:3:HB* - CYS:2:HA | 5,22 | 5,37 | 0,15 | 0,15 | 0,15 |
| 9 | ALA:3:HB* - CYS:2:HN | 5,42 | 5,54 | 0,12 | 0,12 | 0,12 |
| 10 | ALA:3:HB* - THR:4:HN | 3,96 | 3,01 | 0,00 | 0,00 | 0,00 |
| 11 | ALA:3:HB* - APB:0:H1,H5 | 7,02 | 7,61 | 0,57 | 0,57 | 0,57 |
| 12 | THR:4:HA - CYS:5:HN | 2,56 | 2,49 | 0,00 | 0,00 | 0,00 |
| 13 | THR:4:HA - APB:0:H1,H5 | 6,04 | 4,06 | 0,00 | 0,00 | 0,00 |
| 14 | THR:4:HB - CYS:5:HN | 3,09 | 2,92 | 0,00 | 0,00 | 0,00 |
| 15 | THR:4:HB - ASP:6:HN | 4,35 | 6,45 | 2,10 | 2,10 | 2,10 |
| 16 | THR:4:HG2* - ASP:6:HN | 5,09 | 5,69 | 0,60 | 0,60 | 0,60 |
| 17 | THR:4:HG2* - APB:0:H6,H10 | 6,42 | 5,08 | 0,00 | 0,00 | 0,00 |
| 18 | CYS:5:HB1 - ASP:6:HN | 3,28 | 2,64 | 0,00 | 0,00 | 0,00 |
| 19 | CYS:5:HB2 - ASP:6:HN | 3,63 | 2,72 | 0,00 | 0,00 | 0,00 |
| 20 | CYS:5:HB1 - APB:0:H6,H10 | 6,29 | 4,19 | 0,00 | 0,00 | 0,00 |
| 21 | CYS:5:HB2 - APB:0:H6,H10 | 6,06 | 4,45 | 0,00 | 0,00 | 0,00 |
| 22 | CYS:5:HB2 - APB:0:H7,H9 | 6,83 | 5,07 | 0,00 | 0,00 | 0,00 |
| 23 | ASP:6:HA - GLY:7:HN | 2,82 | 2,35 | 0,00 | 0,00 | 0,00 |
| 24 | ASP:6:HB1 - THR:4:HG1 | 4,79 | 7,58 | 2,79 | 2,79 | 2,79 |
| 25 | ASP:6:HB2 - THR:4:HG1 | 4,63 | 6,66 | 2,03 | 2,03 | 2,03 |
| 26 | ASP:6:HB1 - GLY:7:HN | 3,60 | 2,60 | 0,00 | 0,00 | 0,00 |
| 27 | ASP:6:HB2 - GLY:7:HN | 3,64 | 2,94 | 0,00 | 0,00 | 0,00 |
| 28 | ASP:6:HB1 - APB:0:H6,H10 | 7,20 | 7,00 | 0,00 | 0,00 | 0,00 |
| 29 | ASP:6:HB2 - APB:0:H6,H10 | 7,20 | 7,03 | 0,00 | 0,00 | 0,00 |
| 30 | ASP:6:HB1 - APB:0:H7,H9 | 6,80 | 5,80 | 0,00 | 0,00 | 0,00 |
| 31 | ASP:6:HB2 - APB:0:H7,H9 | 6,80 | 6,16 | 0,00 | 0,00 | 0,00 |
| 32 | ASP:6:HB1 - APB:0:HN | 5,62 | 5,31 | 0,00 | 0,00 | 0,00 |
| 33 | GLY:7:HN - ASP:6:HN | 3,86 | 3,40 | 0,00 | 0,00 | 0,00 |
| 34 | GLY:7:HA1 - ASP:6:HA | 4,67 | 4,47 | 0,00 | 0,00 | 0,00 |
| 35 | GLY:7:HA2 - ASP:6:HA | 4,40 | 4,47 | 0,08 | 0,08 | 0,08 |
| 36 | GLY:7:HA1 - PHE:8:HN | 2,71 | 2,25 | 0,00 | 0,00 | 0,00 |
| 37 | GLY:7:HA2 - PHE:8:HN | 3,13 | 3,14 | 0,01 | 0,01 | 0,01 |
| 38 | GLY:7:HA1 - APB:0:HN | 4,28 | 3,92 | 0,00 | 0,00 | 0,00 |
| 39 | GLY:7:HA2 - APB:0:HN | 4,58 | 4,98 | 0,40 | 0,40 | 0,40 |
| 40 | PHE:8:HA - APB:0:HN | 2,50 | 2,70 | 0,20 | 0,20 | 0,20 |
| 41 | PHE:8:HB1 - APB:0:HN | 3,64 | 2,75 | 0,00 | 0,00 | 0,00 |
| 42 | PHE:8:HB2 - APB:0:HN | 3,79 | 2,85 | 0,00 | 0,00 | 0,00 |
| 43 | PHE:8:HN - APB:0:HN | 2,86 | 2,38 | 0,00 | 0,00 | 0,00 |
| 44 | PHE:8:HPHE* - APB:0:HN | 6,50 | 5,26 | 0,00 | 0,00 | 0,00 |

| 3J Nr. | Dihedral | $^3J_{\text{exp}}$ | $^3J_{\text{MD}}$ | Abw. |
|-----------|-------------------|--------------------|-------------------|-------|
| 1 | ALA:1:HN-N-CA-HA | 8,40 | 8,06 | -0,34 |
| 2 | CYS:2:HN-N-CA-HA | 8,10 | 8,08 | -0,02 |
| 3 | ALA:3:HN-N-CA-HA | 5,70 | 7,36 | 1,66 |
| 4 | THR:4:HN-N-CA-HA | 8,80 | 7,57 | -1,23 |
| 5 | CYS:5:HN-N-CA-HA | 8,00 | 7,67 | -0,33 |
| 6 | ASP:6:HN-N-CA-HA | 7,30 | 7,84 | 0,54 |
| 7 | GLY:7:HN-N-CA-HA2 | 5,60 | 6,33 | 0,73 |
| 8 | GLY:7:HN-N-CA-HA1 | 4,20 | 4,16 | -0,04 |

Tabelle D.2: *trans*-cAPB: Protonenabstände und 3J -Kopplungen.

D Tabellen des MD-NMR Vergleichs

| NOE | Atom 1 - Atom2 | | r_{exp} | r_{MD} | Abw. | Abweichung bez. Konformation | | | | | | |
|-----|-----------------|----------------------|------------------|-----------------|------|------------------------------|------|------|------|------|------|------|
| | | | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | ALA:1:HN | - CYS:2:HN | 2,80 | 2,78 | 0,00 | 0,00 | 0,00 | 0,94 | 0,00 | 1,01 | 0,28 | 0,00 |
| 2 | ALA:1:HN | - AMPB:0:H1,H2,H4,H5 | 5,20 | 3,65 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 3 | CYS:2:HN | - ALA:1:HB* | 2,80 | 2,81 | 0,01 | 0,00 | 0,02 | 0,00 | 0,10 | 0,14 | 0,01 | 0,16 |
| 4 | CYS:2:HN | - ALA:3:HN | 3,40 | 2,51 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 5 | CYS:2:HN | - ALA:3:HB* | 5,80 | 5,24 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,35 | 0,00 | 0,00 |
| 6 | ALA:3:HN | - ALA:1:HB* | 4,80 | 4,51 | 0,00 | 0,00 | 0,03 | 0,00 | 0,00 | 1,31 | 0,00 | 0,31 |
| 7 | ALA:3:HN | - CYS:2:HB1 | 3,40 | 2,63 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 8 | ALA:3:HN | - CYS:2:HB2 | 3,40 | 2,92 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,24 | 0,00 | 0,00 |
| 9 | ALA:3:HN | - THR:4:HN | 3,40 | 2,93 | 0,00 | 0,28 | 0,75 | 0,78 | 0,93 | 0,32 | 0,00 | 0,00 |
| 10 | ALA:3:HN | - THR:4:HG* | 5,70 | 5,89 | 0,19 | 0,37 | 0,00 | 0,83 | 0,64 | 0,49 | 0,02 | 0,00 |
| 11 | THR:4:HN | - CYS:2:HB1 | 4,80 | 5,31 | 0,51 | 1,34 | 1,59 | 1,44 | 1,45 | 1,17 | 0,00 | 0,78 |
| 12 | THR:4:HN | - CYS:2:HB2 | 4,80 | 5,78 | 0,98 | 1,53 | 1,61 | 2,27 | 1,70 | 2,34 | 0,49 | 0,98 |
| 13 | THR:4:HN | - ALA:3:HB* | 4,40 | 3,17 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 14 | CYS:5:HN | - THR:4:HN | 3,40 | 2,69 | 0,00 | 0,00 | 0,00 | 0,95 | 0,84 | 0,26 | 0,00 | 0,00 |
| 15 | CYS:5:HN | - THR:4:HG* | 4,30 | 4,07 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,42 | 0,00 | 0,00 |
| 16 | CYS:5:HN | - THR:4:HG1 | 3,40 | 3,38 | 0,00 | 0,00 | 0,00 | 1,15 | 0,74 | 0,76 | 0,19 | 0,09 |
| 17 | ASP:6:HN | - THR:4:HG* | 5,70 | 5,00 | 0,00 | 0,00 | 0,80 | 0,03 | 0,00 | 1,21 | 0,00 | 0,62 |
| 18 | GLY:7:HN | - ASP:6:HB1 | 4,00 | 3,12 | 0,00 | 0,00 | 0,00 | 0,19 | 0,00 | 0,00 | 0,00 | 0,00 |
| 19 | GLY:7:HN | - ASP:6:HB2 | 4,00 | 3,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 20 | AMPB:0:HN | - AMPB:0:HM1 | 2,80 | 2,53 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 21 | AMPB:0:HN | - AMPB:0:HM2 | 2,80 | 2,64 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 22 | AMPB:0:HN | - PHE:8:HA | 3,40 | 2,79 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,00 | 0,00 |
| 23 | AMPB:0:HN | - PHE:8:HB1 | 4,80 | 2,75 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 24 | AMPB:0:HN | - PHE:8:HB2 | 4,80 | 2,44 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 25 | AMPB:0:H7,H9 | - AMPB:0:HM1 | 2,80 | 2,29 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 26 | AMPB:0:H7,H9 | - AMPB:0:HM2 | 2,80 | 2,28 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 27 | AMPB:0:H7,H9 | - PHE:8:HA | 7,20 | 5,43 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 28 | AMPB:0:H7,H9 | - ALA:3:HB* | 8,20 | 5,70 | 0,00 | 0,00 | 0,00 | 4,72 | 1,79 | 0,00 | 0,00 | 0,00 |
| 29 | AMPB:0:H7,H9 | - THR:4:HG* | 8,10 | 5,75 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 30 | AMPB:0:H6,H10 | - AMPB:0:HM1 | 7,20 | 4,64 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 31 | AMPB:0:H6,H10 | - AMPB:0:HM2 | 7,20 | 4,62 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 32 | AMPB:0:H1,2,4,5 | - AMPB:0:H6 | 7,20 | 4,03 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 33 | AMPB:0:H1,2,4,5 | - AMPB:0:H7 | 8,80 | 5,49 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 34 | AMPB:0:H1,2,4,5 | - THR:4:HG* | 8,10 | 5,52 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,63 | 0,00 | 0,00 |

| ^3J Nr. | Dihedral | $^3\text{J}_{\text{exp}}$ | $^3\text{J}_{\text{MD}}$ | Abw. |
|------------------|------------------|---------------------------|--------------------------|-------|
| 1 | ALA:1:HN-N-CA-HA | 6,77 | 7,47 | 0,70 |
| 2 | CYS:2:HN-N-CA-HA | 7,63 | 6,77 | -0,86 |
| 3 | ALA:3:HN-N-CA-HA | 7,14 | 6,61 | -0,53 |
| 4 | PHE:8:HN-N-CA-HA | 8,24 | 7,78 | -0,46 |

Tabelle D.3: *cis*-cAMPB: Protonenabstände und ^3J -Kopplungen.

| NOE | Atom 1 - Atom2 | | r_{exp} | r_{MD} | Abw. | Abw. p. Konf. | | |
|-----|----------------|---------------|------------------|-----------------|------|---------------|------|------|
| | | | | | | 1 | 2 | 3 |
| 1 | ALA:1:HN | - CYS:2:HN | 2,20 | 2,40 | 0,20 | 0,22 | 0,10 | 0,19 |
| 2 | CYS:2:HN | - ALA:1:HB* | 3,70 | 2,82 | 0,00 | 0,00 | 0,00 | 0,00 |
| 3 | ALA:3:HN | - CYS:2:HN | 3,40 | 2,46 | 0,00 | 0,00 | 0,88 | 0,00 |
| 4 | ALA:3:HN | - CYS:2:HA | 3,40 | 2,92 | 0,00 | 0,00 | 0,00 | 0,00 |
| 5 | ALA:3:HN | - CYS:2:HB1 | 2,80 | 2,77 | 0,00 | 0,00 | 0,04 | 0,00 |
| 6 | ALA:3:HN | - CYS:2:HB2 | 2,80 | 2,64 | 0,00 | 0,00 | 0,06 | 0,00 |
| 7 | ALA:3:HN | - ALA:1:HB* | 5,80 | 4,98 | 0,00 | 0,00 | 1,15 | 0,00 |
| 8 | ALA:3:HN | - THR:4:HB | 4,00 | 5,58 | 1,58 | 1,75 | 0,86 | 1,70 |
| 9 | ALA:3:HN | - THR:4:HG* | 5,70 | 6,45 | 0,75 | 0,93 | 0,14 | 0,74 |
| 10 | THR:4:HN | - ALA:3:HB* | 4,30 | 3,04 | 0,00 | 0,00 | 0,00 | 0,00 |
| 11 | CYS:5:HN | - ALA:3:HB* | 4,30 | 5,93 | 1,63 | 1,42 | 2,77 | 1,92 |
| 12 | CYS:5:HN | - THR:4:HB | 2,80 | 3,01 | 0,21 | 0,28 | 0,92 | 0,01 |
| 13 | CYS:5:HN | - THR:4:HG1 | 3,40 | 3,05 | 0,00 | 0,00 | 0,00 | 0,00 |
| 14 | CYS:5:HN | - THR:4:HG* | 4,90 | 3,58 | 0,00 | 0,00 | 0,00 | 0,00 |
| 15 | ASP:6:HN | - CYS:5:HN | 2,80 | 3,12 | 0,32 | 0,15 | 1,29 | 0,73 |
| 16 | ASP:6:HN | - THR:4:HG* | 4,30 | 5,03 | 0,73 | 0,52 | 0,73 | 1,52 |
| 17 | ASP:6:HN | - CYS:5:HB2 | 4,00 | 2,88 | 0,00 | 0,00 | 0,00 | 0,00 |
| 18 | ASP:6:HN | - CYS:5:HB1 | 4,00 | 2,67 | 0,00 | 0,00 | 0,00 | 0,00 |
| 19 | GLY:7:HN | - ASP:6:HA | 3,40 | 2,40 | 0,00 | 0,00 | 0,11 | 0,00 |
| 20 | GLY:7:HN | - ASP:6:HB1 | 2,80 | 2,68 | 0,00 | 0,00 | 0,00 | 0,20 |
| 21 | GLY:7:HN | - ASP:6:HB2 | 2,80 | 2,84 | 0,04 | 0,00 | 0,00 | 0,56 |
| 22 | GLY:7:HN | - PHE:8:HN | 2,80 | 3,00 | 0,20 | 0,13 | 0,00 | 0,90 |
| 23 | PHE:8:HN | - THR:4:HB | 4,80 | 9,42 | 4,62 | 4,37 | 6,49 | 4,81 |
| 24 | PHE:8:HN | - GLY:7:HA1 | 3,70 | 2,76 | 0,00 | 0,00 | 0,00 | 0,00 |
| 25 | PHE:8:HN | - GLY:7:HA2 | 3,40 | 2,38 | 0,00 | 0,00 | 0,00 | 0,00 |
| 26 | PHE:8:HN | - AMPB:0:HN | 2,80 | 2,47 | 0,00 | 0,00 | 0,00 | 1,24 |
| 27 | AMPB:0:HN | - PHE:8:HA | 3,40 | 2,65 | 0,00 | 0,00 | 0,00 | 0,00 |
| 28 | AMPB:0:HN | - AMPB:0:H7,H | 5,10 | 2,89 | 0,00 | 0,00 | 0,00 | 0,00 |
| 29 | AMPB:0:HN | - AMPB:0:HM1 | 2,80 | 2,54 | 0,00 | 0,00 | 0,00 | 0,07 |
| 30 | AMPB:0:HN | - AMPB:0:HM2 | 2,80 | 2,59 | 0,00 | 0,02 | 0,00 | 0,00 |
| 31 | AMPB:0:HN | - GLY:7:HN | 4,80 | 4,37 | 0,00 | 0,00 | 0,00 | 0,10 |
| 32 | AMPB:0:HN | - GLY:7:HA1 | 4,80 | 4,80 | 0,00 | 0,00 | 0,26 | 0,72 |
| 33 | AMPB:0:HN | - GLY:7:HA2 | 4,80 | 4,19 | 0,00 | 0,00 | 0,00 | 0,15 |
| 34 | AMPB:0:HN | - PHE:8:HB1 | 2,80 | 2,73 | 0,00 | 0,00 | 0,00 | 0,74 |
| 35 | AMPB:0:HN | - PHE:8:HB2 | 2,80 | 2,72 | 0,00 | 0,00 | 0,00 | 0,27 |
| 36 | AMPB:0:H7,H9 | - PHE:8:HA | 5,70 | 5,24 | 0,00 | 0,00 | 0,00 | 0,00 |
| 37 | AMPB:0:H7,H9 | - AMPB:0:HM1 | 5,10 | 2,29 | 0,00 | 0,00 | 0,00 | 0,00 |
| 38 | AMPB:0:H7,H9 | - AMPB:0:HM2 | 5,10 | 2,29 | 0,00 | 0,00 | 0,00 | 0,00 |
| 39 | AMPB:0:H6,H10 | - AMPB:0:HN | 5,70 | 4,97 | 0,00 | 0,00 | 0,00 | 0,00 |
| 40 | AMPB:0:H6,H10 | - THR:4:HG* | 6,60 | 5,73 | 0,00 | 0,00 | 0,00 | 0,00 |
| 41 | AMPB:0:H6,H10 | - CYS:5:HB2 | 4,90 | 5,08 | 0,18 | 0,30 | 4,31 | 0,00 |
| 42 | AMPB:0:H6,H10 | - CYS:5:HB1 | 4,90 | 4,85 | 0,00 | 0,00 | 4,41 | 0,00 |

| 3J Nr. | Dihedral | $^3J_{\text{exp}}$ | $^3J_{\text{MD}}$ | Abw. |
|-----------|-------------------|--------------------|-------------------|-------|
| 1 | ALA:1:HN-N-CA-HA | 7,14 | 7,90 | 0,76 |
| 2 | PHE:8:HN-N-CA-HA | 8,12 | 7,67 | -0,45 |
| 3 | ASP:6:HN-N-CA-HA | 8,00 | 7,76 | -0,24 |
| 4 | CYS:5:HN-N-CA-HA | 8,24 | 7,71 | -0,53 |
| 5 | CYS:2:HN-N-CA-HA | 8,00 | 7,89 | -0,11 |
| 6 | ALA:3:HN-N-CA-HA | 6,90 | 7,38 | 0,48 |
| 7 | THR:4:HN-N-CA-HA | 8,86 | 7,51 | -1,35 |
| 8 | GLY:7:HN-N-CA-HA1 | 5,30 | 4,34 | -0,96 |
| 9 | GLY:7:HN-N-CA-HA2 | 5,50 | 5,35 | -0,15 |

Tabelle D.4: *trans*-cAMPB: Protonenabstände und 3J -Kopplungen.

D Tabellen des MD-NMR Vergleichs

| NOE | Atom 1 - Atom2 | r_{exp} | r_{MD} | Abw. | Abw. per Konf. | | | |
|-----|---------------------------|------------------|-----------------|------|----------------|------|------|------|
| | | | | | 1 | 2 | 3 | 4 |
| 1 | AMPB:0:H2,H4 - ALA:1:HN | 4,40 | 2,68 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 2 | CYS:2:HN - ALA:1:HN | 2,80 | 2,42 | 0,00 | 0,00 | 0,00 | 0,00 | 0,60 |
| 3 | ALA:1:HA - CYS:2:HN | 4,00 | 3,14 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 4 | ALA:1:HB* - CYS:2:HN | 4,50 | 2,94 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 5 | CYS:2:HA - ALA:3:HN | 2,80 | 2,25 | 0,00 | 0,00 | 0,00 | 0,13 | 0,00 |
| 6 | CYS:2:HB* - ALA:3:HN | 4,30 | 3,44 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 7 | THR:4:HN - ALA:3:HN | 2,80 | 2,56 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 8 | ALA:3:HB* - THR:4:HN | 2,80 | 3,00 | 0,20 | 0,18 | 0,26 | 0,22 | 0,20 |
| 9 | CYS:5:HN - THR:4:HN | 3,40 | 2,62 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 10 | THR:4:HA - CYS:5:HN | 3,40 | 2,89 | 0,00 | 0,00 | 0,00 | 0,05 | 0,00 |
| 11 | THR:4:HB - CYS:5:HN | 4,00 | 3,38 | 0,00 | 0,00 | 0,00 | 0,00 | 0,14 |
| 12 | THR:4:HG* - CYS:5:HN | 5,90 | 4,59 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 13 | CYS:5:HA - ASP:6:HN | 2,80 | 2,77 | 0,00 | 0,11 | 0,34 | 0,00 | 0,00 |
| 14 | CYS:5:HB* - ASP:6:HN | 3,70 | 2,78 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 15 | THR:4:HN - ASP:6:HN | 4,80 | 4,70 | 0,00 | 0,00 | 0,00 | 1,83 | 2,33 |
| 16 | GLY:7:HN - ASP:6:HN | 4,00 | 2,44 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 17 | ASP:6:HA - GLY:7:HN | 2,80 | 3,03 | 0,23 | 0,32 | 0,09 | 0,74 | 0,00 |
| 18 | ASP:6:HB* - GLY:7:HN | 4,90 | 2,70 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 19 | CYS:5:HB* - GLY:7:HN | 4,80 | 4,39 | 0,00 | 0,00 | 0,77 | 0,53 | 0,00 |
| 20 | GLY:7:HA* - PHE:8:HN | 4,30 | 2,63 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 21 | PHE:8:HA - AMPB:0:HN | 2,80 | 2,46 | 0,00 | 0,00 | 0,74 | 0,75 | 0,67 |
| 22 | PHE:8:HB* - AMPB:0:HN | 4,90 | 2,94 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 23 | AMPB:0:H7,H9 - AMPB:0:HN | 5,60 | 2,83 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 24 | AMPB:0:HM* - AMPB:0:H7,H9 | 5,40 | 2,10 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 25 | PHE:8:HB* - AMPB:0:H7,H9 | 7,10 | 5,57 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 26 | PHE:8:HA - AMPB:0:H7,H9 | 6,20 | 5,12 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 27 | ALA:1:HB* - AMPB:0:H2,H4 | 6,50 | 5,16 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 28 | ALA:1:HB* - AMPB:0:H1,H5 | 8,20 | 7,33 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |

| ^3J Nr. | Dihedral | $^3\text{J}_{\text{exp}}$ | $^3\text{J}_{\text{MD}}$ | Abw. |
|------------------|--------------------|---------------------------|--------------------------|-------|
| 1 | ALA:1:HN-N-CA-HA | 7,50 | 7,67 | 0,17 |
| 2 | CYS:2:HN-N-CA-HA | 7,30 | 7,57 | 0,27 |
| 3 | ALA:3:HN-N-CA-HA | 5,20 | 6,29 | 1,09 |
| 4 | CYS:5:HN-N-CA-HA | 7,50 | 6,44 | -1,06 |
| 5 | ASP:6:HN-N-CA-HA | 7,80 | 7,88 | 0,08 |
| 6 | GLY:7:HN-N-CA-HA1 | 5,60 | 5,98 | 0,38 |
| 7 | GLY:7:HN-N-CA-HA2 | 5,60 | 4,01 | -1,59 |
| 8 | PHE:8:HN-N-CA-HA | 8,00 | 7,64 | -0,36 |
| 9 | AMPB:0:HN-N-CM-HM2 | 5,80 | 4,28 | -1,52 |
| 10 | AMPB:0:HN-N-CM-HM1 | 5,80 | 5,57 | -0,23 |

Tabelle D.5: *cis*-bcAMPB: Protonenabstände und ^3J -Kopplungen.

| NOE | Atom 1 - Atom2 | r_{exp} | r_{MD} | Abw. | Abw. p. Konf. | | |
|-----|---------------------------|------------------|-----------------|------|---------------|------|------|
| | | | | | 1 | 2 | 3 |
| 1 | AMPB:0:H2,H4 - ALA:1:HN | 5,50 | 2,64 | 0,00 | 0,00 | 0,00 | 0,00 |
| 2 | CYS:2:HN - ALA:1:HN | 3,40 | 2,14 | 0,00 | 0,00 | 0,00 | 0,00 |
| 3 | ALA:1:HA - CYS:2:HN | 2,80 | 3,39 | 0,59 | 0,69 | 0,46 | 0,60 |
| 4 | ALA:1:HB* - CYS:2:HN | 4,30 | 3,25 | 0,00 | 0,00 | 0,00 | 0,00 |
| 5 | CYS:2:HB2 - CYS:2:HA | 2,80 | 2,44 | 0,00 | 0,00 | 0,00 | 0,00 |
| 6 | CYS:2:HB1 - CYS:2:HA | 2,80 | 2,87 | 0,07 | 0,20 | 0,00 | 0,04 |
| 7 | CYS:2:HB2 - CYS:2:HN | 4,00 | 2,78 | 0,00 | 0,00 | 0,00 | 0,00 |
| 8 | CYS:2:HB1 - CYS:2:HN | 2,80 | 2,41 | 0,00 | 0,00 | 0,00 | 0,00 |
| 9 | CYS:2:HA - ALA:3:HN | 2,30 | 2,22 | 0,00 | 0,00 | 0,00 | 0,00 |
| 10 | CYS:2:HB2 - ALA:3:HN | 3,40 | 3,31 | 0,00 | 0,31 | 0,00 | 0,00 |
| 11 | THR:4:HN - ALA:3:HN | 2,80 | 2,43 | 0,00 | 0,00 | 0,00 | 0,00 |
| 12 | CYS:5:HN - THR:4:HN | 3,40 | 2,71 | 0,00 | 0,00 | 0,00 | 0,00 |
| 13 | ALA:3:HA - THR:4:HN | 2,80 | 3,56 | 0,76 | 0,76 | 0,76 | 0,76 |
| 14 | CYS:2:HB2 - THR:4:HN | 4,00 | 4,49 | 0,49 | 0,91 | 0,27 | 0,37 |
| 15 | ALA:3:HB* - THR:4:HN | 3,90 | 3,09 | 0,00 | 0,00 | 0,00 | 0,00 |
| 16 | ALA:3:HA - CYS:5:HN | 3,40 | 4,11 | 0,71 | 0,65 | 0,76 | 0,75 |
| 17 | THR:4:HA - CYS:5:HN | 3,40 | 2,78 | 0,00 | 0,00 | 0,00 | 0,00 |
| 18 | THR:4:HB - CYS:5:HN | 4,00 | 3,88 | 0,00 | 0,00 | 0,13 | 0,00 |
| 19 | CYS:5:HA - ASP:6:HN | 2,30 | 2,58 | 0,28 | 0,71 | 0,03 | 0,27 |
| 20 | CYS:5:HB2 - ASP:6:HN | 3,40 | 2,87 | 0,00 | 0,48 | 0,00 | 0,00 |
| 21 | CYS:5:HB1 - ASP:6:HN | 4,90 | 2,91 | 0,00 | 0,00 | 0,00 | 0,00 |
| 22 | ASP:6:HA - GLY:7:HN | 2,30 | 2,48 | 0,18 | 0,23 | 0,11 | 0,20 |
| 23 | CYS:5:HA - GLY:7:HN | 4,00 | 5,74 | 1,74 | 1,52 | 2,16 | 1,78 |
| 24 | ASP:6:HB* - GLY:7:HN | 5,80 | 2,64 | 0,00 | 0,00 | 0,00 | 0,00 |
| 25 | GLY:7:HA1 - GLY:7:HN | 3,40 | 2,68 | 0,00 | 0,00 | 0,00 | 0,00 |
| 26 | GLY:7:HA2 - GLY:7:HN | 3,40 | 2,59 | 0,00 | 0,00 | 0,00 | 0,00 |
| 27 | GLY:7:HA1 - PHE:8:HN | 2,80 | 2,42 | 0,00 | 0,00 | 0,00 | 0,00 |
| 28 | GLY:7:HA2 - PHE:8:HN | 2,80 | 2,71 | 0,00 | 0,45 | 0,00 | 0,00 |
| 29 | PHE:8:HB1 - PHE:8:HN | 2,80 | 2,45 | 0,00 | 0,00 | 0,00 | 0,00 |
| 30 | PHE:8:HB2 - PHE:8:HN | 3,40 | 2,98 | 0,00 | 0,00 | 0,00 | 0,00 |
| 31 | PHE:8:HA - AMPB:0:HN | 2,30 | 2,59 | 0,29 | 0,91 | 1,23 | 0,00 |
| 32 | PHE:8:HB2 - AMPB:0:HN | 3,40 | 2,71 | 0,00 | 0,00 | 0,00 | 0,00 |
| 33 | AMPB:0:HN - AMPB:0:H7,H9 | 5,00 | 2,92 | 0,00 | 0,00 | 0,00 | 0,00 |
| 34 | AMPB:0:HM* - AMPB:0:H7,H9 | 5,90 | 2,12 | 0,00 | 0,00 | 0,00 | 0,00 |
| 35 | ASP:6:HA - AMPB:0:H6,H1 | 6,20 | 5,37 | 0,00 | 0,00 | 0,02 | 0,01 |
| 36 | CYS:5:HB2 - AMPB:0:H6,H1 | 7,00 | 7,51 | 0,51 | 0,81 | 0,18 | 0,51 |
| 37 | ALA:1:HA - AMPB:0:H2,H4 | 5,60 | 4,28 | 0,00 | 0,00 | 0,00 | 0,00 |
| 38 | ALA:1:HB* - AMPB:0:H2,H4 | 7,30 | 5,22 | 0,00 | 0,00 | 0,00 | 0,00 |
| 39 | CYS:5:HB2 - AMPB:0:H2,H4 | 6,20 | 6,86 | 0,66 | 0,91 | 0,13 | 0,98 |
| 40 | CYS:2:HB2 - AMPB:0:H2,H4 | 7,00 | 6,58 | 0,00 | 0,00 | 0,00 | 0,00 |
| 41 | CYS:5:HB2 - AMPB:0:H1,H5 | 7,00 | 6,68 | 0,00 | 0,03 | 0,00 | 0,00 |

| 3J Nr. | Dihedral | $^3J_{\text{exp}}$ | $^3J_{\text{MD}}$ | Abw. |
|-----------|--------------------|--------------------|-------------------|-------|
| 1 | THR:4:HN-N-CA-HA | 9,30 | 6,58 | -2,72 |
| 2 | CYS:5:HN-N-CA-HA | 7,50 | 6,75 | -0,75 |
| 3 | ASP:6:HN-N-CA-HA | 8,20 | 7,75 | -0,45 |
| 4 | GLY:7:HN-N-CA-HA1 | 4,70 | 4,82 | 0,12 |
| 5 | GLY:7:HN-N-CA-HA2 | 4,70 | 4,70 | -0,00 |
| 6 | PHE:8:HN-N-CA-HA | 8,20 | 7,55 | -0,65 |
| 7 | AMPB:0:HN-N-CM-HM2 | 5,80 | 6,33 | 0,53 |
| 8 | AMPB:0:HN-N-CM-HM1 | 5,80 | 4,91 | -0,89 |

Tabelle D.6: *trans*-bcAMPB: Protonenabstände und 3J -Kopplungen.

Literaturverzeichnis

- [1] D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Penguin (1979). 1
- [2] T. E. Creighton, *Proteins: structures and molecular properties*, W. H. Freeman and Company, New York, 2. Auflage (1993). 1, 4, 6, 18
- [3] M. Levitt und S. Lifson, *Refinement of Protein Conformation using a Macromolecular Energy Minimization Procedure*, J. Mol. Biol., 46, 269–279 (1969). 1
- [4] M. Karplus und D. L. Weaver, *Protein Folding Dynamics*, Nature, 260, 404 (1976). 1
- [5] W. F. van Gunsteren und H. J. C. Berendsen, *Algorithms for Macromolecular Dynamics and Constraint Dynamics*, Mol. Phys., 34, 1311–1327 (1977). 1
- [6] J. A. McCammon, B. R. Gelin und M. Karplus, *Dynamics of folded proteins*, Nature, 267, 585–590 (1977). 1, 18
- [7] C. L. Brooks, *Protein and peptide folding explored with molecular simulations*, Acc. Chem. Res., 35, 447–454 (2002). 1, 18, 96
- [8] M. Karplus und J. A. McCammon, *Molecular dynamics simulations of biomolecules*, Nat. Struct. Biol., 9, 646–652 (2002). 1, 18
- [9] G. E. Schulz und R. H. Schirmer, *Principles of Protein Structure*, Springer advanced texts in chemistry, Springer-Verlag, Berlin, Heidelberg, New York (1979). 3, 5, 6
- [10] IUPAC-IUB Commission on Biochemical Nomenclature, *Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains* (1974).
<http://www.chem.qmul.ac.uk/iupac/misc/ppep1.html> 3
- [11] S. C. Lovell, I. W. Davis, W. B. Adrendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson und D. C. Richardson, *Structure validation by C alpha geometry: phi, psi and C beta deviation*, Proteins, 50, 437–450 (2003). 5
- [12] J. A. McCammon und S. C. Harvey, *Dynamics of proteins and nucleic acids*, Cambridge University Press, Cambridge (1987). 5, 7
- [13] G. N. Ramachandran, C. Ramakrishnan und V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*, J. Mol. Biol., 7, 95–99 (1963). 4
- [14] R. Behrendt, M. Schenk, H. J. Musiol und L. Moroder, *Photomodulation of conformational states. Synthesis of cyclic peptides with backbone-azobenzene moieties*, J. Pept. Sci., 5, 519–529 (1999). 6, 24
- [15] D. Voet, V. J. G. und C. W. Pratt, *Fundamentals of Biochemistry*, John Wiley & Sons, New York (1999). 6

- [16] H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain et al., *The Protein Data Bank*, Acta Crystallogr. Sec. D, 58, 899–907 (2002).
www.rcsb.org/pdb 6, 10
- [17] W. Humphrey, A. Dalke und K. Schulten, *VMD - Visual Molecular Dynamics*, J. Mol. Graphics, 14, 33–38 (1996). 6, 204
- [18] C. Branden und J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York (1991). 6
- [19] S. Geibel, J. H. Kaplan, E. Bamberg und T. Friedrich, *Conformational dynamics of the Na⁺/K⁺-ATPase probed by voltage clamp fluorometry*, Proc. Natl. Acad. Sci. USA, 100, 964–969 (2003). 7
- [20] W. A. Eaton, V. Munoz, S. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry und J. Hofrichter, *Fast kinetics and mechanisms in protein folding*, Annu. Rev. Biophys. Biomol. Struct., 29, 327–359 (2000). 8, 14, 23
- [21] L. L. Qiu, S. A. Pabit, A. E. Roitberg und S. J. Hagen, *Smaller and faster: The 20-residue Trp-cage protein folds in 4 μ s*, J. Am. Chem. Soc., 124, 12952–12953 (2002). 8
- [22] H. Grubmüller und P. Tavan, *Molecular-Dynamics of Conformational Substates For a Simplified Protein Model*, J. Chem. Phys., 101, 5047–5057 (1994). 8, 90
- [23] F. A. Hamprecht, C. Peter, X. Daura, W. Thiel und W. F. van Gunsteren, *A strategy for analysis of (molecular) equilibrium simulations: Configuration space density estimation, clustering, and visualization*, J. Chem. Phys., 114, 2079–2089 (2001). 8
- [24] C. Anfinsen, E. Haber, M. Sela und F. J. White, *The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain*, Proc. Natl. Acad. Sci. USA, 47, 1309–1314 (1961). 8
- [25] C. Anfinsen, *Principles that govern the folding of protein chains*, Science, 181, 223–230 (1973). 8
- [26] C. Levinthal, *Are there pathways for protein folding?*, J. Chem. Phys., 65, 44–45 (1968). 9
- [27] R. Zwanzig, A. Szabo und B. Bagghi, *Levinthal's paradox*, Proc. Natl. Acad. Sci. USA, 89, 20–22 (1991). 9
- [28] P. S. Kim und R. L. Baldwin, *Intermediates in the folding reactions of small proteins*, Ann. Rev. Biochem, 59, 631–660 (1990). 9
- [29] H. S. Chan und K. A. Dill, *Protein folding in the landscape perspective: Chevron plots and non-Arrhenius kinetics*, Proteins, 30, 2–33 (1998). 9
- [30] R. L. Baldwin, *The Nature of Protein-Folding Pathways - the Classical Versus the New View*, J. Biomol. NMR, 5, 103–109 (1995). 9
- [31] K. Dill und H. Chan, *From Levinthal to pathways to funnels*, Nat. Struct. Biol., 4, 10–19 (1997). 9

- [32] H. Frauenfelder, P. G. Wolynes und R. H. Austin, *Biological physics*, Rev. Mod. Phys., 71, S419–S430 (1999). 9
- [33] M. Jormakka, S. Tornroth, B. Byrne und S. Iwata, *Molecular basis of proton motive force generation: Structure of formate dehydrogenase-N*, Science, 295, 1863–1868 (2002). 11
- [34] C. R. Cantor und P. R. Schimmel, *Techniques for the study of biological structure and function*, Band 2 von *Biophysical chemistry*, W. H. Freeman and Company, New York (1980). 10, 11, 14
- [35] Z. Ren, B. Perman, V. Srajer, T. Y. Teng, C. Pradervand, D. Bourgeois, F. Schotte, T. Ursby, R. Kort, M. Wulff et al., *A molecular movie at 1.8 angstrom resolution displays the photocycle of photoactive yellow protein, a eubacterial blue-light receptor, from nanoseconds to seconds*, Biochemistry, 40, 13788–13801 (2001). 10
- [36] E. Fanchon, E. Geissler, J. L. Hodeau, J. R. Regnard und P. A. Timmins (Herausgeber), *Neutron and Synchrotron Radiation for Condensed Matter Studies*, Band 4 von *Structure and Dynamics of Biomolecules*, Oxford University Press, Oxford (2000). 11
- [37] C. Renner, R. Behrendt, S. Spörlein, J. Wachtveitl und L. Moroder, *Photomodulation of conformational states. I. Mono- and bicyclic peptides with (4-amino)phenylazobenzoic acid as backbone constituent*, Biopolymers, 54, 489–500 (2000). 12, 25, 26, 28, 93, 129, 131, 133, 147, 149, 166, 168, 173, 221
- [38] C. Renner, J. Cramer, R. Behrendt und L. Moroder, *Photomodulation of conformational states. II. Mono- and bicyclic peptides with (4-aminomethyl)phenylazobenzoic acid as backbone constituent*, Biopolymers, 54, 501–514 (2000). 12, 25, 28, 93, 129, 131, 147, 149, 164, 221
- [39] K. Wüthrich, *NMR of proteins and nucleic acids*, John Wiley & Sons, New York (1986). 12, 156, 158
- [40] J. N. S. Evans, *Biomolecular NMR spectroscopy*, Oxford University Press, Oxford (1995). 12
- [41] M. Saunders, A. Wishnia und J. G. Kirkwood, *The nuclear magnetic resonance spectrum of ribonuclease*, J. Am. Chem. Soc., 79, 3289–3290 (1957). 12
- [42] M. Karplus, *Contact Electron-Spin Coupling of Nuclear Magnetic Moments*, J. Chem. Phys., 30, 11–15 (1959). 12
- [43] D. Neuhaus und M. Williamson, *The nuclear Overhauser effect in structural and conformational analysis*, VCH Verlagsgesellschaft, Weinheim (1989). 12, 13
- [44] L. Nilsson, M. Karplus, A. Brünger, M. Clore und A. Gronenborn, *Determination of 3-D Molecular-Structures from NMR/NOE Distance Data - Application to Nucleic-Acids*, J. Mol. Graphics, 4, 182–182 (1986). 13
- [45] A. G. Palmer, *NMR probes of molecular dynamics: Overview and comparison with other techniques*, Annu. Rev. Biophys. Biomol. Struct., 30, 129–155 (2001). 13

- [46] D. A. Case, *Molecular dynamics and NMR spin relaxation in proteins*, Acc. Chem. Res., 35, 325–331 (2002). 13
- [47] S. Williams, T. P. Causgrove, R. Gilmanshin, K. S. Fang, R. H. Callender, W. H. Woodruff und R. B. Dyer, *Fast events in protein folding: Helix melting and formation in a small peptide*, Biochemistry, 35, 691–697 (1996). 14
- [48] P. Thompson, W. Eaton und J. Hofrichter, *Laser Temperature Jump Study of the HelixCoil Kinetics of an Alanine Peptide Interpreted with a 'Kinetic Zipper' Model*, Biochemistry, 36, 9200–9210 (1997). 14
- [49] I. K. Lednev, A. S. Karnoup, M. C. Sparrow und S. A. Asher, *alpha-helix peptide folding and unfolding activation barriers: A nanosecond UV resonance raman study*, J. Am. Chem. Soc., 121, 8074–8086 (1999). 14
- [50] C. Y. Huang, Z. Getahun, Y. J. Zhu, J. W. Klemke, W. F. DeGrado und F. Gai, *Helix formation via conformation diffusion search*, Proc. Natl. Acad. Sci. USA, 99, 2788–2793 (2002). 14, 15
- [51] M. Jacob, G. Holtermann, D. Perl, J. Reinstein, T. Schindler, M. A. Geeves und F. X. Schmid, *Microsecond folding of the cold shock protein measured by a pressure-jump technique*, Biochemistry, 38, 2882–2891 (1999). 14
- [52] P. Wittung-Stafshede, J. C. Lee, J. R. Winkler und H. B. Gray, *Cytochrome b(562) folding triggered by electron transfer: Approaching the speed limit for formation of a four-helix-bundle protein*, Proc. Natl. Acad. Sci. USA, 96, 6587–6590 (1999). 14
- [53] H. S. M. Lu, M. Volk, Y. Kholodenko, E. Gooding, R. M. Hochstrasser und W. F. DeGrado, *Aminothietyrosine disulfide, an optical trigger for initiation of protein folding*, J. Am. Chem. Soc., 119, 7173–7180 (1997). 14
- [54] M. T. Zanni und R. M. Hochstrasser, *Two-dimensional infrared spectroscopy: a promising new method for the time resolution of structures*, Curr. Opin. Struct. Biol., 11, 516–522 (2001). 15
- [55] S. Woutersen und P. Hamm, *Structure determination of trialanine in water using polarization sensitive two-dimensional vibrational spectroscopy*, J. Phys. Chem. B, 104, 11316–11320 (2000). 15
- [56] S. Woutersen, Y. G. Mu, G. Stock und P. Hamm, *Subpicosecond conformational dynamics of small peptides probed by two-dimensional vibrational spectroscopy*, Proc. Natl. Acad. Sci. USA, 98, 11254–11258 (2001). 15
- [57] J. Dobler, W. Zinth, W. Kaiser und D. Oesterhelt, *Excited-State Reaction Dynamics of Bacteriorhodopsin Studied By Femtosecond Spectroscopy*, Chem. Phys. Lett., 144, 215–220 (1988). 15
- [58] I. Lutz, A. Sieg, A. A. Wegener, M. Engelhard, I. Boche, M. Otsuka, D. Oesterhelt, J. Wachtveitl und W. Zinth, *Primary reactions of sensory rhodopsins*, Proc. Natl. Acad. Sci. USA, 98, 962–967 (2001). 15

- [59] P. Tavan und K. Schulten, *The Low-Lying Electronic Excitations in Long Polyenes – A PPP-MRD-CI Study*, J. Chem. Phys., 85, 6602–6609 (1986). 16
- [60] P. Strodel und P. Tavan, *A revised MRCI-algorithm. I. Efficient combination of spin adaptation with individual configuration selection coupled to an effective valence-shell Hamiltonian*, J. Chem. Phys., 117, 4667–4676 (2002). 16
- [61] The Royal Swedish Academy of Sciences (1998).
<http://www.nobel.se/announcement-98/chemistry98.html> 17
- [62] P. Hohenberg und W. Kohn, *Inhomogeneous Electron Gas*, Phys. Rev., 136, B864–B870 (1964). 17
- [63] W. Kohn und L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev., 140, A1133–A1138 (1965). 17
- [64] A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A, 38, 3098–3100 (1988). 17, 45
- [65] C. Lee, W. Yang und R. G. Parr, *Development of the Colle-Salvetti correlation energy formula into a functional of the electron density*, Phys. Rev. B, 37, 785–789 (1988). 17, 45
- [66] A. D. Becke, *Density-functional thermochemistry. III. The role of exact exchange*, J. Chem. Phys., 98, 5648 (1993). 17, 45
- [67] R. G. Parr und W. Yang, *Density-functional theory of atoms and molecules*, Oxford University Press, New York (1989). 17
- [68] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. Montgomery, J. A., R. E. Stratmann, J. C. Burant et al., *Gaussian 98, revision A.7; Gaussian, Inc.: Pittsburgh, PA* (1998). 17, 46, 74
- [69] J. Hutter, A. Alavi, T. Deutsch, M. Bernasconi, S. Goedecker, D. Marx, M. Tuckerman und M. Parrinello, *CPMD: Car-Parinello Molecular Dynamics, version 3.3a*, MPI für Festkörperforschung und IBM Zurich Research Laboratory 1995-96 (1999). 17, 46
- [70] N. Troullier und J. L. Martins, *Efficient pseudopotentials for plane-wave calculations*, Phys. Rev. B, 43, 1993–2005 (1991). 17, 46
- [71] S. Goedecker, M. Teter und J. Hutter, *Separable dual-space Gaussian pseudopotentials*, Phys. Rev. B, 54, 1703–1710 (1996). 17, 46
- [72] R. Car und M. Parrinello, *Unified Approach for Molecular Dynamics and Density-Functional Theory*, Phys. Rev. Lett., 55, 2471–2474 (1985). 17
- [73] P. Carloni, U. Rothlisberger und M. Parrinello, *The role and perspective of a initio molecular dynamics in the study of biological systems*, Acc. Chem. Res., 35, 455–464 (2002). 17
- [74] M. P. Allen und D. Tildesley, *Computer Simulations of Liquids*, Clarendon Press, Oxford (1987). 18, 19, 20, 21, 37, 57, 59, 64, 132, 215
- [75] D. Frenkel und B. Smit, *Understanding Molecular Simulation*, Band 1 von *Computational Science*, Academic Press, San Diego, 2. Auflage (2002). 18, 89, 91

- [76] W. D. Cornell, P. Cieplak, C. Bayly, I. R. Gould, K. M. J. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell und P. A. Kollman, *A second generation force field for the simulation of proteins and nucleic acids*, J. Am. Chem. Soc., 117, 5179–5197 (1995). 18
- [77] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan und M. Karplus, *CHARMM - a Program For Macromolecular Energy, Minimization, and Dynamics Calculations*, J. Comput. Chem., 4, 187–217 (1983). 18, 38, 44, 117
- [78] W. F. van Gunsteren und H. J. C. Berendsen, *GROMOS Manual*, BIOMOS b.v., Biomolecular Software, Groningen, the Netherlands (). 18
- [79] W. L. Jorgensen, D. S. Maxwell und J. TiradoRives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids*, J. Am. Chem. Soc., 118, 11225–11236 (1996). 18
- [80] G. Mathias, M. Eichinger, H. Carstens, M. Stork, A. Weiss, H. Heller, H. Grubmüller, B. Egwolf, C. Niedermeier und P. Tavan, *EGO-MMII users guide*, Lehrstuhl für BioMolekulare Optik, Ludwig Maximilian Universität München, Oettingenstrasse 67, D-80538 München (), in preparation. 18, 21, 31
- [81] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins*, J. Phys. Chem. B, 102, 3586–3616 (1998). 18, 19, 34, 38, 73, 117, 118, 119, 120, 123, 124, 127, 217
- [82] W. F. van Gunsteren und H. J. C. Berendsen, *Computer Simulation as a Tool for Tracing the Conformational Differences between Proteins in Solution and in the Crystalline State*, J. Mol. Biol., 176, 559–564 (1984). 18
- [83] W. F. van Gunsteren und H. J. C. Berendsen, *Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry*, Angew. Chem. Int. Ed. Engl., 29, 992–1023 (1990). 18, 19
- [84] H. Grubmüller, B. Heymann und P. Tavan, *Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force*, Science, 271, 997–999 (1996). 18
- [85] Y. Duan und P. A. Kollman, *Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution*, Science, 282, 740–744 (1998). 18, 20, 89
- [86] X. Daura, B. Jaun, D. Seebach, W. van Gunsteren und A. Mark, *Reversible Peptide folding in Solution by Molecular Dynamics Simulation*, J. Mol. Biol., 280, 925–932 (1998). 18, 20
- [87] X. Daura, K. Gademann, H. Schafer, B. Jaun, D. Seebach und W. F. van Gunsteren, *The beta-peptide hairpin in solution: Conformational study of a beta-hexapeptide in methanol by NMR spectroscopy and MD simulation*, J. Am. Chem. Soc., 123, 2393–2404 (2001). 18, 129, 155
- [88] G. Hummer, A. E. García und S. Garde, *Helix nucleation kinetics from molecular simulations in explicit solvent*, Proteins, 42, 77–84 (2001). 18, 96

- [89] I. C. Yeh und G. Hummer, *Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent*, J. Am. Chem. Soc., 124, 6563–6568 (2002). 18
- [90] R. Zhou, B. J. Berne und R. Germain, *The free energy landscape for beta hairpin folding in explicit water*, Proc. Natl. Acad. Sci. USA, 98, 14931–14936 (2001). 18, 96
- [91] U. Mayor, N. R. Guydosh, C. M. Johnson, J. G. Grossmann, S. Sato, G. S. Jas, S. M. V. Freund, D. O. V. Alonso, V. Daggett und A. R. Fersht, *The complete folding pathway of a protein from nanoseconds to microseconds*, Nature, 421, 863–867 (2003). 18
- [92] B. Guillot, *A reappraisal of what we have learnt during three decades of computer simulations on water*, J. Mol. Liq., 101, 219–260 (2002). 19
- [93] J. W. Ponder und D. A. Case, *Force Fields for Protein Simulations*, Adv. Prot. Chem, 66, 27–85 (2003). 19, 169
- [94] P. Tavan, H. Carstens und G. Mathias, *Molecular dynamics simulations of proteins and peptides: Problems, achievements, and perspectives*, in J. Buchner und T. Kiefhaber (Herausgeber), *Handbook of Protein Folding*, Wiley-VCH, Weinheim (2004), in press. 19, 169
- [95] P. Soto und A. E. Mark, *The Effect of the Neglect of Electronic Polarization in Peptide Folding Simulations*, J. Phys. Chem. B, 106, 12830–12833 (2002). 19
- [96] G. Mathias, *Elektrostatische Wechselwirkungen in komplexen Flüssigkeiten und ihre Beschreibung mit Molekulardynamiksimulationen*, Dissertation, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (2004). 19, 20, 22, 40
- [97] T. A. Darden, D. York und L. Pedersen, *Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems*, J. Chem. Phys., 98, 10089–10092 (1993). 19, 20
- [98] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee und L. G. Pedersen, *A smooth particle mesh Ewald method*, J. Chem. Phys., 103, 8577–8593 (1995). 19, 20
- [99] B. A. Luty und W. F. van Gunsteren, *Calculating Electrostatic Interactions Using the Particle-Particle Particle-Mesh Method with Nonperiodic Long-Range Interactions*, J. Phys. Chem., 100, 2581–2587 (1996). 19, 20
- [100] G. Mathias, B. Egwolf, M. Nonella und P. Tavan, *A fast multipole method combined with a reaction field for long-range electrostatics in molecular dynamics simulations: The effects of truncation on the properties of water*, J. Chem. Phys., 118, 10847–10860 (2003). 19, 20, 21, 22, 40, 123, 131, 173, 216
- [101] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola und J. R. Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys., 81, 3684–3690 (1984). 19, 123, 132
- [102] I. G. Tironi, R. Sperb, P. E. Smith und W. F. Van Gunsteren, *A Generalized Reaction Field Method For Molecular-Dynamics Simulations*, J. Chem. Phys., 102, 5451–5459 (1995). 20
- [103] P. H. Hünenberger und W. F. van Gunsteren, *Alternative schemes for the inclusion of a reaction-field correction into molecular dynamics simulations: Influence on the simulated energetic, structural, and dielectric properties of liquid water*, J. Chem. Phys., 108, 6117–6134 (1998). 20, 67

- [104] D. Mihailescu und J. C. Smith, *Molecular Dynamics Simulation of the Cyclic Decapeptide Antibiotic, Gramicidin S, in Dimethyl Sulfoxide Solution*, J. Phys. Chem. B, 103, 1586–1594 (1999). 20
- [105] P. P. Ewald, *Die Berechnung optischer und elektrostatischer Gitterpotentiale*, Ann. Phys. (Leipzig), IV, 253–287 (1920). 20
- [106] P. H. Hünenberger und J. A. McCammon, *Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: A continuum electrostatics study*, Biophys. Chem., 78, 69–88 (1999). 21
- [107] C. Niedermeier und P. Tavan, *A Structure Adapted Multipole Method For Electrostatic Interactions in Protein Dynamics*, J. Chem. Phys., 101, 734–748 (1994). 21, 38, 39
- [108] M. Eichinger, H. Grubmüller, H. Heller und P. Tavan, *FAMUSAMM: An algorithm for rapid evaluation of electrostatic interactions in molecular dynamics simulations*, J. Comput. Chem., 18, 1729–1749 (1997). 21, 22, 39
- [109] B. Egwolf, *Proteine in wässriger Umgebung: Kontinuumstheorie der Lösungsmittielektrostatik und ihre effiziente Berechnung*, Dissertation, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (2004). 22
- [110] A. Warshel und M. Levitt, *Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme*, J. Mol. Biol., 103, 227–249 (1976). 22
- [111] M. Eichinger, P. Tavan, J. Hutter und M. Parrinello, *A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields*, J. Chem. Phys., 110, 10452–10467 (1999). 22
- [112] M. Nonella, G. Mathias, M. Eichinger und P. Tavan, *Structures and Vibrational Frequencies of the Quinones in Rb. sphaeroides Derived by a Combined Density Functional/Molecular Mechanics Approach*, J. Phys. Chem. B, 107, 316–322 (2003). 22
- [113] M. Nonella, G. Mathias und P. Tavan, *Infrared Spectrum of p-Benzoquinone in Water Obtained from a QM/MM Hybrid Molecular Dynamics Simulation*, J. Phys. Chem. A, 107, 8638–8647 (2003). 22
- [114] G. Mathias, *Anwendung von QM/MM Hybridmethoden zur Berechnung der Schwingungsspektren von biologischen Chromophoren in situ*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München, München, Germany (2000). 22
- [115] U. Haupts, J. Tittor und D. Oesterhelt, *Closing in on Bacteriorhodopsin: Progress in Understanding the Molecule*, Annu. Rev. Biophys. Biomol. Struct., 28, 367–399 (1999). 23
- [116] W. Zinth und D. Oesterhelt, *The Primary Photochemical Process in Bacteriorhodopsin*, Photobiol., 531–535 (1991). 23
- [117] H. Rau, *Azo Compounds*, in H. Dürr und H. Bouas-Laurent (Herausgeber), *Studies in Organic Chemistry, Photochromism, Molecules and Systems*, Band 40, Kapitel 4, 165–192, Elsevier, Amsterdam (1990). 24

- [118] N. Tamai und H. Miyasaka, *Ultrafast Dynamics of Photochromic Systems*, Chem. Rev., 100, 1875–1890 (2000). 24
- [119] R. Behrendt, C. Renner, M. Schenk, F. Q. Wang, J. Wachtveitl, D. Oesterhelt und L. Moroder, *Photomodulation of the conformation of cyclic peptides with azobenzene moieties in the peptide backbone*, Angew. Chem. Int. Ed. Engl., 38, 2771–2774 (1999). 24
- [120] T. Ikeda und O. Tsutsumi, *Optical Switching and Image Storage by Means of Azobenzene Liquid-Crystal Films*, Science, 268, 1873–1875 (1995). 26
- [121] B. L. Feringa, R. A. van Delden, N. Koumura und E. M. Geertsema, *Chiroptical molecular switches*, Chem. Rev., 100, 1789–1816 (2000). 26
- [122] G. Iftime, F. L. Labarthe, A. Natansohn und P. Rochon, *Control of chirality of an azobenzene liquid crystalline polymer with circularly polarized light*, J. Am. Chem. Soc., 122, 12646–12650 (2000). 26
- [123] R. H. Berg, S. Hvilsted und P. S. Ramanujam, *Peptide oligomers for holographic data storage*, Nature, 383, 505–508 (1996). 26
- [124] A. Natansohn, P. Rochon, X. S. Meng, C. Barrett, T. Buffeteau, S. Bonenfant und M. Pezolet, *Molecular addressing? Selective photoinduced cooperative motion of polar ester groups in copolymers containing azobenzene groups*, Macromolecules, 31, 1155–1161 (1998). 26
- [125] P. H. Rasmussen, P. S. Ramanujam, S. Hvilsted und R. H. Berg, *A remarkably efficient azobenzene peptide for holographic information storage*, J. Am. Chem. Soc., 121, 4738–4743 (1999). 26
- [126] M. S. Vollmer, T. D. Clark, C. Steinem und M. R. Ghadiri, *Photoswitchable hydrogen-bonding in self-organized cylindrical peptide systems*, Angew. Chem. Int. Ed. Engl., 38, 1598–1601 (1999). 26
- [127] D. M. Junge und D. V. McGrath, *Photoresponsive azobenzene-containing dendrimers with multiple discrete states*, J. Am. Chem. Soc., 121, 4912–4913 (1999). 26
- [128] A. Cattani-Scholz, C. Renner, D. Oesterhelt und L. Moroder, *Photoresponsive dendritic azobenzene peptides*, ChemBioChem, 2, 542–549 (2001). 26
- [129] I. Willner und S. Rubin, *Control of the structure and functions of biomaterials by light*, Angew. Chem. Int. Ed. Engl., 35, 367–385 (1996). 26
- [130] L. Ulysse, J. Cubillos und J. Chmielewski, *Photoregulation of Cyclic Peptide Conformation*, J. Am. Chem. Soc., 117, 8466–8467 (1995). 26
- [131] J. R. Kumita, O. S. Smart und G. A. Woolley, *Photo-control of helix content in a short peptide*, Proc. Natl. Acad. Sci. USA, 97, 3803–3808 (2000). 26
- [132] D. G. Flint, J. R. Kumita, O. S. Smart und G. A. Woolley, *Using an azobenzene cross-linker to either increase or decrease peptide helix content upon trans-to-cis photoisomerization*, Chem. Biol., 9, 391–397 (2002). 26
- [133] X. G. Liang, H. Asanuma und M. Komiyama, *Photoregulation of DNA triplex formation by azobenzene*, J. Am. Chem. Soc., 124, 1877–1883 (2002). 26

- [134] T. Hugel, N. B. Holland, A. Cattani, L. Moroder, M. Seitz und H. E. Gaub, *Single-molecule optomechanical cycle*, *Science*, 296, 1103–1106 (2002). 26
- [135] T. Nägele, R. Hoche, W. Zinth und J. Wachtveitl, *Femtosecond photoisomerization of cis-azobenzene*, *Chem. Phys. Lett.*, 272, 489–495 (1997). 26, 179, 181
- [136] P. Hamm, S. Ohline und W. Zinth, *Vibrational cooling after ultrafast photoisomerization of azobenzene by femtosecond infrared spectroscopy*, *J. Chem. Phys.*, 106, 519–529 (1997). 26
- [137] J. Wachtveitl, T. Nägele, B. Puell, W. Zinth, M. Krüger, S. Bohner, D. Oesterhelt und L. Moroder, *Ultrafast photoisomerization of azobenzene compounds*, *J. Photochem. Photobiol. A*, 105, 283–288 (1997). 26, 182
- [138] S. Spörlein, H. Carstens, H. Satzger, C. Renner, R. Behrendt, L. Moroder, P. Tavan, W. Zinth und J. Wachtveitl, *Ultrafast spectroscopy reveals subnanosecond peptide conformational dynamics and validates molecular dynamics simulation*, *Proc. Natl. Acad. Sci. USA*, 99, 7998–8002 (2002). 26, 82, 169, 171, 172, 173, 179, 181, 182, 184, 187, 191, 209, 210
- [139] S. Spörlein, *Femtosekunden-Spektroskopie schnellster Strukturänderungen in Peptid-Chromophor-Komplexen*, Dissertation, Ludwig-Maximilians-Universität München (2001). 26, 170, 181, 182, 183
- [140] N. Biswas und S. Umapathy, *Early time dynamics of trans-azobenzene isomerization in solution from resonance Raman intensity analysis*, *J. Chem. Phys.*, 107, 7849–7858 (1997).
- [141] T. Saito und K. T., *Conformational Change in Azobenzene in Photoisomerization Process Studied with Chirp-Controlled sub-10-fs Pulses*, *J. Phys. Chem. A*, 106, 9436–9441 (2002).
- [142] C. Renner, R. Behrendt, N. Heim und L. Moroder, *Photomodulation of conformational states. III. Water-soluble bis-cysteinyl-peptides with (4-aminomethyl) phenylazobenzoic acid as backbone constituent*, *Biopolymers*, 63, 382–393 (2002). 28
- [143] A. Cattani-Scholz, C. Renner, C. Cabrele, R. Behrendt, D. Oesterhelt und L. Moroder, *Photoresponsive cyclic bis(cysteinyl)peptides as catalysts of oxidative protein folding*, *Angew. Chem. Int. Ed. Engl.*, 41, 289–+ (2001).
- [144] M. Kloppenburg, *Lokale Hauptkomponentenanalyse in hochdimensionalen Merkmalsräumen*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1996). 29
- [145] M. Kloppenburg und P. Tavan, *Deterministic annealing for density estimation by multivariate normal mixtures*, *Phys. Rev. E*, 55, 2089–2092 (1997). 29, 91, 99, 101, 102, 106
- [146] S. Albrecht, *Multivariate GRBF-Netzwerke und Systeme lokaler Experten: Neue Konzepte und ihre Anwendung bei komplexen Regelungsaufgaben*, Dissertation, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1999). 29, 99, 140
- [147] S. Albrecht, J. Busch, M. Kloppenburg, F. Metzke und P. Tavan, *Generalized Radial Basis Function Networks for Classification and Novelty Detection: Self-Organization of Optimal Bayesian Decision*, *Neural Networks*, 13, 1075–1093 (2000). 29, 91, 99, 102, 103, 106, 177, 200

- [148] D. R. Dersch und P. Tavan, *Asymptotic Level Density in Topological Feature Maps*, IEEE Trans. Neural Netw., 6, 230–236 (1995). 29, 91, 106
- [149] D. R. Dersch, *Eigenschaften neuronaler Vektorquantisierer und ihre Anwendung in der Sprachverarbeitung*, Dissertation, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1995). 29, 91, 99
- [150] F. Metze, *Integration neuronaler Methoden zur Phonemerkennung aus fließend gesprochener Sprache*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1998). 29
- [151] V. Groll, *Selektion und Gewichtung selbstorganisierender Merkmalsdetektoren zur Phonemerkennung*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1995).
- [152] F. Schwabl, *Quantenmechanik*, Springer-Verlag, Berlin, Heidelberg, New York, 4. Auflage (1993). 32
- [153] J. Simons und J. Nichols, *Quantum Mechanics in Chemistry*, Oxford Univeristy Press, New York, Oxford (1997). 32
- [154] W. Brenig, *Statistische Theorie der Wärme*, Springer, Berlin, Heidelberg, New York, 3. Auflage (1992). 33
- [155] H. Grubmüller, *Molekulardynamik von Proteinen auf langen Zeitskalen*, Dissertation, Technische Universität München, Germany (1994). 33
- [156] H. Grubmüller und P. Tavan, *Multiple time step algorithms for molecular dynamics simulations of proteins: How good are they?*, J. Comput. Chem., 19, 1534–1552 (1998). 33
- [157] L. Verlet, *Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*, Phys. Rev., 159, 98–103 (1967). 33
- [158] J.-P. Ryckaert, G. Ciccotti und H. J. C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes*, J. Chem. Phys., 23, 327–330 (1977). 34
- [159] V. Krautler, W. F. Van Gunsteren und P. H. Hünenberger, *A fast SHAKE: Algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations*, J. Comput. Chem., 22, 501–508 (2001). 34, 53, 123, 131
- [160] A. T. Brünger, *X-Plor (Version 3.1) A System for X-ray Crystallography and NMR*, Yale University Press, New Haven (1992). 34, 117
- [161] C. Niedermeier und P. Tavan, *Fast version of the structure adapted multipole method- efficient calculation of electrostatic forces in protein dynamics*, Molecular Simulation, 17, 57–66 (1996). 38, 39
- [162] C. Niedermeier, *Modellierung elektrostatischer Wechselwirkungen in Proteinen: Eine strukturadaptierte Multipolmethode*, Dissertation, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1995). 39

- [163] T. M. Martinez und K. J. Schulten, *A 'Neural Gas' Network Learns Topologies.*, in O. Simula (Herausgeber), *Proceedings of the Int. Conf. on Artificial Neural Networks, ICANN-91, Espoo, Finland, 24-28 June 1991*, 397–402, Elsevier Science Publishers, Amsterdam (1991). 39
- [164] M. Eichinger, *Paralleler schneller Multipolalgorithmus mit Mehrschrittverfahren für Molekulardynamiksimulationen*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (1995). 39
- [165] B. G. Rao und U. C. Singh, *A free energy perturbation study of solvation in methanol and dimethyl sulfoxide*, *J. Am. Chem. Soc.*, 112, 3803 – 3811 (1990). 43, 44
- [166] A. Luzar und D. Chandler, *Structure and Hydrogen-Bond Dynamics of Water-Dimethyl Sulfoxide Mixtures By Computer-Simulations*, *J. Chem. Phys.*, 98, 8160–8173 (1993). 43, 44
- [167] H. Y. Liu, F. Müller-Plathe und W. F. van Gunsteren, *A Force-Field For Liquid Dimethyl-Sulfoxide and Physical- Properties of Liquid Dimethyl-Sulfoxide Calculated Using Molecular-Dynamics Simulation*, *J. Am. Chem. Soc.*, 117, 4363–4366 (1995). 43, 44, 57, 62, 65
- [168] P. Bordat, J. Sacristan, D. Reith, S. Girard, A. Glattli und F. Muller-Plathe, *An improved dimethyl sulfoxide force field for molecular dynamics simulations*, *Chem. Phys. Lett.*, 374, 201–205 (2003). 43, 44, 57, 61, 62, 63
- [169] I. I. Vaisman und M. L. Berkowitz, *Local Structural Order and Molecular Associations in Water-DMSO Mixtures. Molecular Dynamics Study*, *J. Am. Chem. Soc.*, 114, 1889–1896 (1992). 43, 65
- [170] A. Luzar, A. K. Soper und D. Chandler, *Combined Neutron-Diffraction and Computer-Simulation Study of Liquid Dimethyl-Sulfoxide*, *J. Chem. Phys.*, 99, 6836–6847 (1993). 43, 68, 70, 71
- [171] I. A. Borin und M. S. Skaf, *Molecular association between water and dimethyl sulfoxide in solution: A molecular dynamics simulation study*, *J. Chem. Phys.*, 110, 6412–6420 (1999). 43
- [172] I. Benjamin, *Structure, thermodynamics, and dynamics of the liquid/vapor interface of water/dimethylsulfoxide mixtures*, *J. Chem. Phys.*, 110, 8070–8079 (1999). 43
- [173] A. Vishnyakov, A. P. Lyubartsev und A. Laaksonen, *Molecular dynamics simulations of dimethyl sulfoxide and dimethyl sulfoxide-water mixture*, *J. Phys. Chem. A*, 105, 1702–1710 (2001). 43
- [174] S. Senapati, *A molecular dynamics simulation study of the dimethyl sulfoxide liquid-vapor interface*, *J. Chem. Phys.*, 117, 1812–1816 (2002). 43
- [175] A. Vishnyakov, G. Widmalm, J. Kowalewski und A. Laaksonen, *Molecular dynamics simulation of the alpha-D-Manp-(1 -> 3)- beta-D-Glcp-OMe disaccharide in water and water DMSO solution*, *J. Am. Chem. Soc.*, 121, 5403–5412 (1999). 43

- [176] A. K. Das und B. L. Tembe, *Molecular dynamics simulations of sodium chloride solutions in water-dimethyl sulphoxide mixtures: Potentials of mean force and solvation structures*, J. Chem. Phys., 111, 7526–7536 (1999). 43
- [177] O. N. Kalugin, M. N. Volobuev, A. V. Ishchenko und A. K. Adya, *Structure and dynamics of Na⁺ and Cl⁻ solvation shells in liquid DMSO: molecular dynamics simulations*, J. Mol. Liq., 85, 299–312 (2000). 43
- [178] A. K. Adya, O. N. Kalugin, M. N. Volobuev und Y. V. Kolesnik, *Microscopic structure of liquid dimethyl sulphoxide and its electrolyte solutions: molecular dynamics simulations*, Mol. Phys., 99, 835–854 (2001). 43
- [179] R. Bürgi, J. Pitera und W. F. van Gunsteren, *Assessing the effect of conformational averaging on the measured values of observables*, J. Biomol. NMR, 19, 305–320 (2001). 43, 129
- [180] Y. J. Zheng und R. L. Ornstein, *A molecular dynamics and quantum mechanics analysis of the effect of DMSO on enzyme structure and dynamics: Subtilisin*, J. Am. Chem. Soc., 118, 4175–4180 (1996). 43
- [181] T. Fox und P. A. Kollman, *Application of the RESP methodology in the parametrization of organic solvents*, J. Phys. Chem. B, 102, 8070–8079 (1998). 44, 56, 62
- [182] M. L. Strader und S. E. Feller, *A flexible all-atom model of dimethyl sulfoxide for molecular dynamics simulations*, J. Phys. Chem. A, 106, 1074–1080 (2002). 44, 56, 62
- [183] F. Müller-Plathe, *How good are molecular local-density methods? Case studies: The quadrupole moment of benzene, geometry and electrostatics of dimethyl sulfoxide, and the conformations of dimethoxy ethane*, Braz. J. Phys., 24, 965–976 (1994). 45, 52, 56
- [184] M. Stener und M. Calligaris, *Density functional study of structural properties and binding energies of dimethylsulfoxide Ru(II) complexes*, Theochem J. Mol. Struct., 497, 91–104 (2000). 45
- [185] B. Kirchner und M. Reiher, *The secret of dimethyl sulfoxide-water mixtures. A quantum chemical study of 1DMSO-nwater clusters*, J. Am. Chem. Soc., 124, 6206–6215 (2002). 45
- [186] B. Kirchner und J. Hutter, *The structure of a DMSO-water mixture from Car-Parrinello simulations*, Chem. Phys. Lett., 364, 497–502 (2002). 45
- [187] W. Feder, H. Dreizler, H. D. Rudolph und V. Z. Typke, *Naturforsch.*, 24, 266–278 (1969). 46
- [188] H. Bertagnolli, E. Schulz und P. Chieux, *Ber. Bunsenges. Phys. Chem.*, 93, 88–95 (1989). 46
- [189] J. A. Glasel, *J. Am. Chem. Soc.*, 92, 372–375 (1970). 46, 52
- [190] J. C. Slater, *The Self-Consistent Field for Molecular and Solids*, Band 4 von *Quantum Theory of Molecular and Solids*, McGraw-Hill, New York (1974). 45
- [191] S. Vosko, L. Wilk und M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Can. J. Phys., 58, 1200 (1980). 45

- [192] J. Perdew und W. Yue, *Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation*, Phys. Rev. B, 33, 8800–8802 (1986). 45
- [193] B. Schrader, *Raman-, Infrared atlas of organic compounds*, VCH Verlagsgesellschaft, Weinheim (1989). 48
- [194] K. G. R. Pachler, *Merck FT IR Atlas*, VCH Verlagsgesellschaft, Weinheim (1988). 48
- [195] P. Pulay, G. Fogarasi, F. Pang und J. E. Boggs, *Systematic ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives*, J. Am. Chem. Soc., 101, 2550–2560 (1979). 49
- [196] K. M. Merz, *Analysis of a Large Database of Electrostatic Potential Derived Atomic Charges*, J. Comput. Chem., 13, 749–771 (1992). 50
- [197] A. Weiss, *Entwicklung von polarisierbaren Wassermolekülen für Molekulardynamik-Simulationen*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (2001). 51
- [198] E. Engel, A. Hock und R. M. Dreizler, *van der Waals bonds in density-functional theory*, Phys. Rev. A, 6103, art. no.–032502 (2000). 59
- [199] W. Press, S. Teukolsky, W. Vetterling und B. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 2. Auflage (1992). 60, 213
- [200] R. Faller, H. Schmitz, O. Biermann und F. Müller-Plathe, *Automatic parameterization of force fields for liquids by simplex optimization*, J. Comput. Chem., 20, 1009–1017 (1999). 60
- [201] M. S. Skaf, *Molecular dynamics simulations of dielectric properties of dimethyl sulfoxide: Comparison between available potentials*, J. Chem. Phys., 107, 7996–8003 (1997). 62
- [202] J. A. Riddick, W. B. Bunger und T. K. Sakand, *Organic Solvents Physical Properties and Methods of Purification*, John Wiley and Sons, New York (1986). 65
- [203] F. Müller-Plathe, *A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity*, J. Chem. Phys., 106, 6082–6085 (1997). 66
- [204] R. Chitra und S. Yashonath, *Estimation of error in the diffusion coefficient from molecular dynamics simulations*, J. Phys. Chem. B, 101, 5437–5445 (1997). 67
- [205] N. Biswas und S. Umapathy, *Density functional calculations of structures, vibrational frequencies, and normal modes of trans- and cis-azobenzene*, J. Phys. Chem. A, 101, 5555–5566 (1997). 74
- [206] N. Kurita, S. Tanaka und S. Itoh, *Ab initio molecular orbital and density functional studies on the stable structures and vibrational properties of trans- and cis-azobenzenes*, J. Phys. Chem. A, 104, 8114–8120 (2000). 74
- [207] V. Stepanic, G. Baranovic und V. Smrecki, *Structure and vibrational spectra of conjugated acids of trans- and cis-azobenzene*, J. Mol. Struct., 569, 89–109 (2001). 74

- [208] T. Tsuji, H. Takashima, H. Takeuchi, T. Egawa und S. Konaka, *Molecular structure and torsional potential of trans- azobenzene. A gas electron diffraction study*, J. Phys. Chem. A, 105, 9347–9353 (2001). 74
- [209] N. Kurita, T. Ikegami und Y. Ishikawa, *Ab initio study of the minimum-energy structure of trans- azobenzene*, Chem. Phys. Lett., 360, 349–354 (2002). 74
- [210] T. Fujino und T. Tahara, *Picosecond time-resolved Raman study of trans-azobenzene*, J. Phys. Chem. A, 104, 4203–4210 (2000). 74
- [211] C. M. Breneman und K. B. Wiberg, *Determining Atom-Centered Monopoles From Molecular Electrostatic Potentials - the Need For High Sampling Density in Formamide Conformational-Analysis*, J. Comput. Chem., 11, 361–373 (1990). 76
- [212] H. A. Stern, G. A. Kaminski, J. L. Banks, R. H. Zhou, B. J. Berne und R. A. Friesner, *Fluctuating charge, polarizable dipole, and combined models: Parameterization from ab initio quantum chemistry*, J. Phys. Chem. B, 103, 4730–4737 (1999). 77
- [213] G. A. Kaminski, H. A. Stern, B. J. Berne, R. A. Friesner, Y. X. X. Cao, R. B. Murphy, R. H. Zhou und T. A. Halgren, *Development of a polarizable force field for proteins via ab initio quantum chemistry: First generation model and gas phase tests*, J. Comput. Chem., 23, 1515–1531 (2002). 77
- [214] G. Tabacchi, C. J. Mundy, J. Hutter und M. Parrinello, *Classical polarizable force fields parametrized from ab initio calculations*, J. Chem. Phys., 117, 1416–1433 (2002). 77
- [215] R. Chelli und P. Procacci, *A transferable polarizable electrostatic force field for molecular mechanics based on the chemical potential equalization principle*, J. Chem. Phys., 117, 9175–9189 (2002). 77
- [216] M. W. Schmidt, K. Baldridge, J. Boatz, S. Elbert, M. Gordin, J. Jensen, S. Koseki, N. Matsunaga, K. Nguyen, S. Su et al., *GAMESS Version 25 June 2001*, J. Comput. Chem., 14, 1347–1363 (1993). 81
- [217] S. Monti, G. Orlandi und P. Palmieri, *Features of the Photochemically Active State Surfaces of Azobenzene*, Chem. Phys., 71, 87–99 (1982). 84, 183
- [218] P. O. Astrand, P. S. Ramanujam, S. Hvilsted, K. L. Bak und S. P. A. Sauer, *Ab initio calculation of the electronic spectrum of azobenzene dyes and its impact on the design of optical data storage materials*, J. Am. Chem. Soc., 122, 3482–3487 (2000). 84
- [219] P. Cattaneo und M. Persico, *An ab initio study of the photochemistry of azobenzene*, Phys. Chem. Chem. Phys., 1, 4739–4743 (1999). 84
- [220] T. Ishikawa, T. Noro und T. Shoda, *Theoretical study on the photoisomerization of azobenzene*, J. Chem. Phys., 115, 7503–7512 (2001). 84
- [221] H. Rau und E. Lüddecke, ???, J. Am. Chem. Soc., 104, 1616–1620 (1982). 85
- [222] I. K. Lednev, T. Q. Ye, L. C. Abbott, R. E. Hester und J. N. Moore, *Photoisomerization of a capped azobenzene in solution probed by ultrafast time-resolved electronic absorption spectroscopy*, J. Phys. Chem. A, 102, 9161–9166 (1998). 85

- [223] T. Fujino, S. Y. Arzhantsev und T. Tahara, *Femtosecond time-resolved fluorescence study of photoisomerization of trans-azobenzene*, J. Phys. Chem. A, 105, 8123–8129 (2001). 85
- [224] W. Domcke und G. Stock, *Theory of ultrafast nonadiabatic excited-state processes and their spectroscopic detection in real time*, Band 100 von *Advances in Chemical Physics*, John Wiley and Sons, New York (1997). 85
- [225] I. T. Joliffe, *Principal Component Analysis*, Springer Series in Statistics, Springer-Verlag, Berlin, Heidelberg, New York (1986). 89, 94, 96
- [226] K. Mardia, J. Kent und J. Bibby, *Multivariate Analysis*, Academic Press, San Diego, 9 Auflage (1994). 89, 94
- [227] C. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford (1997). 89
- [228] R. O. Duda, P. E. Hart und D. G. Stork, *Pattern Classification*, John Wiley and Sons, New York, 2. Auflage (2000). 91, 97, 100, 101, 102
- [229] P. Schnell, *Eine Methode zur Auffindung von Gruppen*, Biometrische Zeitschrift, 6, 47–48 (1964). 91, 106
- [230] P. Tavan, H. Grubmüller und H. Kühnel, *Self-Organization of Associative Memory and Pattern Classification: Recurrent Signal Processing on Topological Feature Maps*, Biol. Cybernetics, 64, 95–105 (1990). 91, 106, 107, 108
- [231] T. Hirschberger, H. Carstens, M. Stork und P. Tavan, *Self-organization of hierarchical classification: Combining density estimation by normal mixtures with scale-space filtering*, IEEE Trans. Pattern Anal. Mach. Intell. (2003), in preparation. 91, 106, 107
- [232] S. R. Jammalamadaka und A. SenGupta, *Topics in circular statistics*, Band 5 von *Series on multivariate analysis*, World Scientific Publishing, Singapore (2001). 91, 112, 113, 114
- [233] K. Pearson, *On lines and planes of closest fit to systems of points in space*, Phil. Mag., 6, 559–572 (1901). 94
- [234] H. Hotelling, *Analysis of a complex of statistical variables into principal components*, J. Educ. Psychol., 24, 417–441 (1933). 94
- [235] A. Kitao und N. Go, *Investigating protein dynamics in collective coordinate space*, Curr. Opin. Struct. Biol., 9, 164–169 (1999). 96
- [236] A. Amadei, A. B. M. Linssen und H. J. C. Berendsen, *Essential Dynamics of Proteins*, Proteins, 17, 412–425 (1993). 96
- [237] A. Amadei, M. A. Ceruso und A. Di Nola, *On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations*, Proteins, 36, 419–424 (1999). 96
- [238] B. L. deGroot, X. Daura, A. E. Mark und H. Grubmüller, *Essential dynamics of reversible peptide folding: Memory-free conformational dynamics governed by internal hydrogen bonds*, J. Mol. Biol., 309, 299–313 (2001). 96

- [239] B. L. deGroot, D. M. F. vanAalten, R. M. Scheek, A. Amadei, G. Vriend und H. J. C. Berendsen, *Prediction of protein conformational freedom from distance constraints*, Proteins, 29, 240–251 (1997). 96
- [240] R. Abseher, L. Horstink, C. W. Hilbers und M. Nilges, *Essential spaces defined by NMR structure ensembles and molecular dynamics simulation show significant overlap*, Proteins, 31, 370–382 (1998). 96, 130
- [241] R. Abseher und M. Nilges, *Efficient sampling in collective coordinate space*, Proteins, 39, 82–88 (2000). 96
- [242] C. J. Margulis, H. A. Stern und B. J. Berne, *Helix unfolding and intramolecular hydrogen bond dynamics in small alpha-helices in explicit solvent*, J. Phys. Chem. B, 106, 10748–10752 (2002). 96
- [243] Z. Y. Guo, C. L. Brooks und E. M. Boczko, *Exploring the folding free energy surface of a three-helix bundle protein*, Proc. Natl. Acad. Sci. USA, 94, 10161–10166 (1997). 96
- [244] B. G. Schulze und J. D. Evanseck, *Cooperative role of Arg45 and His64 in the spectroscopic A(3) state of carbonmonoxy myoglobin: Molecular dynamics simulations, multivariate analysis, and quantum mechanical computations*, J. Am. Chem. Soc., 121, 6444–6454 (1999). 96
- [245] M. A. Balsera, W. Wriggers, Y. Oono und K. Schulten, *Principal component analysis and long time protein dynamics*, J. Phys. Chem., 100, 2567–2572 (1996). 96
- [246] B. Hess, *Convergence of sampling in protein simulations*, Phys. Rev. E, 65, art. no.–031910 (2002). 96
- [247] O. M. Becker, *Geometric versus topological clustering: An insight into conformation mapping*, Proteins, 27, 213–226 (1997). 97
- [248] O. M. Becker, *Principal coordinate maps of molecular potential energy surfaces*, J. Comput. Chem., 19, 1255–1267 (1998). 97
- [249] Y. Levy und O. M. Becker, *Effect of conformational constraints on the topography of complex potential energy surfaces*, Phys. Rev. Lett., 81, 1126–1129 (1998). 97
- [250] Y. Levy und O. M. Becker, *Energy landscapes of conformationally constrained peptides*, J. Chem. Phys., 114, 993–1009 (2001). 97
- [251] E. Parzen, *On Estimation of a Probability Density Function and Mode*, Ann. Math. Stat., 33, 1065–1076 (1962). 97
- [252] D. R. Dersch und P. Tavan, *Load balanced Vector Quantization*, in *Proceedings of the International Conference on Artificial Neural Networks*, 1067–1070, Springer, Berlin (1994). 99, 102, 103, 105
- [253] T. Bayes, *An Essay towards Solving a Problem in the Doctrine of Chances*, Philosophical Transactions of the Royal Society, 53, 370–418 (1763). 100
- [254] P. S. Laplace, *Théorie Analytique des Propabilités*, Courcier, Paris, France (1812). 100

- [255] A. Dempster, N. Laird und D. Rubin, *Maximum Likelihood from Incomplete Data via the EM Algorithm*, J. Royal Stat. Soc., 39, 1–38 (1977). 102
- [256] K. Rose, E. Gurewitz und G. C. Fox, *Statistical Mechanics and Phase Transitions in Clustering*, Phys. Rev. Lett., 65, 945–948 (1990). 102, 103
- [257] H. Ritter, *Asymptotic Level Density for a Class of Vector Quantization Processes*, IEEE Trans. Neural Netw., 1, 173–175 (1991). 106
- [258] T. Kohonen, *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, Heidelberg, New York (1984). 106
- [259] T. Hirschberger, *Hierarchische Klassifikation durch selbstorganisierende neuronale Algorithmen*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (2002). 106
- [260] J. Weickert, S. Ishikawa und A. Imiya, *Linear Scale-Space has First been Proposed in Japan*, J. Math. Imag. Vis., 10, 237–252 (1999). 107
- [261] J. Koenderink, *The Structure of Images*, Biol. Cybernetics, 50, 363–370 (1984). 107
- [262] T. Lindeberg, *Scale-Space Theory: A Basic Tool for Analysing Structures at Different Scales*, J. Appl. Stat., 21, 225–270 (1994). 107
- [263] F. Kowalewski, *A Gradient Procedure for Determining Clusters of Relatively High Point Density*, Pattern Recognition, 28, 1973–1984 (1995). 108
- [264] R. Rojas, *Theorie der neuronalen Netze*, Springer-Verlag, Berlin, Heidelberg, New York (1996). 108
- [265] S. Chakravarthy und J. Ghosh, *Scale-Based Clustering using the Radial Basis Function Network*, IEEE Trans. Neural Netw., 7, 1250–1261 (1996).
- [266] N. Kehtarnavaz und E. Nakamura, *Generalization of the EM algorithm for mixture density estimation*, Pattern Recognition Lett., 19, 133–140 (1998).
- [267] Y. Leung, J.-S. Zhang und Z.-B. Xu, *Clustering by Scale-Space Filtering*, IEEE Trans. Pattern Anal. Mach. Intell. 22, 1396–1410 (2000).
- [268] M. A. Carreira-Perpinan, *Mode-finding for mixtures of Gaussian distributions*, IEEE Trans. Pattern Anal. Mach. Intell. 22, 1318–1323 (2000).
- [269] U. Adolphs, *Selbstorganisierende Perzeptronstrukturen mit radialen Basisfunktionen*, Diplomarbeit, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München, München (1992). 114
- [270] M. Eichinger, H. Grubmüller und H. Heller, *User Manual for EGO_VIII, Release 2.0*, Lehrstuhl für BioMolekulare Optik, Ludwig Maximilian Universität München, Oettingenstrasse 67, D-80538 München (). 117
- [271] I. R. Gould, W. D. Cornell und I. H. Hillier, *A Quantum-Mechanical Investigation of the Conformational Energetics of the Alanine and Glycine Dipeptides in the Gas- Phase and in Aqueous-Solution*, J. Am. Chem. Soc., 116, 9250–9256 (1994). 119, 120

- [272] R. Vargas, J. Garza, B. P. Hay und D. A. Dixon, *Conformational study of the alanine dipeptide at the MP2 and DFT levels*, J. Phys. Chem. A, 106, 3213–3218 (2002). 119, 120
- [273] T. Headgordon, M. Headgordon, M. J. Frisch, C. L. Brooks und J. A. Pople, *Theoretical Study of Blocked Glycine and Alanine Peptide Analogs*, J. Am. Chem. Soc., 113, 5989–5997 (1991). 120
- [274] K. C. Chou, *Prediction of tight turns and their types in proteins*, Analyt. Biochem., 286, 1–16 (2000). 120
- [275] G. Némethy und M. P. Printz, *The γ Turn, a Possible Folded Conformation of the Polypeptide Chain. Comparison with the β Turn*, Macromolecules, 5, 755–758 (1972). 120
- [276] P. J. Rossky und M. Karplus, *Solvation. A Molecular Dynamics Study of a Dipeptide in Water*, J. Am. Chem. Soc., 101, 1913–1937 (1979). 122
- [277] P. J. Rossky und M. Karplus, *A Model for the Simulation of an Aqueous Dipeptide Solution*, Biopolymers, 18, 825–854 (1979). 122
- [278] D. J. Tobias und C. L. Brooks, *Conformational Equilibrium in the Alanine Dipeptide in the Gas- Phase and Aqueous-Solution - a Comparison of Theoretical Results*, J. Phys. Chem., 96, 3864–3870 (1992). 122
- [279] C. L. Brooks und D. A. Case, *Simulations of Peptide Conformational Dynamics and Thermodynamics*, Chem. Rev., 93, 2487–2502 (1993). 122
- [280] J. Kleinjung, P. Bayley und F. Fraternali, *Leap-dynamics: efficient sampling of conformational space of proteins and peptides in solution*, FEBS Lett., 470, 257–262 (2000). 122
- [281] A. Laio und M. Parrinello, *Escaping free-energy minima*, Proc. Natl. Acad. Sci. USA, 99, 12562–12566 (2002). 122
- [282] G. Hummer und I. G. Kevrekidis, *Coarse molecular dynamics of a peptide fragment: Free energy, kinetics, and long-time dynamics computations*, J. Chem. Phys., 118, 10762–10773 (2003). 122
- [283] I. Andricioaei, A. R. Dinner und M. Karplus, *Self-guided enhanced sampling methods for thermodynamic averages*, J. Chem. Phys., 118, 1074–1084 (2003). 122
- [284] R. W. W. Hooft, B. P. Vaneijck und J. Kroon, *An Adaptive Umbrella Sampling Procedure in Conformational-Analysis Using Molecular-Dynamics and Its Application to Glycol*, J. Chem. Phys., 97, 6690–6694 (1992). 122
- [285] C. Bartels und M. Karplus, *Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations*, J. Comput. Chem., 18, 1450–1462 (1997). 122
- [286] P. E. Smith, *The alanine dipeptide free energy surface in solution*, J. Chem. Phys., 111, 5568–5579 (1999). 122, 126, 127
- [287] J. Apostolakis, P. Ferrara und A. Caffisch, *Calculation of conformational transitions and barriers in solvated systems: Application to the alanine dipeptide in water*, J. Chem. Phys., 110, 2099–2108 (1999). 122, 126, 127

- [288] P. G. Bolhuis, C. Dellago und D. Chandler, *Reaction coordinates of biomolecular isomerization*, Proc. Natl. Acad. Sci. USA, 97, 5877–5882 (2000). 122
- [289] H. Hu, M. Elstner und J. Hermans, *Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides"(Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution*, Proteins, 50, 451–463 (2003). 122, 126, 127, 166, 220
- [290] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai und G. Seifert, *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*, Phys. Rev. B, 58, 7260–7268 (1998). 122
- [291] Q. Z. Cui und V. H. Smith, *Solvation structure, thermodynamics, and conformational dependence of alanine dipeptide in aqueous solution analyzed with reference interaction site model theory*, J. Chem. Phys., 118, 279–290 (2003).
- [292] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey und M. L. Klein, *Comparison of simple potential functions for simulating liquid water*, J. Chem. Phys., 79, 926–935 (1983). 123, 217
- [293] A. S. Yang und B. Honig, *Free-Energy Determinants of Secondary Structure Formation .I. Alpha-Helices*, J. Mol. Biol., 252, 351–365 (1995). 127
- [294] R. Bürgi, X. Daura, A. Mark, M. Bellanda, S. Mammi, E. Peggion und W. van Gunsteren, *Folding study of an Aib-rich peptide in DMSO by molecular dynamics simulations*, J. Pept. Res., 57, 107–118 (2001). 129, 155
- [295] X. Daura, I. Antes, W. F. van Gunsteren, W. Thiel und A. E. Mark, *The effect of motional averaging on the calculation of NMR-derived structural properties*, Proteins, 36, 542–555 (1999). 129
- [296] G. Colombo, D. Roccatano und A. E. Mark, *Folding and stability of the three-stranded beta-sheet peptide betanova: Insights from molecular dynamics simulations*, Proteins, 46, 380–392 (2002). 129
- [297] L. J. Smith, A. E. Mark, C. M. Dobson und W. F. Vangunsteren, *Comparison of Md Simulations and Nmr Experiments for Hen Lysozyme - Analysis of Local Fluctuations, Cooperative Motions, and Global Changes*, Biochemistry, 34, 10918–10931 (1995). 129
- [298] U. Stocker und W. F. van Gunsteren, *Molecular dynamics simulation of hen egg white lysozyme: A test of the GROMOS96 force field against nuclear magnetic resonance data*, Proteins, 40, 145–153 (2000). 129
- [299] U. Stocker, K. Spiegel und W. F. van Gunsteren, *On the similarity of properties in solution or in the crystalline state: A molecular dynamics study of hen lysozyme*, J. Biomol. NMR, 18, 1–12 (2000). 129
- [300] C. Peter, X. Daura und W. F. van Gunsteren, *Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations*, J. Biomol. NMR, 20, 297–310 (2001). 129

- [301] K. A. Feenstra, C. Peter, R. M. Scheek, W. F. van Gunsteren und A. E. Mark, *A comparison of methods for calculating NMR cross-relaxation rates (NOESY and ROESY intensities) in small peptides*, J. Biomol. NMR, 23, 181–194 (2002). 129
- [302] H. Arthanari, K. J. McConnell, R. Beger, M. A. Young, D. L. Beveridge und P. H. Bolton, *Assessment of the molecular dynamics structure of DNA in solution based on calculated and observed NMR NOESY volumes and dihedral angles from scalar coupling constants*, Biopolymers, 68, 3–15 (2003). 129
- [303] H. X. Lei und P. E. Smith, *The effects of internal water molecules on the structure and dynamics of chymotrypsin inhibitor 2*, J. Phys. Chem. B, 107, 1395–1402 (2003). 130
- [304] R. H. Stote, A. P. Dejaegere, J. F. Lefevre und M. Karplus, *Multiple conformations of RGDW and DRGDW: A theoretical study and comparison with NMR results*, J. Phys. Chem. B, 104, 1624–1636 (2000). 130
- [305] C. Baysal und H. Meirovitch, *Determination of the stable microstates of a peptide from NOE distance constraints and optimization of atomic solvation parameters*, J. Am. Chem. Soc., 120, 800–812 (1998).
- [306] M. Philippopoulos, A. M. Mandel, A. G. Palmer und C. Lim, *Accuracy and precision of NMR relaxation experiments and MD simulations for characterizing protein dynamics*, Proteins, 28, 481–493 (1997).
- [307] M. Philippopoulos und C. Lim, *Exploring the dynamic information content of a protein NMR structure: Comparison of a molecular dynamics simulation with the NMR and X-ray structures of Escherichia coli ribonuclease HI*, Proteins, 36, 87–110 (1999).
- [308] M. R. Lee und P. A. Kollman, *Free-energy calculations highlight differences in accuracy between X-ray and NMR structures and add value to protein structure prediction*, Structure, 9, 905–916 (2001).
- [309] T. Lazaridis und M. Karplus, *Thermodynamics of protein folding: a microscopic view*, Biophys. Chem., 100, 367–395 (2003). 135
- [310] J. Bredenbeck, J. Helbing, R. Behrendt, C. Renner, L. Moroder, J. Wachtveitl und P. Hamm, *Transient 2D-IR spectroscopy: Snapshots of the nonequilibrium ensemble during the picosecond conformational transition of a small peptide*, J. Phys. Chem. B, 107, 8654–8660 (2003). 168
- [311] J. Bredenbeck, J. Helbing, A. Sieg, T. Schrader, W. Zinth, C. Renner, R. Behrendt, L. Moroder, J. Wachtveitl und P. Hamm, *Picosecond conformational transition and equilibration of a cyclic peptide*, Proc. Natl. Acad. Sci. USA, 100, 6452–6457 (2003). 168
- [312] W. Zinth und W. Kaiser, *Ultrafast Coherent Spectroscopy*, in W. Kaiser (Herausgeber), *Ultrashort Laser Pulses - Generation and Applications*, Band 60 von *Topics in Applied Physics*, 236–277, Springer-Verlag, 2. Auflage (1992). 175
- [313] *IDL, the Interactive Data Language, Version 5.4*, Research Systems, Inc., Boulder (2000). 176

- [314] H. Satzger, *Untersuchung initialer Schritte der Peptidfaltung mit Ultrakurzzeitspektroskopie*, Dissertation, Lehrstuhl für BioMolekulare Optik, Fakultät für Physik, LMU München (2004). 176
- [315] H. Satzger, S. Spörlein, C. Root, J. Wachtveitl, W. Zinth und P. Gilch, *Fluorescence spectra of trans- and cis-azobenzene - mission from the Franck-Condon state*, Chem. Phys. Lett., 372, 216–223 (2003). 181

Danksagung

Allen, die zum Erfolg dieser Arbeit beigetragen haben, möchte ich hiermit ein herzliches Dankeschön aussprechen. Ganz besonders gilt dies für meinem Betreuer Prof. Paul Tavan, der nicht nur das hochinteressante Thema vorschlug, sondern auch stets mit seiner unfehlbaren wissenschaftlichen Intuition, sowie Rat und Tat zur Seite stand, wenn es einmal wirklich knifflig wurde. Seine Einstellung zu der Frage, was anspruchsvolle Forschung und einen guten wissenschaftlichen Schreibstil ausmacht, wird mir stets ein Leitbild sein.

Allen Mitgliedern unserer Arbeitsgruppe möchte ich für die stete Hilfsbereitschaft und gute Laune danken. Eine derart kollegiale Arbeitsatmosphäre ist sicherlich selten. Hierzu beigetragen hat beispielsweise die Nachsichtigkeit von Prof. Marco Nonella mit uns chemisch ungebildeten Physikern, das unbändige Lachen meines Tischkollegens Dr. Paul Strodel und die unerschütterliche Gelassenheit von Gerald Mathias. Zusammen mit Gerald möchte ich Andreas Weiss für die konstruktive Zusammenarbeit beim Aufbau und der Wartung unseres Rechner-Verbundes danken. Andreas ist ferner dafür verantwortlich, dass immer ein gewisser sportlicher Gedanke in der Gruppe herrschte. In diesem Zusammenhang möchte ich Bernhard Egwolf für seine Anregung danken, beim Skilaufen auch mal eine unkontrollierte Schussfahrt zu wagen.

In den Diskussionen mit Martina Storck sind mir so manche Details der statistischen Datenauswertung erst richtig klar geworden. Matthias Schmitz wiederum möchte ich für die engagierte Verteidigung eines klaren grafischen Stils danken, ohne die so manche Dinge viel zu bunt geworden wären. Schließlich sei Thomas Hirschberger dafür gelobt, dass er mich während seiner Diplomarbeit ertragen hat (obwohl ich manchmal nerven kann). Dank auch an die anderen Diplomanden, die zu meiner Zeit die Gruppe bereichert haben. Ein gutes Gelingen an meine designierte Nachfolgerin Verena Schultheis.

In meiner Zeit am BMO bildete das mit Frau Podolski, Frau Bekavac und Frau Klemradt besetzte Sekretariat das Rückgrat des Lehrstuhls. Mein Dank gilt diesen drei Damen und ihrer freundlichen Hilfsbereitschaft.

Bei Prof. Wolfgang Zinth, unserem Lehrstuhlinhaber, möchte ich mich für das große Interesse bedanken, das er meinem Forschungsvorhaben gegenüber gezeigt hat und durch das ich sehr stark motiviert wurde. Auch den anderen Forschern, mit denen ich im Rahmen des SFBs 533 zusammenarbeiten durfte, gilt mein Dank für die einsichtsreichen Diskussionen und die vielen Anregungen. Namentlich sind dies besonders Prof. Josef Wachtveitl und Dr. Christian Renner, sowie die Doktoranden Dr. Sebastian Spörlein und Helmut Satzger.

Einen herzlichen Dank auch an alle anderen „Experimentalos“ für die nette Umgebung und den ganzen ausgegebenen Kuchen. Ich hoffe, mich hier noch revanchieren zu können.

Schließlich hat meine Familie einen wirklich nicht zu unterschätzenden Anteil am Gelingen dieser Arbeit gehabt. Meine Eltern trugen dazu nicht nur durch häufige abendliche Raubtierfütterungen bei, sondern haben es auch auf sich genommen, den nicht in allen Teilen gleichmäßig spannenden Text dieser Arbeit zu korrigieren. Das gleiche gilt auch für die anderen Korrekturleser: Nadine Siemer, Nicole Hacke, Roland Jesseit und Holger Raabe.

Meinem Bruder Dirk und seiner Freundin Michelle habe ich dafür zu danken, dass sie mir geholfen haben, in dieser Zeit den Kontakt zur Realität der weiten Welt da draußen zu halten.

Last but not least, hat sich meine Freundin Nadia stets – und egal was auch kam – hinter mich gestellt. Danke Nadia, ohne Dich hätte ich das hier nicht geschafft!