



Ab initio- und  
Quantenmechanik/Molekularmechanik-  
Molekulardynamik Implementierung für  
wellenfunktionsbasierte quantenchemische  
Methoden

Dissertation

Zur Erlangung des Grades  
DOKTOR DER NATURWISSENSCHAFTEN  
im Promotionsfach Chemie

Am Fachbereich Chemie, Pharmazie, Geographie  
und Geowissenschaften der  
Johannes Gutenberg-Universität Mainz

von

**Till Sven Nils Kirsch**

geboren in Wiesbaden-Dotzheim

Mainz, 2022



1. Berichtstatter: [REDACTED]

2. Berichtstatter: [REDACTED]

Tag der mündlichen Prüfung: 20.09.2022



Die vorliegende Arbeit wurde in der Zeit von November 2017 bis März 2022 am Department Chemie der Johannes Gutenberg-Universität Mainz im Arbeitskreis von [REDACTED] erstellt.

### **Danksagung**

Mein besonderer Dank gilt [REDACTED] für die Möglichkeit, die Dissertation in seinem Arbeitskreis anfertigen zu dürfen, die spannende und herausfordernde Zielsetzung und die gute und hilfreiche Betreuung. Des Weiteren danke ich [REDACTED] und [REDACTED] für die zahlreichen guten und aufschlussreichen Diskussionen.

Dem gesamten Arbeitskreis möchte ich für die wundervolle Zeit danken und insbesondere meinen Bürokolleginnen für die intensiven fachbezogenen sowie nicht fachbezogenen Gespräche.

Ebenso danke ich dem TRR 146 für die Finanzierung der Promotion, das Ermöglichen von Forschungsaufenthalten im Ausland und den Besuch von Konferenzen und Sommerschulen sowie den jährlichen Retreats, bei denen ich viele tolle Menschen kennenlernen konnte.

Abschließend danke ich meinen Freunden, meiner Freundin und insbesondere meinen Eltern, die mich unermüdlich beim Erreichen meiner Ziele unterstützt haben.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Theoretische Grundlagen</b>	<b>5</b>
2.1	Das quantenchemische Problem und Born-Oppenheimer Dynamik . . .	5
2.2	Der QM/MM-Ansatz . . . . .	8
2.2.1	Additiver- und Subtraktiver Ansatz . . . . .	8
2.2.2	Electrostatic Embedding . . . . .	10
2.2.3	Short-Range elektrostatische Wechselwirkungen . . . . .	11
2.2.4	Long-Range elektrostatische Wechselwirkungen . . . . .	12
2.2.5	QM/MM-Grenze durch kovalente Bindungen . . . . .	14
2.3	Molekulardynamik-Simulationen . . . . .	16
2.4	Wahl der QM- und MM-Methoden . . . . .	19
2.4.1	Wahl der MM-Methoden . . . . .	19
2.4.2	Kraftfelder . . . . .	19
2.4.3	Wahl der QM-Methoden . . . . .	21
2.4.4	Hartree-Fock-Theorie (HF) . . . . .	22
2.4.5	Wellenfunktionsbasierte post-Hartree-Fock-Methoden . . . . .	25
2.4.6	Complete Active Space Self Consistent Field (CAS-SCF) . . . . .	29
2.4.7	Dichtefunktionaltheorie (DFT) . . . . .	30
2.5	Multiple-Time-Step Algorithmus . . . . .	33
2.5.1	Der MTS-Algorithmus in CPMD . . . . .	33
<b>3</b>	<b>Implementierung</b>	<b>35</b>
3.1	Level 1: Das Interface . . . . .	36
3.2	Level 2: Die <i>mimic_driver</i> -Routine . . . . .	39
3.3	Level 3: Die QM/MM-Wechselwirkungen . . . . .	41
3.3.1	Berechnung der QM/MM-Wechselwirkungsenergien . . . . .	41
3.3.2	Berechnung der QM/MM-Kräfte . . . . .	43
3.4	Geometrieoptimierung . . . . .	45
3.5	Optimierung von ECPs . . . . .	47
3.6	Keywords und Input-Dateien . . . . .	50
3.6.1	Spezifische Keywords für CFOUR . . . . .	50
3.6.2	Weitere Keywords für CFOUR . . . . .	51
3.6.3	CPMD Keywords zur Benutzung des Interfaces . . . . .	52
3.6.4	Input-Datei für eine QM/MM-Simulation . . . . .	53
3.6.5	Input-Dateien für eine ECP-Optimierung . . . . .	53

<b>4</b>	<b>Simulationsdetails</b>	<b>57</b>
4.1	Aufbau, Equilibrierung und Durchführung der Simulationen . . . . .	57
4.1.1	AIMD-Systeme . . . . .	57
4.1.2	QM/MM-Systeme . . . . .	58
4.2	Liste der QM/MM-Systeme . . . . .	59
4.2.1	Wassersysteme (klein) . . . . .	59
4.2.2	Wassersystem (groß) . . . . .	60
4.2.3	Essigsäuresysteme . . . . .	60
4.2.4	Methanolsystem . . . . .	60
4.2.5	Poly- $\beta$ -alaninsystem . . . . .	61
<b>5</b>	<b>Stabilität und Funktionalität der Implementierung</b>	<b>63</b>
5.1	Energieerhaltung . . . . .	63
5.2	Polarisation des QM-Systems . . . . .	65
5.3	Struktur der Lösungsumgebung . . . . .	67
5.4	Zusammenfassung . . . . .	69
<b>6</b>	<b>Beschleunigung der Simulationen</b>	<b>71</b>
6.1	Long-Range Wechselwirkungen . . . . .	71
6.1.1	Untersuchung der Konvergenz an Snapshot-Rechnungen . . . . .	71
6.1.2	QM/MM-Simulationen mit long-range-Wechselwirkungen . . . . .	74
6.2	Multiple Time Step Algorithmus . . . . .	76
6.3	Zusammenfassung . . . . .	82
<b>7</b>	<b>Optimierung und Verwendung von ECPs</b>	<b>83</b>
7.1	Optimierung der ECPs . . . . .	83
7.2	Test der ECPs in QM/MM-Simulationen von flüssiger Essigsäure . . . . .	88
7.3	Test der ECPs in QM/MM-Simulationen an einem Poly- $\beta$ -alanin . . . . .	93
7.4	Zusammenfassung . . . . .	96
<b>8</b>	<b>Schwingungsfrequenzen aus AIMD- und QM/MM-Simulationen</b>	<b>97</b>
8.1	Untersuchung der Anharmonizität in Schwingungsfrequenzen aus AIMD-Simulationen . . . . .	97
8.2	Schwingungsfrequenzen aus QM/MM-Simulationen . . . . .	106
8.2.1	Berechnung der Schwingungsfrequenzen aus QM/MM-Trajektorien . . . . .	106
8.2.2	Berechnung der Schwingungsfrequenzen aus Snapshots einer QM/MM-Trajektorie . . . . .	111
8.3	Zusammenfassung . . . . .	113
<b>9</b>	<b>Limitationen des Interfaces und Ausblick</b>	<b>115</b>
<b>10</b>	<b>Zusammenfassung</b>	<b>119</b>
<b>A</b>	<b>Ableitungen der Multipolmomente</b>	<b>133</b>

B ECPs	137
C Schwingungsfrequenzen aus den AIMD-Simulationen	145



---

# 1 Einleitung

Computersimulationen gewinnen in vielen verschiedenen Bereichen der Wissenschaft eine immer größere Bedeutung.<sup>[1]</sup> Das zeigt z.B. die Verleihung des Physiknobelpreises 2021 für die Modellierung des Klimas und für bahnbrechende Fortschritte beim Verständnis komplexer Systeme.<sup>[2]</sup> Computerbasierte Simulationen ermöglichen dabei die Untersuchung von Prozessen und Phänomenen, welche experimentell nur schwer oder überhaupt nicht zugänglich sind.<sup>[1,3]</sup> Der wachsende Erfolg von Computersimulationen ist nicht zuletzt der Steigerung der Rechenleistung durch immer schnellere Prozessoren und effizientere Parallelisierungsalgorithmen zu verdanken.<sup>[1]</sup>

Besonders in der Chemie kommen Simulationsmethoden zur Beschreibung von großen und komplexen (bio)chemischen Systemen immer häufiger zum Einsatz.<sup>[1,3-5]</sup> Je nach der Art der zu untersuchenden physikalischen oder chemischen Eigenschaften oder des zu untersuchenden Prozesses und der Zeitskala, auf welcher der Prozess stattfindet, müssen unterschiedliche Methoden verwendet werden. In Abbildung 1.1<sup>[6]</sup> ist die Größen- und Zeitskala der verschiedenen Methoden illustriert. Dabei reicht das Spektrum der zu verwendenden Methoden von einer komplett quantenmechanischen (QM) Beschreibung,<sup>[5,7,8]</sup> über eine atomistische molekularmechanische (MM) Modellierung,<sup>[7]</sup> bis hin zu vergrößernden *coarse-grained* (CG)-Methoden<sup>[9,10]</sup> und feldbasierenden Methoden wie Kontinuum-Methoden<sup>[11]</sup> und hybriden Partikel-Feld (hPF)-Methoden.<sup>[12]</sup>

Bei sogenannten Multiskalenmethoden werden die unterschiedlichen Ansätze so kombiniert, dass jeder Bereich des zu untersuchenden Systems mit der Methode beschrieben wird, welche die für den entsprechenden Bereich ausreichende Genauigkeit besitzt.<sup>[4,5]</sup> So lassen sich große Systeme effizient simulieren, ohne dass das gesamte System mit der genauesten, aber auch rechenaufwendigsten Methode beschrieben werden muss.

Auch in Systemen mit mehreren tausenden Atomen ist die explizite Beschreibung der elektronischen Struktur durch quantenchemische Methoden für die Untersuchung von chemischen Prozessen essentiell.<sup>[3-5,13,14]</sup> Unabhängig von den Fortschritten in der Entwicklung von linear skalierenden Methoden<sup>[15]</sup> und der Parallelisierung<sup>[16]</sup> von quantenchemischen Methoden ist eine komplette Beschreibung auf QM Niveau nicht möglich. Die Entwicklung von kombinierten Quantenmechanik-Molekularmechanik (QM/MM) Methoden durch Warshel und Levitt<sup>[17]</sup> lieferte den Grundstein für eine Vielzahl an QM/MM Implementierungen zur Modellierung von komplexen, chemischen Prozessen.<sup>[4,5]</sup> Dazu gehören (bio)chemische Enzymreaktionen in Lösung,<sup>[18-20]</sup> Katalysereaktionen<sup>[14,21]</sup> und die Beschreibung photoaktiver, biologischer Systeme.<sup>[14,22,23]</sup>

Bei der Anwendung von QM/MM-Methoden spielt die Wahl der verwendeten

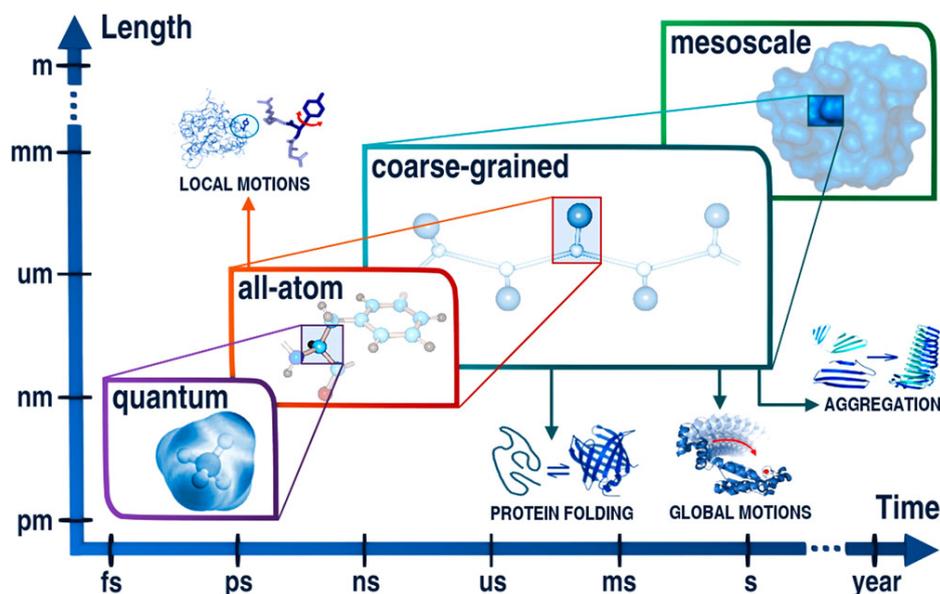


Abbildung 1.1: Anwendungsbereich verschiedener Simulationenethoden (QM, MM, CG und hPF/Kontinuum (mesoscale)) mit der jeweiligen Größenskala und Zeitskala. Multiskalenmethoden verbinden zwei oder mehr Methoden.<sup>[6]</sup>

QM-Methode eine große Rolle. Die Vielfalt an verfügbaren Methoden reicht von sehr genauen, aber rechenaufwändigen Methoden, wie der Coupled-Cluster-Theorie<sup>[24]</sup> (CC), bis hin zu kostengünstigen, aber ungenauen semiempirischen Methoden.<sup>[7]</sup> Die Dichtefunktionaltheorie<sup>[25,26]</sup> (DFT) erweitert das Spektrum an Methoden zusätzlich durch eine große Anzahl an hierarchisch organisierten Funktionalen, anfangend mit der einfachen lokalen Dichtenäherung (LDA) bis zu doppel-hybrid Funktionalen.<sup>[27]</sup> Wie hoch die benötigte Genauigkeit ist, hängt ebenfalls von dem zu untersuchenden System ab. Daher ist es wünschenswert, innerhalb einer QM/MM-Implementierung möglichst einfach und flexibel zwischen den verschiedenen QM-Methoden zu wechseln.

Das Programmpaket CP2K<sup>[28]</sup> ermöglicht beispielsweise neben der Verwendung von DFT mit einer Vielzahl an Funktionalen auch die Verwendung der Hartree-Fock-Methode<sup>[8]</sup> und der Møller-Plesset Störungstheorie.<sup>[29]</sup> Da es allerdings praktisch unmöglich ist, alle verfügbaren QM- und MM-Methoden innerhalb eines Programmpakets zu kombinieren, ist der am häufigsten verwendete Ansatz, unterschiedliche Programmpakete zu verknüpfen. Dieser Ansatz wurde u.a. für die Programmpakete AMBER,<sup>[30]</sup> GROMACS,<sup>[31]</sup> NAMD<sup>[32]</sup> und das PUPIL<sup>[33]</sup> System implementiert. Der Vorteil dieses Ansatzes ist die Flexibilität, die es erlaubt neu entwickelte Methoden durch die Verknüpfung neuer Programme mit relativ niedrigem Aufwand zu verwenden. Der Nachteil ist allerdings, dass die Effizienz aufgrund der Kommunikation der Programme oder durch verschwendete Zeit für die Initialisierung bzw. das Beenden von externen Programmen reduziert wird.<sup>[34]</sup>

Das kürzlich entwickelte MiMiC-Framework<sup>[34,35]</sup> bietet eine hohe Flexibilität und

---

eine hohe Effizienz der QM/MM-Implementierung durch eine effiziente Kommunikation der beteiligten Programme über ein *Message Passing Interface* (MPI). Die ersten Programmpakete, welche über MiMiC verknüpft wurden, sind CPMD<sup>[36]</sup> (QM) und GROMACS<sup>[31,37]</sup> (MM). Die Wahl der QM-Methode war dadurch auf DFT beschränkt. Es ist allerdings bekannt, dass DFT nicht für alle Problemstellungen gute Ergebnisse liefert.<sup>[27]</sup> Beispielsweise führt der Selbstwechselwirkungsfehler<sup>[38]</sup> aufgrund des genäherten Austausch-Korrelations-Funktionalen bei der Berechnung von Bindungsenergien in Radikal-Solvent-Komplexen zu großen Fehlern.<sup>[39]</sup> Auch bei der Berechnung von Ionisierungsenergien und der Beschreibung von komplexen Metall-oxidverbindungen führt der Selbstwechselwirkungsfehler zu Ungenauigkeiten.<sup>[40,41]</sup> Bei der Bestimmung von Anregungsenergien werden DFT-Methoden zwar immer besser,<sup>[42]</sup> allerdings kann besonders bei Charge-Transfer-Systemen nicht die Genauigkeit von CC-Methoden erreicht werden.<sup>[43–45]</sup> Daher sind für diese Fälle genauere Methoden wie die CC-Theorie nötig, um das System adäquat zu beschreiben.<sup>[7,46,47]</sup>

Das CFOUR-Programmpaket<sup>[48]</sup> verfügt neben der Hartree-Fock-SCF-Methode und der Møller-Plesset Störungstheorie über eine Reihe an hochgenauen CC-Methoden. Dabei können Grund- und angeregte Zustände (EOM-CC) sowie Single- und Multi-referenzsysteme (CAS-SCF und MRCC) beschrieben werden. Dabei liegt der Fokus auf der Berechnung von molekularen und spektroskopischen Eigenschaften.<sup>[48]</sup>

Das Ziel dieser Arbeit war es, ein Interface zwischen dem schon bestehenden MiMiC-Framework und dem CFOUR-Programmpaket<sup>[48]</sup> zu implementieren und dadurch das Spektrum der im MiMiC-Framework zugänglichen QM-Methoden zu erweitern. Außerdem ist dies der erste Schritt zur Entwicklung eines neuen Multiskalen-Frameworks, welches sowohl hochgenaue quantenchemischen Methoden als auch CG-Methoden bis hin zur hPF-Theorie beinhaltet.

Dazu werden im zweiten Kapitel zunächst alle relevanten Grundlagen des QM/MM-Ansatzes diskutiert und im Weiteren die theoretischen Details der Implementierung und verwendeten Methoden erläutert. Im dritten Kapitel wird die Implementierung des Interfaces und der Algorithmus zur Geometrieoptimierung erläutert. Weiterhin wird die Implementierung des Algorithmus zur Optimierung von Pseudopotentialen beschrieben, welche verwendet werden, wenn die Grenze zwischen QM- und MM-Bereich durch kovalente Bindungen verläuft. Im vierten Kapitel werden die Details zum Aufbau der Systeme und zur Durchführung der Simulationen beschrieben. Die Funktionalität des implementierten Interfaces wird in Kapitel fünf gezeigt. Dazu wurden Simulationen an flüssigem Wasser durchgeführt und die erhaltenen Eigenschaften mit experimentellen Daten verglichen. Anschließend werden in Kapitel sechs zwei verschiedene Methoden zur Verringerung des Rechenaufwands vorgestellt und die Effektivität durch Simulationen an flüssigem Wasser und gasförmigen Fluorwasserstoff demonstriert. In Kapitel sieben werden die Ergebnisse der Optimierung der Pseudopotentiale vorgestellt und am Beispiel von Poly- $\beta$ -alanin gezeigt, dass sie für die Verwendung in QM/MM-Simulationen von größeren chemischen Systemen geeignet sind. Kapitel acht befasst sich mit der Berechnung von Schwingungsfrequenzen ausgehend von QM/MM-Simulationen und der Frage, wie viel Anharmonizität diese beinhalten. In Kapitel neun werden die

Limitationen des Interfaces diskutiert und Möglichkeiten aufgezeigt, wie diese Limitationen überwunden werden können. Zudem wird ein Ausblick auf zukünftige Erweiterungen und Anwendungen des Interfaces gegeben. Abschließend werden die wichtigsten Ergebnisse der Arbeit zusammengefasst.

---

## 2 Theoretische Grundlagen

In diesem Kapitel werden die Grundlagen der in dieser Arbeit verwendeten Theorien erläutert. Dazu werden zunächst die Konzepte der Born-Oppenheimer Molekulardynamik vorgestellt, sowie die Notwendigkeit und die Grundlagen des QM/MM-Ansatzes erläutert. Anschließend werden die allgemeinen Grundlagen von Molekulardynamik-Simulationen zusammengefasst und die verwendeten Kraftfeld- und quantenchemischen Methoden vorgestellt.

### 2.1 Das quantenchemische Problem und Born-Oppenheimer Dynamik

Zur kompletten Beschreibung eines chemischen Systems muss die quantenmechanische Natur der Elektronen und Kerne berücksichtigt werden. Die zeitliche Entwicklung eines nicht-relativistischen Systems wird durch die zeitabhängige Schrödingergleichung (SG)<sup>[5]</sup>

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

mit der Wellenfunktion  $\Psi$  und dem Hamiltonoperator  $\hat{H}$  beschrieben. Da die zeitabhängige SG für Vielkörpersysteme nicht exakt lösbar ist<sup>[8]</sup> und auch das näherungsweise Lösen zur Beschreibung der Dynamik von Kernen auf quantenmechanischem Level sehr aufwändig ist,<sup>[5,49]</sup> müssen weitere grundlegende Vereinfachungen vorgenommen werden.

Ein weit verbreitetes Vorgehen ist es, ausgehend von der nicht-relativistischen, zeitunabhängigen SG

$$\hat{H} \Psi = E \Psi, \quad (2.2)$$

die Kern- und die Elektronenbewegungen separiert voneinander zu betrachten und die Wellenfunktion in das entsprechende Produkt aus Kern- und elektronischer Wellenfunktion zu überführen. Dies ist möglich, da aufgrund der großen Massendifferenz zwischen Atomkernen und Elektronen die nichtadiabatische Kopplung der Bewegungen sehr klein ist<sup>[5]</sup> und im Rahmen der Born-Oppenheimer-Näherung<sup>[50]</sup> vernachlässigt werden kann. Daraus folgt, dass die Kernpositionen aus Sicht der Elektronen als ortsfest betrachtet werden. Dadurch wird die elektronische SG<sup>[8]</sup>

$$\hat{H}^{\text{el}} \Psi^{\text{el}} = E^{\text{el}} \Psi^{\text{el}} \quad (2.3)$$

erhalten. Gleichung 2.3 stellt ein Eigenwertproblem mit dem elektronischen Hamiltonoperator  $\hat{H}^{\text{el}}$ , den zugehörigen Eigenfunktionen  $\Psi^{\text{el}}$  und dem Energieeigenwert  $E^{\text{el}}$  dar. Der elektronische Hamiltonoperator für ein Molekül besteht aus dem Operator für die kinetische Energie der Elektronen ( $\hat{T}_e$ ), dem Elektron-Elektron-Potential ( $\hat{V}_{ee}$ ) und dem Kern-Elektron-Potential ( $\hat{V}_{Ke}$ ) und hat in atomaren Einheiten die folgende Form:<sup>[8]</sup>

$$\begin{aligned}\hat{H}^{\text{el}} &= \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ke} \\ &= -\sum_i^N \frac{1}{2} \Delta_i + \sum_{i<j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_i^N \sum_I^M \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|}.\end{aligned}\quad (2.4)$$

Hierbei laufen die Summen über alle  $N$  Elektronen bzw.  $M$  Kerne im Molekül.  $Z_I$  ist die Kernladung und der Laplace-Operator  $\Delta$  beinhaltet die zweiten Ableitungen nach allen Ortskoordinaten des Teilchens. Die elektronische Wellenfunktion  $\Psi^{\text{el}}$  beschreibt die Bewegung der Elektronen im Molekül und hängt explizit von den Koordinaten der Elektronen ab, jedoch nur parametrisch von den Koordinaten der Atomkerne. Die Gesamtenergie eines Moleküls (für eine feste Kernanordnung) berücksichtigt zusätzlich auch die elektrostatischen Kern-Kern-Wechselwirkungen:<sup>[8]</sup>

$$E^{\text{tot}} = E^{\text{el}} + \sum_{I<J}^M \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}.\quad (2.5)$$

Wenn nicht anders angegeben, befasst sich dieser Abschnitt ausschließlich mit der Lösung der elektronischen SG sowie mit der Energie des Moleküls für eine feste Kernanordnung, daher ist im Weiteren mit Hamiltonoperator immer der elektronische Hamiltonoperator  $\hat{H}^{\text{el}}$  und mit Energie immer  $E^{\text{tot}}$  bzw.  $E^{\text{BO}}$  gemeint.

Um die Kernbewegung zu beschreiben kann jetzt die Kern-SG gelöst werden. Es genügt meist allerdings, die Kernbewegung von Atom  $I$  mittels klassischer Bewegungsgleichungen<sup>[7]</sup>

$$\mathbf{F}_I = m_I \cdot \mathbf{a}_I,\quad (2.6)$$

mit der Kraft  $\mathbf{F}_I$ , der Masse  $m_I$  und der Beschleunigung  $\mathbf{a}_I$ , zu beschreiben. Dazu wird die *potential energy surface*<sup>[8]</sup> (PES), auf der sich die Kerne bewegen, mittels Lösen der zeitunabhängigen elektronischen SG (Gleichung 2.3) bei entsprechender Konfiguration der Kerne berechnet. Die Kräfte, welche auf die Kerne wirken, werden über den Gradienten der erhaltenen Born-Oppenheimer-Energie<sup>[5]</sup>

$$\mathbf{F}_I = -\nabla_I E^{\text{BO}}(\mathbf{R}_I)\quad (2.7)$$

erhalten. Die Beschreibung der Dynamik der Atome auf diese Weise wird *Born-Oppenheimer* Molekulardynamik<sup>[51]</sup> (BOMD) genannt.

Die Zeitschritte, welche im Algorithmus zum Lösen der Bewegungsgleichung verwendet werden (siehe Abschnitt 2.3), sind typischerweise kleiner als eine Femtosekun-

de.<sup>[5,7]</sup> Daher sind mehrere hunderttausende Schritte nötig, um chemisch relevante Zeitskalen zu simulieren. Für jeden Zeitschritt müssen die zeitunabhängige elektronische SG gelöst und anschließend die Gradienten berechnet werden. Da der Rechenaufwand zur Berechnung der Gradienten auf quantenchemischen Niveau relativ hoch ist und mit der Größe des zu beschreibenden Systems steigt, ist BOMD, je nach Wahl der quantenchemischen Methode, auf nur relativ kleine Systeme (ca. 10-1000 Atome) anwendbar.

Eine weitere Methode, die wie BOMD zur Gruppe der *ab initio*-MD Methoden<sup>[52]</sup> (AIMD) gehört, ist die *Car-Parinello*-MD<sup>[53]</sup> (CPMD) Methode. Diese berücksichtigt zusätzlich eine fiktive klassische Dynamik der Freiheitsgrade der Elektronen.<sup>[52,53]</sup>

## 2.2 Der QM/MM-Ansatz

Die AIMD-Methoden sind aufgrund des hohen Rechenaufwands zur Lösung der elektronischen SG nur auf kleine Systeme und kurze Zeitskalen limitiert. Allerdings kann die explizite Beschreibung der elektronischen Struktur durch quantenchemische Methoden für die Untersuchung von chemischen Prozessen auch in großen Systemen nötig sein, für welche eine reine quantenchemische Beschreibung unmöglich ist.<sup>[4,5,13,14,19]</sup> Oftmals reicht es allerdings, wenn nur ein kleiner Ausschnitt (ca. 10 - 100 Atome) des gesamten Systems quantenchemisch und der Rest des Systems mittels Kraftfeldmethoden beschrieben wird. Diese kombinierten QM/MM-Methoden wurden 1976 von Warshel und Levitt entwickelt<sup>[17]</sup> und erlauben die Simulation von Systemen und die Untersuchung von Prozessen, welche durch AIMD-Simulationen nicht zugänglich sind.<sup>[4,5]</sup>

### 2.2.1 Additiver- und Subtraktiver Ansatz

In allen QM/MM-Methoden wird das Gesamtsystem in ein QM-System und ein MM-System unterteilt (siehe Abbildung 2.1). Die meisten heutzutage verwendeten QM/MM-Methoden verwenden entweder einen additiven oder einen subtraktiven Ansatz.<sup>[4,5]</sup> Im subtraktiven Ansatz wird die Energie des Gesamtsystems über folgende Gleichung bestimmt:<sup>[5]</sup>

$$E = E^{\text{QM}}(\text{QM}) + E^{\text{MM}}(\text{QM} + \text{MM}) - E^{\text{MM}}(\text{QM}). \quad (2.8)$$

Hierbei bedeutet  $E^{\text{QM}}(\text{QM})$ , dass die Energie des QM-Systems auf QM-Level berechnet wird,  $E^{\text{MM}}(\text{QM})$  bedeutet, dass die Energie des QM-Systems auf MM-Level

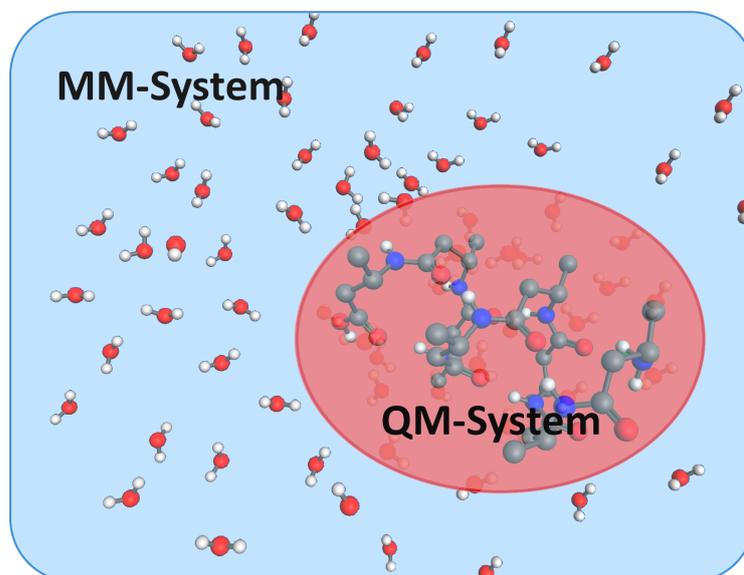


Abbildung 2.1: Schematische Aufteilung eines Systems in QM- und MM-Bereich.

berechnet wird und  $E^{\text{MM}}(\text{QM} + \text{MM})$  bedeutet analog dazu, dass die Energie des Gesamtsystems (QM+MM) auf MM-Level berechnet wird. Im subtraktiven Ansatz werden also nur Rechnungen an isolierten Systemen (QM- oder Gesamtsystem) durchgeführt und die Wechselwirkungen zwischen QM- und MM-System werden über die Differenz  $E^{\text{MM}}(\text{QM} + \text{MM}) - E^{\text{MM}}(\text{QM})$  auf MM-Level beschrieben. Die korrespondierenden Kräfte auf die Atomkerne werden über die Ableitung der jeweiligen Terme bestimmt. Der Vorteil des subtraktiven Ansatzes in dieser Form ist die leichte Implementierung, da keine expliziten QM/MM-Wechselwirkungen berechnet werden müssen.<sup>[5]</sup>

Häufiger wird jedoch ein additiver Ansatz verwendet, in dem die QM/MM-Wechselwirkungen explizit berechnet werden.<sup>[5]</sup> Die Energie des Gesamtsystems ergibt sich dann über

$$E = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM/MM}} \quad (2.9)$$

mit dem zugehörigen Hamiltonoperator

$$\hat{H} = \hat{H}_{\text{QM}} + \hat{H}_{\text{MM}} + \hat{H}_{\text{QM/MM}}, \quad (2.10)$$

der das Gesamtsystem beschreibt. Hierbei ist  $\hat{H}_{\text{QM}}$  der Hamiltonoperator, der das QM-System beschreibt und  $\hat{H}_{\text{MM}}$  der Hamiltonoperator für das MM-System. Die QM/MM-Wechselwirkungen werden über  $\hat{H}_{\text{QM/MM}}$  beschrieben. Es ist zunächst nicht vorgegeben, mit welcher Methode die jeweiligen Beiträge berechnet werden. Der Nachteil des additiven Ansatzes ist, dass hierfür explizite Ausdrücke für die QM/MM-Wechselwirkungen gefunden und implementiert werden müssen.<sup>[4,5,7]</sup> Zudem können bei der Beschreibung der Grenze zwischen QM- und MM-System Probleme auftreten, welche im späteren Verlauf diskutiert werden. Der Vorteil ist, dass die QM/MM-Wechselwirkungen im Prinzip auf jedem theoretischen Level berechnet werden können. In der Regel werden drei verschiedene Methoden unterschiedlicher Komplexität verwendet. Die einfachste Variante ist die Beschreibung über klassische Punktladungen auf MM-Level (*mechanical embedding*<sup>[54]</sup>), ähnlich wie beim subtraktiven Ansatz. Dabei wird die Elektronendichte des QM-Systems allerdings nicht direkt durch die MM-Umgebung beeinflusst, was zu einer falschen Beschreibung der elektronischen Struktur des QM-Systems führen kann. Am häufigsten wird das sogenannte *electrostatic embedding*<sup>[54]</sup> verwendet, dabei werden die MM-Punktladungen als externes Potential zum Hamiltonoperator des QM-Systems hinzugefügt. Dies führt zu einer direkten Polarisierung des QM-Systems durch die MM-Umgebung und dadurch zu einer besseren Beschreibung der elektronischen Struktur des QM-Systems und der QM/MM-Wechselwirkungen. Reicht diese einseitige Polarisierung nicht aus, um die QM/MM-Wechselwirkungen des Systems hinreichend zu beschreiben, können zusätzlich polarisierbare Kraftfelder benutzt werden (*polarizable embedding*<sup>[54,55]</sup>), was dazu führt, dass nicht nur das QM-System polarisiert wird, sondern auch das MM-System durch die Ladungsdichte des QM-Systems.

## 2.2.2 Electrostatic Embedding

Die am häufigsten verwendete Methode, die auch in dieser Arbeit angewendet wurde, ist das *electrostatic embedding*. Hierbei wird der Wechselwirkungshamiltonoperator  $\hat{H}_{\text{QM/MM}}$  in bindende, elektrostatische und van der Waals (vdW) Wechselwirkungen aufgeteilt:<sup>[4]</sup>

$$\hat{H}_{\text{QM/MM}} = \hat{H}_{\text{QM/MM}}^{\text{b}} + \hat{H}_{\text{QM/MM}}^{\text{vdW}} + \hat{H}_{\text{QM/MM}}^{\text{el}}. \quad (2.11)$$

Die bindenden Beiträge  $\hat{H}_{\text{QM/MM}}^{\text{b}}$  sind nur von Null verschieden, wenn die Grenze zwischen QM- und MM-System durch eine kovalente Bindung verläuft. Dieser Fall wird in Abschnitt 2.2.5 diskutiert. Die van der Waals-Beiträge  $\hat{H}_{\text{QM/MM}}^{\text{vdW}}$  werden üblicherweise über ein Lennard-Jones Potential beschrieben.<sup>[4]</sup> Dafür werden meistens die jeweiligen Parameter aus dem MM-Kraftfeld verwendet und die Berechnung wird auch vom jeweiligen MM-Programm übernommen. Dies kann zwar zu Ungenauigkeiten führen, da die Parameter aus dem MM-Kraftfeld in der Regel für bestimmte Systeme wie organische Makromoleküle optimiert sind und nicht zwingend für alle im QM-System vorhandenen Elemente (z.B. Metalle) verwendet werden können bzw. die Parameter nur für MM-MM Wechselwirkungen optimiert wurden. Die dadurch auftretenden Fehler sind in der Regel aber klein und durch die kurzreichweitige Natur der vdW-Wechselwirkungen treten sie lediglich in der Nähe der QM/MM Grenze auf.<sup>[4]</sup> Um zu verhindern, dass diese Fehler die erhaltenen Ergebnisse beeinflussen, kann der QM-Bereich vergrößert werden, sodass die QM/MM Grenze weiter vom relevanten Bereich des QM-Systems entfernt verläuft.

Der elektrostatische Term  $\hat{H}_{\text{QM/MM}}^{\text{el}}$  spielt meist eine dominante Rolle bei den QM/MM-Wechselwirkungen. Die Berechnung des QM-Systems wird in Anwesenheit der MM-Partialladungen durchgeführt, indem die Wechselwirkungsterme zum Eielektronen-Hamiltonoperator des QM-Systems addiert werden, welche folgende Form annehmen:<sup>[4]</sup>

$$\hat{H}_{\text{QM/MM}}^{\text{el}} = - \sum_i^N \sum_A^{M_{\text{MM}}} \frac{q_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_I^{M_{\text{QM}}} \sum_A^{M_{\text{MM}}} \frac{q_A Z_I}{|\mathbf{R}_A - \mathbf{R}_I|}. \quad (2.12)$$

Hierbei beschreibt der erste Term die Wechselwirkung zwischen den Elektronen des QM-Systems mit den Koordinaten  $\mathbf{r}_i$  und den MM-Atomen mit den Koordinaten  $\mathbf{R}_A$  und den Partialladungen  $q_A$ , welche durch das jeweilige MM-Kraftfeld gegeben sind. Der zweite Term beschreibt die Kern-Kern-Wechselwirkung zwischen den QM-Kernen mit den Koordinaten  $\mathbf{R}_I$  und der Kernladung  $Z_I$  und den MM-Atomen. Die Summen laufen über alle  $N$  Elektronen bzw.  $M_{\text{QM}}$  Kerne des QM-Systems und über alle  $M_{\text{MM}}$  MM-Partialladungen.

Die explizite Berechnung der elektrostatischen Wechselwirkungen ist für eine akkurate Beschreibung der Struktur und Eigenschaften des QM-Systems essentiell.<sup>[4]</sup> Die direkte Implementierung des ersten Terms in Gleichung 2.12 ist allerdings mit einem enormen Rechenaufwand verbunden, da die Zahl der zu berechnenden Terme

mit

$$\mathcal{O}(N_p \cdot M_{MM}) \quad (2.13)$$

skaliert, wobei  $N_p$  im Falle einer *plane-wave*-DFT<sup>[52]</sup> Implementierung die Anzahl der Gitterpunkte (ca.  $10^4 - 10^6$ )<sup>[56]</sup> und im Falle einer auf Wellenfunktionen basierenden Implementierung die Anzahl an Einelektronen-Integralen (ca.  $10^3 - 10^5$ ) ist und  $M_{MM}$  die Anzahl an MM-Atomen (ca.  $10^4 - 10^6$  für Biomakromoleküle). Es gibt verschiedene Methoden den Rechenaufwand zu verringern. Der in dieser Arbeit verwendete Ansatz folgt dem von Laio et. al.<sup>[56]</sup>, bei dem der Hamiltonoperator für die QM/MM-Wechselwirkungen ohne Verlust an Genauigkeit in zwei unterschiedliche Beiträge aufgesplittet wird. Dazu werden die Atome des MM-Systems in eine *short-range*-Region (SR) und eine *long-range*-Region (LR) aufgeteilt. Die Einteilung erfolgt über einen cut-off Radius  $r_{cut}$ . Alle MM-Atome, die näher am Zentrum des QM-Systems sind als  $r_{cut}$ , gehören zur SR-Region. Alle MM-Atome, die weiter entfernt sind, gehören zur LR-Region. Analog dazu wird der Wechselwirkungshamiltonoperator  $\hat{H}_{QM/MM}^{el}$  in zwei Terme aufgeteilt:<sup>[56]</sup>

$$\hat{H}_{QM/MM}^{el} = \hat{H}_{sr} + \hat{H}_{lr}. \quad (2.14)$$

### 2.2.3 Short-Range elektrostatische Wechselwirkungen

Die elektrostatischen QM/MM-Wechselwirkungen werden im SR-Bereich explizit berechnet.  $\hat{H}_{sr}$  hat daher die selbe Form wie in Gleichung 2.12:<sup>[56]</sup>

$$\hat{H}_{sr} = - \sum_i^N \sum_A^{M_{sr}} \frac{q_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_I^{M_{QM}} \sum_A^{M_{sr}} \frac{q_A Z_I}{|\mathbf{R}_A - \mathbf{R}_I|}. \quad (2.15)$$

Bei  $\hat{H}_{sr}$  läuft die Summe allerdings ausschließlich über alle  $M_{sr}$  MM-Partialladungen im SR-Bereich. Die für die Lösung der Bewegungsgleichungen nötigen Kräfte werden über den analytischen Gradienten bezüglich der Kernkoordinaten bestimmt, was zu vier verschiedenen Ausdrücken führt. Die ersten beiden Ausdrücke beschreiben die Kräfte, welche aufgrund der Kern-Kern-Wechselwirkungen (2. Term in Gleichung 2.15) auf die QM- bzw. MM-Atome wirken. Die Kraft  $\mathbf{F}_{IA}(A)$  auf ein MM-Atom ist gegeben durch

$$\mathbf{F}_{IA}(A) = \sum_I^{M_{QM}} \frac{q_A Z_I}{|\mathbf{R}_A - \mathbf{R}_I|^3} (\mathbf{R}_A - \mathbf{R}_I) \quad (2.16)$$

und die Kraft  $\mathbf{F}_{IA}(I)$ , die auf ein QM-Atom wirkt, ist gegeben durch

$$\mathbf{F}_{IA}(I) = - \sum_A^{M_{sr}} \frac{q_A Z_I}{|\mathbf{R}_A - \mathbf{R}_I|^3} (\mathbf{R}_A - \mathbf{R}_I). \quad (2.17)$$

Der dritte Ausdruck beschreibt die Kräfte, die auf die MM-Atome aufgrund der Wechselwirkungen mit den QM-Elektronen wirken. Er wird durch den Erwartungswert des Gradienten des ersten Terms in Gleichung 2.15 bezüglich der Koordinaten der MM-Atome  $\mathbf{R}_A$  erhalten. Die Kraft  $\mathbf{F}_{iA}(A)$  auf ein MM-Atom ist durch

$$\mathbf{F}_{iA}(A) = \sum_{\mu\nu} D_{\mu\nu} \left\langle \chi_\mu \left| \frac{q_A}{|\mathbf{R}_A - \mathbf{r}|^3} (\mathbf{R}_A - \mathbf{r}) \right| \chi_\nu \right\rangle \quad (2.18)$$

gegeben. Da die Elektronen nicht explizit beschrieben, sondern über Orbitale dargestellt werden, welche in einer Basis entwickelt werden (siehe Gleichung 2.66), laufen die Summen über alle Basisfunktionen  $\chi_\mu$  und  $\chi_\nu$  mit der Dichtematrix  $D_{\mu\nu}$ . Details zur Basissatzentwicklung der Orbitale und den verwendeten quantenchemischen Methoden sind in Abschnitt 2.4 erläutert. Um den vierten Ausdruck, die korrespondierende Kraft auf die QM-Atome, zu erhalten, wird der Gradient bezüglich der Koordinaten der QM-Atome  $\mathbf{R}_I$ , an denen die Basisfunktionen zentriert sind, gebildet. Die Kraft  $\mathbf{F}_{iA}(I)$  auf ein QM-Atom ist durch

$$\mathbf{F}_{iA}(I) = \sum_A^{M_{sr}} \sum_{\mu\nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_\mu}{\partial \mathbf{R}_I} \left| \frac{q_A}{|\mathbf{R}_A - \mathbf{r}|} \right| \chi_\nu \right\rangle + \left\langle \chi_\mu \left| \frac{q_A}{|\mathbf{R}_A - \mathbf{r}|} \frac{\partial \chi_\nu}{\partial \mathbf{R}_I} \right\rangle \right] \quad (2.19)$$

gegeben.

Zusätzlich entstehen implizite Kräfte auf die QM-Atome, da sich die Molekülorbitale und dadurch auch die Kräfte im QM-System durch die Polarisierung des QM-Systems durch die MM-Partialladungen ändern.

## 2.2.4 Long-Range elektrostatische Wechselwirkungen

Im Gegensatz zu den SR-Wechselwirkungen, werden die QM/MM-Wechselwirkungen für die MM-Partialladungen im LR-Bereich nicht explizit berechnet. Stattdessen werden sie über eine Multipolentwicklung des elektrostatischen Potentials zwischen den MM-Partialladungen und der QM-Elektronendichte beschrieben.<sup>[56]</sup> Dies ist möglich, da sich die QM-Elektronendichte auf den QM-Bereich beschränkt und die Entfernung zu den MM-Partialladungen des LR-Bereichs relativ groß ist. Die Wechselwirkungen mit den QM-Atomen werden analog zu Gleichungen 2.15 bis 2.17 explizit berechnet. Die Multipolentwicklung des elektrostatischen Potentials  $V_{iA}$  bis zur vierten Ordnung mit dem Ursprung Null ist durch

$$\begin{aligned} \sum_i \sum_A \frac{q_A}{|\mathbf{R}_A - \mathbf{r}_i|} &= \sum_A q_A \left[ C \frac{1}{|\mathbf{R}_A|} + \sum_\alpha \mu^\alpha \frac{R_A^\alpha}{|\mathbf{R}_A|^3} + \frac{1}{2} \sum_{\alpha,\beta} Q^{\alpha\beta} \frac{R_A^\alpha R_A^\beta}{|\mathbf{R}_A|^5} \right. \\ &\quad \left. + \frac{1}{6} \sum_{\alpha,\beta,\gamma} O^{\alpha\beta\gamma} \frac{R_A^\alpha R_A^\beta R_A^\gamma}{|\mathbf{R}_A|^7} + \frac{1}{24} \sum_{\alpha,\beta,\gamma,\epsilon} H^{\alpha\beta\gamma\epsilon} \frac{R_A^\alpha R_A^\beta R_A^\gamma R_A^\epsilon}{|\mathbf{R}_A|^9} \right] \quad (2.20) \end{aligned}$$

mit der Ladung  $C$ , dem Dipolmoment  $\mu^\alpha$ , dem Quadrupolmoment  $Q^{\alpha\beta}$ , dem Octupolmoment  $O^{\alpha\beta\gamma}$  und dem Hexadecapolmoment  $H^{\alpha\beta\gamma\epsilon}$  der QM-Elektronendichte gegeben. Dabei beschreiben  $\alpha, \beta$  etc. die kartesischen Komponenten. Für die Multipolmomente ergibt sich aufgrund der Verwendung von Basisfunktionen anstatt der klassischen Ausdrücke eine Integralform:

$$C = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | \chi_\nu \rangle, \quad (2.21)$$

$$\mu^\alpha = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | r^\alpha | \chi_\nu \rangle, \quad (2.22)$$

$$Q^{\alpha\beta} = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | 3r^\alpha r^\beta - |\mathbf{r}|^2 \delta_{\alpha\beta} | \chi_\nu \rangle, \quad (2.23)$$

$$O^{\alpha\beta\gamma} = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | 15r^\alpha r^\beta r^\gamma - 3|\mathbf{r}|^2 (r^\gamma \delta_{\alpha\beta} + r^\beta \delta_{\alpha\gamma} + r^\alpha \delta_{\beta\gamma}) | \chi_\nu \rangle \quad (2.24)$$

und

$$\begin{aligned} H^{\alpha\beta\gamma\epsilon} = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | & 105r^\alpha r^\beta r^\gamma r^\epsilon + 3|\mathbf{r}|^4 (\delta_{\alpha\beta} \delta_{\gamma\epsilon} + \delta_{\alpha\gamma} \delta_{\beta\epsilon} + \delta_{\alpha\epsilon} \delta_{\beta\gamma}) \\ & - 15|\mathbf{r}|^2 (r^\alpha r^\beta \delta_{\gamma\epsilon} + r^\alpha r^\gamma \delta_{\beta\epsilon} + r^\beta r^\gamma \delta_{\alpha\epsilon} + r^\alpha r^\epsilon \delta_{\beta\gamma} + r^\beta r^\epsilon \delta_{\alpha\gamma} + r^\gamma r^\epsilon \delta_{\alpha\beta}) | \chi_\nu \rangle. \end{aligned} \quad (2.25)$$

Die Kraft  $F_{iA}^\alpha(A)$ , welche auf ein MM-Atom wirkt, wird durch den negativen Gradienten von Gleichung 2.20 bezüglich der Koordinaten der MM-Atome  $\mathbf{R}_A$  erhalten und ist durch

$$\begin{aligned} F_{iA}^\alpha(A) = -\frac{\partial V_{iA}}{\partial R_A^\alpha} = q_A \left[ C \frac{R_A^\alpha}{|\mathbf{R}_A|^3} - \mu^\alpha \frac{1}{|\mathbf{R}_A|^3} + 3 \sum_{\beta} \mu^\beta \frac{R_A^\alpha R_A^\beta}{|\mathbf{R}_A|^5} \right. \\ \left. - \sum_{\beta} Q^{\alpha\beta} \frac{R_A^\beta}{|\mathbf{R}_A|^5} + \frac{5}{2} \sum_{\beta,\gamma} Q^{\beta\gamma} \frac{R_A^\alpha R_A^\beta R_A^\gamma}{|\mathbf{R}_A|^7} \right. \\ \left. - \frac{1}{2} \sum_{\beta,\gamma} O^{\alpha\beta\gamma} \frac{R_A^\beta R_A^\gamma}{|\mathbf{R}_A|^7} + \frac{7}{6} \sum_{\beta,\gamma,\epsilon} O^{\beta\gamma\epsilon} \frac{R_A^\alpha R_A^\beta R_A^\gamma R_A^\epsilon}{|\mathbf{R}_A|^9} \right] \end{aligned} \quad (2.26)$$

gegeben. Die korrespondierende Kraft auf die QM-Atome wird durch den negativen Gradienten bezüglich der Koordinaten der QM-Atome  $\mathbf{R}_I$ , an denen die Basisfunk-

tionen zentriert sind, gebildet. Die Kraft  $F_{iA}^\alpha(I)$  auf ein QM-Atom ist durch

$$F_{iA}^\alpha(I) = -\frac{\partial V_{iA}}{\partial R_I^\alpha} = \sum_A q_A \left[ C^{\bar{\alpha}} \frac{1}{|\mathbf{R}_A|} + \sum_\beta \mu^{\beta,\bar{\alpha}} \frac{R_A^\beta}{|\mathbf{R}_A|^3} + \frac{1}{2} \sum_{\beta,\gamma} Q^{\beta\gamma,\bar{\alpha}} \frac{R_A^\beta R_A^\gamma}{|\mathbf{R}_A|^5} + \frac{1}{6} \sum_{\beta,\gamma,\epsilon} O^{\beta\gamma\epsilon,\bar{\alpha}} \frac{R_A^\beta R_A^\gamma R_A^\epsilon}{|\mathbf{R}_A|^7} + \frac{1}{24} \sum_{\beta,\gamma,\epsilon,\kappa} H^{\beta\gamma\epsilon\kappa,\bar{\alpha}} \frac{R_A^\beta R_A^\gamma R_A^\epsilon R_A^\kappa}{|\mathbf{R}_A|^9} \right] \quad (2.27)$$

gegeben. Hierbei sind  $C^{\bar{\alpha}}$ ,  $\mu^{\beta,\bar{\alpha}}$  etc. die Ableitung der Ladung  $C$  bzw. der  $\beta$ -Komponente des Dipolmoments  $\mu^\beta$  nach der  $\alpha$ -Koordinate und sind gegeben durch

$$C^{\bar{\alpha}} = \frac{\partial C}{\partial R_I^\alpha} = \sum_{\mu,\nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_\mu}{\partial R_I^\alpha} \middle| \chi_\nu \right\rangle + \left\langle \chi_\mu \middle| \frac{\partial \chi_\nu}{\partial R_I^\alpha} \right\rangle \right] \quad (2.28)$$

bzw.

$$\mu^{\beta,\bar{\alpha}} = \frac{\partial \mu^\beta}{\partial R_I^\alpha} = \sum_{\mu,\nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_\mu}{\partial R_I^\alpha} \middle| r^\beta \middle| \chi_\nu \right\rangle + \left\langle \chi_\mu \middle| \frac{\partial r^\beta}{\partial R_I^\alpha} \middle| \chi_\nu \right\rangle + \left\langle \chi_\mu \middle| r^\beta \middle| \frac{\partial \chi_\nu}{\partial R_I^\alpha} \right\rangle \right]. \quad (2.29)$$

Die Ableitungen von Quadrupolmoment  $Q^{\alpha\beta}$ , Octupolmoment  $O^{\alpha\beta\gamma}$  und Hexadecapolmoment  $H^{\alpha\beta\gamma\epsilon}$  werden analog berechnet und sind zusammen mit einer möglichen Umformung der Multipolterme aus Gleichung 2.20 und 2.26 für eine einfachere Implementierung in Anhang A angegeben.

Durch die Multipolentwicklung wird die Berechnung der Integrale und deren Ableitungen von der Summation über die MM-Partialladungen entkoppelt. Daher skaliert die Zahl der zu berechnenden Terme nur noch mit

$$\mathcal{O}(N_p + N_{MM}). \quad (2.30)$$

## 2.2.5 QM/MM-Grenze durch kovalente Bindungen

Möchte man den QM-Bereich möglichst klein halten, ist es in biochemischen Systemen mit Makromolekülen häufig unvermeidbar, dass die Grenze zwischen QM- und MM-System durch eine kovalente Bindung verläuft. In diesem Fall verschwindet der bindende Term in Gleichung 2.11 nicht und es entstehen zwei Probleme. Das erste Problem ist die Beschreibung der bindenden Wechselwirkungen an der QM/MM-Grenze. In der Regel wird es dadurch gelöst, dass alle zwei-, drei- oder vieratomigen Wechselwirkungen (siehe Abschnitt 2.4.2), die mindestens ein MM-Atom beinhalten, über die Parameter des MM-Kraftfelds beschrieben werden.<sup>[4]</sup>

Das zweite, kompliziertere Problem ist, dass das QM-System ungesättigte Valenzelektronen enthält, welche die Elektronenstruktur und die chemischen Eigenschaften (Reaktivität) stark beeinflussen. Zur Lösung dieses Problems gibt es drei verschiedene Ansätze. Im häufig verwendeten Link-Atom-Ansatz werden die ungepaarten

Elektronen mit monovalenten Atomen, in der Regel Wasserstoff- oder Fluoratome, abgesättigt (Capping).<sup>[4]</sup> Dieser Ansatz ist zunächst sehr simpel, allerdings werden dem System dadurch zusätzliche Atome mit jeweils drei zusätzlichen Freiheitsgraden hinzugefügt, welche eigentlich nicht vorhanden wären. Diese müssen eingeschränkt werden, indem die Position des Link-Atoms auf der Bindung zwischen QM- und MM-Atom liegen muss. Zudem müssen die zusätzlichen, im eigentlichen System nicht vorhandenen, Wechselwirkungen mit dem Link-Atom aus dem Hamiltonoperator ausgeschlossen werden.<sup>[4]</sup>

Ein zweiter Ansatz ist es, die Orbitale, welche an der QM/MM-Grenze liegen, durch eingefrorene Grenzorbitale zu ersetzen, welche die Elektronendichte der kovalenten Bindung nachahmen.<sup>[4]</sup> Dieser Ansatz hat den Vorteil, dass keine zusätzlichen Atome und Freiheitsgrade eingeführt werden. Allerdings ist der Nachteil, dass er technisch kompliziert ist, da die Orbitale aufgrund der fehlenden Flexibilität vorher an speziellen Modellsystemen entwickelt werden müssen.

Der dritte Ansatz vereint die Vorteile beider vorherigen Ansätze. Beim Verwenden von *Boundary*-Atomen bzw. optimierten *effective core potentials* (ECPs) verläuft die QM/MM-Grenze nicht durch eine Bindung, sondern durch ein Atom.<sup>[4]</sup> Dieses Grenzatom ist entweder ein angepasstes Fluoratom oder besteht aus einem angepassten ECP, welche die Eigenschaften (Bindungslänge, Elektronendichte) einer Methylgruppe nachahmen sollen. Die Grenzatome müssen zwar auch vorher an Modellsystemen optimiert werden, allerdings bleiben sie während einer Simulation flexibel und sie führen keine zusätzlichen Freiheitsgrade ein. Die Optimierung wird mit Hilfe einer Penalty-Funktion durchgeführt.<sup>[57]</sup> Die Form der Penalty-Funktion bestimmt, nach welcher molekularen Eigenschaft das Grenzatom optimiert wird. Beispielsweise eignet sich die Elektronendichte  $\rho(\mathbf{r})$  gut als Referenzeigenschaft für die Optimierung. Die Penalty-Funktion ist dann über

$$P[\rho(\mathbf{r})] = \int d^3r w(\mathbf{r}) |\rho(\mathbf{r})^{ref} - \rho(\mathbf{r})|^2 \quad (2.31)$$

mit der Referenzelektronendichte  $\rho(\mathbf{r})^{ref}$ , welche durch eine QM-Rechnung an einem Referenzsystem erhalten wurde, und einer Gewichtungsfunktion  $w(\mathbf{r})$ , mit der die Integration auf ein bestimmtes Volumen bzw. bestimmte Atome beschränkt werden kann, gegeben.<sup>[57]</sup> In der Optimierung wird die Penalty-Funktion bezüglich der Parameter der ECPs minimiert. Weitere mögliche Referenzeigenschaften sind beispielsweise der Gradient (Molekülstruktur), das Dipolmoment oder Partiaalladungen. Theoretisch können die Grenzatome für jede beliebige Situation angepasst werden, allerdings ist es vorteilhaft eine Bindung zu verwenden, die eine möglichst einfache elektronische Struktur aufweist (Einfachbindung), nicht polar ist und nicht innerhalb einer funktionellen Gruppe im Molekül liegt. Daher hat es sich bewährt, die QM/MM-Grenze durch C-C-Einfachbindungen verlaufen zu lassen.<sup>[4,5]</sup>

Ein ausführlicher Überblick über die vielzähligen Methoden mit den jeweiligen Vor- und Nachteilen ist in Referenz 4 und den Quellen darin zu finden.

## 2.3 Molekulardynamik-Simulationen

Sowohl in klassischen MD-Simulationen als auch bei AIMD- und QM/MM-Simulationen werden die gleichen Algorithmen zur Lösung der klassischen Bewegungsgleichungen verwendet, um die zeitliche Entwicklung des Systems zu beschreiben. Die dazu verwendeten Algorithmen basieren auf einer Taylorentwicklung um den Ort  $\mathbf{R}_I$  eines Teilchens zum Zeitpunkt  $t$ :<sup>[7]</sup>

$$\mathbf{R}_I(t + \Delta t) = \mathbf{R}_I(t) + \mathbf{v}_I(t) + \frac{1}{2}\mathbf{a}_I(t)\Delta t^2 + \frac{1}{6}\frac{\partial^3\mathbf{R}_I(t)}{\partial t^3}\Delta t^3 + \dots, \quad (2.32)$$

mit der Geschwindigkeit  $\mathbf{v}_I$  und der Beschleunigung  $\mathbf{a}_I$  als erste bzw. zweite Ableitung des Ortes  $\mathbf{R}_I$  nach der Zeit. Im Verlet-Algorithmus<sup>[58]</sup> wird die Taylorentwicklung für positive ( $t + \Delta t$ ) und negative ( $t - \Delta t$ ) Zeitschritte nach der zweiten bzw. dritten Ordnung abgebrochen und addiert. Dadurch erhält man die Position  $\mathbf{R}_I(t + \Delta t)$  nach einem Zeitschritt  $\Delta t$  über

$$\mathbf{R}_I(t + \Delta t) = 2\mathbf{R}_I(t) - \mathbf{R}_I(t - \Delta t) + \mathbf{a}_I(t)\Delta t^2, \quad (2.33)$$

welche nur von der aktuellen Position  $\mathbf{R}_I(t)$ , der vorherigen Position  $\mathbf{R}_I(t - \Delta t)$  und der Beschleunigung des Teilchens  $\mathbf{a}_I(t)$ , welche über Gleichung 2.6 erhalten wird, abhängt.<sup>[7]</sup>

Der Verlet-Algorithmus kommt zwar ohne die Berechnung der Teilchengeschwindigkeiten aus, wenn allerdings die Simulationstemperatur kontrolliert werden soll, ist es sinnvoll dazu die Geschwindigkeit der Teilchen zu verwenden und diese ebenfalls zu berechnen. Zudem müssen die vorherige Position  $\mathbf{R}_I(t - \Delta t)$  zum Starten der Simulation anderweitig berechnet werden, da sie nicht bekannt sind. Gleichung 2.33 kann allerdings durch das Einsetzen einer Definition für die Geschwindigkeit<sup>[59]</sup>

$$\mathbf{v}_I(t) = \frac{\mathbf{R}_I(t + \Delta t) - \mathbf{R}_I(t - \Delta t)}{2\Delta t} \quad (2.34)$$

und anschließende Umformung in

$$\mathbf{R}_I(t + \Delta t) = \mathbf{R}_I(t) + \mathbf{v}_I(t)\Delta t + \frac{1}{2}\mathbf{a}_I(t)\Delta t^2 \quad (2.35)$$

überführt werden. Gleichung 2.35 entspricht dem Velocity-Verlet-Algorithmus,<sup>[59]</sup> indem die Position  $\mathbf{R}_I$ , die Geschwindigkeit  $\mathbf{v}_I(t)$  und die Beschleunigung  $\mathbf{a}_I(t)$  des Teilchens zum Zeitpunkt  $t$  benötigt werden, um die neue Position zu erhalten. Die Geschwindigkeit wird im Velocity-Verlet-Algorithmus über

$$\mathbf{v}_I(t + \Delta t) = \mathbf{v}_I(t) + \frac{1}{2}[\mathbf{a}_I(t) + \mathbf{a}_I(t + \Delta t)]\Delta t \quad (2.36)$$

berechnet.<sup>[59]</sup>

Eine weitere häufig verwendete Methode ist der Leap-Frog-Algorithmus,<sup>[60]</sup> bei dem die Geschwindigkeiten allerdings zu  $\frac{1}{2}\Delta t$  versetzten Zeitschritten berechnet werden.

MD-Simulationen werden oftmals im sogenannten mikrokanonischen Ensemble<sup>[7]</sup> durchgeführt. Dies bedeutet, dass die Teilchenanzahl  $N$ , das Volumen der Simulationsbox  $V$  und die Gesamtenergie des Systems  $E$  (NVE-Ensemble) konstant bleiben. In diesem Ensemble ist das System isoliert und es findet keine Wechselwirkung mit der Umgebung statt.

Für realistische Simulationen ist es aber oftmals wünschenswert, dass andere thermodynamische Größen, wie zum Beispiel die Temperatur, konstant gehalten werden (z.B. in Simulationen von Enzymen bei Körpertemperatur).<sup>[7]</sup> Wenn neben der Teilchenanzahl  $N$  und dem Volumen  $V$  die Temperatur  $T$  konstant gehalten werden (NVT), spricht man vom kanonischen Ensemble.<sup>[7]</sup> Da die Temperatur des Systems über

$$T(t) = \frac{1}{3Mk_B} \sum_I^N \frac{|\mathbf{p}_I(t)|^2}{m_I} \quad (2.37)$$

mit der Anzahl der Atome  $N$ , der Boltzmannkonstante  $k_B$ , den Impulsen  $p_I$  und der Masse  $m_I$  mit der kinetischen Energie und damit auch mit den Geschwindigkeiten der Atome zusammenhängt, kann die Temperatur über eine Reskalierung der Geschwindigkeiten konstant gehalten werden.<sup>[7]</sup> Diese Reskalierung muss allerdings langsam durchgeführt werden, da die Trajektorie des Systems andernfalls durch die Faktoren der Skalierung dominiert wird und keine zuverlässigen Ergebnisse erhalten werden. Neben anderen<sup>[7]</sup> ist das Berendsen-Thermostat<sup>[61]</sup> ein sehr häufig verwendetes Schema zur Temperaturkontrolle. Hierbei wird die Beschleunigung eines Atoms  $\mathbf{a}_I$  über

$$\mathbf{a}_I(t) = \frac{\mathbf{F}_I(t)}{m_I} + \frac{\mathbf{p}_I(t)}{m_I\tau} \left( \frac{T_0}{T(t)} - 1 \right) \quad (2.38)$$

reskaliert.<sup>[61]</sup>  $T_0$  ist die gewünschte Temperatur,  $T(t)$  die zum Zeitpunkt  $t$  und  $\tau$  ist die Kopplungskonstante, welche die Stärke der Kopplung bestimmt.

Wird nun anstatt des Volumens der Druck des Systems konstant gehalten (NPT-Ensemble), wird zusätzlich zum Thermostat auch ein Barostat benötigt.<sup>[7]</sup> Das Berendsen-Barostat<sup>[61]</sup> funktioniert auf eine ähnliche Weise wie das Thermostat. Der Druck des Systems ist durch

$$P = \frac{2}{3V} \left( E_{kin} + \frac{1}{2} \sum_{I>J} \mathbf{R}_{IJ} \cdot \mathbf{F}_{IJ} \right) \quad (2.39)$$

mit der kinetischen Energie  $E_{kin}$ , dem Abstand  $\mathbf{R}_{IJ}$  und der Kraft  $\mathbf{F}_{IJ}$  zwischen den zwei Atomen  $I$  und  $J$  gegeben.<sup>[61]</sup> Der Druck kann gemäß Gleichung 2.39 über eine Veränderung der Atomabstände und des Volumens angepasst werden. Dazu werden

die Koordinaten der Atome mit einem konstanten Faktor  $\alpha$  über

$$\frac{\partial \mathbf{R}}{\partial t} = \mathbf{v} + \alpha \mathbf{R} \quad (2.40)$$

angepasst.<sup>[61]</sup> Die Volumenänderung wird entsprechend über

$$\frac{\partial V}{\partial t} = 3\alpha V \quad (2.41)$$

berechnet. Zudem ist die Druckänderung eines Systems mit der isothermen Kompressibilität  $\beta$  über

$$\frac{\partial P}{\partial t} = -\frac{1}{\beta V} \frac{\partial V}{\partial t} = -\frac{3\alpha}{\beta} \quad (2.42)$$

verbunden.<sup>[61]</sup> Der Faktor  $\alpha$  lässt sich über

$$\alpha = -\frac{\beta (P_0 - P(t))}{3\tau} \quad (2.43)$$

mit dem Zieldruck  $P_0$ , dem Druck  $P(t)$  zum Zeitpunkt  $t$  und der Kopplungskonstante  $\tau$  berechnen. Dadurch ergibt sich für die modifizierte Bewegungsgleichung:<sup>[61]</sup>

$$\frac{\partial \mathbf{R}}{\partial t} = \mathbf{v} - \frac{\beta (P_0 - P(t))}{3\tau} \mathbf{R}. \quad (2.44)$$

Andere bekannte Barostate sind das Nosé-Hoover-Barostat<sup>[62,63]</sup> und das Parrinello-Rahman-Barostat.<sup>[64]</sup>

Die verschiedenen Ensembles werden nicht nur für unterschiedliche Simulationsbedingungen beim Produktionslauf verwendet, sie spielen außerdem eine wichtige Rolle bei der Equilibrierung des Systems, in der das System vor der eigentlichen Simulation in einen Gleichgewichtszustand gebracht wird.<sup>[7]</sup>

## 2.4 Wahl der QM- und MM-Methoden

Für die Genauigkeit und Funktionalität von QM/MM-Simulation sind neben der Art der QM/MM-Implementierung die Wahl des QM-Bereichs und die Wahl der verwendeten der QM- und MM-Methoden von essentieller Bedeutung.<sup>[4,5]</sup>

### 2.4.1 Wahl der MM-Methoden

Es gibt eine Vielzahl gut bekannter Kraftfelder zur Beschreibung biochemischer Systeme. Standard-Kraftfelder, wie in Abschnitt 2.4.2 beschrieben, werden am häufigsten in QM/MM-Implementierungen verwendet, da sie aufgrund der Beschreibung der elektrostatischen Wechselwirkungen mit einfachen Punktladungen relativ einfach verwendet werden können. Typische Kraftfelder dieser Art sind OLPS,<sup>[65]</sup> GROMOS,<sup>[66]</sup> CHARMM<sup>[67]</sup> und AMBER.<sup>[68]</sup> Dabei spielt es bei der Implementierung zunächst keine Rolle, ob die *all-atom* oder die *united-atom* Varianten der Kraftfelder verwendet werden. Alle diese Kraftfelder sind gut geeignet, um biochemische Systeme zu beschreiben und ihre Vor- und Nachteile sind gut untersucht.<sup>[69]</sup> Unter bestimmten Bedingungen, wie z.B. bei der Beschreibung von *Charge-Transfer*-Prozessen, liefern sie allerdings keine korrekten Ergebnisse.<sup>[5]</sup> In diesen Fällen müssen polarisierbare Kraftfelder verwendet werden. Die QM/MM-Implementierung mit polarisierbaren Kraftfeldern (*polarizable embedding*), in dem sich QM- und MM-System gegenseitig selbstkonsistent polarisieren, ist allerdings deutlich aufwendiger.<sup>[55]</sup> Polarisierbare Erweiterungen existieren für die Kraftfelder OLPS,<sup>[70,71]</sup> CHARMM<sup>[72]</sup> und AMBER.<sup>[73]</sup> Zusätzlich gibt es eigenständige, polarisierbare Kraftfelder wie AMOEBA.<sup>[74]</sup> Allerdings werden polarisierbare Kraftfelder aufgrund des erhöhten Rechenaufwands häufig nur in statischen QM/MM-Implementierungen verwendet.<sup>[5,75,76]</sup>

### 2.4.2 Kraftfelder

In klassischen MD-Simulationen werden die zur Lösung der klassischen Bewegungsgleichung benötigten Kräfte mittels analytischer Ableitungen von parametrisierten Kraftfeldern berechnet, um dadurch den hohen Rechenaufwand zur Lösung der SG in AIMD-Simulationen zu umgehen. Ein Kraftfeld besteht aus einem Parameterset zur Beschreibung der potentiellen Energie des Systems. Es gibt viele verschiedene empirische Kraftfelder, welche sich in der Art der Parametrisierung und des Anwendungsbereichs unterscheiden.<sup>[7,65,66,69,77]</sup> Die Kraftfelder werden so optimiert, dass sie experimentelle Eigenschaften bestimmter Systeme, wie beispielsweise die Struktur oder Dichte, reproduzieren können. Dabei sind einige Kraftfelder nur für bestimmte Systeme, wie z.B. biologische Makromoleküle, anwendbar.<sup>[7]</sup>

Das Parameterset eines Kraftfelds hat die allgemeine Form

$$V = V_b + V_{nb} \quad (2.45)$$

und besteht aus bindenden ( $V_b$ ) und nicht-bindenden ( $V_{nb}$ ) Beiträgen. Die bindenden

Beiträge setzen sich wie folgt zusammen:<sup>[7]</sup>

$$V_b = V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedral}}. \quad (2.46)$$

Dabei wird das Potential einer Bindung mit dem Bindungsabstand  $R_b$  durch

$$V_{\text{bond}} = \frac{1}{2}k_b(R_b - R_0)^2 \quad (2.47)$$

mit der Kraftkonstante  $k_b$  und dem Gleichgewichtsabstand  $R_0$  beschrieben.<sup>[7]</sup> Das Biegen eines Winkels  $\theta$  zwischen drei Atomen wird über das harmonische Potential

$$V_{\text{angle}} = \frac{1}{2}k_\theta(\theta - \theta_0)^2 \quad (2.48)$$

mit der Kraftkonstante  $k_\theta$  und dem Gleichgewichtswinkel  $\theta_0$  beschrieben.<sup>[7]</sup> Sind vier Atome miteinander verbunden, wird die Winkeländerung des Dihedralwinkels über das periodische Potential

$$V_{\text{dihedral}} = \frac{1}{2} \sum_n^3 V_n [1 + \cos(n\omega)] \quad (2.49)$$

beschrieben.<sup>[77]</sup> Die  $V_n$  sind die Kraftkonstanten der Fourierreihe und  $\omega$  ist der Torsionswinkel.

Wechselwirkungen, welche über vier Atome hinausgehen, werden als nicht-bindende Wechselwirkungen betrachtet. Das Potential  $V_{\text{nb}}$  besteht aus der elektrostatischen Coulombwechselwirkung ( $V_C$ ) zwischen zwei Punktladungen und aus den van der Waals-Wechselwirkungen ( $V_{vdW}$ ) zwischen zwei Atomen unabhängig von ihrer Ladung:<sup>[7]</sup>

$$V_{\text{nb}} = V_C + V_{vdW}. \quad (2.50)$$

Das Coulombpotential zwischen zwei Punktladungen mit dem Abstand  $R_{AB}$  und den jeweiligen Partialladungen der Atome  $q_A$  und  $q_B$  beträgt<sup>[7]</sup>

$$V_C = \frac{q_A q_B}{4\pi\epsilon_0 R_{AB}} \quad (2.51)$$

mit der Dielektrizitätskonstante  $\epsilon_0$  des Vakuums. Die van der Waals-Wechselwirkungen werden über das Lennard-Jones-Potential  $V_{LJ}$  mit der Form

$$V_{LJ}(R) = 4\epsilon_{AB} \left[ \left( \frac{\sigma_{AB}}{R_{AB}} \right)^{12} - \left( \frac{\sigma_{AB}}{R_{AB}} \right)^6 \right] \quad (2.52)$$

mit den atomspezifischen Kraftfeldkonstanten  $\epsilon_{AB}$  und  $\sigma_{AB}$  beschrieben.<sup>[7]</sup> Der erste Term in Gleichung 2.52 beschreibt die Pauli-Abstoßung, der zweite Term die London-Dispersionswechselwirkungen. Je nach Kraftfeld können weitere Parameter

wie Kreuz-Terme, die Bindungslänge und -winkel miteinander verknüpfen oder zusätzliche nicht-bindende Terme hinzukommen.<sup>[7,77]</sup>

Die Anzahl der zu berechnenden nicht-bindenden Wechselwirkungen, welche quadratisch mit der Anzahl an Atomen skaliert ( $\mathcal{O}(N^2)$ ), ist deutlich größer als die Anzahl bindender Wechselwirkungen, welche linear mit der Anzahl an Atomen skaliert ( $\mathcal{O}(N)$ ). Daher ist die Berechnung der nicht-bindenden Wechselwirkungen der zeitbestimmende Schritt in MD-Simulationen. Um den Rechenaufwand zu verringern, gibt es verschiedene Methoden, wie z.B. die Particle-Mesh-Ewald-Methode<sup>[78]</sup> (PME).

### 2.4.3 Wahl der QM-Methoden

Die Wahl der QM-Methode hängt stark von der Wahl des (QM-)Systems und der Problemstellung ab. Theoretisch ist es möglich jede QM-Methode in einer QM/MM-Implementierung zu verwenden.<sup>[4,5]</sup> Entscheidend bei der Wahl der Methode ist immer eine Abwägung zwischen der benötigten Genauigkeit und des Rechenaufwands der entsprechenden Methode sowie der Größe des QM-Systems. Mit semiempirischen Methoden können große QM-Systeme mit mehreren hundert bis tausenden Atomen simuliert werden. Allerdings verfügen sie oftmals nicht über eine ausreichende Genauigkeit.<sup>[5]</sup>

Die in QM/MM-Implementierungen am häufigsten verwendete QM-Methode ist die Dichtefunktionaltheorie (DFT),<sup>[25,26]</sup> da sie ein gutes Verhältnis zwischen Rechenaufwand und Genauigkeit besitzt.<sup>[4,5,7,79]</sup> Mit DFT lassen sich QM-Systeme mit mehreren hundert Atomen und akzeptabler Genauigkeit simulieren. Der Nachteil von DFT-Implementierungen ist die Abhängigkeit von der Wahl des verwendeten Funktionals (siehe Abschnitt 2.4.7). Reicht die Genauigkeit von DFT nicht aus oder sollen Eigenschaften untersucht werden, die mit DFT nicht korrekt beschrieben werden können,<sup>[80]</sup> müssen auf Wellenfunktionen basierende *ab initio* Methoden verwendet werden.<sup>[47,81]</sup> Diese sind post-Hartree-Fock Elektronenkorrelationsmethoden wie Møller-Plesset Störungstheorie<sup>[82]</sup> und Coupled-Cluster<sup>[24]</sup> oder Multireferenz Methoden.<sup>[83,84]</sup> Diese Methoden besitzen zum Teil eine sehr hohe Genauigkeit, allerdings können aufgrund des sehr hohen Rechenaufwands<sup>[7,8]</sup> auch nur relativ kleine QM-Systeme simuliert werden.

Die Wahl des QM-Bereichs basiert immer auf vorheriger Betrachtung des zu untersuchenden Systems.<sup>[5]</sup> Einerseits muss darauf geachtet werden, dass für die elektronische Struktur wichtige Segmente des Moleküls im QM-Bereich mit inbegriffen sind, andererseits versucht man den QM-Bereich aufgrund der Rechenkosten möglichst klein zu halten. Grundsätzlich gilt, je größer der QM-Bereich ist, umso kleiner wird der Fehler im Vergleich zu einer QM-Simulation des gesamten Systems.<sup>[5]</sup> Bezüglich der Wahl der zu verwendenden QM-Methode muss abgewägt werden wie eine möglichst hohe Genauigkeit erreicht werden kann und ob es sinnvoller ist den QM-Bereich zu vergrößern oder eine genauere aber teurere QM-Methode zu verwenden.

Die in dieser Arbeit verwendeten quantenchemischen Methoden werden im Folgenden diskutiert.

### 2.4.4 Hartree-Fock-Theorie (HF)

Da eine exakte Lösung der elektronischen Schrödingergleichung in der Regel nicht möglich ist, muss sie näherungsweise gelöst werden. Die Hartree-Fock-Theorie (HF) ist eine einfache und physikalisch sinnvolle Näherung, welche zudem als Grundlage für eine Vielzahl genauere Methoden zur Beschreibung der Elektronenstruktur dient.<sup>[8]</sup> Ein zentraler Bestandteil ist die Wahl einer Slaterdeterminanten zur Darstellung der Wellenfunktion. Dadurch wird die Antisymmetrie der Wellenfunktion (Pauli-Prinzip) implizit berücksichtigt. Die Wellenfunktion für ein System mit  $N$  Elektronen lässt sich über

$$\Psi(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_i(1) & \varphi_j(1) & \dots & \varphi_N(1) \\ \varphi_i(2) & \varphi_j(2) & \dots & \varphi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_i(N) & \varphi_j(N) & \dots & \varphi_N(N) \end{vmatrix}. \quad (2.53)$$

mit  $N$  Spinorbitalen  $(\varphi_i, \varphi_j, \dots, \varphi_N)$ , welche mit  $N$  Elektronen  $(1, 2, \dots, N)$  besetzt sind, darstellen.<sup>[8]</sup> Ein Spinorbital setzt sich aus einem Ortsorbital  $\phi(\mathbf{r})$  und einer Spinfunktion  $\chi(\sigma)$  über

$$\varphi_i(\mathbf{r}, \sigma) = \phi(\mathbf{r})\chi(\sigma) \quad (2.54)$$

zusammen. Mit dem Hamiltonoperator aus Gleichung 2.4 ergibt sich der HF-Energieerwartungswert<sup>[8]</sup> über

$$E_{HF} = \langle \Psi | \hat{H} | \Psi \rangle = \sum_i^N h_{ii} + \frac{1}{2} \sum_{i,j}^N (\langle ij | ij \rangle - \langle ij | ji \rangle) \quad (2.55)$$

mit den Eielektronenintegralen

$$h_{ii} = \langle i | \hat{h} | i \rangle \quad (2.56)$$

und dem Eielektronen-Hamiltonoperator

$$\hat{h} = -\frac{1}{2}\Delta - \sum_I^M \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|}. \quad (2.57)$$

Die Zweielektronenintegrale werden als Coulombintegral

$$\langle ij | ij \rangle = J_{ij} = \langle ij | \frac{1}{r_{12}} | ij \rangle, \quad (2.58)$$

welches die klassische Coulombwechselwirkung zwischen den Elektronen beschreibt, bzw. als Austauschintegral

$$\langle ij|ji\rangle = K_{ij} = \langle ij|\frac{1}{r_{12}}|ji\rangle, \quad (2.59)$$

welches einen rein quantenmechanischen Beitrag aufgrund des Pauli-Prinzips beschreibt, bezeichnet. Im geschlossenschaligen Fall, d.h. wenn es keine ungepaarten Elektronen sondern nur doppelt besetzte Orbitale gibt, erhält man nach Spinintegration die Restricted Hartree-Fock-Energie (RHF):<sup>[8]</sup>

$$E_{RHF} = \langle \Psi|\hat{H}|\Psi\rangle = \sum_i^{N/2} 2h_{ii} + \frac{1}{2} \sum_{i,j}^{N/2} (4J_{ij} - 2K_{ij}). \quad (2.60)$$

Hierbei wird über alle  $N/2$  besetzte Ortsorbitale summiert. Der Erwartungswert der Energie des Systems wird durch die Slaterdeterminante, also durch die verwendeten Orbitale bestimmt. Um die bestmöglichen Orbitale zu erhalten, wird der Erwartungswert unter Variation der Orbitale minimiert. Dabei muss berücksichtigt werden, dass die Orthogonalität der Orbitale

$$\langle \phi_i|\phi_j\rangle = \delta_{ij} \quad (2.61)$$

weiterhin gewährleistet ist. Durch die Minimierung werden die kanonischen Hartree-Fock-Gleichungen<sup>[8]</sup>

$$\hat{F}\phi_i = \epsilon_i\phi_i \quad (2.62)$$

erhalten. Hierbei ist  $\epsilon_i$  die Orbitalenergie des optimierten Molekülorbitals  $\phi_i$ . Im RHF-Fall hat der Fockoperator  $\hat{F}$  folgende Form:

$$\hat{F} = \hat{h} + \sum_j^{N/2} (2\hat{J}_j - \hat{K}_j). \quad (2.63)$$

Hierbei sind  $\hat{J}_j$  bzw.  $\hat{K}_j$  der Coulomb- bzw. der Austauschoperator und sind durch ihre Wirkung auf Elektron (1) in Orbital  $\phi_i$  über

$$\hat{J}_j\phi_i(1) = \int d^3r_2 \phi_j^*(2) \frac{1}{r_{12}} \phi_j(2) \phi_i(1) \quad (2.64)$$

bzw.

$$\hat{K}_j\phi_i(1) = \int d^3r_2 \phi_j^*(2) \frac{1}{r_{12}} \phi_j(1) \phi_i(2) \quad (2.65)$$

definiert. Da der Coulomb-Operator die Wechselwirkung zwischen den Ladungsverteilungen von zwei Elektronen beschreibt, aber nicht die explizite Wechselwirkung, wird der HF-Ansatz auch als *mean-field* Methode bezeichnet.<sup>[8]</sup> Der Fockoperator selbst hängt über Coulomb- und Austausch-Operator von den Molekülorbitalen  $\phi_j$  ab, weshalb die HF-Gleichungen nur über ein iteratives *self-consistent-field*-Verfahren<sup>[85]</sup> (SCF-Verfahren) gelöst werden können.

### Roothaan-Hall-Gleichungen

Entscheidend für die Güte der Wellenfunktion ist die Form der Molekülorbitale (MO), welche allerdings bisher unbekannt ist. Der Ansatz von Roothaan<sup>[86]</sup> und Hall<sup>[87]</sup> ist, die Molekülorbitale in einer Basis aus bekannten Atomorbitalen (AO) darzustellen. Dadurch werden die Hartree-Fock-Gleichungen in eine algebraische Form überführt, sodass sie durch Standardtechniken der Matrizenrechnung gelöst werden können. Die MOs werden über

$$\phi_i = \sum_{\mu}^M \chi_{\mu} c_{\mu i} \quad (2.66)$$

als Linearkombination aus  $M$  Basisfunktionen  $\chi_{\mu}$ , in der Regel wasserstoffähnliche AOs aus kontrahierten Gaußfunktionen (GTOs), erhalten (LCAO-Ansatz).<sup>[8]</sup> Hierbei sind  $c_{\mu i}$  die Entwicklungskoeffizienten. Die Genauigkeit der Methode hängt von der Art und der Größe (mehr Variationsmöglichkeiten) der verwendeten Basis ab.<sup>[88]</sup> Wird Gleichung 2.66 in die HF-Gleichungen eingesetzt und von links mit  $\langle \chi_{\nu} |$  multipliziert, erhält man die Roothaan-Hall-Gleichungen in Matrixdarstellung:<sup>[8]</sup>

$$\mathbf{FC} = \mathbf{SC}\epsilon. \quad (2.67)$$

Hierbei ist  $\mathbf{F}$  die Fockmatrix mit den Matrixelementen

$$F_{\nu\mu} = \langle \chi_{\nu} | \hat{F} | \chi_{\mu} \rangle, \quad (2.68)$$

$\mathbf{S}$  die Überlappungsmatrix mit den Matrixelementen

$$S_{\nu\mu} = \langle \chi_{\nu} | \chi_{\mu} \rangle, \quad (2.69)$$

$\mathbf{C}$  die Koeffizientenmatrix und  $\epsilon$  eine Diagonalmatrix mit den Orbitalenergien als Einträge. Der Fockoperator ist nun nicht mehr explizit von den Molekülorbitalen sondern von den Koeffizienten abhängig. Als Pseudoeigenwertproblem werden die Roothaan-Hall-Gleichungen ebenfalls über ein iteratives SCF-Verfahren gelöst.

Der HF-Ansatz ist eine qualitativ gute Näherung und aufgrund des verwendeten Variationsprinzips bildet die HF-Energie eine obere Schranke für die exakte Energie.<sup>[8]</sup> Allerdings ist die Genauigkeit aufgrund der fehlenden, expliziten Coulombwechselwirkung der Elektronen, der sogenannten Coulomb-Korrelation, quantitativ nicht ausreichend. Die Korrelationsenergie ist als Differenz zwischen der HF-Energie und

der exakten, nicht-relativistischen Energie über

$$E_{\text{korr}} = E_{\text{exakt}}^{\text{nr}} - E_{\text{HF}}. \quad (2.70)$$

definiert.<sup>[89]</sup> Die Korrelationsenergie ist zwar im Vergleich zur Gesamtenergie eines Systems relativ klein, für die Beschreibung von molekularen Eigenschaften und Reaktionsenergien allerdings oftmals essentiell. Die Entwicklung geeigneter Methoden zur Berechnung chemischer Eigenschaften unter expliziter Berücksichtigung der Elektronenkorrelation ist daher seit jeher ein wichtiger Bereich der theoretischen Chemie.<sup>[7,8,90]</sup> Im Folgenden werden zwei unterschiedliche Gruppen von Methoden zur Beschreibung der Elektronenkorrelation vorgestellt. Das sind zum einen Methoden wie die Møller-Plesset-Störungstheorie, die Coupled-Cluster-Theorie und CAS-SCF, welche auf einer Wellenfunktion basieren. Zum anderen ist es die auf Elektronendichte basierende Dichtefunktionaltheorie (DFT).

### 2.4.5 Wellenfunktionsbasierte post-Hartree-Fock-Methoden

Da die HF-Wellenfunktion bereits eine gute Näherung für die exakte Lösung der Schrödingergleichung ist, wird sie häufig als Ausgangspunkt für komplexere, wellenfunktionsbasierende Methoden zur Beschreibung der Elektronenkorrelation verwendet. Diese Methoden haben gemeinsam, dass sie, ausgehend von der HF-Wellenfunktion, angeregte Determinanten berücksichtigen.<sup>[7]</sup>

#### Møller-Plesset-Störungstheorie

Da die Korrelationsenergie im Vergleich zur Gesamtenergie relativ klein ist, ist es eine gute Möglichkeit sie als Störung zu berechnen. Eine häufig verwendete Methode dieser Art ist die Møller-Plesset-Störungstheorie<sup>[29,82]</sup> (MP) welche eine spezielle Form der allgemeinen Rayleigh-Schrödinger-Störungstheorie<sup>[7]</sup> ist. Allgemein wird in der Störungstheorie der Hamiltonoperator in einen ungestörten Teil  $\hat{H}^{(0)}$  und einen gestörten Teil  $\hat{H}'$  über

$$\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}' \quad (2.71)$$

aufgeteilt.<sup>[7]</sup> Der Störparameter  $\lambda$  kann dabei zwischen 0 und 1 variieren. Hierbei ist es wichtig, dass der ungestörte Hamiltonoperator sowie seine Eigenwerte und Eigenfunktionen bekannt sind. Werden die Wellenfunktion des Grundzustandes  $\Psi$  und die Energie  $E$  in einer Taylorreihe nach  $\lambda$  entwickelt, erhält man

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \quad (2.72)$$

bzw.

$$\Psi = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \dots \quad (2.73)$$

mit den Eigenwerten  $E^{(0)}$  und der Eigenfunktion  $\Psi^{(0)}$  des ungestörten Hamiltonoperators  $\hat{H}^{(0)}$ . Werden die Entwicklungen aus Gleichung 2.72 und 2.73 in die SG

$$\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)|\Psi\rangle = E|\Psi\rangle \quad (2.74)$$

eingesetzt und die Terme nach der Potenz von  $\lambda$  sortiert, erhält man eine Reihe von Gleichungen für die verschiedenen Ordnungen der Energiekorrekturen, welche über Multiplikation mit  $\langle\Psi^{(0)}|$  und Ausnutzen der Orthogonalitätsbeziehung weiter vereinfacht werden können.<sup>[7]</sup> Für die nullte- bis zweite Ordnung erhält man

$$E^{(0)} = \langle\Psi^{(0)}|\hat{H}^{(0)}|\Psi^{(0)}\rangle, \quad (2.75)$$

$$E^{(1)} = \langle\Psi^{(0)}|\hat{H}'|\Psi^{(0)}\rangle \quad (2.76)$$

und

$$E^{(2)} = \sum_{j>0} \frac{|\langle\Psi_j^{(0)}|\hat{H}'|\Psi^{(0)}\rangle|^2}{E^{(0)} - E_j^{(0)}} \quad (2.77)$$

mit der Summe über alle Eigenfunktionen  $\Psi_j^{(0)}$  des ungestörten Hamiltonoperators  $\hat{H}^{(0)}$ . In dem Ansatz von Møller und Plesset<sup>[29,82]</sup> wird für den ungestörten Hamiltonoperator die Summe der Fockoperatoren (Gleichung 2.63)

$$\hat{H}^{(0)} = \sum_i^N \hat{F}_i \quad (2.78)$$

und für die Grundzustandswellenfunktion  $\Psi^{(0)}$  die HF-Wellenfunktion  $\Psi_{HF}$  gewählt. Damit ergibt sich für den gestörten Hamiltonoperator über Gleichung 2.71

$$\hat{H}' = \hat{H} - \hat{H}^{(0)} = \sum_{i<j}^N \frac{1}{r_{ij}} - \sum_i^N \sum_j^N \left(\hat{J}_j - \hat{K}_j\right). \quad (2.79)$$

Durch Einsetzen von  $\hat{H}^{(0)}$  und  $\hat{H}'$  in Gleichungen 2.75 bis 2.77 erhält man für die Energie nullter Ordnung

$$E^{(0)} = \langle\Psi^{(0)}|\sum_i^N \hat{F}_i|\Psi^{(0)}\rangle = \sum_i^N \epsilon_i, \quad (2.80)$$

was der Summe der HF-Orbitalenergien entspricht. Die Summe aus der Energie nullter und erster Ordnung entspricht über

$$\begin{aligned} E^{(0)} + E^{(1)} &= \langle \Psi^{(0)} | \hat{H}^{(0)} | \Psi^{(0)} \rangle + \langle \Psi^{(0)} | \hat{H}' | \Psi^{(0)} \rangle \\ &= \langle \Psi^{(0)} | \hat{H}^{(0)} + \hat{H}' | \Psi^{(0)} \rangle \\ &= E_{HF} \end{aligned} \quad (2.81)$$

der HF-Energie.<sup>[7]</sup> Damit liefert erst die Energie zweiter Ordnung den ersten Beitrag zur Korrelationsenergie und ist über

$$E^{(2)} = \frac{1}{4} \sum_{ij} \sum_{ab} \frac{|\langle ij|ab \rangle - \langle ia|jb \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \quad (2.82)$$

gegeben.<sup>[7]</sup> Die Indizes  $i$  und  $j$  bzw.  $a$  und  $b$  stehen dabei für besetzte bzw. unbesetzte Orbitale. In Gleichung 2.82 tragen nur zweifach angeregte Determinanten zur Korrelationsenergie bei, da aufgrund des *Brillouin-Theorems*<sup>[8]</sup> und der *Slater-Condon-Regeln*<sup>[91,92]</sup> alle Terme mit einfach bzw. dreifach und höher angeregten Determinanten verschwinden. Bei Energiekorrekturen höherer Ordnung liefern hingegen auch höher angeregte Determinanten einen Beitrag zur Korrelationsenergie.<sup>[7]</sup> Aufgrund der benötigten Transformation der Zweielektronenintegrale von AO- in MO-Basis für Gleichung 2.82 skaliert die MP2-Methode mit  $\mathcal{O}(N^5)$  und ist damit rechenaufwendiger als die HF-Methode ( $\mathcal{O}(N^4)$ ). Die MP2-Methode ist nicht variationell und es ist nicht gesichert, dass die Energie bei der Berücksichtigung höherer Ordnungen trotz steigender Rechenkosten konvergiert.<sup>[93,94]</sup> Daher werden heutzutage nur noch selten höhere MP-Korrekturen verwendet.

### Coupled-Cluster-Theorie

Eine weitere Methode zur Beschreibung der Elektronenkorrelation, welche sich durch sehr hohe Genauigkeit auszeichnet, ist die Coupled-Cluster-Theorie<sup>[24,95–97]</sup> (CC). Auch bei der CC-Theorie wird die Elektronenkorrelation über angeregte Determinanten, welche ausgehend von der HF-Grundzustandswellenfunktion  $\Psi_{HF}$  gebildet werden, beschrieben. Die CC-Wellenfunktion wird über einen exponentiellen Ansatz durch

$$|\Psi_{CC}\rangle = e^{\hat{T}} |\Psi_0\rangle \quad (2.83)$$

gebildet.<sup>[95]</sup> Der Clusteroperator  $\hat{T}$  hat die Form

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_N \quad (2.84)$$

und besteht aus einer Summe von Anregungsoperatoren  $\hat{T}_n$ , welche  $n$  Elektronen aus besetzten in unbesetzte Orbitale anregen. Wird der exponentielle Ansatz in einer

Taylorreihe entwickelt, erhält man

$$e^{\hat{T}} = \left( 1 + \hat{T} + \frac{1}{2!}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + \dots \right). \quad (2.85)$$

Die Anregungsoperatoren sind im Rahmen des Formalismus der *zweiten Quantisierung*<sup>[8]</sup> über

$$\hat{T}_n = \left( \frac{1}{n!} \right)^2 \sum_{ij\dots} \sum_{ab\dots} t_{ij\dots}^{ab\dots} \hat{a}_a^\dagger \hat{a}_b^\dagger \dots \hat{a}_j \hat{a}_i, \quad (2.86)$$

mit den Amplituden  $t_{ij\dots}^{ab\dots}$ , den Vernichtungsoperatoren  $\hat{a}_i$  und den Erzeugungsoperatoren  $\hat{a}_a^\dagger$  gegeben.<sup>[95]</sup> Auf die Grundzustandswellenfunktion  $\Psi_0$  angewendet ergibt sich beispielsweise mit  $\hat{T}_2$  über

$$\hat{T}_2 \Psi_0 = \frac{1}{4} \sum_{ij} \sum_{ab} t_{ij}^{ab} \Phi_{ij}^{ab} \quad (2.87)$$

eine zweifach angeregte Determinante  $\Phi_{ij}^{ab}$ . Für die SG ergibt sich mit der CC-Wellenfunktion

$$\hat{H} e^{\hat{T}} |\Psi_0\rangle = E_{CC} e^{\hat{T}} |\Psi_0\rangle. \quad (2.88)$$

Eine variationelle Lösung von Gleichung 2.88 ist nicht praktikabel,<sup>[95]</sup> da die Potenzreihenentwicklung des Operators nicht abbricht. Daher wird zur Bestimmung der CC-Energie und der Amplituden von links mit  $e^{-\hat{T}}$  multipliziert und auf die Grundzustandsdeterminante  $\Psi_0$  bzw. auf die angeregten Determinanten  $\Psi_P$  projiziert. Dadurch erhält man die CC-Gleichungen<sup>[95]</sup>

$$E_{CC} = \langle \Psi_0 | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle \quad (2.89)$$

bzw. die Amplitudengleichungen

$$0 = \langle \Psi_P | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle. \quad (2.90)$$

Der Operator  $e^{-\hat{T}} \hat{H} e^{\hat{T}}$  aus den Gleichungen 2.89 und 2.90 kann über die *Baker-Campbell-Hausdorff*-Entwicklung<sup>[95]</sup> ausgewertet werden, wodurch die Energie und die Amplituden bestimmt werden. Die Auswertung der einzelnen Terme ist relativ aufwendig und kann in Referenz 95 nachgelesen werden.

Die Summe im Clusteroperator wird in der Regel ab einer bestimmten Anregung abgebrochen, da eine Berücksichtigung aller Anregungen zu kostenintensiv wäre. Werden beispielsweise nur Einfach- und Zweifachanregungen berücksichtigt ( $\hat{T} = \hat{T}_1 + \hat{T}_2$ ) spricht man von der *coupled-cluster singles and doubles*-Methode<sup>[98]</sup> (CCSD). Durch den exponentiellen Ansatz der Wellenfunktion werden implizit auch höhere Anregungen berücksichtigt, weshalb abgebrochene CC-Methoden größenkonsistent sind.<sup>[95]</sup>

Der Rechenaufwand der CC-Methoden steigt deutlich, wenn höhere Anregungen im Clusteroperator berücksichtigt werden. Während CCSD mit  $\mathcal{O}(N^6)$  skaliert, steigt bei der Berücksichtigung von Dreifachanregungen<sup>[99]</sup> (CCSDT) der Rechenaufwand auf  $\mathcal{O}(N^8)$ . Die CCSD(T)-Methode,<sup>[100]</sup> welche die Dreifachanregungen über eine störungstheoretische Behandlung im Anschluss einer CCSD-Rechnung berücksichtigt, skaliert mit  $\mathcal{O}(N^7)$ . Die Genauigkeit der CCSD(T)-Methode ist allerdings meist so hoch, dass man vom *Goldstandard* der Quantenchemie spricht.

### 2.4.6 Complete Active Space Self Consistent Field (CAS-SCF)

Die bisher vorgestellten Methoden zur Beschreibung der Elektronenkorrelation, welche auf der HF-Determinante beruhen, berücksichtigen die sogenannte *dynamische* Elektronenkorrelation,<sup>[7]</sup> basierend auf der Elektron-Elektron Wechselwirkung. Es gibt allerdings chemische Systeme, bei denen für die Beschreibung des Grundzustandes eine HF-Determinante schon qualitativ nicht ausreicht. In diesen Fällen wird die Wellenfunktion wiederum als Linearkombination mehrerer Determinanten über

$$\Psi = c_0\Psi_0 + c_1\Psi_1 + c_2\Psi_2 + \dots \quad (2.91)$$

dargestellt. In diesem Fall spricht man von *statischer* Elektronenkorrelation. Systeme, bei denen der Grundzustand von mehreren Determinanten beschrieben wird, nennt man *Multireferenz*-Systeme.<sup>[7]</sup> Wird die Wellenfunktion aus Gleichung 2.91 optimiert, müssen zum einen die Koeffizienten  $c_0, c_1, \dots$  und zum anderen die MOs (bzw. die Koeffizienten des LCAO Ansatzes), aus denen die Determinanten aufgebaut sind, optimiert werden.<sup>[7]</sup>

Dadurch werden Multi-Referenz-Rechnungen deutlich aufwändiger als vergleichbare Single-Referenz-Rechnungen. Dabei ist die Auswahl an Determinanten, welche in der Rechnung berücksichtigt werden, essentiell. Da es zu aufwändig (und auch nicht nötig) wäre alle möglichen Determinanten in der Rechnung zu berücksichtigen, muss eine Vorauswahl getroffen werden. Bei einer sogenannten *Complete Active Space*<sup>[101]</sup> SCF Rechnung (CAS-SCF) wird ein aktiver Raum vorgegeben, welcher aus einer bestimmten Anzahl an Elektronen und Orbitalen besteht. Innerhalb dieses aktiven Raums werden alle möglichen Kombinationen an besetzten und unbesetzten Orbitalen erzeugt, welche zusammen mit den Orbitalen des inaktiven Raums die jeweiligen Determinanten bilden.<sup>[7]</sup> Hierbei ist es wichtig den aktiven Raum (und damit auch die Anzahl der Determinanten) möglichst klein zu halten und dennoch alle relevanten Determinanten zu berücksichtigen. Dafür gibt es allerdings keine ideale und auf alle Systeme anwendbare Vorgehensweise. Daher ist ein gewisses chemisches Verständnis vorteilhaft, um die relevanten Orbitale zu identifizieren, oftmals müssen jedoch mehrere aktive Räume getestet werden.<sup>[7]</sup>

Ausgehend von CAS-SCF Determinanten wurden auch post-HF-Methoden wie *Complete Active Space Second-Order Perturbation Theory* (CASPT2) und Multi-Referenz-Coupled-Cluster zur kompletten Beschreibung von statischer und dynami-

scher Elektronenkorrelation formuliert.<sup>[102]</sup>

### 2.4.7 Dichtefunktionaltheorie (DFT)

Eine andere Methode, welche nicht auf einer Wellenfunktion basiert, ist die Dichtefunktionaltheorie.<sup>[25,26,103,104]</sup> Sie ist aufgrund des guten Kosten-Genauigkeit-Verhältnis häufig die Methode der Wahl, besonders bei größeren Systemen, die mit hochgenauen CC-Methoden nicht zu berechnen sind.<sup>[4,5,52]</sup> Im Folgenden werden die Grundlagen der DFT kurz erläutert, welche für AIMD- und QM/MM-MD-Simulationen relevant sind. Tieferegehende Ausführungen der DFT können in der Literatur (Ref. 7,26,105,106) und im Speziellen die Theorie und Implementierung im Programmcode CPMD im Buch von Marx und Hutter<sup>[52]</sup> nachgelesen werden.

Im Rahmen der Kohn-Sham-Dichtefunktionaltheorie<sup>[104]</sup> wird die Energie eines Systems über

$$E^{\text{KS}}[\{\phi_i\}] = T_s[\{\phi_i\}] + \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int V_{\text{H}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_{\text{xc}}[\rho] \quad (2.92)$$

erhalten.<sup>[52]</sup> Dieses Kohn-Sham-Funktional hängt von den Kohn-Sham-Orbitalen  $\phi_i(\mathbf{r})$  ab, einem Satz von Hilfsfunktionen, welche die Orthogonalitätsbeziehung  $\langle \phi_i | \phi_j \rangle = \delta_{ij}$  erfüllen. Die Elektronendichte  $\rho(\mathbf{r})$  ist über

$$\rho(\mathbf{r}) = \sum_i^{\text{occ}} f_i |\phi_i(\mathbf{r})|^2 \quad (2.93)$$

gegeben, wobei  $f_i$  ganzzahlige Besetzungszahlen sind.<sup>[52]</sup> Gleichung 2.92 weist einige Analogien zum Hamiltonoperator (Gleichung 2.4) auf. Der erste Term entspricht der kinetischen Energie eines nicht wechselwirkenden Referenzsystems,<sup>[52]</sup> beschrieben über die Kohn-Sham-Orbitale und ist gegeben durch

$$T_s[\{\phi_i\}] = \sum_i^{\text{occ}} f_i \left\langle \phi_i \left| -\frac{1}{2}\Delta \right| \phi_i \right\rangle. \quad (2.94)$$

Der zweite Term entspricht der Coulombwechselwirkung der Elektronen mit den Kernen und der Kern-Kern Wechselwirkung. Das externe Potential  $V_{\text{ext}}(\mathbf{r})$  hat die folgende Form:<sup>[52]</sup>

$$V_{\text{ext}}(\mathbf{r}) = - \sum_I \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|} + \sum_{I < J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (2.95)$$

Der dritte Term entspricht der Coulombwechselwirkung zwischen zwei Elektronendichten und wird über das Hartree-Potential

$$V_{\text{H}} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (2.96)$$

erhalten.<sup>[52]</sup> Der vierte Term in Gleichung 2.92 ist das sogenannte Austausch-Korrelations-Funktional<sup>[52]</sup>  $E_{xc}[\rho]$ . Es beinhaltet die Austausch- und Korrelationseffekte und ist definiert über die Differenz zwischen exakter Energie und den ersten drei Termen des Kohn-Sham-Funktional. Die exakte Form von  $E_{ex}[\rho]$  ist allerdings nicht bekannt. Daher ist die Suche nach einer möglichst guten Näherung für das Funktional der zentrale Punkt der Dichtefunktionaltheorie.<sup>[25,26,52]</sup> Mittlerweile gibt es mehrere Hundert verschiedene Funktionale unterschiedlicher Komplexität und Genauigkeit. Ein Überblick über die verschiedenen Arten der Funktionale und deren Genauigkeit, sowie Benchmark-Rechnungen haben Mardirossian und Head-Gordon zusammengestellt.<sup>[27]</sup>

Die minimale Energie wird erhalten, indem das Kohn-Sham-Funktional bezüglich der Elektronendichte  $\rho(\mathbf{r})$  bzw. bezüglich der Kohn-Sham-Orbitale  $\phi_i(\mathbf{r})$  minimiert wird. Dadurch werden, analog zu den Hartree-Fock-Gleichungen, die kanonischen Kohn-Sham-Gleichungen

$$H_e^{\text{KS}} \phi_i = \epsilon_i \phi_i \quad (2.97)$$

mit dem Kohn-Sham-Hamiltonoperator

$$H_e^{\text{KS}} = -\frac{1}{2}\Delta + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + \frac{\partial E_{xc}[\rho]}{\partial \rho(\mathbf{r})} \quad (2.98)$$

erhalten.<sup>[52]</sup>

### Plane-Wave Dichtefunktionaltheorie (pw-DFT)

Die Kohn-Sham-Orbitale können ebenfalls über den LCAO-Ansatz aus einer festen Basis aus Atomorbitalen gebildet werden. Ein anderer Ansatz ist es, eine Basis aus ebenen Wellen<sup>[52,106]</sup> (*plane-waves*) zu verwenden (pw-DFT). Die ebenen Wellen sind über

$$\chi_{\mathbf{G}}^{\text{PW}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{G}\mathbf{r}} \quad (2.99)$$

mit  $\Omega$ , dem Volumen der Einheitszelle des QM-Systems und dem reziproken Raumvektor  $\mathbf{G}$  definiert.<sup>[52]</sup> Die Orbitale werden als Linearkombination aus einer im Prinzip unendlichen Anzahl an Wellen gebildet:<sup>[52]</sup>

$$\phi_i^{\text{KS}}(\mathbf{r}, \mathbf{k}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{j\mathbf{G}}(\mathbf{k}) e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}}. \quad (2.100)$$

Hierbei sind  $c_{j\mathbf{G}}$  die komplexen Koeffizienten und  $\mathbf{k}$  ist der Wellenvektor. In der Praxis muss die Summe über  $\mathbf{G}$  abgebrochen werden. Dafür wird ein Energie-Cutoff  $E_{\text{cut}}$  eingeführt und es werden ausschließlich solche ebenen Wellen in der Summe berücksichtigt, deren Energie niedriger als der Energie-Cutoff ist. Dadurch wird die Größe der Basis und damit die Genauigkeit der Rechnung über den Energie-Cutoff

bestimmt.<sup>[52]</sup>

Die Entwicklung der Orbitale in ebenen Wellen hat aufgrund ihrer Eigenschaften gewisse Vor- und Nachteile. Die ebenen Wellen sind im Gegensatz zu AOs nicht an Atomen zentriert und haben keinen Ursprung. Dies ist vorteilhaft bei der Beschreibung von delokalisierten Elektronen, da keine diffusen Basisfunktionen zur AO-Basis hinzugefügt werden müssen.<sup>[52,106]</sup> Allerdings ist eine pw-Basis anfällig für einen *electron spill-out*<sup>[4,5,52]</sup> in QM/MM Simulationen, was bedeutet, dass sich Elektronendichte an positiv geladenen MM-Atomen bildet. Zudem ist es bei der Verwendung einer pw-Basis zwingend notwendig eine Pseudopotential-Näherung<sup>[52,107]</sup> für alle kernnahen Elektronen zu verwenden. Dies liegt daran, dass für eine genaue Beschreibung der kernnahen Orbitale, welche zum Teil große Wertänderungen auf kleinen Längenskalen aufweisen, sehr viele Wellen nötig sind, wodurch die Berechnung von schwereren Elementen praktisch unmöglich ist.<sup>[52]</sup> Durch die Verwendung von Pseudopotentialen und die Entwicklung von *ultraweichen* Pseudopotentialen<sup>[108]</sup> wurde der pw-DFT-Ansatz zu einer effektiven Methode für AIMD und QM/MM-MD Simulationen.

## 2.5 Multiple-Time-Step Algorithmus

Ein Problem, welches besonders bei AIMD- und QM/MM-Simulationen, aber auch bei klassischen MD-Simulationen auftritt, ist der hohe Rechenaufwand bei längeren Simulationszeiten. Die Anzahl an nötigen Rechnungen lässt sich zwar durch die Verwendung größerer Zeitschritte  $\Delta t$  im Algorithmus zur Lösung der Bewegungsgleichungen verringern, allerdings können die Zeitschritte nicht beliebig groß gewählt werden. Um numerische Stabilität zu gewährleisten, sollte der gewählte Zeitschritt deutlich kleiner als die schnellste Bewegung (in der Regel C-H- oder O-H-Bindungsschwingung) im System sein und beträgt daher meist weniger als eine Femtosekunde.<sup>[5,7]</sup>

Eine Möglichkeit, um den verwendeten Zeitschritt zu vergrößern, ist das Einschränken dieser schnellen Freiheitsgrade.<sup>[109,110]</sup> Eine andere Möglichkeit ist die Verwendung von Multiple-Time-Step (MTS) Algorithmen, bei denen unterschiedlich große Zeitschritte in den Simulationen verwendet werden.<sup>[111,112]</sup> Ursprünglich wurden MTS Algorithmen für die Separation von schnellen und langsamen Freiheitsgraden verwendet, welche unterschiedlich oft propagiert werden müssen.<sup>[112–115]</sup> Der rRESPA-Algorithmus von Tuckerman et al.<sup>[116]</sup> wurde allerdings auch in verschiedene MTS-Schemata für AIMD-Simulationen angewandt, bei denen die Separation der verschiedenen Zeitschritte nicht physikalisch begründet ist (z.B. langsam entwickelnde long-range Kräfte<sup>[113,117]</sup>), sondern artifiziell eingeführt wurden.<sup>[118–120]</sup> Dabei werden unterschiedliche quantenchemische Methoden, wie z.B. Hartree-Fock und MP2, für die unterschiedlichen Kräfte verwendet.<sup>[118]</sup>

### 2.5.1 Der MTS-Algorithmus in CPMD

Der im Programm CPMD implementierte MTS-Algorithmus<sup>[117]</sup> erlaubt die Verwendung von unterschiedlichen Funktionalen zur Berechnung von schnellen Referenz- und langsamen Korrekturkräften auf pw-DFT Niveau. Die 'schnellen' Kräfte  $\mathbf{F}_{low}$  werden in jedem Schritt der Simulation auf dem niedrigeren quantenchemischen Niveau berechnet, während die 'langsamen' Kräfte  $\mathbf{F}_{high}$  nur alle  $n$  Schritte auf dem höheren quantenchemischen Niveau berechnet werden. Die Kräfte, die zur Lösung der klassischen Bewegungsgleichungen verwendet werden, ergeben sich in den Referenzschritten nur aus  $\mathbf{F}_{low}$ . In den Korrekturschritten ergibt sich die Kraft über

$$\mathbf{F} = n\mathbf{F}_{high} - (n - 1)\mathbf{F}_{low}. \quad (2.101)$$

Wenn der MTS-Faktor  $n$  nicht zu groß gewählt wird (siehe Kapitel 6.2) entspricht die Genauigkeit der MTS-Simulation einer Simulation auf dem höheren quantenchemischen Niveau, bei einer deutlichen Reduktion der benötigten Rechenzeit. Liberatore et al.<sup>[117]</sup> zeigen, dass diese Separation möglich ist, da die Differenz der auf unterschiedlichen Niveaus berechneten Kräfte, im Vergleich zu den absoluten Kräften, deutlich langsamer fluktuiert.

Durch die Verwendung des in dieser Arbeit implementierten Interfaces ist es möglich nicht nur verschiedene Funktionale für die Berechnung der Referenz- und Korrekturkräfte zu verwenden, sondern eine Kombination von unterschiedlichen quantenchemischen Methoden. Zum Beispiel kann für die Berechnung der Referenzkräfte die kostengünstige pw-DFT Methode gewählt werden und für die Berechnung der Korrekturkräfte die hochgenaue, aber kostenintensive CCSD(T)-Methode.

---

## 3 Implementierung

Im Rahmen dieser Arbeit wurde ein Interface zwischen dem quantenchemischen Programmpaket CFOUR und dem AIMD Programm CPMD bzw. dem klassischen MD-Programm GROMACS mit Hilfe des MiMiC-Frameworks geschrieben. Ziel war es, die vorhandene QM/MM Implementierung des MiMiC-Frameworks zu erweitern, um sowohl AIMD- als auch QM/MM-Simulationen mit auf Wellenfunktionen basierenden quantenchemischen Methoden zu ermöglichen. Zudem wurden Algorithmen zur Geometrieoptimierung des QM-Bereichs in Anwesenheit des MM-Systems und zur Optimierung von ECPs, welche verwendet werden, wenn die Grenze zwischen QM- und MM-Bereich durch für kovalente Bindungen verläuft, implementiert. Mit dem neuen Interface sind folgende Simulationen und Rechnungen durchführbar:

- AIMD Simulationen mit CFOUR<sup>a</sup>
- AIMD-MTS Simulationen mit CFOUR<sup>a</sup> und CPMD<sup>b</sup>
- QM/MM-Simulationen mit CFOUR<sup>a</sup> und GROMACS
- QM/MM-MTS-Simulationen mit CFOUR<sup>a</sup>, CPMD<sup>b</sup> und GROMACS
- QM/MM-Simulationen mit und ohne long-range Wechselwirkungen
- Geometrieoptimierung des QM-Bereichs in Anwesenheit der MM-Atome mit CFOUR<sup>a</sup> und CPMD<sup>b</sup>
- Optimierung von ECPs mit CFOUR, welche verwendet werden, wenn die Grenze zwischen QM- und MM-Bereich durch für kovalente Bindungen verläuft

Die Kommunikation zwischen den einzelnen Programmen wird durch MiMiC über eine Kommunikationsbibliothek<sup>[34]</sup> übernommen. Die einzelnen Programme kommunizieren über ein *Message Passing Interface* (MPI), welches erlaubt, dass die einzelnen Programme separat und unabhängig voneinander ausgeführt werden können. Dadurch können die Hardware-Ressourcen (CPUs und Speicherplatz) effizient genutzt werden, da die einzelnen Programme optimiert für ihre spezielle Aufgabe ausgeführt werden. Die Implementierung des Interfaces mit CFOUR ist in drei Level unterteilt. Das oberste Level beinhaltet den Ablauf der Simulation und das Prinzip der Kommunikation bzw. des Datenaustauschs von CFOUR mit MiMiC. Das zweite

---

<sup>a</sup>Die Methoden, welche mit CFOUR nutzbar sind und mit denen das Interface bereits getestet wurde, sind HF, MP2, CC und CAS-SCF.

<sup>b</sup>Plane-wave-DFT mit Funktionalen unterschiedlicher Qualität.

Level beinhaltet die Implementierung der MD-Schleife innerhalb von CFOUR, die Verarbeitung der Daten, welche über MiMiC ausgetauscht werden und das Aufrufen der benötigten Module von CFOUR. Das dritte Level bezieht sich auf die konkrete Implementierung der elektrostatischen QM/MM-Wechselwirkungen zur Berechnung der Energie und der Kräfte innerhalb der entsprechenden Module von CFOUR.

In diesem Kapitel werden die drei Level des Interfaces erläutert. Anschließend wird die Funktionsweise der Geometrieoptimierung und der Optimierung von ECPs beschrieben. Abschließend werden die benötigten Keywords und die Input-Dateien erklärt.

## 3.1 Level 1: Das Interface

Im ersten Level des Interfaces wird der Ablauf der Simulation geregelt. Das Programm CPMD wird als MD-Driver verwendet, d.h. es löst die Bewegungsgleichungen für die Atomkerne und führt den MD-Schritt aus (siehe Abschnitt 2.3). Wird z.B. eine kombinierte Simulation aus DFT und CCSD(T) durchgeführt, übernimmt CPMD auch die Berechnung des QM-Bereichs für die DFT-Teilschritte (siehe Abschnitt 2.5). GROMACS berechnet die Energie bzw. die Kräfte des MM-Systems sowie die van der Waals QM/MM-Wechselwirkungen. Die Energie und die Kräfte des QM-Systems, sowie die elektrostatischen QM/MM-Wechselwirkungen werden von CFOUR berechnet. Dies ist ein Unterschied zur ursprünglichen Implementierung des MiMiC-Frameworks. Dort wurden die QM/MM-Wechselwirkungen von MiMiC berechnet (siehe transparente Box in Abbildung 3.1). Dies liegt an der Integralform der elektrostatischen QM/MM-Wechselwirkungen (siehe Gleichung 2.18 und 2.19), welche vom QM-Programm (CFOUR) berechnet werden müssen. In Abbildung 3.1 ist der Ablauf der Simulation und die Aufgabe der einzelnen Programme dargestellt. Alle drei Programme werden unabhängig voneinander ausgeführt und kommunizieren via MiMiC miteinander. Bei der Initialisierung wird von jedem Programm die jeweilige Input-Datei eingelesen, MiMiC sammelt alle notwendigen Informationen der Programme und sendet diese an CPMD. Nach Beendigung der Initialisierung beginnt die MD-Schleife und MiMiC sendet alle relevanten Informationen an die jeweiligen Programme. Diese berechnen gleichzeitig die jeweiligen Energien und Kräfte, welche anschließend erneut von MiMiC gesammelt und an CPMD weitergegeben werden, damit CPMD die Bewegungsgleichungen lösen und die entsprechenden Systemeigenschaften (Temperatur, Atomgeschwindigkeiten, etc.) berechnen kann. Anschließend wird der nächste Schritt in der MD-Schleife gestartet.

Um die maximale Flexibilität des Interfaces zu gewährleisten und gleichzeitig die hohe Effizienz der Implementierungen der einzelnen Programme beizubehalten, basiert das MiMiC-Framework auf dem Multiple-Program Multiple-Data Ansatz (MPMD).<sup>[34,35,121]</sup> Das bedeutet, dass die Programme einzeln ausgeführt werden und über eine vordefinierte Programmierschnittstelle (API) verbunden werden. Über diese API kann MiMiC mit einem Host-Programm, welches der MD-Driver ist und mit mehreren Client-Programmen, welche die jeweiligen Energien und Kräfte berechnen,

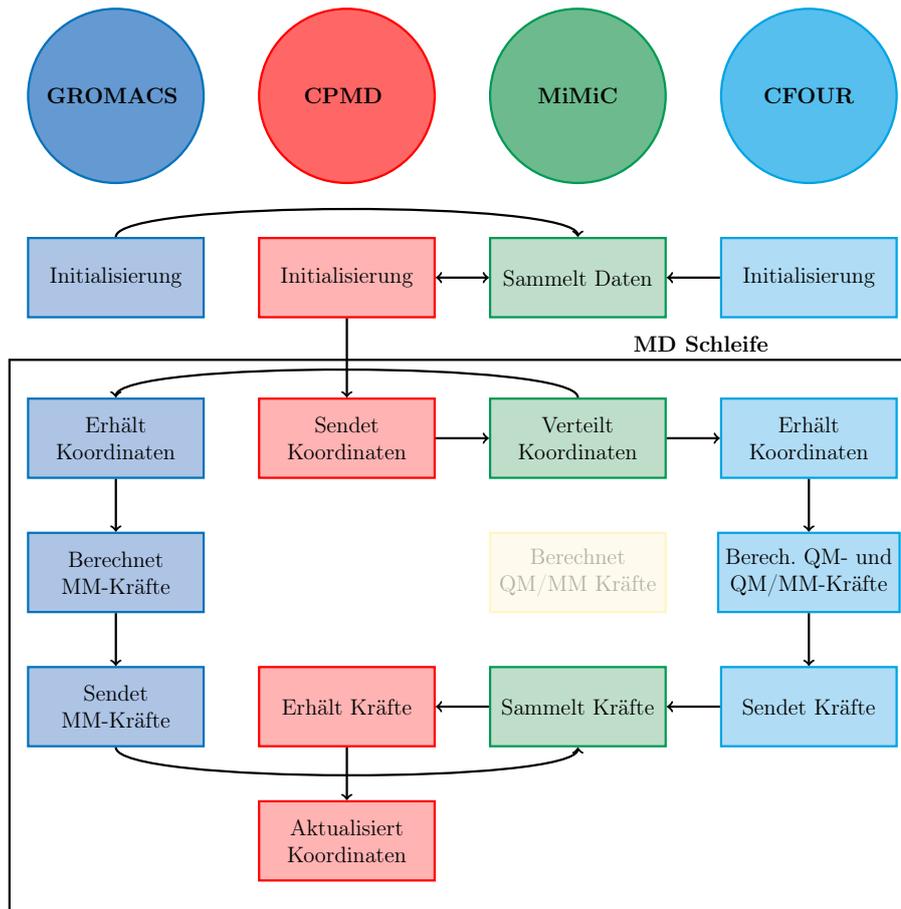


Abbildung 3.1: Schematischer Ablauf einer QM/MM-Simulation mit den Programmen CFOUR (QM-System), GROMACS (MM-System) und CPMD (MD-Driver) über das MiMiC Framework. Die transparente Box bedeutet, dass im Falle einer MTS-Simulation die QM/MM-Kräfte in den pw-DFT-Teilschritten von MiMiC berechnet werden.

verbunden werden. Dadurch wird erreicht, dass die Client-Programme möglichst wenig verändert werden müssen, wenn sie mit MiMiC verknüpft werden.<sup>[34]</sup> Die Kommunikation zwischen den Programmen funktioniert über MPI Kommunikationsroutinen.<sup>[34]</sup> Dieser Ansatz erlaubt die optimale Nutzung von Ressourcen, da jedes Programm seinen eigenen Speicher und eine individuelle Anzahl von Nodes und CPUs verwenden kann. Dadurch sind massiv parallelisierte Anwendungen mit der höchst möglichen Effizienz für die einzelnen Programme möglich.<sup>[35]</sup>

Für die Kommunikation zwischen MiMiC und CFOUR sind drei Kommunikationsroutinen nötig. Zunächst eine Routine zum Aufbauen der Verknüpfung, durch die CFOUR als Client-Programm initialisiert wird. Diese wird bei der Initialisierung der Programme ausgeführt, sodass Client- und Host-Programm "voneinander wissen" und ab dem Moment miteinander kommunizieren können. Anschließend werden jeweils eine Routine zum Versenden von Informationen und zum Erhalten von Informationen

an bzw. vom Host-Programm verwendet. Dabei ist zu beachten, dass das Format der Informationen (Größe des Arrays, Art der Zahlen, etc.) angegeben werden muss.

### 3.2 Level 2: Die *mimic\_driver*-Routine

In der *driver*-Routine in CFOUR wird die Input-Datei gelesen und ausgehend davon werden die benötigten Module von CFOUR, z.B. zur Berechnung der Integrale oder zum Lösen des SCF-Problems, in der richtigen Reihenfolge aufgerufen. Für die Durchführung von QM/MM-MD Simulationen ist es nötig mehrere Tausend QM-Rechnungen hintereinander durchzuführen. Daher wurde eine modifizierte Version der *driver*-Routine erstellt, in der sowohl die Kommunikation mit MiMiC als auch die Durchführung der MD-Schleife innerhalb von CFOUR stattfindet, welche im Falle einer QM/MM-MD Simulation anstatt der ursprünglichen *driver*-Routine aufgerufen wird. Der Aufbau der *mimic\_driver*-Routine ist in Abbildung 3.2 dargestellt und wird im Folgenden detaillierter erklärt. Dabei wird sich allerdings nur auf die Modifikationen bezüglich des implementierten QM/MM-Interfaces und damit auf die Unterschiede zur ursprünglichen *driver*-Routine bezogen. Alle dafür benötigten Routinen gehören zum Modul *mimic\_mod.f90*.

Als erstes wird die MPI-Umgebung initialisiert und die Verknüpfung mit MiMiC aufgebaut, um eine Kommunikation zwischen CFOUR und MiMiC zu ermöglichen. Dies erfolgt durch die Routine *mimic\_initialize*. Diese ruft die entsprechenden MPI-Routinen bzw. MiMiC-CommLib<sup>[34]</sup> Routinen auf. Der zweite Schritt ist die Initialisierung der CFOUR-Rechnung selbst und das Erhalten aller relevanten Informationen von MiMiC. Dafür wird die Routine *mimic\_init\_data* verwendet, welche zunächst einige Arrays mit grundlegenden Informationen wie Elementsymbol, Masse

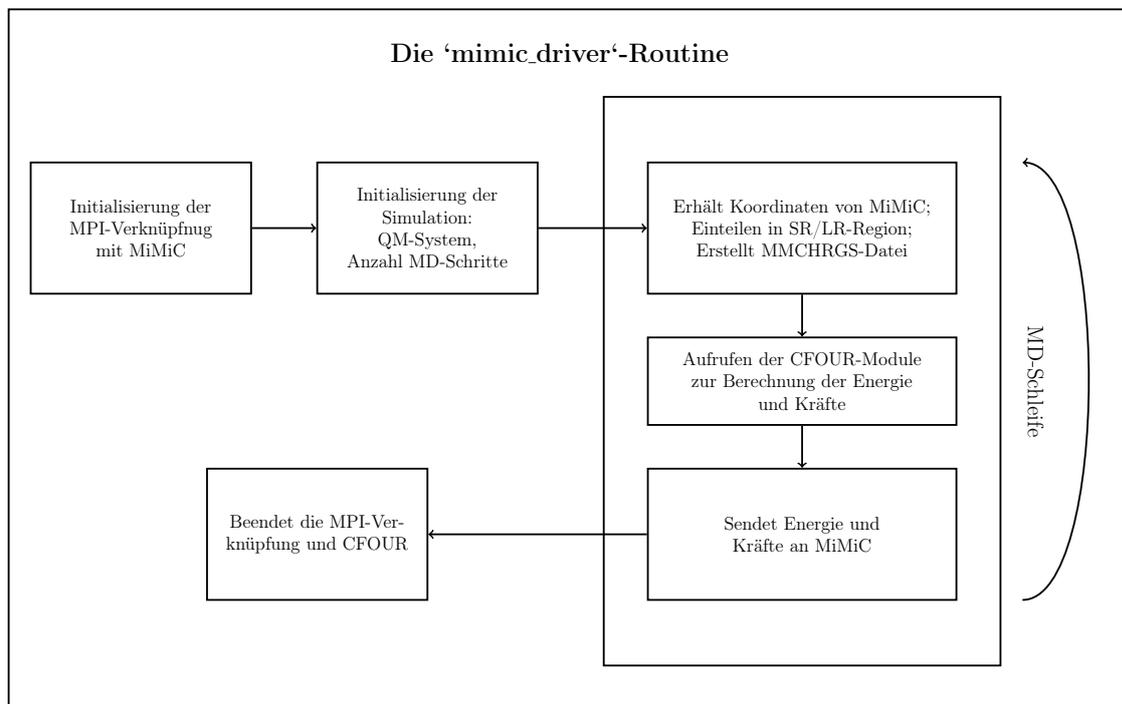


Abbildung 3.2: Schematischer Aufbau der Routine *mimic\_driver*.

und Ordnungszahl initialisiert. Anschließend wird die Input-Datei eingelesen und damit die Informationen über die Größe und enthaltenen Elemente des QM-Bereichs erhalten. Zuletzt werden eine Reihe von Dummy-Informationen, welche allesamt Null sind, da sie sich auf MM-Client-Programme beziehen, an MiMiC gesendet und die Anzahl der MD-Schritte und der MTS-Faktor von MiMiC erhalten.

Damit ist die Initialisierung abgeschlossen und die MD-Schleife beginnt. Als erstes werden innerhalb der MD-Schleife die Koordinaten des gesamten Systems und die MM-Partialladungen zur Berechnung der QM/MM-Wechselwirkungen erhalten und auf einen Array geschrieben. Dafür wird die Routine *mimic\_receive\_coordinates* verwendet. Bei der Verwendung von long-range Wechselwirkungen werden in der Routine *sr\_lr\_com* zunächst alle Koordinaten so verschoben, dass das Massenzentrum (COM) des QM-Bereichs im Ursprung liegt. Anschließend werden die MM-Partialladungen nach Abstand zum COM sortiert und entsprechend des in der Input-Datei angegebenen cut-off Radius in short-range und long-range Region eingeteilt. Bevor die entsprechenden Module von CFOUR zur Berechnung des QM-Systems aufgerufen werden können müssen die Koordinaten in den Dateien *ZMAT*, *MOL* und *BASIS* aktualisiert werden, welche Informationen über das QM-System und die verwendete Basis beinhalten. Zusätzlich werden alle Koordinaten und Ladungen der QM- und MM-Atome in die Datei *MMCHRGs* geschrieben. Diese Datei wird von den verschiedenen Modulen von CFOUR zur Berechnung der QM/MM-Wechselwirkungen benötigt.

Nun werden die benötigten Module von CFOUR je nach QM-Level aufgerufen und die Energie und Kräfte für das QM-System und die QM/MM-Wechselwirkungen berechnet. Die dafür nötigen Modifikationen in den entsprechenden Modulen werden in Abschnitt 3.3 erläutert. Anschließend wird die berechnete Energie zusammen mit den Kräften auf die jeweiligen Atome in der Routine *mimic\_send\_forces* an MiMiC gesendet und die nächste Iteration kann gestartet werden.

Am Ende der Simulation werden relevante Größen, wie z.B. das Dipolmoment des QM-Systems der jeweiligen Schritte, auf Dateien geschrieben und anschließend wird die Verknüpfung mit MiMiC bzw. die MPI-Umgebung beendet.

## 3.3 Level 3: Die QM/MM-Wechselwirkungen

In den ersten beiden Leveln des Interfaces wird der Ablauf der QM/MM-Simulation bzw. die Kommunikation und der Datentransfer zwischen den Programmen geregelt. Im dritten Level des Interfaces werden die benötigten QM/MM-Wechselwirkungen berechnet. Dazu gehören die Energie des QM-Systems, die QM/MM-Wechselwirkungsenergie sowie die Kräfte auf die QM- und MM-Atome durch die QM/MM-Wechselwirkungen. Bei der Berechnung der Energie und der Kräfte muss unterschieden werden, ob es sich um Kern-Kern-Wechselwirkungen (siehe Gleichung 2.16 und 2.17) oder Elektron-Kern-Wechselwirkungen (siehe Gleichung 2.18 und 2.19) handelt und bei letzteren, ob sich die MM-Atome in der short-range oder long-range Region befinden.

### 3.3.1 Berechnung der QM/MM-Wechselwirkungsenergien

Für die einzelnen Wechselwirkungen wurden entsprechende Modifikationen in unterschiedlichen Modulen von CFOUR durchgeführt. Für die Berechnung der Energie sind Modifikationen in drei Modulen nötig: in *vmol*, *mint* und in *vsf*. Der Ablauf zur Berechnung der QM/MM-Energien ist in Abbildung 3.3 dargestellt.

Als erstes wird das Modul *vmol* aufgerufen. Hier werden die Kern-Kern-Wechselwirkungsenergie sowie die Einelektronenintegrale berechnet. Zunächst müssen die Koordinaten und Ladungen aller Atome, sowohl QM- als auch MM-Atome, aus der MMCHRGs-Datei eingelesen und auf einen entsprechenden Array geschrieben werden. Dies wird durch die Routine *get\_chrgs* am Anfang des Moduls durchgeführt. In der Routine *readin* wird unter anderem die Kern-Kern-Wechselwirkungsenergie berechnet. Im Falle einer QM/MM-Simulation wird hier die Routine *calc\_nucrep* aufgerufen, welche sowohl die Kern-Kern-Wechselwirkungen des QM-Systems als auch die QM/MM-Kern-Kern-Wechselwirkungen berechnet.

Die Einelektronenintegrale werden in der Routine *onelh* berechnet und anschließend wird der Einelektronen-Hamiltonoperator gebildet. Durch das Aufrufen der Routine *add\_oneh* werden hier die entsprechenden QM/MM-Integrale (siehe Gleichung 2.15) für alle MM-Atome in der short-range Region berechnet und zum Einelektronen-Hamiltonoperator hinzugefügt. Wenn alle MM-Atome explizit berücksichtigt werden, sich also in der short-range Region befinden, sind die Modifikationen in *vmol* die einzigen nötigen Modifikationen. Für die Wechselwirkungen mit den MM-Atomen aus der long-range Region bedarf es noch weiterer Modifikationen.

Da die QM/MM-Wechselwirkungen mit den Atomen aus der long-range Region über eine Multipolentwicklung des Coulombpotentials (siehe Gleichung 2.20) berechnet werden, werden für die QM/MM-Energie die Integrale der 0. bis 4. Ordnung (Überlappungsintegrale bis Hexadecapolintegrale) benötigt. Diese werden in dem Modul *mint* zusammen mit den jeweiligen Ableitungen, welche für die Berechnung der Kräfte benötigt werden, berechnet und in entsprechende Dateien geschrieben. Da nur die Integrale des QM-Systems benötigt werden, müssen in *mint* nicht die Koordinaten bzw. Ladungen der MM-Atome eingelesen werden.

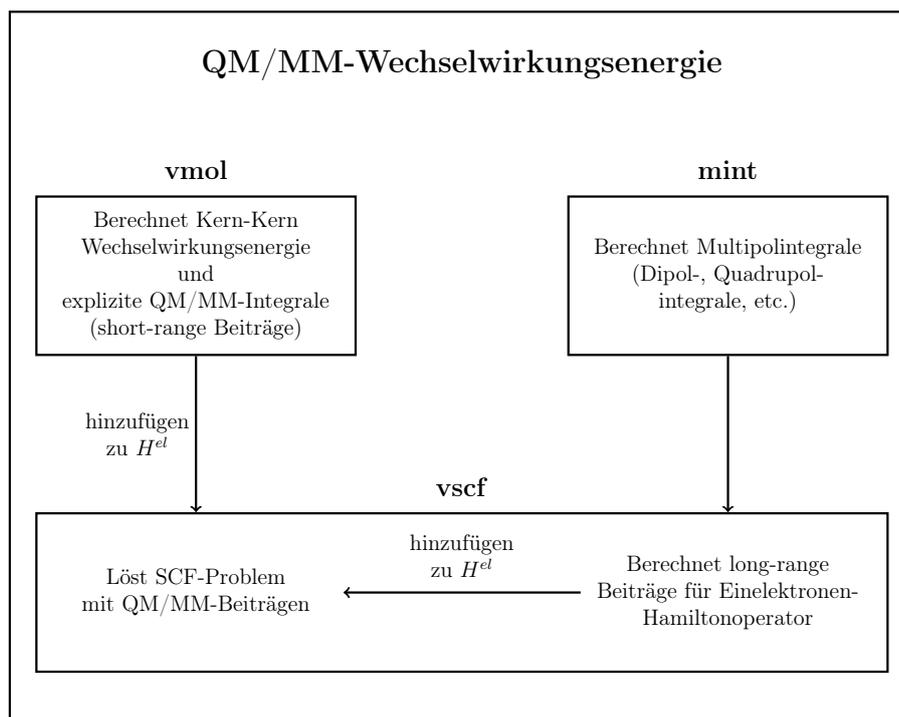


Abbildung 3.3: Schematischer Ablauf der Berechnung der QM/MM-Wechselwirkungsenergie.

Die long-range QM/MM-Beiträge zum Einelektronen-Hamiltonoperator werden in dem Modul *vscf*, in dem auch das SCF-Problem gelöst wird, berechnet und aufaddiert. Dazu wird in der Routine *get1e*, in welcher der Hamiltonoperator gebildet wird, erneut die entsprechende Routine *get\_chrgs* aufgerufen. Die Version dieser Routine in dem Modul *vscf* liest allerdings nicht nur die Koordinaten und Ladungen der MM-Atome ein, sie berechnet zusätzlich die long-range QM/MM-Wechselwirkungen. Nachdem die Koordinaten und Ladungen aller MM-Atome eingelesen werden, können damit die Vorfaktoren aus Gleichung 2.20, welche nur von den Koordinaten der MM-Atomen abhängen, berechnet werden. Anschließend werden die in *mint* berechneten Integrale eingelesen und in die richtige Form gebracht, da sich die Orbitalreihenfolge in *mint* und *vmol* unterscheidet. Aus den Vorfaktoren und den Integralen werden anschließend die benötigten Beiträge berechnet und zum Einelektronen-Hamiltonoperator hinzugefügt, sodass das SCF-Problem unter Einbeziehung der zusätzlichen elektrostatischen QM/MM-Beiträge gelöst werden kann.

### 3.3.2 Berechnung der QM/MM-Kräfte

Für die Berechnung der Kräfte sind zusätzliche Modifikationen in dem Modul *vdint* nötig. Sowohl die QM/MM-Kräfte für MM-Atome aus der short-range Region als auch für die Atome aus der long-range Region werden hier berechnet. Der Ablauf zur Berechnung der entsprechenden Kräfte ist in Abbildung 3.4 dargestellt. Ebenso wie für die Berechnung der Energie werden zunächst in der Routine *get\_chrgs* die Koordinaten und Ladungen aller Atome eingelesen. Anschließend werden in der Routine *qmmm\_nucrep* die Kräfte berechnet, welche aus der Kern-Kern-Wechselwirkung resultieren (siehe Gleichung 2.16 und 2.17). Zur Berechnung der Kräfte aufgrund der expliziten Wechselwirkungen zwischen short-range MM-Atomen und QM-Elektronen wurden die entsprechenden, bereits vorhandenen Routinen verwendet und lediglich die Arraygrößen und die Summationsgrenzen verändert, sodass alle MM-Atome aus der short-range Region berücksichtigt werden.

Die Kräfte, welche aus den long-range Wechselwirkungen resultieren, werden in der Routine *lr\_force\_mint* berechnet. Dort werden zunächst die in *mint* berechneten Multipolintegrale eingelesen, in die richtige Form gebracht und mit der Dichtematrix kontrahiert, um Ladung, Dipolmoment, Quadrupolmoment etc. und die entsprechenden Ableitungen zu erhalten, welche zur Berechnung der Kräfte auf die MM-Atome

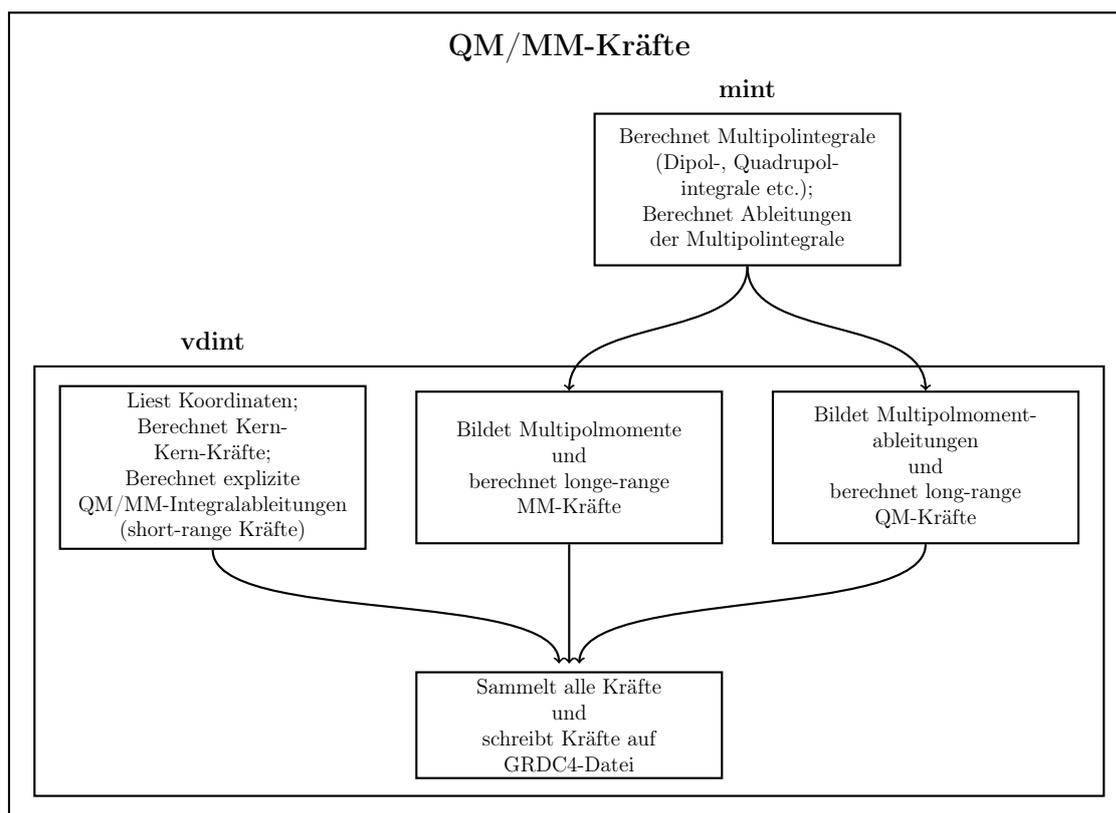


Abbildung 3.4: Schematischer Ablauf der Berechnung der Kräfte aufgrund der QM/MM-Wechselwirkungen.

(siehe Gleichung 2.26) bzw. auf die QM-Atome (siehe Gleichung 2.27) benötigt werden. Anschließend werden die Kräfte auf die MM-Atome berechnet. Bevor die Kräfte auf die QM-Atome berechnet werden können, müssen die bereits in *uscf* berechneten Vorfaktoren eingelesen werden. Anschließend werden mit diesen Vorfaktoren und den Ableitungen der Multipole die Kräfte auf die QM-Atome berechnet.

Alle berechneten Kräfte, d.h. die Kräfte für den QM-Bereich, die Kräfte aufgrund der Kern-Kern-Wechselwirkung und die short-range und die long-range QM/MM-Beiträge, werden zum Schluss von der Routine *collect\_force* gesammelt und auf die Datei GRDC4 geschrieben, welche dann in der *mimic\_driver*-Routine gelesen und an MiMiC gesendet werden können.

### 3.4 Geometrieoptimierung

Um Eigenschaften des QM-Bereichs in Anwesenheit der MM-Atome zu bestimmen, ist es oftmals nötig während der Simulation Snapshots des Systems, d.h. die Koordinaten aller Atome zu einem bestimmten Zeitpunkt, zu speichern, um anschließend, ausgehend von diesen Snapshots, die Eigenschaften zu berechnen.<sup>[122]</sup> Dazu ist es manchmal nötig, dass sich der QM-Bereich in einem Gleichgewichtszustand befindet.<sup>[122]</sup> Dazu muss eine Geometrieoptimierung des QM-Bereichs unter Berücksichtigung aller QM/MM-Beiträge durchgeführt werden. Da der Algorithmus zur Geometrieoptimierung in CFOUR nur mit internen Koordinaten funktioniert, musste eine neue Geometrieoptimierung implementiert werden. Diese ist eine gradientenbasierende (*steepest descent*) Geometrieoptimierung, bei welcher der Schritt in

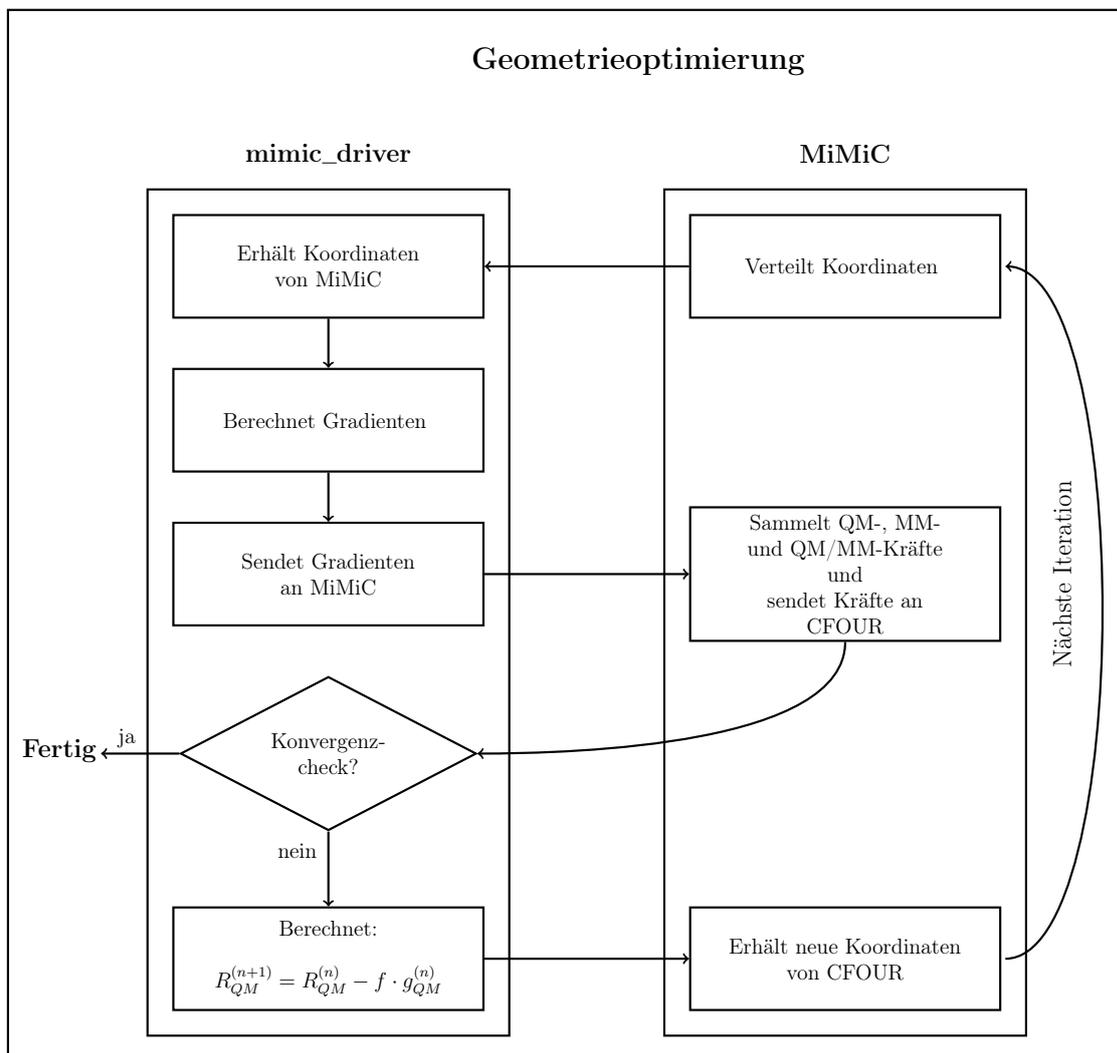


Abbildung 3.5: Schematischer Ablauf einer Geometrieoptimierung des QM-Bereichs unter Berücksichtigung der QM/MM-Kräfte.

jeder Iteration von dem Gradienten bestimmt wird. Die Geschwindigkeit, mit der die Geometrieoptimierung konvergiert, wird nur über einen Faktor verbessert, mit dem der Gradient multipliziert wird. Die Geometrieoptimierung wird in der Routine *mimic\_driver* durchgeführt. Der Ablauf ist in Abbildung 3.5 dargestellt.

Zu Beginn der Iteration werden die Koordinaten des Systems von MiMiC erhalten. Anschließend werden die Kräfte auf die gleiche Weise wie bei einer QM/MM-Simulation berechnet und wieder an MiMiC gesendet. MiMiC sammelt nun alle Kräfte, welche von den unterschiedlichen Programmen berechnet wurden und sendet diejenigen Kräfte, die auf die QM-Atome wirken, wieder an CFOUR zurück. Am Ende der Iteration wird die Routine *do\_geo\_opt* aufgerufen, welche zunächst die Konvergenz der letzten fünf Schritte berechnet und anhand dieser den Skalierungsfaktor  $f$  für den Gradienten bestimmt. Wenn das angegebene Konvergenzkriterium erreicht wird, wird die Geometrieoptimierung abgebrochen. Ist das Konvergenzkriterium nicht erreicht, werden die neuen Koordinaten über

$$R_{QM}^{(n+1)} = R_{QM}^{(n)} - f \cdot g_{QM}^{(n)} \quad (3.1)$$

berechnet und erneut an MiMiC gesendet. Anschließend wird die nächste Iteration gestartet.

### 3.5 Optimierung von ECPs

Die Verwendung von optimierten *effective core potentials* (ECPs) für den Fall, dass die Grenze zwischen QM- und MM-Bereich durch kovalente Bindungen verläuft, hat sich in vergangenen Studien als sehr geeignet erwiesen.<sup>[57,123]</sup> Ein Vorteil dieser Methode ist der relativ geringe Aufwand für das Erstellen von optimierten ECPs für verschiedene Situationen. Aus diesem Grund und aufgrund der Tatsache, dass die meisten bereits verfügbaren optimierten ECPs nur für DFT bzw. plane-wave DFT Implementierungen zu verwenden sind, wurde in CFOUR ein Algorithmus zur Optimierung von ECPs implementiert. Das Vorgehen der Optimierung wird im Folgenden erläutert.

Für den QM/MM-Bindungsbruch sollten kovalent gebundene Kohlenstoffatome verwendet werden. Das Kohlenstoffatom, welches den QM-Bereich abschließt, wäre demnach nur einfach gebunden und würde ungepaarte Elektronen besitzen, welche die Elektronenstruktur des QM-Bereichs stark stören würden. Daher wird für dieses Kohlenstoffatom ein ECP verwendet, welches nicht nur die kernnahen Elektronen, sondern alle Elektronen, bis auf das in der Bindung beteiligte Elektron, beschreibt. Dadurch wird das Kohlenstoffatom quasi monovalent und der QM-Bereich wird möglichst wenig gestört.

Um die verbleibende Störung möglichst gering zu halten, werden die Parameter eines bereits vorhandene ECPs so optimiert, dass die Eigenschaften eines zur Optimierung gewählten Testsystems möglichst exakt reproduziert werden. Dazu wird eine Penalty-Funktion<sup>[57]</sup>

$$P[n(\mathbf{r})] = \int d\mathbf{r} |n(\mathbf{r})^{ref} - n(\mathbf{r})|^2 \quad (3.2)$$

mit der Referenzeigenschaft  $n(\mathbf{r})^{ref}$  verwendet, welche minimiert wird. Als Referenzeigenschaften können in dieser Implementierung das Dipolmoment, der Gradient bzgl. der Kernverschiebung, eine Kombination aus Dipolmoment und Gradient und die Ladungsdichte des Moleküls verwendet werden. Grundsätzlich kann aber jede weitere, geeignete Eigenschaft wie z.B. Partiaalladungen zur Optimierung verwendet werden.

Die Minimierung funktioniert über einen *steepest descent* Algorithmus, daher wird die Ableitung der Penalty-Funktion  $P[n(\mathbf{r})]$  bezüglich der ECP-Parameter  $\sigma_i$  benötigt, welche über

$$\frac{dP[n(\mathbf{r})]}{d\sigma_i} = 2 \int d\mathbf{r} |n(\mathbf{r}) - n(\mathbf{r})^{ref}| \cdot n'_i(\mathbf{r}) \quad (3.3)$$

mit  $n'_i(\mathbf{r})$ , der Ableitung der Eigenschaft  $n(\mathbf{r})$  bezüglich des ECP-Parameters  $\sigma_i$ , gegeben ist. Die Ableitung der Eigenschaft  $n'_i(\mathbf{r})$  wird mittels numerischer Differentiation

über

$$n'_i(\mathbf{r}) = \frac{n_i(\mathbf{r}, \sigma_i + x) - n_i(\mathbf{r}, \sigma_i - x)}{2x} \quad (3.4)$$

erhalten und die neuen ECP-Parameter werden wie bei der Geometrieoptimierung über

$$\sigma_i^{(n+1)} = \sigma_i^{(n)} - f \cdot \frac{dP[n(\mathbf{r})]}{d\sigma_i^{(n)}} \quad (3.5)$$

berechnet.

Der Ablauf der ECP-Optimierung ist in Abbildung 3.6 dargestellt. Die Routine *ecp\_fit*, welche die ECP-Optimierung durchführt, wird anstelle der *driver*-Routine aufgerufen.

Zunächst wird die Optimierung initialisiert. Dazu werden alle Informationen der Input-Datei eingelesen. Dazu gehört die Eigenschaft, bezüglich der das ECP optimiert werden soll, die maximale Anzahl an Iterationen, die Art des ECPs und der Startwert des Faktors  $f$ , mit der der Gradient in der Optimierung multipliziert wird, um die Minimierung zu beschleunigen. Anschließend wird die entsprechende Referenzeigenschaft von einer Datei eingelesen. Jetzt wird das erste Mal die *driver*-Routine aufgerufen, wodurch eine erste single-point-Rechnung an dem Testsystem mit dem ursprünglichen ECP durchgeführt wird, um den Ausgangswert der entsprechenden Eigenschaften zu erhalten. Als letztes werden die Parameter des ECPs eingelesen.

Anschließend an die Initialisierung wird die Optimierungsschleife gestartet. Der erste Schritt innerhalb der Optimierung ist die Berechnung der Ableitung der entsprechenden Eigenschaft bezüglich der ECP-Parameter (siehe Gleichungen 3.3 und 3.4). Dazu ist für jeden der zu optimierenden ECP-Parameter je eine single-point-Rechnung mit positiver ( $\sigma_i + x$ ) und negativer ( $\sigma_i - x$ ) Abweichung des Parameterwerts nötig. Daher wird eine zweite Schleife über alle zu optimierenden ECP-Parameter gestartet. Darin wird zuerst eine neue ECP-Datei mit positiver Abweichung eines Parameters geschrieben und anschließend die *driver*-Routine aufgerufen, wodurch der entsprechende Wert der Eigenschaft erhalten wird. Anschließend wird dasselbe mit einer negativen Abweichung durchgeführt und mit diesen Werten die Ableitung der Penalty-Funktion bezüglich des entsprechenden Parameters berechnet.

Nachdem dies für jeden der Parameter geschehen ist, werden die neuen ECP-Parameter über Gleichung 3.5 berechnet und die neue ECP-Datei geschrieben. Zuletzt wird erneut die *driver*-Routine aufgerufen, um den neuen Wert der Eigenschaft, bezüglich der optimiert wird, zu erhalten. Mit dem neuen Wert kann nun die Konvergenz überprüft werden und gegebenenfalls ein neuer Faktor  $f$  berechnet werden. Ist das Konvergenzkriterium oder die maximale Anzahl an Iterationen erreicht, wird die Optimierung abgebrochen, ansonsten startet die nächste Iteration.

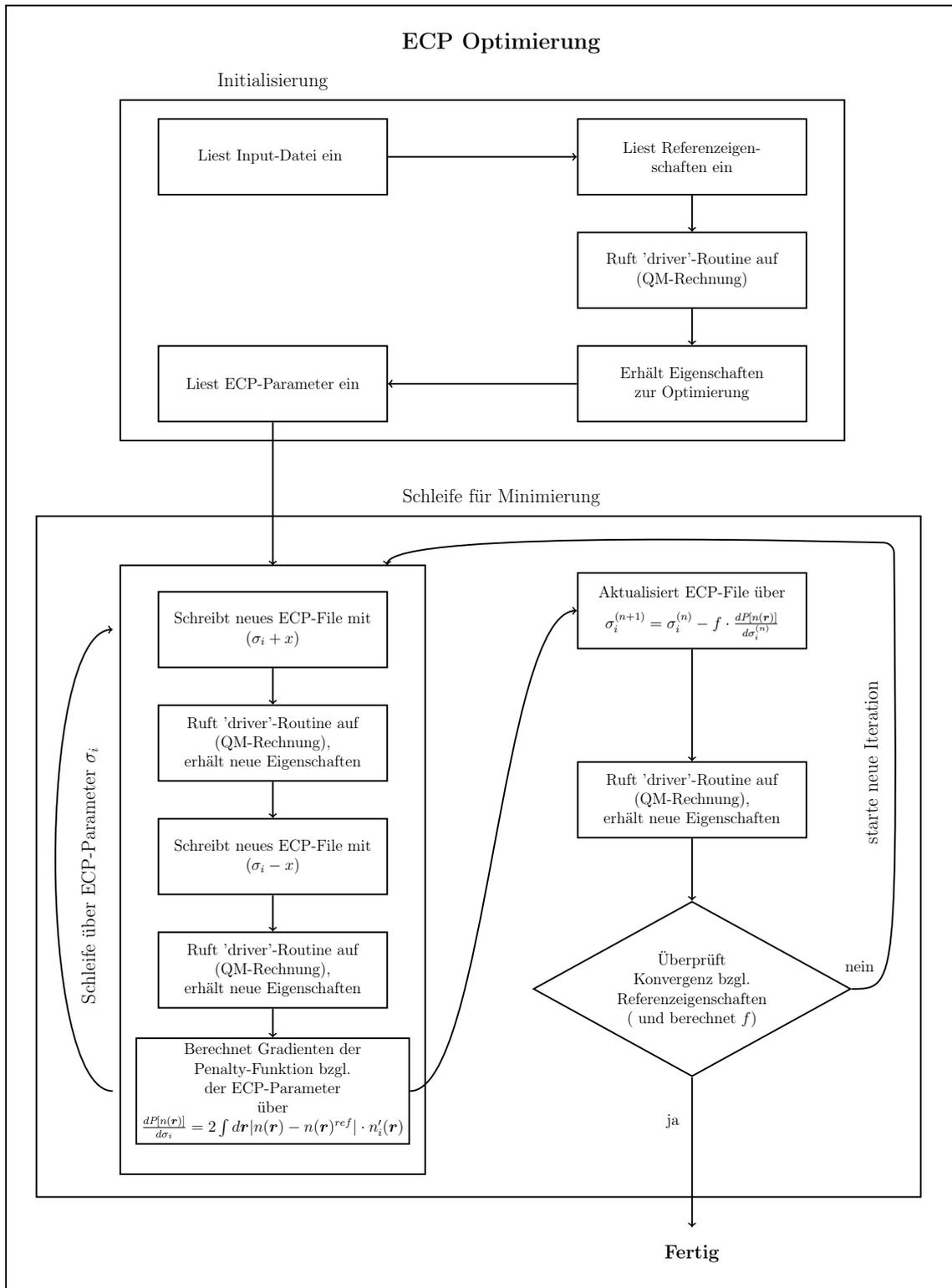


Abbildung 3.6: Schematischer Ablauf der Optimierung eines ECP.

## 3.6 Keywords und Input-Dateien

Im Folgenden werden alle Keywords erläutert, die für die Verwendung des Interfaces nötig sind. Neben den allgemeinen Keywords für CFOUR zur Wahl der Methode, des Basissatzes oder des Speicherplatzes gibt es zwei Gruppen an Keywords, welche für die Benutzung des Interfaces relevant sind. Zum einen gibt es spezifische Keywords für das Interface, welche für die Durchführung der Simulation (bzw. der Optimierungen) verantwortlich sind. Zum anderen gibt es Keywords, welche im Grunde unabhängig vom Interface sind, allerdings müssen sie für eine fehlerfreie Durchführung einer AIMD- bzw. QM/MM-Simulation einen bestimmten Wert annehmen. Der allgemeine Aufbau der ZMAT-Datei für eine Rechnung mit CFOUR kann im CFOUR-Manual<sup>[124]</sup> nachgelesen werden.

### 3.6.1 Spezifische Keywords für CFOUR

#### **%MIMIC**

Mit diesem Keyword wird die Verwendung der Routine *mimic\_driver* initiiert und damit die Kommunikation mit MiMiC. Für das Durchführen einer AIMD- oder QM/MM-Simulation ist dieses Keyword nötig. Soll hingegen nur eine einzelne QM-Rechnung in Anwesenheit der MM-Partialladungen durchgeführt werden, ist dieses Keyword nicht anzugeben. Das Keyword wird in der Sektion für zusätzlichen Input in der ZMAT-Datei angegeben (siehe Abbildung 3.7).

#### **cutoff**

Mit diesem Keyword wird der Wert des Cutoff-Radius für die Verwendung von elektrostatischen long-range Wechselwirkungen angegeben. Der Wert wird von der nachfolgenden Zeile eingelesen. Die Einheit ist Bohr und der Defaultwert ist 0. Das Keyword muss unter dem Keyword **%MIMIC** angegeben werden.

#### **pri\_mmch**

Mit diesem Keyword wird angegeben, dass alle  $n$  Schritte die Koordinaten und Ladungen des Systems herausgeschrieben werden sollen, damit anschließend weiterführende Rechnungen mit diesen Snapshots durchgeführt werden können. Der ganzzahlige Wert  $n$  wird von der nachfolgenden Zeile eingelesen. Der Defaultwert ist 0. Das Keyword muss unter dem Keyword **%MIMIC** angegeben werden.

**qmmm\_geo**

Mit diesem Keyword wird angegeben, dass eine Geometrieoptimierung des QM-Bereichs unter Berücksichtigung aller QM/MM-Wechselwirkungen durchgeführt wird. In der nachfolgenden Zeile wird das Konvergenzkriterium  $n$  für die Geometrieoptimierung angegeben. Die Geometrieoptimierung endet, wenn die maximale Kraft, die auf eines der QM-Atome wirkt, kleiner als  $10^{-n}$  ist. Der Defaultwert ist  $n = 7$ . Das Keyword muss unter dem Keyword **%MIMIC** angegeben werden.

**%ECP**

Mit diesem Keyword wird eine ECP-Optimierung initiiert, mit der ECPs zur Verwendung für QM/MM-Bindungsbrüche angepasst und optimiert werden können. Für weitere Informationen siehe Abschnitt 3.5. Das Keyword wird in der Sektion für zusätzlichen Input in der ZMAT-Datei angegeben (siehe Abbildung 3.7).

**EXTERN\_POT**

Mit diesem Keyword wird die Berechnung der elektrostatischen QM/MM-Wechselwirkungen zwischen QM-Bereich und den in der Datei MMCHRGs gespeicherten MM-Atomen aktiviert. Die Optionen sind (=0), bei der keine QM/MM-Wechselwirkungen berücksichtigt werden (z.B. bei AIMD-Simulationen) und (=2), wodurch die Berechnung der QM/MM-Wechselwirkungen aktiviert wird. Die Option (=1) ermöglicht auch die Berechnung des QM-Bereichs in Anwesenheit von externen Punktladungen, allerdings ist die Anzahl der externen Ladungen auf 150 limitiert.<sup>[124]</sup> Das Keyword wird in der Keyword-Sektion in der ZMAT-Datei angegeben.

**3.6.2 Weitere Keywords für CFOUR**

Die folgenden Keywords müssen angegeben werden bzw. den angegebenen Wert annehmen, damit eine fehlerfreie Durchführung einer AIMD- bzw. QM/MM-Simulation möglich ist. Alle Keywords werden in der Keyword-Sektion der ZMAT-Datei angegeben (siehe Abbildung 3.7).

**COORD=CARTESIAN**

Dieses Keyword ist nötig, da alle Dateien, die das Interface schreibt, kartesische Koordinaten verwenden. Durch dieses Keyword wird die Verwendung von kartesischen Koordinaten in der ZMAT-Datei aktiviert.

#### **UNITS=BOHR**

Dieses Keyword ist nötig, da alle Koordinaten, die von MiMiC erhalten werden, in Bohr angegeben sind.

#### **SYM=OFF**

Dieses Keyword zwingt CFOUR bei der Berechnung keine Symmetrie (Punktgruppe  $C_1$ ) zu verwenden. Dies ist zusammen mit dem Keyword **FIXGEOM=ON** nötig, damit der QM-Bereich nicht unabhängig von den MM-Atomen gedreht oder verschoben wird, sondern die Koordinaten exakt so bleiben, wie sie von MiMiC erhalten werden.

#### **FIXGEOM=ON**

Dieses Keyword sorgt zusammen mit dem Keyword **SYM=OFF** dafür, dass der QM-Bereich nicht unabhängig von den MM-Atomen gedreht oder verschoben wird, sondern die Koordinaten exakt so bleiben, wie sie von MiMiC erhalten werden.

#### **DERIV\_LEVEL=1**

Dieses Keyword aktiviert die Berechnung von Kräften durch CFOUR. Dies ist nötig, da die Kräfte zum Lösen der Bewegungsgleichungen benötigt werden.

### **3.6.3 CPMD Keywords zur Benutzung des Interfaces**

Um das Interface zu CFOUR benutzen zu können, müssen in der Input-Datei für CPMD ebenfalls entsprechende Keywords angegeben werden. Generelle Informationen zum Ablauf und Input für eine Rechnung mit CPMD können dem CPMD-Manual<sup>[125]</sup> entnommen werden.

#### **CPMD [AIMD, GEO]**

Dieses Keyword aktiviert die Benutzung des Interfaces zu CFOUR. Es muss in der *&CPMD*-Sektion der Input-Datei für CPMD angegeben werden. Die Erweiterungen **AIMD** und **GEO** müssen angegeben werden, wenn eine AIMD-Simulation bzw. eine Geometrieoptimierung mit CFOUR durchgeführt werden soll.

### 3.6.4 Input-Datei für eine QM/MM-Simulation

Die ZMAT-Datei, welche alle Informationen für CFOUR enthält, ist in drei Bereiche eingeteilt. Im ersten Bereich wird die Geometrie des Systems angegeben, anschließend folgen die Keyword-Sektion und die Sektion für zusätzlichen Input. Die in Abschnitt 3.6.1 und 3.6.2 erläuterten Keywords werden in den letzten beiden Bereichen angegeben. Die Geometrie des QM-Systems muss nicht zu Beginn der Simulation in der ZMAT-Datei angegeben werden. Es müssen lediglich die richtigen Atome in der selben Reihenfolge, wie sie in CPMD verwendet werden, angegeben werden (siehe CPMD-Manual<sup>[125]</sup>). Die Koordinaten betragen dabei 0.0. Eine Beispiel-Datei für eine QM/MM-Simulation für ein Wassermolekül im QM-Bereich auf CCSD(T)/cc-pVTZ Niveau mit einem long-range Cutoff-Radius von 50 a.u. ist in Abbildung 3.7 dargestellt.

```

+ ZMATint
1 H2O .....
2 O    0.0  0.0  0.0
3 H    0.0  0.0  0.0
4 H    0.0  0.0  0.0
5 .....
6 *CFOUR(CALC=CCSD(T) .....
7 BASIS=PVTZ .....
8 COORD=CARTESIAN .....
9 UNITS=BOHR .....
10 SCF_CONV=9 .....
11 SYM=OFF .....
12 FIXGEOM=ON .....
13 DERIV_LEVEL=1 .....
14 EXTERN_POT=2 .....
15 GEO_CONV=9 .....
16 MEMORY=15 .....
17 MEM_UNIT=GB) .....
18
19 %MIMIC
20 cutoff
21 50
22

```

Geometrie-Sektion

Keyword-Sektion

Sektion für zusätzlichen Input

Abbildung 3.7: Beispiel einer ZMAT-Datei für eine QM/MM-Simulation für ein System mit einem Wassermolekül im QM-Bereich.

### 3.6.5 Input-Dateien für eine ECP-Optimierung

Für eine ECP-Optimierung werden folgende Dateien benötigt: Die Input-Datei "ecp\_info", welche die benötigten Informationen für die Initialisierung der Optimierung enthält, die Datei, welche die Referenzeigenschaft zur Optimierung beinhaltet und die ECP-Datei "ECP\_FILE" selbst, die das zu optimierende ECP enthält. Die verschiedenen Dateien werden im Folgenden erläutert.

**eCP\_info**

Die Datei besteht aus 4 Zeilen. In der ersten Zeile wird angegeben, welche Art von Eigenschaft zur Optimierung verwendet werden soll. Die möglichen Optionen sind **dipo**, für das Dipolmoment, **grad**, für den Gradienten, **comb**, für eine Kombination aus Dipolmoment und Gradienten und **dens**, für die Ladungsdichte. In der zweiten Zeile wird die maximale Anzahl an Iterationen angegeben. In der dritten Zeile wird die Art des ECPs angegeben. Zur Auswahl stehen **sma**, für ein kleines ECP mit 6 zu optimierenden Parametern und **big**, für ein großes ECP mit 28 zu optimierenden Parametern. In der vierten und letzten Zeile wird der Startwert des Faktors  $f$  angegeben.

```
+ ecp_info
1 dens
2 600
3 sma
4 2.00
```

Abbildung 3.8: Beispiel der Input-Datei für eine ECP-Optimierung.

**dip\_ref und ch\_ref**

Die Dateien "dip\_ref" und "ch\_ref" enthalten das Dipolmoment bzw. die Ladungsdichten des Referenzsystems in atomaren Einheiten. Für das Dipolmoment wird in der ersten Zeile die x-Komponente und in den weiteren die y- und z-Komponente des Dipolmoments angegeben. Für die Ladungsdichten wird in der ersten Zeile die Anzahl an Atomen, für welche die Ladungsdichte berücksichtigt werden soll, angegeben. In den folgenden Zeilen werden jeweils die entsprechenden Ladungsdichten der Atome angegeben.

```
+ dip_ref          + ch_ref
1 0.3087978438    1 4
2 -0.6380683530  2 120.768823730208
3 0.00000000006  3 297.586445541641
                  4 297.262952779445
                  5 0.359864970985
```

Abbildung 3.9: Beispiele der Dateien für die Referenzeigenschaften.

## ECP\_FILE

Das Format der ECP-Datei muss dem CFOUR-Standard<sup>[124]</sup> entsprechen. In den ersten beiden Zeilen ist der Name und der Verwendungszweck des ECPs angegeben. In Zeile vier sind die Zahl der Elektronen, welche durch das ECP beschrieben werden (NCORE=5) und die maximale Drehimpulsquantenzahl (LMAX) angegeben. Die Parameter des ECPs für Gaußfunktionen<sup>[124,126,127]</sup>

$$U_i(r) = \sum_m c_m e^{\alpha_m r^2} r^{N_m}, \quad (3.6)$$

welche das ECP bilden, sind in den folgenden Zeilen in der Reihenfolge  $c_m$ ,  $N_m$  und  $\alpha_m$  angegeben. Momentan stehen zwei verschiedene ECPs zum Optimieren zur Verfügung. Ein kleines ECP<sup>[128,129]</sup> mit 6 zu optimierenden Parametern und ein großes ECP<sup>[128,130]</sup> mit 28 zu optimierenden Parametern. Die ECP-Dateien sind in Abbildung 3.10 dargestellt.

```

a) + ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4      NCORE = 5      LMAX = 3
5 f
6      0.0000000      2      1.0000000
7 s-f
8      33.1216380      2      6.4010520
9 p-f
10     -1.9862570      2      7.3077470
11 d-f
12     -9.4543180      2      5.9617960
13 *
14

b) + ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4      NCORE = 5      LMAX = 4
5 g
6      0.0000000      2      1.0000000
7 s-g
8      370.2360900      2      27.2422250
9      9.2146290      2      7.4679480
10 p-g
11     99.1350340      2      45.5679910
12     198.2714540      2      45.5175050
13     28.3881210      2      18.2457690
14     56.8034710      2      17.6510780
15 d-g
16    -17.9670850      2      52.7605980
17    -26.8155390      2      53.1366630
18     -1.3838930      2      14.5640790
19     -2.0726630      2      12.6318570
20      0.0803340      2      3.9230420
21      0.2385290      2      3.7387740
22 f-g
23     -0.7056800      2      11.4073780
24     -2.4759510      2      18.4290540
25 *
26

```

Abbildung 3.10: Die zwei ECP-Dateien, welche mit CFOUR optimiert werden können.



---

## 4 Simulationsdetails

Im folgenden Kapitel werden die Simulationdetails für alle in dieser Arbeit durchgeführten AIMD- und QM/MM-Simulationen angegeben. Zudem wird eine Übersicht aller verwendeten QM/MM-Systeme abgegeben und die Vorgehensweise zum Erstellen dieser Systeme erläutert. Alle Simulationen wurden mit dem in dieser Arbeit implementierten Interface für AIMD- und QM/MM-Simulationen durchgeführt, welches modifizierte Versionen von GROMACS (Version 2018) und CPMD (Version 4.3) mit einer modifizierten Entwicklerversion von CFOUR über das MiMiC-Framework verknüpft. Es wurden auf Wellenfunktionen basierende quantenchemische Methoden (HF, MP2, CCSD(T) und CAS-SCF(6,6)<sup>a</sup>) verwendet sowie plane-wave-DFT. Dabei wurden die Basissätze<sup>[131]</sup> cc-pVDZ und cc-pVTZ, bzw. die Funktionale BLYP<sup>[132,133]</sup> und PBE<sup>[134]</sup> verwendet.

### 4.1 Aufbau, Equilibrierung und Durchführung der Simulationen

Die für die AIMD- und QM/MM-Simulationen verwendeten Systeme wurden, wenn nicht anders angegeben, alle auf die selbe Weise erstellt und equilibriert. Das allgemeine Vorgehen der Equilibrierung und die Simulationsdetails werden im Folgenden erläutert.

#### 4.1.1 AIMD-Systeme

Für AIMD-Simulationen ist es notwendig, dass sich das System zum Start der Simulation im Gleichgewichtszustand befindet, da es ansonsten nicht möglich ist gezielt eine Temperatur einzustellen. Befindet sich das System nicht im Gleichgewichtszustand, beinhaltet das System potentielle Energie, welche während der Simulation in Form von kinetischer Energie frei wird und somit das System aufheizt. Daher wurde die Startgeometrie des Systems bei allen AIMD-Simulationen durch eine vorherige Geometrieoptimierung auf dem gleichen quantenchemischen Niveau berechnet, auf dem auch die AIMD-Simulation durchgeführt wurde. Schwingungsfrequenzen zum Vergleich wurden ebenfalls jeweils auf dem gleichen Niveau berechnet.

Für alle AIMD-Simulationen wurde der BO-MD-Ansatz mit einem Zeitschritt von 10 a.u. ( $\approx 0.24$  fs) verwendet. Die Simulationen wurden im NVE-Ensemble

---

<sup>a</sup> CAS-SCF(6,6) bedeutet, dass der aktive Raum 6 Orbitale umfasst, die mit 6 Elektronen besetzt sind.

ohne Temperaturkontrolle durchgeführt und als Integrator wurde der Velocity-Verlet-Algorithmus<sup>[59]</sup> verwendet.

Bei den auf Wellenfunktionen basierenden quantenchemischen Methoden wurde ein Konvergenzkriterium von  $10^{-9}$  a.u. für die Optimierung der Wellenfunktion verwendet. Im Falle der CCSD(T)-Rechnungen wurde ein Konvergenzkriterium von  $10^{-7}$  a.u. für die Lösung der CC- und Lambdagleichungen verwendet. Die Geometrieoptimierungen wurden mit einem Konvergenzkriterium von  $10^{-9}$  a.u. für das quadratische Mittel des Gradienten durchgeführt.

Für die pw-DFT-Simulationen wurde ein Konvergenzkriterium von  $10^{-5}$  a.u. für die Optimierung der Wellenfunktion, eine quadratische Box mit einer Seitenlänge von 30 a.u. und ein plane-wave-Cutoff von 100 Ry verwendet. Bei der Geometrieoptimierung wurde ein Konvergenzkriterium von  $10^{-7}$  a.u. für die Optimierung der Wellenfunktion und von  $10^{-5}$  a.u. für den größten Wert des Gradienten verwendet. Die Simulationen und Geometrieoptimierungen wurden mit den Austausch-Korrelations-Funktionalen BLYP oder PBE und mit norm-erhaltenden Pseudopotentialen von Troullier und Martins<sup>[107]</sup> durchgeführt. Dabei wurde das Tuckermann und Maryna Schema<sup>[135]</sup> zur Lösung der Poisson-Gleichung für isolierte Systeme verwendet.

### 4.1.2 QM/MM-Systeme

Die QM/MM-Systeme wurden über rein klassische MD-Simulationen erstellt und equilibriert. Dazu wurde das Molekül, welches sich später im QM-Bereich befinden sollte, in die Mitte einer quadratischen Simulationsbox platziert und die Box anschließend mit dem jeweiligen Lösungsmittel gefüllt. Die Struktur des Systems wurde zunächst über ein *steepest decent* Verfahren minimiert bis die maximale Kraft, welche auf ein Teilchen wirkt, weniger als  $1000 \text{ kJ} \cdot (\text{mol} \cdot \text{nm})^{-1}$  betrug. Anschließend wurde eine 100 ps Simulation mit Zeitschritten von 2 fs im NVT-Ensemble durchgeführt. Die Ausgangsgeschwindigkeiten wurden über eine Maxwell-Verteilung bei 300 K bestimmt. Dabei wurde die Temperatur über *velocity rescaling*<sup>[136]</sup> mit einer Kopplungskonstante von 0.1 ps kontrolliert. Als nächstes wurde für das System eine 100 ps Simulation mit Zeitschritten von 2 fs im NPT-Ensemble bei 1.0 Bar durchgeführt, wobei ein Parrinello-Rahman-Barostat<sup>[64]</sup> mit einer Kopplungskonstante von 2 ps verwendet wurde. Die Temperatur wurde wie in der NVT-Equilibrierung über *velocity rescaling*<sup>[136]</sup> bei 300 K gehalten. Als letzter Teil der Equilibrierung wurde eine 1 ns Simulation mit Zeitschritten von 2 fs wieder im NVT-Ensemble bei 300 K durchgeführt.

In allen Equilibrierungsschritten wurden periodische Randbedingungen und ein Cutoff-Radius von 1.0 nm für das Verlet-Schema<sup>[137]</sup> für die short-range und van der Waals Wechselwirkungen verwendet. Die Struktur der Wassermoleküle wurde über den SETTLE-Algorithmus<sup>[109]</sup> eingeschränkt, während für alle anderen Bindungen, welche H-Atome beinhalten, der LINCS-Algorithmus<sup>[110]</sup> verwendet wurde. Die long-range Wechselwirkungen wurden über die PME-Methode<sup>[78]</sup> berechnet und als Integrator wurde der leap-frog-Algorithmus<sup>[60]</sup> verwendet.

Für alle QM/MM-Simulationen wurde ebenfalls der BO-MD-Ansatz verwendet

und die Simulationen wurden im NVE-Ensemble ohne Temperaturkontrolle durchgeführt. Als Integrator wurde der Velocity-Verlet-Algorithmus<sup>[59]</sup> verwendet. Die elektrostatischen QM/MM-Wechselwirkungen wurden innerhalb der *Minimum Image Convention* berechnet und es wurden keine periodischen Randbedingungen verwendet. Für die Berechnung der MM-Wechselwirkungen wurden periodischen Randbedingungen verwendet und für die elektrostatischen short-range und van der Waals Wechselwirkungen wurde ein Cutoff-Radius von 1.5 nm angewandt. Die long-range Wechselwirkungen wurden über die PME-Methode berechnet.

Bei den auf Wellenfunktionen basierenden quantenchemischen Methoden wurde ein Konvergenzkriterium von  $10^{-9}$  a.u. für die Optimierung der Wellenfunktion verwendet. Im Falle der CCSD(T)-Rechnungen wurde ein Konvergenzkriterium von  $10^{-7}$  a.u. für die Lösung der CC- und Lambdagleichungen verwendet.

Für die pw-DFT-Simulationen wurde ein Konvergenzkriterium von  $10^{-5}$  a.u. für die Optimierung der Wellenfunktion, eine quadratische Box mit 30 a.u. Seitenlänge und einem plane-wave-Cutoff von 100 Ry verwendet. Die Simulationen wurden mit den Austausch-Korrelations-Funktional BLYP und mit norm-erhaltenden Pseudopotentialen von Trullier und Martins<sup>[107]</sup> durchgeführt. Dabei wurde das Tuckermann und Maryna Schema<sup>[135]</sup> zur Lösung der Poisson-Gleichung für isolierte Systeme verwendet.

## 4.2 Liste der QM/MM-Systeme

### 4.2.1 Wassersysteme (klein)

Zur Verifizierung des Interfaces (Kapitel 5) wurden zwei Systeme erstellt, welche aus jeweils einem QM- und 1011 MM-Wassermolekülen bestehen. Für ein System wurde das SPC/E<sup>[138]</sup> Wassermode, für das andere das TIP3P<sup>[139]</sup> Wassermode verwendet. Die Seitenlänge der Simulationsbox des SPC/E-Systems beträgt 3.116 nm und die des TIP3P-Systems beträgt 3.129 nm.

Für beide Systeme wurden jeweils QM/MM-Simulationen mit den quantenchemischen Methoden HF, MP2, CCSD(T) und CAS-SCF(6,6) mit der cc-pVTZ-Basis und pw-DFT mit dem BLYP Funktional durchgeführt. Die Simulationszeit betrug jeweils  $\approx 12.1$  ps (50000 Zeitschritte je 10 a.u.), wobei für die Auswertung nur die letzten 30000 Schritte verwendet wurden, da die ersten 20000 Schritte als Teil der Equilibrierung genutzt wurden. Für die Systeme wurden keine long-range-Wechselwirkungen verwendet.

Für die Berechnung von Schwingungsfrequenzen (Kapitel 8) aus einzelnen Snapshots einer QM/MM-Trajektorie wurde für das SPC/E-System eine Simulation für  $\approx 7.3$  ps (30000 Zeitschritte je 10 a.u.) auf HF/cc-pVTZ Niveau durchgeführt, welche mit den Endkoordinaten und Endgeschwindigkeiten der ersten HF/cc-pVTZ-Simulation gestartet wurde.

### 4.2.2 Wassersystem (groß)

Für die Untersuchung der long-range-Wechselwirkungen (Kapitel 6.1) wurde ein Wassersystem verwendet, welches aus einem QM- und 12000 MM-Wassermolekülen besteht. Im MM-Bereich wurde das SPC/E Wassermolekülmodell verwendet. Die Seitenlänge der Simulationsbox beträgt 7.119 nm. Die QM/MM-Simulationen wurden auf HF/cc-pVTZ Niveau unter Verwendung verschiedener Cutoff-Radien für die long-range-Wechselwirkungen durchgeführt. Die Simulationszeit betrug jeweils  $\approx 12.1$  ps (50000 Zeitschritte je 10 a.u.), wobei für die Auswertung nur die letzten 30000 Schritte verwendet wurden, da die ersten 20000 Schritte als Teil der Equilibrierung genutzt wurden.

### 4.2.3 Essigsäuresysteme

Es wurden zwei Essigsäuresysteme zur Überprüfung der Anwendbarkeit der optimierten ECPs (Kapitel 7) erstellt. Ein System verwendet ein *united-atom*-Kraftfeld für den MM-Bereich, das andere ein *all-atom*-Kraftfeld. Das verwendete Gromos54a7-Kraftfeld<sup>[140]</sup> und die initialen Strukturen für die Equilibrierung wurden über die Automated Topology Builder (ATB) Website<sup>[141-143]</sup> erhalten (ATB molid: 305617). Die Systeme bestehen jeweils aus 493 Essigsäuremolekülen, wobei sich ein Molekül im QM-Bereich befindet. Die Seitenlänge der Simulationsbox des *united-atom*-Systems beträgt 3.555 nm und die des *all-atom*-Systems beträgt 3.596 nm.

Es wurden QM/MM-Simulationen mit der HF-Methode und den Basissätzen cc-pVDZ und cc-pVTZ durchgeführt. Die Simulationszeit der Simulationen mit der cc-pVDZ-Basis betrug  $\approx 4.8$  ps (10000 Zeitschritte je 20 a.u.) und die der Simulationen mit der cc-pVTZ-Basis betrug  $\approx 0.97$  ps (2000 Zeitschritte je 20 a.u.). Für die Systeme wurden keine long-range-Wechselwirkungen verwendet.

### 4.2.4 Methanolsystem

Für die Berechnung von Schwingungsfrequenzen (Kapitel 8) aus einer QM/MM-Trajektorie wurde ein Methanolsystem erstellt, welches aus einem QM- und 1108 MM-Methanolsmolekülen besteht. Der MM-Bereich wird mit dem *united-atom*-Kraftfeld 'B3' von Walser et al.<sup>[144]</sup> beschrieben. Die Seitenlänge der Simulationsbox beträgt 4.198 nm.

Als Teil der Equilibrierung wurde eine MTS-QM/MM-Simulation für  $\approx 2.4$  ps (10000 Zeitschritte je 10 a.u.) auf MP2/cc-pVTZ/pw-DFT-D2/PBE Niveau mit einem MTS-Faktor von  $n = 10$  durchgeführt. Dabei wurde ein long-range Cutoff-Radius von 40 a.u. verwendet und die Temperatur wurde mit einem Nöse-Hoover Thermostat<sup>[62,63]</sup> bei 300 K gehalten. Die Simulation, welche zur Auswertung verwendet wurde, wurde anschließend auf dem gleichen Niveau durchgeführt. Die Simulationszeit betrug  $\approx 12.1$  ps (50000 Zeitschritte je 10 a.u.) und es wurde kein Thermostat verwendet (NVE-Ensemble).

### 4.2.5 Poly- $\beta$ -alaninsystem

Das in Kapitel 7 verwendete  $\beta$ -Alaninsystem wurde von Uribe et al.<sup>[145,146]</sup> erstellt und besteht aus einem Poly- $\beta$ -alanin, welches aus acht  $\beta$ -Alanin-Bausteinen besteht und 6354 Methanolkolekülen. Für den MM-Bereich wird das *united-atom* Gromos53a6-Kraftfeld<sup>[66]</sup> für das  $\beta$ -Alanin bzw. das *united-atom*-Kraftfeld 'B3'<sup>[144]</sup> für Methanol verwendet. Die elektrostatischen long-range Wechselwirkungen im MM-Bereich wurden über die Reaktion-Field-Methode<sup>[147,148]</sup> mit einem Cutoff-Radius von 1.2 nm und einer Dielektrizitätskonstante von 32.6<sup>[144]</sup> für Methanol berechnet. Die Seitenlänge der Simulationsbox beträgt 4.536 nm.

Es wurden QM/MM-Simulationen mit unterschiedlich großen QM-Bereichen auf HF/cc-pVDZ Niveau durchgeführt. Die Simulationszeit der Simulationen mit den beiden kleinen QM-Bereichen (4 bzw. 10 Atome im QM-Bereich) betrug  $\approx 4.8$  ps (10000 Zeitschritte je 20 a.u.). Die der Simulationen mit den beiden größeren QM-Bereichen (17 bzw. 23 Atome im QM-Bereich) betrug  $\approx 0.97$  ps (2000 Zeitschritte je 20 a.u.). Für alle Simulationen wurde ein long-range Cutoff-Radius von 50 a.u. verwendet.



---

# 5 Stabilität und Funktionalität der Implementierung

Um das implementierte Interface zwischen CFOUR, CPMD und MiMiC für AIMD- und QM/MM-Simulationen auf numerische Stabilität und Funktionalität zu testen, wurden AIMD-Simulationen an einem Wassermolekül und QM/MM-Simulationen an einem System aus einem QM-Wassermolekül umgeben von 1011 klassischen (SPC/E) Wassermolekülen durchgeführt. Dabei wurde zunächst untersucht, ob die Energie bei Simulationen im NVE-Ensemble über den zeitlichen Verlauf konstant bleibt. Dies ist eine unerlässliche Voraussetzung für die Durchführung numerisch stabiler Simulationen. Ein weiterer Aspekt ist die Funktionalität des implementierten QM/MM-Ansatzes. Dies bedeutet, dass durch QM/MM-Simulationen strukturelle und chemische Eigenschaften eines Systems korrekt beschrieben und im besten Fall experimentelle Ergebnisse reproduziert werden können. Um zu zeigen, dass die Solvatation bzw. die Polarisierung des QM-Systems durch das Lösungsmittel in QM/MM-Simulationen korrekt beschrieben wird, wurden das Dipolmoment des QM-Wassermoleküls und die Lösungsumgebung über die radiale Verteilungsfunktion<sup>[149]</sup> (RDF) untersucht. Dabei wurden die Ergebnisse jeweils mit experimentellen Werten und mit QM/MM-Simulationen auf pw-DFT Niveau, welche mit dem schon vorhandenen CPMD/MiMiC-Interface durchgeführt wurden, verglichen.

## 5.1 Energieerhaltung

Die Energieerhaltung während AIMD- und QM/MM-Simulationen ist eine Voraussetzung für die Durchführung numerisch und thermodynamisch stabiler Simulationen. Geringe Fluktuationen aufgrund des Diskretisierungsfehlers der verwendeten Zeitschritte und der thermischen Fluktuation sind hierbei normal.

Es wurden jeweils AIMD- und QM/MM-Simulationen mit den Methoden HF, MP2, CCSD(T) und CAS-SCF(6,6) mit der cc-pVTZ Basis<sup>[131]</sup> durchgeführt. Untersucht wurde jeweils die Gesamtenergie des Systems sowie die Abweichung der Energie  $\Delta E(t) = E(t) - \langle E \rangle$  pro Atom. In Abbildung 5.1 ist die Fluktuation der Energie pro Atom dargestellt. Im oberen Teil der Abbildung ist die Energie der AIMD-Simulationen zu sehen. Die Fluktuation der Energie pro Atom ist für alle verwendeten Methoden sehr klein. Die Fluktuationen für CAS-SCF(6,6) sind etwas größer als für die anderen Methoden, mit einer Standardabweichung von  $2.4 \cdot 10^{-8}$  a.u. sind sie dennoch sehr klein. Die Simulationen verlaufen sehr stabil, da bei keiner der Methoden ein Drift in der Energie zu erkennen ist.

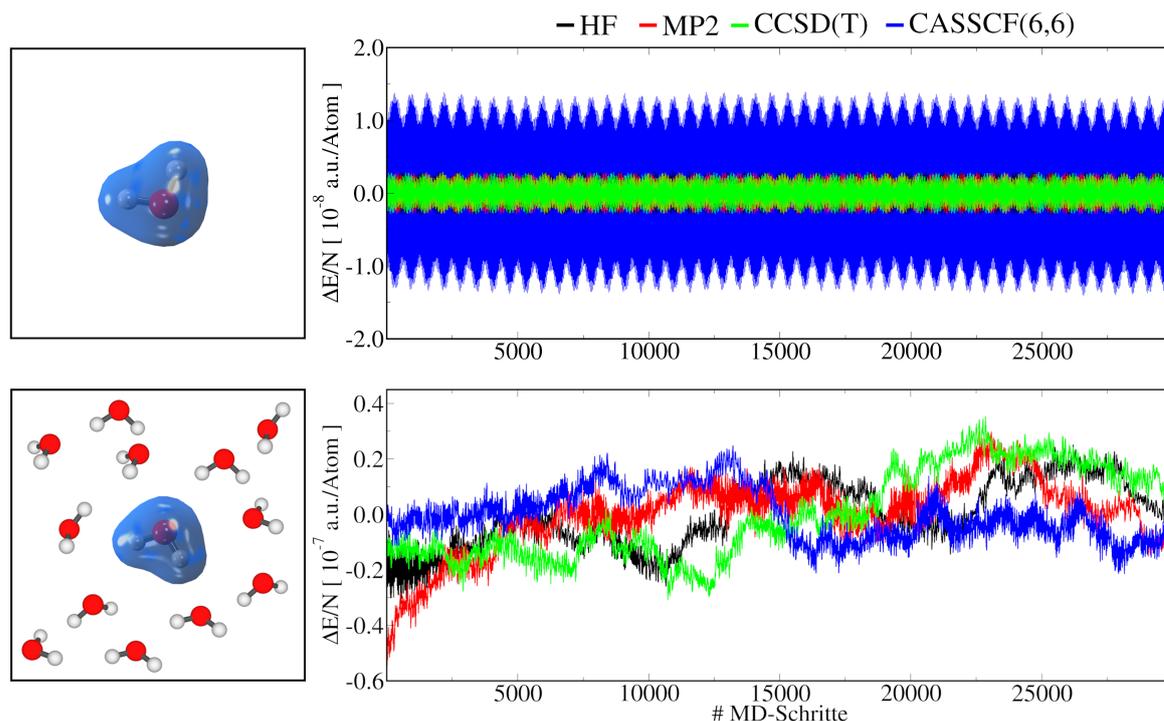


Abbildung 5.1: Fluktuation der Energie pro Atom für ein einzelnes Wassermolekül in einer AIMD-Simulation (oberer Abschnitt) bzw. für ein System aus 1012 Wassermolekülen in einer QM/MM-Simulation (unterer Abschnitt) im NVE-Ensemble. Die Simulationen wurden mit den Methoden HF, MP2, CCSD(T) und CAS-SCF(6,6) (jeweils cc-pVTZ Basis) für den QM-Bereich und dem SPC/E Wassermodell für den MM-Bereich durchgeführt.

Im unteren Teil der Abbildung 5.1 ist der Energieverlauf der QM/MM-Simulationen abgebildet. Die Fluktuation der Energie pro Atom ist für alle Methoden etwas größer als bei den AIMD-Simulationen. Neben den schnellen Fluktuationen sind auch kleine Drifts der Energie zu erkennen, welche sich im Verlauf der Simulation allerdings gegenseitig aufheben. Die Standardabweichungen der Energie sind mit  $3.7 \cdot 10^{-5}$  a.u. für HF,  $4.0 \cdot 10^{-5}$  a.u. für MP2,  $4.9 \cdot 10^{-5}$  a.u. für CCSD(T) und mit  $2.7 \cdot 10^{-5}$  a.u. für CAS-SCF(6,6) klein und in der typischen Größenordnung<sup>[5]</sup> für QM/MM-Simulationen. Das bedeutet, dass mit dem neuen Interface numerisch stabile AIMD- und QM/MM-Simulationen mit verschiedenen wellenfunktionsbasierten quantenchemischen Methoden durchgeführt werden können.

## 5.2 Polarisation des QM-Systems

Eine relevante chemische Eigenschaft, welche in QM/MM-Simulationen korrekt beschrieben werden sollte, ist die Polarisierung der elektronischen Struktur des QM-Systems durch das Lösungsmittel. Dazu wurde der Unterschied des Dipolmoments zwischen den AIMD- und QM/MM-Simulationen und der Einfluss von unterschiedlichen Wassermodellen (SPC/E und TIP3P) auf das Dipolmoment in QM/MM-Simulationen untersucht. Experimentelle und theoretische Untersuchungen in der Literatur zeigen, dass die Polarisierung der elektronischen Struktur in flüssigem Wasser zu einer Vergrößerung des Dipolmoments führt.<sup>[150–155]</sup>

Um die Polarisierung der elektronischen Struktur des QM-Systems zu untersuchen, wurden die mittleren Dipolmomente, welche aus AIMD- und QM/MM-Simulationen sowie aus einzelnen quantenchemischen Rechnungen auf den selben Niveaus erhalten wurden, miteinander verglichen. Zusätzlich wurden AIMD- und QM/MM-Simulationen auf pw-DFT/BLYP-Niveau durchgeführt. Die erhaltenen Ergebnisse sind in Tabelle 5.1 gegenübergestellt und werden mit experimentellen Werten verglichen.

Den Erwartungen entsprechend ergeben die AIMD-Simulationen die gleichen Werte wie die einzelnen quantenchemischen Rechnungen. Nur bei den Methoden CAS-SCF(6,6) und pw-DFT treten kleine Abweichungen auf. Die Abweichungen zu den experimentellen Werten sind auf die Ungenauigkeit der verwendeten Metho-

Tabelle 5.1: Dipolmoment in Debye eines QM-Wassermoleküls. Die Dipolmomente wurden aus einzelnen quantenchemischen Rechnungen, AIMD- und QM/MM-Simulationen erhalten. Als QM-Methode wurden HF, MP2, CCSD(T) und CAS-SCF(6,6) (jeweils cc-pVTZ Basis) und pw-DFT/BLYP verwendet. Der MM-Bereich wurde mit den Wassermodellen SPC/E bzw. TIP3P beschrieben.

Methode	HF	MP2	CCSD(T)
QM	1.988	1.936	1.916
AIMD	1.988 ± 0.004	1.936 ± 0.004	1.916 ± 0.004
QM/MM (SPC/E)	2.694 ± 0.160	2.742 ± 0.145	2.765 ± 0.166
QM/MM (TIP3P)	2.732 ± 0.146	2.692 ± 0.166	2.632 ± 0.168
Methode	CAS-SCF(6,6)	pw-DFT	Exp.
QM	1.935	1.814	1.847 ± 0.001 <sup>a</sup>
AIMD	1.909 ± 0.031	1.846 ± 0.005	
QM/MM (SPC/E)	2.624 ± 0.177	2.720 ± 0.169	2.9 ± 0.660 <sup>b</sup>
QM/MM (TIP3P)	2.755 ± 0.157	2.699 ± 0.200	

<sup>a</sup> in Gasphase<sup>[150]</sup>

<sup>b</sup> für flüssiges Wasser<sup>[151]</sup>

den zurückzuführen und werden von HF zu CCSD(T) aufgrund der zusätzlichen Berücksichtigung der Elektronenkorrelation kleiner. Die Übereinstimmung mit den einzelnen quantenchemischen Rechnungen und den experimentellen Werten zeigt, dass die AIMD-Simulationen Moleküle in der Gasphase korrekt beschreiben können.

Die Dipolmomente, welche über die QM/MM-Simulationen erhalten wurden, sind alle deutlich größer als die entsprechenden AIMD-Werte, was auf die Polarisierung durch die Lösungsumgebung zurückzuführen ist. Zwischen den beiden klassischen Wassermodellen sind keine großen Unterschiede zu erkennen. Beide Wassermodelle polarisieren die elektronische Struktur des QM-Bereichs gleichermaßen. Die erhaltenen Werte sind zwar etwas kleiner als der experimentelle Wert, liegen aber innerhalb der experimentellen Standardabweichung und stimmen mit den Werten anderer theoretischer Studien überein.<sup>[151–155]</sup>

Ein Grund für die etwas zu schwache Polarisierung des QM-Bereichs ist, dass die elektronische Struktur nur von den klassischen Wassermolekülen polarisiert wird, welche selbst ein festes und vergleichsweise kleines Dipolmoment besitzen (2.351 D für SPC/E<sup>[138]</sup> und 2.344 D für TIP3P<sup>[139]</sup>). Die verwendete QM/MM-Implementierung berücksichtigt nur die einseitige Polarisierung des QM-Bereichs durch die Atome des MM-Systems und keine gegenseitige Polarisierung. Bessere Resultate würde man beispielsweise erhalten, wenn eine gegenseitige Polarisierung durch die Verwendung von polarisierbaren Kraftfeldern wie MB-pol<sup>[156]</sup> berücksichtigt würde. Ebenso könnte man den QM-Bereich vergrößern und mehrere Wassermoleküle quantenchemisch beschreiben, sodass sich diese gegenseitig polarisieren können.

Die deutliche Vergrößerung des Dipolmoments des Wassermoleküls in Lösung verglichen zum Dipolmoment im Vakuum zeigt jedoch, dass die QM/MM-Implementierung die starke Wechselwirkung zwischen dem QM-Wassermolekül und der Lösungsumgebung korrekt beschreibt.

## 5.3 Struktur der Lösungsumgebung

Neben der Beschreibung der Polarisierung des QM-Bereichs durch die Wechselwirkung mit dem Lösungsmittel ist auch die korrekte Beschreibung der Struktur der Lösungsumgebung relevant. Um die Struktur der Solvationshülle des QM-Wassermoleküls zu untersuchen, wurde die radiale Verteilungsfunktion bezüglich des QM-Sauerstoffs und der klassischen Sauerstoffmoleküle des Lösungsmittels untersucht. Die RDFs der QM/MM-Simulationen mit verschiedenen QM-Methoden, einer rein klassischen MD-Simulation mit dem SPC/E Wassermodell sowie die experimentelle RDF sind in Abbildung 5.2 dargestellt. Die Positionen der ersten beiden Maxima der RDFs sind in Tabelle 5.2 aufgelistet.

Die Position und die Form des ersten Maximums stimmt bei allen QM-Methoden gut mit den experimentellen Werten<sup>[149]</sup> überein. Die Form und die Position des zweiten Maximums variieren für die verschiedenen QM-Methoden und die Überein-

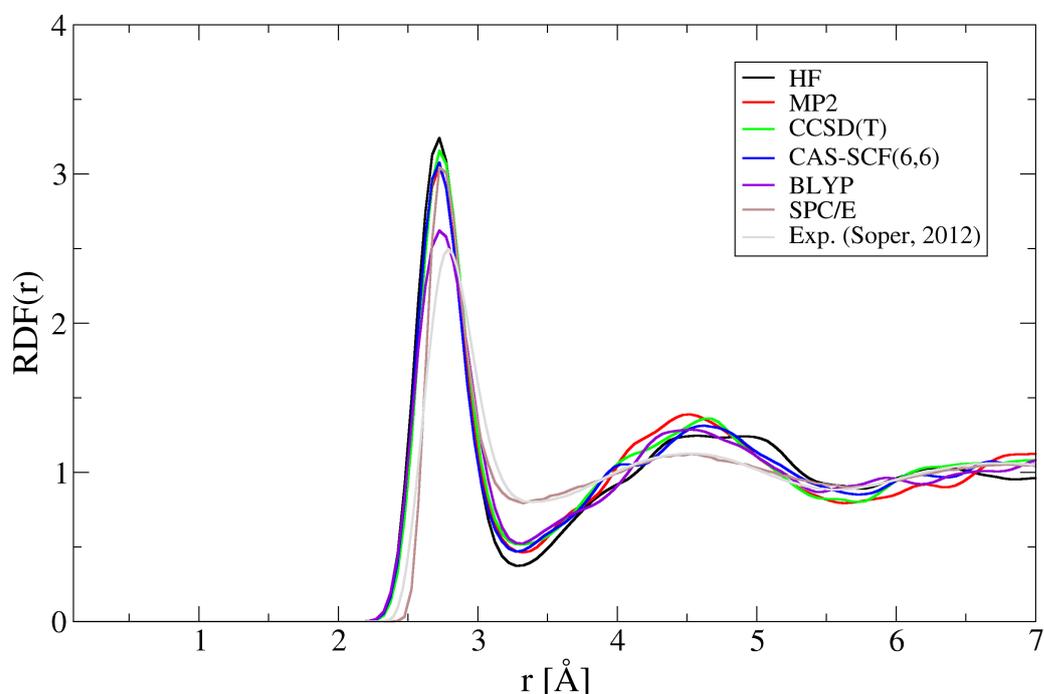


Abbildung 5.2: Die radiale Verteilungsfunktion zwischen dem QM-Sauerstoff und den klassischen Sauerstoffatomen erhalten aus QM/MM-Simulationen bei 300 K im NVE-Ensemble. Als QM-Methode wurden HF, MP2, CCSD(T) und CAS-SCF(6,6) (jeweils cc-pVTZ Basis) und pwDFT/BLYP verwendet. Der MM-Bereich wurde mit dem SPC/E Wassermodell beschrieben. Die RDF aus einer rein klassischen MD-Simulation mit dem SPC/E Wassermodell und experimentelle Werte<sup>[149]</sup> sind zum Vergleich angegeben.

Tabelle 5.2: Die Positionen in Å für die ersten beiden Maxima der radialen Verteilungsfunktion bezüglich des QM-Sauerstoffs und der klassischen Sauerstoffmoleküle bei 300 K, welche aus QM/MM-Simulationen und einer rein klassischen MD-Simulation im NVE-Ensemble erhalten wurden. Als QM-Methode wurden HF, MP2, CCSD(T) und CAS-SCF(6,6) (jeweils cc-pVTZ Basis) und pw-DFT/BLYP verwendet. Der MM-Bereich wurde mit dem SPC/E Wassermolekülmodell beschrieben.

Maximum	HF	MP2	CCSD(T)	CAS-SCF(6,6)	pw-DFT	SPC/E	Exp. <sup>[149]</sup>
1. $r_{\text{OO}}$	2.75	2.74	2.77	2.76	2.76	2.82	2.8
2. $r_{\text{OO}}$	4.72	4.65	4.67	4.50	4.56	4.55	4.5

stimmung mit dem Experiment ist schlechter als für das erste Maximum. Dies kann daran liegen, dass das zweite Maximum deutlich flacher ist als das erste und die RDFs der QM/MM-Simulationen aufgrund der kurzen Trajektorien relativ grob sind.

Verglichen mit dem Experiment und den reinen SCP/E Werten fällt auf, dass die QM/MM-Simulationen die erste Solvatationsschale etwas überstrukturiert beschreiben. Dies ist an einem höheren ersten Maximum und einem niedrigeren ersten Minimum zu erkennen und ist in Übereinstimmung mit theoretischen AIMD-Studien an flüssigem Wasser.<sup>[157–159]</sup>

Insgesamt zeigt der Vergleich mit dem Experiment, dass die QM/MM-Implementierung die Lösungsumgebung des QM-Wassermoleküls qualitativ korrekt beschreibt.

## 5.4 Zusammenfassung

In diesem Kapitel wurde die numerische Stabilität und die Funktionalität des implementierten Interfaces für AIMD- und QM/MM-Simulationen durch Simulationen an einem Wassersystem untersucht. Dabei wurden drei verschiedene Aspekte diskutiert. Als erstes wurde gezeigt, dass AIMD- und QM/MM-Simulationen unter Erhalt der Energie und ohne große Fluktuationen durchgeführt werden können. Dadurch ist sichergestellt, dass die Simulationen numerisch und thermodynamisch stabil verlaufen.

Zudem wurde die Beschreibung der Polarisierung des QM-Systems durch die klassische Lösungsmittelumgebung untersucht. Dabei wurde gezeigt, dass die erhaltenen Werte für das Dipolmoment von gasförmigem und flüssigem Wasser mit den experimentellen Werten übereinstimmen und die durch das Lösungsmittel induzierte Vergrößerung des Dipolmoments für flüssiges Wasser korrekt beschrieben wird.

Als letztes wurde auch die Struktur des Lösungsmittels anhand der radialen Verteilungsfunktion untersucht. Auch hier wurde gezeigt, dass die experimentellen Werte im Rahmen der verwendeten Methoden korrekt reproduziert werden können.

Anhand dieser Parameter konnte gezeigt werden, dass das implementierte Interface für AIMD- und QM/MM-Simulationen numerisch Stabilität garantiert und dazu in der Lage ist chemische und strukturelle Eigenschaften von Solut-Solvens-Systemen korrekt zu beschreiben.



---

## 6 Beschleunigung der Simulationen

Die Verwendung von hochgenauen quantenchemischen Methoden, wie der Coupled-Cluster-Theorie, in AIMD- oder QM/MM-Simulationen ist im Vergleich zu DFT-Methoden sehr aufwändig und es ist im Prinzip nicht möglich mit diesen QM-Methoden Simulationen über mehrere hundert Pikosekunden oder sogar Nanosekunden durchzuführen. Daher ist es wichtig Wege zu finden die Effizienz der Simulationen zu verbessern, um so die benötigte Rechenzeit zu verringern. Dazu gibt es verschiedene Ansätze, welche in dieser Implementierung umgesetzt wurden und im folgenden Abschnitt diskutiert werden.

Eine Möglichkeit ist es, die Rechenzeit der QM-Rechnung bzw. der Berechnung der QM/MM-Wechselwirkungen zu reduzieren. Dazu wurde eine hierarchische Berechnung der QM/MM-Wechselwirkungen mit verschiedenen Bereichen (short-range und long-range Region) implementiert. Ein weiterer Ansatz ist die Reduktion der Anzahl der kostenintensiven QM-Rechnungen. Durch das Interface zwischen CFOUR und CPMD ist es möglich, das in CPMD enthaltene *multiple time step* (MTS) Schema zu verwenden.<sup>[117]</sup>

### 6.1 Long-Range Wechselwirkungen

Bei dem ersten Ansatz zur Reduzierung der Zeit zur Berechnung der elektrostatischen QM/MM-Wechselwirkungen werden die Atome des MM-Bereichs über einen Cutoff-Radius in eine short-range und eine long-range Region eingeteilt. Die elektrostatischen QM/MM-Wechselwirkungen in der short-range Region werden explizit berechnet, während die Wechselwirkungen im long-range Bereich über eine Multipolentwicklung berechnet werden (siehe Abschnitt 2.2.4).

#### 6.1.1 Untersuchung der Konvergenz an Snapshot-Rechnungen

Die Effektivität der Reduzierung der Rechenzeit durch die Verwendung von long-range-Wechselwirkungen hängt von der Anzahl an MM-Atomen in der short-range Region ab. Allerdings nimmt die Genauigkeit der Multipolentwicklung ab, je weniger MM-Atome sich in der short-range Region befinden, da die long-range Atome dann zunehmend näher am QM-Bereich liegen. Um die Effektivität der Implementierung und die Abhängigkeit der Genauigkeit bzw. der Reduzierung der Rechenzeit von der

Anzahl an MM-Atomen in der short-range Region zu untersuchen, wurden zunächst einzelne QM-Rechnungen auf HF/cc-pVTZ Niveau mit verschiedenen Cutoff-Radien durchgeführt. Dazu wurde ein Snapshot einer QM/MM-Simulation eines Systems aus einem QM-Wassermolekül, welches von 12000 klassisch beschriebenen (SPC/E) Wassermolekülen umgeben ist, verwendet. Mit insgesamt 36003 Atomen handelt es sich hierbei um eine mittlere Systemgröße, z.B. im Vergleich zu Studien von Enzymreaktionen.<sup>[5]</sup> Es wurden die Energie (QM- und QM/MM-Energie), das Dipolmoment und die Kraft auf verschiedene Atome untersucht, da dies alle relevante Parameter zur stabilen Durchführung von MD-Simulationen sind bzw. charakteristische Eigenschaften des Moleküls. Das Konvergenzkriterium für die Lösung der Hartree-Fock Gleichungen wurde auf  $10^{-12}$  a.u. gesetzt, obwohl für QM/MM-Simulationen  $10^{-7}$  bis  $10^{-9}$  a.u. ausreichend sind. Damit wurde sichergestellt, dass die jeweils auftretenden Abweichungen durch die Ungenauigkeit der Multipolentwicklung begründet sind und nicht durch numerische Ungenauigkeiten bei der Lösung der Hartree-Fock Gleichungen.

In Abbildung 6.1 sind die Ergebnisse der Rechnungen und die jeweilige Abweichung von den Referenzwerten einer Rechnung, in der sich alle MM-Atome in der short-range Region befinden, dargestellt. Die Abweichungen sind bereits für eine sehr kleine Anzahl an short-range Atomen relativ klein und nehmen mit steigender Zahl an MM-Atomen in der short-range Region immer weiter ab. Für die Energie wird bereits bei ungefähr 10% der MM-Atome in der short-range Region eine Genauigkeit von  $10^{-9}$  a.u. erreicht, welche bei einer weiteren Zunahme an MM-Atomen in den short-range Bereich nur langsam verbessert wird. Eine Genauigkeit von  $10^{-9}$  a.u. ist für eine QM/MM-Simulation ausreichend, da der Fehler im Vergleich zum Diskretisierungsfehler der verwendeten Zeitschritte und der thermischen Fluktuation zu vernachlässigen ist (siehe Abschnitt 5.1). Vollständige Konvergenz wird nur erreicht, wenn sich nahezu alle MM-Atome in der short-range Region befinden.

In Abbildung 6.1 b wird der mittlere, absolute Fehler der Kraft auf vier verschiedene Atomgruppen dargestellt. Dies sind die drei QM-Atome, die drei short-range MM-Atome, welche sich am nächsten am Massenmittelpunkt des QM-Bereichs befinden, die drei long-range MM-Atome, welche sich am nächsten am Massenmittelpunkt befinden und die drei long-range MM-Atome, welche am weitesten vom Massenmittelpunkt entfernt sind. Abgesehen von den am weitesten entfernten long-range MM-Atomen, bei denen die Kraft schon bei sehr wenigen short-range Atomen nahezu exakt ist, ist der gleiche Trend wie für die Energie zu erkennen. Allerdings wird eine Genauigkeit von  $10^{-9}$  a.u. erst bei ca. 20% der MM-Atome in der short-range Region erreicht. Beim Dipolmoment wird bei 10-20% der MM-Atome in der short-range Region eine Genauigkeit von ca.  $10^{-8}$  Debye erreicht.

Die Genauigkeit, die bei 10% der MM-Atome in der short-range Region erreicht wird, ist für QM/MM-Simulationen ausreichend, da in der Regel ein Konvergenzkriterium von  $10^{-7}$  bis  $10^{-9}$  a.u. verwendet wird. Dies bedeutet, dass die Anzahl an MM-Atomen in der short-range Region, welche für eine numerisch stabile QM/MM-Simulation mit der gleichen Genauigkeit wie eine Simulation mit allen MM-Atomen in der short-range Region benötigt wird, für dieses System bei ca. 10% bzw. 3600

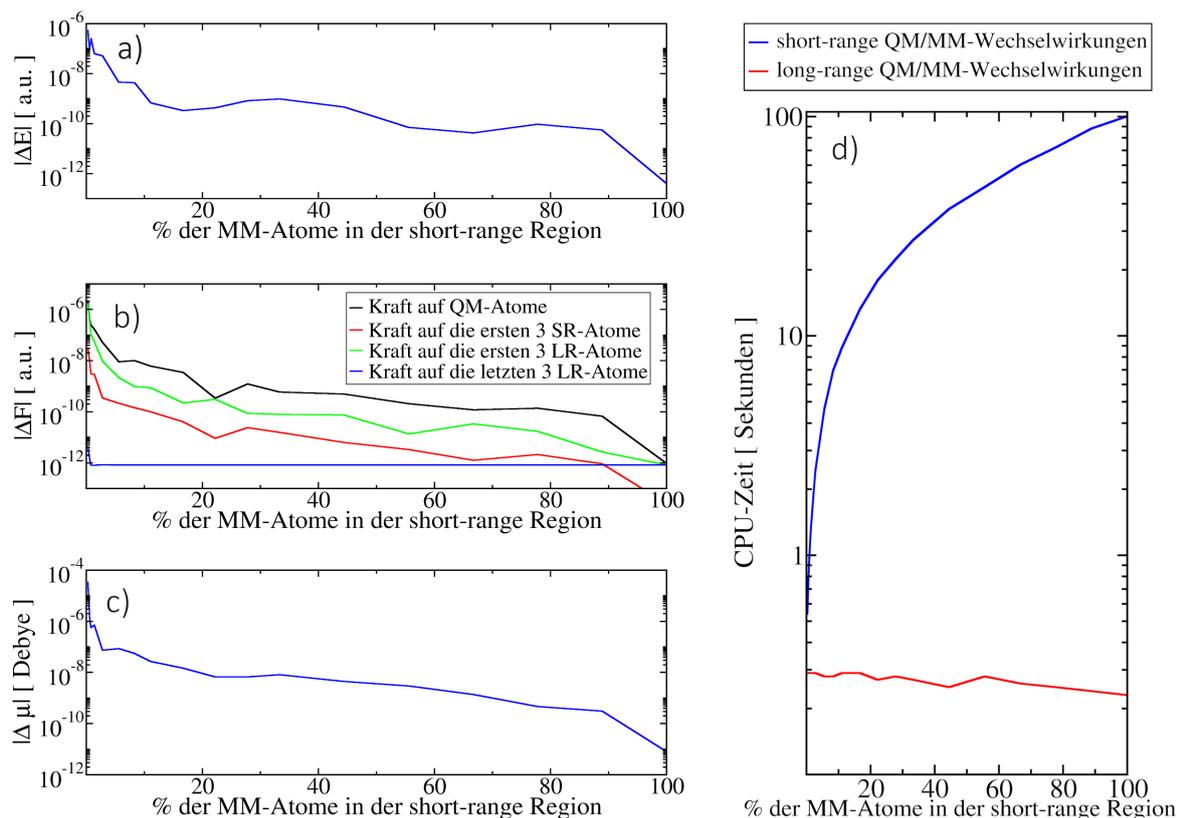


Abbildung 6.1: Konvergenz der Energie (a), der Kräfte (b) und des Dipolmoments (c) sowie die absolute CPU-Zeit (d) zur Berechnung der elektrostatischen short- und long-range Wechselwirkungen in Abhängigkeit der Anzahl an MM-Atomen in der short-range Region. Die verschiedenen Atomgruppen, auf die die Kräfte wirken, sind im Text näher erläutert. Alle Rechnungen wurden für ein Snapshot eines Systems aus 12001 Wassermolekülen auf HF/cc-pVTZ Niveau durchgeführt. Als Referenz dienen die jeweiligen Werte einer Rechnung, in der sich alle MM-Atome in der short-range Region befinden. Alle Rechnungen wurden auf einem Intel Xeon X5675 CPU mit 3.07 GHz durchgeführt.

Atomen liegt.

Die Zeitersparnis für die Berechnung der QM/MM-Wechselwirkungen ist in Abbildung 6.1 d dargestellt. Während die benötigte Zeit für die Berechnung der short-range Wechselwirkungen linear mit der Anzahl an short-range Atomen zunimmt (siehe Gleichung 2.13), bleibt die benötigte Zeit für die Berechnung der long-range Wechselwirkungen nahezu konstant. Im Vergleich mit der benötigten Zeit für die Berechnung der short-range Wechselwirkungen ist die benötigte Zeit für die long-range Wechselwirkungen vernachlässigbar klein. Die Berechnung der QM/MM-Wechselwirkungen mit allen MM-Atomen in der short-range Region benötigt 102 s. Wenn sich nur ca. 11% (4000) der MM-Atome in der short-range Region befinden, benötigt die Berech-

nung der short-range Wechselwirkungen 8.9s und die Berechnung der long-range Wechselwirkungen nur 0.3s. Dies bedeutet eine Reduzierung der benötigten Zeit um ca. 90%.

### 6.1.2 QM/MM-Simulationen mit long-range-Wechselwirkungen

Um die numerische Stabilität der Simulationen und die Genauigkeit der erhaltenen Eigenschaften zu untersuchen, wurden für das Wassersystem aus 12001 Wassermolekülen QM/MM-Simulationen mit verschiedenen Cutoff-Radien durchgeführt. Die Energiefluktuationen der Simulationen sind in Abbildung 6.2 dargestellt. Außer bei der Simulation mit einem Cutoff-Radius von nur 5 a.u. (ca. 2.65 Å), ist bei keiner der Simulationen ein Drift der Energie zu erkennen. Ebenso sind die Fluktuationen der Energie pro Atom für alle Cutoff-Radien (außer für den Cutoff-Radius von 5 a.u.) mit  $3.0 \cdot 10^{-4}$  bis  $7.0 \cdot 10^{-5}$  a.u. in der gleichen Größenordnung wie bei der Simulation mit allen MM-Atomen in der short-range Region ( $1.1 \cdot 10^{-4}$  a.u.). Dies zeigt, dass selbst bei der Wahl eines kleinen Cutoff-Radius von 7 a.u., was lediglich ca. 35 bis 50 MM-Atome in der short-range Region entspricht, eine numerisch stabile Simulation möglich ist. Um ca. 10% short-range Atome zu erhalten, benötigt man einen Cutoff-Radius von 35 bis 40 a.u. (ca. 20 Å).

Die gemittelten Dipolmomente aus den Simulationen und die jeweiligen Standardabweichungen sind in Tabelle 6.1 angegeben. Alle Dipolmomente der Simulationen mit Cutoff-Radius liegen im selben Bereich und innerhalb der Standardabweichung der Simulation mit allen MM-Atomen in der short-range Region. Das bedeutet, dass das Dipolmoment ausreichend gut reproduziert wird. Die Dipolmomente für die kleineren Cutoff-Radien weichen etwas stärker von dem Referenzwert ab als die Werte für die größeren Cutoff-Radien. Ab einem Cutoff-Radius von 30 a.u. liegt die Abweichung unterhalb von 0.04 Debye. Die aus den Snapshot-Rechnungen vorhergesagte

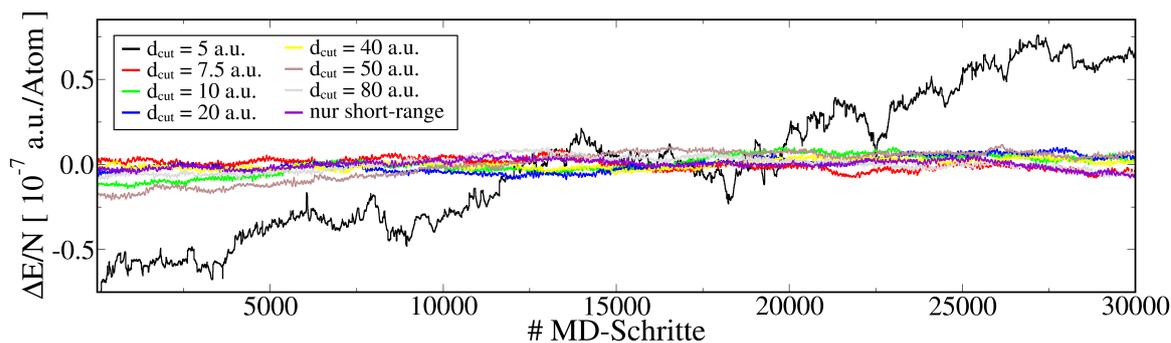


Abbildung 6.2: Energiefluktuation pro Atom für die verschiedenen Cutoff-Radien zur Benutzung der long-range Wechselwirkungen bei der Simulation von einem QM-Wasser (HF/cc-pVTZ) umgeben von 12000 klassischen Wassermolekülen (SPC/E) im NVE Ensemble.

Tabelle 6.1: Das mittlere Dipolmoment und die entsprechende Standardabweichung in Debye des QM-Wassermoleküls für verschiedene Cutoff-Radien zur Benutzung der long-range Wechselwirkungen bei der Simulation von einem QM-Wasser (HF/cc-pVTZ) umgeben von 12000 klassischen Wassermolekülen (SPC/E) im NVE Ensemble.

$r_{\text{cut}}$	5	7.5	10	20	30	40	50	80	nur SR
$\bar{\mu}$	2.811	2.639	2.753	2.637	2.748	2.692	2.706	2.719	2.711
$\sigma$	0.147	0.150	0.150	0.144	0.181	0.185	0.154	0.171	0.178

Genauigkeit des Dipolmoments von  $\leq 10^{-7}$  D liegt zwar deutlich darunter, allerdings ist die Standardabweichung mit  $\pm 0.178$  D relativ groß und die Simulationsdauer für die Mittelung des Dipolmoments relativ kurz (ca. 7.3 ps). Daher ist es anzunehmen, dass die Abweichung der Dipolmomente der Simulationen mit Cutoff-Radius vom Referenzwert bei genügend langer Simulationsdauer kleiner wird.

Die Snapshot-Rechnungen und die QM/MM-Simulationen zeigen, dass die Verwendung von long-range Wechselwirkungen die benötigte Zeit zur Berechnung der elektrostatischen QM/MM-Wechselwirkungen drastisch reduzieren kann, ohne dass die Genauigkeit der erhaltenen Ergebnisse dabei abnimmt.

## 6.2 Multiple Time Step Algorithmus

Die benötigte Zeit zur Berechnung der elektrostatischen QM/MM-Wechselwirkungen kann durch die Verwendung der long-range Wechselwirkungen reduziert werden. Die eigentliche QM-Rechnung wird dadurch allerdings nicht beschleunigt. Daher ist eine weitere Möglichkeit zur Reduzierung der Simulationszeit, die Anzahl an nötigen QM-Rechnungen auf hochgenauem Niveau zu reduzieren. Die Verwendung des in CPMD implementierten MTS Algorithmus<sup>[117]</sup> erlaubt es, verschiedene quantenchemische Methoden innerhalb einer Simulation zu kombinieren. Dabei wird die genauere, kostenintensivere Methode (hier CCSD(T)/cc-pVTZ) in den Korrekturschritten verwendet, während die ungenauere aber kostengünstigere Methode (hier pw-DFT/PBE) in den Referenzschritten verwendet wird. Dadurch ist es möglich, bei deutlich reduzierter Rechendauer die gleichen hochgenauen Ergebnisse zu erhalten, wie bei einer Simulation, die nur mit der genaueren Methode durchgeführt wird. Zur Demonstration der Methode wurden AIMD-Simulationen an Fluorwasserstoff durchgeführt und die Schwingungsfrequenz dieses Moleküls bestimmt. Da Fluorwasserstoff nur eine Schwingungsmode besitzt, kann die Qualität der MTS-Simulationen an dieser Schwingungsmode ohne andere, störende Effekte untersucht werden. Um die erhaltenen Werte der MTS-Simulationen bewerten zu können, wurden jeweils AIMD-Simulationen auf reinem CCSD(T)/cc-pVTZ bzw. pw-DFT/PBE Niveau durchgeführt. Die Simulationsdauer betrug jeweils ca. 12 ps (50.000 Schritte mit je 10 a.u.). Anschließend wurden Simulationen auf pw-DFT/PBE Niveau durchgeführt, bei denen in jedem  $n$ -ten Zeitschritt eine CCSD(T)/cc-pVTZ Rechnung als Korrektur durchgeführt wurde. Dabei wurden unterschiedliche MTS-Faktoren von  $n = 2, 4, 6, 10, 15$  und  $20$  verwendet. Um zu demonstrieren, dass es ohne Referenzschritte nicht möglich ist, die Anzahl an CCSD(T)-Rechnungen zu reduzieren, wurde auch eine AIMD-Simulation mit 10.000 Zeitschritten mit je 50 a.u. auf CCSD(T)/cc-pVTZ Niveau durchgeführt.

Die Fluktuation der Energie für die MTS-Simulationen ist in Abbildung 6.3 dargestellt. Die Fluktuationen sind klein und es ist bei keinem der MTS-Faktoren ein Drift der Energie zu erkennen. Mit steigendem MTS-Faktor werden die Fluktuationen größer, da die Simulationen mit dem auf CCSD(T)/cc-pVTZ Niveau berechnetem Gleichgewichtsabstand gestartet wurden und sich das System durch die Verwendung von pw-DFT/PBE nicht mehr im Gleichgewicht befindet und daher potentielle Energie frei wird. In den vergrößerten Ausschnitten sind diskrete Energiesprünge alle  $n$  Schritte zu erkennen, die durch die CCSD(T)-Korrekturen entstehen.

Die aus den Simulationen erhaltenen Schwingungsspektren sind in Abbildung 6.4 dargestellt. Die Schwingungsfrequenzen der reinen CCSD(T)/cc-pVTZ bzw. pw-DFT/PBE Simulationen unterscheiden sich, wie erwartet, deutlich. Mit Ausnahme der Simulation mit dem größten MTS-Faktor  $n = 20$  reproduzieren hingegen alle MTS-Simulationen die reine CCSD(T)/cc-pVTZ Schwingungsfrequenz gut. Die Abweichungen liegen alle im Rahmen von  $\pm 4 \text{ cm}^{-1}$ . Die MTS-Simulation mit dem größten MTS-Faktor und die CCSD(T)/cc-pVTZ Simulation mit dem Zeitschritt von 50 a.u. ergeben falsche Ergebnisse, da der Zeitschritt zwischen den einzel-

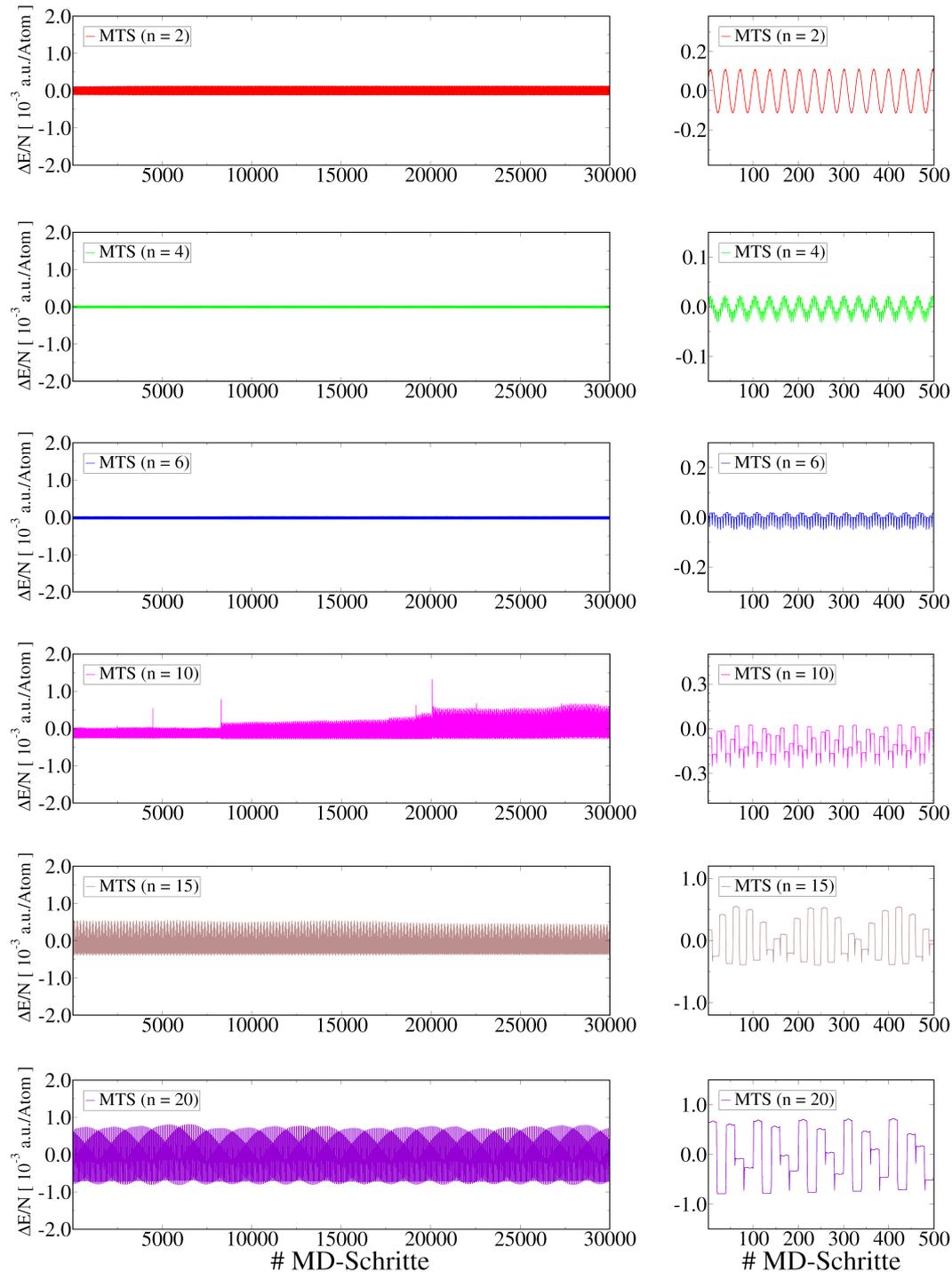


Abbildung 6.3: Fluktuation der Gesamtenergie pro Atom für Fluorwasserstoff im Vakuum bei AIMD-Simulationen auf dem kombinierten CCSD(T)/pw-DFT Niveau mit CCSD(T)/cc-pVTZ Korrekturrechnungen alle  $n = 2, 4, 6, 10, 15$  und 20 Zeitschritte. Auf der linken Seite der Abbildung ist die Energie für die gesamte Simulationsdauer dargestellt. Im rechten Teil sind nur die ersten 500 MD-Schritte abgebildet, wodurch die verschiedenen Fluktuationismuster zu erkennen sind.

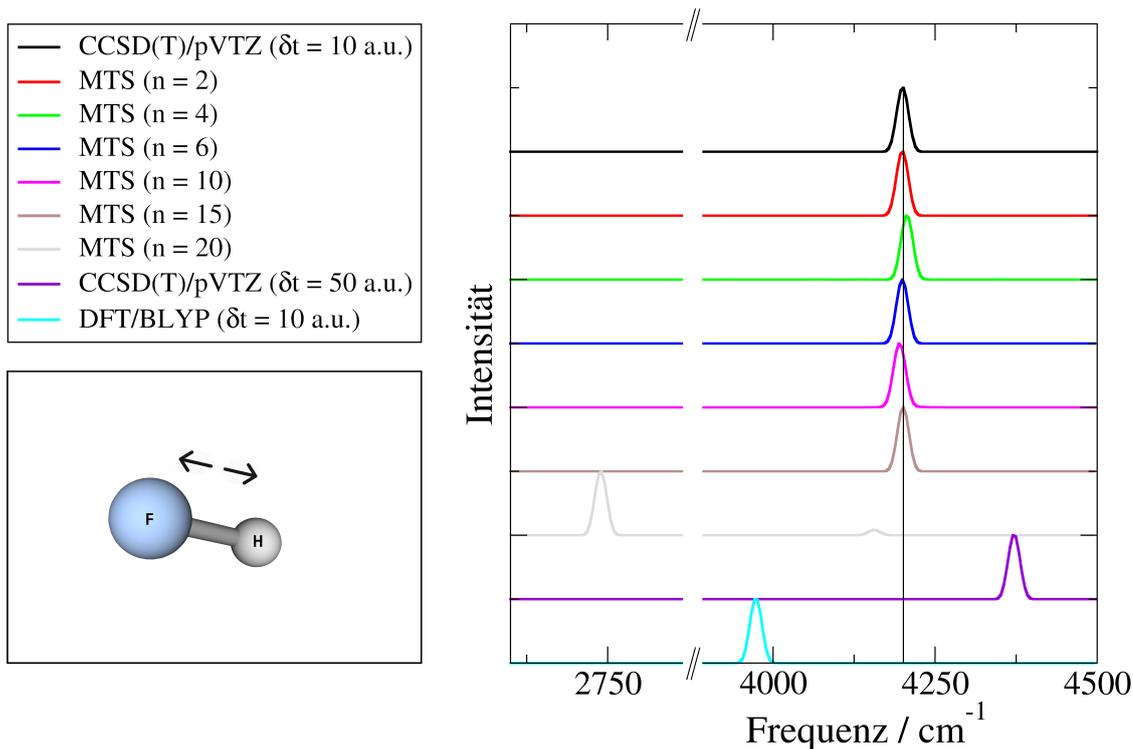


Abbildung 6.4: Schwingungsspektrum von Fluorwasserstoff im Vakuum berechnet aus AIMD-Simulationen auf dem kombinierten CCSD(T)/pw-DFT Niveau mit CCSD(T)/cc-pVTZ Korrekturrechnungen alle  $n = 2, 4, 6, 10, 15$  und  $20$  Zeitschritte. Reine CCSD(T)/cc-pVTZ- und pw-DFT/PBE-Spektren sind zum Vergleich angegeben.

nen CCSD(T)-Rechnungen zu groß ist, um die relativ schnelle Schwingung des Fluorwasserstoff-Moleküls richtig zu beschreiben. Auch die mittlere CCSD(T)/cc-pVTZ-Bindungslänge des Fluorwasserstoffs von  $0.916 \pm 0.002 \text{ \AA}$  wird von allem MTS-Simulationen gut reproduziert.

Um genauer zu untersuchen, warum die MTS-Simulationen die CCSD(T)-Ergebnisse so gut reproduzieren können, bzw. warum die Simulation mit dem MTS-Faktor  $n = 20$  es nicht kann, wurde der zeitliche Verlauf der Bindungslänge während der MTS-Simulationen genauer untersucht. Abbildung 6.5 zeigt die Oszillation der Bindungslänge für die reinen CCSD(T)/cc-pVTZ bzw. pw-DFT/PBE Simulationen und für verschiedene MTS-Simulationen. Im oberen Teil der Abbildung ist zu sehen, dass sich die mittlere Bindungslänge von CCSD(T)- und pw-DFT-Simulationen um ca.  $0.017 \text{ \AA}$  unterscheiden und die Schwingung bei der pw-DFT Simulation eine etwas niedrigere Schwingungsfrequenz besitzt (siehe Abbildung 6.4). Bei den MTS-Simulationen mit MTS-Faktor  $n = 2$  bzw.  $n = 4$  sind die Korrekturen durch die CCSD(T)-Rechnungen deutlich als schnelle Oszillation entlang der reinen CCSD(T)-Trajektorie zu erkennen. Es ist ersichtlich, dass der MTS-Faktor nicht zu groß sein

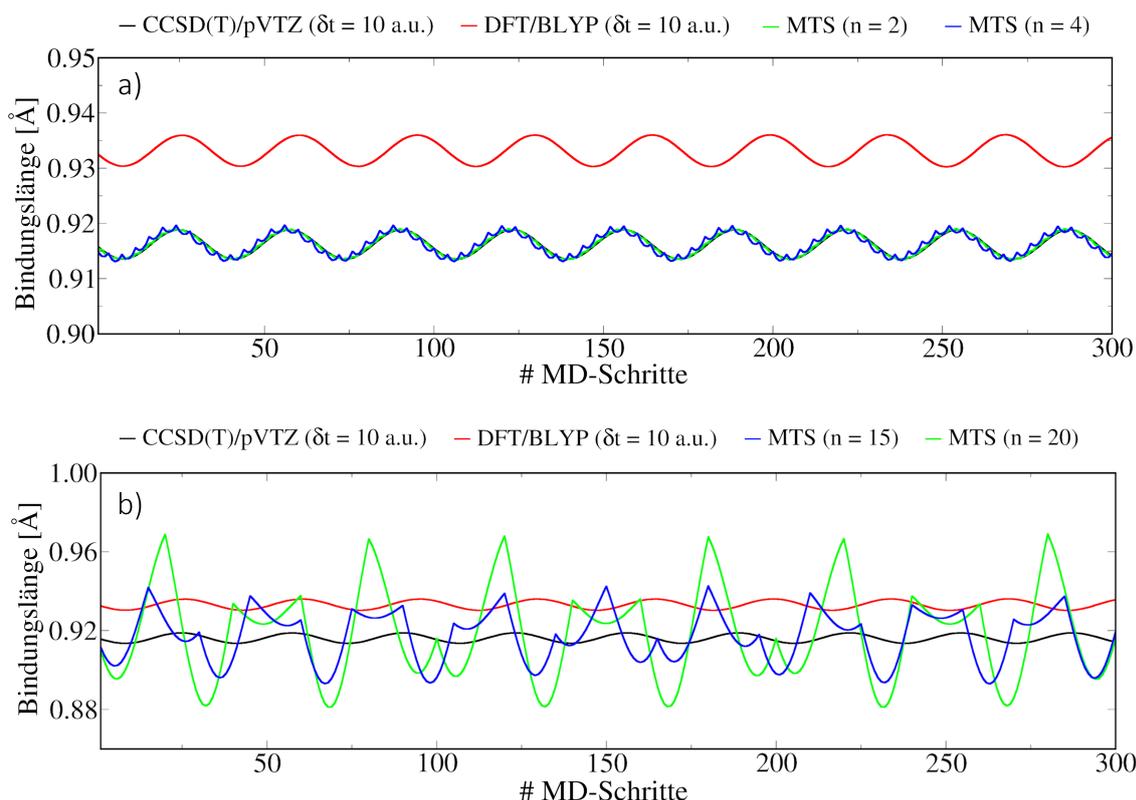


Abbildung 6.5: Oszillation der Bindungslänge von Fluorwasserstoff im Vakuum während AIMD-Simulationen auf CCSD(T)/cc-pVTZ, pw-DFT/PBE und dem kombinierten MTS-CCSD(T)/pw-DFT Niveau. In Teil a) der Abbildung sind die MTS-Faktoren  $n = 2$  bzw.  $n = 4$  zu sehen, in Teil b) die MTS-Faktoren  $n = 15$  bzw.  $n = 20$ .

darf, da er die Frequenz der Korrekturschwingung festlegt und bei einer zu geringen Korrekturfrequenz kann der CCSD(T)-Trajektorie nicht mehr gefolgt werden. Dies ist bei den Simulationen mit den MTS-Faktoren  $n = 15$  bzw.  $n = 20$  zu erkennen. Die Amplitude der Oszillation ist deutlich größer als für kleinere MTS-Faktoren und die Trajektorie der Bindungslänge ist unregelmäßig mit großen Sprüngen.

Die Korrekturschwingung, welche durch die CCSD(T)-Rechnungen entsteht, ist auch in den Schwingungsspektren zu sehen (siehe Abbildung 6.6). Während die Korrekturschwingung für den MTS-Faktor  $n = 2$  eine sehr hohe Frequenz von  $68900 \text{ cm}^{-1}$  besitzt und damit gut von der physikalischen Molekülschwingung separiert ist, besitzen die Korrekturschwingungen für die MTS-Faktoren  $n = 15$  bzw.  $n = 20$  Frequenzen von  $9187 \text{ cm}^{-1}$  bzw.  $6890 \text{ cm}^{-1}$  und sind damit nahe an der Schwingungsfrequenz von Fluorwasserstoff. Die Simulation mit einem MTS-Faktor von  $n = 15$  kann zwar die CCSD(T)-Schwingungsfrequenz reproduzieren, aber die Dynamik der Schwingung verläuft nicht entlang der CCSD(T)-Trajektorie. Daher wird in dem Fall von Fluorwasserstoff ein MTS-Faktor von  $n = 15$  als Grenze angenommen, bis

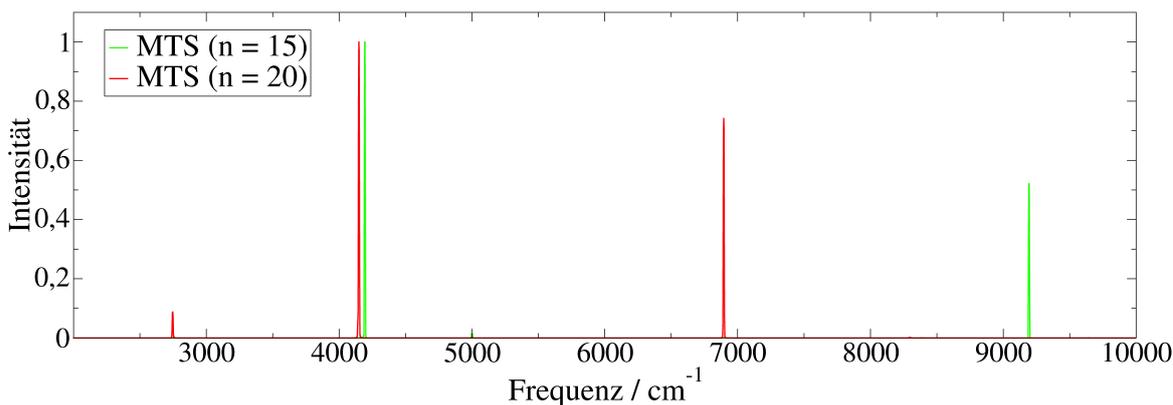


Abbildung 6.6: Schwingungsspektrum von Fluorwasserstoff im Vakuum berechnet aus AIMD-Simulationen auf dem kombinierten CCSD(T)/pw-DFT Niveau mit CCSD(T)/cc-pVTZ Korrekturrechnungen alle  $n = 15$  und  $n = 20$  Zeitschritte.

zu der MTS-Simulationen die CCSD(T)-Ergebnisse sinnvoll reproduzieren können. Das ist auch zu erwarten, da ein Durchlauf der Schwingung von Fluorwasserstoff ca. 33 Zeitschritte benötigt. Der MTS-Faktor sollte daher kleiner als die Hälfte einer Schwingungsperiode des zu untersuchenden Systems sein. Diese Erkenntnis stimmt auch mit den in der Literatur angegebenen Limitationen von MTS-Algorithmen und dem Auftreten von Resonanzproblemen überein.<sup>[117,160]</sup>

Eine weitere von Liberatore et al.<sup>[117]</sup> beschriebene Limitation zur Reduktion der Simulationsdauer ist die ansteigende Nummer an SCF-Zyklen, die bei der Bestimmung der Wellenfunktion für die Korrekturmethode benötigt werden, wenn der MTS-Faktor größer wird. Dies liegt daran, dass der Anfangswert bei der Bestimmung der Wellenfunktion in der Regel aus den Wellenfunktionen der vorangegangenen MD-Schritte extrapoliert wird. Je mehr Referenzschritte zwischen zwei Korrekturschritten liegen, umso schlechter ist der extrapolierte Startwert. In dem hier gezeigten Beispiel ist dieser Effekt allerdings vernachlässigbar. Die relative Beschleunigung für die Simulation mit dem MTS-Faktor  $n = 20$  liegt mit 19.8 sehr nah an dem theoretischen Wert von 20.0. Der Grund dafür ist, dass die Anzahl an benötigten SCF-Zyklen nur minimal ansteigt (9-10 für eine reine CCSD(T)-Simulation bzw. 10-11 für die Simulation mit dem MTS-Faktor  $n = 20$ ). Außerdem ist der zeitbestimmende Schritt einer CCSD(T)-Rechnung das Lösen der CC-Gleichungen, welche unabhängig von dem Anfangswert für die Wellenfunktion ist, da in CFOUR die MP2-Amplituden als Startwert verwendet werden. Es wurde zwar versucht, die Startwerte aus den CC-Amplituden der vorherigen MD-Schritte zu extrapolieren, die MP2-Amplituden stellten jedoch immer einen besserer Startwert dar als die extrapolierten Amplituden. Bei der Berechnung der relativen Beschleunigung ist allerdings nicht die zusätzlich benötigte Zeit der Referenzschritte mit einberechnet, welche für eine reine CCSD(T)-Simulation nicht benötigt werden, da diese einen konstanten Wert bildet, der unabhängig vom verwendeten MTS-Faktor ist. Sie hängt von verschiedenen

Faktoren, wie der verwendeten Methode oder der Anzahl an verwendeten CPUs für CPMD, ab und verringert die praktische Reduktion der Simulationsdauer weiter.

Die Beispielrechnungen an Fluorwasserstoff zeigen, dass die Anzahl an benötigten CCSD(T)-Rechnungen für eine AIMD-Simulation durch die Verwendung des MTS-Algorithmus deutlich reduziert werden kann, was eine Reduktion der Simulationsdauer um einen Faktor von bis zu 15 bedeuten kann. Dabei werden für die untersuchten Eigenschaften wie der Schwingungsfrequenz oder der Bindungslänge trotzdem die gleichen Werte erhalten wie bei einer reinen CCSD(T)-Simulation.

## 6.3 Zusammenfassung

In diesem Kapitel wurden zwei verschiedene Methoden zur Beschleunigung von QM/MM-Simulationen vorgestellt und die Anwendbarkeit an verschiedenen Systemen demonstriert.

Im ersten Teil wurde die Beschreibung von elektrostatischen long-range Wechselwirkungen über eine Multipolentwicklung getestet. Dafür wurde die erreichbare Genauigkeit in Abhängigkeit von der Anzahl an short-range und long-range Atomen in einzelnen quantenchemischen Rechnungen eines Wassersystems mit über 36000 Atomen untersucht. Es wurde gezeigt, dass eine für QM/MM-Simulationen ausreichende Genauigkeit von  $10^{-8}$  a.u. bis  $10^{-9}$  a.u. für die Energie, die Kräfte und für das Dipolmoment bei nur jeweils 10-15% der MM-Atome in der short-range-Region bei einer Reduktion der Rechenzeit von ca. 90% erreicht wird. Anhand von QM/MM-Simulationen des gleichen Systems mit unterschiedlichen Cutoff-Radien wurde die numerische Stabilität der Simulationen bewiesen sowie die Genauigkeit des Dipolmoments, welches durch die unterschiedlichen Simulationen erhalten wurde, untersucht.

Die zweite vorgestellte Methode ist die Anwendung des in CPMD implementierten MTS-Algorithmus. Dafür wurden AIMD-Simulationen mit unterschiedlichen MTS-Faktoren auf dem kombinierten CCSD(T)/cc-pVTZ/pw-DFT/PBE Niveau durchgeführt und die erhaltenen Schwingungsfrequenzen miteinander verglichen. Dabei wurde gezeigt, dass die reine CCSD(T)/cc-pVTZ Schwingungsfrequenz durch MTS-Simulationen mit einem MTS-Faktor von bis zu  $n = 15$  reproduziert werden kann, wodurch die Zahl an nötigen CCSD(T)-Rechnungen um bis zu 93% reduziert werden kann. Allerdings zeigt die Untersuchung der Trajektorie des Bindungsabstands, dass die Schwingungsfrequenz mit einem MTS-Faktor von  $n = 15$  zwar reproduziert werden kann, die Dynamik der Schwingung allerdings nicht entlang der CCSD(T)-Trajektorie verläuft.

---

# 7 Optimierung und Verwendung von ECPs

Bei Simulationen an großen Molekülen, wie Enzymen oder Peptiden in Lösung, ist es essentiell, dass die Grenze zwischen QM- und MM-Bereich innerhalb des Moleküls durch kovalente Bindungen verläuft, damit nicht das gesamte Makromolekül durch eine QM-Methode beschrieben werden muss.<sup>[4,5]</sup> Dadurch kann die Größe des QM-Bereichs verkleinert werden sodass nur die chemisch relevante Gruppe quantenchemisch beschrieben wird. Damit die Elektronenstruktur des QM-Bereichs dabei möglichst wenig gestört wird, wählt man meistens kovalente C-C-Bindungen, durch welche die Grenze zwischen QM- und MM-Bereich verläuft und sättigt den QM-Bereich mit einem Link-Atom oder einem für diese Anwendung optimierten ECP ab. Im Folgenden wird zunächst das Vorgehen der Optimierung verschiedener ECPs bezüglich Essigsäure als Referenzsystem erläutert. Anschließend werden die optimierten ECPs in QM/MM-Simulationen an einem System von flüssiger Essigsäure und eines Poly- $\beta$ -alanins getestet.

## 7.1 Optimierung der ECPs

Wenn die Grenze zwischen QM- und MM-Bereich durch kovalente Bindungen verlaufen muss, wird dafür in der Regel eine C-C-Bindung verwendet. Um die Störung der Elektronenstruktur durch eine offene Valenz an dem Kohlenstoffatom im QM-Bereich zu minimieren, wird an dem Grenzatom ein ECP verwendet. Das ECP wird zur

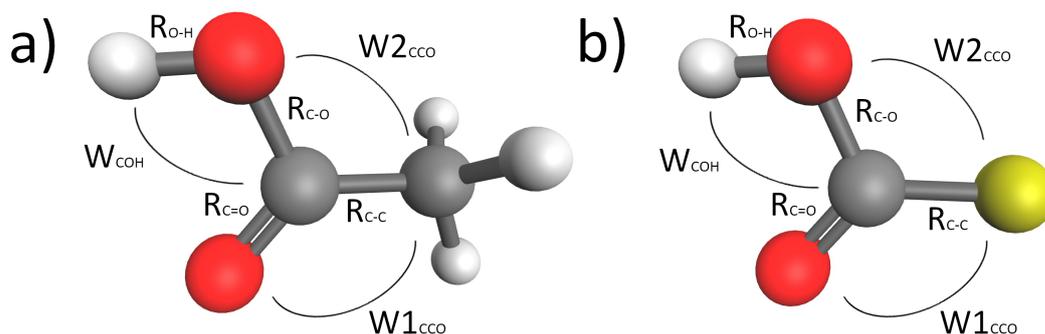


Abbildung 7.1: Struktur von Essigsäure ohne ECP (a) und mit ECP am Kohlenstoffatom (b), welches die Methylgruppe ersetzt (gelbes Atom).

Beschreibung aller Elektronen bis auf eins verwendet, sodass das Kohlenstoffatom nur noch ein freies Elektron besitzt, welches den QM-Bereich absättigen kann. Als Referenzsystem für die Optimierung wurde Essigsäure gewählt, bei dem das Kohlenstoffatom mit dem ECP die Methylgruppe ersetzen soll (siehe Abbildung 7.1). Dadurch ist das optimierte ECP in der Lage den QM-Bereich in komplexeren Systemen an einer Alkylgruppe abzuschließen. Es ist jedoch auch möglich, die Optimierung für andere Bindungssituationen zu optimieren.

Für die Optimierung von ECPs wurden mehrere Varianten implementiert. Als ECPs, welche optimiert werden können, stehen zwei Grundformen, ein großes und ein kleines ECP, zur Verfügung, welche beide dem CFOUR-Standard entsprechen. Das kleine ECP ist ursprünglich ein für Kohlenstoff entwickeltes ECP.<sup>[128,129]</sup> Das große ECP wurde ursprünglich für die Berücksichtigung relativistischer Effekte in Germanium entwickelt<sup>[128,130]</sup> und wurde zur Verwendung für Kohlenstoff entsprechend angepasst. Die ECPs sind in Kapitel 3.6 angegeben.

Vor der Optimierung wurde für Essigsäure als Referenzsystem eine Geometrieoptimierung auf HF/cc-pVDZ bzw. HF/cc-pVTZ Niveau durchgeführt, um die entsprechenden Referenzeigenschaften zu erhalten bezüglich der die ECPs optimiert wurden. Beide ECPs wurden jeweils bezüglich des Dipolmoments, des Gradienten und der Kombination aus beiden unter Verwendung der Basissätze cc-pVDZ und cc-pVTZ optimiert. Bezüglich der Ladungsdichte wurde nur das kleine ECP mit der cc-pVDZ-Basis optimiert. Da die Resultate in etwa die gleichen sind wie bezüglich des Dipolmoments (siehe Anhang B), die Optimierung jedoch viel langsamer konvergiert ist, wurde darauf verzichtet weitere ECPs bezüglich der Ladungsdichte zu optimieren.

Die optimierten ECPs sind in Anhang B angegeben. Während bei der Optimierung bezüglich des Dipolmoments bei beiden ECPs jeweils eine Konvergenz von ca.  $10^{-5}$  a.u. erreicht wurde, konnte bei der Optimierung bezüglich des Gradienten nur eine Konvergenz von  $10^{-3}$ -  $10^{-4}$  a.u. erreicht werden.

Im Anschluss an die Optimierung der ECPs wurden ebenfalls Geometrieoptimierungen der Essigsäure mit den optimierten ECPs durchgeführt, um die resultierenden Eigenschaften vergleichen zu können. In Tabelle 7.1 bis 7.4 sind die Abweichungen der Geometrie, des Dipolmoments und der Ladungsdichte bezüglich des Referenzsystems nach der Geometrieoptimierung angegeben.

Bei der Optimierung bezüglich des Dipolmoments fällt auf, dass besonders die C-C-Bindung, welche direkt an das ECP grenzt, in allen Fällen um  $-0.1971$  bis  $-0.2467$  Å zu kurz ist. Alle anderen Bindungen weisen nur kleine Abweichungen von ca.  $10^{-3}$  Å auf. Der Fehler der Winkel liegt in allen Fällen unter 1 Grad. Das Dipolmoment wird auf  $10^{-2}$  D genau erhalten und die Abweichung der Ladungsdichte an den jeweiligen Atomen beträgt jeweils nur ca.  $10^{-2}$  a.u..

Bei der Optimierung bezüglich des Gradienten werden deutlich bessere Geometrien erhalten. Besonders die C-C-Bindung, welche bei der Optimierung bezüglich des Dipolmoments deutlich zu kurz ist, weist nur noch einen Fehler von ca.  $10^{-3}$  Å auf. Auch die Fehler der Winkel sind mit  $10^{-1}$  bis  $10^{-2}$  Grad etwas kleiner. Dafür ist der Fehler des Dipolmoments mit 0.421 bis 0.681 D deutlich größer. Der Fehler der Ladungsdichten ist nur unwesentlich größer als bei der Optimierung bezüglich des

Tabelle 7.1: Ergebnisse der Optimierung des kleinen ECPs mit cc-pVDZ Basis mit Essigsäure als Referenzsystem. Für die Optimierungen sind jeweils die Differenzen bzgl. der Referenzwerte angegeben. Die Bindungslängen sind in Å angegeben, die Winkel in Grad, das Dipolmoment  $\mu$  in Debye und die Ladungsdichten  $\delta$  in atomaren Einheiten.

Parameter	Ref.	Optimierung bezüglich:		
		Dipolmoment	Gradient	Kombiniert
$R_{C-C}$	1.4996	-0.2467	0.0026	-0.0128
$R_{C=O}$	1.1853	-0.0033	-0.0011	-0.0046
$R_{C-O}$	1.3317	-0.0034	0.0046	-0.0008
$R_{O-H}$	0.9511	0.0024	0.0037	0.0037
$W1_{CCO}$	125.72	-0.61	0.29	0.03
$W2_{CCO}$	111.94	-0.60	-0.05	-0.86
$W_{COH}$	107.77	0.32	-0.08	0.06
$\mu$	1.802	-0.043	0.660	0.228
$\delta_C$	120.769	0.054	0.111	0.072
$\delta_O$	297.586	0.034	0.038	0.044
$\delta_O$	297.263	0.044	0.084	0.076
$\delta_H$	0.360	-0.001	-0.001	-0.001

Tabelle 7.2: Ergebnisse der Optimierung des großen ECPs mit cc-pVDZ Basis mit Essigsäure als Referenzsystem. Für die Optimierungen sind jeweils die Differenzen bzgl. der Referenzwerte angegeben. Die Bindungslängen sind in Å angegeben, die Winkel in Grad, das Dipolmoment  $\mu$  in Debye und die Ladungsdichten  $\delta$  in atomaren Einheiten.

Parameter	Ref.	Optimierung bezüglich:		
		Dipolmoment	Gradient	Kombiniert
$R_{C-C}$	1.4996	-0.2122	0.0030	-0.0145
$R_{C=O}$	1.1853	-0.0042	-0.0011	-0.0046
$R_{C-O}$	1.3317	-0.0040	0.0042	-0.0014
$R_{O-H}$	0.9511	0.0026	0.0036	0.0035
$W1_{CCO}$	125.72	-0.66	0.17	-0.16
$W2_{CCO}$	111.94	-0.70	0.12	-0.58
$W_{COH}$	107.77	0.30	-0.12	0.00
$\mu$	1.802	-0.061	0.421	0.258
$\delta_C$	120.769	0.059	0.108	0.053
$\delta_O$	297.586	0.037	0.036	0.041
$\delta_O$	297.263	0.049	0.083	0.074
$\delta_H$	0.360	-0.001	-0.001	-0.001

Tabelle 7.3: Ergebnisse der Optimierung des kleinen ECPs mit cc-pVTZ Basis mit Essigsäure als Referenzsystem. Für die Optimierungen sind jeweils die Differenzen bzgl. der Referenzwerte angegeben. Die Bindungslängen sind in Å angegeben, die Winkel in Grad, das Dipolmoment  $\mu$  in Debye und die Ladungsdichten  $\delta$  in atomaren Einheiten.

Parameter	Ref.	Optimierung bezüglich:		
		Dipolmoment	Gradient	Kombiniert
$R_{C-C}$	1.4995	-0.2348	0.0028	-0.0135
$R_{C=O}$	1.1817	-0.0036	-0.0011	-0.0049
$R_{C-O}$	1.3293	-0.0033	0.0048	-0.0007
$R_{O-H}$	0.9455	0.0025	0.0036	0.0037
$W1_{CCO}$	125.62	-0.71	-0.06	-0.26
$W2_{CCO}$	112.01	-0.45	0.30	-0.59
$W_{COH}$	108.47	0.23	-0.02	0.09
$\mu$	1.897	-0.045	0.681	0.212
$\delta_C$	121.256	0.011	0.070	0.045
$\delta_O$	298.819	0.051	0.055	0.061
$\delta_O$	298.502	0.062	0.097	0.092
$\delta_H$	0.411	-0.002	-0.001	-0.002

Tabelle 7.4: Ergebnisse der Optimierung des großen ECPs mit cc-pVTZ Basis mit Essigsäure als Referenzsystem. Für die Optimierungen sind jeweils die Differenzen bzgl. der Referenzwerte angegeben. Die Bindungslängen sind in Å angegeben, die Winkel in Grad, das Dipolmoment  $\mu$  in Debye und die Ladungsdichten  $\delta$  in atomaren Einheiten.

Parameter	Ref.	Optimierung bezüglich:		
		Dipolmoment	Gradient	Kombiniert
$R_{C-C}$	1.4995	-0.1971	0.0050	-0.0135
$R_{C=O}$	1.1817	-0.0046	-0.0007	-0.0048
$R_{C-O}$	1.3293	-0.0038	0.0053	-0.0010
$R_{O-H}$	0.9455	0.0027	0.0035	0.0035
$W1_{CCO}$	125.62	-0.64	-0.03	-0.28
$W2_{CCO}$	112.01	-0.66	0.43	-0.41
$W_{COH}$	108.47	0.26	-0.05	0.04
$\mu$	1.897	-0.057	0.465	0.277
$\delta_C$	121.256	0.025	0.074	0.036
$\delta_O$	298.819	0.053	0.053	0.057
$\delta_O$	298.502	0.065	0.096	0.089
$\delta_H$	0.411	-0.002	-0.001	-0.002

Dipolmoments.

Bei der kombinierten Optimierung wird im Vergleich zur Optimierung bezüglich des Gradienten ein größerer Fehler für die C-C-Bindung zum ECP-Kohlenstoffatom von  $-0.0128$  bis  $-0.0145$  Å erhalten, was deutlich besser ist als bei der Optimierung bezüglich des Dipolmoments. Die Fehler der anderen Bindungslängen und Winkel sind von der gleichen Größenordnung wie bei der Optimierung bezüglich des Gradienten. Beim Dipolmoment ist der Fehler etwas kleiner als im Vergleich zur Optimierung bezüglich des Gradienten, aber mit  $0.212$  bis  $0.277$  D größer als bei der Optimierung bezüglich des Dipolmoments. Die Fehler in den Ladungsdichten ist bei allen drei Optimierung in etwa gleich groß.

Zwischen den beiden ECPs und den Basissätzen können keine wesentlichen Unterschiede in den Ergebnissen beobachtet werden. Lediglich das Dipolmoment ist bei der Optimierung bezüglich des Gradienten bei den kleinen ECPs etwas größer als bei den großen ECPs. Bei der kombinierten Optimierung wiederum ist der Fehler des Dipolmoments bei den kleinen ECPs etwas kleiner.

Da die C-C-Bindung bei der Optimierung bezüglich des Dipolmoments deutlich zu kurz ist und die anderen beiden Optimierungen insgesamt wesentlich bessere Ergebnisse liefern, wurden die so optimierten ECPs nicht weiter verwendet.

## 7.2 Test der ECPs in QM/MM-Simulationen von flüssiger Essigsäure

Um die Anwendbarkeit der optimierten ECPs in QM/MM-Simulationen zu untersuchen, wurden sie zunächst in QM/MM-Simulationen von flüssiger Essigsäure getestet. Dazu wurden zwei verschiedene Systeme verwendet. Im ersten System wird der klassische Bereich von einem *all-atom*-Kraftfeld beschrieben, bei dem alle Atome explizit berücksichtigt werden (Abbildung 7.2 a). Im zweiten System wird der klassische Bereich hingegen von einem *united-atom*-Kraftfeld beschrieben, bei dem die aliphatischen Wasserstoffatome nicht explizit beschrieben werden, sondern zusammen mit den jeweiligen Kohlenstoffatomen zu einem Atom zusammengefasst werden (Abbildung 7.2 b). Dies macht bei der Verwendung von ECPs insoweit einen Unterschied, da sich alle Atome jenseits des ECP-Kohlenstoffatoms im MM-Bereich befinden und daher vom Kraftfeld beschrieben werden. Sind diese Atome Kohlenstoffatome, gibt es keinen Unterschied zwischen *all-atom*- und *united-atom*-Kraftfeld. Sind diese Atome

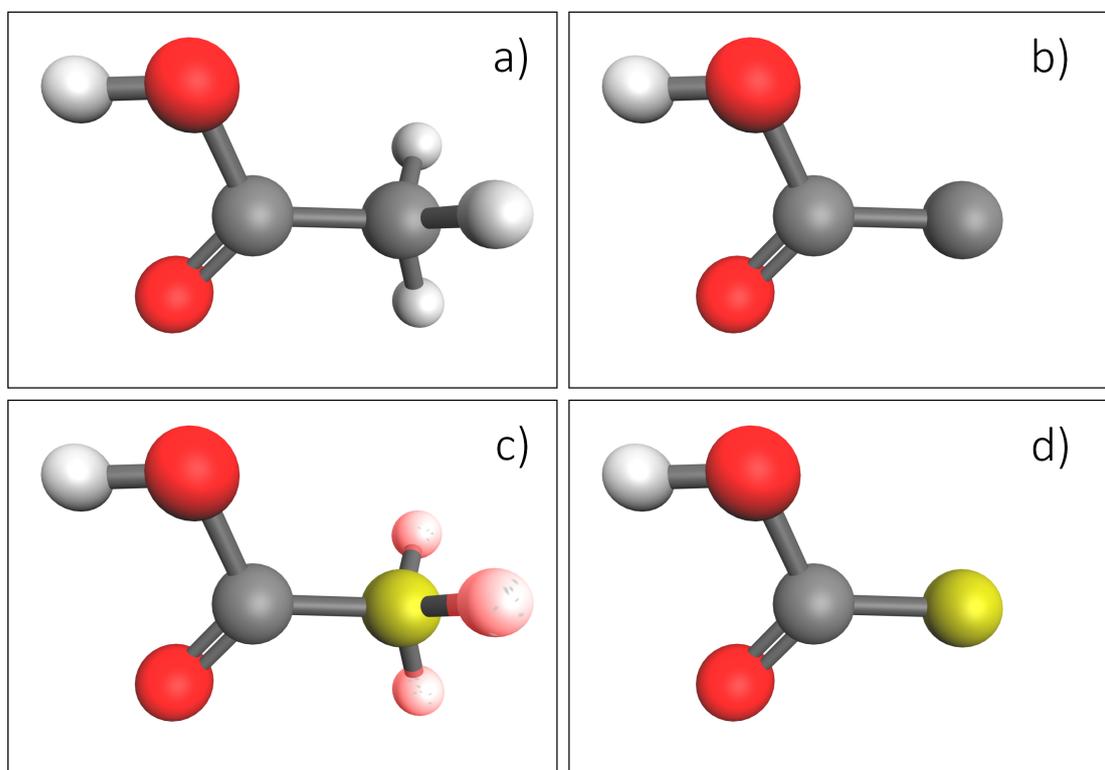


Abbildung 7.2: Struktur von Essigsäure in einem *all-atom*-Kraftfeld (a) und in einem *united-atom*-Kraftfeld (b). In Abbildung c) ersetzt das ECP den Methyl-Kohlenstoff, die Wasserstoffatome der Methylgruppe (farblich markiert) werden durch das *all-atom*-Kraftfeld beschrieben. In Abbildung d) ersetzt das ECP den Methyl-Kohlenstoff des *united-atom*-Essigsäuremoleküls.

allerdings Wasserstoffatome, so müssen sie im Falle eines *all-atom*-Kraftfeld vom Kraftfeld beschrieben werden (Abbildung 7.2 c). Im Falle eines *united-atom*-Kraftfelds werden die Wasserstoffatome hingegen nicht berücksichtigt und werden weder im QM-Bereich noch im MM-Bereich beschrieben (Abbildung 7.2 d).

Zum Vergleich der QM/MM-Simulationen mit den optimierten ECPs wurde jeweils eine QM/MM-Simulation ohne Verwendung eines ECPs durchgeführt, in dem sich alle acht Atome der Essigsäure im QM-Bereich befinden. Getestet wurden die bezüglich des Gradienten optimierten ECPs sowie die ECPs der kombinierten Optimierung jeweils auf HF/cc-pVDZ bzw. HF/cc-pVTZ Niveau. Als Vergleichsparameter wurden die gleichen Bindungslängen und Winkel wie in Abschnitt 7.1 sowie das Dipolmoment verwendet. Es wurden jeweils die über die gesamte Simulation gemittelten Werte verglichen. Die erhaltenen Werte bzw. Abweichungen zu den Referenzwerten sowie die jeweilige Standardabweichungen sind in Tabelle 7.5 bis 7.8 zusammengefasst.

Bei den QM/MM-Simulation mit dem *all-atom*-Kraftfeld auf HF/cc-pVDZ Niveau sieht man für die Bindungslängen und Winkel zwischen dem großen und kleinen ECP und zwischen den beiden Optimierungsarten keine signifikanten Unterschiede. Die Abweichungen haben jeweils immer das gleiche Vorzeichen und sind in etwa gleich groß. Eine Ausnahme bildet der Winkel  $W_{\text{COH}}$ , der bei dem kleinen ECP, welches durch die kombinierte Optimierung erhalten wurde, deutlich kleiner als bei den anderen ECPs und dem Referenzsystem ist. Alle Abweichungen der Bindungslängen und Winkel sind etwas größer als die der Geometrieoptimierungen (siehe Tabellen 7.1 und 7.2), liegen aber alle innerhalb der Standardabweichung und auch in den meisten Fällen innerhalb der Standardabweichung des Referenzsystems. Insgesamt sind die Abweichungen relativ klein und die Geometrie kann zufriedenstellend reproduziert werden.

Bei dem Dipolmoment hingegen können nur die durch die kombinierte Optimierung erhaltenen ECPs die Referenzwerte gut reproduzieren. Die bezüglich des Gradienten optimierten ECPs liefern um  $-0.376$  bzw.  $-0.279$  Debye zu kleine Werte, was allerdings trotzdem innerhalb der Standardabweichung liegt, welche mit  $\pm 0.420$  bzw.  $\pm 0.461$  Debye relativ groß ist.

Bei den QM/MM-Simulationen mit dem *all-atom*-Kraftfeld auf HF/cc-pVTZ Niveau sind für die Bindungslängen und Winkel ebenfalls keine signifikanten Unterschiede zwischen den verschiedenen ECPs zu erkennen. Alle Abweichungen sind ebenfalls relativ klein und die Geometrie wird gut reproduziert.

Für das Dipolmoment allerdings ergeben alle ECPs einen um  $-0.270$  bis  $-0.361$  Debye zu kleinen Wert, wobei die Werte der kombiniert optimierten ECPs etwas besser sind als die der ECPs, die bezüglich des Gradienten optimiert wurden.

Verglichen mit den HF/cc-pVDZ-Simulationen sind die einzigen Unterschiede, dass die Bindung  $R_{\text{O-H}}$  und der Winkel  $W1_{\text{CCO}}$  jeweils bei den HF/cc-pVTZ-Simulationen um ca. eine Größenordnung besser reproduziert werden können. Außerdem ist bei Winkel  $W2_{\text{CCO}}$  das Vorzeichen der Abweichung unterschiedlich. Alle anderen Parameter werden ähnlich gut reproduziert.

Bei den QM/MM-Simulationen mit dem *united-atom*-Kraftfeld sind die Resultate ähnlich. Die Abweichungen der Bindungslängen und Winkel sind in manchen Fällen

Tabelle 7.5: Abweichungen der mittleren Bindungslängen, Winkel und des Dipolmoments bezüglich des Referenzsystems und Standardabweichungen für die über die mit den optimierten ECPs durchgeführten QM/MM-Simulationen auf HF/cc-pVDZ Niveau mit *all-atom*-Kraftfeld. Die Bindungslängen sind in Å angegeben, die Winkel in Grad und das Dipolmoment  $\mu$  in Debye.

Parameter	Ref.	kleines ECP		großes ECP	
		Gradient	Kombiniert	Gradient	Kombiniert
R <sub>C-C</sub>	1.502±0.038	0.025±0.053	0.013±0.051	0.024±0.048	0.019±0.057
R <sub>C=O</sub>	1.195±0.012	-0.016±0.011	-0.017±0.021	-0.014±0.011	-0.017±0.018
R <sub>C-O</sub>	1.327±0.033	-0.013±0.025	-0.024±0.029	-0.014±0.025	-0.028±0.033
R <sub>O-H</sub>	0.964±0.009	0.013±0.043	0.018±0.053	0.016±0.052	0.018±0.055
W <sub>1CCO</sub>	124.78±2.88	-0.78±4.83	-1.69±4.44	-1.25±3.96	-2.43±4.49
W <sub>2CCO</sub>	112.73±2.76	-2.15±4.43	-1.61±4.03	-1.85±4.03	-1.12±4.15
W <sub>COH</sub>	110.42±4.87	-1.08±4.60	-11.70±13.59	-0.41±4.00	-0.93±4.18
$\mu$	2.499±0.298	-0.376±0.420	-0.007±0.411	-0.279±0.461	0.019±0.435

Tabelle 7.6: Abweichungen der mittleren Bindungslängen, Winkel und des Dipolmoments bezüglich des Referenzsystems und Standardabweichungen für die über die mit den optimierten ECPs durchgeführten QM/MM-Simulationen auf HF/cc-pVTZ Niveau mit *all-atom*-Kraftfeld. Die Bindungslängen sind in Å angegeben, die Winkel in Grad und das Dipolmoment  $\mu$  in Debye.

Parameter	Ref.	kleines ECP		großes ECP	
		Gradient	Kombiniert	Gradient	Kombiniert
R <sub>C-C</sub>	1.504±0.051	0.024±0.068	0.013±0.074	0.026±0.067	0.015±0.082
R <sub>C=O</sub>	1.192±0.022	-0.015±0.011	-0.019±0.014	-0.015±0.010	-0.018±0.014
R <sub>C-O</sub>	1.328±0.043	-0.003±0.035	-0.002±0.035	-0.001±0.034	0.001±0.039
R <sub>O-H</sub>	0.960±0.010	0.012±0.055	0.011±0.055	0.013±0.056	0.010±0.052
W <sub>1CCO</sub>	125.06±3.07	0.04±4.83	-0.09±6.09	0.17±4.80	-0.16±5.61
W <sub>2CCO</sub>	112.05±3.90	1.08±4.56	2.06±6.70	1.31±4.63	1.92±6.84
W <sub>COH</sub>	112.84±3.46	-3.08±4.42	-4.24±5.41	-2.85±4.45	-3.95±5.16
$\mu$	2.638±0.416	-0.361±0.437	-0.305±0.430	-0.344±0.432	-0.270±0.416

Tabelle 7.7: Abweichungen der mittleren Bindungslängen, Winkel und des Dipolmoments bezüglich des Referenzsystems und Standardabweichungen für die über die mit den optimierten ECPs durchgeführten QM/MM-Simulationen auf HF/cc-pVDZ Niveau mit *united-atom*-Kraftfeld. Die Bindungslängen sind in Å angegeben, die Winkel in Grad und das Dipolmoment  $\mu$  in Debye.

Parameter	Ref.	kleines ECP		großes ECP	
		Gradient	Kombiniert	Gradient	Kombiniert
R <sub>C-C</sub>	1.502±0.033	0.011±0.046	-0.010±0.044	0.010±0.045	-0.012±0.044
R <sub>C=O</sub>	1.200±0.022	-0.002±0.027	-0.005±0.029	-0.004±0.020	-0.007±0.023
R <sub>C-O</sub>	1.314±0.024	0.014±0.026	0.010±0.017	0.017±0.035	0.008±0.020
R <sub>O-H</sub>	0.966±0.021	0.005±0.033	0.005±0.029	0.006±0.037	0.005±0.027
W <sub>1CCO</sub>	124.06±2.46	1.16±4.40	1.17±4.09	1.22±4.99	1.01±3.23
W <sub>2CCO</sub>	112.94±2.40	-0.26±4.43	-0.88±4.06	-0.15±4.98	-0.95±3.20
W <sub>COH</sub>	110.04±3.44	-0.25±4.88	-0.38±3.97	-0.42±4.63	-0.14±4.15
$\mu$	2.656±0.334	0.454±0.333	0.092±0.342	0.415±0.307	0.097±0.293

Tabelle 7.8: Abweichungen der mittleren Bindungslängen, Winkel und des Dipolmoments bezüglich des Referenzsystems und Standardabweichungen für die über die mit den optimierten ECPs durchgeführten QM/MM-Simulationen auf HF/cc-pVTZ Niveau mit *united-atom*-Kraftfeld. Die Bindungslängen sind in Å angegeben, die Winkel in Grad und das Dipolmoment  $\mu$  in Debye.

Parameter	Ref.	kleines ECP		großes ECP	
		Gradient	Kombiniert	Gradient	Kombiniert
R <sub>C-C</sub>	1.512±0.025	0.003±0.032	-0.015±0.057	0.003±0.033	-0.018±0.053
R <sub>C=O</sub>	1.187±0.022	0.007±0.023	0.006±0.027	0.007±0.023	0.006±0.028
R <sub>C-O</sub>	1.293±0.023	0.037±0.026	0.029±0.029	0.037±0.026	0.029±0.027
R <sub>O-H</sub>	0.958±0.025	0.007±0.032	0.008±0.030	0.007±0.031	0.007±0.029
W <sub>1CCO</sub>	123.19±2.17	2.36±5.42	2.00±3.86	2.44±5.30	2.00±3.66
W <sub>2CCO</sub>	113.23±1.97	0.37±5.11	-0.43±3.50	0.48±4.98	-0.25±3.38
W <sub>COH</sub>	114.43±4.37	-1.08±6.64	-3.39±4.82	-1.48±6.54	-3.52±4.55
$\mu$	2.752±0.282	0.691±0.313	0.261±0.318	0.757±0.305	0.303±0.322

zwar etwas kleiner als bei den Simulationen mit dem *all-atom*-Kraftfeld, allerdings sind die Unterschiede nur minimal. Die durch die kombinierte Optimierung erhaltenen ECPs liefern jeweils bessere Werte für das Dipolmoment als die bezüglich des Gradienten optimierten ECPs, allerdings sind die Abweichungen bei den HF/cc-pVTZ-Simulationen deutlich größer als bei den Simulationen mit der cc-pVDZ-Basis.

Obwohl das große ECP deutlich mehr Parameter als das kleine ECP besitzt, welche optimiert werden können und die Flexibilität deutlich größer sein sollte, ist die erzielte Genauigkeit bei beiden ECPs in etwa die gleiche. Da die Optimierung des großen ECPs aufgrund der größeren Anzahl an Parameter jedoch deutlich länger dauert als die Optimierung des kleinen ECPs, ist es sinnvoll ausschließlich das kleine ECP zu verwenden, wenn erneut ein ECP für andere Bindungsverhältnisse optimiert werden soll. Außerdem sollte die kombinierte Optimierung bzw. das durch die kombinierte Optimierung erhaltene ECP verwendet werden, da die Geometrie mit allen ECPs ähnlich gut reproduziert werden kann, die kombinierte Optimierung aber im Vergleich zu den bezüglich des Gradienten optimierten ECPs beim Dipolmoment eine höhere Genauigkeit besitzt.

## 7.3 Test der ECPs in QM/MM-Simulationen an einem Poly- $\beta$ -alanin

Die optimierten ECPs können die Struktur von Essigsäure in QM/MM-Simulationen gut reproduzieren. Allerdings ist Essigsäure so klein, dass es vollständig mit quantenchemischen Methoden beschrieben werden kann. Um die Anwendbarkeit für komplexere Systeme zu testen, welche zu groß sind um sie komplett mit quantenchemischen Methoden zu beschreiben, wurde das kleine ECP, welches durch die kombinierte Optimierung erhalten wurde, an einem Poly- $\beta$ -alanin-System getestet. Das System besteht aus einem Poly- $\beta$ -alanin, welches aus acht Segmenten besteht und 6254 Methanolmolekülen als Lösungsmittel. Um die Anwendbarkeit des ECPs zu testen, wurden vier unterschiedlich große QM-Bereiche gewählt und die mittleren Bindungsabstände der Atome C, N und O miteinander verglichen. Da es nicht möglich ist das gesamte Molekül quantenchemisch zu beschreiben, wurde der größte gewählte QM-Bereich als Referenz verwendet. Das Poly- $\beta$ -alanin ist in Abbildung 7.3 dargestellt, dabei werden nur im QM-Bereich alle aliphatischen Wasserstoffatome explizit berücksichtigt. Der Rest des Moleküls wird durch ein *united-atom*-Kraftfeld beschrieben. Die Simulationen wurden jeweils auf HF/cc-pVDZ Niveau unter Verwendung eines long-range-Cutoffs von 50 a.u. (ca. 26.5 Å) verwendet. Die Verwendung

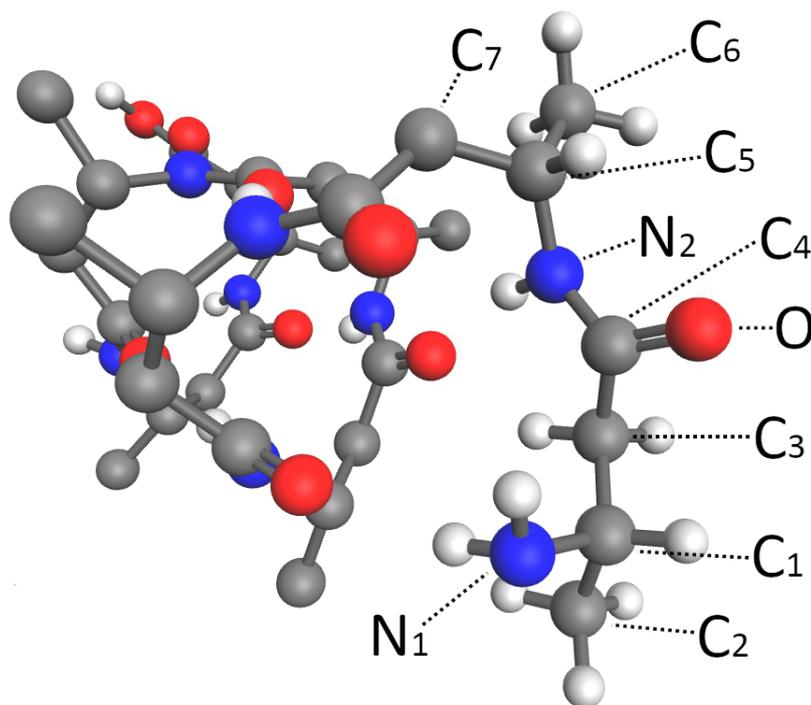


Abbildung 7.3: Struktur des Poly- $\beta$ -alanins, bei dem sich alle Atome bis einschließlich Atom C<sub>7</sub> im QM-Bereich befinden. Der linke Teil des Moleküls gehört zum MM-Bereich.

Tabelle 7.9: Mittlere Bindungslängen und Standardabweichung in Å für unterschiedliche QM-Bereiche in QM/MM-Simulationen von Poly- $\beta$ -alanin in Methanol auf HF/cc-pVDZ Niveau bei 300K. Für die QM-Bereiche 1-3 ist die jeweilige Differenz zum entsprechenden Referenzwert angegeben.

Bindung	Referenz	3. QM-Bereich	2. QM-Bereich	1. QM-Bereich
$R_{N_1-C_1}$	$1.4617 \pm 0.0245$	$0.0020 \pm 0.0301$	$-0.0130 \pm 0.0293$	$-0.0508 \pm 0.0381$
$R_{C_1-C_2}$	$1.5319 \pm 0.0310$	$0.0005 \pm 0.0358$	$-0.0088 \pm 0.0319$	-
$R_{C_2-C_3}$	$1.5392 \pm 0.0300$	$-0.0038 \pm 0.0345$	$-0.0200 \pm 0.0533$	-
$R_{C_3-C_4}$	$1.5215 \pm 0.0313$	$0.0062 \pm 0.0397$	-	-
$R_{C_4=O}$	$1.2135 \pm 0.0141$	$0.0036 \pm 0.0151$	-	-
$R_{C_4-N_2}$	$1.3500 \pm 0.0273$	$-0.0026 \pm 0.0312$	-	-
$R_{N_2-C_5}$	$1.4507 \pm 0.0295$	$-0.0399 \pm 0.0480$	-	-
$R_{C_5-C_6}$	$1.5271 \pm 0.0264$	-	-	-
$R_{C_6-C_7}$	$1.5090 \pm 0.0640$	-	-	-

von long-range-Wechselwirkungen setzt voraus, dass die elektronische Struktur nicht stark gestört wird, da Energie und Kräfte über eine Multipolentwicklung berechnet werden. Eine direkte Beurteilung des Dipolmoments ist nicht möglich, da die Dipolmomente der unterschiedlichen QM-Bereiche nicht miteinander verglichen werden können.

Der in Abbildung 7.3 hervorgehobene Bereich bis einschließlich Kohlenstoffatom  $C_7$  zeigt den größten gewählten QM-Bereich, bei dem das ECP am  $C_7$ -Atom sitzt. Beim zweitgrößten QM-Bereich sitzt das ECP an Atom  $C_5$ . Bei dem mittleren und dem kleinen QM-Bereich sitzt das ECP an Atom  $C_3$  bzw. an Atom  $C_1$ . Dabei ist zu berücksichtigen, dass das ECP so optimiert wurde, dass das ECP den QM-Bereich an einer kovalenten C-C-Bindung abschließt. Bei QM-Bereich 1 und 3 sitzt das ECP an den Atomen  $C_5$  und  $C_1$ , sodass der QM-Bereich an einer C-N-Bindung abgeschlossen wird.

Die mittleren Bindungslängen und die Standardabweichungen, welche über die QM/MM-Simulationen erhalten wurden, sind in Tabelle 7.9 angegeben. Die Bindung zum ECP-Kohlenstoff ist in allen drei kleinen QM-Bereichen kürzer als bei der Referenzsimulation. Die Abweichungen der Bindungen  $R_{N_2-C_5}$  im 3. QM-Bereich bzw.  $R_{N_1-C_1}$  im 1. QM-Bereich sind mit  $-0.0399 \pm 0.0480$  Å bzw.  $-0.0508 \pm 0.0381$  Å doppelt so groß wie die Abweichung von Bindung  $R_{C_2-C_3}$  im 2. QM-Bereich. Dies liegt daran, dass diese beiden Bindungen C-N-Bindungen sind, für welche das ECP nicht optimiert wurde. Trotzdem sind die Abweichungen selbst in diesen Fällen relativ klein, sodass die Struktur gut reproduziert wird.

Für alle anderen Bindungen sind die Abweichungen bezüglich der Referenzsimulation in der Größenordnung von  $10^{-3}$  Å und damit sehr klein. Selbst die Bindungen, die nur ein Atom vom ECP-Kohlenstoff entfernt sind, können ausreichend gut beschrie-

ben werden und die Standardabweichung der Bindungslängen sind ebenfalls in der gleichen Größenordnung wie bei der Referenzsimulation. Dies bedeutet für Anwendungen, dass das ECP-Kohlenstoff nicht weit vom relevanten Zentrum entfernt sein muss, da sich die Störung der Elektronenstruktur nur auf die an das ECP-Kohlenstoff angrenzende Bindung auswirkt und selbst dort nicht besonders groß ist.

Mit den optimierten ECPs kann der QM-Bereich deutlich verkleinert und der Rechenaufwand gesenkt werden, ohne dass die relevante elektronische Struktur im QM-Bereich stark gestört wird. Zudem ist die Verwendung der ECPs sehr einfach, da sie sich am letzten Kohlenstoffatom des QM-Bereichs befinden. Bei allen durchgeführten Simulationen wurden keine Restriktionen, wie z.B. das Festhalten von Bindungen, verwendet, wie es in anderen Implementierungen üblich ist. <sup>[4,123]</sup>

## 7.4 Zusammenfassung

In diesem Kapitel wurde die Optimierung und Anwendung von ECPs zur Verwendung an kovalenten Bindungsbrüchen entlang der QM/MM-Grenze vorgestellt.

Dabei wurden zunächst die Ergebnisse der Optimierung von verschiedenen ECPs bezüglich der Eigenschaften des Referenzsystems Essigsäure dargestellt, welche zeigen, dass die Struktur, das Dipolmoment und die Ladungsdichte von Essigsäure unter Verwendung der optimierten ECPs in quantenchemischen Rechnungen gut reproduziert werden können.

Anschließend wurden die optimierten ECPs in unterschiedlichen QM/MM-Simulationen von flüssiger Essigsäure und einem in Methanol gelösten Poly- $\beta$ -alanin verwendet. Die Ergebnisse zeigen, dass die optimierten ECPs nicht nur für die Anwendung in QM/MM-Simulationen des Systems, für welche sie optimiert wurden, geeignet sind. Auch in QM/MM-Simulationen mit Poly- $\beta$ -alanin mit unterschiedlichen QM-Bereichen wird die QM-Struktur mit ECPs gut reproduziert. Selbst die Bindungen, die direkt an das ECP-Kohlenstoffatom angrenzen, weisen nur kleine Abweichungen bezüglich des Referenzsystems auf.

---

# 8 Schwingungsfrequenzen aus AIMD- und QM/MM-Simulationen

Die Berechnung von Schwingungsfrequenzen aus den Trajektorien von MD-Simulationen ist gängige Praxis.<sup>[161–168]</sup> Dabei spielt es zunächst keine Rolle, ob es sich um Trajektorien aus AIMD-<sup>[166–169]</sup>, QM/MM-<sup>[170]</sup> oder rein klassischen<sup>[162,164,171]</sup> Simulationen handelt. Meistens werden die Spektren über eine Fouriertransformation einer Autokorrelationsfunktion<sup>[162,163,165,169,172]</sup> oder über eine Hauptkomponentenanalyse<sup>[162,163,173]</sup> (PCA) erhalten. Bei QM/MM-Simulationen können die Schwingungsfrequenzen des QM-Bereichs auch aus einzelnen Snapshots der Trajektorie berechnet werden.<sup>[122]</sup>

Im folgenden Abschnitt wird zunächst der Anteil an Anharmonizität in den aus AIMD-Simulationen erhaltenen Schwingungsspektren untersucht. Anschließend wird die Übertragbarkeit der erhaltenen Erkenntnisse auf die unterschiedlichen Berechnungen von Schwingungsfrequenzen aus QM/MM-Simulationen diskutiert.

## 8.1 Untersuchung der Anharmonizität in Schwingungsfrequenzen aus AIMD-Simulationen

Zur Berechnung von Schwingungsfrequenzen wird in der Quantenchemie in der Regel eine harmonische Näherung verwendet.<sup>[7,174]</sup> Dazu wird die Potentialfläche (PES) um die Gleichgewichtsstruktur herum in einer Taylorreihe entwickelt und nach dem quadratischen Term abgebrochen. In Abbildung 8.1 ist die Potentialkurve von Fluorwasserstoff auf CCSD(T)/cc-pVTZ Niveau berechnet und die harmonische Näherung im Minimum dargestellt. Es ist zu erkennen, dass die harmonische Näherung in der Nähe der Gleichgewichtsstruktur die exakte Potentialkurve gut darstellt. Je weiter der Bindungsabstand jedoch von der Gleichgewichtsstruktur abweicht, umso ungenauer wird die harmonische Näherung. Durch eine Transformation in Normalkoordinaten kann die Bewegung der Kerne nun durch einen (multidimensionalen) harmonischen Oszillator beschrieben werden.<sup>[7,175]</sup> Die daraus resultierenden Schwingungsfrequenzen werden über

$$\nu = \frac{1}{2\pi} \sqrt{k} \quad (8.1)$$

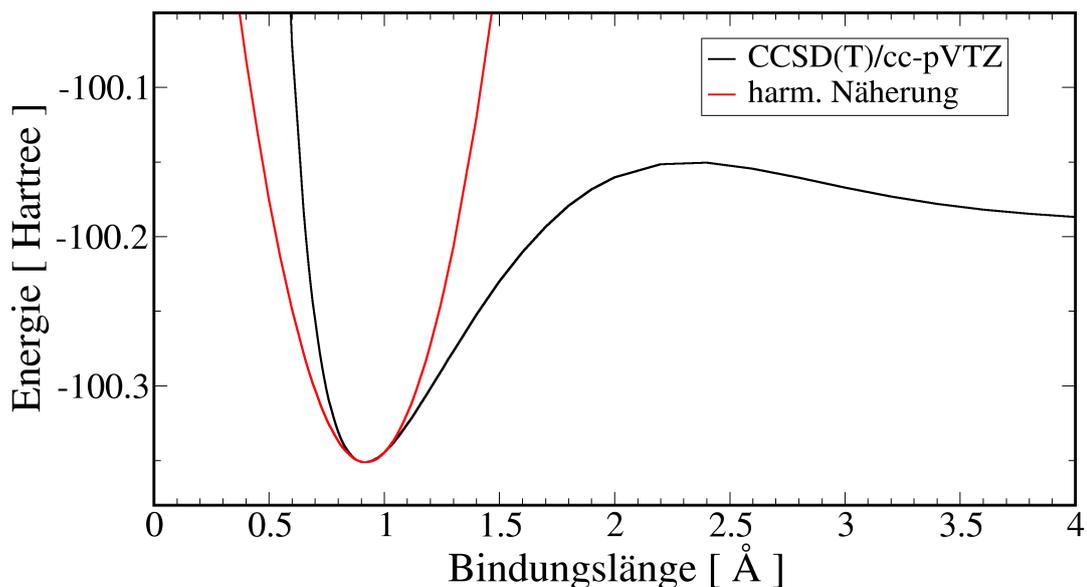


Abbildung 8.1: Potentialkurve von Fluorwasserstoff auf CCSD(T)/cc-pVTZ- Niveau berechnet und die harmonische Näherung am Minimum.

mit der Kraftkonstanten  $k$  berechnet.

Die in der harmonischen Näherung berechneten Schwingungsfrequenzen sind zwar eine gute Näherung, die Fehler aufgrund fehlender Anharmonizität sind allerdings oftmals signifikant<sup>[7]</sup> (siehe Tabelle 8.1) und erlauben keine hochgenauen Vorhersagen. Um die Genauigkeit der Berechnungen zu verbessern gibt es unterschiedliche Ansätze. Zum einen wurde ausgenutzt, dass der Fehler der harmonischen Näherung zwar signifikant, aber meistens systematisch ist. So wurden anhand von Testsystemen Skalierungsfaktoren<sup>[7,176]</sup> für die harmonischen Frequenzen für unterschiedliche Methoden berechnet, wodurch die Genauigkeit zwar zunimmt, aber weiterhin keine zufriedenstellende Werte annimmt. Die heute am weitesten verbreitete Methode ist allerdings die Berücksichtigung von höheren Termen in der Taylorentwicklung zur Beschreibung der Anharmonizität, z.B. durch Störungstheorie (VPT2).<sup>[177–179]</sup> Diese Berechnungen ergeben eine deutlich verbesserte Genauigkeit verglichen zu den harmonischen Werten.

Für einzelne Moleküle in der Gasphase ist es einfach die Schwingungsfrequenzen über einzelne quantenchemische Rechnungen zu erhalten. Für mehrere Moleküle in Clustern oder für Moleküle in Lösung ist dies nicht praktikabel. Daher werden die Schwingungsfrequenzen für größere Systeme oftmals aus der Trajektorie einer AIMD- oder QM/MM-Simulation berechnet. In der Literatur wird in diesem Fall häufig davon gesprochen, dass es sich bei den so berechneten Schwingungsfrequenzen um anharmonische<sup>[168–170]</sup> oder zumindest teilweise anharmonische<sup>[163,166]</sup> Schwingungsfrequenzen handelt. Die Idee dahinter ist, dass sich die Atome in der AIMD-Simulation (oder im QM-Bereich einer QM/MM-Simulation) auf der exakten und damit auch anharmonischen PES (siehe Abbildung 8.1) bewegen und somit auch die Schwingungsfrequenzen

anharmonisch sind.

Um die Anharmonizität zu untersuchen, wurde eine AIMD-Simulation mit 50.000 Schritten ( $\approx 12$  ps) mit Fluorwasserstoff auf CCSD(T)/cc-pVTZ Niveau bei 0.5 K durchgeführt. Die Schwingungsfrequenzen können entweder direkt über die Autokorrelationsfunktion der Bindungslänge von Fluorwasserstoff oder über die des Dipolmoments berechnet werden. In diesem Fall macht es keinen Unterschied, da es nur eine Bindung gibt. Für größere Moleküle wird meistens nur die Autokorrelationsfunktion des Dipolmoments verwendet, da das Dipolmoment über

$$\boldsymbol{\mu} = \int \rho(\mathbf{r}) \cdot \mathbf{r} d^3\mathbf{r} \quad (8.2)$$

mit der Ladungsdichte  $\rho(\mathbf{r})$  von der Struktur abhängt und alle Schwingungen des Moleküls über die Autokorrelationsfunktion des Dipolmoments zugänglich sind. In Abbildung 8.2 ist die zeitliche Entwicklung der Bindungslänge und des Dipolmoments gegenübergestellt. Man kann sehen, dass die Oszillationen jeweils den gleichen Verlauf nehmen und auch die erhaltenen Frequenzen von  $4199.8 \text{ cm}^{-1}$  für das Dipolmoment bzw.  $4199.7 \text{ cm}^{-1}$  für die Bindungslänge gleich sind.

Der Vergleich mit den aus einzelnen quantenchemischen Rechnungen erhaltenen Schwingungsfrequenzen (Tabelle 8.1) zeigt, dass die aus der AIMD-Simulation erhal-

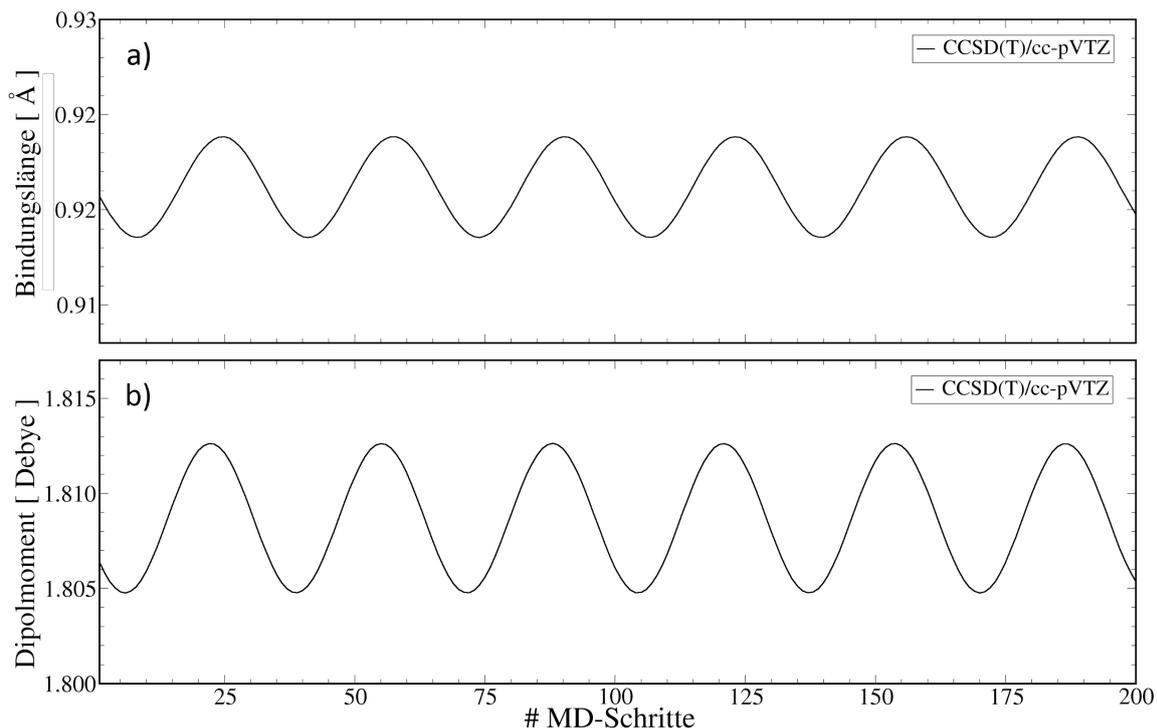


Abbildung 8.2: Zeitlicher Verlauf der Bindungslänge (a) und des Dipolmoments (b) von Fluorwasserstoff in einer AIMD-Simulation bei 0.5 K auf CCSD(T)/cc-pVTZ Niveau durchgeführt.

Tabelle 8.1: Harmonische und anharmonische (VPT2-Näherung) Schwingungsfrequenzen von gasförmigen Fluorwasserstoff auf CCSD(T)/cc-pVTZ Niveau berechnet sowie aus AIMD-Simulationen bei 0.5 K und 300 K (ebenfalls CCSD(T)/cc-pVTZ Niveau) erhaltene Schwingungsfrequenzen verglichen mit dem Experiment.<sup>[180]</sup>

Methode	Exp.	QM harm.	QM anharm.	AIMD [0.5 K]	AIMD [300 K]
Frequenz [ $\text{cm}^{-1}$ ]	3961.4 <sup>[180]</sup>	4193.7	4015.2	4199.8	4190.9

tene Schwingungsfrequenz keine Anharmonizität aufweist. Das war zu erwarten, da die Simulation bei 0.5 K durchgeführt wurde und die Schwingung der Moleküle daher sehr nah am Minimum der PES und damit im harmonischen Bereich stattfindet. Thomas et al.<sup>[166]</sup> berichten, dass der Samplingbereich bei höheren Simulationstemperaturen aufgrund der größeren Amplituden mehr vom anharmonischen Bereich der PES beinhaltet und damit auch die Schwingungsfrequenzen mehr Anharmonizität aufweisen. Es gibt allerdings keine quantitative Aussage darüber, wie viel Anharmonizität bei welchen Simulationstemperaturen erhalten wird. Eine Simulation bei 300 K zeigt, dass für Fluorwasserstoff selbst bei Raumtemperatur nur minimale anharmonische Effekte erhalten werden, da die Schwingungsfrequenz im Vergleich zu 0.5 K nur um  $9 \text{ cm}^{-1}$  verschoben wird. Abbildung 8.3 zeigt, dass die Oszillation der Bindungslänge bei 300 K im Vergleich zu der Simulation bei 0.5 K eine größere Amplitude aufweist. Dies bestätigt, dass der Samplingbereich der PES größer ist und damit auch mehr Anharmonizität erhalten wird.

Um die Temperaturabhängigkeit der Schwingungsfrequenz genauer zu untersuchen,

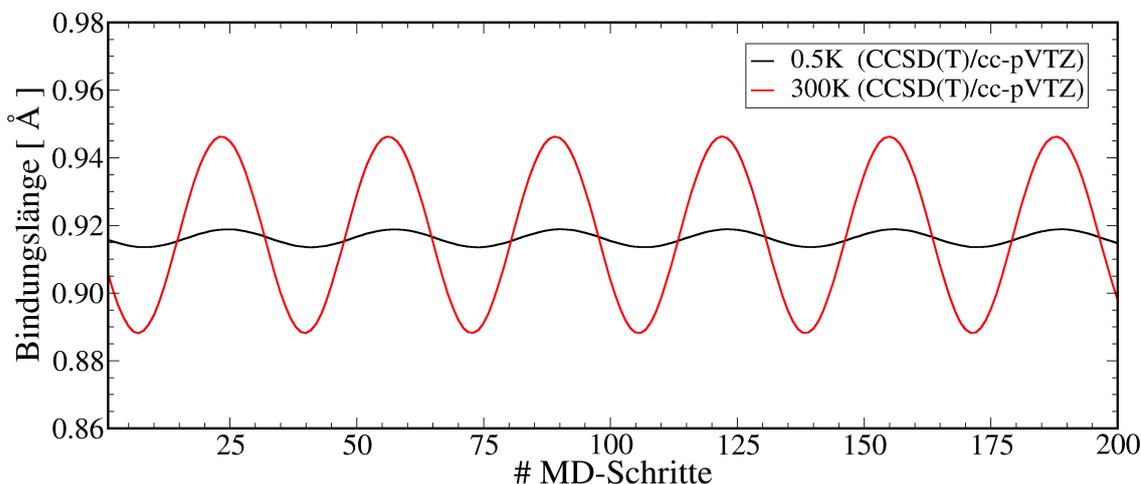


Abbildung 8.3: Zeitlicher Verlauf der Bindungslänge von Fluorwasserstoff in einer AIMD-Simulation bei 0.5 K und 300 K auf CCSD(T)/cc-pVTZ Niveau.

Tabelle 8.2: Schwingungsfrequenzen von gasförmigen Fluorwasserstoff aus AIMD-Simulationen auf CCSD(T)/cc-pVTZ Niveau bei verschiedenen Temperaturen.

Temp.	[K]	0.5	50	150	300	499	996	2477
Frequenz	[ $\text{cm}^{-1}$ ]	4199.8	4198.3	4195.3	4190.9	4184.9	4169.9	4124.9
Temp.	[K]	4909	7291	9625	14131	18410	26171	
Frequenz	[ $\text{cm}^{-1}$ ]	4049.1	3972.4	3894.9	3737.0	3575.1	3235.5	

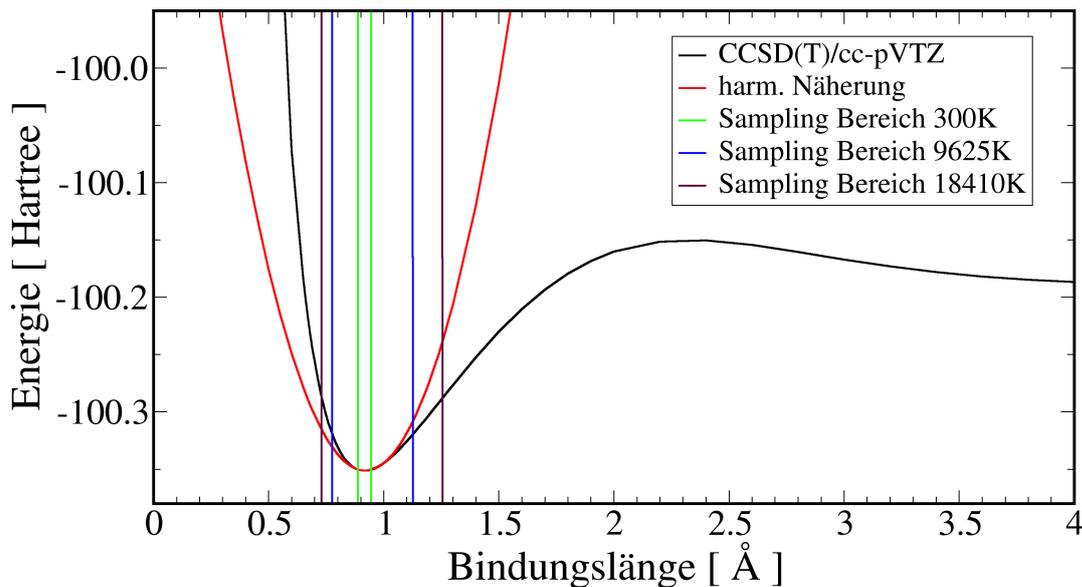


Abbildung 8.4: Potentialkurve von Fluorwasserstoff auf CCSD(T)/cc-pVTZ- Niveau berechnet und die harmonische Näherung am Minimum. Die maximalen Auslenkungen der Fluorwasserstoffbindung bei AIMD-Simulationen bei verschiedenen Temperaturen sind farblich markiert.

wurden weitere Simulationen bei unterschiedlichen Temperaturen bis zu 26171 K durchgeführt. Die daraus erhaltenen Schwingungsfrequenzen sind in Tabelle 8.2 angegeben. Bei immer größeren Simulationstemperaturen verschiebt sich die Schwingungsfrequenz zu immer kleineren Wellenzahlen. Abbildung 8.4 zeigt, wie groß die Auslenkung der Bindung bei verschiedenen Temperaturen ist und welche Teile der PES im Samplingbereich enthalten sind.

Bei 300 K ist nur der Bereich nahe des Minimums, d.h. nur der harmonische Teil der PES, enthalten, während die Schwingung bei 18410 K einen deutlich größerer Bereich der PES einschließt und die Schwingungsfrequenz dadurch zu deutlich kleineren Wellenzahlen verschoben wird.

Allerdings wird die Schwingungsfrequenz von Fluorwasserstoff bei hohen Tempera-

turen deutlich weiter verschoben als es durch die in einer einzelnen quantenchemischen Rechnung erhaltenen anharmonischen Frequenz zu erwarten wäre. Daher stellt sich die Frage, ob es möglich ist aus der beobachteten Temperaturabhängigkeit der Schwingungsfrequenz eine Vorhersage für die experimentelle bzw. anharmonische Schwingungsfrequenz zu erhalten. Da die Temperaturabhängigkeit annähernd linear ist, wie in Abbildung 8.5 zu sehen ist, ist es möglich die Schwingungsfrequenzen für bestimmte Temperaturen zu extrapolieren. Ein Versuch die Temperatur zu bestimmen, bei der die anharmonischen Anteile der Schwingungsfrequenz genau so groß sind, sodass die experimentelle bzw. anharmonische Schwingungsfrequenz erhalten wird, ist die Verwendung der Schwingungstemperatur  $\Theta_{vib}$ .<sup>[175]</sup> Diese wird in der statistischen Thermodynamik verwendet um die Zustandssumme des Systems zu beschreiben<sup>[175]</sup> und ergibt für Fluorwasserstoff

$$\Theta_{vib} = \frac{hc\nu_{HF}}{k_B} \approx 6042.6K, \quad (8.3)$$

mit dem Planckschen Wirkungsquantum  $h$ , der Lichtgeschwindigkeit  $c$ , der Boltzmann-Konstante  $k_B$  und der für 0 K extrapolierten Schwingungsfrequenz  $\nu_{HF} = 4199.8\text{cm}^{-1}$ . Extrapoliert man nun mittels des linearen Fits aus Abbildung 8.5 die Schwingungsfrequenz für 6042.6 K erhält man eine Schwingungsfrequenz von  $4019.5\text{cm}^{-1}$ , was ziemlich gut der anharmonischen Schwingungsfrequenz der VPT2-Näherung entspricht. Die Übereinstimmung mit den Werten der VPT2-Näherung ist hier von

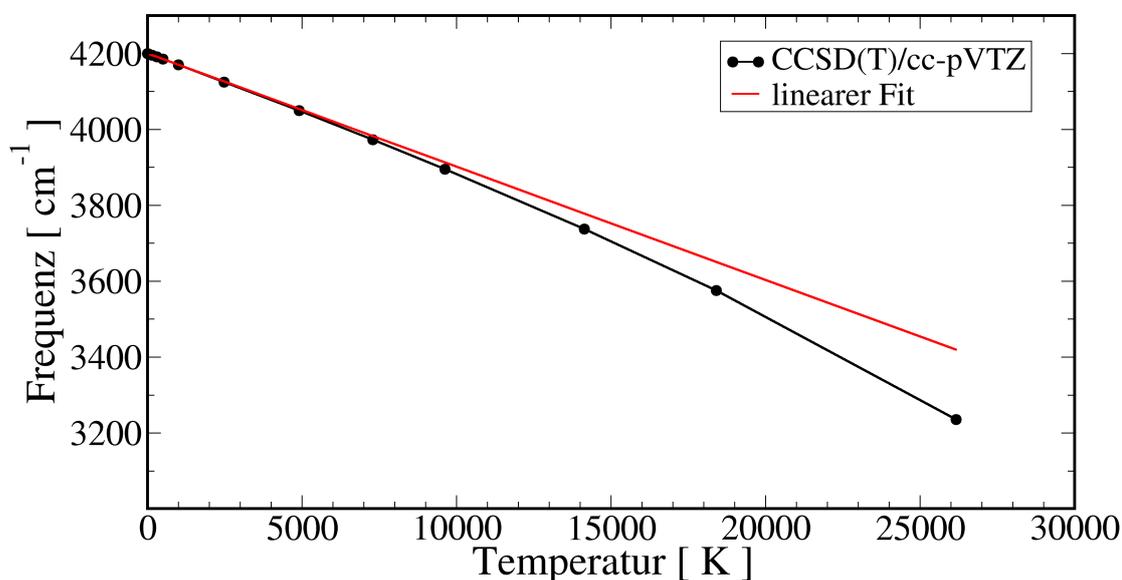


Abbildung 8.5: Temperaturabhängigkeit der Schwingungsfrequenz von Fluorwasserstoff erhalten aus AIMD-Simulationen auf CCSD(T)/cc-pVTZ Niveau.

größerer Relevanz als der Vergleich mit dem Experiment, da Abweichungen bzgl. des Experiments auch aufgrund der Genauigkeit der verwendeten Methode und der Basis (CCSD(T)/cc-pVTZ) zurückzuführen sind und nicht zwingend auf die fehlende Anharmonizität.

Nachdem der erste Versuch der Extrapolation zum Erhalten von anharmonischen Schwingungsfrequenzen am Beispiel von Fluorwasserstoff vielversprechende Ergebnisse geliefert hat, wurde das Verfahren an weiteren Molekülen getestet. Dazu wurden sechs weitere Moleküle mit mehr als zwei Atomen und damit auch mehr als einer Schwingungsfrequenz untersucht. Die untersuchten Moleküle sind Wasser, Wasserstoffperoxid, Kohlenstoffdioxid, Ammoniak, Methan und Ethen. Für jedes Molekül

Tabelle 8.3: Anharmonische Schwingungsfrequenzen für verschiedene Moleküle in Gasphase. Die Schwingungsfrequenzen wurden auf CCSD(T)/cc-pVTZ Niveau berechnet (VPT2-Näherung) bzw. über eine Extrapolation der Schwingungsfrequenzen aus AIMD-Simulationen (CCSD(T)/cc-pVTZ) erhalten. Experimentelle Werte sind zum Vergleich angegeben.

Molekül	HF	CH <sub>4</sub>		CO <sub>2</sub>		
Mode	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	
QM VPT2	4015.2	3018.4	1320.2	2366.9	664.8	
AIMD extrap.	4019.5	2753.1	1286.7	2323.9	656.0	
Experiment	3961 <sup>[180]</sup>	3019 <sup>[181]</sup>	1306 <sup>[181]</sup>	2349 <sup>[181]</sup>	667 <sup>[181]</sup>	
	✓	✗	✓	✓	(✓)	
Molekül	H <sub>2</sub> O			H <sub>2</sub> O <sub>2</sub>		
Mode	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	$\nu_3$ [cm <sup>-1</sup> ]	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	$\nu_3$ [cm <sup>-1</sup> ]
QM VPT2	3774.2	3682.7	1616.5	3636.1	1277.0	319.5
AIMD extrap.	3762.8	3646.7	1591.8	3740.1	1090.1	153.0
Experiment	3756 <sup>[182]</sup>	3657 <sup>[182]</sup>	1595 <sup>[182]</sup>	3608 <sup>[183]</sup>	1266 <sup>[183]</sup>	317 <sup>[183]</sup>
	✓	✓	✓	✗	✗	✗
Molekül	NH <sub>3</sub>					
Mode	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	$\nu_3$ [cm <sup>-1</sup> ]	$\nu_4$ [cm <sup>-1</sup> ]		
QM VPT2	3446.9	3353.5	1645.6	1025.0		
AIMD extrap.	3174.9	3106.2	1630.3	1110.4		
Experiment	3444 <sup>[181]</sup>	3337 <sup>[181]</sup>	1627 <sup>[181]</sup>	950 <sup>[181]</sup>		
	✗	✗	✓	✗		
Molekül	C <sub>2</sub> H <sub>4</sub>					
Mode	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	$\nu_3$ [cm <sup>-1</sup> ]	$\nu_4$ [cm <sup>-1</sup> ]		
QM VPT2	3105.5	2877.9	1451.9	958.2		
AIMD extrap.	2829.1	2548.9	1404.3	950.9		
Experiment	3105.5 <sup>[181]</sup>	2988.7 <sup>[181]</sup>	1443.5 <sup>[181]</sup>	949.3 <sup>[181]</sup>		
	(✗)	(✗)	✓	✓		

wurde je eine AIMD-Simulation bei 0.5 K bzw. 300 K durchgeführt und mittels der damit erhaltenen Schwingungsfrequenzen für jede Normalmode eine harmonische Schwingungsfrequenz (0 K Limit) extrapoliert. Mit dieser harmonischen Schwingungsfrequenz wurde über Gleichung 8.3 die entsprechende Schwingungstemperatur berechnet und anschließend die anharmonische Schwingungsfrequenz extrapoliert. Die Ergebnisse sind in Tabelle 8.3 zusammengefasst. Bei der Bewertung der extrapolierten Werte ist, wie bereits erwähnt, die Übereinstimmung mit den Werten der VPT2-Rechnung ausschlaggebend, welche in den meisten Fällen allerdings auch gut mit dem Experiment übereinstimmen.

Zunächst ist zu erwähnen, dass nicht alle Schwingungsmoden der Moleküle angegeben sind, da einige Moden entartet sind und somit die gleiche Frequenz besitzen. Außerdem sind einige Moden IR-inaktiv und besitzen daher bei 0.5 K eine Intensität von 0. Methan hat z.B. grundsätzlich 9 Schwingungsmoden ( $3 \cdot N - 6 = 9$ , mit  $N = 5$ ), dabei sind aber drei IR-inaktiv und die restlichen sechs bestehen aus zwei jeweils dreifach entarteten Moden. Daher sind in Tabelle 8.3 für Methan nur zwei Schwingungsfrequenzen angegeben.

Für alle Moleküle gilt, dass alle Werte, die aus den AIMD-Simulationen bei 0.5 K erhalten wurden bzw. die zum 0 K Limit extrapoliert wurden, gut mit den Werten der harmonischen Näherung der einzelnen, quantenchemischen Rechnungen übereinstimmen (siehe Anhang C).

Die Extrapolation der Schwingungsfrequenzen funktioniert nur für Fluorwasserstoff und Wasser mit zufriedenstellender Genauigkeit für alle Moden. Bei Methan und Ammoniak hingegen stimmen die extrapolierten Werte nur bei einer von zwei bzw. einer von vier Schwingungsmoden mit den Werten der VPT2-Näherung überein. Bei Wasserstoffperoxid funktioniert die Extrapolation bei keiner der drei Schwingungsmoden. Eine kleine Ausnahme bildet Kohlenstoffdioxid, bei dem es zwar so aussieht, als ob die Extrapolation bei beiden Schwingungsmoden gut funktioniert hat, allerdings ist der extrapolierte Wert von  $\nu_2$  schlechter als der harmonische Wert (siehe Anhang C), weshalb diese Mode aus der Bewertung herausgenommen wurde. Bei Ethen stimmen die extrapolierten Werte bei zwei von vier Moden gut mit den Werten der VPT2-Näherung überein. Die anderen beiden Werte stimmen zwar nicht überein, aber da die Simulationszeit bei Ethen kürzer als bei den anderen Molekülen war und die Bestimmung der Schwingungsfrequenzen bei der Simulation bei 300 K aufgrund der relativ nahen Banden schwierig war, wurden diese Moden ebenfalls aus der Bewertung herausgenommen.

Insgesamt ergibt die Extrapolation nur bei 9 von 16 Moden eine gute Übereinstimmung mit den VPT2-Werten. Eine zuverlässige Vorhersage der anharmonischen Schwingungsfrequenzen ist somit nicht möglich. Eine Erfolgsquote von 56% ist aber eine gute Grundlage um die Methode weiter zu verbessern und zu untersuchen, weshalb eine Extrapolation in manchen Fällen nicht funktioniert.

Ein möglicher Grund, weshalb die Methode nicht in allen Fällen funktioniert ist, dass das System für eine sinnvolle Extrapolation sehr gut equilibriert sein muss.<sup>[163]</sup> Das heißt, dass die Energie, welche in das System gesteckt wird, um eine Temperatur von 0.5 K zu erreichen, gleichmäßig auf alle Schwingungsmoden verteilt sein muss.

Andernfalls funktioniert die Extrapolation, welche für jede Schwingungsmode einzeln vorgenommen wird, nicht, da nicht die exakte Temperatur der jeweiligen Schwingung bekannt ist, sondern nur die des Gesamtsystems.

Zudem könnte es sinnvoll sein, andere Temperaturen zum Extrapolieren zu verwenden, da bei 0.5 K einige Moden nicht zu sehen sind (IR-inaktiv) und bei 300 K einige Moden bereits relativ breit sind und die Bestimmung der Schwingungsfrequenz dadurch ungenau wird.

Ein weiterer Grund könnte die lineare Extrapolation sein. Auch wenn die berechneten Schwingungstemperaturen bei allen Moden im Bereich von 1000 – 6000 K liegt und die Temperaturabhängigkeit in diesem Temperaturbereich größtenteils linear sein sollte, könnte ein genauerer Fit, bei dem die Werte von drei oder noch mehr verschiedenen Simulationstemperaturen berücksichtigt werden, bessere Resultate ergeben.

Zuletzt ist zu diskutieren, weshalb eine Simulation (oder Extrapolation) bei sehr hohen Temperaturen überhaupt notwendig ist, um die richtigen anharmonischen Frequenzen zu erhalten, obwohl die experimentellen Schwingungsfrequenzen bei deutlich geringeren Temperaturen gemessen werden (Raumtemperatur oder niedriger<sup>[180,183]</sup>) und sich die Atome auf der richtigen PES bewegen. Es handelt sich zwar um die richtige PES auf der sich die Atome bewegen, allerdings bewegen sich die Atome aufgrund des Lösen der Newtonschen Bewegungsgleichungen als klassische Teilchen auf dieser PES. Das bedeutet, dass sich die Atome bei (nahezu) 0 K im Minimum befinden und sich nur minimal bewegen. In einem realen System muss allerdings die quantenmechanische Natur der Teilchen berücksichtigt werden. Aufgrund der Nullpunktsenergie<sup>[7]</sup> (ZPE) liegt die Energie des Systems höher als es eine klassische Beschreibung für das Minimum der PES vorhersagt. Daher muss dem System im Fall der AIMD-Simulationen Energie hinzugefügt werden, um den quantenmechanischen Effekt der ZPE zu imitieren.

## 8.2 Schwingungsfrequenzen aus QM/MM-Simulationen

Die Berechnung von Schwingungsfrequenzen aus AIMD-Simulationen von einzelnen Molekülen in der Gasphase ist zwar konzeptionell interessant, hat jedoch nur begrenzt Anwendungsmöglichkeiten, da der Aufwand im Vergleich zu einzelnen quantenchemischen Rechnungen zur Berechnung der Schwingungsfrequenzen deutlich größer ist. Von größerer Relevanz sind allerdings Simulationen an Systemen wie Lösungen, da der Lösungsmiteleinfluss durch einzelne quantenchemische Rechnungen nicht zu untersuchen ist. Mit QM/MM-Simulationen, in denen der QM-Bereich aus einem einzelnen Molekül besteht, welches von tausenden anderen Molekülen im MM-Bereich umgeben ist, kann der Einfluss der Lösungsumgebung untersucht werden. Im Folgenden werden zwei Methoden zur Berechnung der Schwingungsfrequenzen aus QM/MM-Simulationen untersucht und der jeweilige Anteil der Anharmonizität diskutiert.

### 8.2.1 Berechnung der Schwingungsfrequenzen aus QM/MM-Trajektorien

Ebenso wie bei AIMD-Simulationen können die Schwingungsfrequenzen aus den QM/MM-Trajektorien bzw. aus der Autokorrelationsfunktion des Dipolmoments des QM-Bereichs berechnet werden.<sup>[163,170]</sup> Auch hierbei stellt sich die Frage, wie viel Anharmonizität in diesen Schwingungsfrequenzen enthalten ist. Ein analoges Vorgehen zur Extrapolation der Schwingungsfrequenzen wie in Abschnitt 8.1 ist aus folgenden Gründen allerdings nicht möglich:

- Bei der Untersuchung der Schwingungsfrequenzen von Molekülen in Lösung spielen hauptsächlich zwei Faktoren eine Rolle. Zum einen der Einfluss des Lösungsmittels auf die Struktur des Moleküls und damit auch auf die Schwingungsfrequenzen und zum anderen der Anteil der Anharmonizität. Allerdings lassen sich beide Effekte nicht unabhängig voneinander untersuchen, da die Lösungsumgebung einen direkten Einfluss auf die PES besitzt und damit auch auf die Anharmonizität der PES.
- Zur Extrapolation der Schwingungsfrequenzen sind zwei oder mehrere Simulationen bei verschiedenen Temperaturen notwendig. Anders als bei einzelnen Gasmolekülen sind die Eigenschaften des Systems zum Teil stark von der Temperatur abhängig. Dynamische Lösungsmittelleffekte sind bei Temperaturen nahe 0 K im Prinzip nicht vorhanden. Bei höheren Temperaturen steigt die kinetische Energie der Moleküle und die Struktur des Lösungsmittels geht verloren.<sup>[184,185]</sup> Daher sind für die Extrapolation, wenn überhaupt, nur Simulationen in einem sehr kleinen Temperaturbereich sinnvoll.
- Der Vergleich mit anharmonischen Schwingungsfrequenzen, welche mit einzelnen quantenchemischen Rechnungen berechnet wurden, ist nicht oder nur

bedingt möglich. Die einzige Möglichkeit dazu ist die Berechnung von statischen Schwingungsfrequenzen aus mehreren Snapshots aus der QM/MM-Trajektorie und die Mittelung darüber (siehe Abschnitt 8.2.2), was allerdings recht aufwendig ist.

Aus den Erkenntnissen aus Abschnitt 8.1 kann allerdings geschlossen werden, dass bei der Simulation von Lösungen bei 300 K in gewissem Maße anharmonische Anteile berücksichtigt werden. Zudem ist es wahrscheinlich, dass der Anteil der Anharmonizität größer sein wird als im Beispiel mit Fluorwasserstoff, da das Lösungsmittel die PES abflacht und somit weniger Energie nötig ist, um mit der Atombewegung in die anharmonischen Teile der PES zu gelangen.

Das Schwingungsspektrum von flüssigem Wasser, welches aus der Trajektorie einer QM/MM-Simulation erhalten wurde, ist in Abbildung 8.6 dargestellt. Zum Vergleich ist ebenfalls das Schwingungsspektrum von gasförmigem Wasser bei 300 K und das experimentelle<sup>[186]</sup> Schwingungsspektrum angegeben. Im Vergleich zum gasförmigen Spektrum sieht man eine deutliche Verschiebung der Streckerschwingungen von Wasser zu kleineren Wellenzahlen. Ebenso ist eine deutliche Verbreiterung der Schwingungsbanden zu erkennen. In der Annahme, dass der Anteil an Anharmonizität im QM/MM-Spektrum genauso groß ist wie im AIMD-Spektrum, sind diese Änderungen auf die Lösungsmittelleffekte zurückzuführen. Verglichen mit dem Experiment liegen die Banden der Streckerschwingungen allerdings bei zu hohen Wellenzahlen, wobei die Verbreiterung der Schwingungsbanden ähnlich ausgeprägt ist. Die Biegeschwingung

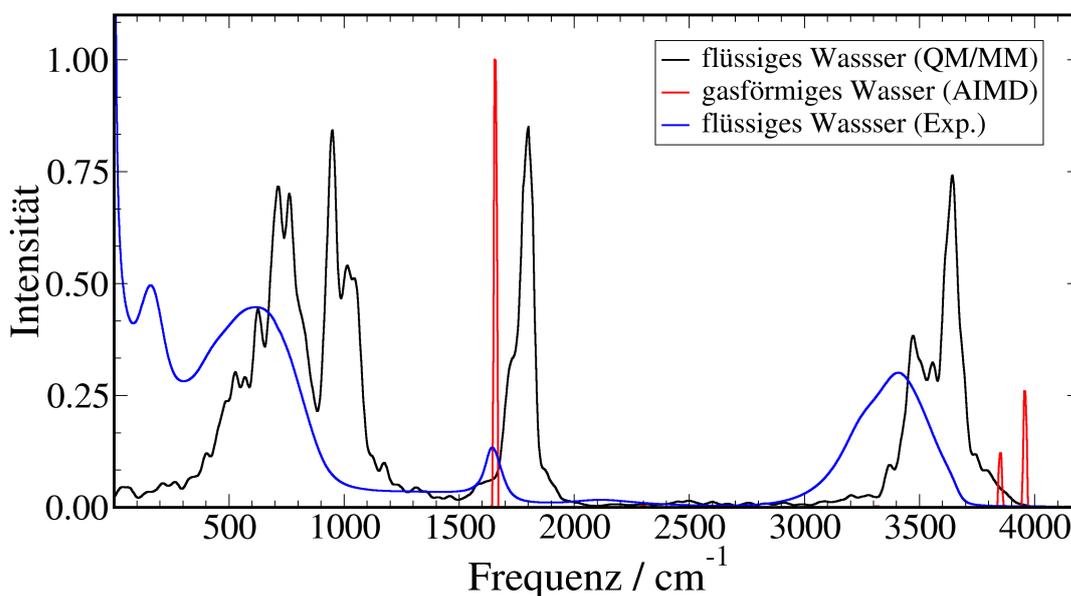


Abbildung 8.6: Schwingungsspektrum von gasförmigen und flüssigem Wasser erhalten aus AIMD- und QM/MM-Simulationen bei jeweils 300 K auf CCSD(T)/cc-pVTZ Niveau verglichen mit dem experimentellen Schwingungsspektrum.<sup>[186]</sup>

sowie auch die Translations- und Librationsbanden im niedrigen Wellenzahlenbereich liegen ebenfalls bei etwas zu großen Wellenzahlen. Diese in etwa konstante Verschiebung des QM/MM-Spektrums bezüglich des experimentellen Spektrums deutet auf die fehlende Anharmonizität hin.

Ein weiteres System, wofür die Schwingungsfrequenzen bestimmt wurden, ist Methanol. In Tabelle 8.4 sind die Schwingungsfrequenzen für gasförmiges Methanol angegeben, welche wie in Abschnitt 8.1 über AIMD-Simulationen auf MP2/cc-pVTZ Niveau berechnet bzw. extrapoliert wurden. In Abbildung 8.7 ist das über eine MTS-QM/MM-Simulation bei 300 K erhaltenes Schwingungsspektrum mit experimentellen Werten<sup>[187]</sup> gegenübergestellt. Die MTS-Simulation wurde auf MP2/cc-pVTZ/pw-DFT/PBE Niveau mit einem MTS-Faktor  $N = 10$  durchgeführt. Bei der Berechnung der elektrostatischen QM/MM-Wechselwirkungen wurde ein long-range Cutoff von 40 a.u. verwendet.

Bei der Extrapolation der anharmonischen Schwingungsfrequenzen für gasförmiges Methanol stimmen die Werte bei 7 von 10 Schwingungsmoden mit den VPT2-Werten überein. Zudem liegen die Schwingungsmoden um  $3000 \text{ cm}^{-1}$  sehr nah beieinander, sodass die Extrapolation bzw. die Bestimmung der Schwingungsfrequenzen schwierig war. Trotz der dicht beieinander liegenden Moden und der, im Vergleich zu den in Abschnitt 8.1 untersuchten Molekülen, größeren Anzahl von Schwingungsmoden, funktioniert die Extrapolation bei Methanol sehr gut.

Bei dem aus einer QM/MM-Simulation erhaltenen Schwingungsspektrum für flüssiges Methanol sind die gleichen Trends wie für das Schwingungsspektrum von Wasser zu erkennen. Die Form des Spektrums stimmt relativ gut mit dem experimentellen Spektrum überein. Die deutlich verbreiterte Form der OH-Streckschwingungsbande

Tabelle 8.4: Schwingungsfrequenzen für Methanol in Gasphase. Die Schwingungsfrequenzen wurden auf MP2/cc-pVTZ Niveau berechnet (VPT2-Näherung) bzw. über eine Extrapolation der Schwingungsfrequenzen aus AIMD-Simulationen (MP2/cc-pVTZ) erhalten. Experimentelle Werte<sup>[181]</sup> sind zum Vergleich angegeben.

Mode	$\nu_1$ [ $\text{cm}^{-1}$ ]	$\nu_2$ [ $\text{cm}^{-1}$ ]	$\nu_3$ [ $\text{cm}^{-1}$ ]	$\nu_4$ [ $\text{cm}^{-1}$ ]	$\nu_5$ [ $\text{cm}^{-1}$ ]
QM VPT2	3727.0	3036.5	3015.4	3114.4	1499.7
AIMD extrap.	3776.1	2961.7	2566.1	2300.5	1418.7
Experiment <sup>[181]</sup>	3681	3000	2960	2844	1477
	✓	✓	✗	✗	✓
Mode	$\nu_6$ [ $\text{cm}^{-1}$ ]	$\nu_7$ [ $\text{cm}^{-1}$ ]	$\nu_8$ [ $\text{cm}^{-1}$ ]	$\nu_9$ [ $\text{cm}^{-1}$ ]	$\nu_{10}$ [ $\text{cm}^{-1}$ ]
QM VPT2	1334.8	1171.8	1097.2	1032.6	210.4
AIMD extrap.	1369.5	1173.1	1095.0	1066.4	314.5
Experiment <sup>[181]</sup>	1345	1165	1060	1033	200
	✓	✓	✓	✓	✗

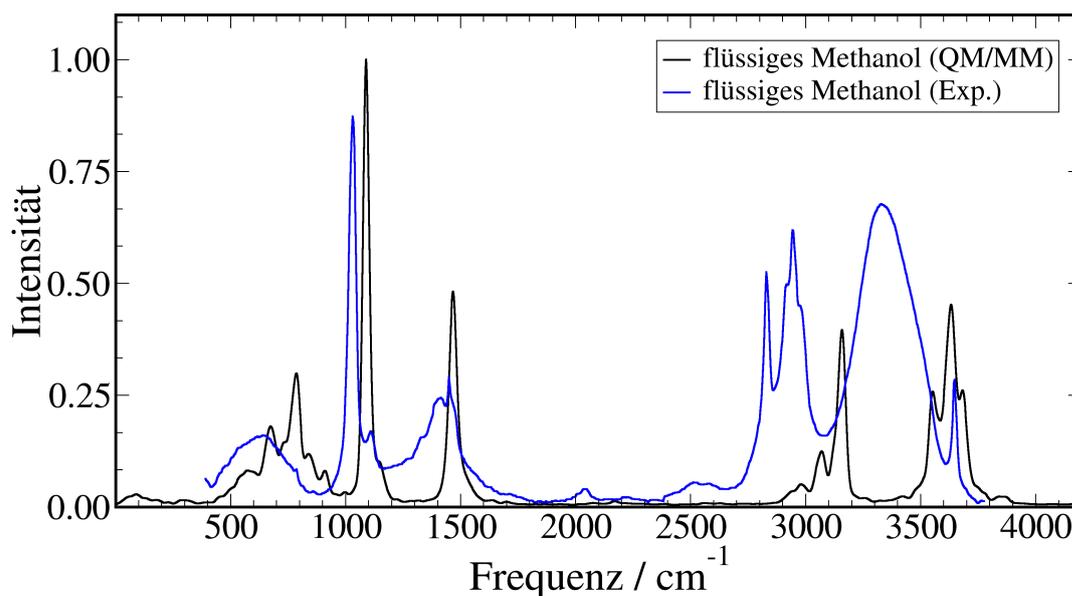


Abbildung 8.7: Schwingungsspektrum von flüssigem Methanol erhalten aus einer QM/MM-Simulation bei 300 K auf MP2/cc-pVTZ Niveau verglichen mit dem experimentellen Schwingungsspektrum.<sup>[187]</sup>

um  $3500\text{ cm}^{-1}$  ist gut reproduziert. Ebenso ist die starke, aber schmale CO-Streckschwingungsbande bei  $1100\text{ cm}^{-1}$  in beiden Spektren ähnlich ausgeprägt. Allerdings sind alle Banden wiederum konstant zu höheren Wellenzahlen als im experimentellen Spektrum verschoben.

Beide Schwingungsspektren, welche direkt aus QM/MM-Simulationen erhalten wurden, stimmen in der Form gut mit dem Experiment überein, allerdings liegen ausnahmslos alle Banden bei etwas zu hohen Wellenzahlen. Die Untersuchungen an einzelnen gasförmigen Molekülen in Abschnitt 8.1 sowie an Methanol (Tabelle 8.4) zeigen, dass bei AIMD-Simulationen bei Temperaturen nahe 0 K in etwa die gleichen Schwingungsfrequenz erhalten werden, wie bei einzelnen harmonischen, quantenchemischen Rechnungen. Auch bei AIMD-Simulationen bei etwas höheren Temperaturen, wie 300 K, werden nur geringfügige anharmonische Effekte erhalten und die Schwingungsfrequenzen liegen sehr nah an den rein harmonischen Werten. Da die anharmonischen Schwingungsfrequenzen der VPT2-Näherung der gasförmigen Moleküle gut mit den entsprechenden experimentellen Werten übereinstimmen, ist es unwahrscheinlich, dass die deutlichen Abweichungen in den simulierten Schwingungsspektren von flüssigem Wasser bzw. Methanol im Vergleich zum Experiment auf einen Fehler der Methode bzw. des Basissatzes zurückzuführen sind. Vielmehr ist anzunehmen, dass die Abweichungen auf die fehlende Anharmonizität zurückzuführen ist.

Eine weitere Möglichkeit ist, dass die Abweichungen durch eine nicht ausreichende Beschreibung der Lösungsumgebung durch den QM/MM-Ansatz bzw. durch eine nicht ausreichende Qualität der klassischen Kraftfelder verursacht werden. Um

dies auszuschließen müssten AIMD-Simulationen an den jeweiligen Systemen durchgeführt werden, was aufgrund der Größe der Systeme aber derzeit nicht möglich ist. In Kapitel 5 wurde allerdings gezeigt, dass die Polarisierung des QM-Systems durch die Lösungsumgebung im QM/MM-Ansatz qualitativ und quantitativ gut beschrieben wird. Daher ist es unwahrscheinlich, dass die Abweichungen in den Spektren darauf zurückzuführen sind.

### 8.2.2 Berechnung der Schwingungsfrequenzen aus Snapshots einer QM/MM-Trajektorie

Eine weitere Möglichkeit zur Berechnung von Schwingungsfrequenzen in QM/MM-Simulationen ist die Benutzung von Snapshots einer Trajektorie. Dieses Vorgehen ist allerdings deutlich aufwändiger als die Berechnung über die Fouriertransformation der Dipolmomentautokorrelationsfunktion, da sich das QM-System zur Berechnung der Schwingungsfrequenzen im Minimum befinden muss. Daher muss für jeden Snapshot eine Geometrieminimierung in Anwesenheit der MM-Atome durchgeführt werden. Anschließend können für diese Geometrien die Schwingungsfrequenzen berechnet werden. Durch eine Mittelung über alle Snapshots werden die Schwingungsfrequenzen und eine zugehörige Standardabweichung erhalten, welche Aufschluss über die Breite der jeweiligen Schwingungsbande gibt. Ebenso wie für die Geometrieoptimierung gilt für die Berechnung der Schwingungsfrequenzen, dass die Ladungen der MM-Atome in der Berechnung berücksichtigt werden müssen. Für die Berechnung von harmonischen Schwingungsfrequenzen bedeutet dies, dass die Berechnung der zweiten Ableitung der Energie nach den Kernkoordinaten in Anwesenheit der MM-Atome implementiert werden mussten<sup>a</sup> (siehe Ref. 188).

Tabelle 8.5: Schwingungsfrequenzen für flüssiges Wasser. Die Schwingungsfrequenzen wurden über die Dipolmomentautokorrelationsfunktion und über einzelne Snapshotrechnungen einer QM/MM-Simulation bei 300 K auf HF/cc-pVTZ Niveau berechnet.

Methode	$\nu_1$ [cm <sup>-1</sup> ]	$\nu_2$ [cm <sup>-1</sup> ]	$\nu_3$ [cm <sup>-1</sup> ]
QM/MM Dipolmoment	3810.0	3719.6	1891.4
QM/MM Snapshots	3840.7 ± 106.4	3691.3 ± 73.8	1890.6 ± 29.0

Für die Berechnung der Schwingungsfrequenzen wurde alle 1000 Schritte einer 30000-Schritte QM/MM-Simulation von flüssigem Wasser ein Snapshot erstellt. Da die Berechnung der Schwingungsfrequenzen in Anwesenheit der MM-Atome bisher nur für die HF-Methode implementiert wurde,<sup>[188]</sup> wurde die QM/MM-Simulation auf HF/cc-pVTZ Niveau durchgeführt. Für diese 30 Snapshots wurde die Geometrie des QM-Wassermoleküls optimiert<sup>b</sup> und die harmonischen Schwingungsfrequenzen berechnet. Die gemittelten Werte und die jeweiligen Standardabweichungen sind in Tabelle 8.5 zusammen mit den über die Dipolmomentautokorrelationsfunktion erhaltenen Schwingungsfrequenzen derselben QM/MM-Simulation angegeben. In Abbildung 8.8 sind die jeweils erhaltenen Schwingungsspektren abgebildet. Um aus den über die Snapshots erhaltenen Schwingungsfrequenzen ein Spektrum zu erhalten,

<sup>a</sup> Für die Berechnung der harmonischen Schwingungsfrequenzen werden nur die elektrostatischen Beiträge der QM/MM-Wechselwirkungen bei der Berechnung der Hessematrix berücksichtigt.

<sup>b</sup> Für die Geometrieoptimierungen wurde jeweils ein Konvergenzkriterium von 10<sup>-8</sup> a.u. für den größten Gradienten verwendet.

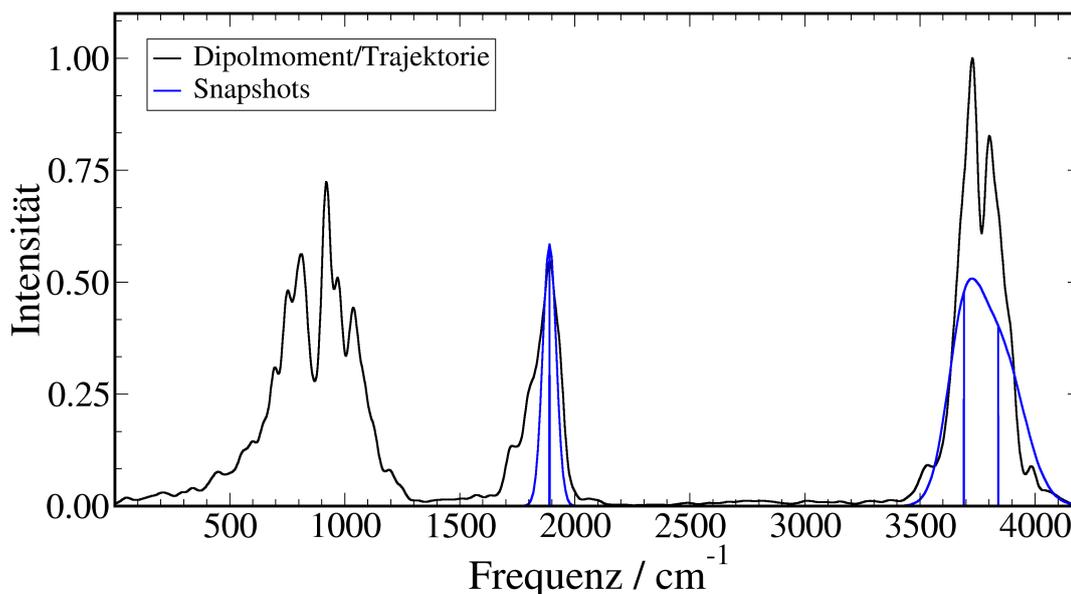


Abbildung 8.8: Schwingungsspektrum von flüssigem Wasser erhalten über die Dipolmomentautokorrelationsfunktion und über einzelne Snapshotrechnungen einer QM/MM-Simulation bei 300 K auf HF/cc-pVTZ Niveau. Die Translation- und Librationsbanden unterhalb von  $1000\text{ cm}^{-1}$  können durch die harmonische Näherung der Snapshotrechnungen nicht berechnet werden.

wurden aus den gemittelten Werten und den jeweiligen Standardabweichungen die Dichtefunktionen der Gaußverteilungen gebildet und übereinander gelegt. Die Intensität des Spektrums wurde so skaliert, dass die Biegeschwingungen die gleiche Intensität besitzen.

Wie in Tabelle 8.5 und Abbildung 8.8 zu sehen ist, ergeben beide Methoden sehr ähnliche Schwingungsfrequenzen und auch die Form der Spektren ist gleich. Gerade die große Standardabweichung der Streckschwingungen und die dadurch verbreiterte Form und Überlagerung der Banden ist bei beiden Methoden ähnlich ausgeprägt.

Bei der Berechnung der Schwingungsfrequenzen aus den Snapshots der QM/MM-Trajektorie wurde die harmonische Näherung verwendet. Ein weiterer Schritt wäre es, für die Snapshots die anharmonischen Frequenzen zu berechnen und zu untersuchen, ob die so erhaltenen Spektren im Vergleich zum Experiment keine Verschiebung zu höheren Wellenzahlen aufweisen.

## 8.3 Zusammenfassung

In diesem Kapitel wurden verschiedene Methoden zur Berechnung von Schwingungsfrequenzen aus AIMD- und QM/MM-Simulationen vorgestellt und untersucht, wie viel Anharmonizität die Schwingungsfrequenzen beinhalten.

Dazu wurde zunächst die Temperaturabhängigkeit der Schwingungsfrequenzen für Fluorwasserstoff untersucht und eine Methode entwickelt, mit der aus den Schwingungsfrequenzen, die aus AIMD-Simulationen bei zwei verschiedenen Temperaturen (0.5 und 300 K) erhalten wurden, die anharmonischen Schwingungsfrequenzen extrapoliert werden können. Diese Methode wurde für verschiedene Moleküle angewendet und es konnte für 56% der Schwingungsmoden die anharmonische Frequenz extrapoliert werden. Als mögliche Gründe, weshalb die Methode nur bei ca. der Hälfte der Schwingungsmoden funktioniert, wurde eine nicht gleichmäßige Verteilung der Energie auf alle Schwingungsmoden und eine nicht ausreichende Qualität der linearen Extrapolation diskutiert.

Anschließend wurden am Beispiel von flüssigem Wasser und Methanol zwei Methoden vorgestellt, wie ein Schwingungsspektrum aus QM/MM-Simulationen berechnet werden kann. Bei der einen Methode wird das Schwingungsspektrum über die Dipolmomentautokorrelationsfunktion berechnet. Bei der anderen werden für mehrere Snapshots einer QM/MM-Trajektorie die Schwingungsfrequenzen über statische quantenchemische Rechnungen berechnet und gemittelt. Beide Methoden ergeben die gleichen Ergebnisse.

Bei allen durchgeführten Simulationen wurde gezeigt, dass bei typischen Simulationstemperaturen von bis zu 300 K hauptsächlich harmonische Schwingungsfrequenzen erhalten werden und nur minimale anharmonischen Effekte zu erkennen sind.



## 9 Limitationen des Interfaces und Ausblick

Das Interface ermöglicht die Verwendung von verschiedenen quantenchemischen Methoden in AIMD- und QM/MM-Simulationen, darunter auch die hochgenauen CC-Methoden. Die Anwendbarkeit auf Systeme unterschiedlicher Größe wurde in dieser Arbeit gezeigt und der Aufwand zur Berechnung der elektrostatischen QM/MM-Wechselwirkungen kann durch die Verwendung eines long-range-Cutoffs deutlich verringert werden. Allerdings hängt der benötigte Rechenaufwand nicht nur von der Anzahl an Atomen im MM-Bereich ab, sondern auch maßgeblich von der Anzahl an Atomen im QM-Bereich bzw. von der Anzahl an Basisfunktionen. Dabei skaliert der Rechenaufwand exponentiell mit der Größe des QM-Systems. Für die HF-Methode skaliert der Rechenaufwand mit  $\mathcal{O}(N^4)$  und für die CCSD(T)-Methode schon mit  $\mathcal{O}(N^7)$  mit  $N$  der Anzahl an Basisfunktionen. Dies limitiert die Anwendbarkeit des Interfaces auf Systeme mit (sehr) kleinem QM-Bereich.

Zudem stellt sich die Frage, welcher der geschwindigkeitsbestimmende Teil in einer QM/MM-Simulation ist und ob das für verschiedene Systemgrößen unterschiedlich ist. Um diese Frage zu untersuchen wurde die benötigte Zeit für die Berechnung der short-range- und long-range-Wechselwirkungen sowie für die Berechnung des QM-Systems für unterschiedliche QM-Bereiche bestimmt. Dazu wurde das in Kapitel 7.3 verwendete System mit Poly- $\beta$ -alanin verwendet und einzelne Rechnungen auf HF/cc-pVDZ Niveau für fünf unterschiedlich große QM-Bereiche durchgeführt. Die Anzahl an Atomen des Poly- $\beta$ -alanins in den QM-Bereichen beträgt 4, 10, 17, 23 und 51. Die Anzahl der short-range Atome beträgt in allen Rechnungen 3500, während

Tabelle 9.1: Benötigte Rechenzeit für QM/MM-Rechnungen auf HF/cc-pVDZ Niveau mit unterschiedlich großen QM-Bereichen eines Systems mit Poly- $\beta$ -alanin in Methanol.

#BF	Atome je Bereich:				Sekunden benötigt für:			
	QM	MM	SR	LR	Gesamt	SR	LR	QM
38	4	19117	3500	15621	6.15	4.67	0.19	1.29
86	10	19115	3500	15625	39.76	19.37	0.73	19.66
157	17	19110	3500	15627	268.80	62.83	3.47	202.50
205	23	19108	3500	15631	639.00	102.33	7.68	528.99
480	51	19092	3500	15543	15809.60	543.53	154.78	15111.29

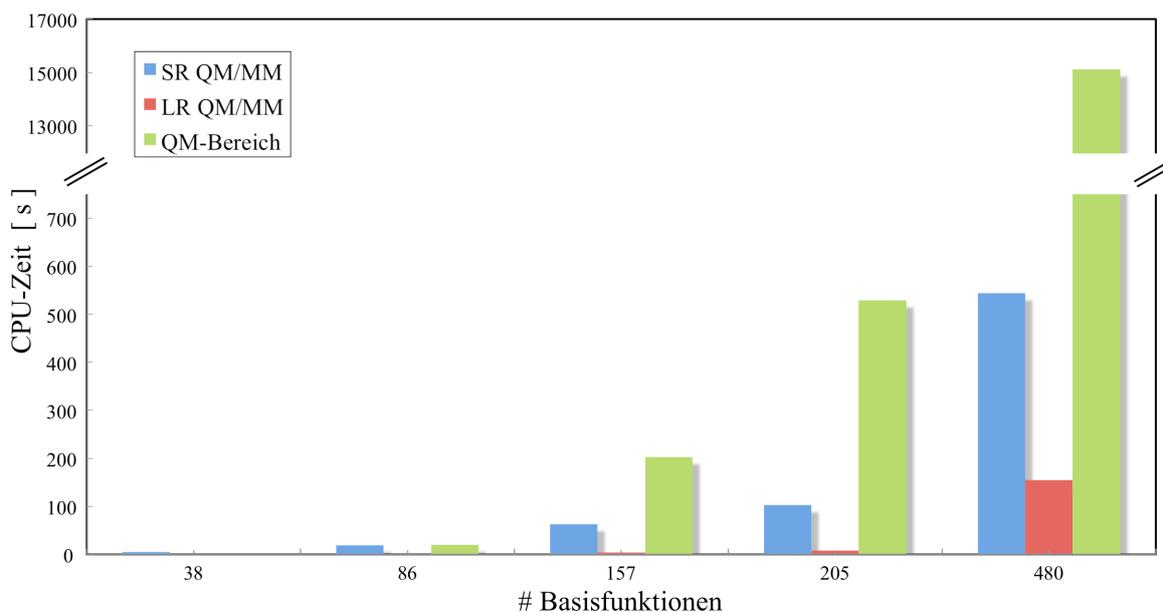


Abbildung 9.1: Benötigte Rechenzeit für QM/MM-Rechnungen auf HF/cc-pVDZ Niveau in Abhängigkeit von der Anzahl an Basisfunktionen im QM-Bereich.

sich der Rest der Atome in der long-range-Region befindet. Alle Rechnungen wurden auf einem Intel Xeon X5675 CPU mit 3.07 GHz durchgeführt.

Die benötigten Rechenzeiten sind in Tabelle 9.1 und Abbildung 9.1 angegeben. Für den kleinsten QM-Bereich mit nur 4 Atomen und 38 Basisfunktionen ist der geschwindigkeitsbestimmende Schritt die Berechnung der short-range-Wechselwirkungen, welcher das 3.5-fache der Zeit zur Berechnung des QM-Systems benötigt. Die Zeit zur Berechnung der long-range-Wechselwirkungen ist vernachlässigbar klein. Bei dem nächst größeren QM-Bereich mit 10 Atomen bzw. 86 Basisfunktionen benötigt die Berechnung der short-range-Wechselwirkungen mit knapp 20 s schon genauso lange wie die Berechnung des QM-Systems. Die Zeit für die long-range-Wechselwirkungen ist mit unter einer Sekunde erneut vernachlässigbar. Für die Rechnungen mit den QM-Bereichen mit 17 und 23 QM-Atomen bzw. 157 und 205 Basisfunktionen benötigt die Berechnung des QM-Bereichs schon deutlich länger als die Berechnung der short-range-Wechselwirkungen. Am größten ist der Unterschied für den größten QM-Bereich mit 51 Atomen bzw. 480 Basisfunktionen. Hier benötigt die Berechnung des QM-Bereichs mit 15111 s (4 h 20 min) das ca. 28-fache der benötigten Zeit zur Berechnung der short-range-Wechselwirkungen. Die Berechnung der long-range-Wechselwirkungen benötigt für alle Rechnungen am wenigsten Zeit und ist im Vergleich zu der Gesamtzeit immer vernachlässigbar.

Die unterschiedliche Skalierung der benötigten Rechenzeiten ist darin begründet, dass die Berechnung der QM/MM-Wechselwirkungen nur von der Anzahl an Einelektronenintegralen abhängt, die mit  $\mathcal{O}(N^2)$  skaliert, während die Zahl der Zweielek-

---

tronenintegrale mit  $\mathcal{O}(N^4)$  skaliert. Dadurch wird die Berechnung des QM-Systems schon bei kleinen bis mittleren QM-Systemen der geschwindigkeitsbestimmende Schritt. Für genauere Methoden, wie MP2 ( $\mathcal{O}(N^5)$ ) oder CCSD(T) ( $\mathcal{O}(N^7)$ ), ist dieses Verhalten noch deutlich ausgeprägter und die Berechnung des QM-Systems wird schnell der limitierende Faktor zur Durchführung der QM/MM-Simulation.

Das bedeutet, dass die Verwendung von long-range-Wechselwirkungen zwar für kleine QM-Systeme eine deutliche Reduktion der benötigten Rechenzeit ermöglicht, bei größeren QM-Systemen aber an Wirkung verliert. Ebenso kann durch die Verwendung des MTS-Algorithmus zwar die Anzahl an teuren QM-Rechnungen reduziert werden, allerdings ist eine QM/MM-Simulation über eine größere Anzahl an Zeitschritten nicht möglich, wenn diese teuren QM-Rechnungen 4 h oder länger dauern. Daher ist es für die Anwendbarkeit für Systeme mit großem QM-Bereich notwendig die benötigte Zeit zur Berechnung des QM-Bereichs weiter zu reduzieren.

Eine Methode zur Reduktion der Rechenzeit ist die Verwendung der Cholesky-Zerlegung für die Zweielektronenintegrale.<sup>[189,190]</sup> Diese wurde vor Kurzem für CAS-SCF<sup>[191]</sup> und für die Berechnung von molekularen Eigenschaften im Magnetfeld für MP2<sup>[192,193]</sup> im Programmpaket CFOUR implementiert und auch für die Berechnung von Gradienten gibt es bereits Implementierungen.<sup>[194]</sup> Durch die Verwendung der Cholesky-Zerlegung lässt sich die Rechenzeit und der benötigte Cache-Speicher verringern. Zudem erlaubt es eine einfache Parallelisierung des Programmcodes, was für QM/MM-Simulationen an größeren Systemen essentiell ist.<sup>[35]</sup>

Daher wird es ein wichtiger Schritt sein, das Interface durch die Implementierung der Cholesky-Zerlegung für die Berechnung der Zweielektronenintegrale und durch ein Parallelisierungsschema zu erweitern, um so die Anwendung für Systeme mit mittleren bis großen QM-Bereichen zu ermöglichen.

Alternative Methoden, welche die Rechenzeit ebenfalls verringern und deren Entwicklung in den vergangenen Jahren große Fortschritte gemacht hat, sind linear skalierende quantenchemische Methoden. Dazu zählen *density-fitting*-Methoden wie DF-MP2 und DF-LMP2<sup>[195–197]</sup> und CC-Methoden wie *domain based local pair natural orbital*-CCSD(T).<sup>[198–200]</sup>

Das implementierte Interface wurde in dieser Arbeit für die Methoden HF, MP2, CCSD(T) und CAS-SCF getestet. Ein weiterer Schritt wäre es, das Interface für weitere in CFOUR verfügbare quantenchemischen Methoden zu testen. Dabei sollte der Fokus besonders auf der Berechnung von Eigenschaften wie NMR-Verschiebungen oder der Berechnung von angeregten Zuständen mit EOM-CC-Methoden,<sup>[201,202]</sup> welche in der Spektroskopie und Photochemie relevant sind, liegen.

Zudem ist geplant, dass das Interface um CG-<sup>[9,10]</sup> und hPF-Methoden<sup>[12]</sup> erweitert wird, sodass große Systeme mit dem vollen Spektrum an verfügbaren Methoden beschrieben werden können. Dabei hilft die Flexibilität der Implementierung, welche durch Verwendung von MiMiC ermöglicht wird.

Außerdem ist es durch die flexible Implementierung und die Verwendung von CPMD möglich, eine Schema aus hierarchischen QM/QM Schichten aus hochgenauen quantenchemischen Methoden und pw-DFT zu implementieren, wodurch die Berechnung von großen QM-Systemen mit CFOUR umgangen werden kann.



---

# 10 Zusammenfassung

In dieser Arbeit wurde die Implementierung eines Interfaces zwischen dem Programmpaket CFOUR und dem MiMiC-Framework vorgestellt, welches AIMD- und QM/MM-Simulationen mit verschiedenen quantenchemischen Methoden bis hin zur CC-Theorie ermöglicht. Dazu wurden zunächst alle relevanten theoretischen Grundlagen erläutert (Kapitel 2) und anschließend die praktische Umsetzung der Implementierung diskutiert (Kapitel 3). Zusätzlich zum Interface für AIMD- und QM/MM-Simulationen wurden Algorithmen zur Geometrieoptimierung des QM-Bereichs in QM/MM-Systemen und zur Optimierung von ECPs zur Absättigung des QM-Bereichs in dem Fall, dass die Grenze zwischen QM- und MM-Bereich durch kovalente Bindungen verläuft, implementiert.

Im vierten Kapitel wurden die in der Arbeit verwendeten chemischen Systeme vorgestellt und erläutert, wie diese equilibriert wurden. Außerdem wurden die Details der verschiedenen durchgeführten Simulationen erläutert.

Im fünften Kapitel wurde die numerische Stabilität und die Funktionalität des implementierten Interfaces getestet. Dazu wurden AIMD- und QM/MM-Simulationen an einem einzelnen Wassermolekül und einem Wassersystem aus 1012 Wassermolekülen durchgeführt. Es wurde gezeigt, dass die Energie während der Simulationen konstant bleibt und nur leichten Fluktuationen ausgesetzt ist. Zudem wurde gezeigt, dass die erhaltenen Werte für das Dipolmoment von gasförmigen und flüssigen Wasser gut mit experimentellen Werten übereinstimmen und die durch das Lösungsmittel induzierte Vergrößerung des Dipolmoments für flüssiges Wasser korrekt beschrieben wird. Ebenso kann die Struktur der Lösungsumgebung bzw. die radiale Verteilungsfunktion durch QM/MM-Simulationen korrekt reproduziert werden.

Ein zentraler Punkt bei der Durchführung von QM/MM-Simulationen ist die benötigte Rechenzeit. Daher wurden im sechsten Kapitel zwei Methoden zur Beschleunigung der Simulationen vorgestellt. Durch die Verwendung von long-range-Wechselwirkungen zur Beschreibung der elektrostatischen QM/MM-Wechselwirkungen in einem System aus über 36000 Atomen wird eine Reduktion der dafür benötigten Rechenzeit von bis zu 90% erreicht. Dabei zeigen QM/MM-Simulationen, dass sich die Verwendung der long-range-Wechselwirkungen nicht auf die Stabilität und die Genauigkeit der Simulationen auswirkt. Die zweite vorgestellte Methode ist die Anwendung des in CPMD implementierten MTS-Algorithmus, welcher AIMD- und QM/MM-Simulationen mit verschiedenen quantenchemischen Methoden ermöglicht. Es wurden Simulationen auf dem kombinierten CCSD(T)/cc-pVTZ/pw-DFT/PBE Niveau mit unterschiedlichen MTS-Faktoren durchgeführt und die erhaltenen Schwingungsfrequenzen miteinander verglichen. Es wurde gezeigt, dass die reine CCSD(T)/cc-pVTZ Schwingungsfrequenz durch MTS-Simulationen mit einem MTS-Faktor von bis zu

$n = 15$  reproduziert werden kann, wodurch die Zahl an nötigen CCSD(T)-Rechnungen um bis zu 93% reduziert werden kann.

Im siebten Kapitel wurde die Optimierung und Anwendung von ECPs zur Absättigung des QM-Bereichs diskutiert. Dafür wurde gezeigt, dass die chemischen Eigenschaften des Essigsäuremoleküls, welches als Referenzsystem verwendet wurde, in quantenchemischen Rechnungen unter Verwendung der optimierten ECPs gut reproduziert werden können. Zudem wurde gezeigt, dass die optimierten ECPs auch für die Anwendung in QM/MM-Simulationen von verschiedenen Systemen geeignet sind und die Störung der elektronischen Struktur nur gering ist.

In Kapitel acht wurden verschiedene Methoden zur Berechnung von Schwingungsfrequenzen aus AIMD- und QM/MM-Simulationen vorgestellt und diskutiert, wie viel Anharmonizität die Schwingungsfrequenzen beinhalten. Im ersten Teil wurde die Temperaturabhängigkeit der Schwingungsfrequenzen für Fluorwasserstoff in AIMD-Simulationen untersucht. Weiterhin wurde versucht für verschiedene Moleküle die anharmonischen Schwingungsfrequenzen zu extrapolieren, was bei ca. 56% der Schwingungsmoden funktioniert hat. Anschließend wurden am Beispiel von flüssigem Wasser und Methanol zwei Methoden vorgestellt, wie ein Schwingungsspektrum aus QM/MM-Simulationen berechnet werden kann. Bei der ersten Methode wird das Schwingungsspektrum über die Dipolmomentautokorrelationsfunktion berechnet. Bei der zweiten werden für mehrere Snapshots einer QM/MM-Trajektorie die Schwingungsfrequenzen über statische quantenchemische Rechnungen berechnet und gemittelt. Beide Methoden liefern übereinstimmende Ergebnisse. Zudem wurde gezeigt, dass bei typischen Simulationstemperaturen von bis zu 300 K hauptsächlich harmonische Schwingungsfrequenzen erhalten werden und nur minimale anharmonischen Effekte zu erkennen sind.

Das neunte Kapitel befasst sich mit den Limitationen des implementierten Interfaces in der aktuellen Form. Dazu gehört der hohe Rechenaufwand für die Berechnung des QM-Bereichs für eine große Anzahl an Basisfunktionen. Eine mögliche Methode zur Reduzierung der Rechenzeit ist die Implementierung der Cholesky-Zerlegung und eines Parallelisierungsschemas für die Berechnung der Zweielektronenintegrale. Zudem wurden verschiedene zukünftige Erweiterungen des Interfaces wie die Implementierung hierarchischer QM/QM Schichten und die Verwendung von CG- und hPF-Methoden diskutiert.

# Literaturverzeichnis

- [1] Horstemeyer, M. F. *Multiscale Modeling: A Review*. In: Leszczynski J., Shukla M. (eds) *Practical Aspects of Computational Chemistry*.; Springer: Dordrecht, 2009.
- [2] Press release: The Nobel Prize in Physics 2021. <https://www.nobelprize.org/prizes/physics/2021/press-release/>.
- [3] Thiel, W.; Hummer, G. *Nature* **2013**, *504*, 96–97.
- [4] Senn, H. M.; Thiel, W. *QM/MM Methods for Biological Systems*. In: Reiher M. (eds) *Atomistic Approaches in Modern Biology*.; Springer Spektrum: Berlin, Heidelberg, 2006.
- [5] Brunk, E.; Rothlisberger, U. *Chem. Rev.* **2015**, *115*, 6217–6263.
- [6] Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. *Chem. Rev.* **2016**, *116*, 7898–7936.
- [7] Cramer, C. J. *Essentials of Computational Chemistry*, 2nd ed.; Wiley, 2004.
- [8] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications Inc.: New York, 1996.
- [9] Saunders, M. G.; Voth, G. A. *Annu. Rev. Biophys.* **2013**, *42*, 73–93.
- [10] Klein, M. L.; Shinoda, W. *Science* **2008**, *8*, 798–800.
- [11] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.
- [12] Milano, G.; Kawakatsu, T. *J. Chem. Phys.* **2009**, *130*, 214106.
- [13] Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.
- [14] Campomanes, P.; Neri, M.; Horta, B. A. C.; Röhrig, U. F.; Vanni, S.; Tavernelli, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2015**, *136*, 3842–3851.
- [15] Zalesny, R.; Papadopoulos, M. G.; Mezey, P. G.; Leszczynski, J. *Linear-Scaling Techniques in Computational Chemistry and Physics*; Springer: Dordrecht, 2011.
- [16] Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. *J. Chem. Phys.* **2005**, *122*, 084119.

- [17] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.
- [18] Adhireksan, Z.; Davey, G. E.; Campomanes, P.; Groessl, M.; Clavel, C. M.; Yu, H.; Nazarov, A. A.; Yeo, C. H.; Ang, W. H.; Dröge, P.; Rothlisberger, U.; Dyson, P. J.; Davey, C. A. *Nat. Commun.* **2014**, *18*, 3462.
- [19] Campomanes, P.; Rothlisberger, U.; Alfonso-Prieto, M.; Rovira, C. *J. Am. Chem. Soc.* **2015**, *137*, 11170–11178.
- [20] Genna, V.; Vidossich, P.; Ippoliti, E.; Carloni, P.; ; Vivo, M. D. *J. Am. Chem. Soc.* **2016**, *138*, 14592–14598.
- [21] Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.
- [22] Cupellini, L.; Caprasecca, S.; Guido, C. A.; Müh, F.; Renger, T.; Mennucci, B. *J. Phys. Chem. Lett.* **2018**, *9*, 6892–6899.
- [23] Loco, D.; Buda, F.; Lugtenburg, J.; Mennucci, B. *J. Phys. Chem. Lett.* **2018**, *9*, 2404–2410.
- [24] Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge University Press: Cambridge, 2009.
- [25] Robert G. Parr, Y. W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1994.
- [26] Becke, A. D. *J. Chem. Phys.* **2014**, *140*, 18A301.
- [27] Mardirossian, N.; Head-Gordon, M. *Mol. Phys.* **2017**, *115*, 2315–2372.
- [28] Kühne, T. D. et al. *J. Chem. Phys.* **2020**, *152*, 194103.
- [29] Møller, C.; Plesset, M. S. *Phys. Rev.* **2034**, *46*, 618–622.
- [30] Götz, A. W.; Clark, M. A.; Walker, R. C. *J Comput. Chem.* **2014**, *15*, 95–108.
- [31] van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J Comput. Chem.* **2005**, *26*, 1701–1718.
- [32] Melo, M. C. R.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. B.; Ribeiro, J. V.; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z. *Nat. Methods* **2018**, *15*, 351–354.
- [33] Torras, J.; Roberts, B. P.; Seabra, G. M.; Trickey, S. B. *Adv. Protein Chem. Struct. Biol.* **2015**, *100*, 1–31.
- [34] Olsen, J. M. H.; Bolnykh, V.; Meloni, S.; Ippoliti, E.; Bircher, M. P.; Carloni, P.; Rothlisberger, U. *J. Chem. Theory Comput.* **2019**, *15*, 3810–3823.

- [35] Bolnykh, V.; Olsen, J. M. H.; Meloni, S.; Bircher, M. P.; Ippoliti, E.; Carloni, P.; Rothlisberger, U. *J. Chem. Theory Comput.* **2019**, *15*, 5601–5613.
- [36] CPMD, <http://www.cpmc.org/>, Copyright IBM Corp 1990-2019. Copyright MPI für Festkörperforschung Stuttgart 1997-2001.
- [37] Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.; Smith, J.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845–854.
- [38] Mori-Sánchez, P.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2006**, *125*, 201102.
- [39] Tentscher, P. R.; Arey, J. S. *J. Chem. Theory Comput.* **2013**, *9*, 1568–1679.
- [40] Ranasinghe, D. S.; Margraf, J. T.; Jin, Y.; Bartlett, R. J. *J. Chem. Phys.* **2017**, *146*, 034102.
- [41] Lundberg, M.; Siegbahn, P. E. M. *J. Chem. Phys.* **2005**, *122*, 224103.
- [42] Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.
- [43] Stein, T.; Kronik, L.; Baer, R. *J. Am. Chem. Soc.* **2009**, *131*, 2818–2820.
- [44] Momeni, M. R.; Brown, A. *J. Chem. Theory Comput.* **2015**, *11*, 2619–2632.
- [45] Haugland, T. S.; Ronca, E.; Kjønstad, E. F.; Rubio, A.; Koch, H. *Phys. Rev. X* **2020**, *10*, 041043.
- [46] Kozma, B.; Tajti, A.; Demoulin, B.; Izsák, R.; Nooijen, M.; Szalay, P. G. *J. Chem. Theory Comput.* **2020**, *16*, 4213–4225.
- [47] Claeysens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. *Angew. Chem. Int. Ed.* **2006**, *45*, 6856–6859.
- [48] Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2020**, *152*, 214108.
- [49] Kosloff, R. *J. Phys. Chem.* **1988**, *92*, 2087–2100.
- [50] Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *389*, 457–484.
- [51] Payne, M. C.; Teter, M. P.; Joannopoulos, J. D. *Rev. Mod. Phys.* **1992**, *64*, 1045–1097.
- [52] Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press, 2009.
- [53] Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

- [54] Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- [55] Nottoli, M.; Lipparini, F. *J. Chem. Phys.* **2020**, *153*, 224108.
- [56] Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941–6947.
- [57] von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *J. Chem. Phys.* **2005**, *122*, 014113.
- [58] Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
- [59] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637–649.
- [60] Hockney, R. W.; Goel, S. P.; Eastwood, J. *J. Chem. Phys.* **1974**, *14*, 148–158.
- [61] Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [62] Nosé, S. *Mol. Phys.* **1984**, *52*, 255–268.
- [63] Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- [64] Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [65] Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- [66] Oostenbrink, C.; Villa, A.; Mark, A. E.; Gunsteren, W. F. V. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- [67] Vanommeslaeghe, K.; Hatcher, E.; Acharay, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. V.; MacKerell Jr., A. D. *J. Comput. Chem.* **2010**, *31*, 671–690.
- [68] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.
- [69] Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27–85.
- [70] Kaminski, A. G.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2004**, *108*, 621–627.
- [71] Harder, E.; Kim, B.; Friesner, R. A.; Berne, B. J. *J. Chem. Theory Comput.* **2005**, *1*, 169–180.
- [72] Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell Jr., A. D. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

- [73] Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. *J. Comput. Chem.* **2006**, *27*, 781–790.
- [74] Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieiders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio Jr., R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
- [75] Thompson, M. A.; Schenter, G. K. *J. Phys. Chem.* **1995**, *99*, 6374–6386.
- [76] Olsen, J. M.; Aidas, K.; Kongsted, J. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.
- [77] Ramachandran, K. I.; Deepa, G.; Namboori, K. *Computational Chemistry and Molecular Modeling: Principles and Applications*; Springer, 2008.
- [78] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- [79] Eichinger, M.; Tavana, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452–10467.
- [80] Burke, K. *J. Chem. Phys.* **2012**, *136*, 150901.
- [81] Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. *J. Phys. Chem. A* **2003**, *107*, 2578–2588.
- [82] Cremer, D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.
- [83] Werner, H.-J. *Adv. Chem. Phys.* **1987**, *69*, 1–62.
- [84] Shepard, R. *Adv. Chem. Phys.* **1987**, *69*, 63–200.
- [85] Hartree, D. R. *Math. Proc. Camb. Philos. Soc.* **1928**, *24*, 89–110.
- [86] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- [87] Hall, G. G. *Proc. R. Soc. A* **1951**, *205*, 541–552.
- [88] Jensen, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 273–295.
- [89] Löwdin, P.-O. *Phys. Rev.* **1955**, *97*, 1509–1520.
- [90] Kutzelnigg, W. *Theory of Electron Correlation. In: Rychlewski J. (eds) Explicitly Correlated Wave Functions in Chemistry and Physics. Progress in Theoretical Chemistry and Physics*; Springer: Dordrecht, 2003; Vol. 13.
- [91] Slater, J. C. *Phys. Rev.* **1929**, *34*, 1293–1322.
- [92] Condon, E. U. *Phys. Rev.* **1930**, *36*, 1121–1133.

- [93] Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem.* **1976**, *10*, 1–19.
- [94] Olsen, J.; Jørgensen, P.; Helgaker, T.; Christiansen, O. *J. Chem. Phys.* **2000**, *112*, 9736–9748.
- [95] Crawford, T. D.; Schaefer III, H. F. *Rev. Comput. Chem.* **2000**, *14*, 33.
- [96] Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- [97] Čížek, J. *Adv. Chem. Phys.* **1969**, *14*, 35–89.
- [98] Purvis III, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- [99] Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041–7050.
- [100] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- [101] Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157–173.
- [102] Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. *Chem. Rev.* **2020**, *120*, 5878–5909.
- [103] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–871.
- [104] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–1138.
- [105] Eschrig, H. *The Fundamentals of Density Functional Theory*; Springer Vieweg: Wiesbaden, 1996.
- [106] Kohanoff, J. *Electronic Structure Calculations for Solids and Molecules: Theory and Computational Methods*; Cambridge University Press: Cambridge, 2006.
- [107] Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2006.
- [108] Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892–7895.
- [109] Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
- [110] Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [111] Streett, W. B.; Tildesley, D. J.; Saville, G. *Mol. Phys.* **1978**, *35*, 639–648.
- [112] Tuckerman, M. E.; Martyna, G. J.; Berne, B. J. *J. Chem. Phys.* **1990**, *93*, 1287–1291.
- [113] Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1991**, *94*, 6811–6815.
- [114] Tuckerman, M. E.; Berne, B. J.; Rossi, A. *J. Chem. Phys.* **1991**, *94*, 1465–1469.

- [115] Tuckerman, M. E.; Berne, B. J. *J. Chem. Phys.* **1991**, *95*, 8362–8364.
- [116] Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990–2001.
- [117] Liberatore, E.; Meli, R.; Rothlisberger, U. *J. Chem. Theory Comput.* **2018**, *14*, 2834–2842.
- [118] Steele, R. P. *J. Chem. Phys.* **2013**, *139*, 011102.
- [119] Luehr, N.; Markland, T. E.; Martinez, T. J. *J. Chem. Phys.* **2014**, *140*, 084116.
- [120] Fatehi, S.; Steele, R. P. *J. Chem. Theory Comput.* **2015**, *11*, 884–898.
- [121] Chang, C.-C.; Czajkowski, G.; von Eicken, T.; Kesselman, C. *Evaluating the Performance Limitations of MPMD Communication*. Proceedings of the 1997 ACM/IEEE Conference on Supercomputing. New York, NY, USA, 1997; p 1–10.
- [122] Giovannini, T.; Egidi, F.; Cappelli, C. *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.
- [123] DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. *J. Chem. Phys.* **2002**, *116*, 9578–9584.
- [124] CFOUR-Manual. <http://slater.chemie.uni-mainz.de/cfour/index.php?n=Main.Manual>.
- [125] CPMD-Manual. <https://www.cpmd.org/wordpress/CPMD/getFile.php?file=manual.pdf>.
- [126] Kahn, L. R.; Baybutt, P.; Truhlar, D. G. *J. Chem. Phys.* **1976**, *65*, 3826–3853.
- [127] McMurchie, L. E.; Davidson, E. R. *J. Comput. Phys.* **1981**, *44*, 289–301.
- [128] Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibsom, T. D.; Windus, T. L. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820.
- [129] Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuß, H. *Mol. Phys.* **1993**, *80*, 1431–1441.
- [130] Peterson, K. A. *J. Chem. Phys.* **2003**, *119*, 11099–11112.
- [131] Dunning Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [132] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [133] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- [134] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

- [135] Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810–2821.
- [136] Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
- [137] Páll, S.; Hess, B. *Comput. Phys. Commun.* **2013**, *184*, 2641–2650.
- [138] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- [139] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926–935.
- [140] Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur. Biophys. J.* **2011**, *40*, 843–856.
- [141] Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. *J. Chem. Theory Comput.* **2011**, *7*, 4026–4037.
- [142] Stroet, M.; Caron, B.; Visscher, K. M.; Geerke, D. P.; Malde, A. K.; Mark, A. E. *J. Chem. Theory Comput.* **2018**, *14*, 5834–5845.
- [143] Automated Topology Builder (ATB). <https://atb.uq.edu.au>.
- [144] Walser, R.; Mark, A. E.; van Gunsteren, W. F. *J. Chem. Phys.* **2000**, *112*, 10450–10459.
- [145] Uribe, L.; Gauss, J.; Diezemann, G. *J. Phys. Chem. B* **2015**, *119*, 8313–8320.
- [146] Uribe, L.; Gauss, J.; Diezemann, G. *J. Phys. Chem. B* **2016**, *120*, 10433–10441.
- [147] Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- [148] Riniker, S.; Kunz, A.-P. E.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2011**, *7*, 1469–1475.
- [149] Soper, A. K. *ISRN Phys. Chem.* **2013**, *2013*, 1–67.
- [150] Clough, S. A.; Beers, Y.; Klein, G. P.; Rothman, L. S. *J. Chem. Phys.* **1973**, *59*, 2254–2259.
- [151] Badyal, Y. S.; Saboungi, M.-L.; Price, D. L.; Shastri, S. D.; Haeffner, D. R.; Soper, A. K. *J. Chem. Phys.* **2000**, *112*, 9206–9208.
- [152] Silvestrelli, P. L.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 3572–3580.
- [153] Gubskaya, A. V.; Kusalik, P. G. *J. Chem. Phys.* **2002**, *117*, 5290–5302.
- [154] Liu, H.; Wang, Y.; Bowman, J. M. *J. Phys. Chem. B* **2016**, *120*, 1735–1742.

- [155] Scipioni, R.; Schmidt, D. A.; Boero, M. *J. Chem. Phys.* **2009**, *130*, 024502.
- [156] Reddy, S. K.; Straight, S. C.; Bajaj, P.; Pham, C. H.; Riera, M.; Moberg, D. R.; Morales, M. A.; Knight, C.; Götz, A. W.; Paesani, F. *J. Chem. Phys.* **2016**, *145*, 194504.
- [157] Lin, I.-C.; Seitsonen, A. P.; Coutinho-Neto, M. D.; Tavernelli, I.; Rothlisberger, U. *J. Phys. Chem. B* **2009**, *113*, 1127–1131.
- [158] Lin, I.-C.; Seitsonen, A. P.; Tavernelli, I.; Rothlisberger, U. *J. Chem. Theory Comput.* **2012**, *8*, 3902–3910.
- [159] Miceli, G.; Hutter, J.; Pasquarello, A. *J. Chem. Theory Comput.* **2016**, *12*, 3456–3462.
- [160] Ma, Q.; Izaguirre, J. A.; Skeel, R. D. *SIAM J. Sci. Comput.* **2003**, *26*, 1951–1973.
- [161] Bowman, J. M.; Zhang, X.; Brown, A. *J. Chem. Phys.* **2003**, *119*, 646–650.
- [162] Wheeler, R. A.; Dong, H. *ChemPhysChem* **2003**, *4*, 1227–1230.
- [163] Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233–12246.
- [164] Horníček, J.; Kaprálová, P.; Bouřa, P. *J. Chem. Phys.* **2007**, *127*, 084502.
- [165] Praprotnik, M.; Janežič, D.; Mavri, J. *J. Phys. Chem. A* **2004**, *108*, 11056–11062.
- [166] Thomas, M.; Brehm, M.; Fligg, R.; Vöhringer, P.; Kirchner, B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622.
- [167] Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478–482.
- [168] Pagliai, M.; Cavazzoni, C.; Cardini, G.; Erbacci, G.; Parrinello, M.; Schettino, V. *J. Chem. Phys.* **2008**, *128*, 224514.
- [169] Wang, S. *ACS Omega* **2019**, *4*, 9271–9283.
- [170] Vitale, V.; Dziedzic, J.; Dubois, S. M.-M.; Fangohr, H.; Skylaris, C.-K. *J. Chem. Theory Comput.* **2015**, *11*, 3321–3332.
- [171] Medders, G. R.; Paesani, F. *J. Chem. Theory Comput.* **2015**, *11*, 1145–1154.
- [172] Semrouni, D.; Sharma, A.; Dognon, J.-P.; Ohanessian, G.; Clavaguéra, C. *J. Chem. Theory Comput.* **2014**, *10*, 3190–3199.
- [173] Brooks, B. R.; Janežič, D.; Karplus, M. *J. Comput. Chem.* **1995**, *16*, 1522–1542.

- [174] Nielsen, H. H. *Rev. Mod. Phys.* **1951**, *23*, 90–136.
- [175] Findenegg, G. H.; Hellweg, T. *Statistische Thermodynamik*; Springer Spektrum: Berlin, Heidelberg, 2015.
- [176] Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502–16513.
- [177] Clabo Jr., D. A.; Allen, W. D.; Remington, R. B.; Yamaguchi, Y.; Schaefer III, H. F. *Chem. Phys.* **1988**, *123*, 187–239.
- [178] Schneider, W.; Thiel, W. *Chem. Phys. Lett.* **1989**, *157*, 367–373.
- [179] Bloino, J.; Barone, V. *J. Chem. Phys.* **2012**, *136*, 124108.
- [180] von Puttkamer, K.; Quack, M. *Chem. Phys.* **1989**, *139*, 31–53.
- [181] Shimanouchi, T. *Tables of Molecular Vibrational Frequencies Consolidated Volume I*; National Bureau of Standards: Gaithersburg, 1972.
- [182] Lemus, R. *J. Mol. Spectrosc.* **2004**, *225*, 73–92.
- [183] Geguère, P. A.; Srinivasan, T. K. *J. Raman Spectrosc.* **1974**, *2*, 125–132.
- [184] Pathak, H.; Späh, A.; Kim, K. H.; Tsironi, I.; Mariedahl, D.; Blanco, M.; Huotari, S.; Honkimäki, V.; Nilsson, A. *J. Chem. Phys.* **2019**, *150*, 224506.
- [185] Piskulich, Z. A.; Thompson, W. H. *J. Chem. Phys.* **2020**, *152*, 011102.
- [186] Max, J.-J.; Chapados, C. *J. Chem. Phys.* **2009**, *131*, 184505.
- [187] Smith, A. L. *The Coblentz Society Desk Book of Infrared Spectra*; The Coblentz Society: Kirkwood, 1982.
- [188] Schmidt, Y. *Implementierung analytischer Ableitungen in QM/MM-Ansätzen zur Berechnung von Schwingungsfrequenzen in biologisch relevanten Molekülen*. M.Sc. thesis, Johannes Gutenberg-Universität Mainz, Mainz, 2020.
- [189] Beebe, N. H. F.; Linderberg, J. *Int. J. Quantum Chem.* **1977**, *17*, 683–705.
- [190] Koch, H.; de Merás, A. S.; Pedersen, T. B. *J. Chem. Phys.* **2003**, *118*, 9481–9484.
- [191] Nottoli, T.; Gauss, J.; Lipparini, F. *J. Chem. Theory Comput.* **2021**, *17*, 6819–6831.
- [192] Burger, S.; Lipparini, F.; Gauss, J.; Stopkowicz, S. *J. Chem. Phys.* **2021**, *155*, 074105.
- [193] Blaschke, S.; Stopkowicz, S. *J. Chem. Phys.* **2022**, *156*, 044115.

- [194] Feng, X.; Epifanovsky, E.; Gauss, J.; Krylov, A. I. *J. Chem. Phys.* **2019**, *151*, 014110.
- [195] Kutzelnigg, W.; Klopper, W. *J. Chem. Phys.* **1991**, *94*, 1985–2001.
- [196] Werner, H.-J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149–8160.
- [197] Schütz, M.; Werner, H.-J.; Lindh, R.; Manby, F. R. *J. Chem. Phys.* **2004**, *121*, 737–750.
- [198] Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009**, *130*, 114108.
- [199] Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2016**, *144*, 024109.
- [200] Liakos, D. G.; Guo, Y.; Neese, F. *J. Phys. Chem. A* **2020**, *124*, 90–100.
- [201] Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029–7039.
- [202] Hampe, F.; Stopkowicz, S. *J. Chem. Phys.* **2007**, *146*, 154105.



---

# A Ableitungen der Multipolmomente

## Ableitungen der Multipolmomente

Im Folgenden sind alle Multipolmomente und deren Ableitungen angegeben, welche in den Gleichungen 2.20, 2.26 und 2.27 verwendet werden. Die Indizes  $\alpha$ ,  $\beta$ ,  $\gamma$  etc. entsprechen den x-, y- und z-Komponenten der kartesischen Koordinaten.

$$C = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | \chi_\nu \rangle, \quad (\text{A.1})$$

$$\mu^\alpha = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | r^\alpha | \chi_\nu \rangle, \quad (\text{A.2})$$

$$Q^{\alpha\beta} = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | 3r^\alpha r^\beta - |\mathbf{r}|^2 \delta_{\alpha\beta} | \chi_\nu \rangle, \quad (\text{A.3})$$

$$O^{\alpha\beta\gamma} = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | 15r^\alpha r^\beta r^\gamma - 3|\mathbf{r}|^2 (r^\gamma \delta_{\alpha\beta} + r^\beta \delta_{\alpha\gamma} + r^\alpha \delta_{\beta\gamma}) | \chi_\nu \rangle, \quad (\text{A.4})$$

$$\begin{aligned} H^{\alpha\beta\gamma\epsilon} = \sum_{\mu,\nu} D_{\mu\nu} \langle \chi_\mu | & 105r^\alpha r^\beta r^\gamma r^\epsilon + 3|\mathbf{r}|^4 (\delta_{\alpha\beta} \delta_{\gamma\epsilon} + \delta_{\alpha\gamma} \delta_{\beta\epsilon} + \delta_{\alpha\epsilon} \delta_{\beta\gamma}) \\ & - 15|\mathbf{r}|^2 (r^\alpha r^\beta \delta_{\gamma\epsilon} + r^\alpha r^\gamma \delta_{\beta\epsilon} + r^\beta r^\gamma \delta_{\alpha\epsilon} + r^\alpha r^\epsilon \delta_{\beta\gamma} + r^\beta r^\epsilon \delta_{\alpha\gamma} + r^\gamma r^\epsilon \delta_{\alpha\beta}) | \chi_\nu \rangle. \end{aligned} \quad (\text{A.5})$$

Die Ableitungen sind:

$$C^{\bar{\alpha}} = \frac{\partial C}{\partial R_I^\alpha} = \sum_{\mu,\nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_\mu}{\partial R_I^\alpha} \middle| \chi_\nu \right\rangle + \left\langle \chi_\mu \middle| \frac{\partial \chi_\nu}{\partial R_I^\alpha} \right\rangle \right], \quad (\text{A.6})$$

$$\mu^{\beta, \bar{\alpha}} = \frac{\partial \mu^{\beta}}{\partial R_I^{\alpha}} = \sum_{\mu, \nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_{\mu}}{\partial R_I^{\alpha}} \middle| r^{\beta} \middle| \chi_{\nu} \right\rangle + \left\langle \chi_{\mu} \middle| \frac{\partial r^{\beta}}{\partial R_I^{\alpha}} \middle| \chi_{\nu} \right\rangle + \left\langle \chi_{\mu} \middle| r^{\beta} \middle| \frac{\partial \chi_{\nu}}{\partial R_I^{\alpha}} \right\rangle \right], \quad (\text{A.7})$$

$$Q^{\beta\gamma, \bar{\alpha}} = \frac{\partial Q^{\beta\gamma}}{\partial R_I^{\alpha}} = \sum_{\mu, \nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_{\mu}}{\partial R_I^{\alpha}} \middle| 3r^{\beta}r^{\gamma} - |\mathbf{r}|^2\delta_{\beta\gamma} \middle| \chi_{\nu} \right\rangle + \left\langle \chi_{\mu} \middle| \frac{\partial (3r^{\beta}r^{\gamma} - |\mathbf{r}|^2\delta_{\beta\gamma})}{\partial R_I^{\alpha}} \middle| \chi_{\nu} \right\rangle + \left\langle \chi_{\mu} \middle| 3r^{\beta}r^{\gamma} - |\mathbf{r}|^2\delta_{\beta\gamma} \middle| \frac{\partial \chi_{\nu}}{\partial R_I^{\alpha}} \right\rangle \right], \quad (\text{A.8})$$

$$O^{\beta\gamma\epsilon, \bar{\alpha}} = \frac{\partial O^{\beta\gamma\epsilon}}{\partial R_I^{\alpha}} = \sum_{\mu, \nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_{\mu}}{\partial R_I^{\alpha}} \middle| 15r^{\beta}r^{\gamma}r^{\epsilon} - 3|\mathbf{r}|^2(r^{\epsilon}\delta_{\beta\gamma} + r^{\gamma}\delta_{\beta\epsilon} + r^{\beta}\delta_{\gamma\epsilon}) \middle| \chi_{\nu} \right\rangle + \left\langle \chi_{\mu} \middle| \frac{\partial [15r^{\beta}r^{\gamma}r^{\epsilon} - 3|\mathbf{r}|^2(r^{\epsilon}\delta_{\beta\gamma} + r^{\gamma}\delta_{\beta\epsilon} + r^{\beta}\delta_{\gamma\epsilon})]}{\partial R_I^{\alpha}} \middle| \chi_{\nu} \right\rangle + \left\langle \chi_{\mu} \middle| 15r^{\beta}r^{\gamma}r^{\epsilon} - 3|\mathbf{r}|^2(r^{\epsilon}\delta_{\beta\gamma} + r^{\gamma}\delta_{\beta\epsilon} + r^{\beta}\delta_{\gamma\epsilon}) \middle| \frac{\partial \chi_{\nu}}{\partial R_I^{\alpha}} \right\rangle \right], \quad (\text{A.9})$$

$$\begin{aligned} H^{\beta\gamma\epsilon\zeta, \bar{\alpha}} &= \frac{\partial H^{\beta\gamma\epsilon\zeta}}{\partial R_I^{\alpha}} \\ &= \sum_{\mu, \nu} D_{\mu\nu} \left[ \left\langle \frac{\partial \chi_{\mu}}{\partial R_I^{\alpha}} \middle| 105r^{\beta}r^{\gamma}r^{\epsilon}r^{\zeta} + 3|\mathbf{r}|^4(\delta_{\beta\gamma}\delta_{\epsilon\zeta} + \delta_{\beta\epsilon}\delta_{\gamma\zeta} + \delta_{\beta\zeta}\delta_{\gamma\epsilon}) \right. \right. \\ &\quad \left. \left. - 15|\mathbf{r}|^2(r^{\beta}r^{\gamma}\delta_{\epsilon\zeta} + r^{\beta}r^{\epsilon}\delta_{\gamma\zeta} + r^{\gamma}r^{\epsilon}\delta_{\beta\zeta} + r^{\beta}r^{\zeta}\delta_{\gamma\epsilon} + r^{\gamma}r^{\zeta}\delta_{\beta\epsilon} + r^{\epsilon}r^{\zeta}\delta_{\beta\gamma}) \middle| \chi_{\nu} \right\rangle \right. \\ &\quad + \left\langle \chi_{\mu} \middle| \frac{\partial [105r^{\beta}r^{\gamma}r^{\epsilon}r^{\zeta} + 3|\mathbf{r}|^4(\delta_{\beta\gamma}\delta_{\epsilon\zeta} + \delta_{\beta\epsilon}\delta_{\gamma\zeta} + \delta_{\beta\zeta}\delta_{\gamma\epsilon})]}{\partial R_I^{\alpha}} \middle| \chi_{\nu} \right\rangle \\ &\quad + \left\langle \chi_{\mu} \middle| \frac{\partial [-15|\mathbf{r}|^2(r^{\beta}r^{\gamma}\delta_{\epsilon\zeta} + r^{\beta}r^{\epsilon}\delta_{\gamma\zeta} + r^{\gamma}r^{\epsilon}\delta_{\beta\zeta} + r^{\beta}r^{\zeta}\delta_{\gamma\epsilon} + r^{\gamma}r^{\zeta}\delta_{\beta\epsilon} + r^{\epsilon}r^{\zeta}\delta_{\beta\gamma})]}{\partial R_I^{\alpha}} \middle| \chi_{\nu} \right\rangle \\ &\quad + \left\langle \chi_{\mu} \middle| 105r^{\beta}r^{\gamma}r^{\epsilon}r^{\zeta} + 3|\mathbf{r}|^4(\delta_{\beta\gamma}\delta_{\epsilon\zeta} + \delta_{\beta\epsilon}\delta_{\gamma\zeta} + \delta_{\beta\zeta}\delta_{\gamma\epsilon}) \right. \\ &\quad \left. - 15|\mathbf{r}|^2(r^{\beta}r^{\gamma}\delta_{\epsilon\zeta} + r^{\beta}r^{\epsilon}\delta_{\gamma\zeta} + r^{\gamma}r^{\epsilon}\delta_{\beta\zeta} + r^{\beta}r^{\zeta}\delta_{\gamma\epsilon} + r^{\gamma}r^{\zeta}\delta_{\beta\epsilon} + r^{\epsilon}r^{\zeta}\delta_{\beta\gamma}) \middle| \frac{\partial \chi_{\nu}}{\partial R_I^{\alpha}} \right\rangle \right], \quad (\text{A.10}) \end{aligned}$$

## Umformung der Multipolentwicklung

Für die Implementierung der elektrostatischen QM/MM-Wechselwirkungen kann es nützlich sein, die Gleichungen 2.20, 2.26 und 2.27 so umzuformen, dass die Multipolterme eine einfachere Form annehmen. Die Umformung wird im Folgenden anhand des Quadrupolterms

$$\Theta = \frac{1}{2} \sum_{\alpha,\beta} Q^{\alpha\beta} \frac{R_A^\alpha R_A^\beta}{|\mathbf{R}_A|^5} \quad (\text{A.11})$$

aus Gleichung 2.20 demonstriert. Es soll gezeigt werden, dass

$$\frac{1}{2} \sum_{\alpha,\beta} r_i^\alpha r_i^\beta \left( \frac{3R_A^\alpha R_A^\beta - |\mathbf{R}_A|^2 \delta_{\alpha\beta}}{|\mathbf{R}_A|^5} \right) = \frac{1}{2} \sum_{\alpha,\beta} Q^{\alpha\beta} \frac{R_A^\alpha R_A^\beta}{|\mathbf{R}_A|^5} \quad (\text{A.12})$$

gilt. Auf die Integralform des Quadrupolmoments aus Gleichung A.3 wurde aus Gründen der Übersicht verzichtet, die Umformungen können aber analog auch mit der Integralform durchgeführt werden.

Zunächst wird der Term

$$\sum_{\alpha,\beta} r_i^\alpha r_i^\beta \left( 3R_A^\alpha R_A^\beta - |\mathbf{R}_A|^2 \delta_{\alpha\beta} \right) = \sum_{\alpha,\beta} 3r_i^\alpha r_i^\beta R_A^\alpha R_A^\beta - \sum_{\alpha,\beta} r_i^\alpha r_i^\beta |\mathbf{R}_A|^2 \delta_{\alpha\beta} \quad (\text{A.13})$$

der linken Seite von Gleichung A.12 betrachtet. Dort werden die beiden Relationen

$$|\mathbf{R}_A|^2 = R_{A,x}^2 + R_{A,y}^2 + R_{A,z}^2 = \sum_{\sigma,\mu} R_A^\sigma R_A^\mu \delta_{\sigma\mu} \quad (\text{A.14})$$

bzw.

$$\sum_{\alpha,\beta} r_i^\alpha r_i^\beta \delta_{\alpha\beta} = r_{i,x}^2 + r_{i,y}^2 + r_{i,z}^2 = |\mathbf{r}_i|^2 \quad (\text{A.15})$$

eingesetzt. Dadurch wird der Ausdruck

$$\begin{aligned} \sum_{\alpha,\beta} 3r_i^\alpha r_i^\beta R_A^\alpha R_A^\beta - \sum_{\alpha,\beta} r_i^\alpha r_i^\beta |\mathbf{R}_A|^2 \delta_{\alpha\beta} &= \sum_{\alpha,\beta} 3r_i^\alpha r_i^\beta R_A^\alpha R_A^\beta - \sum_{\alpha,\beta} r_i^\alpha r_i^\beta \delta_{\alpha\beta} \sum_{\sigma,\mu} R_A^\sigma R_A^\mu \delta_{\sigma\mu} \\ &= \sum_{\alpha,\beta} 3r_i^\alpha r_i^\beta R_A^\alpha R_A^\beta - |\mathbf{r}_i|^2 \sum_{\sigma,\mu} R_A^\sigma R_A^\mu \delta_{\sigma\mu} \end{aligned} \quad (\text{A.16})$$

erhalten. Jetzt können die Indizes  $\sigma$  und  $\mu$  des zweiten Terms in Gleichung A.16 in

$\alpha$  und  $\beta$  umbenannt werden, sodass der Ausdruck

$$\begin{aligned} \sum_{\alpha,\beta} 3r_i^\alpha r_i^\beta R_A^\alpha R_A^\beta - |\mathbf{r}_i|^2 \sum_{\sigma,\mu} R_A^\sigma R_A^\mu \delta_{\sigma\mu} &= \sum_{\alpha,\beta} 3r_i^\alpha r_i^\beta R_A^\alpha R_A^\beta - |\mathbf{r}_i|^2 \sum_{\alpha,\beta} R_A^\alpha R_A^\beta \delta_{\alpha\beta} \\ &= \left( 3r_i^\alpha r_i^\beta - |\mathbf{r}_i|^2 \delta_{\alpha\beta} \right) R_A^\alpha R_A^\beta. \end{aligned} \tag{A.17}$$

erhalten wird. Wieder eingesetzt in Gleichung A.12 wird so der gewünschte Term

$$\Theta = \frac{1}{2} \sum_{\alpha,\beta} Q^{\alpha\beta} \frac{R_A^\alpha R_A^\beta}{|\mathbf{R}_A|^5} \tag{A.18}$$

mit dem Quadrupolmoment  $Q^{\alpha\beta}$  erhalten. Die Umformung der höheren Multipolmomente erfolgt analog zum Quadrupolmoment.

## B ECPs

Im Folgenden sind die Ergebnisse der Optimierung des kleinen ECPs mit cc-pVDZ Basis bezüglich des Dipolmoments, der Gradienten, einer Kombination aus Dipolmoment und Gradienten und der Ladungsdichten an den jeweiligen Atomen sowie die durch die Optimierung erhaltenen ECPs aufgelistet. Es wurden ein kleines und ein großes ECP jeweils für die cc-pVDZ und cc-pVTZ Basis bezüglich des Dipolmoments, der Gradienten und einer Kombination aus Dipolmoment und Gradienten optimiert. Das kleine ECP wurde die cc-pVDZ Basis auch bezüglich der Ladungsdichten an den jeweiligen Atomen optimiert.

Tabelle B.1: Ergebnisse der Optimierung des kleinen ECPs mit cc-pVDZ Basis mit Essigsäure als Referenzsystem. Für die Optimierungen sind jeweils die Differenzen bzgl. der Referenzwerte angegeben. Die Bindungslängen sind in Å angegeben, die Winkel in Grad, das Dipolmoment  $\mu$  in Debye und die Ladungsdichten  $\delta$  in atomaren Einheiten.

Parameter	Ref.	Optimierung bezüglich:			
		Dipolmoment	Gradient	Kombiniert	Ladungsdichte
R <sub>C-C</sub>	1.4996	-0.2467	0.0026	-0.0128	-0.2271
R <sub>C=O</sub>	1.1853	-0.0033	-0.0011	-0.0046	-0.0031
R <sub>C-O</sub>	1.3317	-0.0034	0.0046	-0.0008	-0.0028
R <sub>O-H</sub>	0.9511	0.0024	0.0037	0.0037	0.0025
W <sub>1CCO</sub>	125.72	-0.61	0.29	0.03	-0.57
W <sub>2CCO</sub>	111.94	-0.60	-0.05	-0.86	-0.53
W <sub>COH</sub>	107.77	0.32	-0.08	0.06	0.29
$\mu$	1.802	-0.043	0.660	0.228	-0.010
$\delta_C$	120.769	0.054	0.111	0.072	0.056
$\delta_O$	297.586	0.034	0.038	0.044	0.035
$\delta_O$	297.263	0.044	0.084	0.076	0.047
$\delta_H$	0.360	-0.001	-0.001	-0.001	-0.001

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     20.7923622    2     36.6438568
9 p-f
10    0.6852457    2     7.8622337
11 d-f
12    -9.3833460    2     6.3253280
13 *
14

```

Abbildung B.1: Kleines ECP bezüglich des Dipolmoments optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     30.9934751    2     15.5627047
9 p-f
10    4.7481044    2     2.5969638
11 d-f
12    -9.1712392    2     7.2864808
13 *
14

```

Abbildung B.2: Kleines ECP bezüglich der Gradienten optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     30.0492953    2     19.6217735
9 p-f
10    6.0247177    2     2.1667719
11 d-f
12    -9.1981358    2     7.1767509
13 *
14

```

Abbildung B.3: Kleines ECP bezüglich des Dipolmoments und der Gradienten optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     9.1337296    2     17.5831044
9 p-f
10    11.4981071   2     20.7898232
11 d-f
12    -8.8856082   2     22.4028416
13 *
14

```

Abbildung B.4: Kleines ECP bezüglich der Ladungsdichten optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 4
5 g
6     0.0000000    2     1.0000000
7 s-g
8     369.0775120  2     54.7596113
9     -9.4086442   2     8.1366062
10 p-g
11     99.1573994   2     45.4475928
12     198.2939537  2     45.2746742
13     28.6129239   2     17.3643711
14     57.0718190   2     15.3879725
15 d-g
16    -17.9670434   2     52.7606260
17    -26.8155231   2     53.1367276
18     -1.3795023   2     14.5655100
19     -2.0655017   2     12.6358812
20     0.4542451    2     3.8396226
21     0.6934883    2     3.5512188
22 f-g
23     -0.7053231   2     11.4074726
24     -2.4759150   2     18.4290752
25 *
26

```

Abbildung B.5: Großes ECP bezüglich des Dipolmoments optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 4
5 g
6     0.0000000    2     1.0000000
7 s-g
8     369.8219990    2     39.5734706
9     3.8904100    2     14.5112477
10 p-g
11     99.1346267    2     45.5723681
12     198.2710447    2     45.5263022
13     28.3741287    2     18.3174540
14     56.7879941    2     17.8130167
15 d-g
16     -17.9671029    2     52.7605763
17     -26.8155570    2     53.1366324
18     -1.3854773    2     14.5635546
19     -2.0752617    2     12.6303741
20     -0.0613670    2     3.9240236
21     0.0739366    2     3.7613446
22 f-g
23     -0.7060111    2     11.4072912
24     -2.4759923    2     18.4290288
25 *
26

```

Abbildung B.6: Großes ECP bezüglich der Gradienten optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVDZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 4
5 g
6     0.0000000    2     1.0000000
7 s-g
8     369.1324103    2     53.7135697
9     -6.9378929    2     11.7443968
10 p-g
11     99.1670824    2     45.3940917
12     198.3037452    2     45.1659783
13     28.7277003    2     16.8760295
14     57.2456557    2     13.6455629
15 d-g
16     -17.9669652    2     52.7607370
17     -26.8154253    2     53.1368697
18     -1.3731810    2     14.5675646
19     -2.0552220    2     12.6416148
20     1.1076986    2     3.3707829
21     5.1283721    2     1.1287765
22 f-g
23     -0.7045974    2     11.4076490
24     -2.4758247    2     18.4291273
25 *
26

```

Abbildung B.7: Großes ECP bezüglich des Dipolmoments und der Gradienten optimiert (cc-pVDZ Basis).

```

+ ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     25.5232782    2     31.6765540
9 p-f
10    -0.0481717    2     7.9287309
11 d-f
12    -9.4054685    2     6.2160887
13 *
14

```

Abbildung B.8: Kleines ECP bezüglich des Dipolmoments optimiert (cc-pVTZ Basis).

```

+ ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     31.1391103    2     14.9561874
9 p-f
10    4.8228996     2     3.1603834
11 d-f
12    -9.1747164    2     7.2703416
13 *
14

```

Abbildung B.9: Kleines ECP bezüglich der Gradienten optimiert (cc-pVTZ Basis).

```

+ ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 3
5 f
6     0.0000000    2     1.0000000
7 s-f
8     29.9425537    2     18.9731306
9 p-f
10    4.9238460     2     2.1392842
11 d-f
12    -9.1934118    2     7.1906637
13 *
14

```

Abbildung B.10: Kleines ECP bezüglich des Dipolmoments und der Gradienten optimiert (cc-pVTZ Basis).

```

+ ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4      NCORE = 5      LMAX = 4
5 g
6      0.0000000      2      1.0000000
7 s-g
8      369.0311410      2      55.7274789
9      -9.2371514      2      9.3046662
10 p-g
11      99.1578085      2      45.4453312
12      198.2943653      2      45.2701018
13      28.6176267      2      17.3445918
14      57.0780025      2      15.3301435
15 d-g
16      -17.9670405      2      52.7606362
17      -26.8155124      2      53.1367340
18      -1.3794435      2      14.5655288
19      -2.0654030      2      12.6359383
20      0.4612420      2      3.8371135
21      0.7026612      2      3.5454363
22 f-g
23      -0.7053508      2      11.4074621
24      -2.4759177      2      18.4290731
25 *
26

```

Abbildung B.11: Großes ECP bezüglich des Dipolmoments optimiert (cc-pVTZ Basis).

```

+ ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4      NCORE = 5      LMAX = 4
5 g
6      0.0000000      2      1.0000000
7 s-g
8      369.6100811      2      44.2423967
9      1.6715220      2      15.2934734
10 p-g
11      99.1590684      2      45.4385989
12      198.2956398      2      45.2564143
13      28.6313543      2      17.2861991
14      57.0976071      2      15.1343082
15 d-g
16      -17.9669585      2      52.7607510
17      -26.8154096      2      53.1368872
18      -1.3729733      2      14.5676188
19      -2.0548944      2      12.6417797
20      1.1011729      2      3.3902356
21      2.3500324      2      1.4564224
22 f-g
23      -0.7046993      2      11.4076404
24      -2.4758296      2      18.4291272
25 *
26

```

Abbildung B.12: Großes ECP bezüglich der Gradienten optimiert (cc-pVTZ Basis).

```

+ ECP_FILE
1 CC:PVTZ-QMMM
2 # ECP for QM/MM bond break
3 *
4     NCORE = 5     LMAX = 4
5 g
6     0.0000000    2     1.0000000
7 s-g
8     369.1595456    2     53.2793011
9     -6.1647596    2     12.6800396
10 p-g
11     99.1634021    2     45.4113091
12     198.3000268    2     45.2009212
13     28.7024443    2     16.9505348
14     57.2101264    2     13.8955164
15 d-g
16     -17.9669590    2     52.7607445
17     -26.8154186    2     53.1368774
18     -1.3729036    2     14.5676456
19     -2.0547775    2     12.6418486
20     1.1234153    2     3.3627977
21     4.2835180    2     1.1035762
22 f-g
23     -0.7046197    2     11.4076552
24     -2.4758224    2     18.4291289
25 *
26

```

Abbildung B.13: Großes ECP bezüglich des Dipolmoments und der Gradienten optimiert (cc-pVTZ Basis).



# C Schwingungsfrequenzen aus den AIMD-Simulationen

Im Folgenden sind alle erhaltenen Schwingungsfrequenzen aus den AIMD-Simulationen und die entsprechenden Parameter der Extrapolation für die verschiedenen Moleküle aus Kapitel 8 angegeben.

Tabelle C.1: Schwingungsfrequenzen für das  $\text{H}_2\text{O}_2$  Molekül in Gasphase. Die Schwingungsfrequenzen wurden auf CCSD(T)/cc-pVTZ Niveau berechnet (harmonische und VPT2-Näherung) bzw. aus AIMD-Simulationen bei 0.5 und 300 K erhalten. Die Schwingungsfrequenzen bei anderen Temperaturen wurden über Extrapolation erhalten. Die Schwingungstemperatur  $\Theta_{vib}$  ist in K angegeben und die Steigung  $m$  ist dimensionslos. Die Schwingungsfrequenzen sind in  $\text{cm}^{-1}$  angegeben. Experimentelle Werte sind zum Vergleich angegeben.

Molekül	$\text{H}_2\text{O}_2$		
	$\nu_1$	$\nu_2$	$\nu_3$
$\nu(0.5K)$	3828.8	1331.2	375.2
$\nu(300K)$	3824.0	1293.5	251.9
$m$	-0.0161	-0.1259	-0.4118
$\Theta_{vib}$	5508.9	1915.5	540.1
$\nu(0K)$	3828.9	1331.3	375.4
$\nu(\Theta_{vib})$	3740.1	1090.1	153.0
QM harm.	3824.2	1330.3	374.3
QM VPT2	3636.1	1277.0	319.5
Experiment	3608 <sup>[183]</sup>	1393.5 <sup>[183]</sup>	317 <sup>[183]</sup>

Tabelle C.2: Schwingungsfrequenzen für verschiedene Moleküle in Gasphase. Die Schwingungsfrequenzen wurden auf CCSD(T)/cc-pVTZ Niveau berechnet (harmonische und VPT2-Näherung) bzw. aus AIMD-Simulationen bei 0.5 und 300 K erhalten. Die Schwingungsfrequenzen bei anderen Temperaturen wurden über Extrapolation erhalten. Die Schwingungstemperatur  $\Theta_{vib}$  ist in K angegeben und die Steigung  $m$  ist dimensionslos. Die Schwingungsfrequenzen sind in  $cm^{-1}$  angegeben. Experimentelle Werte sind zum Vergleich angegeben.

Molekül	HF		CH <sub>4</sub>		CO <sub>2</sub>		H <sub>2</sub> O		
	$\nu_1$	$\nu_2$	$\nu_1$	$\nu_2$	$\nu_1$	$\nu_2$	$\nu_1$	$\nu_2$	$\nu_3$
$\nu(0.5K)$	4199.8	3160.2	1354.0	2413.3	665.7	3969.1	3863.1	1667.6	
$\nu(300K)$	4190.9	3133.4	1343.6	2405.6	662.7	3958.3	3851.4	1658.1	
$m$	-0.0298	-0.0895	-0.0346	-0.0257	-0.0102	-0.0361	-0.0389	-0.0316	
$\Theta_{vib}$	6042.6	4546.9	1948.1	3472.2	957.8	5710.7	5558.2	2399.3	
$\nu(0K)$	4199.8	3160.2	1354.0	2413.3	665.7	3969.1	3863.1	1667.6	
$\nu(\Theta_{vib})$	4019.5	2753.1	1286.7	2323.9	656.0	3762.8	3646.7	1591.8	
QM harm.	4193.7	3157.8	1353.9	2412.8	665.4	3964.0	3858.4	1667.3	
QM VPT2	4015.2	3018.4	1320.2	2366.9	664.8	3774.2	3682.7	1616.5	
Experiment	3961 [180]	3019 [181]	1306 [181]	2349 [181]	667 [181]	3756 [182]	3657 [182]	1595 [182]	
Molekül	NH <sub>3</sub>								
Mode	$\nu_1$	$\nu_2$	$\nu_3$	$\nu_4$	$\nu_1$	$\nu_2$	$\nu_3$	$\nu_4$	
$\nu(0.5K)$	3624.9	3503.4	1693.6	1098.5	3252.6	3169.3	1487.1	979.7	
$\nu(300K)$	3599.1	3479.8	1685.8	1100.8	3225.6	3128.7	1475.6	973.6	
$m$	-0.0863	-0.0788	-0.0260	0.0075	-0.0905	-0.1361	-0.0387	-0.0204	
$\Theta_{vib}$	5215.5	5040.7	2436.7	1580.5	4679.8	4560.0	2139.7	1409.5	
$\nu(0K)$	3625.0	3503.5	1693.6	1098.5	3252.6	3169.4	1487.2	979.7	
$\nu(\Theta_{vib})$	3174.9	3106.2	1630.3	1110.4	2829.1	2548.9	1404.3	950.9	
QM harm.	3621.1	3500.0	1693.3	1098.5	3248.9	3158.3	1488.0	979.3	
QM VPT2	3446.9	3353.5	1645.6	1025.0	3105.5	2877.9	1451.9	958.2	
Experiment	3444 [181]	3337 [181]	1627 [181]	950 [181]	3105.5 [181]	2988.7 [181]	1443.5 [181]	949.3 [181]	

Tabelle C.3: Schwingungsfrequenzen für Methanol in Gasphase. Die Schwingungsfrequenzen wurden auf MP2/cc-pVTZ Niveau berechnet (harmonische und VPT2-Näherung) bzw. aus AIMD-Simulationen bei 0.5 und 300 K erhalten. Die Schwingungsfrequenzen bei anderen Temperaturen wurden über Extrapolation erhalten. Die Schwingungstemperatur  $\Theta_{vib}$  ist in K angegeben und die Steigung  $m$  ist dimensionslos. Die Schwingungsfrequenzen sind in  $cm^{-1}$  angegeben. Experimentelle Werte sind zum Vergleich angegeben.

Mode	$\nu_1$	$\nu_2$	$\nu_3$	$\nu_4$	$\nu_5$
$\nu(0.5K)$	3904.8	3188.1	3126.1	3075.6	1546.8
$\nu(300K)$	3898.0	3173.3	3088.9	3023.2	1529.6
$m$	-0.0230	-0.0495	-0.1248	-0.1755	-0.0579
$\Theta_{vib}$	5618.6	4587.9	4500.2	4428.5	2226.7
$\nu(0K)$	3905.1	3188.8	3127.8	3078.0	1547.6
$\nu(\Theta_{vib})$	3776.1	2961.7	2566.1	2300.5	1418.7
QM harm.	3906.6	3168.7	3138.2	3075.9	1545.1
QM VPT2	3727.0	3036.5	3015.4	3114.4	1499.7
Experiment <sup>[181]</sup>	3681	3000	2960	2844	1477
Mode	$\nu_6$	$\nu_7$	$\nu_8$	$\nu_9$	$\nu_{10}$
$\nu(0.5K)$	1387.9	1194.4	1109.4	1072.5	308.4
$\nu(300K)$	1385.1	1190.7	1106.7	1071.3	312.6
$m$	-0.0093	-0.0125	-0.0091	-0.0040	0.0142
$\Theta_{vib}$	1997.0	1718.7	1596.4	1543.1	443.4
$\nu(0K)$	1388.0	1194.5	1109.5	1072.5	308.2
$\nu(\Theta_{vib})$	1369.5	1173.1	1095.0	1066.4	314.5
QM harm.	1386.3	1201.9	1119.8	1054.1	277.2
QM VPT2	1334.8	1171.8	1097.2	1032.6	210.4
Experiment <sup>[181]</sup>	1345	1165	1060	1033	200

## Publikationen

1. Kirsch, T.; Olsen, J. M. H.; Bolnykh, V.; Meloni, S.; Ippoliti, E.; Rothlisberger, U.; Cascella, M.; Gauss, J. *J. Chem. Theory Comput.* **2022**, 18, 13-24