

Atomistische Modellierung der Struktur und Funktion von Biomolekülen

Marcus Elstner

Habilitationsschrift zum Habilitationsgesuch
am
Fachbereich Physik
der
Universität Paderborn
Mai 2003

Inhaltsverzeichnis

1	Einführung	5
1.1	Methoden	6
1.2	Problemstellung am Beispiel Bakteriorhodopsin	13
2	DFTB	19
2.1	Tests	21
2.2	Einbeziehung weiterer Atomtypen	22
2.3	Erweiterung fuer H-Bruecken	24
2.4	Dispersion	25
2.5	Struktur von Peptiden	26
2.6	Struktur des Retinals in Bakteriorhodopsin	29
3	Grosse Systeme	31
3.1	Kombination von Methoden:QM/MM	31
3.2	O(N) Methoden	33
3.3	Kombinierte QM/MM O(N)	35
4	Stabilität von Protein Sekundärstrukturen	37
5	Optische Eigenschaften	41
6	Zeitabhängige Phänomene	43
7	Zusammenfassung und Ausblick	45
A	Relevante Veröffentlichungen	51

Kapitel 1

Einführung

In den letzten Jahren sind theoretische Studien realer Materialien ein wichtiger Teil der Materialforschung geworden, sei es in der Physik, Chemie oder Biologie. Dies liegt einerseits an den grossen Fortschritten in der Computertechnik und andererseits an der Entwicklung effizienter numerischer Simulationsmethoden.

Untersucht werden neben kristallinen und amorphen Festkörpermodifikationen auch Cluster, Fullerene, organische Moleküle, Polymere und biologische Moleküle. Ein Hauptschwerpunkt liegt dabei in dem atomistischen Verständnis der Struktur und Dynamik dieser Materialien, aber auch weitergehend in der Aufklärung des Verhältnisses von Struktur, Eigenschaft und Funktion. Ferner kann durch die theoretische Vorhersage experimentell relevanter Daten Hilfestellung bei der Strukturaufklärung und bei der Interpretation experimenteller Ergebnisse gegeben werden. Über Beiträge zum Verständnis von Materialien hinaus können aber auch gezielte Vorschläge zur Modifikation, zum Design oder Optimierung gemacht werden. Ein prominentes Beispiel ist hier das sogenannte 'computer aided drug design', wo durch theoretische Simulation Struktur und Eigenschaften neuer Medikamente gezielt untersucht werden können, bevor sie synthetisiert und getestet werden. Ferner können Informationen auf theoretischem Wege gewonnen werden, die dem Experiment nicht, oder nicht ohne weiteres zugänglich sind. Dies betrifft etwa die Modellierung chemischer Reaktionen, die Simulation von Prozessen der Anlagerung von Molekülen, oder auch Ladungs- und Energietransportprozesse, die auf solch kurzen Zeitskalen stattfinden, dass sie vom Experiment z.Z. nicht aufgelöst werden können.

Die Simulation biomolekularer Prozesse, stellt eine besondere Herausforderung für theoretische Methoden dar. Dies liegt an der Zusammensetzung, Ausdehnung und spezifischen Art der Wechselwirkung dieser Systeme mit ihrer Umgebung. Ausgehend von Methoden, die in der Cluster- und Festkörperphysik inzwischen eine breite Verwendung finden, liegt der Schwerpunkt dieser Arbeit deshalb in der Entwicklung neuer, und in der Weiterentwicklung und Kombination bestehender

Methoden. Ziel ist es, ein Methodenspektrum zu entwickeln, das den Herausforderungen bei der Modellierung biochemischer Prozesse gewachsen ist.

Dazu wird im Folgenden zunächst ein Überblick über die verschiedenen Simulationsmethoden gegeben, die in der theoretischen Materialforschung eingesetzt werden (Kapitel 1.1). Daran anschliessend wird skizziert, welche besonderen Anforderungen biologische Strukturen und Prozesse an diese Methoden stellen (Kapitel 1.2). Der Hauptteil dieser Arbeit beschäftigt sich mit der Entwicklung und dem Testen von Methoden, die eine Beschreibung komplexer biologischer Strukturen erlauben. Ausgangspunkt der Entwicklung ist ein quantenmechanisches Rechenverfahren, das bisher auf organische Moleküle, Cluster und Festkörperstrukturen sehr erfolgreich angewendet wurde. Dieses wird in Kapitel 2 zunächst kurz eingeführt, und erste Erweiterungen des Schemas und Anwendungen und Tests der Methode für biologische Modellstrukturen werden vorgestellt. Kapitel 3 beschäftigt sich mit der Beschreibung grosser Systeme. Hier werden verschiedene Algorithmen vorgestellt, die eine quantenmechanische Behandlung grosser biologischer Moleküle erlauben. In Kapitel 4 wird eine Anwendung vorgestellt, die Simulation der Stabilität und Dynamik von Polypeptiden und eines Proteins, welche erst durch die vorgestellten methodischen Neuerungen möglich geworden ist. Kapitel 5 widmet sich der Beschreibung von optischen Anregungen innerhalb der hier verwendeten Methode, und Kapitel 6 generalisiert dieses Vorgehen auf die Beschreibung von zeitabhängigen Prozessen, die durch eine Wechselwirkung der Moleküle mit Licht ausgelöst werden können. Kapitel 7 fasst den gegenwärtigen Stand der Methode zusammen und gibt einen Ausblick auf weitere Entwicklungsrichtungen. Die Originalarbeiten zu den Themen in den jeweiligen Kapiteln sind in dem Anhang beigefügt.

1.1 Methoden

Zur Bestimmung der Eigenschaften und der Dynamik molekularer Systeme muss im allgemeinen Fall das zeitabhängige quantenmechanische Vielteilchenproblem für die elektronischen und nuklearen Freiheitsgrade gelöst werden. Dieses Vorgehen ist jedoch sehr aufwendig. Um zu praktisch realisierbaren Algorithmen zu kommen wird zunächst der Spezialfall des zeitunabhängigen Problems betrachtet. Ferner wird angenommen, daß die Dynamik der elektronischen Freiheitsgrade auf wesentlich kürzeren Zeitskalen stattfindet, als die Dynamik der Kernfreiheitsgrade. Im Rahmen der sogenannten Born-Oppenheimer (BO) Näherung wird die Schrödingergleichung (SG) separiert in einen Teil, der die elektronischen Freiheitsgrade bei festgehaltenen Kernfreiheitsgraden beschreibt und einen Teil, der die Kernwellenfunktion in einem Potential beschreibt, das aus der Lösung der elektronischen SG resultiert. Die elektronische Wellenfunktion hängt dann nur

noch parametrisch von den Kernkoordinaten ab (und nicht noch zusätzlich etwa von den Kernimpulsen), es entsteht das Bild von Potentialenergieflächen (PES: 'potential energy surface'), auf denen sich die Kerne bewegen. Der Fehler durch Einführen der BO Approximation ist i.A. klein gegenüber den nachfolgend diskutierten weiteren Näherungen. Es gibt jedoch chemische Situationen, in denen sie 'zusammenbricht', etwa bei hochenergetischen Stößen von Teilchen, Wechselwirkung mit starken elektromagnetischen Feldern, oder in bestimmten molekularen Konformationen, in denen sich zwei PES entlang einer Reaktionskoordinate (etwa elektronischer Grund- und angeregter Zustand) kreuzen. In diesem Fall müssen alternative Zugänge gewählt werden, wie in Kapitel 6 ausgeführt. In einigen Fällen ist auch die klassische Beschreibung der Atomkerne nicht angemessen. Die wird akut, wenn chemische Reaktionen untersucht werden, die den Transfer von Protonen beinhalten.

In einem weiteren Näherungsschritt werden die Atomkerne als klassische Partikel behandelt. Die Quanteneffekte der Atomkerne können jedoch mit Kenntnis der PES entlang einer Reaktionskoordinate im Nachhinein berücksichtigt werden. Hierfür stehen in der Literatur eine Reihe von Zugängen zur Verfügung, auf die aber in dieser Arbeit nicht weiter eingegangen wird.

Viele Eigenschaften molekularer Systeme sind demnach durch die Kenntnis der elektronischen Gesamtenergie des Systems zugänglich. Chemische Informationen können aus der Kenntnis der Energie in Abhängigkeit von den Kernkoordinaten gewonnen werden. Dies betrifft etwa molekulare Strukturen, wie aber auch Energiebarrieren für chemische Reaktionen etc. Weitere Eigenschaften können aus der Kenntnis der Energieänderungen in Abhängigkeit von äusseren Störungen gewonnen werden. Auf diese Weise können spektroskopische Daten, wie etwa optische Eigenschaften, NMR/ESR Spektren, Schwingungsspektren (IR, Raman) berechnet werden.

Zur Bestimmung der elektronischen Gesamtenergie ist aber die exakte Lösung der Schrödingergleichung (SG) für Mehrelektronensysteme wegen des hohen Rechenaufwandes i.A. nicht praktikabel. Die Lösung der SG etwa durch Monte Carlo Integration [1] ist auf die Beschreibung von Systemen mit wenigen Atomen beschränkt. Praktikable numerische Rechenverfahren für grössere Systeme basieren daher auf weitergehenden geeigneten Näherungen der elektronischen SG.

Die Methoden, die ausgehend von der molekularen Struktur die elektronische Gesamtenergie komplexer Systeme zu bestimmen erlauben, lassen sich in drei Kategorien einteilen: Die wellenfunktionsbasierten *ab initio* (1.1.1) und Dichtefunktional Methoden (1.1.2), die semi-empirischen Verfahren (1.1.3) und die klassischen Kraftfeldmethoden (1.1.4).

Wellenfunktionsbasierte *ab initio* Methoden

Einen quantenmechanischen Zugang zu grossen Systemen erlauben erst Wellenfunktionsbasierte Methoden, die eine näherungsweise Lösung der SG durch einen Ansatz für die Vielteilchen-Wellenfunktion erreichen. Startpunkt ist die Hartree-Fock (HF) Methode, die als Ansatz ein antisymmetrisiertes Produkt aus Einzelteilchenwellenfunktionen zu Grunde legt. Davon ausgehend kann die Lösung der HF Methode durch Einbeziehung der Elektronenkorrelationen verbessert werden (siehe etwa [2]). Zum Einen durch störungstheoretische Methoden (benannt nach Möller und Plesset, die als erste die Raileigh-Schrödingersche Störungstheorie in diesem Kontext anwendeten: MP2, MP4), zum Anderen durch Methoden, die ausgehend von der HF Wellenfunktion erweiterte Wellenfunktionsansätze verwenden (CI, CASSCF, ...).

Die HF Methode ist heute ein Standardverfahren, das insbesondere zur Beschreibung von Struktur und Schwingungseigenschaften von Molekülen sehr erfolgreich eingesetzt wird. Mit ihr können Systeme von bis zu etwa 100 Atomen routinemässig behandelt werden. Die Energetik der chemischen und die Struktur und Energetik der H-Brücken und VdW Bindungen können mit HF nicht, oder nicht mit ausreichender Genauigkeit berechnet werden. Hierzu ist die Einbeziehung der Elektronenkorrelation unbedingt erforderlich. Die entsprechenden Methoden sind aber wesentlich aufwendiger, mit MP2 etwa können nur noch Systeme mit bis zu etwa 30 Atomen betrachtet werden.

Die Dichtefunktional Theorie (DFT)

Als Alternative zu den *ab initio* Verfahren hat sich in den letzten Jahren die sogenannte Dichtefunktionaltheorie (DFT) etabliert (siehe etwa [3, 4]). Ihre Attraktivität beruht darin, daß sie viele relevante molekulare Systeme mit der Genauigkeit von post-HF Methoden (etwa MP2) beschreibt, bei gleichzeitig wesentlich geringerer Rechenzeit.

In diesem Abschnitt sollen einige grundlegende Gleichungen der Dichtefunktionaltheorie eingeführt werden ([3, 4]), die als Startpunkt für die Entwicklung eines approximativen Verfahrens zur Bestimmung der elektronischen Struktur molekularer Systeme dienen.

Der (nicht relativistische) Hamiltonoperator für ein System von N Elektronen in dem Coulomb-Potential der Atomkerne ist (in atomaren Einheiten) durch

$$\hat{H} = \sum_i^N \left[-\frac{\Delta_i}{2} + V_{Kern}(\mathbf{r}_i) + \sum_{j>i}^N \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \right] \quad (1.1)$$

gegeben. $V_{Kern}(\mathbf{r}_i)$ ist das auf die Elektronen wirkende Potential der Atomrümpfe.

Die Eigenzustände $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ des Systems können durch die Lösung der Schrödingergleichung erhalten werden:

$$\hat{H}\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (1.2)$$

Die Grundidee der Dichtefunktionaltheorie ist, die komplizierte N-Elektronenwellenfunktion $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ und die dazugehörige Schrödingergleichung durch die viel einfachere Elektronendichte $\rho(\mathbf{r})$ und ein entsprechendes Schema zu ihrer Berechnung zu ersetzen. Grundlegend für die Dichtefunktionaltheorie ist ein Theorem von Hohenberg und Kohn [5], nach dem eine eindeutige Abbildung zwischen einem externen Potential $V_{ext}(\mathbf{r})$ (z.B. das Coulomb-Potential der Atomkerne, $V_{ext}(\mathbf{r}) = V_{Kern}(\mathbf{r})$) und der Elektronendichte $\rho(\mathbf{r})$ besteht. Nach dem ersten Hohenberg-Kohn Theorem ist $\rho(\mathbf{r})$ durch das externe Potential $V_{ext}(\mathbf{r})$ eindeutig bestimmt.

Durch das zweite Hohenberg-Kohn Theorem ist eine Beziehung der Grundzustandsenergie zu der N-Elektronendichte gegeben: Die exakte Grundzustandsenergie E_0 ist durch die Grundzustandsdichte $\rho_G(\mathbf{r})$ eindeutig festgelegt. Für eine Testfunktion $\rho'(\mathbf{r})$, mit $\rho'(\mathbf{r}) > 0$ und

$$\int \rho'(\mathbf{r}) d\mathbf{r} = N$$

gilt:

$$E[\rho'(\mathbf{r})] \geq E_0.$$

Das Energiefunktional läßt sich in folgende Teilausdrücke zerlegen:

$$E[\rho(\mathbf{r})] = T_s[\rho(\mathbf{r})] + \int V_{Kern}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int V_H[\rho(\mathbf{r})]\rho(\mathbf{r})d\mathbf{r} + E_{xc}[\rho(\mathbf{r})] \quad (1.3)$$

$$V_{Kern}(\mathbf{r}) = - \sum_A \frac{Z^A}{|\mathbf{R}_A - \mathbf{r}|}, V_H[\rho(\mathbf{r})] = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (1.4)$$

$T_s[\rho(\mathbf{r})]$ ist die kinetische Energie eines nicht wechselwirkenden Elektronengases, $E_{xc}[\rho(\mathbf{r})]$ ist die sogenannte Austausch- und Korrelationsenergie und Z^A sind die Kernladungszahlen der Atome mit den Koordinaten \mathbf{R}_A .

Kohn und Sham [6] haben für die Bestimmung von $\rho_G(\mathbf{r})$ ein auf der Einführung von Eielektronenorbitalen $\Psi_i(\mathbf{r})$ basierendes Schema vorgeschlagen. Hierbei wird ein System nicht wechselwirkender Elektronen betrachtet, die sich in einem externen Potential V_s bewegen, wobei sich die Dichte durch die Kohn-Sham (KS) Orbitale $\Psi_i(\mathbf{r})$ und Besetzungszahlen n_i gemäß

$$\rho(\mathbf{r}) = \sum_{i=1}^{occ} n_i \langle \Psi_i(\mathbf{r}) | \Psi_i(\mathbf{r}) \rangle$$

darstellen läßt und die kinetische Energie durch

$$T_s[\rho(\mathbf{r})] = \sum_{i=1}^{occ} n_i \langle \Psi_i(\mathbf{r}) | -\nabla^2 | \Psi_i(\mathbf{r}) \rangle$$

gegeben ist.

Die Grundzustandsenergie kann durch Anwendung des Variationsprinzips erhalten werden. Im Grundzustand des Systems gilt, daß die Variation der Energie unter der Nebenbedingung der Orthogonalität der Wellenfunktionen

$$\int |\Psi_i(\mathbf{r})|^2 d\mathbf{r} = 1,$$

verschwindet:

$$\frac{\delta}{\delta \Psi_i^*(\mathbf{r})} \left[E[\rho(\mathbf{r})] - \sum_j n_j \epsilon_j \left(\int |\Psi_j(\mathbf{r})|^2 d\mathbf{r} - 1 \right) \right] = 0$$

Dies führt zu dem Eigenwertproblem (Kohn-Sham Gleichungen)

$$\left[-\nabla^2 + V_s[\rho(\mathbf{r})] \right] \Psi_i(\mathbf{r}) = \epsilon_i \Psi_i(\mathbf{r}), \quad (1.5)$$

mit

$$V_s[\rho(\mathbf{r})] = V_{Kern}(\mathbf{r}) + V_H[\rho(\mathbf{r})] + V_{xc}[\rho(\mathbf{r})]$$

und

$$V_{xc}[\rho(\mathbf{r})] = \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})}$$

$V_s[\rho(\mathbf{r})]$ ist ein effektives Einteilchenpotential, das die Wechselwirkung eines Elektrons mit den übrigen Elektronen und den Atomrümpfen beschreibt. Das Eigenwertproblem muß, beginnend mit einer Startdichte, iterativ gelöst werden, da der Hamiltonoperator die Dichte selbst enthält. Die Iteration wird bis zur Selbstkonsistenz durchgeführt, d.h., bis sich die Dichten zwischen zwei Zyklen nur noch innerhalb einer festgelegten Toleranz unterscheiden.

Die Bestimmung des Austausch- und Korrelationsfunktionals E_{xc} ist eine große Herausforderung im Rahmen der Dichtefunktionaltheorie. Die Hauptdefizite dieser Theorie liegen immer noch darin, daß eine angemessene Repräsentation der Austausch- und Korrelationseffekte durch entsprechende Funktionale im großen und ganzen nicht befriedigend gelöst ist. Auf einige Defizite wird auch in diesen Ausführungen noch eingegangen. Die Form von E_{xc} ist analytisch nicht bekannt, wodurch jedes Funktional auf Näherungen beruht. Die Bestimmung von guten Approximationen für E_{xc} ist Gegenstand intensiver Forschung und kann hier nicht weitergehend behandelt werden.

Die Effizienz der DFT Methoden ist in etwa mit der der HF Methode vergleichbar, wie oben schon erwähnt, es können in etwa Systeme mit 100 Atomen routinemässig behandelt werden. Mit Hilfe der Eigenwertgleichung kann die Gesamtenergie auch folgendermaßen geschrieben werden:

$$E[\rho(\mathbf{r})] = \sum_i n_i \epsilon_i - \frac{1}{2} \int V_H[\rho(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r} - \int V_{xc}[\rho(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r} + E_{xc}[\rho(\mathbf{r})] \quad (1.6)$$

$$+ \frac{1}{2} \sum_{A \neq B} \frac{Z^A Z^B}{|\mathbf{R}_A - \mathbf{R}_B|}$$

Diese Gleichung ist der Ausgangspunkt für Näherungen zur vereinfachten Berechnung elektronischer Eigenschaften molekularer Systeme, von denen die nächsten Abschnitte handeln.

Die meisten heutigen Verfahren zur Bestimmung der elektronischen Struktur molekularer Systeme basieren auf einer Einelektronenorbitalformulierung, wie sie am Beispiel der Dichtefunktionaltheorie gerade vorgestellt wurde. In einem solchen Schema wird der Grundzustand E_0 eines molekularen Systems durch eine iterative Lösung der Eigenwertgleichungen bestimmt. Die Lösung beinhaltet damit zwei Rechenschritte: Zunächst müssen die Matrixelemente des Hamilton Operators und der Überlappmatrix in einem System von Basisfunktionen berechnet werden. Dazu werde die Wellenfunktionen Ψ_i durch die Basisfunktionen ϕ_μ

$$\Psi_i = \sum_\mu c_\mu^i \phi_\mu$$

dargestellt. Damit können dann die Hamilton- und Überlappmatrixelemente in dieser Basis berechnet werden. Im nächsten Schritt werden dann die Eigenwerte und Eigenfunktionen der Hamiltonmatrix bestimmt (Dieser Ablauf gilt auch für die oben angesprochenen HF und post-HF Methoden). Für den ersten Schritt, sind numerisch aufwendige Integrationen durchzuführen. Die auftretenden Integrale sind i. A. Mehrzentrenintegrale, d.h. sie hängen von den Kernkoordinaten mehrerer Atome ab. Daher skaliert der Aufwand zu ihrer Berechnung auch mit der dritten (und bei den *ab initio* Methoden auch der vierten oder fünften) Potenz der Anzahl der im System vorhandenen Atome. Der zweite Schritt besteht in der Diagonalisierung der Hamiltonmatrix. Der Zeitaufwand für diesen Schritt steigt mit der dritten Potenz der Dimension dieser Matrix. Die Dimension wiederum hängt von der Anzahl der Basisfunktionen ab. Bei Verwendung von Basissätzen, die an den Atomkernen lokalisiert sind, hängt somit die Rechenzeit von der Anzahl der Basisfunktionen pro Atom und der Anzahl der Atome ab. Die Anzahl der Atome ist durch die Systemgrösse gegeben. Somit lässt sich die Rechenzeit

durch die Basissatzgrösse (Anzahl der Basisfunktionen pro Atom) beeinflussen. Allerdings verschlechtert eine kleinere Basis die Qualität der Ergebnisse, so daß eine Balance zwischen Qualität der Rechnung und Rechenzeit gesucht werden muß.

Es lassen sich zwei Ansatzpunkte für schnellere Verfahren benennen. Der Eine bezieht sich auf die Bestimmung der Matrixelemente. Hier gibt es vereinfachte, semi-empirische und empirische Verfahren, die die Matrixelemente im Voraus und möglichst unabhängig von den zu betrachteten Systemen zu bestimmen versuchen. Damit fällt eine Berechnung der Matrixelemente weg, was einen signifikanten Rechenzeitgewinn bedeutet, allerdings auch Abstriche bei der Präzision der Ergebnisse nach sich zieht. Andererseits kann man versuchen, die aufwendige Diagonalisierung der Hamiltonmatrix zu umgehen. Die sogenannten $O(N)$ Verfahren versuchen genau dies, die Bestimmung der Grundzustandsenergie bei einem Rechenaufwand, der linear mit der Systemgröße wächst.

Semi-empirische Methoden

Semi-empirische Verfahren vernachlässigen und nähern die aufwendig zu berechnenden Integrale der *ab initio* Verfahren und ersetzen sie anschließend durch zu bestimmende Parameter (siehe etwa [7]). Ausgehend von der HF Theorie wurden in den letzten 40 Jahren eine Reihe von semi-empirischen Verfahren entwickelt. Bei den erfolgreichsten Vertretern dieser Klasse von Methoden, den AM1 [8] und PM3 [9] Methoden werden die Parameter an experimentelle Daten von organischen Molekülen gefittet. Eine zweite Klasse von semi-empirischen Verfahren, die sogenannten tight-binding (TB) Methoden können als stationäre Approximation an die DFT verstanden werden [10]. Charakteristisch ist die Darstellung der Gesamtenergie aus Gl. 1.6, welche aus einem Bandstrukturterm E_{bs} und einem repulsiven Energiebeitrag, E_{rep} hervorgeht. E_{bs} repräsentiert den ersten Term auf der rechten Seite von Gl. 1.6, die restlichen Terme werden in E_{rep} zusammengefasst. Die Kohn-Sham Orbitale werden dann in einer lokalen Basis dargestellt, und die resultierenden Kohn-Sham Gleichungen werden nicht-selbstkonsistent gelöst.

Da bei diesen beiden Klassen von semi-empirischen Methoden die Wechselwirkungsintegrale nicht mehr berechnet werden müssen, ist der zeitlimitierende Schritt durch die Diagonalisierung der Hamiltonmatrix gegeben. Die Rechenzeit dieser Methoden skaliert daher kubisch mit der Systemgröße. Mit semi-empirischen Verfahren sind Systeme bis zu 1000 Atomen und Simulationszeiten von 10-100 ps möglich, sie können damit etwa 10 fach grössere Systeme behandeln als DFT.

Empirische Kraftfeld Methoden

Kraftfeldmethoden modellieren die Wechselwirkungen zwischen den Atomen durch klassische Potentiale [11, 12, 13, 14]. Dies führt zu einer erheblichen Vereinfachung in der Berechnung der interatomaren Wechselwirkungen. Der zeitlimitierende Faktor besteht in der Berechnung der Coulomb-Wechselwirkung zwischen den Atomen, die durch Punktladungen repräsentiert werden. Daher wächst die Rechenzeit quadratisch mit der Systemgrösse, solange nicht weitere Näherungen, wie etwa das Abschneiden der Coulomb-WW für große Abstände zwischen den Atomen oder Multipol-Ansätze zur Verwendung kommen. Kraftfeldmethoden sind allerdings aufwendig zu parametrisieren, da die Wechselwirkung der Atome von der chemischen Umgebung abhängt. So müssen beispielsweise verschiedene Typen eines Elements unterschieden werden, je nach funktionaler Gruppe, in der es auftritt. Eine sorgfältige Kalibrierung hat jedoch zu erfolgreichen Anwendungen dieser Verfahren geführt, so daß sie inzwischen zum Standardrepertoire bei der Modellierung komplexer Strukturen gehören. Mit ihnen lassen sich Systeme bis zu einigen 10000 Atomen und Simulationszeiten bis zu Nanosekunden realisieren.

Empirische Kraftfelder werden sehr erfolgreich zur Modellierung der Struktur und Dynamik von Proteinen und DNA eingesetzt. Bestimmte Probleme, wie etwa das der Proteinfaltung können dennoch zur Zeit auch mit diesen nicht Methoden behandelt werden, da die hierfür relevanten Zeitskalen die mit den Kraftfeldern nicht erreicht werden.

Neben dem Problem ihrer Genauigkeit und Übertragbarkeit zwischen verschiedenen chemischen Umgebungen liegt die massgebliche Einschränkung dieser Methoden darin, chemische Reaktionen nicht beschreiben zu können. Dies betrifft etwa enzymatische Katalyse, Protonentransport, Elektronentransport, die Wechselwirkung von Materie mit elektromagnetischer Strahlung oder elektronisch angeregte Zustände. Hier wird eine quantenmechanische Methode erforderlich

Die Herausforderungen an Simulationsmethoden bei der Modellierung solcher Prozesse sollen stellvertretend am Beispiel des Protonentransfers in dem Membranprotein Bakteriorhodopsin (bR) skizziert werden.

1.2 Problemstellung am Beispiel Bakteriorhodopsin

bR stellt das einfachste bekannte photosynthetische Zentrum dar. Es dient damit als Modellsubstanz für das Studium grundlegender photoinduzierter Prozesse in der Biologie und wird sowohl experimentell als auch theoretisch intensiv untersucht (einen zusammenfassenden Überblick findet man in [15]). Darüber hinaus

werden aber auch bedeutende Anwendungspotentiale dieses biologischen Systems in der molekularen Elektronik gesehen.

Trifft ein Photon auf das Retinal Chromophor, welches das aktive Zentrum in bR darstellt, so wird ein Protonentransferprozess in Gang gesetzt. Der dadurch erzeugte pH Gradient ist der Ausgangspunkt für die ATP Synthese in der Zelle. Sowohl das Retinal, als auch die Beschaffenheit der Proteinumgebung des Retinals sind wesentlich für das Funktionieren dieses Prozesses verantwortlich. Durch gezielten Austausch von Aminosäuren des bR mit Hilfe gentechnischer Methoden konnte dies im Detail gezeigt werden. Die optischen Eigenschaften, aber auch der Protonentransport hängen sensitiv von den elektrostatischen und sterischen Wechselwirkungen mit der Proteinumgebung ab. Eine genaue Aufklärung dieser Wechselwirkungen ist somit nicht nur für das prinzipielle Verständnis von bR, sondern auch im Hinblick auf dessen Modifizierung für technologische Anwendungen wichtig. Die Proteinumgebung hat somit einen großen Einfluß auf die räumliche Struktur und die elektronischen Eigenschaften des Retinals. Ohne diese Umgebung könnte das Retinal seine Rolle in der Lichtspeicherung und umwandlung nicht ausführen. Daher ist die Berücksichtigung der Proteinumgebung in der theoretischen Modellierung des Retinals unumgänglich. Die Grösse des Proteins inklusiver seiner Hydrathülle (etwa 10000 Atome) würde für den Einsatz empirischer Karstoffmethoden sprechen. Allerdings kann das konjugierte π -Elektronensystem des Retinals, dessen elektronisch angeregten Zustände wie auch der Protonentransfer nur mit Hilfe von quantenmechanischen Methoden beschrieben werden.

Diese zwei Anforderungen, die Notwendigkeit der quantenmechanischen Beschreibung des Retinals zusammen mit der Grösse des gesamten bR Moleküls machen die Anwendung neuer Techniken zwingend.

Die hier beschriebene Problemstellung ist typisch für eine Reihe biologisch interessanter Fragen, die eine besondere Herausforderung an theoretische Methoden stellen:

- Gegenüber anorganischen und organischen Systemen heben sich die biologischen durch die zu behandelnde Systemgrösse ab. Sie umfassen von einigen 100 bis zu einige 1000 Atomen. Die Behandlung grosser Systeme kann üblicherweise durch Ausnutzung der Systemsymmetrie (etwa bei Festkörperproblemen) vereinfacht werden, es müssen dann nur noch kleine Ausschnitte des Systems behandelt werden. Dieser Ausweg ist jedoch aufgrund der fehlenden Symmetrie bei Biomolekülen nicht gegeben.
- Biomoleküle enthalten eine Vielzahl von verschiedenen Atomtypen, von Wasserstoff bis zu Metallen und Übergangsmetallelementen. Dies erschwert die theoretische Beschreibung im Vergleich zu den oft homogeneren Festkörpersystemen erheblich.

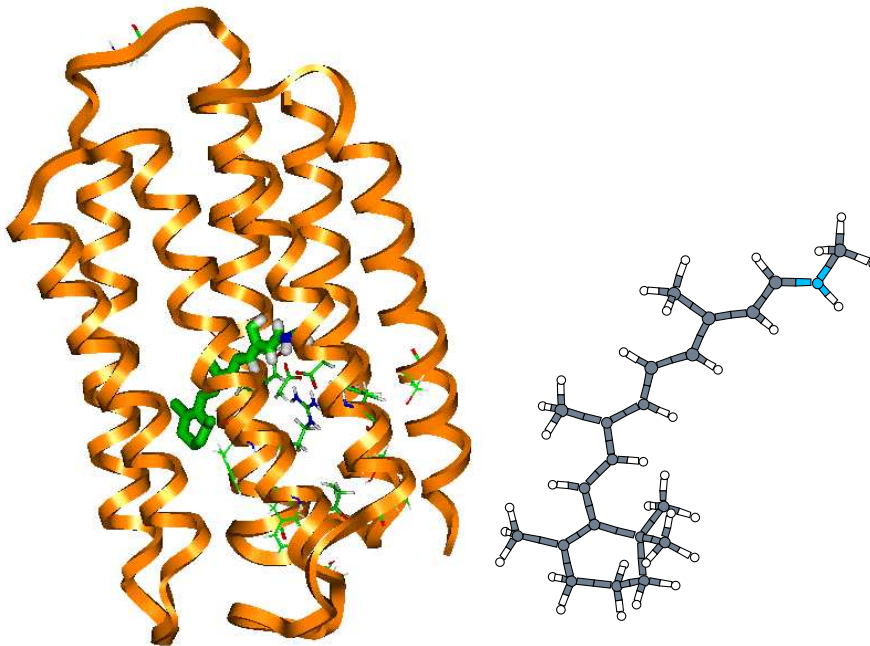


Abbildung 1.1: Das bR Molekül und das Retinal

- Biomoleküle liegen meist nicht in der Gasphase vor. Die Wechselwirkung mit dem sie umgebenden Lösungsmittel, das zudem auch teilweise in sie eindringt, verändert Struktur, Eigenschaften und Funktion dieser Moleküle signifikant. Daher muss eine Beschreibung die Umgebung stets mit berücksichtigen. Diese bedeutet sowohl methodisch, als auch im Rechenzeitbedarf einen signifikanten Mehraufwand.
- Die Stabilität biomolekularer Strukturen beruht zum grossen Teil auf Wasserstoffbrücken und Van der Waals Bindungen, die hohe Anforderungen an die Genauigkeit der Simulationsmethoden stellen, wie weiter unten noch näher diskutiert wird.
- Auch ist eine genaue Beschreibung der Struktur und Energetik der chemischen Bindungen erforderlich. Zum Einen, um etwaige chemische Reaktionen beschreiben zu können, zum Anderen aber, um die Stabilität der Konformationen des Proteins zu gewährleisten.
- Eine zusätzliche Schwierigkeit bei der Beschreibung von Br liegt darin, daß die Reaktion Photoinduziert wird. Die Beschreibung von Prozessen im elektronisch angeregten Zustand ist theoretisch sehr aufwendig, viele Methoden erreichen die hierzu erforderlichen Genauigkeitsanforderungen nicht. In den

letzten Jahren lag daher auch ein Schwerpunkt auf der Entwicklung neuer Methoden zur Beschreibung angeregter Zustände.

Die Systemgrösse und die Anforderungen bestimmen demnach die Wahl der Simulationsmethode. Aus dem bisher gesagten dürfte ersichtlich werden, dass keine der unter 1.1 beschriebenen Methoden per se in der Lage ist, das skizzierte Problem zu anzugehen.

Das quantenmechanisch zu beschreibenden Problem (Chromophor und einige Aminosäuren aus der Proteinumgebung) umfasst etwa 100 Atome, ist daher zu groß um mit *ab initio* Verfahren behandelt zu werden. Für derartige Systemgrößen müssen quantenmechanische Methoden bereitgestellt werden, welche die chemische Genauigkeit und Übertragbarkeit von Post-Hartree-Fock bzw. DFT-Rechnungen mit der Effizienz und Einfachheit von stärker genäherten semiempirischen Verfahren verbinden. Im Rahmen meiner Dissertation [16] habe ich eine solche Methode, das sogenannte SCC-DFTB Verfahren ('Self-consistent charge density-functional tight-binding'; im Folgenden wird jedoch der Lesbarkeit halber die Methode nur noch mit DFTB abgekürzt) entwickelt und anhand von Tests und ersten Anwendungen für organische Moleküle und biologische Systeme eingeführt [17].

In den letzten Jahren wurde die Parametrisierung der Methode um neue Atomtypen erweitert und die Methode vor allem in Hinblick auf die beschriebenen Anforderungen getestet. Desweiteren wurden verschiedenen Möglichkeiten untersucht, die Beschreibung von H-Brücken im DFTB Formalismus zu verbessern. Van der Waals Wechselwirkungen werden aufgrund der approximativen Behandlung der Elektronenkorrelationen in der Dichtefunktionaltheorie nicht berücksichtigt. Da das DFTB Verfahren als Näherung an die DFT auch ihre Probleme übernimmt, wurde ein empirischer Ansatz zur Einbeziehung von VdW (Dispersions) Kräften in die DFTB Methode verfolgt. Damit wurden die Grundsteine gelegt, mit Hilfe des DFTB biologische Strukturen, wie das hier skizzierte bR, erfolgreich behandeln zu können (Kapitel 2).

Die Prozesse im reaktiven Zentrum sind allerdings sensitiv abhängig von der Proteinumgebung, eine isolierte Behandlung des reaktiven Zentrums macht daher wenig Sinn. Auf Workstations können mit der DFTB Methode komplexe Atomordnungen mit bis zu ca. 500 Atomen behandelt werden. Die Bearbeitung größerer Strukturen, wie sie bei der hier anvisierten Problemstellung auftreten, kann durch drei Vorgehensweisen erreicht werden: eine massive Parallelisierung der Methode [19], die Verwendung von sogenannten *Order-N*-Verfahren oder ein sogenanntes QM/MM Verfahren.

- Die massive Parallelisierung erlaubt zwar eine Vergrößerung der Systeme, bedeutet aber nicht eine effizientere Nutzung der Computerressourcen, da die Gesamtrechenzeit nur auf mehrere Prozessoren aufgeteilt wird.

- Eine vielversprechende Entwicklung stellen die sogenannten $O(N)$ Algorithmen dar. Die Lösung der quantenmechanischen Eigenwertgleichungen skaliert mit $O(N^3)$, d.h. die Rechenzeit steigt kubisch mit wachsender Systemgröße (N : Anzahl der Basisfunktionen), während unter Verwendung von $O(N)$ Algorithmen ein lineares Anwachsen erreicht werden kann.
- Die Idee von QM/MM Methoden besteht darin, einen Teil des Systems quantenmechanisch zu beschreiben und den Rest des Systems in der Ankopplung mit einem empirischen Kraftfeld zu berücksichtigen.

In den letzten Jahren wurden alle drei Optionen implementiert und kürzlich sogar eine Kombination der drei Strategien realisiert. Diese Entwicklungen, ihre Implementation und Anwendung werden in Kapitel 3 ausgeführt. In Kapitel 5 werden die Entwicklungen zur Beschreibung elektronisch angeregter Zustände und in Kapitel 6 Methoden, wie sie etwa zur Behandlung der Wechselwirkung von Licht mit Materie notwendig sind, dargestellt. Zunächst soll jedoch kurz auf die charakteristischen Merkmale der DFTB Methode und deren Erweiterungen und Test eingegangen werden.

Kapitel 2

DFTB

Eine ausführliche Diskussion der DFTB Methode findet sich in meiner Dissertation [16] und in Ref. [17, 18]. Hier sollen nur knapp die wichtigsten Charakteristika der Methode angesprochen werden.

Die DFTB Methode basiert auf einer Entwicklung des effektiven potentials in dem Kohn-Sham Energiefunktional nach Ladungsdichtefluktuationen δn an einer gegebenen Referenz-Elektronendichte n_0 [16]. Die Gesamtenergie in zweiter Ordnung in den Ladungsfluktuation ist:

$$\begin{aligned} E &= \sum_i^{occ} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \iint \left(\frac{1}{|\vec{r} - \vec{r}'|} + \left. \frac{\delta^2 E_{xc}}{\delta n(\vec{r}) \delta n(\vec{r}')} \right|_{n_0} \right) \delta n(\vec{r}) \delta n(\vec{r}') d\vec{r} d\vec{r}' \\ &- \frac{1}{2} \iint' \frac{n_0(\vec{r}) n_0(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' + \int [E_{xc}[n_0] - V_{xc}[n_0]] n_0(\vec{r}) d\vec{r} + E_{ii} \end{aligned} \quad (2.1)$$

$\hat{H}^0 = \hat{H}[n_0]$ ist der effektive Kohn-Sham Hamiltonian, ausgewertet an der Referenzdichte, und die ψ_i sind die sogenannten Kohn-Sham Orbitale. E_{xc} und V_{xc} bezeichnen die Austausch- und Korrelationsenergie- und Potentialbeiträge und E_{ii} ist die Kern-Kern Abstoßung. Um den Ausdruck für die Gesamtenergie in der DFTB Methode zu erhalten, werden noch folgende Näherungsschritte ausgeführt:

- Die Hamiltonmatrixelemente $\langle \psi_i | \hat{H}^0 | \psi_i \rangle$ werden in einer minimalen Basis optimierter atomarer Wellenfunktionen ϕ_μ dargestellt,

$$\psi_i = \sum_\mu c_\mu^i \phi_\mu.$$

Um die ϕ_μ zu erhalten werden die atomaren Kohn-Sham Gleichungen unter Verwendung eines harmonischen Zusatzpotentials gelöst [20], welches die Aufgabe hat, die Wellenfunktionen räumlich stärker zu lokalisieren. Die Hamiltonmatrixelemente in dieser Basis $H_{\mu\nu}^0$ werden dann wie folgt berechnet. Als Diagonalbeiträge $H_{\mu\mu}^0$ werden die atomare Eigenwerte angesetzt

und die nicht-Diagonalelemente $H_{\mu\nu}^0$ werden durch eine Zweizentrennäherung erhalten.

$$H_{\mu\nu}^0 = \langle \phi_\mu | \hat{T} + V_{eff}[n_\alpha^0 + n_\beta^0] | \phi_\nu \rangle \quad \mu \in \alpha, \nu \in \beta,$$

Diese werden zusammen mit den Überlappmatrixelementen $S_{\mu\nu}$ vorab berechnet und in Abhängigkeit vom interatomaren Abstand $R_{\alpha\beta}$ tabelliert. V_{eff} ist das effektive Kohn-Sham Potential und die n_α^0 sind die Elektronendichten der neutralen Atome α ; \hat{T} ist der Operator der kinetischen Energie.

- Die Ladungsdichtefluktuationen δn werden als Superposition atomarer Beiträge δn_α geschrieben,

$$\delta n = \sum_\alpha \delta n_\alpha,$$

welche durch die Ladungsdichtefluktuationen $\Delta q_\alpha = q_\alpha - q_\alpha^0$ ausgedrückt werden. q_α^0 ist die Anzahl der Valenelektronen des neutralen Atoms α und die Ladungen q_α werden durch eine Mulliken Populationsanalyse erhalten. Der Term zweiter Ordnung in δn in Gl. (2.1) wird dann durch $\sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta$ genähert. Für $\alpha \neq \beta$, wird $\gamma_{\alpha\beta}$ durch eine analytische Funktion bestimmt, die die Coulomb Wechselwirkung spärlicher Ladungsverteilungen an den Kernorten R_α und R_β repräsentiert. Die Diagonalterme $\gamma_{\alpha\alpha}$ werden aus $\frac{\partial E_\alpha^2}{\partial^2 q_\alpha}$ erhalten.

- Die 'double counting' Beiträge und die Kern-Kern Abstoßung werden in einem Term E_{rep} zusammengefaßt. Dieser wird als Summe von Zweizentrenpotentialen $U[R_{\alpha\beta}]$ genähert.

$$E_{rep} = \sum_{\alpha \neq \beta} U[R_{\alpha\beta}].$$

Mit diesen Näherungen und Definitionen lässt sich dann die Gesamtenergie folgendermaßen ausdrücken:

$$E_{tot} = \sum_{i\mu\nu} c_\mu^i c_\nu^i H_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta + E_{rep} \quad (2.2)$$

Zur Berechnung der Hamiltonmatrixelemente wird das Austausch- und Korrelationspotential von Perdew, Burcke und Ernzerhof verwendet [21].

Unter Anwendung des Variationsprinzips erhält man genäherte Kohn-Sham Gleichungen, die nun iterativ gelöst werden müssen, da die Matrixelemente selbst von der Elektronendichte über die Mulliken Ladungen abhängen.

	DFTB	AM1	PM3	LDA	GGA
Reaktionen (33)					
$\bar{\mathbf{F}}(E)(kcal/mol)$	5.0	-	-	11.9	4.6
Geometrien (65 Moleküle getestet)					
$\bar{\mathbf{F}}(R) (\text{\AA})$	0.013	0.017	0.012	-	-
$\bar{\mathbf{F}}(\theta)(^\circ)$	1.9	2.0	2.4	-	-
Frequenzen (33 Moleküle getestet)					
$\mathbf{F}(\nu)$	6%			3-5%	3-5%

Tabelle 2.1: Reaktionsenergien, Bindungsparameter und Schwingungsfrequenzen. Vergleich der semi-empirischen und DFT basierten Methoden mit Bezug auf experimentelle Werte. LDA bezeichnet DFT in der lokalen Dichtenäherung, GGA in der gradientenkoriigierten Variante. $\bar{\mathbf{F}}$ bezeichnet den mittleren absoluten Fehler in Bezug auf die experimentellen Daten.

Für die atomaren Kräfte wurden analytische Ausdrücke hergeleitet, die zweiten Ableitungen nach den Kernkoordinaten werden numerisch ausgeführt.

Die Zweizentrenpotentiale $U[R_{\alpha\beta}]$ erhält man durch Subtraktion der DFT Gesamtenergie von der elektronischen Energie des DFTB Formalismus in Abhängigkeit vom zwischenatomaren Abstand $r = R_{\alpha\beta}$.

2.1 Tests

Für Anwendungen auf organische und biologische Systeme wurde die DFTB Methode zunächst für die Atomtypen H, C, N, und O parametrisiert und an einem umfangreichen Satz organischer Moleküle getestet. Zur Evaluierung wurden insbesondere 33 Reaktionen zwischen kleinen organischen Molekülen untersucht und die erhaltenen Reaktionsenergien mit experimentellen Werten und den Resultaten anderer Methoden verglichen (Daten von Anzelm und Wimmer [22]). Weiter wurden die Geometrien von 65 (Moleküle aus dem Satz von Ref [8]) und die Schwingungsfrequenzen von 33 Molekülen im Vergleich mit anderen Methoden und experimentellen Daten untersucht. Es wurde jeweils der mittlere Fehler der verschiedenen Methoden bezüglich der experimentellen Daten ermittelt, wie in Tabelle 2.1 wiedergegeben. Bezüglich der Reaktionsenergien weist DFTB eine ähnliche Abweichung von den experimentellen Daten auf, wie DFT-GGA [22], was als exzellentes Ergebniss zu werten ist. DFT-LDA [22] schneidet hier durch das wohlbekannte 'überbinden' (overbinding) wesentlich schlechter ab. Für die

semi-empirischen Methoden AM1 [8] und PM3 [9] liegen für dieses Set von Reaktionen explizit keine Daten vor. Allerdings liegen die Fehler für molekulare Bindungsenthalpien für diese Methoden bei 4-6 kcal/mol, der Fehler in den Reaktionsenergien sollte daher von der gleichen Grösse sein.

Die Abweichungen in den Bindungslängen $\bar{F}(R)$ sind bei den drei approximativen Methoden sehr klein, und sind vergleichbar mit dem Ergebniss von *ab initio* MP2 und DFT Methoden, die auch Abweichungen von etwa 0.015 (Å) aufweisen. Ähnliches gilt für die Vorhersage der Bindungswinkel θ . Allerdings liegen DFT und MP2 Daten nicht für einen solch grossen Satz von Molekülen vor, dennoch kann man aus den vorhandenen Daten einen Eindruck von deren Genauigkeit gewinnen [23].

Die Schwingungsfrequenzen werden mit einem mittleren Fehler von 6% etwas schlechter beschrieben als durch DFT-GGA und DFT-LDA, aber wesentlich besser als bei den semi-empirischen AM1 und PM3 Methoden, die mittlere Fehler von über 10% aufweisen (siehe etwa [24]).

In den untersuchten Molekülen treten Bindungssituationen auf, wie sie auch in grösseren organischen und biologischen Molekülen typisch sind. Die bisher vorgestellten Tests lassen damit auf eine sehr gute Beschreibung der kovalenten Bindungen schließen. Allerdings müssen ähnliche Tests für jeden weiteren Atomtyp, der in die Berechnungen einbezogen werden soll, wiederholt werden. Dies betrifft bei biologischen Anwendungen vor allem die Elemente Schwefel und Phosphor. Desweiteren aber auch Alkalimetalle und Übergangsmetallelemente. Weitere Elemente für die Methode zur Verfügung zu stellen, war daher ein Schwerpunkt der Weiterentwicklung der Methode.

Darüberhinaus geben die Tests keinen Aufschluß über die Qualität der Methode bei der Beschreibung von nicht-bindenden Wechselwirkungen ('non-bonding' interactions), wie etwa bei Wasserstoffbrücken oder Stapelungs Wechselwirkungen (stacking).

2.2 Einbeziehung weiterer Atomtypen

Zur Beschreibung von Proteinen wird die Parametrisierung von Schwefel in DFTB notwendig, da einige Aminosäuren, wie etwa Cystein oder Methionin dieses Element enthalten. Schwefelatome finden ihre fundamentale Bedeutung bei der Stabilisierung der räumlichen Struktur der Proteine durch die Bildung von Disulfidbrücken. Bei der Parameterentwicklung für DFTB stellte sich heraus [24], daß es zur Beschreibung der sogenannten hypervalenten Schwefelverbindungen wichtig ist, die minimale sp-Basis um die unbesetzten 3d Orbitale des Schwefels zu erweitern. Probleme treten in der minimalen Basis auch schon bei SO₂ auf, dessen Bindungswinkel ohne die d-Funktionen um 13° unterschätzt wird. Oder auch

bei der S-C Bindung in $(\text{C}_6\text{H}_5)_2\text{SO}_2$, die ohne d-Basisfunktionen um 0.15 Å überschätzt wird. In der minimalen Basis findet im Allgemeinen ein zu starker Ladungstransfer zwischen S und O statt (was zu der Bindungsverlängerung führt), der durch die sogenannte 'back-donation' von Elektronen in die Schwefel d-Funktionen abgeschwächt wird. Daher ist die Einbeziehung von d-Orbitalen in die Basis unbedingt notwendig. Zwei Parameter sind hierfür zu festzulegen, den atomaren Energieeigenwert des d-Orbitals, ϵ_d , und den Kompressionsradius für das d Orbital, r_0^d . Bei der Bestimmung der Kompressionsradien ist folgendes zu beachten: Die unbesetzten Orbitale sind sehr langreichweitig, so daß für das d-Orbital eine stärkere Kompression als bei den besetzten Orbitalen nötig sein könnte. Andererseits können die d-Orbital die minimale Basis aber auch als 'diffuse' Funktionen erweitern, was für eher grössere Kontraktionsradien sprechen würde (und so auch in der Parametrisierung realisiert). Festzuhalten bleibt jedoch, daß eine vorsichtige Optimierung der Kontraktionsradien nötig ist und das sonst robuste Schema

$$r_0 = 2 * r_{cov}$$

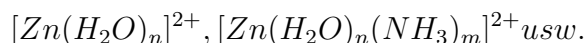
in diesem Fall nicht umstandslos anzuwenden ist. Der atomare Eigenwert ϵ_d bestimmt, wie stark die unbesetzten Orbitale in der Verbindung mit den besetzten Orbitalen mischen. Da die unbesetzten Orbitale durch die DFT Rechnungen i.A. energetisch zu niedrig liegend berechnet werden, ist gegebenenfalls auch eine Nachjustierung des Energiewerts notwendig, wovon im Falle des Schwefels abgesehen wurde, z.B. aber bei Silizium wichtig zu sein scheint.

Die Parametrisierung wurde an einem großen Satz organischer Moleküle, analog zu der Parametrisierung für O, N, C, H, getestet. Die Abweichungen von den experimentellen Daten für Reaktionsenergien, Geometrien, Schwingungsfrequenzen etc. sind durchaus mit den entsprechenden Werten aus DFT Rechnungen in der Qualität vergleichbar, vor allem aber zeigt sich eine deutliche Verbesserung gegenüber den getesteten semi-empirischen quantenchemischen Verfahren.

Die Erfahrungen bei der Parametrisierung des Schwefels konnten unmittelbar auf die Parametrisierung von Phosphor übertragen werden, wo ähnliche Parameterwerte verwendet wurden und eine vergleichbar gute Beschreibung molekularer Eigenschaften erhalten wurde [25].

Desweiteren wurden Parameter für die Wechselwirkung mit Zink entwickelt [26]. Zink ist zwar durch die gefüllte 3d und 4s Schale nicht kompliziert zu parametrisieren, doch waren einige Unsicherheiten bezüglich der Parametrisierungsprozedur zu bedenken, deren Lösung paradigmatisch für die Einbeziehung von 3d Elementen ist. Zink ist ein biologisch wichtiges Ion, es liegt meist zweifach positiv geladen vor. Zum Einen hat es katalytische, zum Anderen wichtige strukturelle Funktionen, indem es Proteinkonformationen durch seine Anlagerung stabilisieren kann. Obwohl es hauptsächlich im doppelt positivierten Zustand vorliegt, wurde als Startdichte n_0 die Dichte es neutralen Zink Atoms verwendet. Dies ist deswe-

gen empfehlenswert, da dann die Energiebeiträge, die in E_{rep} zusammengefasst sind, keine langreichweitigen Coulombanteile aufweisen, damit E_{rep} insgesamt also eher kurzreichweitig wird. Desweiteren wurde an neutralen, fiktiven Molekülen der Form H_nZnXH_m ($X=O, N, C, S, P, H$) angepasst und die freien Valenzen durch eine variable Anzahl von Wassertoffatomen n und m abgesättigt. Angepasst wurde E_{rep} demnach für eine 'kovalente' Bindung, was nicht repräsentativ für die Bindungssituationen in Proteinen ist. Daher wurden umfangreiche Tests an für die Biologie relevanten Modellkomplexen vorgenommen, z.B. an:



Sowohl die Struktur, als auch die Energetik (Bindungsenergien, Deprotonierungsenergien) werden hervorragend im Vergleich zu aufwendigen DFT Rechnungen wiedergegeben. Mit dieser Entwicklung können nun eine Vielzahl von katalytischen Prozessen in Zink Proteinen beschrieben werden. In Zukunft werden weitere 3d Elemente (Cu, Fe) und Alkali Metalle in Anlehnung an die hier beschriebene Prozedur parametrisiert, was eine Beschreibung vieler der interessanten katalytischen Prozesse in der Biologie erlaubt.

2.3 Erweiterung fuer H-Bruecken

Im Vordergrund der Entwicklung steht die Anwendbarkeit der DFTB Methode auf Biomoleküle, speziell zunächst auf Proteine und DNA Fragmente. Wichtig für deren Struktur und Funktion sind die langreichweitigen Wechselwirkungen, die über Wasserstoff Brücken (H-Brücken) und Van der Waals (VdW) Wechselwirkungen vermittelt werden. Bisher wurden nur die Eigenschaften der kovalenten Bindungen, Bindungslängen, Frequenzen und die Potentialtiefen (Energetik) getestet. Die DNA Doppelhelix wird durch H-Brücken zwischen den beiden Strängen der Helix stabilisiert, ebenso Proteine, die durch H-Brücken entlang des Protein Rückrats in ihrer dreidimensionalen Struktur stabilisiert werden. Daher ist aber eine genaue Beschreibung der Struktur als auch der Energetik von Wasserstoffbrücken (H-Brücken) gebundenen Systemen unabdingbar. Die DFTB Methode wurde deshalb an einem grossen Satz von H-Brücken gebundenen Modellsystemen getestet [27, 28, 29, 30]. Die erhaltenen Geometrien aller Komplexe sind exzellent, wobei die Differenzen zu *ab-initio* Ergebnissen lediglich innerhalb deren Streubreite (je nach Methode und verwendetem Basissatz) liegen. Die energetische Ordnung verschiedener Konformere der H-Brücken-gebundenen Komplexe wird ebenfalls ausnahmslos korrekt wiedergegeben. Eine leichte Unterschätzung aller Bindungsenergien auf Grund der verwendeten minimalen Basis konnte durch eine Erweiterung der Basis für die H-Atome beseitigt werden [27].

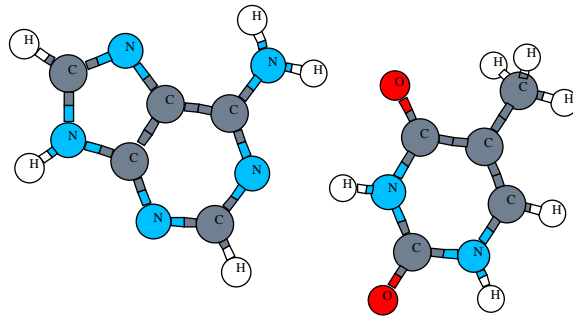


Abbildung 2.1:

2.4 Dispersion

Neben der gerade angesprochenen Stabilisierung der DNA durch H-Brücken zwischen den Basen auf verschiedenen Strängen, wie in Abbildung 2.1 dargestellt, ist die Wechselwirkung zwischen den Stufen der DNA durch Van der Waals (VdW) Kräfte elementar (Abbildung 2.2). Diese sogenannte Stapelungswechselwirkung resultiert aus den rein quantenmechanischen Energiebeiträgen zur Gesamtenergie, der Austausch- und Korrelationsenergie. Die Austauschwechselwirkung führt zur Repulsion, wenn sich die Elektronendichten der zwei molekularen Fragmente überlappen (Pauli Repulsion). Der attraktive Teil, die sogenannte Dispersionswechselwirkung resultiert aus der Elektronenkorrelation, die auf HF Niveau nicht berücksichtigt ist. Zahlreiche DFT Studien zu VdW Komplexen in den letzten Jahren zeigten eine ernstzunehmende Schwäche von DFT in der Beschreibung dieser Wechselwirkung, sie wird beispielsweise mit LDA Funktionalen stark überschätzt, mit einigen GGA Funktionalen wiederum treten starke repulsive Effekte auf, so daß DFT zum gegenwärtigen Zeitpunkt nicht ohne weiteres auf solche Komplexe anwendbar ist.

Da in der HF Methode nicht enthalten, ist die Dispersion erst auf dem MP2 Niveau beschreibbar. Größere DNA Fragmente sind aber mit MP2 aufgrund des Rechenaufwandes nicht zu behandeln. Im Prinzip wäre es denkbar, daß semiempirische Methoden wie AM1 und PM3 hier erfolgreich einsetzbar sind, da sie über die Parametrisierung Elektronenkorrelationen enthalten. Es stellt sich aber heraus, daß diese Methoden anstatt einer Attraktion der Basenpaare fälschlicherweise eine Repulsion beschreiben, was zu einer Instabilität der DNA Doppelhelix führt. Damit sind diese Methoden in ihrer gegenwärtigen Parametrisierung für diese Anwendung unbrauchbar.

Da die DFTB Methode auf der DFT basiert, hat sie auch deren Schwächen. Dies betrifft insbesondere auch die fehlerhafte Berücksichtigung der VdW (Dispersions) Kräfte. Innerhalb DFT liegt das an dem approximativen Charakter der Austausch- und Korrelationsfunktionale. Mit Kenntnis der exakten Funktio-

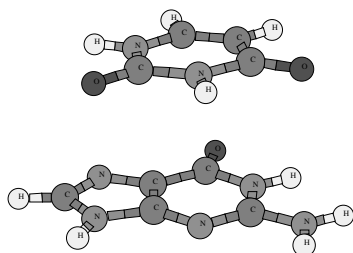


Abbildung 2.2:

nale wäre auch eine korrekte Behandlung der VdW Wechselwirkung möglich.

Durch die Parametrisierungsprozedur des DFTB werden aber genau die für die Dispersionswechselwirkung verantwortlichen Energiebeiträge vernachlässigt, selbst wenn sie exakt bekannt wären. Somit besteht die Möglichkeit, diese Wechselwirkung in empirischer Weise auf das DFTB Schema aufzusetzen, ohne die Korrelationsbeiträge doppelt zu berücksichtigen, wie es bei dem analogen Vorgehen für DFT der Fall wäre (da hier die Korrelationsenergie teilweise berücksichtigt ist). Hierzu wird eine empirische Formel für die Dispersion, die eine $1/R^6$ Abhängigkeit vom zwischenatomaren Abstand R aufweist und nur von atomaren Polarisierbarkeiten abhängt, zu dem DFTB Gesamtenergieausdruck addiert (DFTB+VdW) [31]. Allerdings wird diese empirische Formel für kurze interatomare Abstände R , bei denen die Elektronendichten der Atome zu überlappen beginnen, ungültig. Eine bloße Addition der Formel führt daher für kleine R zu zu großer Attraktion, weshalb ein Dämpfungsterm eingeführt wurde, der den Dispersionsanteil für kleine R zu Null führt.

In dieser Weise können die Energetik und Struktur der Basenpaarstapelung in hervorragender Übereinstimmung mit MP2 Rechnungen beschrieben werden [31]. Damit wird der Weg zu einer quantenmechanischen Beschreibung von Fragmenten der DNA Doppelhelix erstmals geebnet.

2.5 Struktur von Peptiden

Die Bausteine, aus denen Proteine und Polypeptide bestehen, sind die Aminosäuren. Die Abfolge der Aminosäuren entlang des Proteinrückgrats, die Sequenz, wird dessen Primärstruktur genannt. Die dreidimensionale Struktur von Polypeptiden und Proteinen entsteht durch die räumliche Anordnung von sogenannten Sekundärstrukturelementen, wie etwa α Helizes und β -Faltblattstruktur. Diese Sekundärstrukturen entstehen durch ein Wechselspiel zwischen den Torsionsbarrieren für Drehungen um einzelne Bindungen entlang des Proteinrückgrats und internen Wasserstoffbrücken, die zwischen den einzelnen Aminosäuren der jeweiligen Sekundärstruktur bestehen. Die Stabilität der Tertiärstruktur resultiert aus

den Wasserstoffbrücken oder Disulfidbrücken die zwischen den Sekundärstrukturelementen ausgebildet werden.

Die Abbildung 2.3 zeigt zwei solcher Sekundärstrukturen, die α Helix und die β Faltblattstruktur. Diese beiden Strukturen können durch Drehungen der Bindungswinkel im Peptid Rückgrat ineinander überführt werden.

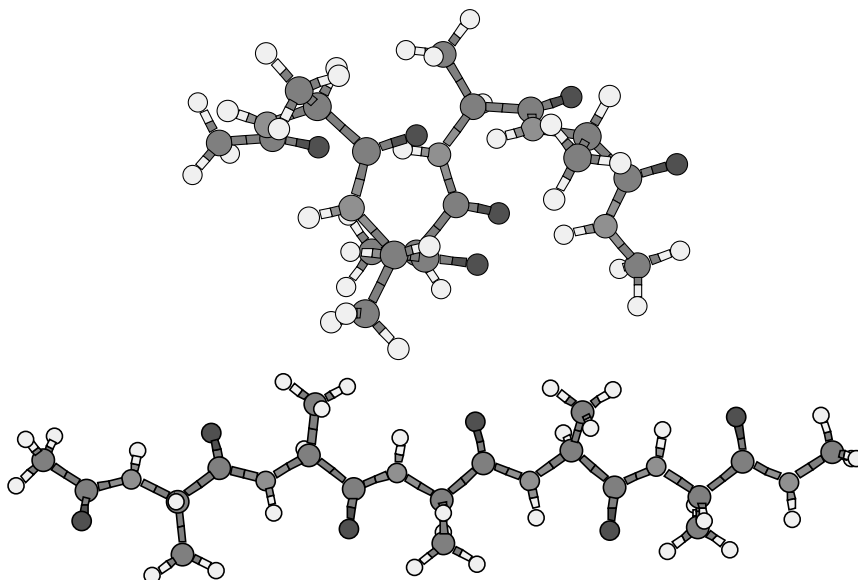


Abbildung 2.3:

Wie von anderen approximativen Methoden (etwa AM1 und PM3) bekannt, geben Tests an kleineren Molekülen nur wenig Aufschluss über die Zuverlässigkeit der Methode für ausgedehnte komplexe biologische Strukturen. AM1 und PM3 reproduzieren die Energetik und Struktur kleiner organischer Moleküle mit guter Genauigkeit, können aber die Peptidbindung nicht adäquat beschreiben. Diese Methoden können dadurch die Energetik und Struktur verschiedener Konformere von Polypeptiden nur unzureichend wiedergeben.

Daher ist eine genaue Evaluierung approximativer Verfahren (das bezieht sich im engeren Sinne auf semi-empirische Methoden, generell aber auch auf *ab initio* und DFT Methoden) für die Anwendung auf Polypeptide und Proteine unbedingt erforderlich.

Um die Zuverlässigkeit des DFTB in der Beschreibung von Protein-Sekundärstrukturen zu dokumentieren, wurden poly-Alanin Ketten im Vergleich zu *ab initio*, DFT und semi-empirischen (AM1, PM3) Methoden untersucht. Die atomaren Strukturen, Dipolmomente und relativen Energien verschiedener Konformere werden auf der Grundlage des DFTB Verfahren mit einer hervorragenden Genauigkeit wiedergegeben [30, 27, 32].

In einer weiteren Arbeit [34] wurden die Geometrien und relativen Energien zweier Helixtypen untersucht, die in Proteinen am häufigsten auftreten: der α und der 3_{10} Helix [34]. Die α Helix wird durch interne Wasserstoffbrücken stabilisiert, die zwischen der i -ten und der $i+4$ -ten Aminosäure geformt werden, bei der 3_{10} Helix bestehen diese zwischen der i -ten und $i+3$ -ten Aminosäure entlang des Proteinrückgrats. Aufgrund der unterschiedlichen Verknüpfungen ist die 3_{10} Helix dünner und länger als die α Helix, wie in Bild 2.4 zu sehen ist. Entgegen Rechnungen mit klassischen Kraftfeldern, wurden mit DFTB und DFT die 3_{10} Helizes als energetisch stabiler beschrieben. Auch sind die α Helizes für Ketten mit weniger als 8 Einheiten strukturell nicht stabil, sie relaxieren in 3_{10} Helizes. Ebenfalls ein Widerspruch zu den Rechnungen mit empirischen Kraftfeldern. Die DFTB und DFT Energien und Strukturen stimmen sehr gut überein, eine hervorragende Bestätigung für die DFTB Methode, die damit in der Lage ist, diese feinen Energieunterschiede, wie sie bei Konformationsänderungen auftreten, zu beschreiben. Die α Helizes werden erst durch die Einbeziehung von Lösungsmittelleffekten stabiler als die 3_{10} Helizes. Diese Problemstellung wird weiter unten noch fortgeführt.

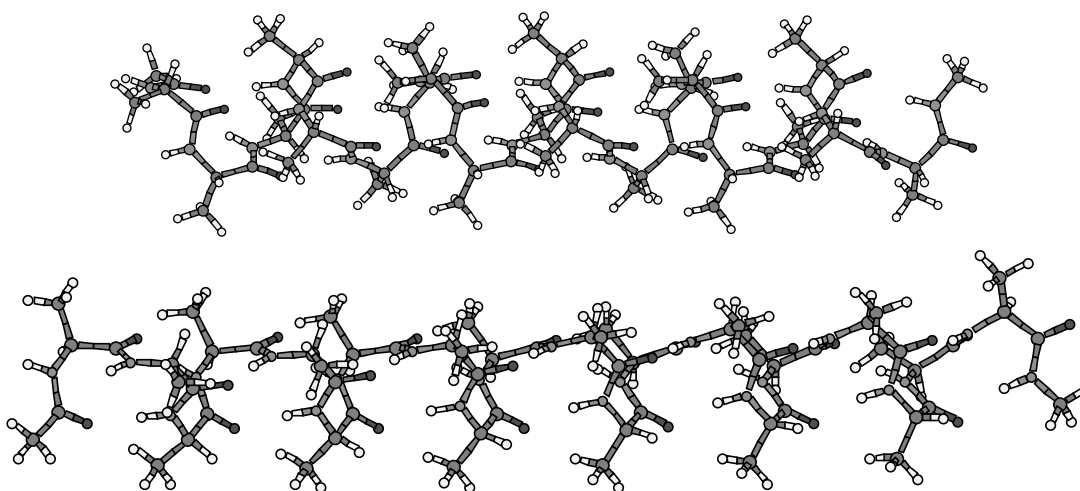


Abbildung 2.4:

Um die Anwendbarkeit für Molekulardynamiksimulationen zu evaluieren, wurden die Schwingungsfrequenzen von kleinen Peptiden in Abhängigkeit von der Konformationen mit der DFTB-Methode im Vergleich mit MP2 und DFT untersucht. Für die Standardabweichungen der Frequenzen von den experimentellen Werten mit MP2 und DFT werden 3.0%, und 4.4% ermittelt, während das DFTB Verfahren mit 6.7 % nur unwesentlich schlechter abschneidet [33]. Speziell ist die Methode in der Lage, die sehr feinen Differenzen in den Frequenzen zwischen den verschiedenen Konformeren zu reproduzieren.

Desweiteren wurden IR und Raman Intensitäten mit Hilfe des DFTB auf zwei Näherungsstufen berechnet. Zum Einen wurden die Geometrien und Hessische mit dem DFTB Verfahren gewonnen, und die IR und Raman Tensoren für die DFTB-Geometrie mit Hilfe der DFT berechnet. Dieses Hybridverfahren ermöglicht die Berechnung von IR und Raman Intensitäten, die sehr gut mit den vollen *ab initio* Werten übereinstimmen [33]. Ein weitergehender Näherungsschritt besteht darin, die IR Intensitäten ganz in der DFTB Näherung unter Verwendung der Mulliken Näherung für die (Ableitung der) Dipolmomente zu berechnen. Dieses Vorgehen ermöglicht sehr effizient die Berechnung von IR Spektren. Qualitativ werden die Spektren reproduziert, man findet allerdings eine Überschätzung der Intensität der intensiven Moden relativ zu den Moden mit geringer IR Intensität. Dennoch zeigen die Spektren, eventuell nach einer Skalierung der intensiven Moden, alle qualitativ wichtigen Merkmale der 'fingerprint' Regionen, die zur Identifikation der Konformation notwendig sind [27].

2.6 Struktur des Retinals in Bakteriorhodopsin

Für eine Vertiefung des Verständnisses der Effekte der Proteinumgebung auf die strukturelle und elektronische Eigenschaften des Chromophors in dem Protein Bakteriorhodopsin (siehe Einleitung) müssen alle Einflußfaktoren aus der Proteinumgebung mit einer hinreichenden Genauigkeit in den Rechnungen berücksichtigt werden. Die alleinige Einbeziehung einzelner Aminosäuren in der Bindungstasche kann den Einfluß des Gesamtproteins einschließlich seiner Lösungsmoleküle auf die RSB nur unvollständig beschreiben. Die Wechselwirkung der Proteinumgebung mit dem Chromophor ist vielmehr durch kollektive Effekte geprägt, die durch die Dynamik einer Vielzahl von geladenen, polaren bzw. polarisierbaren molekularen Gruppen hervorgerufen werden. Die Einbeziehung all dieser Gruppen in dynamische Simulationen wird auch in der nahen Zukunft bei weitem die Leistungsfähigkeit von *ab-initio* Rechnungen übersteigen. Andererseits wären die traditionellen semi-empirischen AM1 und PM3 Verfahren zwar aufgrund ihrer Rechengeschwindigkeit für die Behandlung entsprechender Fragestellungen geeignet, versagen aber vollständig durch zu große Defizite in der erreichbaren Genauigkeit.

Deshalb wurden *ab-initio* und DFT Rechnungen als Datenbasis für umfangreiche Tests der DFTB Methode genutzt. Zunächst wurden die Geometrien, Protonenaffinitäten und Isomerisierungsbarrieren für einige Modelle des Chromophors berechnet. Desweiteren wurde der Einfluß einzelner Aminosäuren und Wassermoleküle der Proteinumgebung auf die Geometrie und Ladungsverteilung des Chromophors untersucht. Alle Ergebnisse der DFT Rechnungen wurden dabei hervorragend reproduziert. Deprotonierungsenergien, Rotationsbarrieren und Geometrieänderungen werden ebenso mit hoher Präzision wiedergegeben wie Trends

dieser Größen in Abhängigkeit von der Länge des betrachteten Polyenes oder von der Einbeziehung einzelner geladener Gruppen in der Umgebung [35]. Die DFTB Methode ist damit gegenwärtig das einzige Näherungsschema, mit dem in einer mit semiempirischen Verfahren vergleichbaren Effizienz die Wechselwirkung von Retinoiden in ihrer Proteinumgebung voll quantenmechanisch mit hoher Genauigkeit behandelt werden kann.

Kapitel 3

Grosse Systeme

3.1 Kombination von Methoden:QM/MM

Die Idee von QM/MM Methoden besteht darin, einen Teil des Systems quantenmechanisch zu beschreiben (etwa das Reaktionszentrum, wenn eine genauere Beschreibung nötig ist oder wenn chemische Reaktionen modelliert werden sollen), und den Rest des Systems mit einem empirischen Kraftfeld.

Die QM/MM Methodik wurde ursprünglich eingeführt, um die enzymatische Katalyse untersuchen zu können [44, 45]. In den letzten zwei Jahrzehnten wurden diese Methodik von verschiedenen Gruppen weiterentwickelt, wobei im QM Bereich hauptsächlich semi-empirische Methoden eingesetzt wurden. Erst in den letzten Jahren werden vermehrt auch *ab initio* (vornehmlich HF) und DFT Methoden im QM Bereich verwendet.

Am Beispiel des bR (siehe Einführung) soll die QM/MM Methode kurz erläutert werden. Die Proteinumgebung hat einen großen Einfluß auf die räumliche Struktur und die elektronischen Eigenschaften des Retinals. Ohne diese Umgebung könnte das Retinal seine Rolle in der Lichtabsorption- und Umwandlung nicht ausführen. Daher ist die Berücksichtigung der Proteinumgebung in der theoretischen Modellierung des Retinals unumgänglich. Das konjugierte π -Elektronensystem des Retinals kann aber nur mit Hilfe von quantenmechanischen Methoden beschrieben werden. Diese zwei Anforderungen, die Notwendigkeit der quantenmechanischen Beschreibung des Retinals zusammen mit der Grösse des gesamten bR Moleküls machen die Anwendung moderner QM/MM Techniken, wie oben vorgestellt, zwingend. Hierbei kann das Retinal zusammen mit einigen wichtigen funktionalen Gruppen der Proteinumgebung quantenmechanisch, der Rest des Moleküls aber mit den wesentlich weniger aufwendigen empirischen Kraftfeldmethoden behandelt werden. Auf diese Weise kann der Einfluß der Proteinumgebung auf das Retinal, sowie der Protonentransportprozess untersucht werden.

Die Kopplung zwischen beiden Beschreibungsweisen erfolgt über ein *link-*

atom-Konzept und berücksichtigt sowohl elektrostatische als auch Van-der-Waals Wechselwirkungen.

Die Gesamtenergie für den QM/MM Formalismus lautet:

$$E_{tot} = E_{QM} + E_{MM} + E_{QM-MM}, \quad (3.1)$$

wobei E_{QM} als Energie des quantenmechanischen Teils durch die DFTB Gesamtenergie repräsentiert wird, der Energiebeitrag E_{MM} durch die Energiefunktion des klassischen empirischen Kraftfeldes gegeben ist. E_{QM-MM} beschreibt die Kopplung der beiden Teilgebiete.

Das Schema wurde zuerst für den Fall implementiert [28], daß durch die Systemgrenzen keine kovalenten Bindungen verlaufen. In diesem Fall besteht E_{QM-MM} nur aus Coulomb und Van der Waals Wechselwirkungen. Im DFTB Schema lässt sich dies besonders einfach ausdrücken:

$$E_{QM-MM} = \sum_{Ak} \frac{Q_A \Delta q_k}{\mathbf{R}_{Ak}} + E_{VdW}$$

Q_A sind die Ladungen auf den MM Atomen, Δq_k die Mulliken Ladungen aus dem DFTB und \mathbf{R}_{Ak} ist der Richtungsvektor zwischen Atom A und Atom k. E_{VdW} gibt die VdW Wechselwirkung zwischen den Atomen aus den beiden Bereichen wieder, wobei auch für den QM Bereich die Parameter des MM Kraftfeldes verwendet werden.

Rechnungen für H-Brücken gebundene Komplexe zeigen eine sehr gute Übereinstimmung mit *ab-initio* Resultaten [28] für Geometrien, Bindungsenergien und die energetischen Ordnung verschiedener Konformere. Es konnten auch erstmals die oben diskutierten Polyalanin Helices in einer Umgebung von Wassermolekülen untersucht werden, wobei die Wassermoleküle in dem MM Bereich liegen und die Helix Atome durch QM beschrieben werden. Allerdings wurden nur Geometrienoptimierungen durchgeführt, was die Aussage der Rechnungen sehr limitiert hat. Dennoch konnte man eine energetischer Favorisierung der α Helix feststellen. Die Analyse ergab, daß die etwas größeren Dipolmomente der α Helix die Energieerniedrigung im Vergleich zur 3_{10} Helix erklären können.

Durchschneidet die QM-MM Grenze eine kovalente Bindung, wird das Problem etwas komplexer. Die einfachste Lösung besteht in dem sogenannten 'link atom' Konzept. Hierbei wird die QM Region durch ein Wasserstoffatom abgesättigt, das entlang der Achse der durchgeschnittenen Bindung im entsprechenden Abstand von dem QM Atom plaziert wird. Die Wechselwirkung zwischen den Atomen der durchgeschnittenen Bindung wird durch das empirische Kraftfeld vermittelt. Das 'link atom' Wasserstoff ist dabei ein Fremdkörper, dessen Einfluß auf das System in vielen Fällen vernachlässigt werden kann. Dies wurde auch für die DFTB Methode intensiv getestet [29]. Es wurden eine Reihe von Protonentransferprozessen in dem Enzym Triosephosphate Isomerase (TIM) durchgeführt [29].

Im Vergleich mit DFT-Rechnungen ergibt sich eine mittlere Abweichung von 2-4 kcal/mol, während die semi-empirischen Methoden AM1 und PM3 mittlere Abweichungen von 11 kcal/mol aufweisen. Zusätzlich hervorzuheben ist, daß die Geometrien der Übergangszustände sehr gut mit den DFT-Geometrien übereinstimmen. In der gleichen Arbeit wurde auch die Stabilität der Polyalanin Helices in Lösung untersucht, diesmal allerdings mit Hilfe einer QM/MM MD Simulation. Hier zeigte sich, daß die 3_{10} Helix mit acht Aminosäureeinheiten in Lösung dynamisch instabil ist, sie geht innerhalb von 10 ps Simulationszeit in eine α Helix über.

Leztere Simulationen dokumentieren die Begrenztheit der QM/MM Methoden. Simulationen mit einem QM Bereich dieser Größe werden sehr rechenzeintensiv, wobei die meiste Rechenzeit im QM Teil bei der Diagonalisierung verbraucht wird. Für grössere Systeme empfiehlt sich daher die Verwendung von $O(N)$ Algorithmen in der QM Region.

3.2 $O(N)$ Methoden

Die Entwicklung von linear skalierenden Methoden stellt einen Durchbruch in der Modellierung großer biologischer Systeme dar. W. Yang entwickelte als erster einen solchen Algorithmus, das sogenannte 'divide-and-conquer' (DAC) Schema [38, 39]. In den folgenden Jahren wurden weitere $O(N)$ Verfahren vorgeschlagen, die auf ähnlichen Prinzipien beruhen. Einen zusammenfassenden Überblick über die Methoden geben Yang und Perez-Jorda [40]. Geodecker [41] und Galli [42]. Kombiniert man $O(N)$ mit approximativen Verfahren wie AM1, PM3 oder DFTB, wird die Behandlung von Systemen mit mehreren 1000 Atomen möglich [43].

In den meisten $O(N)$ Methoden wird die räumliche Lokalisierung der Elektronendichte ausgenutzt, um zu einer linearen Abhängigkeit der Rechenzeit von der Systemgröße zu kommen. In diesem Fall läßt sich die elektronische Struktur des Systems aus lokalen Größen, etwa aus den lokalisierten molekularen Orbitalen oder einer lokalisierten Dichtematrix aufbauen. Eine Dichtematrixformulierung der Methode von Yang und Lee vereinfacht den Algorithmus weitergehend [39]. Wenn die Dichtematrix

$$\rho_{\mu\nu} = \sum_i n_i c_\mu^i c_\nu^i \quad (3.2)$$

räumlich lokalisiert ist (nicht in Metallen), läßt diese sich aus Beiträgen von lokal ermittelten Dichtematrizen aufbauen. Dies geschieht folgendermassen: Man teilt das zu behandelnde System in M Untersysteme, die sich überlappen. Dann wird das Eigenwertproblem durch Diagonalisierung für die Untersysteme gelöst, man erhält die c_μ^i für jedes Untersystem. Anstatt also das Eigenwertproblem für das Gesamtsystem zu lösen, löst man es für die M Untersysteme. Aus den lokal

berechneten Koeffizienten c_μ^i wird dann die Dichtematrix Gl. (3.2) konstruiert. Um die Besetzungszahlen der Eigenzustände der Subsysteme zu finden, wird das chemische Potential in die Rechnung eingeführt. Die Eigenzustände werden bis zu diesem Energiewert besetzt, und das chemische Potential wird solange iteriert, bis die berechnete Elektronenzahl mit der Gesamtelektronenzahl des Systems identisch ist. Da mit wachsender Systemgröße die Anzahl der Subsysteme linear anwächst, wächst somit der Gesamtrechenaufwand linear. Für kleine Systeme ist allerdings die DAC Methode aufwendiger als die direkte Diagonalisierung, da bei DAC mehrere Subsysteme diagonalisiert werden müssen. Erst ab einem 'cross over' Punkt wird die DAC Methode effizienter. Dieser liegt bei einer Kombination mit DFTB bei etwa 200 Atomen (bei etwa 40% Wasserstoffatomen, der Rest der Atome wird mit minimaler sp Basis beschrieben). Ein Problem entsteht nur an den Grenzen der Subsysteme, wenn die Überlappungen zu klein gewählt sind, so daß Beiträge zur Dichtematrix vernachlässigt werden, die noch signifikant für deren Darstellung sind. Die Qualität der Methode wird somit über die Größe der Überlappungsregion bestimmt. Ist diese sehr groß so wird die Methode sehr genau, aber die Rechenzeit steigt. Ist sie zu klein, werden die Energien und atomaren Kräfte ungenau. Für die DFTB Methode wurde der Überlappbereich durch ausprobieren bestimmt. Dabei wurden Alanin Polypeptide mit 20 Aminosäureeinheiten (212 Atome) verwendet und die Energien und Kräfte für verschiedene Konformationen mit verschiedenen großen Überlappbereichen ermittelt. Ein Subsystem wurde als eine Aminosäure definiert. Zur Bestimmung der Überlappbereiche wurden zwei Kriterien verwendet: Zum Einen wurden N Nachbarn jedes Atoms in Subsystem M berücksichtigt, zum Anderen alle Atome im Abstand R_b eines jeden Atoms aus dem Subsystem M. Die Vereinigung dieser beiden Mengen ergab die Überlappmenge. In Tabelle 3.1 sind die Ergebnisse der Testrechnungen dargestellt. Zwischen den Konformationen, wie durch Diagonalisierung erhalten, bestehen Energieunterschiede von etwa 10 kcal/mol. Eine Kombination $R_b = 3.5 \text{ \AA}$ und $N=3$ ergibt einen Fehler, der in der Größenordnung der Energieidifferenzen zwischen den Konformeren liegt. Erst die Kombination $R_b = 4.0 \text{ \AA}$ und $N=4$ führt zu akzeptabel kleinen Fehlern. Ein Fehler von 5 kcal/mol entspricht einem Fehler von 1meV pro Atom im System, was 0.03 % der kohäsiven Energie entspricht.

Die verschiedenen Konformationen benötigen unterschiedliche Rechenzeiten im DAC Formalismus, da je nach Kompaktheit des Moleküls die Überlappregionen verschieden groß ausfallen. Bei der gewählten Größe der Überlappregionen erhält man eine Halbierung der Rechenzeit im Vergleich zur Diagonalisierung.

Als erste Anwendung wurde das Protein Crambin betrachtet (Abb. 3.1). Dieses Molekül enthält 46 Aminosäuren (=46 Subsysteme in der Rechnung) und insgesamt 639 Atome.

Die volle Diagonalisierung benötigt etwa 15 Minuten (auf einer Dec alpha

Konformation	C_7^{eq}	C_5^{ext}	3_{10}	α	Rechenzeit[s]
Diagonalisierung	0.0	34.5	-11.9	-32.0	85
$R_b(\text{Å})$ N					
3.5 3	16.2	13.4	7.1	5.3	16-37
4.0 4	3.1	2.3	1.6	3.2	22-57
5.0 4	0.4	2.1	0.7	1.4	35-110

Tabelle 3.1: Die relativen Energien (kcal/mol) der verschiedenen Konformere, wie sie durch Diagonalisierung erhalten werden (2. Zeile). In den folgenden Zeilen werden die Fehler in Bezug auf die exakten Energien in Abhängigkeit von R_b und N wiedergegeben. Die letzte Spalte gibt die Rechenzeiten (s) wieder.

ev6), während die DAC Rechnung mit den obigen Parametern nur eine Minute benötigt. Die Geometrie des Moleküls wurde einer Geometrieoptimierung unterzogen, wobei die Startstruktur, die der Kristallstruktur entspricht, sehr gut erhalten wird.

3.3 Kombinierte QM/MM O(N)

Da die biologischen Moleküle meist in (wässriger) Lösung oder in der Zellumgebung sind, machen Rechnungen an isolierten Molekülen wenig Sinn: Umgebungseinflüsse können die elektronische und geometrische Struktur stark beeinflussen. Um auch Simulationen zu Systemen mit grossen QM Regionen durchführen zu können, wurde das DAC Verfahren in einen QM/MM Algorithmus implementiert.

Als erstes Testbeispiel wurde das Protein Crambin, umgeben von etwa 3000 Wassermolekülen betrachtet. Das Crambin Molekül wurde hierbei mit der linear skalierenden DFTB-Methode behandelt, die Wassermoleküle mit einem empirischen Kraftfeld. Für 2 Picosekunden Molekulardynamiksimulation dieses Systems von 7000 Atomen, benötigt eine DEC/alpha Workstation ca. 3 Tage [46]. Dieses Programmsystem wurde auch parallelisiert, so daß nun für Systeme der oben genannten Grössenordnung MD Simulationen über 100ps in etwa 2 Wochen auf 16 Prozessoren durchgeführt werden können.

Die Simulation ergab, daß Crambin strukturell nicht stabil ist. Crambin enthält zwei α Helices, und beide gehen während der MD in 3_{10} Helices über. Nachfolgend wurden Simulationen durchgeführt, in denen die Dispersionswechselwirkung mit in die Rechnung einbezogen wurde. Damit war die Struktur von Crambin stabil. Im Detail wird dieses Ergebnis weiter unten diskutiert. Für die geometrischen Parameter wurde eine hervorragende Übereinstimmung mit den experimentellen Daten gefunden, eine bessere Übereinstimmung sogar als die empirischen Kraftfeldmethoden erzielten, die zum Vergleich herangezogen wurden. Damit wurde

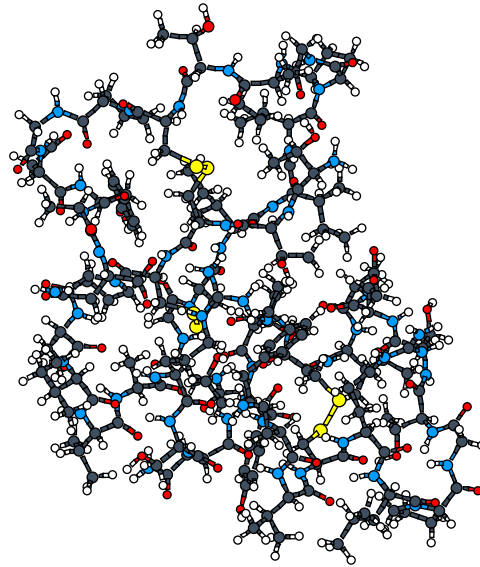


Abbildung 3.1: Das Protein Crambin

gezeigt, daß die DFTB + Dispersion Methode sehr gut geeignet ist, Proteinstrukturen zu beschreiben. Weiter wurde ein nichtlokaler Ladungstransfer innerhalb des Proteins gefunden, der von empirischen Potentialen nicht beschreiben werden kann. Inwieweit dieser Ladungstransfer für die Struktur wichtig ist, oder ob es auch Ladungstransfer zwischen Protein und Wasser gibt, bleibt nachfolgenden Untersuchungen vorbehalten.

Kapitel 4

Stabilität von Protein Sekundärstrukturen

Eine Reihe von Arbeiten haben sich mit der Stabilität von Sekundärstrukturelementen kurzer Polypeptide beschäftigt [34, 32, 27, 28, 29], insbesondere mit der relativen Stabilität der 3_{10} vs. der α_R Helix.

Unter Verwendung von DFT und DFTB Methoden wurden die Geometrien von 3_{10} und α_R Helices optimiert. Dabei stellte sich heraus, daß zwar die 3_{10} Strukturen stabil für alle untersuchten Kettenlängen sind, die α_R Helices aber erst ab einer bestimmten Länge von etwa acht Aminosäureeinheiten. Für $N=8$ (N : Anzahl der Aminosäureeinheiten) ist nur eine α_R typische $i \rightarrow i+4$ H-Brücke zu finden, so daß der Großteil der Helix, nämlich die beiden Enden, in der 3_{10} Konformation vorliegt.

In Proteinen, wie etwa Crambin, findet man aber α_R Fragmente von der Länge $N=6-8$. Daher wurde die relative Stabilität der zwei Helices für Dihedralwinkel in dem Proteinrückgrat untersucht, die typischerweise in Proteinen aufzufinden sind. Während der Geometrieoptimierung wurden die Dihedralwinkel bei den idealen Werten festgehalten. Dabei stellte sich heraus, daß die α_R erst für $N \gg 10$ stabiler als die 3_{10} Helix wird. Die Rechnungen beziehen sich auf Differenzen von Gesamtenergien, nicht von freien Energien. Allerdings scheinen die Entropiebeiträge zusätzlich die 3_{10} Helix zu favorisieren. Dies steht im Gegensatz zu den experimentellen Ergebnissen.

Die Stabilisierung der α_R gegenüber der 3_{10} Helix findet durch Umgebungseinflüsse statt. So fanden wir durch die oben beschriebenen QM/MM Rechnungen, daß für $N=8$ die α_R Helix in wässriger Lösung stabiler wird. Der Wert der dielektrischen Konstante von Wasser liegt bei $\epsilon=78$, die Vakuumrechnungen unterstellen einen Wert von $\epsilon=1$, während in der Proteinumgebung ein Wert von etwa $\epsilon=4-8$ angenommen werden kann. Unterstellt man eine monotone Abhängigkeit der relativen Stabilitäten von dem Wert des Dielektrikums, so bewirkt die Pro-

teinumgebung in jedem Fall eine Stabilisierung der α_R Konformation relativ zur 3_{10} Helix. Die Simulation an dem Crambin Molekül hat jedoch ergeben, daß der Einfluß der Proteinumgebung nicht ausreicht, um die zwei α_R Helices in Crambin zu stabilisieren. Diese gehen im Verlauf der Dynamik innerhalb von 30 ps in 3_{10} Konformationen über, im Gegensatz zu den QM/MM Rechnungen in Wasser [29].

Auf der anderen Seite ergab eine MD Simulation unter Einbeziehung der Dispersionskräfte [46] in dem DFTB Formalismus strukturell stabile α Helices in Crambin. Die Dispersion (Korrelation) scheint daher die α Helix gegenüber der 3_{10} Helix zu favorisieren. Die Dispersionskräfte haben eine reine $1/R^6$ Abhängigkeit, kompaktere Moleküle, wie die α_R Strukturen weisen daher eine grössere Dispersionswechselwirkung auf als die eher längliche und dünnere 3_{10} Konformation.

Um dieses Ergebnis etwas genauer zu verstehen, wurden abermals Rechnungen an Polyalanin Helices durchgeführt. Wiederum wurden die Geometrien beider Helices für $N=3,5,8$ bei festgehaltenen Dihedralwinkeln des Peptidrückrades optimiert. Es wurden HF, MP2, DFTB und DFTB+Dis Energien für die jeweilige Konformation bestimmt. Mit MP2 sind wegen der langen Rechenzeit Geometrieoptimierungen schwer, bzw. z. Z. nicht durchführbar. Die Ergebnisse sind in Tabelle 4 wiedergegeben.

N	MP2	HF	B3LYP	DFTB	DFTB+vdW
2	3.5	3.3	3.6	2.9	2.9
3	4.8	5.4	4.9	3.8	3.3
5	5.1	7.6	6.3	5.2	3.3
8	2.6	8.9	6.1	5.9	1.4

Tabelle 4.1: Die relativen Stabilitäten von α_R und 3_{10} (in Kcal/mole) für verschiedene Kettenlängen N . Die positiven Werte bezeichnen den Energieunterschied zwischen der stabileren 3_{10} Helix zu der α_R Helix.

MP2 und DFTB+vdW beschreiben eine wesentlich schnellere Stabilisierung der α_R Helix, ein crossover ist für $N=9,10$ zu erwarten. Auffallend sind die größeren Energieunterschiede in der MP2 Methode. In vorhergehenden Arbeiten wurde aber eine generelle Tendenz des DFTB gefunden, kleinere Energiedifferenzen zwischen den Konformationen vorherzusagen als die *ab initio* und DFT Methoden. Dies muß nicht ein Defizit der DFTB Methode sein, es gibt Hinweise, daß die Energiedifferenzen aufgrund des internen 'basis set superposition error' (BSSE) Effektes überschätzt werden [32]. Die 3_{10} Helix hat eine interne Wasserstoffbrücke mehr als die α_R Helix, und BSSE führt zu einer Überschätzung der Bindungsenergie von typischerweise 1-2 kcal/mole für Basissätze mittlerer Qualität. wie den

hier verwendeten. Daher kann man annehmen, daß die Energie der 3_{10} Helix etwa um diesen Betrag unterschätzt wird, die Energiedifferenz also um diesen Betrag zu groß ist. Die DFTB+VdW Werte stünden dann aber in sehr guter Übereinstimmung mit solcherart 'BSSE-korrigierten' *ab initio* Werten. Leider ist zur Zeit keine Methode verfügbar, die den internen BSSE korrigieren kann. Man kann nur auf Rechnungen zu kleinen Modellpeptiden zurückgreifen, die diese Diskussion zu unterstützen scheinen [32].

Den Werten in Tabelle 4 liegen die Geometrien aus DFT Rechnungen mit festgehaltenen Dihedralwinkeln des Peptidrückgrades zu Grunde. Mit dieser Vorgehensweise ist der absolute Kreuzungspunkt der relativen Stabilität nicht genau bestimmbar, dazu sind volle Geometrieoptimierungen nötig. Um dies zu bewerkstelligen, wurden die Geometrien für N=4-12 mit der DFTB+VdW Methode optimiert. Für N=4 findet noch ein Übergang der α_R in die 3_{10} Konformation statt. Aber schon für N= 5 bilden sich stabile $i \rightarrow i+4$ H-Brücken aus, also partiell stabile α_R Helices, wobei aber erst bei N=7-8 der Kreuzungspunkt in den Stabilitäten liegt. Im Vergleich zu der Situation ohne Dispersionswechselwirkung bedeutet dies, daß der Kreuzungspunkt bei sehr viel kleineren Peptiden zu finden ist, als vormalig angenommen.

Dieser Wert lässt sich sehr gut mit dem Wert N=6 aus der Analyse der Strukturen in den Proteindatenbanken vergleichen. Bis N=6 ist die 3_{10} Konformation stabil, für längere Ketten findet man α_R . Die Umgebungseinflüsse werden die α_R Konformation zudem zusätzlich begünstigen.

Zur Berechnung von Proteinstrukturen ist also eine korrekte Einbeziehung der Dispersionskräfte unumgänglich. HF, aber auch DFT Rechnungen sind damit nicht in der Lage, Proteinstrukturen i. A. vorherzusagen, es ist zu erwarten, daß auch bei diesen Methoden α_R im Protein nicht stabil sind.

Dagegen scheint die Einbeziehung der Dispersionskräfte in den DFTB sehr vielversprechend zu sein, die gegenwärtigen Ergebnisse lassen eine zuverlässige Beschreibung von Proteinstrukturen erwarten. Da HF und DFT Methoden z.Z. nicht uneingeschränkt anwendbar scheinen, andere semi-empirische Methoden wie AM1 und PM3 Proteinstrukturen sehr schlecht beschreiben, scheint DFTB+VdW die z.Z. einzige Alternative zu empirischen Kraftfeldmethoden zu sein.

Kapitel 5

Optische Eigenschaften

Die theoretische Beschreibung optischer Eigenschaften ist sehr aufwendig. Wellenfunktionsbasierte Methoden sowie Propagatorstechniken erzielen zwar eine sehr gute Übereinstimmung mit dem Experiment, sind gleichzeitig aber extrem rechenintensiv und können nur für sehr kleine Systeme angewendet werden.

Die DFT, ursprünglich nur streng gültig für den Grundzustand, ermöglicht in ihrer Erweiterung auf zeitabhängige Dichten ('time dependent' DFT - TD-DFT) einen robusten und einfachen Zugang zur Berechnung von Anregungsenergien im Rahmen der 'linear response' (LR) Theorie.

Dazu ist in einem ersten Schritt eine DFT Grundzustandsrechnung zur Bestimmung der Kohn-Sham Orbitale $\psi_i(\mathbf{r})$ und Kohn-Sham Energien ϵ_i nötig. In einem zweiten Schritt wird die sogenannte Kopplungsmatrix $K_{ij\sigma,kl\tau}$ berechnet:

$$K_{ij\sigma,kl\tau} = \iint \psi_i(\mathbf{r})\psi_j(\mathbf{r}) \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E_{xc}}{\delta n_\sigma(\mathbf{r}) \delta n_\tau(\mathbf{r}')} \Big|_{n_0} \right) \psi_k(\mathbf{r}')\psi_l(\mathbf{r}') d\mathbf{r}d\mathbf{r}' \quad (5.1)$$

Die Differenzen von Kohn-Sham Energien $\omega_{ij} = \epsilon_i - \epsilon_j$ repräsentieren nicht die Anregungsenergien des Systems. Um diese zu erhalten, muss die im Angeregten Zustand modifizierte Elektron-Elektron Wechselwirkung, wie sie durch $K_{ij\sigma,kl\tau}$ berechnet wird, berücksichtigt werden. Dies geschieht durch Lösung des folgenden Eigenwertproblems:

$$\sum_{ij\sigma} \left[\omega_{ij}^2 \delta_{ik} \delta_{jl} \delta_{\sigma\tau} + 2\sqrt{\omega_{ij}} K_{ij\sigma,kl\tau} \sqrt{\omega_{ij}} \right] F_{ij\sigma}^I = \omega_I^2 F_{kl\tau}^I \quad (5.2)$$

Die ω_I sind die Anregungsenergien des Systems, aus den Eigenvektoren $F_{ij\sigma}^I$ können Oszillatorstärken berechnet werden.

Obwohl die TD-DFT wesentlich weniger rechenzeitintensiv ist als die herkömmlichen Wellenfunktionsbasierten Methoden, ist dennoch die Untersuchung von

Nanomaterialien mit mehreren 100 Atomen nicht möglich. Hierzu ist ein Formalismus nötig, der in etwa die Rechenzeiteffizienz von semi-empirischen Methoden hat, also zwei bis drei Grössenordnungen schneller ist, als DFT. Deshalb wurde der TD-DFT Formalismus auch im Rahmen des DFTB Schmas implementiert. Die DFTB Methode stellt die Kohn-Sham Orbitale und Energien zur Verfügung, ein Problem besteht in der effizienten Bestimmung von $K_{ij\sigma,kl\tau}$.

Im Rahmen der DFTB Methode wurde jedoch schon eine Näherung für ein Integral gefunden, das die gleiche Form wie $K_{ij\sigma,kl\tau}$ aufweist: Der zweite Term auf der rechten Seite in Gl. 2.1, die zweite Ableitung der Coulomb- und Austausch-Korrelationsenergie nach der Elektronendichte wurde durch die Funktion $\gamma_{\alpha\beta}$ approximiert.

Diese Näherung kann ebenfalls zur Bestimmung von $K_{ij\sigma,kl\tau}$ verwendet werden. Daraus resultiert ein Schema, das sehr effizient die Anregungsenergien von Molekülen und Clustern im Rahmen der DFTB Methoden bestimmen kann [47, 48]. Damit steht mit der TD-DFTB Methode eine sehr schnelle und genaue Methode zur Berechnung von Anregungsenergie biologischer Moleküle zur Verfügung. Durch Implementation von numerischen Energiegradienten sind in Zukunft Geometrieoptimierungen und MD Simulationen von Molekülen im elektronisch angeregten Zustand möglich.

Kapitel 6

Zeitabhängige Phänomene

Alle bisher diskutierten Methoden beruhen auf der Born Oppenheimer Näherung. In einigen Anwendungsfällen kann jedoch die Entkopplung der Dynamik der Elektronen von der der Atomkerne nicht mehr aufrecht erhalten werden. Dies betrifft zum Einen Stossprozesse, zum Anderen die Wechselwirkung von Materie mit elektromagnetischer Strahlung aber auch strahlungslose Übergänge beispielsweise in photochemischen Reaktionen, wie am Beispiel bR in der Einleitung beschrieben. Gemeinsam ist allen drei Prozessen, daß im Verlauf der Dynamik der Atomkerne elektronische Übergänge stattfinden. Um diese in die Beschreibung einbeziehen zu können, ist die Berücksichtigung der Zeitabhängigkeit des quantenmechanischen Zustandes notwendig. Hierzu kann man von einer modifizierten Lagrange-Funktion ausgehen, die über das Variationsprinzip gekoppelte Differentialgleichungen für die Elektronen und Kerne ergibt [48]. Die resultierende Schrödingergleichung für die Kohn-Sham-Orbitale innerhalb des DFTB-Schemas wird dann mit Hilfe eines modifizierten Cayley-Algorithmus' gelöst.

Bisher wurde das beschriebene Verfahren an *sys-anti*-Isomerisierungsreaktionen verschiedener Imine getestet und ergab sowohl bezüglich des Anregungsmechanismus' als auch der Beschreibung der angeregten Potentialhyperfläche gute Resultate. Da die benutzte Ankopplung des elektromagnetischen Feldes hinsichtlich der eingestrahlten Intensität keine Annahmen macht, erscheint eine weitere Anwendung der Methode auf Photanregungen von Schwingungsmoden und Photofragmentationsprozesse aussichtsreich. So wurde mit dieser Feldankopplung das Fullerene C₆₀ untersucht [49]. In Zukunft sollen mit dieser Methode die photoinduzierte Dynamik des Retinals in dem bR Protein untersucht werden. Mit dieser Methode können dann sowohl der Anregungsprozess, als auch der Übergang vom angeregten Zustand in den Grundzustand beschrieben werden.

Kapitel 7

Zusammenfassung und Ausblick

Ausgehend von der Entwicklung des DFTB wurden hier wesentliche Erweiterungen der Methode vorgestellt, die eine Anwendung auf biologische Probleme erst ermöglicht hat. Zum Einen wurden

- Parameter für neue Atomtypen entwickelt, die im biologischen Kontext wichtig sind. Dies betrifft etwa Schwefel, Phosphor, Zink und weitere Metalle und Übergangsmetalle.
- Auch wurde die Qualität der Beschreibung der schwachen Wechselwirkungen, der Wasserstoffbrücken und der VdW Wechselwirkung wesentlich verbessert, was erst eine stabile Beschreibung vieler wichtiger biologischer Strukturen ermöglicht

Desweiteren wurden zur Behandlung großer Strukturen, die die Möglichkeiten der DFTB Methode übersteigen, neue Algorithmen implementiert. DFTB wurde

- in einen QM/MM Algorithmus implementiert und damit mit einem empirischen Kraftfeld kombiniert
- in einen linear skalierenden ($O(N)$) Algorithmus implementiert
- in einen kombinierten $O(N)$ -QM/MM implementiert, der zudem parallelisiert wurde, was die Anwendung auf sehr große Strukturen und lange Zeitskalen erlaubt.

Damit sind alle z.Z. gängigen Methoden zur Verbesserung der Rechenzeiteffizienz für den DFTB verfügbar. Auch wurde die Methode zur Berechnung spektroskopischer Daten erweitert:

- IR Spektren können sehr effizient ermittelt werden. Die Qualität der Ergebnisse ist als semi-quantitativ zu bezeichnen, eine Skalierung der Ergebnisse würde die Übereinstimmung mit dem Experiment wesentlich verbessern

- Die Methode wurde für die Berechnung von optischen Anregungsspektren erweitert.

Es wurden erfolgreiche Anwendungen auf Protontransfer- und Katalyseprozesse, auf die Stabilität von Proteinsekundärstrukturen und die Dynamik eines Proteins in wässriger Lösung vorgestellt. Die Vorhersage von Geometrien und Schwingungsfrequenzen mit Hilfe des DFTB ist excellent, in Bezug auf die Energetik könnte man sich eine Verbesserung insofern vorstellen, da das z. Z. vorhandene 'overbinding' eine genaue Ermittlung von Übergangszustandsenergien nicht erlaubt. Desweiteren wäre die Entwicklung weiterer spektroskopischer Methoden, wie etwa die Vorhersage des Spektrums des Raman- oder resonanten Ramaneffekts, sehr hilfreich.

Insgesamt ist mit der hier vorgestellten Entwicklung ein Methodenspektrum entstanden, das eine echte Erweiterung der Methoden der Quantenchemie im Blick auf Nanostrukturen darstellt. Die DFTB Methode ist den *ab initio* und DFT Verfahren an Rechenaufwand überlegen, den quantenchemischen semiempirischen Methoden an Genauigkeit. Sie füllt damit eine Lücke im Methodenspektrum, die zur Behandlung nanoskaliger Materialien dringend benötigt wird.

Literaturverzeichnis

- [1] Foulkes W M C, Mitas L, Needs R J and Rajagopal G 2001 *Rev. Mod. Phys.* **73** 33.
- [2] F. Jensen, Introduction to Computational Chemistry, John Wiley and Sons, 1999
- [3] R. G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford 1989.
- [4] R. M. Dreizler, E. K. U. Gross, *Density Functional Theory*, Springer-Verlag, Berlin, 1990.
- [5] P. Hohenberg, W. Kohn, *Phys. Rev.* **B 136** (1964) 864.
- [6] W. Kohn, L. J. Sham, *Phys. Rev.* **A 140** (1965) 1133.
- [7] M. C. Zerner, Semiempirical Molecular Orbital Methods, in: Lipkowitz, Boyd (Hrg.), *Reviews in Computational Chemistry*, Vol.2, VCH Publishers, New York 1991.
- [8] M. J. S. Dewar et al., *J. Am. Chem. soc.* **107** (1985) 3902
- [9] J. J. P. Stewart, *J. Comp. Chem.* **10** (1989).
- [10] W. M. Foulkes, R. Haydock, *Phys. Rev.* **B 39** (1989) 12520.
- [11] W. D. Cornell et al., *J. Am. Chem. Soc.*, **117** (1995) 5179.
- [12] A. D. MacKerell et al., *J. Am. Chem. Soc.*, **117** (1995) 11946.
- [13] J. R. Maple et al., *J. Am. Chem. Soc.*, **115** (1994) 162.
- [14] W. L. Jorgensen, J. T. Rives, *J. Am. Chem. Soc.*, **110** (1988) 1657.

- [15] Janos K. Lanyi, Bacteriorhodopsin. *International Journal of Cytology* vol. 187, (1999) 161-202.
Janos K. Lanyi, Molecular mechanism of ion transport in bacteriorhodopsin: insights from crystallographic, spectroscopic, and mutational studies. *J. Phys. Chem. B* vol. 194(48), (2000) 11441-11448.
- [16] M. Elstner, Ph.D. Thesis, University Paderborn: Paderborn, Germany 1998.
- [17] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, *Phys. Rev. B* **58** (1998) 7260.
- [18] M. Elstner, D. Porezag, G. Jungnickel, T. Frauenheim, S. Suhai, and G. Seifert, in *Tight-Binding Approach to Computational Materials Science* (P. E. A. Turchi, A. Gonis, and L. Colombo, eds.), vol. 491, p. 131, MRS, 1998.
- [19] M. Haugk, J. Elsner, Th. Heine, Th. Frauenheim, G. Seifert, *J. Comp. Mat. Science* **13**, 239 (1999).
- [20] D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert, R. Kaschner, *Phys. Rev. B* **51**, 12947 (1995).
- [21] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77** (1996) 3865
- [22] J. Andzelm, E. Wimmer, *J. Chem. Phys.* **96** (1992) 1280 221.
- [23] B. G. Johnson et al., *J. Chem. Phys.* **98** (1993) 5612
- [24] T. Niehaus, M. Elstner, T. Frauenheim, and S. Suhai, *J. Mol. Struct., (THEOCHEM)* 541 (2001) 185.
- [25] M. Elstner, Q. Cui, T. Frauenheim, S. Suhai, Manuskript.
- [26] M. Elstner, C. Qiang, T. Frauenheim, T. Kaxiras, M. Karplus, *J. Chem. Phys.*, zur Publikation eingereicht.
- [27] M. Elstner, T. Frauenheim, E. Kaxiras, G. Seifert, and S. Suhai, *phys. stat. sol.(b)*, **217/1** (2000) 375.
- [28] W. Han, M. Elstner, K. Jalkanen, T. Frauenheim, S. Suhai, *Int. J. Quant. Chem.*, **78** (2000) 459
- [29] Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, M. Karplus, *J. Phys. Chem. B* 105 (2001) 569
- [30] M. Elstner, D. Porezag, G. Seifert, T. Frauenheim, and S. Suhai, In: *Multiscale Modelling of Materials Symp.*, pp. 541–546, MRS, 1999.

- [31] M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, and E. Kaxiras, *J. Chem. Phys.* **114** (2001) 5149
- [32] M. Elstner, K. Jalkanen, M. Knapp-Mohammady, T. Frauenheim, S. Suhai, *Chem. Phys.* **263** (2001) 203
- [33] H. G. Bohr, K. J. Jalkanen, M. Elstner, K. Frimand, S. Suhai, *Chem. Phys.* **246** (1999) 13.
- [34] M. Elstner, K. Jalkanen, M. Knapp-Mohammady, T. Frauenheim, S. Suhai, *Chem. Phys.* **256** (2000) 15
- [35] H. Zhou, E. Taikhorshid, T. Frauenheim, S. Suhai, M. Elstner, *J. Mol. Struct. (THEOMCHEM)*, **277** (2002) 91.
- [36] C. Alemán, R. Roca, F.J. Luque, and M. Orozco, *Proteins* **28** (1997) 83.
- [37] W. Han, K. Jalkanen, M. Elstner S. Suhai, *J. Phys. Chem. B* **102** (1998) 2587
- [38] W. Yang, *Phys. Rev. Lett.* **66**, 1438 (1991).
- [39] W. Yang, T. S. Lee, *J. Chem. Phys.* **163** (1995) 5674.
- [40] W. Yang, J. M. Pérez-Jordá, *Linear Scaling Methods for Electronic Structure Calculations*. In *Encyclopedia of Computational Chemistry*; P. Schleyer, Ed.; John Wiley & Sons: New York, 1998.
- [41] j S. Goedecker *Rev. Mod. Phys.* **71** (1999) 1085.
- [42] G. Galli, *Phys. Status Solidi B-Basic Res.* **217** (2000) 231.
- [43] T. S. Lee, D. York, W. Yang, *J. Chem. Phys.* **105** (1996) 2744
- [44] A. Warshel, M. Levitt, *J. Mol. Bio.* **103** (1976) 227.
- [45] M. J. Field, P. A. Bash, M. Karplus, *J. Comp. Chem.* **11** (1990) 700.
- [46] H. Y. Liu, M. Elstner, E. Kaxiras, T. Frauenheim, J. Hermans, W. Yang, *PROTEINS* **44** (2001) 484.
- [47] T. Niehaus, S. Suhai, F. Della Sala, P. Lugli, M. Elstner, G. Seifert, T. Frauenheim, *Phys. Rev. B* **63** (2001) 085108.
- [48] T. Niehaus, PhD Thesis, Universität Paderborn 2001.
- [49] B. Torralva, T. Niehaus, M. Elstner, S. Suhai, T. Frauenheim, R. E. Allen, *Phys. Rev. B* **64** (2001) 153105

Anhang A

Relevante Veröffentlichungen

- T. Niehaus, M. Elstner, T. Frauenheim, Sándor Suhai, Application of an approximate density functional method to sulfur containing compounds, *J. Mol. Struct. (THEOCHEM)*, **541** (2001) 185.
- M. Elstner, Th. Frauenheim, E. Kaxiras, G. Seifert, S. Suhai, A self-consistent charge density-functional based tight-binding scheme for large biomolecules, *phys. stat. sol. (b)* **217** (2000) 357.
- M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, and E. Kaxiras, Hydrogen bonding and stacking interactions of nucleic acid base pairs: a density-functional-theory based treatment, *J. Chem. Phys.* **114** (2001) 5149.
- M. Elstner, K. J. Jalkanen, M. Knapp-Mohammady, T. Frauenheim and S. Suhai, Energetics and structure of glycine and alanine based model peptides: Approximate SCC-DFTB, AM1 and PM3 methods in comparison with DFT, HF and MP2 calculations, *Chem. Phys.* **263** (2001) 203.
- M. Elstner, K. Jalkanen, M. Knapp-Mohammady, T. Frauenheim, S. Suhai, DFT studies on Helix Formation in N-Acetyl-(L-Alanyl)_n-N'-Methylamide for n = 1-20 *Chem. Phys.* **256** (2000) 15.
- H. Zhou, E. Tajkhorshid, T. Frauenheim, S. Suhai, M. Elstner, Applicability of SCC-DFTB method for studying the conjugated Schiff base molecules, *J. Mol. Struct. (THEOCHEM)*, **277** (2002) 91.
- W. Han, M. Elstner, K. J. Jalkanen, T. Frauenheim and S. Suhai, Hybrid SCC-DFTB/Molecular Mechanical Studies of H-Bonded Systems and of N-Acetyl-(L-Ala)_n-N'-Methylamide Helices in Water Solution *Int. J. Quant. Chem.*, **78** (2000) 459.

- Q. Cui, M. Elstner, T. Frauenheim, E. Kaxiras, M. Karplus, Combined self-consistent charge density functional tight-binding (SCC-DFTB) and CHARMM, *J. Phys. Chem. B* **105** (2001) 569.
- H. Liu, M. Elstner, E. Kaxiras, T. Frauenheim, J. Hermans, W. Yang, Quantum Mechanics Simulation of Protein Dynamics on Long Time Scale, *PROTEINS* **44** (2001) 484.
- T. Niehaus, S. Suhai, F. Della-Sala, P. Lugli, M. Elstner, G. Seifert T Frauenheim, A simple tight-binding approach to Time-Dependent Density-Functional response-Theory, *Phys. Rev.* **B 63** (2001) 5108
- B. Torralva, T. Niehaus, M. Elstner, S. Suhai, T. Frauenheim, and R.E. Allen, Response of C-60 and C-n to ultrashort laser pulses, *Phys. Rev.* **B 64** (2001) 3105.