

Diss. ETH No. 13450

Molecular Dynamics Simulations with a Quantum-Chemical Core: Methodology and Applications in Photochemistry and Bioinorganic Chemistry

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Natural Sciences

presented by
CHRISTIAN D. BERWEGER
Dipl. Chem. ETH

born December 28, 1971
citizen of Herisau, Switzerland

accepted on the recommendation of
Prof. Dr. Wilfred F. van Gunsteren, examiner
Prof. Dr. Ursula Röthlisberger, PD Dr. Florian Müller-Plathe, co-examiners

2000

Laß die Moleküle rasen
was sie auch zusammenknobeln!
Laß das Tüfteln, laß das Hobeln,
heilig halte die Ekstasen!

Christian Morgenstern

Meinen Eltern

Acknowledgments

I wish to thank Wilfred van Gunsteren for giving me the opportunity to do my thesis with him. His profound scientific knowledge and unique personal style rendered my four-years stay in his group a pleasant and stimulating period of my life. His indisputable skills in obtaining computer power are amazing and removed any obstacles in terms of computing bottlenecks. I very much enjoyed the freedom in working hours and scientific topics and bios' undoubted generosity in providing coffee, *borrels* and dinners with guests.

I thank Florian Müller-Plathe, who had the original idea of introducing finite-element interpolation in molecular dynamics simulation. Collaboration with him has been especially fruitful and motivating.

I'm very grateful to Prof. Dr. Walter Thiel, who provided me the source code of the newest version of the MNDO program. Its many features made the development of the zumos program much easier than initially planned.

I thank all members of the group for informatikgestützte Chemie at ETH Zürich. They all have provided a seething environment both scientifically and personally.

I thank Prof. Dr. Kurt Kremer the Max-Planck-Institute für Polymerforschung who generously made it possible for me to stay at his institute for three months. This was a very intense and productive time. I thank all the members of the AK Kremer for their care in rendering my stay so cordial and pleasant.

Much of the work presented here has been enabled by free software. I especially mention Linux and its accompanying software, who established in my home office the same powerful and efficient computer environment as in the lab. Thanks to Linus Torvalds! Thanks also to the mostly unknown people from GNU and the Free Software Foundation. Most of my keystrokes echo in emacs, their ultimate editor, and the fecomd program compiles in their gcc and g++. Their tool gmake greatly facilitated my work and was actually the best way get my projects organised. It and gawk helped many times in my goal to let computers do the work. Many of the graphics in this work are made with xmgr and xfig. Finally, \LaTeX is probably the only text processor to write publications and theses without horror. Thanks to all who contributed to these programs!

I especially thank Salomon Billeter, who was my primary mentor in using the aforementioned programs. Thanks to Walter Scott, Thomas Huber, Philippe Hünenberger, Harald Bopp, Heiko Schäfer, Roland Bürgi, Alexandre Bonvin, Urs Stocker, Tomas Hansson and Fred Hamprecht for keeping the computers running. Thanks to Prisca Cerutti for keeping most of the administrative concerns away from me.

Contents

Kurzfassung	11
Summary	13
Publications	14
1 Introduction	17
1.1 Computers and Chemistry	17
1.2 Problems Addressed	19
1.2.1 Photochemistry: Photoisomerisation of <i>cis</i> -Stilbene	19
1.2.2 Bioinorganic Chemistry: Metallothionein	21
1.3 Theory	22
1.3.1 Quantum Chemistry	22
<i>Ab initio</i> Quantum Chemistry	22
Hartree-Fock Method	23
Configuration Interaction	24
Calculation of Electronically Excited States	25
Semi-Empirical Quantum Chemistry	26
MNDO	26
MNDO/d	26
1.3.2 Classical Molecular Dynamics	27
Newton's Equations of Motion	27
Force Field Basics	27
1.3.3 Combining Quantum Chemistry and Classical Molecular Dynamics	28
Embedding a Quantum-Chemical System into a Classical Environment	28
A Simple Model	28
Polarising the Quantum-Chemical System	29
Split Quantum-Classical Molecules	29
Quantum Topology	30
Saturation of the Quantum-Chemical System	30
Directly Bonded Atoms	32
2 Molecular Dynamics Simulation with an <i>ab initio</i> Potential Energy Function and Finite Element Interpolation: Method and Validation	33
2.1 Abstract	33
2.2 Introduction	33
2.3 Methods	34

2.3.1	An Analog	34
2.3.2	Finite Element Interpolation for Molecular Dynamics Simulations	35
2.3.3	Algorithm Outline	36
2.3.4	The Regular Grid	37
2.3.5	The Interpolation	38
2.3.6	The Quantum/Classical Combination Model	40
2.3.7	The Quantum Chemistry Method	40
2.3.8	Computational Details	40
	Stilbene Geometry	40
	Simulation Parameters	41
2.4	Results and Discussion	42
2.4.1	Preliminary Investigation of the Potential Energy Surface of Photoex- cited Stilbene	42
2.4.2	Simulations	42
2.4.3	Accuracy of the interpolation	42
2.4.4	Efficiency	44
2.5	Conclusions	45
3	The Photoisomerisation of <i>cis</i>-Stilbene Does not Follow the Minimum Energy Path	49
3.1	Summary	49
3.2	Introduction	49
3.3	Potential Energy Surface	50
3.4	Kinetic Activation	50
3.5	Solvent Effect	51
3.6	Molecular Shape Changes	52
3.7	Conclusions	52
4	Viscosity Dependence and Solvent Effects in the Photoisomerisation of <i>cis</i>-Stilbene	53
4.1	Abstract	53
4.2	Introduction	53
4.3	Methods	55
4.3.1	Computational Details	55
4.3.2	Activation Energies	56
4.3.3	Solvent Properties	56
4.3.4	Estimation of Reaction Rate Constants	57
4.4	Results and Discussion	58
4.4.1	Potential Energy Surface	58
4.4.2	Dependence on Temperature and Pressure	62
4.4.3	Viscosity Dependence	64
4.4.4	Average Trajectories	68
4.4.5	Some Individual Dihedral Angle Trajectories	70
4.4.6	Reason for the Barrier-Recrossings	72
4.4.7	Behaviour on the Barrier	74
4.4.8	Barrier Close-ups	74
4.5	Conclusions	76

5	Simulation of the β Domain of Metallothionein	79
5.1	Summary	79
5.2	Introduction	79
5.3	Methods	80
5.3.1	Computational Details	80
5.3.2	Estimation of Van-der-Waals Interaction Parameters for Cadmium	82
5.4	Results and Discussion	83
5.4.1	Comparison of the Cd Zn ₂ X-Ray Crystal Structure with the Cd ₃ NMR Solution Structure	83
5.4.2	The Cd Zn ₂ MDc Simulation Compared to the X-Ray Structure	87
5.4.3	The Cd Zn ₂ MDq Simulation Compared to the X-Ray Structure	89
5.4.4	Comparison of the Cd ₃ MDc Simulation with NMR Data	92
5.4.5	Comparison of the Cd ₃ MDq Simulation with NMR Data	94
5.4.6	Comparison of the Classical MDc and Quantum-Chemical MDq Simulations	95
5.4.7	Comparison of the Simulations of the Cd ₃ , Cd Zn ₂ and Zn ₃ Variants	96
5.5	Conclusions	98
6	Outlook	99
6.1	Photoisomerisation of Stilbene	99
6.1.1	Photoisomerisation of <i>trans</i> -Stilbene	99
6.1.2	Quantum Dynamics with Surface Hopping	99
6.1.3	Interpolation in More Dimensions	99
6.1.4	Another System	100
6.2	Metallothionein	100
6.2.1	Other Metals in the β Domain	100
6.2.2	The α Domain	100
6.2.3	Other Proteins	100
A	The fecomd Implementation	101
A.1	Features	101
A.2	Input File	102
A.3	Output Files	103
A.4	Auxiliary Programs	103
B	The zumos Implementation	105
B.1	The Quantum Topology	105
B.2	Running zumos	106
	Bibliography	109
	Curriculum Vitae	117

Kurzfassung

Der Einsatz von kombiniert quantenchemisch-klassischen Methoden ist populär geworden zur Computersimulation von grossen Systemen, die ein reagierendes Molekül enthalten, oder, allgemein gesagt, deren entscheidender Teil mit klassischen Theorien schwierig zu erfassen ist. Typische Anwendungen sind beispielsweise kleinere reagierende Moleküle in Lösung, wobei die reagierenden Moleküle quantenchemisch, und das Lösemittel klassisch beschrieben wird, oder aber Proteine, deren aktives Zentrum quantenchemisch beschrieben wird und der Rest klassisch. Dabei hat das Lösemittel oder der Rest des Proteins einen entscheidenden Einfluss auf das Zentrum und kann daher nicht einfach weg gelassen werden. Kapitel 1 gibt neben einer allgemeinen Einführung einen Überblick über die verwendeten Methoden.

Die vorliegende Arbeit beschreibt zwei Neuerungen auf diesem Gebiet. Der erste Teil befasst sich mit dem Problem, dass genaue quantenchemische Berechnungen häufig so rechenzeitintensiv sind, dass molekulardynamische Simulationen fast nicht möglich sind, weil für etwa jede Femtosekunde simulierter Zeit eine solche Berechnung nötig ist. Kapitel 2 beschreibt eine Interpolationsmethode, die während der Simulation nach Bedarf die Energiefläche des reagierenden Moleküls aufspannt. Dabei kommt ein relativ grobes Gitter zum Einsatz, dessen Stützpunkte quantenchemisch berechnet werden. Zwischen den Gitterpunkten liegende Punkte werden mit finiten Elementen interpoliert und so die Energien und Gradienten (Kräfte) erhalten. Durch das Gitter und die Tatsache, dass einmal berechnete Stützpunkte immer wieder verwendet werden können, wenn mehrere Trajektorien simuliert werden, auch unter verschiedenen Drücken und Temperaturen, kann der Aufwand an Computerzeit für die quantenchemischen Berechnungen enorm gesenkt werden. So wird auch die molekulardynamische Simulation eines mittelgrossen Moleküls im elektronisch angeregten Zustand machbar.

Kapitel 3 und 4 beschreiben die Anwendung dieser Interpolationsmethode auf die Photoisomerisierung von *cis*-Stilben. Diese Reaktion wird auch experimentell intensiv untersucht. Kapitel 3 erwähnt einige besonders bemerkenswerte Ergebnisse. Beispielsweise erfolgt die Isomerisierung nicht dem Weg minimaler Energie auf der Potentialfläche. Die Annahme des Pfads minimaler Energie wird häufig gemacht und ist Voraussetzung für eine ganze Reihe von Theorien über Reaktionsdynamik. Die Gültigkeit dieser Annahme wird daher in Frage gestellt.

Kapitel 4 behandelt ausführlich die Lösemittelleffekte, die bei der Photoisomerisierung von *cis*-Stilben von Bedeutung sind. Der experimentelle Befund, dass die Reaktionsgeschwindigkeitskonstante kaum von der Temperatur, dafür deutlich vom Druck des Systems abhängt, konnte reproduziert und erklärt werden.

Der zweite Teil der Arbeit beschreibt den Einsatz der semiempirischen Methode MNDO/d innerhalb einer Molekulardynamiksimulation. Die semiempirische Quantenchemie ist rechenzeitgünstig und kann daher ohne spezielle Interpolationsmethoden in jedem Zeitschritt angewendet werden. Die verwendete Methode MNDO/d eignet sich auch für Schwermetalle wie Zink, Cadmium und Quecksilber. Dies ermöglicht die Simulation von Metallothionein, einem Protein, das grosse Mengen an Metallionen enthält, auch ohne erst ein Kraftfeld für die Metalle zu entwickeln. Auch stossen klassische, empirische Kraftfelder rasch an ihre Grenzen, wenn Schwermetallkomplexe zuverlässig beschrieben werden sollen.

In Kapitel 5 zeigt sich, dass das klassische Standard-Kraftfeld von GROMOS die Form des Metallzentrums einigermassen zu bewahren vermag, allerdings mit deutlich zu kurzen Bin-

dingungslängen und teilweise falschen Bindungswinkeln. Der zu kompakte Metallkomplex wirkt sich negativ auf die gesamte Proteinstruktur aus. Während die Bindungslängen durch Anpassung der Kraftfeldparameter korrigiert werden könnten, stellt sich bei den Bindungswinkeln ein fundamentaleres Problem. Diese Probleme konnten durch den Einsatz einer quantenchemischen Beschreibung des Metallzentrums vermieden werden: MNDO/d liefert Strukturen, die gut mit experimentellen Daten übereinstimmen. Insbesondere erfüllt der Cd_3 -Komplex die experimentellen NOE-Schranken gut. Es zeigt sich, dass die metallgebundenen Cysteine sehr stabil sind, während die Peptidschleifen dazwischen ausserordentlich flexibel sind. MNDO ohne Erweiterung auf d-Orbitale hingegen eignet sich nicht: Der Metallkomplex zerfällt bereits nach kurzer Simulation.

Kapitel 6 gibt einen Ausblick in mögliche zukünftige Erweiterungen der vorgestellten Methodologien und nennt weitere mögliche Anwendungsbeispiele. Im Anhang schliesslich werden die beiden zu diesen Studien entwickelten Programmsammlungen kurz vorgestellt und ihre Bedienung erläutert.

Summary

The combination of quantum-chemical and classical methods has become popular in recent years. It is useful for the simulation of large systems with a core that is hard to describe by classical methodology, for example a reacting molecule in solution. In this case, the molecule is described by quantum chemistry, and the solvent by classical force fields. Another typical application are proteins, whose active site is treated quantum-chemically, and all the rest classically. An important aspect is that the rest of the protein and the solvent has an essential influence on the core. Thus they cannot simply be neglected. Chapter 1 gives a general introduction and an overview of these methods.

This work presents two innovations in this field. The first part addresses the problem of the large computational expense of accurate quantum-chemical calculations. Their use in molecular dynamics simulations is nearly impossible, because every time step such a calculation has to be performed. Chapter 2 describes an interpolation method designed to solve this problem. The potential energy surface is constructed “on the fly” when required during the simulation. The surface is represented by a regular grid. The mesh points are calculated by quantum chemistry, and in-between the required energies and gradients (forces) are interpolated using finite elements. By means of the coarse grid, much fewer quantum-chemical calculations are required. The efficiency is greatly improved further when many trajectories are simulated, e. g. under different temperatures and pressures. Now, even molecular dynamics simulations of a medium-sized molecule in its first excited state are feasible.

Chapters 3 and 4 describe the application of the interpolation method to the photoisomerisation of *cis*-stilbene, which is also subject to extended experimental investigations. Chapter 3 presents some remarkable results. For example, the isomerisation does not follow the path of minimum energy on the surface. Such a minimum-energy path is often assumed and it is the basis for several theories about reaction dynamics. Thus the validity of this assumption is questionable.

Chapter 4 presents an in-depth investigation of the solvent effects that occur in the photoisomerisation of *cis*-stilbene. In experiment, the reaction rate constant hardly depends on the temperature, but strongly depends on the pressure of the solvent. The simulation is able to reproduce and explain these findings.

The second part of this work describes the inclusion of the semi-empirical method MNDO/d in a molecular dynamics simulation. Semi-empirical quantum chemistry is computationally cheap and can be applied every time step without any interpolation procedure. MNDO/d is suitable for treating heavy metals such as zinc, cadmium and mercury. This is required for the simulation of metallothionein, a protein capable of binding large amounts of these metals. Doing so, it is not necessary to develop a force field for the metal ions. Moreover, classical force fields often have difficulties in describing metal clusters.

Chapter 5 shows that the classical standard GROMOS force field is able to maintain the overall form of the metal cluster, albeit with bond lengths that are much too short. This problem could be solved by scaling the Lennard-Jones parameters, however, it would not remedy some incorrect bond angles. The too compact structure of the metal core also affects the whole enfolding protein. In contrast, MNDO/d reproduces the experimental structures quite well. In particular, the experimental NOE bounds are well satisfied. The metal-bound cysteines are stable, while the

peptide loops between them are extraordinarily flexible. MNDO without extension to d orbitals is not suitable, as the metal cluster disintegrates quickly in the simulations.

Chapter 6 gives an outlook to possible extensions to the presented methodologies and mentions further potential applications. Finally, the appendices briefly present the two program packages that have been developed for these studies. Some implementation details are given and their usage is explained.

Publications

This thesis is based on the following publications:

Chapter 2:

Christian D. Berweger, Florian Müller-Plathe, and Wilfred F. van Gunsteren,
“Molecular dynamics simulation with an *ab initio* potential energy function and finite element interpolation: The photoisomerisation of *cis*-stilbene in solution”
The Journal of Chemical Physics, **108**, (1998) 8773–8781.

Chapter 3:

Christian D. Berweger, Wilfred F. van Gunsteren, and Florian Müller-Plathe,
“The photoisomerisation of *cis*-stilbene does not follow the minimum energy path”
Angewandte Chemie International Edition in English, **38**, (1999) 2609–2611.

German Translation:

Christian D. Berweger, Wilfred F. van Gunsteren, und Florian Müller-Plathe,
“Die Photoisomerisierung von *cis*-Stilben folgt nicht dem Weg minimaler Energie”
Angewandte Chemie, **111**, (1999) 2771–2773.

Chapter 4:

Christian D. Berweger, Wilfred F. van Gunsteren, and Florian Müller-Plathe,
“Viscosity dependence and solvent effects in the photoisomerisation of *cis*-stilbene: Insight from a molecular dynamics study with an *ab initio* potential energy function”
The Journal of Chemical Physics, **111**, (1999) 8987–8999.

Chapter 5:

Christian D. Berweger and Wilfred F. van Gunsteren.,
“Simulation of the β Domain of Metallothionein”
Proteins: Structure, Function and Genetics, (2000) submitted for publication

a preliminary version of the interpolation method:

Christian D. Berweger, Wilfred F. van Gunsteren, and Florian Müller-Plathe,
“Finite element interpolation for combined classical / quantum-mechanical molecular dynamics simulations”
The Journal of Computational Chemistry, **18**, (1997) 1484–1495.

Chapter 1

Introduction

*“The underlying physical laws for the mathematical theory of ...
the whole of chemistry are completely known.”* P. A. M. Dirac

In a popular sense, chemists are dealing with the production of new substances. Indeed, this is the main field of a synthesist, and research has come far in the development of highly specific synthesis reactions. Though surprising, little is known about these reactions. The fundamental reason of why a reaction occurs along a certain pathway is often unclear. The pathway itself is often unknown, and the explanations are often hard to verify. This is why one needs a synthesist with large knowledge in specific reaction types and good chemical intuition. However, the basic theories which fundamentally describe the whole chemistry have been known for many decades. Indeed, quantum mechanics and quantum dynamics provide all that is needed for describing matter and change. In theory.

In practice, chemical systems are mostly so complex that these theories' equations cannot be solved. There is need for simplification, approximation and assumption, and for enormous computer power.

1.1 Computers and Chemistry

“We can calculate everything.” E. Clementi

In the past decades, computing facilities became much more powerful and much less costly. This led to the evolution of a new branch of theoretical chemistry: Computational chemistry. It can be roughly divided into two branches.

- **Quantum chemistry**, based on quantum-mechanical equations, is able to calculate the electronic structure of a molecule. From that, energies, energy gradients and many other molecular properties can be derived. However, computational expense increases immensely with increasing system size and accuracy to be achieved. Thus, only relatively small systems can be treated. Also, dynamical information is out of reach for all but the smallest systems. This theory is described in more detail in Section 1.3.1.
- **Molecular dynamics simulations** employ an empirical force field to describe the interactions in a system and Newton's classical equation of motion to propagate it in time. Using statistical mechanics, macroscopic properties and dynamic information are available. Computational expense is moderate and limits only the length of simulation elapsed

time, currently in the range of nanoseconds. However, chemical reactions are entirely disabled, and a force field for a specific task is sometimes difficult or impossible to develop. More details concerning molecular dynamics can be found in Section 1.3.2.

So, both these methodologies have their advantages and drawbacks, their strengths and limitations. Still the wish remains to simulate chemical reactions, taking into account the influence of the environment such as a solvent, instead of investigating isolated molecules only. Sometimes, it is also desirable to study a large system without having to develop a force field first.

The idea is over twenty years old [1] to combine both methods, taking advantage of both but eliminating their disadvantages: The interesting, reacting part, is described by quantum chemistry, while its surroundings, being less interesting but still important, are described by a classical force field. Both parts are propagated classically. This partitioning is schematically shown in Figure 1.1. The innermost core, containing a reacting molecule in this case, is entirely treated by quantum chemistry. It is surrounded by a shell which is treated by a force field, but directly influences the core. The outer area completes the system.

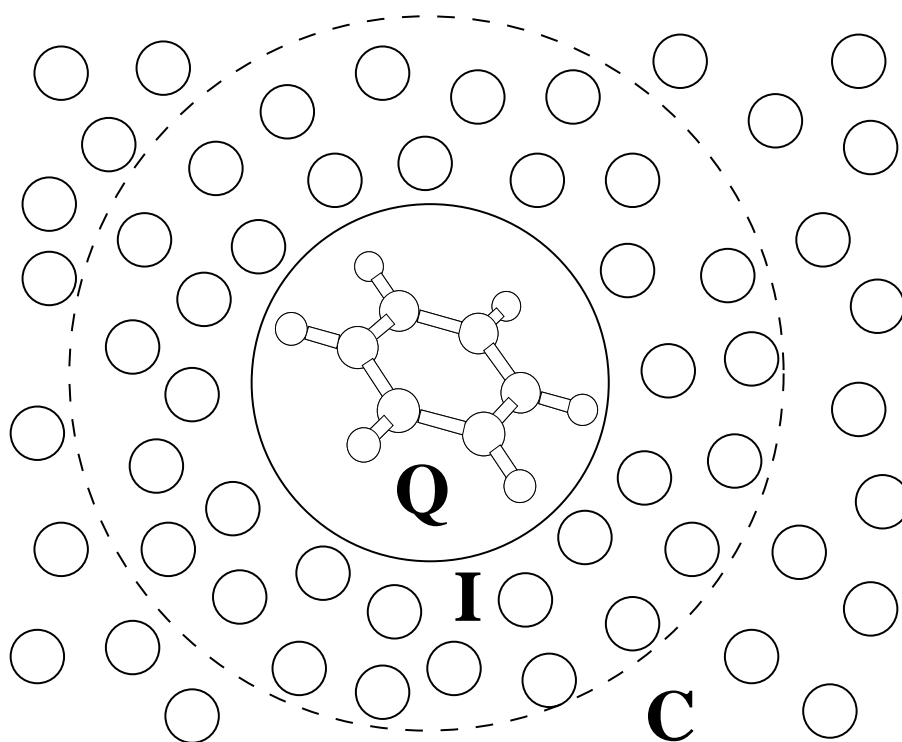


Figure 1.1: Partitioning of the system. **Q** denotes the quantum-chemical core, **I** the shell influencing the inner core, and **C** the classical rest of the system.

The basic concepts originate from Warshel and Levitt [1], and elaborate models have been presented [2–7]. New in the current work is the inclusion of an interpolated intramolecular potential energy surface for increased computational efficiency: Doing so makes it possible to perform molecular dynamics simulations of the photoexcited state of a medium-sized molecule in solution. The surface is based on accurate quantum-chemical calculations, thus avoiding any bias from a force field or experimental results aimed to be reproduced. The great advantage is that the interpolated surface is based on a relatively small number of explicit quantum-chemical calculations (a few thousands), compared to the number of molecular dynamics time steps performed

(several millions). The efficiency of such a method is especially high if the following three conditions are fulfilled. (i) The accessible configuration space of the quantum-chemical part of the molecular system is limited, i. e. there are a few low-energy regions in which the molecular system resides most of the time. (ii) Many trajectories are simulated, using the interpolated surface over and over again. (iii) The computational expense of the quantum-chemical calculation is very large compared to the computational expense of the classical part of the simulation. This is demonstrated by the photoisomerisation of *cis*-stilbene, which uses an *ab initio* configuration interaction treatment for the first electronically excited state. Chapters 3 and 4 present the results. A detailed description of the interpolation method is given in Chapter 2, including tests on accuracy.

There are other methods for interpolating potential energy surfaces [8–10]. However, these methods are not well suited for molecular dynamics simulations. Either the computational expense for many repeated trajectories is still too high [8] or the methods require the potential energy surface to be calculated and refined prior to any dynamics simulations [9, 10]. A method very similar to the one described here employs rectangular finite elements [11].

The second new aspect concerns the inclusion of the semi-empirical method MNDO/d. This quantum-chemical method includes d orbitals on some elements, thus making the method suitable for the calculation of molecular systems containing sulphur, phosphorus and transition metals such as zinc, cadmium or mercury. This methodology is applied to the simulation of the protein metallothionein. It contains clusters of varying transition metals, which are simulated without requiring a force field. This application is described in Chapter 5.

1.2 Problems Addressed

1.2.1 Photochemistry: Photoisomerisation of *cis*-Stilbene

Stilbene (Figure 1.2) has ever been a system of special interest, both for experimentalists and theoreticians [12]. It occurs in two conformations: the *trans* conformation is nearly planar, with only both phenyl rings twisted a little out of planarity, and *cis*-stilbene, which has the phenyl rings towards the same side but tilted against each other by about 45° . There is a large energy barrier between these two conformations so the molecule cannot easily isomerise unless heated or photoexcited. Upon photoexcitation by a laser pulse, the molecule is lifted to its first electronically excited state S_1 . There the potential energy surface differs dramatically from the ground state S_0 , as shown schematically in Figure 1.3. The *cis* conformation is on a high hill, while

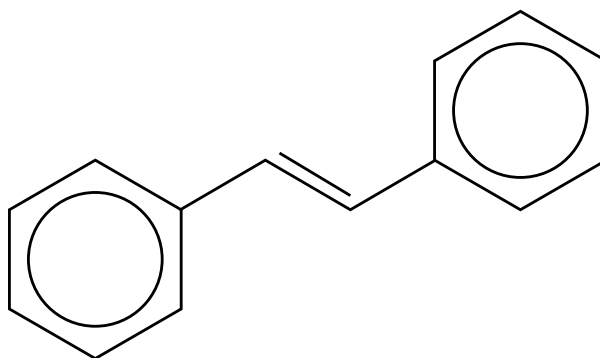


Figure 1.2: Structural formula of *trans*-stilbene.

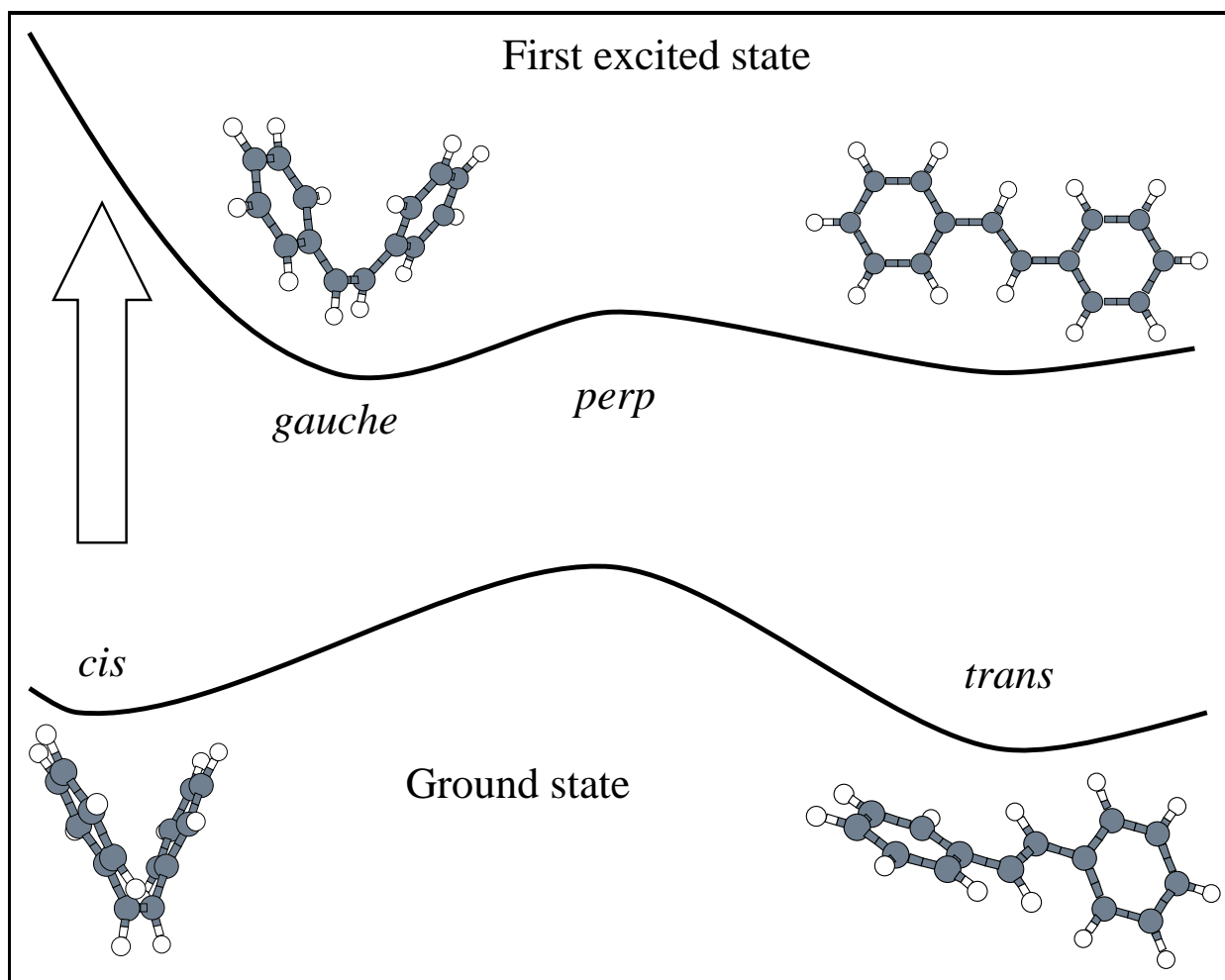


Figure 1.3: Schematic representation of the potential energy surface of the ground state and the first electronically excited state of stilbene.

there is still a minimum in the *trans* region. There is another minimum, slightly deeper, near the *gauche* conformation, and a small energy barrier in the perpendicular conformation.

A similar process enables vision in biological systems. The primary process is the photon-induced *cis-trans* isomerisation of retinal. The conformational change of retinal triggers a series of reactions which eventually lead to the transmission of a nerve impulse.

In experiment, ground-state *cis*-stilbene in solution is photoexcited by an ultrashort laser pulse. Finding itself on a steep flank of an energy hill, the molecule starts moving towards the *gauche* minimum. This photoreaction is extremely fast and occurs in a few picoseconds [13–15]. Using modern ultrafast laser equipment, the ongoing reaction can be tracked in realtime. Eventually the molecule ends up in the ground state again. Depending on the solvent, the temperature and the pressure, this can primarily be *cis*-stilbene, or *trans*-stilbene, if the molecule was able to cross the small barrier in the excited state. There is also an alternative reaction channel which leads to the photocyclisation to dihydrophenanthrene (which is not considered in this work).

The basic experimental results are decay curves [13–16] or time-dependent spectra [17] at best. These results then have to be interpreted. In this phase, assumptions and models cannot be avoided, and often influence the conclusions. Although the system has been heavily investigated under a vast variety of conditions, many details are left up to speculation and are sometimes

still discussed controversially. Here, computer simulation approaches from a different direction. Making assumption and using models completely different to interpretation, the simulation provides a reaction dynamics at atomic detail. Still the computer simulation can be validated by reproducing some experimental results. In the ideal case, experiment and simulation complement each other and, taken together, lead to a deeper understanding of matter and change.

1.2.2 Bioinorganic Chemistry: Metallothionein

The role of metals in biological systems has gained increased interest in recent years [18]. Iron as part of hemoglobin in blood cells is well known, and sodium and potassium are known for their importance to maintain the osmotic balance. However, the functions of metals are much more versatile. Calcium, for example, serves as a nerve pulse messenger. Cobalt and zinc are used as a catalytic centre in enzymes. Structures are maintained by magnesium, calcium, manganese and zinc. Electron transport is the field of copper and iron. Special tasks such as nitrogen fixation are performed by special metals, such as vanadium and molybdenum.

One of the most important metal ions in biological systems is *zinc*. Its structural function is for example important in zinc fingers, and it enables special catalytic reactions in enzymes such as alcohol dehydrogenase, which is important in the degradation of ethanol. However, *cadmium* has similar binding properties as zinc, but without the same structural functions or catalytic capabilities. This is a reason why cadmium is toxic: it competes with zinc in binding to biomolecules, but does not provide the special properties of zinc.

A remarkable metalloprotein is *metallothionein*. It is a class of abundant small proteins of about sixty residues. It contains twenty cysteines and is able to bind up to seven metal ions, sometimes even twelve of them. Binding to copper, zinc and cadmium seems to be biologically

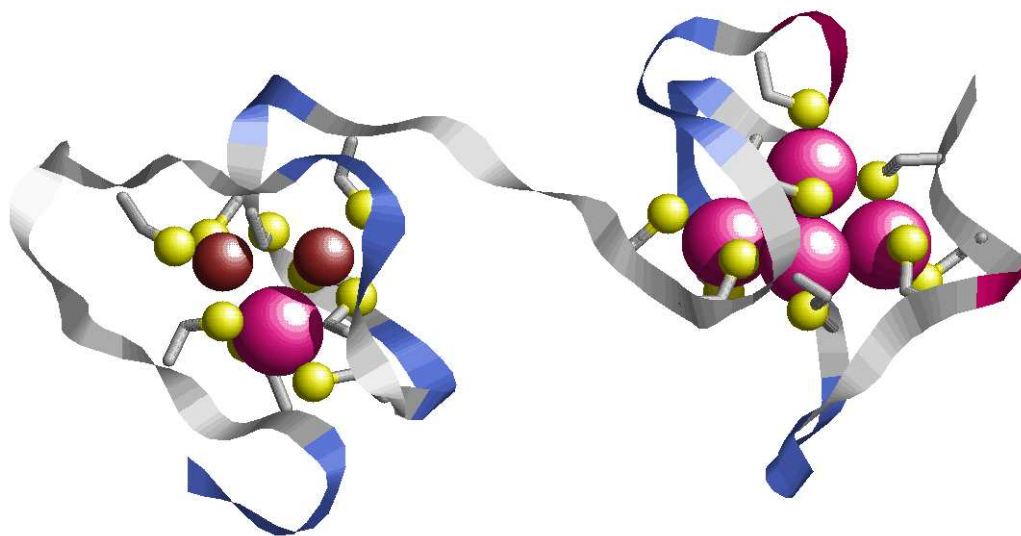


Figure 1.4: Crystal structure of rat liver metallothionein. The large balls are the cadmium ions, the dark medium-sized balls the zinc ions, and the small light balls the cysteine's sulphurs. Cysteine side chains are represented as thick sticks. The ribbon displays the backbone, which has no regular secondary structure. The β domain is to the left and contains three metal ions, the α domain with four metal ions is to the right.

most important, but it is also found to bind to mercury and platinum. Metal binding occurs in two independent, separate domains. The β domain binds three metal ions to nine deprotonated cysteinic sulphurs, the α domain binds four metal ions to eleven cysteines. Figure 1.4 shows the crystal structure of rat liver metallothionein.

The function of metallothionein is unclear. The most likely explanations are supplying metal ions, primarily copper and zinc, for the biosynthesis of other metalloproteins, and removal of toxic metals such as cadmium and mercury. An interplay with zinc fingers and thus gene expression is also possible.

The three-metal β domain of rat liver metallothionein is simulated by means of a purely classical force field and a combined quantum-chemical/force-field approach. Different combinations of zinc and cadmium (Zn_3 , CdZn_2 and Cd_3) contents are investigated. For the purely classical simulations, the standard GROMOS96 [19] force-field parameters are used, and parameters are estimated for cadmium. The results are compared to simulations with the metal clusters described by semi-empirical quantum chemistry (MNDO and MNDO/d), and to experimental data.

1.3 Theory

1.3.1 Quantum Chemistry

Ab initio Quantum Chemistry

The state of a chemical system is quantum-mechanically defined by a wave function Ψ which satisfies the (time-independent) Schrödinger Equation

$$\mathcal{H}\Psi = E\Psi \quad (1.1)$$

where the Hamilton operator \mathcal{H} describes the system in a formal fashion, and E is the energy. However, this equation cannot be solved but for the most simple systems such as the hydrogen atom. Therefore, a commonly made simplification is the Born-Oppenheimer approximation, which separates the motions of the nuclei and the electrons. This is reasonable because the mass of an electron is more than three orders of magnitude smaller than the mass of a nucleus. Thus the electrons are assumed to move in the field of the fixed nuclei, the kinetic energy of the nuclei is neglected, and the nuclear repulsion is constant. The Born-Oppenheimer Hamilton operator \mathcal{H}^{BO} for a molecular system with N electrons and M nuclei reads

$$\mathcal{H}^{\text{BO}} = -\frac{1}{2} \sum_i^N \nabla_i^2 - \sum_i^N \sum_A^M \frac{Z_A}{r_{iA}} + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \quad (1.2)$$

with the distance between particles r or R and the nuclear charges Z_A . Formally, the Born-Oppenheimer approximation makes the electronic wave function depend parametrically on the nuclear positions, while it explicitly depends on the electronic coordinates. The electronic energy only depends on the nuclear positions. This creates the Born-Oppenheimer surface, which is essential for moving the nuclei by classical dynamics. The Born-Oppenheimer approximation also includes that the electrons instantaneously adapt to motions of the nuclei, as happens during the molecular dynamics simulations.

The next step in simplification concerns the interaction between electrons. Instead of explicitly taking into account all instantaneous electronic positions, it is assumed that every single

electron moves in the average field of all other electrons. In this way, the electrons produce a self-consistent field (SCF). It is described by the Hartree-Fock (HF) equation [20]

$$f_i \chi_i = \varepsilon_i \chi_i, \quad i = 1..N \quad (1.3)$$

where i counts through all N electrons and χ_i is an orbital (a one-electron wave function) with energy ε_i . The Fock operator

$$f_i = -\frac{1}{2} \nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}} + v_i^{\text{SCF}} \quad (1.4)$$

where the first term on the right hand side represents the kinetic energy of the electron, and the second term represents the potential energy of the electron in the field of the nuclei (M nuclei with nuclear charges Z_A). The sum of these two terms is referred to as core-Hamiltonian. The third term v_i^{SCF} represents the potential energy in the self-consistent field of the other electrons. This term v_i^{SCF} itself is dependent on the orbitals χ_i , thus the Fock operator f_i depends on its solutions χ_i . The basic idea of the Hartree-Fock method is to make an initial guess on the orbitals χ_i , from that calculate the field v_i^{SCF} and solve the Hartree-Fock equation for better orbitals. This procedure is iterated until self-consistency is obtained.

Hartree-Fock Method To make the Hartree-Fock equation solvable on a computer, a finite basis is introduced to represent the orbitals. For closed-shell systems, this yields the Roothaan equation, a matrix equation

$$\mathbf{FC} = \mathbf{SC}\varepsilon \quad (1.5)$$

where \mathbf{F} is the Fock matrix, \mathbf{C} contains the expansion coefficients which build the orbitals from the basis functions, \mathbf{S} is the overlap matrix, and ε is a diagonal matrix containing the orbital energies. The Roothaan equation can computationally be solved by matrix manipulations.

Often the density matrix \mathbf{P} is calculated from the expansion coefficients

$$\mathbf{P}_{\mu\nu} = 2 \sum_i^{N/2} \mathbf{C}_{\mu i} \mathbf{C}_{\nu i}^* \quad (1.6)$$

In conjunction with the basis set, the density matrix completely specifies the electron density in the molecule. The overlap matrix \mathbf{S} has the following elements:

$$\mathbf{S}_{\mu\nu} = \int \varphi_\mu^*(\vec{r}) \varphi_\nu(\vec{r}) d\vec{r} \quad (1.7)$$

where φ are the basis functions. As the basis functions are normally not orthogonal, the overlap matrix is not the identity matrix. The Fock matrix \mathbf{F} is the matrix representation of the Fock operator and has the following elements:

$$\mathbf{F}_{\mu\nu} = \mathbf{H}_{\mu\nu}^{\text{core}} + \sum_\lambda \sum_\sigma \mathbf{P}_{\lambda\sigma} \left[(\mu\nu|\sigma\lambda) - \frac{1}{2} (\mu\lambda|\sigma\nu) \right] \quad (1.8)$$

with the core-Hamiltonian matrix

$$\mathbf{H}_{\mu\nu}^{\text{core}} = \int \varphi_\mu^*(\vec{r}) \left(-\frac{1}{2} \nabla_{\vec{r}}^2 \right) \varphi_\nu(\vec{r}) d\vec{r} + \int \varphi_\mu^*(\vec{r}) \left(-\sum_A^M \frac{Z_A}{|\vec{r} - \vec{R}_A|} \right) \varphi_\nu(\vec{r}) d\vec{r} \quad (1.9)$$

where the first term are the kinetic energy integrals, and the second term represent the nuclear attraction integrals. The core-Hamiltonian is also called the one-electron part of the Fock matrix. The sum in Equation 1.8 is the expression for the self-consistent field and is called the two-electron part. The expressions $(\mu\nu|\lambda\sigma)$ are short-hand notations for the two-electron integrals

$$(\mu\nu|\lambda\sigma) = \int \int \Phi_{\mu}^*(\vec{r}_1) \Phi_{\nu}(\vec{r}_1) \frac{1}{|\vec{r}_1 - \vec{r}_2|} \Phi_{\lambda}^*(\vec{r}_2) \Phi_{\sigma}(\vec{r}_2) d\vec{r}_1 d\vec{r}_2 \quad (1.10)$$

where the symbols μ , ν , λ and σ denote basis functions. Computation of these two-electron integrals makes up the major part, because of their large number. Because the two-electron part depends on the density matrix, it has to be recalculated in every iteration step. In contrast, the core-Hamiltonian is constant and has to be calculated once at the beginning.

The most important quantity for performing molecular dynamics on a quantum-chemical potential energy surface, is the molecular energy E^{HF} and its gradients, which correspond to the forces needed for propagating the nuclei in a classical way. The Hartree-Fock energy is given by

$$E^{\text{HF}} = \frac{1}{2} \sum_{\mu} \sum_{\nu} \mathbf{P}_{\mu\nu} (\mathbf{H}_{\mu\nu}^{\text{core}} + \mathbf{F}_{\mu\nu}) + \sum_A \sum_{B>A} \frac{Z_A Z_B}{r_{AB}} \quad (1.11)$$

where the first term is the electronic energy and the second term represents the nuclear repulsion energy. The gradient of the Hartree-Fock energy with respect to a nuclear coordinate X reads [21]

$$\frac{\partial E^{\text{HF}}}{\partial X} = \sum_{\mu} \sum_{\nu} \mathbf{P}_{\mu\nu} \frac{\partial \mathbf{H}_{\mu\nu}}{\partial X} + \frac{1}{2} \sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} \mathbf{P}_{\mu\nu} \mathbf{P}_{\lambda\sigma} \frac{\partial (\mu\lambda||\nu\sigma)}{\partial X} - \sum_{\mu} \sum_{\nu} \mathbf{W}_{\mu\nu} \frac{\partial \mathbf{S}_{\mu\nu}}{\partial X} + \frac{\partial V_{\text{nuclear}}}{\partial X} \quad (1.12)$$

where $(\mu\lambda||\nu\sigma)$ is an abbreviated notation for the antisymmetrised two-electron integrals

$$(\mu\lambda||\nu\sigma) = (\mu\lambda|\nu\sigma) - (\mu\lambda|\sigma\nu) \quad (1.13)$$

and \mathbf{W} is an energy-weighted density matrix

$$\mathbf{W}_{\mu\nu} = \sum_i \varepsilon_i \mathbf{C}_{\mu i}^* \mathbf{C}_{\nu i} \quad (1.14)$$

The Hartree-Fock method gives the lowest-energy single-determinantal result for the electronic ground state in the given basis set. However, the assumption of an average field of the electron neglects correlation between the electrons. So the result may be inaccurate. The Hartree-Fock method is a good starting point for either refined methods which yield more accurate results at the cost of increased computational effort, or for more simplification in order to reduce the computational expense.

Configuration Interaction The basic concept of the configuration interaction (CI) method is similar to that of the Hartree-Fock method. The major difference is that instead of an orbital basis, an N -electron basis is employed. It takes advantage of the fact that the exact wave function Ψ can be expanded in a basis of all possible N -electron Slater determinants ψ , which in turn are formed from a set of orbitals χ .

The principal procedure is as follows. A Hartree-Fock calculation is performed. From the resulting determinant, the N -electron basis is constructed by “exciting” the determinant in all possible ways. Here, exciting means promoting electrons from occupied to virtual orbitals. If

only one electron is moved, then singly excited determinants are obtained, if two electrons are promoted, then doubly excited determinants result, and so on. The Hamilton matrix is set up in the basis of these determinants. The matrix is diagonalised to obtain the eigenvalues (energies) and eigenfunctions (wave functions). Thus, by solving the CI equation, we obtain as a result not only the ground state, but all electronic states together with their energies!

However, this procedure is not feasible in practice because the complete basis is enormously large. For example, in the case of stilbene which is calculated later, there are 96 electrons to be distributed among 150 basis functions (with the relatively small basis set 6-31G). The number of all possible distributions is given by the binomial $\binom{150}{96}$, which results in a CI matrix larger than $10^{41} \times 10^{41}$! One is clearly forced to restrict oneself to a small selection of excitations, based on the goal of the calculations. For example, for the calculation of electronically excited states, the singly excited determinants are most important, so only those enter the basis (configuration interaction with single excitations, CIS). To reduce the basis still further, the lowest-lying electrons are not excited (frozen core), and the highest-lying virtual orbitals are not occupied. In the case of stilbene as it was actually calculated, the active window ranged from orbitals 27 through 80, which includes 22 occupied and 32 virtual orbitals. This gives an affordable basis set size of 704 configurations.

Calculation of Electronically Excited States The calculation of electronically excited states is described in detail by Foresman *et al.* [22]. Here it is sufficient to recall the most important results. It is stated that configuration interaction with single excitations (CIS) is an adequate approximation for the calculation of excited states, at an affordable computational effort. The quality of a CIS calculation of the first excited state is comparable to that of a ground state at the HF level. The wave function Ψ^{CIS} is expanded into singly excited Slater determinants ψ_{ia} with the expansion coefficients d_{ia}

$$\Psi^{\text{CIS}} = \sum_i \sum_a d_{ia} \psi_{ia} \quad (1.15)$$

where the index i runs over the occupied molecular orbitals, and a runs over the virtual molecular orbitals. The molecular energy is given by

$$E^{\text{CIS}} = E^{\text{HF}} + \sum_i \sum_a d_{ia}^2 (\epsilon_a - \epsilon_i) - \sum_i \sum_j \sum_a \sum_b d_{ia} d_{jb} (ja||ib) \quad (1.16)$$

where the ϵ denote the energies of the molecular orbitals, and indices i and j run over occupied molecular orbitals, and indices a and b stand for virtual molecular orbitals. The forces on the nuclei are determined by the derivative of the energy,

$$\frac{\partial E^{\text{CIS}}}{\partial X} = \sum_{\mu} \sum_{\nu} \sum_{\lambda} \sum_{\sigma} \Gamma_{\mu\nu\lambda\sigma}^{\text{CIS}} \frac{\partial(\mu\nu|\lambda\sigma)}{\partial X} + \sum_{\mu} \sum_{\nu} \mathbf{P}_{\mu\nu}^{\text{CIS}} \frac{\partial \mathbf{H}_{\mu\nu}}{\partial X} + \sum_{\mu} \sum_{\nu} \mathbf{W}_{\mu\nu}^{\text{CIS}} \frac{\partial \mathbf{S}_{\mu\nu}}{\partial X} + \frac{\partial V_{\text{nuclear}}}{\partial X} \quad (1.17)$$

where Γ^{CIS} is the two-particle CIS density matrix, \mathbf{P}^{CIS} is the CIS density matrix, \mathbf{H} is the one-electron core-Hamiltonian matrix, \mathbf{W}^{CIS} is an energy-weighted density matrix, \mathbf{S} is the overlap matrix, and the last term is the derivative of the nuclear repulsion energy.

Semi-Empirical Quantum Chemistry

*“Die numerische Quantenchemie ist ein Königreich hässlicher
Abkürzungen.”*

Hans Primas

The semi-empirical methods are an attempt to reduce the computational effort of the Hartree-Fock method. The following steps are usually taken:

- Reduction of the basis set. Only the valence shells are explicitly treated, and the inner shells are described by a so-called semi-empirical atom core. For the valence electrons, a minimal basis set is used.
- Neglect of differential overlap, meaning the basis functions do not overlap under certain circumstances: Many of the cumbersome two-electron integrals are neglected. The exact definition depends on the semi-empirical model and will be given later.
- Replacement of remaining integrals by simple parameterised functions. A sensible parameterisation should compensate for the simplifications made before.

MNDO The MNDO method is one of the most successful and most used semi-empirical methods. The abbreviation MNDO stands for “modified neglect of diatomic overlap” and belongs to the NDDO family of semi-empirical methods (“neglect of diatomic differential overlap”). This family neglects the overlap of basis functions χ if they belong to different atoms. Formally,

$$\chi_i^A \chi_j^B = \chi_i^A \chi_j^B \delta_{AB} \quad (1.18)$$

with the Kronecker delta δ_{AB} . The Roothaan-Hall equation takes the form [23]

$$\sum_{\nu} (\mathbf{F}_{\mu\nu} - E_i \delta_{\mu\nu}) \mathbf{C}_{\nu i} = 0 \quad (1.19)$$

The expression for the electronic energy looks the same as for the Hartree-Fock method (Equation 1.11), but the elements for the core-Hamiltonian and the Fock matrix are different, as the integrals are replaced by parameterised functions. The nuclear repulsion is replaced by the repulsion between semi-empirical cores

$$V_{AB}^{\text{core}} = Z_A Z_B (s_A s_A | s_B s_B) [1 + \exp(-\alpha_A R_{AB}) + \exp(-\alpha_B R_{AB})] \quad (1.20)$$

where the term $(s_A s_A | s_B s_B)$ shows that the core-core interaction is modeled as interaction between s orbitals, and the α are examples of semi-empirical parameters. There are up to seven adjustable parameters per element, which are optimized using *ab initio* and experimental data.

MNDO/d For heavier elements, the MNDO standard basis of s and p orbitals, is not sufficient. Therefore, d orbitals are included [24] for third-row elements (sodium and heavier elements). This leads to a significant improvement in the description of molecules containing these elements. It was found that some metals still perform well with an sp basis [25], but need reparameterisation for the balance with elements with an spd basis. Thus for example sodium, magnesium, zinc, cadmium and mercury are retained with an sp basis. For elements with an spd basis, such as aluminum, silicon, phosphorus, sulphur and the halogens, the core-core repulsion reads

$$E_{AB}^{\text{core}} = Z_A Z_B \frac{1}{\sqrt{R_{AB}^2 + (\rho_A + \rho_B)^2}} [1 + \exp(-\alpha_A R_{AB}) + \exp(-\alpha_B R_{AB})] \quad (1.21)$$

with the additional adjustable parameters ρ . In total, there are 14 adjustable parameters for an element with an spd basis to be optimised.

1.3.2 Classical Molecular Dynamics

Newton's Equations of Motion

The aim of molecular dynamics (MD) is to simulate the evolution in time of a molecular system. In order to do so, Newton's equation of motion is integrated,

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{x}_i} = m_i \frac{d^2 \vec{x}_i}{dt^2} \quad (1.22)$$

where F_i is the force acting on atom i , which has position x_i and mass m_i , and V is the potential energy of the system. The integration is accomplished by discretisation of the time into time steps, which are usually in the range of a femtosecond. There exist several algorithms for integrating Newton's equation of motion [26], of which the leap-frog algorithm is widely used. The general scheme works as follows.

1. Calculate the potential energy of the system's configuration and its gradient, corresponding to the forces acting on the atoms.
2. Using the forces, accelerate the atoms to obtain their velocities at the next half-time step.
3. Using these velocities, displace the atoms to obtain their positions at the next time step.
4. Proceed with step 1.

The potential energy is usually determined by a force field, but in principle any differentiable function of the system's configuration can serve as a potential energy function.

Force Field Basics

A force field is used to describe the interactions in a chemical system in terms of classical, empirical, parameterised interactions. These interactions can be grouped into three categories.

- **Coulomb interaction** between the partial charges q_A and q_B on the atoms A and B at the distance R_{AB}

$$E^{\text{Coulomb}} = \frac{1}{4\pi\epsilon_0\epsilon_1} \frac{q_A q_B}{R_{AB}} \quad (1.23)$$

This interaction takes account of interactions between charged species, dipoles and higher multipoles. The corresponding force-field parameters are the partial charges q_A .

- **Lennard-Jones interaction** (also called van-der-Waals interaction), an empirical potential energy function between the atoms A and B

$$E_{AB}^{\text{LJ}} = \frac{C_{AB}^{\text{twelve}}}{R_{AB}^{12}} - \frac{C_{AB}^{\text{six}}}{R_{AB}^6} \quad (1.24)$$

$$\text{with } C_{AB}^{\text{twelve}} = \sqrt{C_A^{\text{twelve}}} \sqrt{C_B^{\text{twelve}}}, \quad C_{AB}^{\text{six}} = \sqrt{C_A^{\text{six}}} \sqrt{C_B^{\text{six}}} \quad (1.25)$$

where, sloppily speaking, the pairwise parameter C_{AB}^{twelve} specifies the repulsion at small interatomic distances R_{AB} due to electron shell overlap, and C_{AB}^{six} determines the attraction at intermediate distances due to instantaneous dipole induction in the electron shell. In the force field, the parameters are generally specified in an atom-wise fashion by the parameters $\sqrt{C_A^{\text{twelve}}}$ and $\sqrt{C_A^{\text{six}}}$.

The two aforementioned interactions are collectively referred to as non-bonded interactions. For simplicity and reduction of computational effort, these interactions are generally neglected for atom pairs that are more than a certain cut-off distance apart. For computational efficiency, a pair-list is established containing pairs within the cutoff distance, for which the non-bonded forces have to be calculated.

- **Bonded interactions**, namely for chemical bonds, bond angles, improper dihedrals (to handle planar or chiral atoms) and torsional angles. These interactions are empirical functions which usually contain an ideal value plus a force constant which specifies how strongly the ideal value is enforced. The set of all bonded interactions defines the connectivity of a molecule.

The set of all atomic parameters q_A , $\sqrt{C_A^{\text{twelve}}}$ and $\sqrt{C_A^{\text{six}}}$ together with all bonded interactions defines how a molecular system, and its environment, interact with themselves and each other. It is referred to as molecular topology.

The development of a force field is rather tedious. The parameters are sometimes guessed based on chemical intuition, derived from extensive quantum-chemical calculations, and carefully fine-tuned and optimized to yield bulk condensed-phase properties known from experiment.

A force field naturally has its limitations. For example, its formal framework prevents it from describing phenomena which involve chemical bond formation and cleavage, or large and varying polarisation. It is difficult to describe complex bonding situations as occur in transition-metal complexes, in which variable oxidation states and coordination numbers can occur. Sometimes, appropriate experimental data to parameterise against is lacking for a system of interest. This happens for instance for short-lived species such as excited states.

1.3.3 Combining Quantum Chemistry and Classical Molecular Dynamics

Embedding a Quantum-Chemical System into a Classical Environment

The situation of a combined system based on quantum chemistry and a force field, as depicted in Figure 1.1, can be described in a formal and superficial way by

$$H = H^{\text{FF}} + \mathcal{H}^{\text{QC}} + \overline{\mathcal{H}}^{\text{QC/FF}} \quad (1.26)$$

with the plain Hamiltonian H^{FF} denoting that this part is treated classically, the calligraphic Hamiltonian \mathcal{H}^{QC} indicates a part treated by quantum-chemistry, and the overstrike character of the coupling Hamiltonian $\overline{\mathcal{H}}^{\text{QC/FF}}$ demonstrates that it is not yet clear which method is used for the coupling. This will be made clear in the following paragraphs.

A Simple Model Let us consider a quantum-chemical molecule, the solute, in a classical non-polar solvent. The Born-Oppenheimer potential energy surface of the solute is described by quantum chemistry as if it were in the gas phase, and the electronic wave function is not perturbed

by the solvent. The solvent mainly provides friction to molecular motions. The explicit solvent allows for dynamic solvent effects, which will prove important in a later application (Chapter 4).

All intra-solute interactions are covered by quantum chemistry. The solvent-solvent interactions are treated by the force field, as are the solvent-solute interactions. For this reason, Lennard-Jones spheres are assigned to the quantum atoms in order to inhibit overlap of the classical and the quantum-chemical atoms. For simplicity, the Lennard-Jones parameters are directly taken from the corresponding atoms of the force field. This procedure may be inappropriate if the character of a quantum-chemical atom changes drastically during the simulation, for example when bonds are formed or cleaved; or for electronically excited states of small molecules if the excited electron occupies a very extended orbital. However, such effects are considered to be of minor importance in the applications reported later.

The model can formally be described as

$$H = H_{\text{solvent-solvent}}^{\text{LJ}} + H_{\text{solute-solvent}}^{\text{LJ}} + \mathcal{H}_{\text{intra-solute}}^{\text{QC}} \quad (1.27)$$

The intra-solute forces acting on the quantum-chemical atoms are obtained by Equation 1.12 and added to the forces originating from the force field. This model has been used for simulating the photoisomerisation of stilbene in non-polar solution (Chapters 2, 3 and 4).

Polarising the Quantum-Chemical System If the solvent is polar, then the assumption of the solute's wave function being unperturbed is no longer valid. The solvent's partial charges influence the solute's electron density. This effect can be included into the quantum-chemical calculation by means of so-called background charges. The concept is simple. Recall the core-Hamiltonian (Equation 1.9) described in Section 1.3.1. The second term contains the nuclear charges, which determine the electric field in which the electrons move. Here, the background charges can be inserted in the same way as the nuclear charges. The background charges differ from the nuclear charges insofar that they are mostly fractional, and that they do not possess basis functions. So the electrons will still gather around the real nuclei, but the wave function is polarised by the background charges. The forces acting on the nuclei as well as on the background charges are added to the classical forces. There are no classical Coulomb interactions between the classical and the quantum-chemical part.

Thus the Hamiltonian of this model reads

$$H = H_{\text{solvent-solvent}}^{\text{FF}} + H_{\text{solute-solvent}}^{\text{LJ}} + \mathcal{H}_{\text{intra-solute}}^{\text{QC}} + \mathcal{H}_{\text{solute-solvent}}^{\text{QC-Coulomb}} \quad (1.28)$$

Similar to the non-bonded forces in the force field, only partial charges within the cutoff distance to any quantum atoms are included in the quantum-chemical calculation. These atoms are called the neighbour atoms. In a molecular dynamics simulation, the interface atoms change with time. So there has to be a mechanism which dynamically builds up a list of neighbour atoms. This is relatively easily implemented by scanning through the non-bonded pair-list. This model has been used for the simulation of metallothionein (Chapter 5).

Split Quantum-Classical Molecules

If a large solute such as a protein is to be studied, then it is not possible to include the whole solute in the quantum-chemical calculations, due to the computational expense. It is necessary to split the solute into a classical and a quantum-chemical part. Abandoning the clear separation between the quantum-chemical solute and the classical solvent has two major implications.

Firstly, because the boundary between quantum and classical part goes through a chemical bond, the quantum part is not a complete molecule any more, it has some dangling bonds. Therefore, a commonly used approach is to attach additional atoms, mostly hydrogen, to saturate the dangling bonds. These auxiliary atoms are called link atoms. Secondly, atoms which are bonded across the boundary lie so close to each other that they have to be excluded from the normal cross-boundary interaction, similar to the exclusion of first and second bond neighbours from the non-bonded interactions in a force-field [19].

Quantum Topology For a purely classical simulation, all classical interactions that occur in the system are listed in the classical molecular topology. If, however, part of the system is treated quantum-chemically, the interactions covered by quantum-chemistry have to be removed, or the corresponding interaction has to be explicitly excluded. For example, no classical Lennard-Jones interactions should occur between any two quantum atoms. So all quantum atoms are mutually excluded by putting them into the exclusions list. Moreover, there should not be any classical electrostatic interaction between a quantum atom and any other atom. So all quantum atoms have their partial charge set to zero. Also, when considering which bonded force-field terms should be treated, it is easiest to think of the system fully classical first, from which part of the classical interactions are removed and replaced by the quantum-chemical description. The resulting topology is then called a quantum topology.

Saturation of the Quantum-Chemical System For ease of description, let us introduce the following nomenclature (Figure 1.5): The quantum-chemical atom which is bonded to a classical atom is called join atom **J**, its classical bond partner is called connect atom **C**, and the link atom between the two is designated **L**. The neighbouring atoms on the quantum side are labeled with **Q₁**, **Q₂** and so forth, the bonded atom on the classical side are labeled **N₁**, **N₂** and so on.

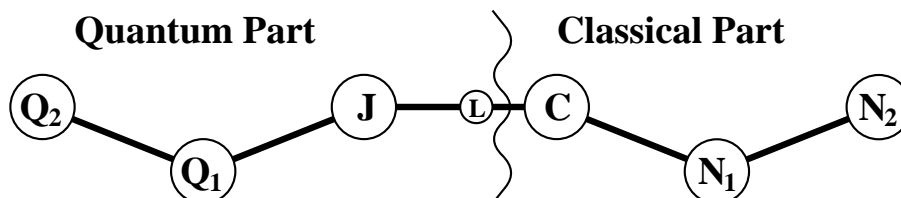


Figure 1.5: Nomenclature used in the description of the link atom approaches.

Link Atoms In the widely used concept of link atoms [2], auxiliary hydrogen atoms are added to the quantum-chemical part to saturate bonds across the quantum-classical boundary. The link atoms do not have any other interactions, meaning they have no Lennard-Jones sphere so they will not interact with the classical part. The link atom is therefore floating around freely. The bond situation is achieved by classical force field terms that span over the boundary. Specifically, there is a force-field bond between the connect atom and the join atom. In this approach, generally all bonded interactions involving exclusively quantum atoms have to be removed from the classical topology. However, all bonded interactions, bond, bond angles and torsion angles, that involve both the join atom and the connect atom, are treated by classical force-field terms. A list of force-field interactions required in this traditional link-atom approach is given in Table 1.1.

The fundamental problem about this link atom approach is that spurious atoms are introduced into the simulation, thus introducing unphysical degrees of freedom.

Link Atom Approach	Classical Force-Field Terms		
	Bonds	Bond Angles	Torsion angles
Traditional	J-C, C-N₁, N₁-N₂	Q₁-J-C, J-C-N₁, C-N₁-N₂	Q₂-Q₁-J-C, Q₁-J-C-N₁, J-C-N₁-N₂
Bond-Constrained	C-N₁, N₁-N₂	J-C-N₁, C-N₁-N₂	Q₁-J-C-N₁, J-C-N₁-N₂

Table 1.1: List of force-field terms near the quantum-classical boundary that have to be included in either link atom approach.

Bond-Constrained Link Atoms The above-mentioned problem can be avoided if bond-constrained link atoms are employed. In this approach, the link atom is placed between the join atom and the connect atom in every time step. The exact location of the link atom \vec{x}_L is determined by a constant ratio s of atom distances

$$\vec{x}_L = \vec{x}_J + s(\vec{x}_C - \vec{x}_J) \quad (1.29)$$

where \vec{x}_J and \vec{x}_C are the positions of the join atom and the connect atom respectively. The ratio s is chosen such as to reflect the ratio between the J-L and J-C standard bonds lengths. For example, if the quantum-classical boundary crosses a carbon-carbon bond (0.154 nm), which is replaced by a carbon-hydrogen bond (0.107 nm) [27] in the quantum-chemical calculation, then the ratio s is 0.695. A very similar approach is known as the scaled position link atom method (SPLAM) [28].

So the link atom does not move freely. In fact, it does not even exist as atom which is propagated in time. Instead, the forces acting on it are distributed onto the join atom and the connect atom, in such a way that the total force and the total torque is conserved.

$$\vec{F}'_J = \vec{F}_J + (1-s)\vec{F}_L \quad (1.30)$$

$$\vec{F}'_C = \vec{F}_C + s\vec{F}_L \quad (1.31)$$

This procedure gives the bond across the boundary quantum character, as the bond characteristics from the join atom to the link atom is transferred to the bond with the connect atom. Consequently, no classical bond force-field term is used to describe the bond across the boundary. However, this bond may not be very accurately described. This does not matter much, since the focus of interest is usually at the center of the quantum-chemical core, not at its boundary. Fewer force-field terms across the boundary are needed, thus making fuller use of the quantum-chemical calculations. In general, all bonded force-field terms that involve any quantum atom, are removed from the fully classical topology, except if it involves the join atom, the connect atom plus at least one additional classical atom. Table 1.1 gives a list of force-field terms still required.

It should be mentioned that the bond-constrained link atom approach is compatible to a new development called adjusted connection atom (ACA) [29]. This approach employs a special semi-empirical parameterisation for an atom that mimics a carbon atom, but is monovalent. Such an artificial atom is then used as a link atom, which is at the same time the connect atom. This corresponds to a bond length ratio s equal to unity.

It has been argued that the traditional link atom approach was superior to other methods using fixed link atoms for energy minimisation [4]. However, the differences to fixed link atoms was found to be small [29]. In a molecular dynamics simulations, the unwanted artificial increase of

the degrees of freedom is more important. Moreover, the presence of moving artificial link atoms affects the kinetic energy. This is why the constrained link atom scheme has been used in the application reported in Chapter 5.

Directly Bonded Atoms Covalently bonded atoms are normally too close to each other, so any non-bonded interaction, either Lennard-Jones or electrostatic, would be much too strong. In fact, due to the chemical bond present, these interactions are physically absent. So bonded atoms (first neighbours) are usually excluded from non-bonded interactions. The same applies for atoms connected by two bonds. An exception is the chemical bond with a predominant electrostatic character, as present in the binding of metal cations to negatively charged species. Such a situation is often modeled by the balance between the attractive electrostatic interaction, and the repulsive part of the Lennard-Jones interaction. However, it is sometimes difficult to reproduce such a situation, and auxiliary bonded force-field terms can be used in addition to the non-bonded interactions.

Neglect Conventions Similar to first and second neighbours being excluded from non-bonded interactions, these exclusions are retained if a molecule crosses the quantum-classical boundary. For the Lennard-Jones interaction, which is treated classically across the boundary, the following three atom pairs are excluded (see Figure 1.5): $\mathbf{J-C}$, $\mathbf{J-N_1}$, $\mathbf{Q_1-C}$. If a quantum topology is derived from a classical topology, the required exclusions are already present, so no changes are necessary.

As there is no classical electrostatic interaction across the boundary, there is no need to change the topology in this respect. However, the problem manifests itself by the background charges that enter the quantum-chemical calculation. Especially the link atom and the background charge from the connect atom would lie very close in space, thus introducing an unacceptably large distortion. So the partial charge on the connect atom is neglected. Any other partial charges are included. A comparison of different options for treating shortest-range charge interaction is presented by Antes [29]. The differences are rather small for the present embedding models and for semi-empirical wave functions.

It should be stressed that the boundary model cannot replace a sensible choice of the boundary. For example, the boundary should not cut a dipolar bond, which would make the dipole disappear. There should neither be a dipole in a bond between the connect and the $\mathbf{N_1}$ atom, which would leave a monopole after neglecting the charge on the connect atom. It is certainly good advice to place the boundary in a region of uncharged or weakly charged atoms. Multiple bonds should never be split, as the link atom approach cannot account for such special bonding situations.

Chapter 2

Molecular Dynamics Simulation with an *ab initio* Potential Energy Function and Finite Element Interpolation: Method and Validation

2.1 Abstract

An interpolation scheme for potential energy surfaces is presented. It employs a regular grid and finite element interpolation. The aim is the reduction of the computational expense for molecular dynamics simulation with a quantum chemical potential energy function. The methods used are described in detail. The feasibility is demonstrated and efficiency and accuracy are evaluated for the photoisomerisation of *cis*-stilbene in supercritical argon, using an *ab initio* configuration interaction treatment for the first electronically excited state of the stilbene molecule and classical force fields for the solvent-solute interactions (quantum mechanical / molecular mechanical molecular dynamics). The number of required quantum chemical calculations of energy and gradients was substantially reduced compared to a simulation not using the interpolation scheme. On the other hand, the impact on the accuracy is insignificant.

2.2 Introduction

The photoisomerisation of stilbene (1,2-diphenyl ethene, Figure 2.2) continues to attract interest from both experimentalists and theoreticians. The reaction dynamics of the isolated molecule is reasonably well understood [30, 31], but the shape of the potential energy surface of the first excited state is uncertain. While the isomerisation from the *cis* conformation to the *gauche* minimum seems to be a barrierless process [32], a barrier of approximately 14.5 kJ/mol for the *trans*-to-*gauche* process is considered to be experimentally evident (see [30] and references therein).

The situation in solution is more complex. The reaction has been investigated by femtosecond pump-probe absorption spectroscopy and fluorescence decay measurements under various conditions in many different solvents [13–15, 33–38]. There are two major additional effects for the reaction in solution: a solvent-induced modification of the potential energy surface, and a

viscous reduction of the flux due to mechanical friction. These effects are described by different theoretical models, but are hard to separate in experiment, thus a verification of the theory is difficult.

In most cases, an exponential decay of the reactant's concentration is observed. This finding is usually attributed to the presence of an energy barrier, which creates a bottleneck in the reaction pathway. In particular, an exponential decay is observed in the photoisomerisation of *cis*-stilbene. There is, however, no temperature dependence of the reaction rate except that caused by the viscosity of the solvent, which suggests a barrierless process. This contradictory situation is interpreted as the consequence of a small energy barrier [14]. However, Schroeder *et al.* report experiments [36] in which the barrier vanished, but the observed decay remained exponential.

Recently, several empirical force-field based models of the first excited state of stilbene and molecular dynamics (MD) simulations thereof were presented [39–42]. However, these models suffer from the deficiency that they are explicitly adjusted to reproduce the experimental findings, and thus have no predictive and little explanatory power.

We describe the potential energy surface of the reacting species (stilbene) in an unbiased way by quantum chemistry, using a combined quantum/classical model [2, 6]: The isomerising molecule (solute) is described using *ab initio* quantum chemistry (accurate, but expensive), while the solvent and the solvent-solute interaction are modelled purely by a classical force field (more approximate and cheaper). In order to reduce the computational expense of the quantum chemistry calculations, we recently developed an interpolation scheme for the potential energy surface of the reacting solute [43], which employs finite element interpolation. We now use an extension of it, which allows a greatly enhanced efficiency with a minor loss in accuracy. This chapter describes the new method. We use the photoisomerisation of *cis*-stilbene to demonstrate the feasibility of conducting molecular dynamics simulation with an *ab initio* potential energy surface for electronically excited states, for which classical force fields are notoriously difficult to parametrise. We report first results of the simulation of the photoisomerisation of *cis*-stilbene in supercritical argon. The side reaction, the photocyclisation that leads to dihydrophenanthrene, is not considered.

2.3 Methods

In this section, we describe the methods that we use for the interpolation of the potential energy surface of the isomerising molecule. Although for the present example of stilbene, the potential energy surface is reduced to three dimensions, we took care that all of the methodology is easily generalisable to multiple dimensions. However, all tests and applications have been carried out in three dimensions so far. Some formulae are given in three dimensions for clarity, according to the current application.

For ease of description, the system is split into two parts, the solute and the solvent. The solute consists of the stilbene molecule and is described by *ab initio* quantum chemistry. The argon atoms are the solvent. Both solvent-solute and solvent-solvent interactions are treated by classical force fields.

2.3.1 An Analog

Let us think of a large art museum with many rooms and many paintings, and with signposts that guide through the exhibition. Light bulbs are present in every corner of the rooms, in such

a way that a single bulb can shine in all the rooms that share this corner. Moreover, every room is equipped with a motion sensor that lights all the bulbs in the room as soon as someone enters the room. Once a bulb is lit, it is never extinguished. At the beginning, all rooms are dark.

Let us then think of an art connoisseur who visits the museum. Any room she enters will be fully illuminated because of the motion sensors. As she can read the signposts, she will be led to the rooms which exhibit art that interests her, and avoid rooms with paintings that she does not like. Very likely, the art lover will come back to the interesting rooms which are already illuminated, and will stay most of the time in these rooms. If, later, a second visitor with a similar art gusto enters the museum, he will find the majority of the interesting rooms already illuminated.

Let us now transform this picture to the interpolation scheme. The rooms of the museum are the finite elements that cover the conformation space. The light bulbs are the vertices, and the lighting of a bulb corresponds to the quantum chemical calculation. Note that, as the visitor (molecular system) enters an adjacent room, due to the special lighting design, some bulbs are already burning, and only a few will light up additionally. The quality of the paintings is related to the energy. The signposts are the energy gradients which show the way to the more interesting paintings. These low energy regions will be frequently visited, in contrast to the dull paintings (high energy) which will not be visited at all and will not be illuminated (not calculated).

2.3.2 Finite Element Interpolation for Molecular Dynamics Simulations

The molecule of interest is highly constrained and the remaining degrees of freedom span the conformation space in which the interpolation takes place. In the case of stilbene, all degrees of freedom are frozen except the central ethylenic torsion angle and the two adjacent phenyl torsion angles. For all conformations, energy and energy gradient (i. e. forces) are calculated as needed for the MD, and stored for later use. These so-called *vertices* then define the finite elements. Within an element, the potential energy surface is approximated by an interpolation polynomial, which is derived from information at the vertices.

In our original method of finite element interpolation of the potential energy surface [43], we used the points of the MD trajectory as vertices in the finite element mesh. An element was formed if enough vertices lied closer than a given maximum edge length from each other and fulfilled a set of other conditions. The maximum element edge length is the basic parameter which determines accuracy and efficiency of the interpolation method. It turned out that the elements may be rather large but still allow an accurate interpolation. The element's maximum edge length can be made much greater than the distance in conformation space between two subsequent time steps. Hence, it is advantageous to place a large element just in the region where the trajectory presumably will evolve in the next steps, and then interpolate these steps. A simple implementation of this concept uses a predefined, but not precalculated, grid which is built up from elements of exactly the size by which the desired accuracy is achieved. The energy at the vertices and its gradient is calculated as needed, namely as soon as the trajectory steps into a new element. This approach takes advantage of the fact that many simulation steps take place within the same element, which are covered by a few quantum-chemical calculations and interpolations at every step. As the vertices are shared among several elements, only few (usually one) additional vertices have to be calculated to cover the next couple of MD steps. Moreover, no computer time will be spent to regions of high energy that are not visited by the molecular

system. As the information at the vertices is permanently stored, additional trajectories can profit from the prior simulations.

We investigated the performance of this method with the same test system as before [43]. Comparing the largest energy difference between calculated and interpolated value that occurred during the simulation (as a measure of accuracy), the interpolation per calculation ratio (as a measure of efficiency) was increased by a factor of about 10–25 [44]. Moreover, the handling of the finite elements becomes simpler, faster and requires less storage.

2.3.3 Algorithm Outline

Formally, the procedure can be described as follows.

1. Choose which solute degrees of freedom of the molecular system to freeze, which to let evolve freely, and choose the maximum size of the elements. The number of active degrees of freedom defines the dimensionality of the interpolation grid.
2. Define the regular grid, without calculating any vertex yet.
3. Determine in which element the actual trajectory point lies.
4. Calculate those vertices of the element which are not yet known. Store this information for later use.
5. Interpolate energy and forces of the trajectory point from information at the vertices.
6. Calculate the additional classical forces for the solvent.
7. Propagate the system by an MD time step, taking into account the constraints for the frozen degrees of freedom, where necessary.
8. Proceed to step 3.

Thus, in the very first step of our three-dimensional example (and using simplicial elements), four vertices have to be calculated. Provided the element is much larger than the distance moved in a time step, the next few steps still lie in the same element and no additional quantum calculations are required. If the trajectory crosses the boundary of an element, one further vertex needs to be calculated, the three other vertices are shared with the first element. The next few trajectory points will then lie in the second element and no more quantum calculations are needed to interpolate the energy and forces for them. If many steps take place within the same element, this procedure will be very efficient. Needless to say, that the elements can be reused if the molecule returns to the same region in conformational space again, which was the basic idea for the interpolation. The same applies if a later trajectory follows a similar path as a previous one, or if the system is simulated under different conditions, such as different temperature, pressure, or type of solvent.

2.3.4 The Regular Grid

The regular grid is organised as follows. First, the conformation space is divided into equally sized orthorhombic subunits, which we call *bricks*. The size of the bricks must be specified for every dimension of the conformation space, and basically determines efficiency and accuracy of the method. Larger bricks allow increased efficiency, but are less accurate. If the conformation space only consists of dimensions of equal type, e. g. three dihedral angles, then the bricks are chosen to be cubic and their size is determined by one single parameter, the maximum element edge length l^{\max} , which is the space diagonal of a brick. So the brick edge l^{brick} is given by

$$l^{\text{brick}} = l^{\max} / \sqrt{n} \quad (2.1)$$

where n is the dimensionality of the conformation space. Every brick is then divided further into simplicial elements. The condition is that the triangulation of the entire conformation space is assured. If the bricks are translationally replicated in every dimension, then this condition is fulfilled if the element edges are parallel on opposite faces of the brick. One possible solution for the three-dimensional case, the one we used, is shown in Figure 2.1.

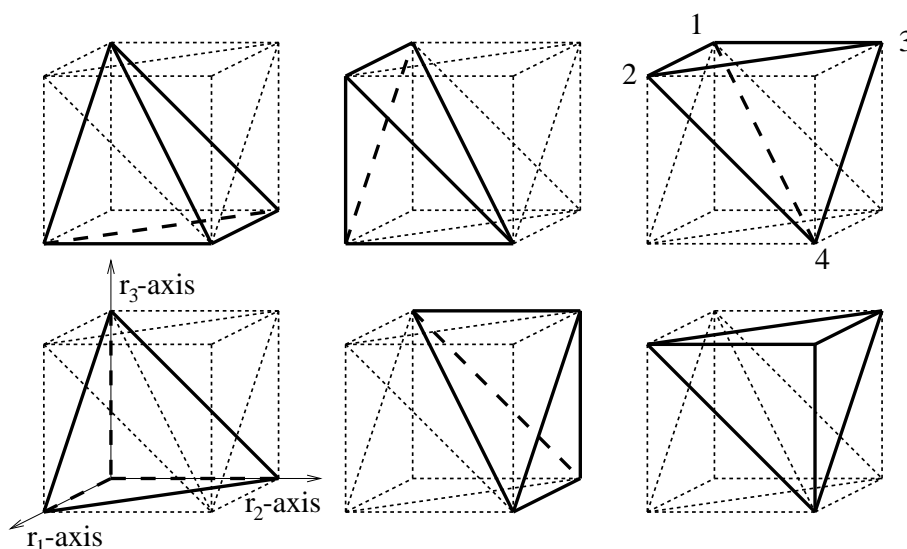


Figure 2.1: Example of six simplicial elements forming a cubic brick.

The interpolation actually takes place in the so-called *master brick*. This is a special virtual brick of the same size as in the grid, but translated to have one of its corners at the origin. The finite elements in the master brick may be characterised by a vector function $\vec{\Xi}^{s,v}$ which gives the coordinates of all the vertices v of the element s . For example, the characterisation of the element shown in the upper right corner in Figure 2.1 reads

$$\begin{aligned} \vec{\Xi}^{s,1} &= (0, 0, l_3^{\text{brick}}) & \vec{\Xi}^{s,2} &= (l_1^{\text{brick}}, 0, l_3^{\text{brick}}) \\ \vec{\Xi}^{s,3} &= (0, l_2^{\text{brick}}, l_3^{\text{brick}}) & \vec{\Xi}^{s,4} &= (l_1^{\text{brick}}, l_2^{\text{brick}}, 0) \end{aligned} \quad (2.2)$$

The numbering of the elements and the vertices is arbitrary. The bricks are then periodically replicated to cover the entire conformation space.

In the present implementation, a finite “active region” of the conformation space has to be selected in which the vertices are collected and the interpolations can take place. If non-bounded

coordinates are used for the interpolation, such as bond lengths, a finite sub-space has to be used. This should not give rise to problems, as bond lengths normally will not extend over a very wide range. However, making the active region much larger than actually required does not result in increased costs, since the expensive calculations are only carried out if the information is needed during the simulation. For convenience, the same implementation is also used for periodic coordinates such as the dihedral angles. It is ensured that the active region covers one full period of the coordinates. The restriction to a predefined active region is not of principal nature. With little extra book-keeping one can also implement an active region that expands as needed, thereby allowing also the treatment of truly non-bounded coordinates.

Because of the regularity of the grid it is quick and easy to determine the brick containing a newly found point \vec{r} , and to find the corresponding position \vec{r}' in the master brick. Once the appropriate brick is found, the element that contains the point is easily determined, by comparing the mapped position \vec{r}'_i of the point to the element's faces, which are described for the present case (see Figure 2.1) by the four planes

$$\begin{aligned} \frac{r_1}{l_1^{\text{brick}}} + \frac{r_3}{l_3^{\text{brick}}} &= 1 \\ \frac{r_2}{l_2^{\text{brick}}} + \frac{r_3}{l_3^{\text{brick}}} &= 1 \\ \frac{r_1}{l_1^{\text{brick}}} + \frac{r_2}{l_2^{\text{brick}}} + \frac{r_3}{l_3^{\text{brick}}} &= 1 \\ \frac{r_1}{l_1^{\text{brick}}} + \frac{r_2}{l_2^{\text{brick}}} + \frac{r_3}{l_3^{\text{brick}}} &= 2. \end{aligned} \quad (2.3)$$

In more than three dimensions, these equations may look more complicated. However, the general scheme described in [43] can be applied in any case.

The indices I_i^{vertex} of the required vertices are obtained by adding the coordinates of the vertices in the master brick to the index of the brick I_i^{brick}

$$I_i^{\text{vertex } v} = I_i^{\text{brick}} + \Xi_i^{s,v} / l_i^{\text{brick}} \quad (2.4)$$

The energy and gradients at the vertices are looked up in a table and plugged into the vertices of the element in the brick. If the required data is not yet available, it is calculated by quantum chemistry.

It should be noted that, at extra implementation effort, the regular grid could be given up in favour of an adaptive scheme allowing the partitioning of bricks into 2^n sub-bricks and so on (sometimes called oct-tree) if at a certain place a higher accuracy is required. However, some special kind of bricks is needed to interface between bricks of different subdivision for the sake of a correct finite element triangulation.

2.3.5 The Interpolation

In the current implementation, we employ a quadratic interpolation polynomial, which reads in the case of three dimensions

$$P(\vec{r}) = C_1 + r_1 C_2 + r_1^2 C_3 + r_2 C_4 + r_2^2 C_5 + r_1 r_2 C_6 + r_3 C_7 + r_3^2 C_8 + r_1 r_3 C_9 + r_2 r_3 C_{10} \quad (2.5)$$

where the r_i are the coordinates in the conformation space, and the C_1 through C_{10} are ten unknown coefficients that are determined by the interpolation procedure. Using this quadratic interpolation polynomial, it is convenient to use the energies at the vertices and additionally the

energies at every mid-point of the edges of the element. The energies E_m of the mid-points are interpolated between vertex positions \vec{p}_i and \vec{p}_j by a third-order polynomial, using the energies E_i and E_j and directional derivatives $\left(\frac{\partial E}{\partial p}\right)_{\vec{p}_i}$ and $\left(\frac{\partial E}{\partial p}\right)_{\vec{p}_j}$ along the edge $\vec{p} = \vec{p}_i - \vec{p}_j$

$$E_m = \frac{1}{8} \left(\left(\frac{\partial E}{\partial p}\right)_{\vec{p}_j} - \left(\frac{\partial E}{\partial p}\right)_{\vec{p}_i} \right) |\vec{p}_i - \vec{p}_j| + \frac{1}{2}(E_i + E_j) \quad (2.6)$$

By this procedure we obtain ten energy values (four from the vertices directly, and six from the mid-points), from which the ten unknown coefficients of Equation 2.5 can be calculated. The vertices and the mid-points are collectively referred to as *mesh points*.

Note that the use of a quadratic interpolation polynomial guarantees the continuity of energy over the element boundaries, whereas the gradients may be discontinuous, which is unphysical. However, if the potential energy surface is well-behaved and the elements are not too large, then the discontinuities will be small. Indeed, our previous study [43] demonstrated that the energetics and dynamics are not affected by this effect, except for maximum element lengths greater than 0.6 rad.

The interpolation of the potential energy surface and its gradient follows common finite element practice. The element in which the interpolation takes place, is transformed into the so-called *master element*. Energy and gradients are then interpolated by the aid of *shape functions* and the result is transformed back to the original position. In the rest of the subsection, a more elaborate description of this process is given.

The master element is a special virtual element, which has the origin and unity on every axis as vertices. Any irregular (triclinic) finite element can be transformed into the master element by a coordinate transform. The transformation of the point \vec{p}_t subjected to interpolation is achieved by the matrix operation

$$\vec{r} = \mathbf{M}^{-1}(\vec{p}_t - \vec{p}_0) \quad (2.7)$$

with the transformation matrix \mathbf{M} defined by the translated vertices of the triclinic element

$$\mathbf{M} = \left((\vec{p}_1 - \vec{p}_0) \quad (\vec{p}_2 - \vec{p}_0) \quad (\vec{p}_3 - \vec{p}_0) \right) \quad (2.8)$$

where the \vec{p}_0 to \vec{p}_3 are the positions of the vertices of the finite element. The point \vec{r} then has the same relative position in the master element as the point \vec{p}_t has in the actual element. As the energies are invariant under coordinate transformation, they can be taken directly from the actual element.

Inside the master element, the shape functions are defined. These shape functions have the noteworthy property to take the value of unity at exactly one of the mesh points, while being zero at all the other mesh points. From that condition, the shape functions Ψ_m relevant for us can be derived (see [43] for details). The interpolation function $P(\vec{r})$ is obtained by linear combination of the shape functions with the energies of the corresponding mesh points m as coefficients

$$P(\vec{r}) = \sum_m E_m \Psi_m(\vec{r}) \quad (2.9)$$

Energy and gradient are calculated from the interpolation function $P(\vec{r})$ and its derivatives with respect to the components of \vec{r} . While the energy again can be taken directly back to the actual element, the gradient requires back-transformation

$$\nabla \vec{p}_t = (\mathbf{M}^T)^{-1} \nabla \vec{p} \quad (2.10)$$

2.3.6 The Quantum/Classical Combination Model

In the present simulation, the solvent is modelled purely by a force field, while the solute is described entirely by *ab initio* quantum chemistry. According to the Born-Oppenheimer approximation, the potential energy for the solute equals the total electronic energy plus the Coulomb repulsion of the nuclei. The coupling of the two parts follows very much the general outline of Field *et al.* [2] and Liu *et al.* [6]. The Lennard-Jones interaction between solvent and solute is also described in terms of a force field. Solvent effects occur solely via irregular collisions between solvent and solute which are modelled as Lennard-Jones interactions, thus exerting some kind of molecular friction. As the solvent is uncharged and apolar, intermolecular Coulomb interactions do not occur. Polarisation effects are neglected. Hence, the potential energy surface of the solute is equal to that of the isolated molecule.

2.3.7 The Quantum Chemistry Method

For the determination of the potential energy of the first excited state of stilbene, a configuration interaction calculation with single excitations (CIS) was performed using the Gaussian 94 package [45]. We used the 6-31G basis set and observed that the inclusion of polarisation functions did not significantly change the results of several single points on the expected reaction coordinate, but was much more expensive computationally and considered unaffordable. A preliminary investigation revealed that the orbitals 27–80 contributed most to the CI matrix. So the other orbitals were not included in the CI in the main calculations. At this level of approximation, a single point evaluation of energy and forces took around 25 minutes on a DEC Alpha 440 MHz processor.

2.3.8 Computational Details

Stilbene Geometry

The geometry of the isomerising molecule was kept rigid except for the central ethylenic dihedral angle and the two adjacent phenyl ring torsion angles. The values for the remaining geometry parameters were obtained from an optimisation of the *gauche*-conformation of the first excited state, calculated with the above-mentioned *ab initio* method, with certain similar coordinates constrained to have the same value. These constraints are visible in Figure 2.2, and the corre-

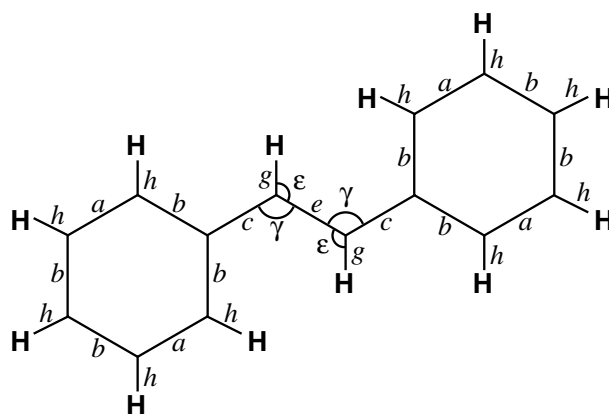


Figure 2.2: Geometry definition for the stilbene molecule.

sponding geometric parameters are given in Table 2.1. Unlabelled bond angles were all fixed at 120° . The phenyl rings and the atoms bonded to them are constrained to planarity. These constraints assume a fast internal vibrational relaxation upon excitation compared to the relatively slow isomerisation reaction, as suggested by Syage *et al.* [31], and a negligible coupling of the fixed degrees of freedoms to the active dihedrals. The interpolated conformation space is spanned by the three remaining dihedral angles. The central dihedral angle around the *e* bond is labelled r_2 , and the two adjacent phenyl torsion angles around the *c* bonds r_1 and r_3 respectively.

parameter	value
<i>a</i>	0.13788 nm
<i>b</i>	0.14084 nm
<i>c</i>	0.14148 nm
<i>e</i>	0.14215 nm
<i>g</i>	0.10824 nm
<i>h</i>	0.10724 nm
γ	127.4°
ϵ	116.6°

Table 2.1: Geometry parameters for the stilbene molecule. Lowercase Latin characters indicate bond lengths, Greek characters denote bond angles.

Simulation Parameters

The Lennard-Jones parameters for the involved atoms were taken from the GROMOS96 force field 43A1 [46] and are given in Table 2.2. The interaction was cut off at a distance of 0.9 nm. The time step of the leap-frog algorithm was 1 fs. The temperature was weakly coupled [47] to a bath with 0.1 ps relaxation time. Geometry fixing of the stilbene molecule was achieved by distance constraints (SHAKE, [48]) with a relative tolerance of 10^{-6} and dihedral angle constraints [49] with a tolerance of 10^{-6} rad.

The computational box with cubic periodic boundary conditions (5.035 nm edge length) contained one stilbene molecule and 2744 argon atoms. Initially, the solvent was equilibrated around a fully rigid *cis*-stilbene for 20 ps. Taking into account the Franck-Condon principle, the initial values for the free dihedral angles were taken from *ab initio* geometry optimisations of the *ground state* at the HF/6-31G** level. They were 4.5° for r_2 and 43.5° for r_1 and r_3 for the *cis* conformation. These values agree with neutron scattering experiments [50]. To obtain several different starting configuration, the solvent around the fixed molecule was equilibrated further and coordinate snapshots every 1 ps were used.

Atom type	GROMOS type	ϵ [kJ/mol]	σ [nm]
Argon	AR	0.996	0.341
Carbon	C	0.40587	0.33611
Hydrogen	HC	0.11838	0.23734

Table 2.2: Lennard-Jones parameters for the atoms involved.

2.4 Results and Discussion

2.4.1 Preliminary Investigation of the Potential Energy Surface of Photoexcited Stilbene

We first investigated the potential energy surface of photoexcited stilbene by a few single point calculations (CIS/6-31G**). We note the following observations:

- There are two minima. One is the planar *trans* conformation ($r_2 = 180^\circ$), the other is near a *gauche* conformation, with the central dihedral $r_2 = 44.1^\circ$ and the side dihedrals $r_1 = r_3 = 8.1^\circ$, which is 3.4 kJ/mol lower than the *trans* minimum. However, the exact location of the *gauche* minimum is very susceptible to small changes of the geometric constraints.
- There is a barrier between the two minima, located at the perpendicular (*perp*) conformation with $r_2 = 90^\circ$ and $r_1 = r_3 = 0^\circ$. The barrier is 15 kJ/mol above the *trans* minimum, which is in line with experimental observations [30].

2.4.2 Simulations

Twenty-one NVT-simulations of the photoisomerisation of *cis*-stilbene, differing in the initial solvent conformations, were performed at 236.7 K. Two of the simulations resulted in isomerisation to the *trans* minimum, while all other ended up in the *gauche* minimum. Figure 2.3 shows three examples of internal coordinate and solute potential energy trajectories. They all exhibit the same course in the first 50 femtoseconds and then diverge. The solid line represents an example of the majority of trajectories resulting in the *gauche* minimum and the broken lines show simulations that led to the *trans* conformation. While the dashed line obviously represents a case in which the barrier is crossed using the (kinetic) energy gained from the rapid downhill movement from the *cis* Franck-Condon region, the dot-dashed line seems to represent a thermally activated barrier-crossing process.

Twenty simulations of 5 ps length were performed at increased temperature of 348.0 K. Five of these simulations ended up in the *trans* minimum, two crossed the barrier to the *trans* conformation and recrossed back to the *gauche* minimum (see Figure 2.4). The other thirteen simulations did not show any barrier crossings and remained *gauche*.

2.4.3 Accuracy of the interpolation

To test the accuracy of the interpolation scheme, two sets of simulations have been performed using different maximum element edge lengths of 0.5 rad and 0.25 rad, thereby increasing the density of the mesh points by a factor of 8. By halving the brick size, the mesh points of the coarse grid can be reused for the fine grid. An exhaustive simulation of the first 100 steps with a quantum-chemical evaluation of energy and forces in every step was also performed.

Figure 2.5 shows the deviation of the trajectories with interpolations from the “true” trajectory without interpolations. The largest deviation of the solute potential energy is 1.21 kJ/mol ($\approx 0.6 k_B T$) for the coarse grid, while it is 0.18 kJ/mol ($\approx 0.09 k_B T$) for the interpolation with a fine grid, which is about 7 times more accurate. The deviations in the dihedral angles are very small

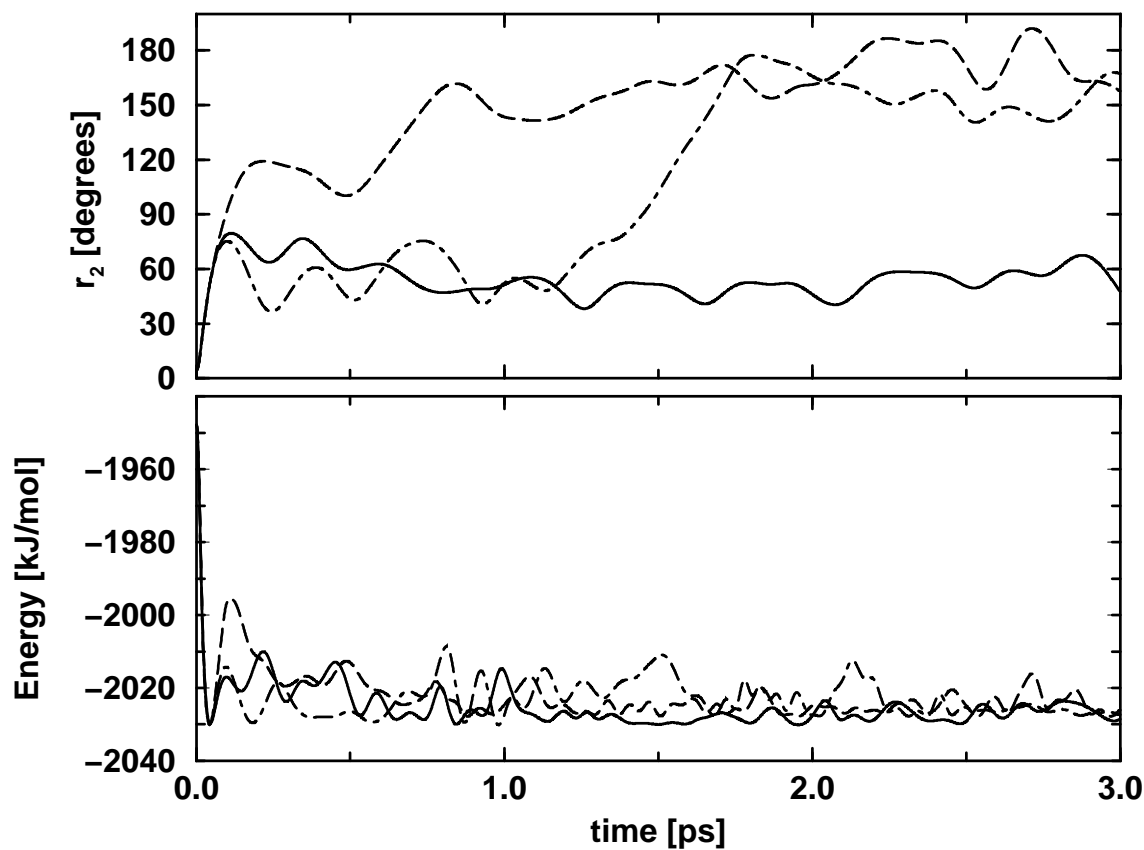


Figure 2.3: Solute potential energy and internal coordinate trajectories of the central dihedral for three simulations of the photoisomerisation of cis stilbene at 236.7 K.

	r_1 / degrees	r_2 / degrees	r_3 / degrees	E / kJmol^{-1}
coarse grid (0.50 rad)				
average	0.074	0.080	0.060	0.41
largest	0.11	-0.13	0.16	-1.21
fine grid (0.25 rad)				
average	0.007	0.012	0.004	0.062
largest	0.013	0.030	0.012	-0.18

Table 2.3: Deviations of the trajectories with interpolations from the one without interpolations within the first 100 steps of the simulation.

and are summarised in Table 2.3. They are reduced by approximately one order of magnitude by going from the coarse grid to the fine grid.

Figure 2.6 shows the dihedral angles and energy trajectories obtained with finite element interpolation using a coarse grid (dashed) and a fine grid (solid line). Both trajectories show qualitatively the same pathway up to approximately 1.5 ps, and then diverge exponentially. The molecule rapidly falls down a steep hill, crosses near the *gauche* minimum and climbs up the *perp* barrier without crossing it. Then the molecule relaxes into the *gauche* minimum. Note that,

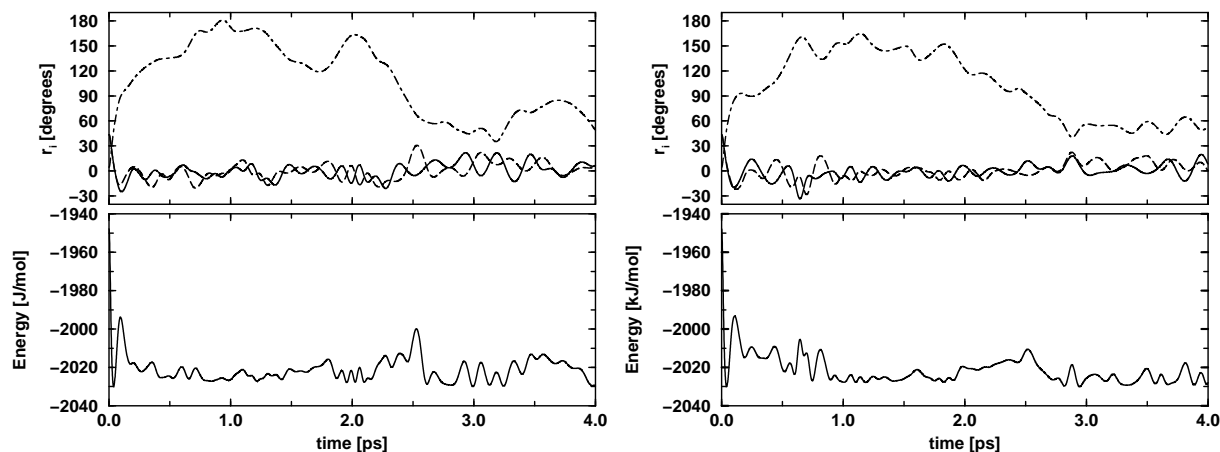


Figure 2.4: Two trajectories of the photoisomerisations of *cis*-stilbene at 348 K that show barrier recrossing. Internal coordinate trajectories are shown in the upper part (dot-dashed line: centre dihedral angle, solid and dashed lines: phenyl torsion angles). The lower parts show the molecule’s potential energy trajectories.

in spite of the symmetric starting configuration, the two phenyl torsion angles evolve differently due to the microscopically anisotropic friction exerted by the solvent atoms.

There is no sign of an energy barrier near the starting configuration. Potentially, there could be small localised features on the potential energy surface that are poorly represented by the interpolated surface. However, this is clearly not the case because the trajectories with and without interpolations perfectly coincide in the region where a barrier could be expected. It is not clear whether the limited quantum chemical method is not able to reproduce an existing small barrier, whether the barrier is induced by very specific solvent-solute interaction which are neglected in the present study, or whether there is indeed no barrier at all.

2.4.4 Efficiency

The power of rendering MD simulations with a quantum chemical potential affordable by finite element interpolation is demonstrated by Table 2.4. Compared to an exhaustive all-point calculation simulation, the interpolation method reduces the amount of explicit quantum chemical calculations to 1% using the coarse grid, and to 2.4% for the fine grid for the first trajectory. The efficiency of the method is greatly amplified if multiple trajectories are generated which follow a similar pathway as previous ones, since vertices already calculated can be reused. This is demonstrated by the “all trajectories” rows of Table 2.4.

Table 2.4 also shows that the efficiency clearly depends on how much of the conformation space is explored, as the simulations that lead to the molecule relaxing quickly in the *gauche* minimum close to the *cis* starting configuration are considerably more efficient than the ones that exhibit a barrier crossing to the *trans* conformer. The efficiency of the simulations at higher temperature is not quite as high. This is clearly a consequence of the fact that more barrier crossings occur, and the general effect that a larger part of the conformation space is accessible, which both require more vertices to be evaluated.

The same applies if the grid is refined by dividing the grid size by an integer factor k . For a densely populated coarse grid, after refining already $1/k^n$ of the vertices are known. On the other

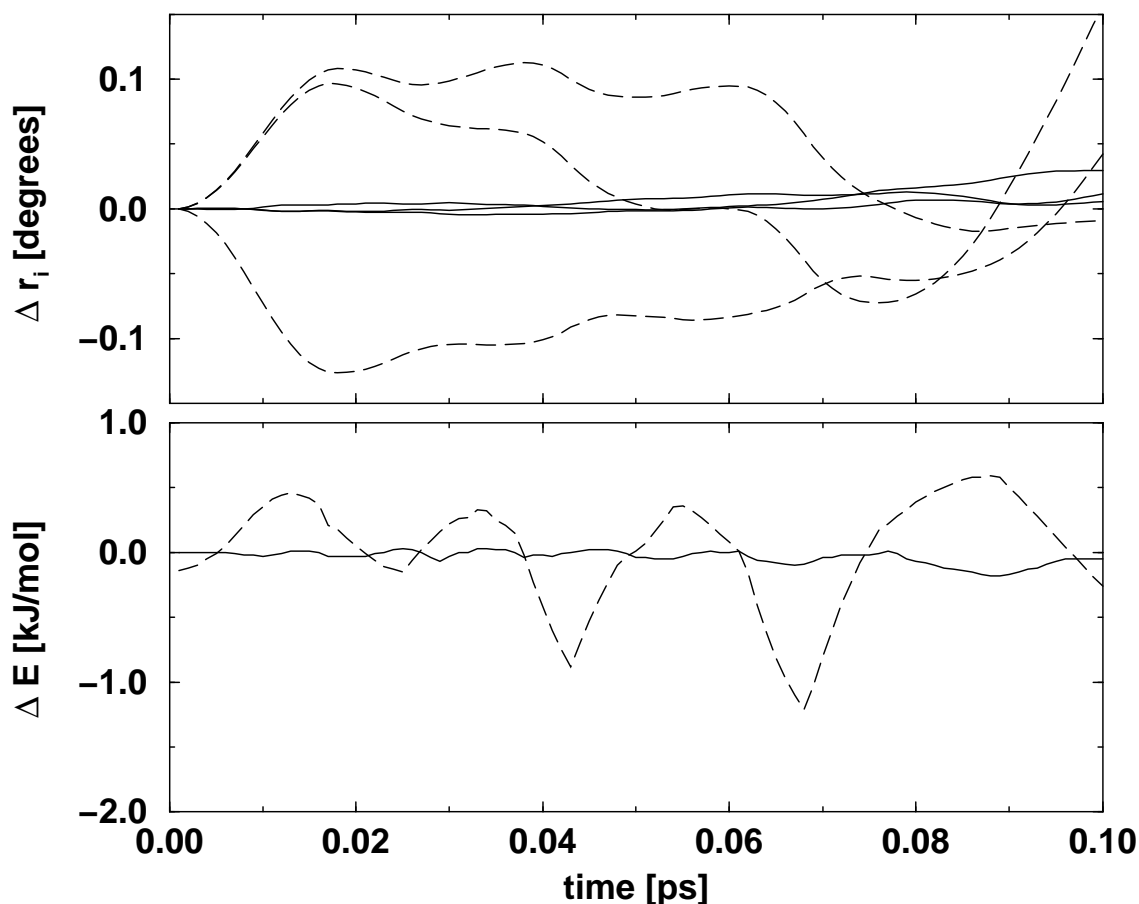


Figure 2.5: Deviation of the trajectories with interpolation from the trajectory without interpolations within the first 100 steps of the simulation. All simulations started from the same initial conditions. Above: Active dihedrals trajectories, below: solute potential energy trajectories. Solid lines: fine-grained grid interpolation; dashed lines: coarse-grained grid interpolation.

hand, the expense for running the simulation are multiplied by a factor k^n in principle. We found, however, that the computational cost only increases by a factor of 2.7 instead of theoretically 8. This is most probably due to the one-dimensional character of the short trajectories.

2.5 Conclusions

We have simulated the photoisomerisation of *cis*-stilbene using *ab initio* quantum-chemical description of the potential energy surface of the molecule and a classical description for the motion of all atoms and the intermolecular interactions. The computational expense of the quantum chemical part of the simulation is greatly reduced by a finite element interpolation scheme. The method employs a regularly bricked grid and the required information is calculated as needed.

The method is shown to yield accurate results while being very efficient. The number of quantum-mechanical evaluations of points of the potential energy could be reduced to 1% to 2.4% of the MD steps, thus rendering the simulation of large photoexcited molecules at an *ab initio* level affordable. Errors of the interpolation were below 1.5% in energy for the simulation with the coarse grid, and deviations in the dihedral angles were negligible. The accuracy could

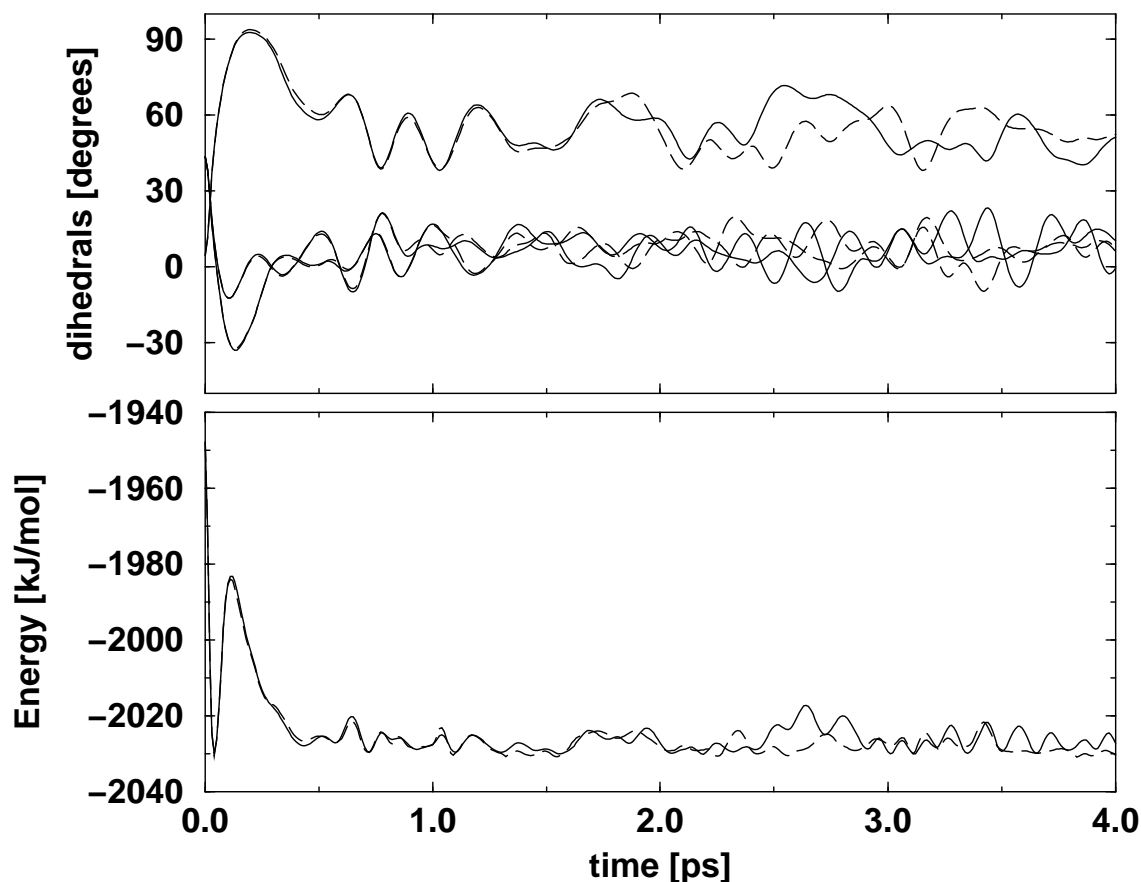


Figure 2.6: Internal coordinate and solute potential energy trajectories. Comparison between simulation with fine grid (solid lines) and coarse grid (dashed lines) interpolation.

max. edge length	Temp [K]	number of vertices used			number of steps	enhancement factor
		min	max	average		
0.5 rad	236.7	32	100 ^a 57 ^b	45.5	5000	109.8
		all trajectories			148	125000
0.25 rad	236.7	74	256 ^a 136 ^b	120.3	5000	41.6
		all trajectories			497	110000
0.25 rad	348.0	99	317 ^a 174 ^b	180.9	5000	27.6
		all trajectories			761	100000

Table 2.4: Efficiency of the finite element interpolation method. The number of vertices required by 21 individual trajectories, all 5000 steps long, are given as minimum, maximum and average values. The enhancement factor is the ratio of MD steps per vertex required for an individual trajectory. The “all trajectories” values concern the expense of all simulations with the indicated parameters together. The numbers do not take into account the vertices in the fine grid already known from the previous simulations with the coarse grid. *a*: with isomerisation to trans, *b*: without isomerisation to trans.

be increased by a factor of about 8 by halving the size of the bricks, at the cost of 2.5 times more explicit quantum chemical evaluations.

A further advantage is the possibility of reusing the known parts of the potential energy surface for simulating more trajectories, which substantially increases the overall efficiency of the method. Also, the desired accuracy or affordable expense may be tuned by selecting an appropriate brick size. Accuracy can be improved by dividing the bricks up. Again, information from previous simulation can be recycled.

As this is our first application of *ab initio* molecular dynamics to excited states, this chapter focuses on the feasibility of such calculations. Clearly, the chosen *ab initio* method can be improved and a treatment at higher level of theory (MCSCF for example) is desirable. However, with all due caveats, we can report first results on the stilbene system, which suggest that there is no potential energy barrier between the *cis* and the *gauche* conformation of stilbene in the first excited state. For the barrier crossing process to the *trans* conformer, two mechanisms are important: Thermal activation and barrier climbing enforced by the inertia of the motion. At higher temperatures, even barrier recrossings occur. The further study of the photoisomerisation of stilbene in more detail and under different conditions is presented in the following chapters.

Chapter 3

The Photoisomerisation of *cis*-Stilbene Does not Follow the Minimum Energy Path

3.1 Summary

Computer simulation of the photoisomerization of *cis*-stilbene demonstrates that barrier crossing reactions can occur without thermal activation, but with excess energy from the photoexcitation. Moreover, the reaction proceeds with large energy transfers but small conformational changes. This has an impact on the reaction dynamics.

3.2 Introduction

The development of femtosecond spectroscopy has made it possible to monitor chemical reactions in realtime. As a prototype example, the photoisomerization of stilbene in solvent has been examined under various conditions (temperature, pressure, solvent) [14, 33–37, 51]. However, important issues such as the atomic detail of the reaction dynamics or the shape of the potential energy landscape still remain experimentally unresolved. Computational methods allow the detailed study of the time-resolved dynamics of these systems. On the one hand, molecular dynamics [26, 46], provides the time-dependent evolution of the system and allows for effects due to temperature and solvent. On the other hand, accurate potential energy surfaces are provided by quantum chemistry [45]. By combining the two methods [1, 2, 6, 52], the best from both worlds can be used: The effect of the solvent and the evolution in time by molecular dynamics, and the unbiased description of the reacting molecule by quantum chemistry. However, the latter is computationally rather demanding. Evaluation of potential energy and forces for photoexcited stilbene takes about half an hour on a modern microprocessor. As thousands, or rather millions of such evaluations are required for a meaningful molecular dynamics simulation, a straightforward implementation of this concept is not feasible. For this reason, we recently developed an interpolation method based on finite elements [43, 53]. The method represents the part of the potential energy surface that is required during the simulation without losing the accuracy of the quantum chemical method. This reduces the number of quantum chemical evaluations by a factor of about 2000 compared to a straightforward implementation. Computational details are given in Chapter 2. Here, it is sufficient to say that stilbene in its first excited state is treated by

a singly-excited configuration interaction calculation in a 6-31G basis set. Supercritical argon is used as a model solvent at the appropriate density.

3.3 Potential Energy Surface

Figure 3.1 shows the potential energy surface as a representation in internal coordinates, namely the central ethylenic dihedral angle and the phenyl torsion angle. Both phenyl torsion angles behave similarly, so one of them is omitted for clarity. The trajectory of the isomerizing system is drawn as a white line. Starting point is the *cis* conformation in the upper right corner. Being a minimum in the ground state, this conformation has a relatively high potential energy in the excited state (82 kJ/mol above the *gauche* minimum). The energy difference is provided by the photoexcitation. After a rapid downhill motion, the minimum in the primary *gauche* region (central dihedral angle $\approx 50^\circ$, side dihedral angle $\approx 10^\circ$) is crossed. Then the barrier is climbed. It is located at a central dihedral angle of $\approx 90^\circ$. The barrier is crossed, and the system relaxes into the wide 8-shaped minimum in the *trans* region. This trajectory is representative for most barrier crossings that we observed in our study.

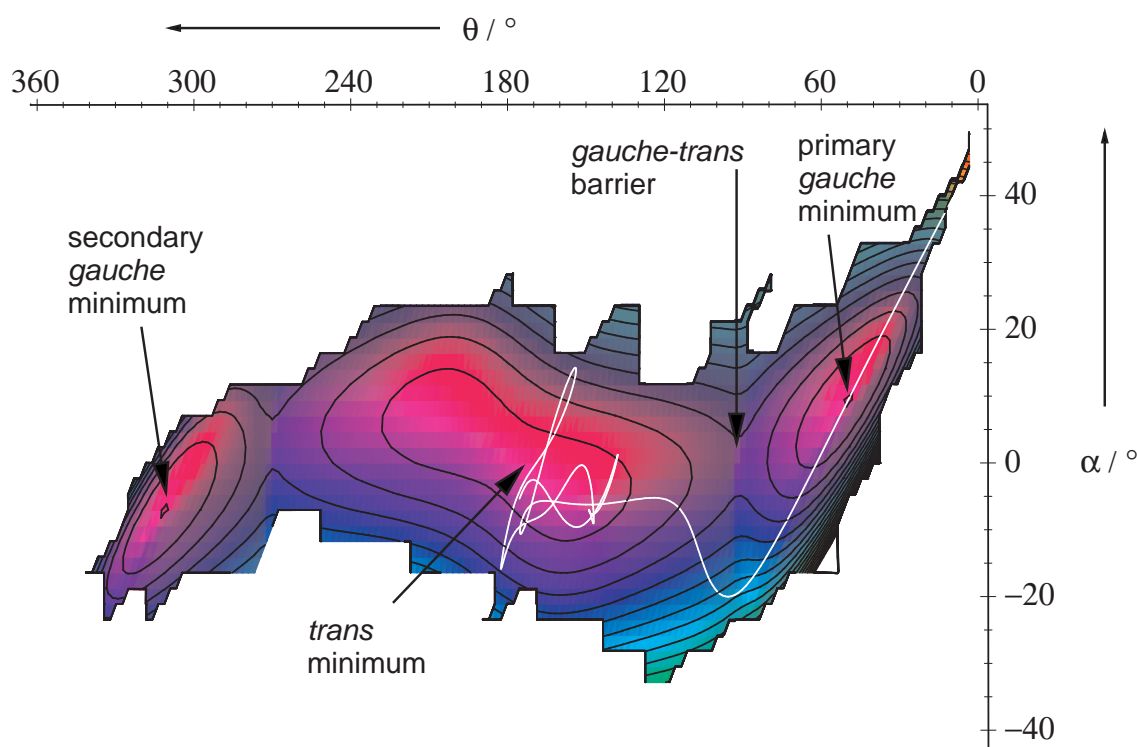


Figure 3.1: The potential energy surface of stilbene in its first excited state, as calculated during the simulations. An example trajectory which crosses the barrier is drawn as a white line. The contour lines are 5 kJ/mol apart.

3.4 Kinetic Activation

Two conclusions are drawn from the above. (i) The primary *gauche* minimum is passed in one go. No relaxation takes place, and no thermal activation is required to leave it. The activation

energy is provided by the kinetic energy gained from the initial downhill motion. (ii) The barrier is not crossed at the saddle point, which is at a position of 0° of the phenyl torsion angles and 14.2 kJ/mol above the *gauche* minimum. In fact, the average barrier crossing location is -20° for both phenyl torsion angles, and the corresponding average energy is more than 19.3 kJ/mol above the saddle point. Thus, the concept of the reaction taking the “minimum energy path” is clearly not admissible. We note that in the few thermally activated events that we observed, the barrier is crossed over the saddlepoint *on average*, *but still* the average deviation from the saddle point is 7° for the phenyl torsion angles, and the energy is 4.2 kJ/mol above the saddle point.

3.5 Solvent Effect

The solvent has a profound effect on this reaction. Figure 3.2 shows averaged trajectories of the system at different pressures. The time evolution of the central dihedral angle and one of the phenyl torsion angles is displayed. A clear separation strictly according to pressure is obvious. Such a behavior is expected, as at higher pressure the motion is quenched more effectively, and energy is dissipated to the solvent environment. The solvent influence is clearly visible beyond 40 fs . At low pressure, most trajectories migrate inertially to the *trans* minimum. With increasing solvent pressure, the majority of molecules is quenched before reaching the barrier, so they relax to the *gauche* minimum. This effect influences the rate constant measured by femtosecond

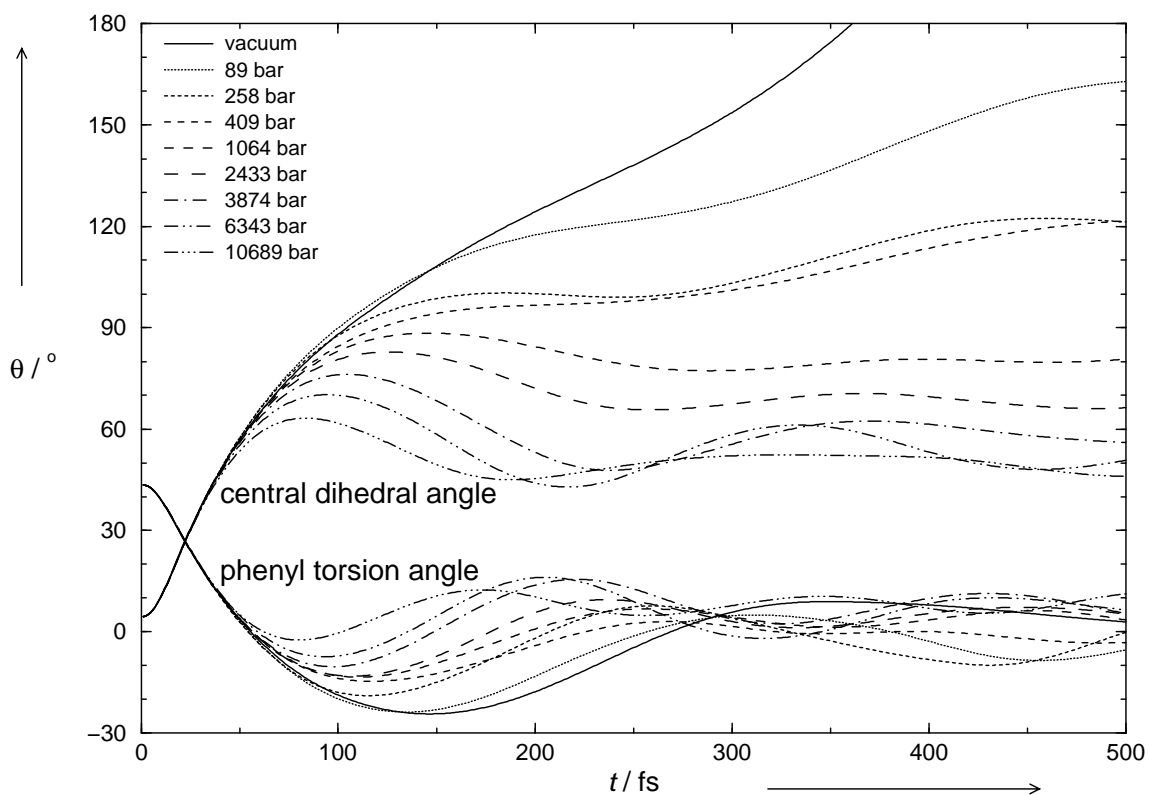


Figure 3.2: Averaged dihedral angle trajectories from simulations at 190 K at different pressures. All trajectories of a series of 20 simulations at the same state point, but differing in the initial solvent configurations, were averaged.

experiments [14, 35]. Investigations to connect this behavior quantitatively to the solvent shear viscosity are presented in the next Chapter.

3.6 Molecular Shape Changes

It is, however, rather surprising that within the first 40 fs all trajectories coincide in spite of the environmental conditions ranging from vacuum over liquid to solid solvent. The solvent seems to have no effect on this short-time behavior. In the same time span, both dihedral angles change by more than 40° . However, despite the drastic change in internal coordinates, the overall shape of the molecule does not change much. Figure 3.3 demonstrates this by showing the initial conformation, and the conformation after 40 fs. Within this time, the two ethylenic carbon atoms and the hydrogen atoms bonded to them are displaced (bottom in Figure 3.3), whereas the bulky phenyl rings nearly remain in place. Thus, it is not required to move any solvent atom, and there is no influence by the solvent.

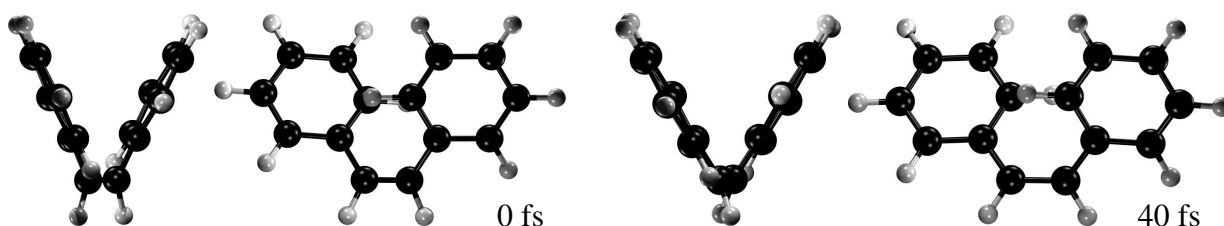


Figure 3.3: Initial conformation of *cis*-stilbene and conformation after 40 fs of reaction. Views from the left and from the front are shown for both conformations.

3.7 Conclusions

Our detailed analysis of the dynamics of photoexcited *cis*-stilbene shows that this reaction does not proceed by thermal activation. Rather, the excessive internal potential energy after the photoexcitation is used to overcome the barrier to the *trans*-minimum. The use of initial energy for the transition is facilitated by the fact that the first phase of the reaction can occur without large motion of the phenyl rings. There is no relaxation prior to the barrier transition, and the barrier is crossed far from the minimum energy path. Thus, the prerequisites for common concepts of reaction dynamics such as transition state theory or Rice-Ramsperger-Kassel-Marcus (RRKM) theory are not satisfied.

Chapter 4

Viscosity Dependence and Solvent Effects in the Photoisomerisation of *cis*-Stilbene

4.1 Abstract

Molecular dynamics simulations of the photoisomerisation of *cis*-stilbene in supercritical argon were performed. The stilbene molecule is represented by *ab initio* quantum chemistry, while the solvent, the interaction with solvent, and the time evolution were described by classical mechanics. Reaction rate constants are estimated and their dependence on temperature, pressure and viscosity are investigated. Agreement with available experimental data was obtained. Our simulations strongly suggest a minimum on the excited state potential energy surface at a *gauche* conformation which is very rapidly reached after excitation, which leads to non-equilibrium barrier transitions. Specific solvent effects were identified. Implications on the current opinion on stilbene photoisomerisation are discussed.

4.2 Introduction

In the past twenty years, new developments in laser technology and in spectroscopy techniques, made it possible to observe ultrafast chemical reactions in real-time. The two most powerful and versatile techniques are fluorescence decay measurements and pump-probe spectroscopy. Their main application areas are unimolecular photoreactions, such as photo-induced dissociation or photoisomerisation. Using such techniques, it is possible to directly investigate chemical reaction dynamics on the femtosecond time scale. For example, it is possible to spectroscopically observe the dynamics of the dissociation of a simple molecule in gas phase [54, 55]. However, the many details of more complex processes, such as a reaction in solution, still remain unrevealed. The solvent effect in the photoisomerisation of *cis*-stilbene is dramatic: while the lifetime of an isolated excited molecule is 0.32 ps [32], in solution it ranges from 0.5 ps in methanol [15], 1.0 ps in isopentane, 1.6 ps in hexadecane [13], to 2.1 ps in cyclohexane [15]. Over many years, Jürgen Troe and his coworkers have investigated the photoisomerisation of stilbene in solution by pump-probe spectroscopy [14, 33, 35–37, 51]. Yet, important features of the potential energy surface are unclear and left open to speculation, and the detailed dynamics are very hard to investigate in experiment.

Nikowa et al. [14] have found that the isomerisation rate constant depends linearly on the inverse solvent viscosity. The proportionality constant A depends on the solvent type. However, Todd and Fleming [34] suggest a more general approach using molecular friction as a reference, which they claim to be independent of the solvent. However, there is no clear definition of such a molecular friction.

Parallel to the experimental development mentioned above, computational chemistry has evolved. It can be roughly divided into the following two methodologies: (i) Quantum chemistry is capable of calculating the electronic energy, its gradients, and a great variety of other molecular properties at a high level of accuracy, if desired. However, it is limited to rather small molecular systems. (ii) Classical molecular dynamics uses an empirical force field and is capable of providing dynamic information of large systems at an atomic resolution. However, it is unable to simulate chemical phenomena such as bond cleavage and formation.

So why not perform molecular dynamics simulations of the photoisomerisation of stilbene? The problem is twofold. As crucial point, a force field of a photoexcited molecule is not easy to obtain. The general procedure of fitting the force field parameters to macroscopic properties of the liquid species, is not applicable. Another approach, to construct a potential energy surface suitable to reproduce selected spectroscopic data, has recently been conducted [39–42]. However, this procedure has little predictive power, since the desired results are put in previously, and then reproduced. Moreover, Eli Pollak and co-workers [56–59] apply sophisticated theories, based on transition state theory, to the photoisomerisation of *trans*-stilbene. This reaction differs from *cis*-stilbene insofar that it is most probably thermally activated. We demonstrate that this is not the case for *cis*-stilbene, and thus the prerequisites for transition state theory are not satisfied. Coming back to the construction of a potential energy surface, one could escape to quantum chemistry as a last resort. A quantum-chemical potential energy surface for stilbene [60, 61], embedded in a classical environment, seems to be an appropriate solution. By use of an interpolation scheme, designed for the reduction of computational expense in such a situation [43, 53], such a task is indeed feasible. The results are presented in this chapter. We note that non-adiabatic quantum effects such as couplings between states have not been taken into account. Lastly, can the system be adequately described by classical dynamics, or should one rather apply quantum dynamics? If we take the rule of thumb and say if $\hbar\omega \ll k_B T$, then the motion is classical, we must admit that both quantities are almost equal in the case of the *cis*-stilbene photoisomerisation. However, Gershinsky and Pollak [59] report that the influence of quantum effects is small, especially in condensed systems, and that classical molecular dynamics can be relied upon.

The following questions are addressed in this chapter:

- What does the potential energy surface look like?
- Which trends accompany variations in temperature and in pressure?
- Can the experimental reaction rates be reproduced by simulation?
- Is the shear viscosity as a macroscopic bulk property a good measure for the reaction rate constant as a molecular quantity?
- How can the solvent effect be described?

4.3 Methods

4.3.1 Computational Details

The photoreaction of *cis*-stilbene in argon solution has been simulated by means of a combined classical/quantum-mechanical model [1, 2, 6]. The solvent is described by the classical GROMOS96 [19] force field, while the potential energy surface of the reacting stilbene is obtained by *ab initio* quantum chemistry. A finite element interpolation scheme [43, 53] is used to reduce the computational expense of the quantum-chemical calculations. The method has been described in detail and the feasibility and efficiency for the same system as treated here has been demonstrated in Chapter 2. It is sufficient to know the following. A conformation of the molecule is fed in, and energy and gradients are returned. These quantities are calculated from an interpolated surface which is spanned by a fixed finite element grid. As soon as needed, the quantities at the grid points are calculated by *ab initio* quantum chemistry at the desired level. The results are stored for later use. Because the potential energy surface of the molecule only depends on its conformation, the same grid points can be reused through many series of simulation, also at different temperatures or pressures. This makes the method extremely efficient. However, the method requires the molecule to be constrained to a few degrees of freedom. The central ethylenic dihedral angle (labeled r_2) and the two phenyl torsional angles (r_1 and r_3) were the three degrees of freedom that spanned the potential energy surface. The other geometric parameters are optimized for the *gauche* minimum of the potential energy surface of the first excited state (S_1) of stilbene, and were constrained during the simulations. See Figure 2.2 and Table 2.1.

For the quantum chemical calculations, a configuration interaction including single excitations (CIS) in a restricted window of orbitals (from orbital number 27 to 80) has been used with the 6-31G basis set. The evaluation of the energy and the gradients with the Gaussian 94 program [45] took half an hour on average on a 440 MHz DEC Alpha processor. A higher level of theory or a larger basis set, while desirable, was still considered unaffordable. Nevertheless, the potential energy surface obtained by the above-mentioned method was found to be fairly reasonable: The height of the *trans-gauche* barrier is in agreement with common opinion amongst experimentalists [30], and the shape of the surface is in good agreement with spectroscopy-derived data reported by Frederick et al. [62]. A recent configuration interaction study [63] unfortunately does not list appropriate data for comparison.

The solvent-solvent and solvent-solute interaction is modeled by standard classical force fields. The Lennard-Jones parameters for the involved atoms were taken from the GROMOS96 force field [46] ($\sigma_{\text{Ar}} = 0.3410$ nm, $\epsilon_{\text{Ar}} = 0.9964$ kJ/mol, $\sigma_{\text{C}} = 0.3361$ nm, $\epsilon_{\text{C}} = 0.4059$ kJ/mol, $\sigma_{\text{H}} = 0.2373$ nm, $\epsilon_{\text{H}} = 0.1184$ kJ/mol). As combination rule for the interaction between different types of atoms, the arithmetic mean of the single atom type σ values and the geometric mean of the single atom type ϵ values are employed (Lorentz-Berthelot mixing rule).

The time step of the leap frog algorithm was 1 fs. The Lennard-Jones interaction was cut off at 0.9 nm. The temperature was weakly coupled [47] to a bath with 0.1 ps relaxation time. Geometry fixing of the stilbene molecule was achieved by distance constraints (SHAKE [48]) with a relative tolerance of 10^{-6} and dihedral angle constraints [49] with a tolerance of 10^{-6} rad. The computational box with cubic periodic boundary conditions contained one stilbene molecule and 2744 argon atoms. Several box sizes were used for simulations at different pressures, but the volume of the box was constant during the individual simulations.

According to the Franck-Condon principle, the initial conformations of the active dihedral angles of the stilbene molecule corresponded to the *cis* minimum of the ground state (HF 6-31G**).

The initial values for the free dihedral angles were 4.5° for r_2 and 43.5° for r_1 and r_3 and agree with neutron scattering experiments [50]. Initial configurations for the entire boxes were obtained by equilibrating the solvent atoms around a stilbene molecule which was held completely rigid by constraints. By using coordinates from snapshots every 1 ps in an equilibration simulation, several different starting configurations for the same state point were obtained. For the investigation of the dependence on viscosity, and some derived properties, series of simulations were performed at different temperatures and pressures.

4.3.2 Activation Energies

We can calculate the activation energy E_A for the reaction in an approximate way. Starting point is the Arrhenius equation,

$$k = F \exp(-E_A/k_B T) \quad (4.1)$$

with the pre-exponential factor F , the temperature T and Boltzmann's constant k_B . Linearized for E_A , we obtain

$$\ln(k) = -\frac{E_A}{k_B T} + \ln(F). \quad (4.2)$$

The reaction rate constant k is obtained from the outcome of the reaction: After a certain time interval τ , a certain ratio of the reactant molecules has already reached the final state, while the complementary ratio $s = I_a/I_0$ is still in the initial state. Assuming an exponential decay of reactants, we obtain for the reaction rate constant

$$k = -\frac{\ln(I_a/I_0)}{\tau}. \quad (4.3)$$

Insertion into Equation 4.2 yields

$$\ln(-\ln(I_a/I_0)) - \ln(\tau) = -\frac{E_A}{k_B T} + \ln(F). \quad (4.4)$$

Thus, in an $\ln(-\ln(s))$ vs. $-1/k_B T$ plot the activation energy E_A can be obtained from the slope of the regression line without knowing the value of τ , which only influences the intercept of the regression line.

4.3.3 Solvent Properties

The shear viscosity of the solvent was calculated from separate simulations of the solvent only. Simulation boxes were set up in such a way that they match the average pressures obtained from the simulations including solvent and solute. Simulation parameters were equal to the ones of the solution simulations. The pressure was sampled over 250 ps. The viscosity η was then obtained by the relation [26]

$$\eta_{\alpha\beta} = \frac{V}{k_B T} \int_0^\infty \langle \delta P_{\alpha\beta}(t) \delta P_{\alpha\beta}(0) \rangle dt \quad (4.5)$$

where V is the volume of the computational box. The integrand is the time autocorrelation function of the fluctuation of an off-diagonal element of the pressure tensor. The correlation time was obtained by fitting the numerically calculated normalized correlation function to a Lorentzian f

$$f(t) = \frac{1}{1 + at^2} \quad (4.6)$$

with the adjustable parameter a . The Lorentzian was found to fit very well, compared to exponentials or Gaussians, for example. The function f has the analytical integral

$$\int_0^{\infty} f(t) dt = \frac{1}{2} \frac{\pi}{\sqrt{a}} \quad (4.7)$$

which yields altogether

$$\eta_{\alpha\beta} = \frac{V}{2k_B T} \sigma^2(P_{\alpha\beta}) \frac{\pi}{\sqrt{a_{\alpha\beta}}} \quad (4.8)$$

with $\sigma^2(P_{\alpha\beta})$ being the variance of the pressure tensor element. The viscosity η is obtained by averaging the $\eta_{\alpha\beta}$ values obtained for the three different off-diagonal elements of the pressure tensor.

The diffusion coefficient D is calculated as

$$D = \frac{1}{6tN} \sum_{i=1}^N |\vec{x}_i(t) - \vec{x}_i(0)|^2 \quad (4.9)$$

with the elapsed time t after the starting configuration $\vec{x}_i(0)$.

4.3.4 Estimation of Reaction Rate Constants

In most of the experimental work [14, 15, 34, 35, 64], an exponential decay of the signal is observed. This finding is usually attributed to an energy barrier which creates a bottleneck in the reaction pathway [14]. In this case, a small barrier is assumed near the initial *cis* region. However, as we noted in our previous work [53], no such barrier is present in our *ab initio* potential energy surface. In this subsection, we discuss how this dilemma might be resolved.

In pump-probe spectroscopy, it is generally assumed that only conformations close to the Franck-Condon excitation region is spectroscopically visible. Nikowa et al. [14] estimate that in the case of stilbene, this small barrier is between $r_2 = 7^\circ$ and 14° for non-polar solvents. With this assumption and our simulated trajectories, however, the signal would abruptly disappear after a few femtoseconds, and would not decay on a picosecond time scale as observed experimentally. If we assume, in contrast, that the spectroscopically active region is more extended, then a different picture of the photoisomerisation kinetics is possible. It might also be that the probed molecule is not in the Franck-Condon region any more. Abrash et al. [13] find that there is no spectral shift after 100 fs which is their experimental resolution. Such a spectral shift would be likely upon conformational change. They conclude that a spectral diffusion is taking place faster than 100 fs. This interpretation is consistent with our study.

If we assume that a molecule leaves the region of spectroscopically visible conformations when it crosses the barrier between the *gauche* and the *trans* minimum, then the molecules that are caught in the *gauche* minimum remain visible, while the molecules that isomerise to the *trans* conformation disappear. The point is to find this ‘‘spectroscopic threshold’’. It need not be exactly at the barrier, and the exact determination is extremely difficult. Making a non-restrictive assumption, any value in the range between, say, 70° and 130° seems reasonable. There we have indeed an energy barrier between the spectroscopically active and inactive regions. We note that this does not truly lead to an exponential decay of the signal, because the process is predominantly kinetically activated, not thermally activated. In experiment, many different

effects (such as fluorescence, internal conversion, escape through a conical intersection or the photocyclisation to dihydrophenanthrene) may occur that influence the decay curve, which are not accounted for in our simulation. However, for simplicity, we assume that we can fit an exponential curve to our calculated data.

In an exponential decay with rate constant k , the ratio I_a/I_0 of the initial amount I_0 is still active after a time t :

$$I_a/I_0 = \exp(-kt). \quad (4.10)$$

From a set of simulations at a given state point, we can easily determine the ratio of molecules that remain active, as well as the average time t_T that is required to reach the spectroscopic threshold. From that, we can estimate the reaction rate constant

$$k = -\frac{\ln(I_a/I_0)}{t_T} \quad (4.11)$$

where I_a is the number of simulations remaining in the active region (in the *gauche* conformation in the case of stilbene) of a total number of performed simulations I_0 .

Nikowa et al. [14] state that the non-radiative rate constant k_{nr} can be decomposed into the rate k_{DHP} of the photocyclisation to dihydrophenanthrene, and a viscosity dependent term with the parameter A . The parameter A is solvent-specific, but temperature-independent.

$$k_{nr} = k_{\text{DHP}} + A/\eta \quad (4.12)$$

As the photocyclisation is not possible in the way we set up our simulation, we are left with the second term. Having calculated the rate constant k from Equation 4.11 and the solvent viscosity from Equation 4.8, we are able to calculate the parameter A

$$A = k \cdot \eta. \quad (4.13)$$

4.4 Results and Discussion

4.4.1 Potential Energy Surface

The interesting region of the potential energy surface, i. e. the regions which were at least once visited in all the simulations, is obtained as a by-product of the interpolation scheme. Table 4.1 gives an overview of special points on the potential energy surface, with their location and energy. For example, the initial downhill energy gain is 82.4 kJ/mol and the *trans-gauche* energy barrier is 11.9 kJ/mol.

Figure 4.1 shows a picture of the potential energy surface of stilbene in the first excited state. It shows a two-dimensional cut of the three-dimensional surface, with the condition $r_1 = r_3$ (both phenyl torsion angles have the same value). Only the part of the whole surface which was known after all the simulations is shown, so virtually the space that is accessible during the photoisomerisation. The high peak in the back of the picture is the initial *cis* conformation from where the simulations were started. Figure 4.2 shows a top view of a symmetric cut like Figure 4.1. In addition, an example trajectory is shown as a white line. The coarse-grained boundary shape of the surface originates from the grid used in the interpolation scheme. Clearly visible is the *gauche* minimum and the wide, shallow 8-shaped minimum in the *trans* region, as well as the barrier in between. The top views allow easy location of the barrier. The path

Feature	Location / degree		Energy kJ/mol
	central dihedral	phenyl torsions	
<i>cis</i> Franck-Condon region	4.5	43.5	82.4
<i>gauche</i> minimum	49.8	7.9	$\stackrel{!}{=} 0$
<i>perp</i> barrier saddlepoint	91.6	-1.6	14.2
<i>trans</i> minimum	157.3	-5.1	2.3
barrier between the <i>trans</i> minima	180.0	0.0	3.5

Table 4.1: Selected features of the potential energy surface of the first electronically excited state of stilbene. Locations in dihedral angle space and energy are shown. The energy origin is set to zero for the *gauche* minimum.

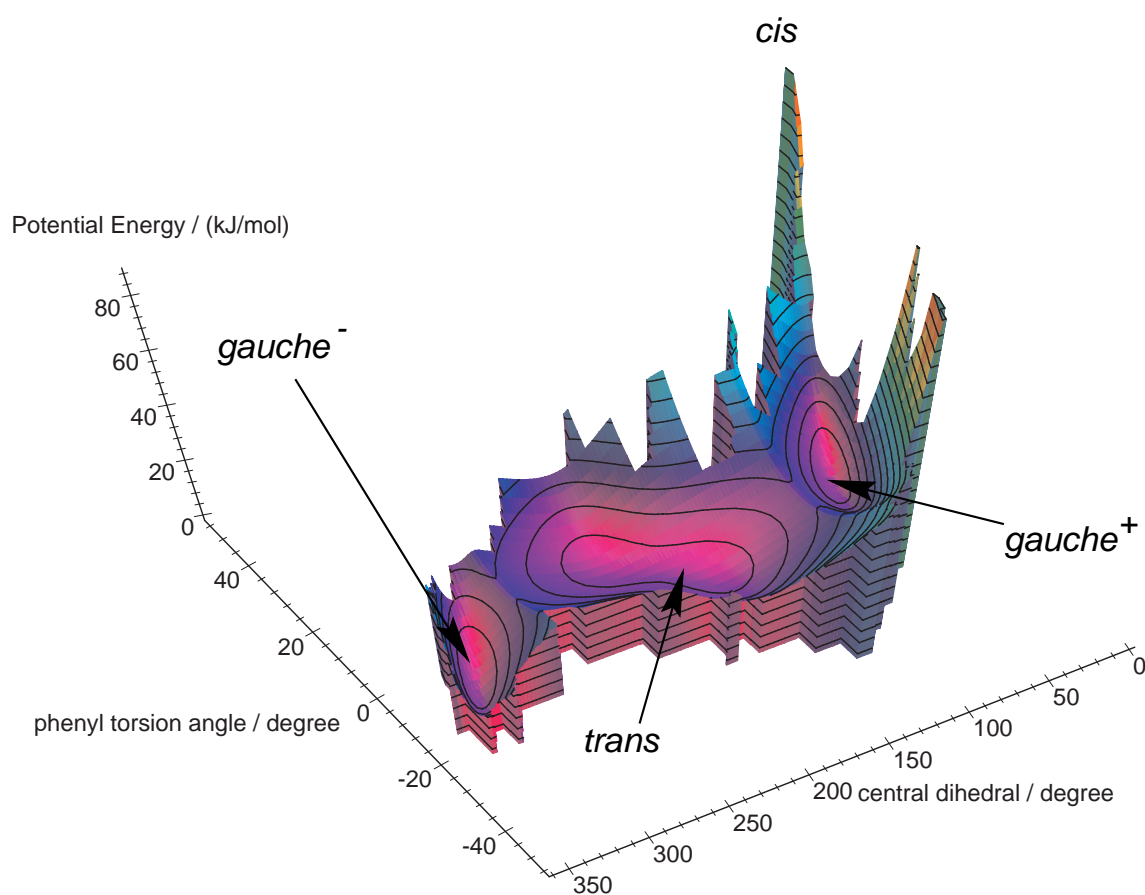


Figure 4.1: 3D view of the potential energy surface of the first excited state of stilbene. The picture shows a cut of the surface with both the phenyl torsion angles constrained to the same value. Only the regions that have been visited during the simulation series are known and displayed. The distance between the contour lines is 5 kJ/mol. The high peak in the back of the figure is the *cis* Franck-Condon region from where the simulations are started. Clearly visible are the extended shallow minimum in the *trans* region, the two *gauche* minima, and the barriers in between.

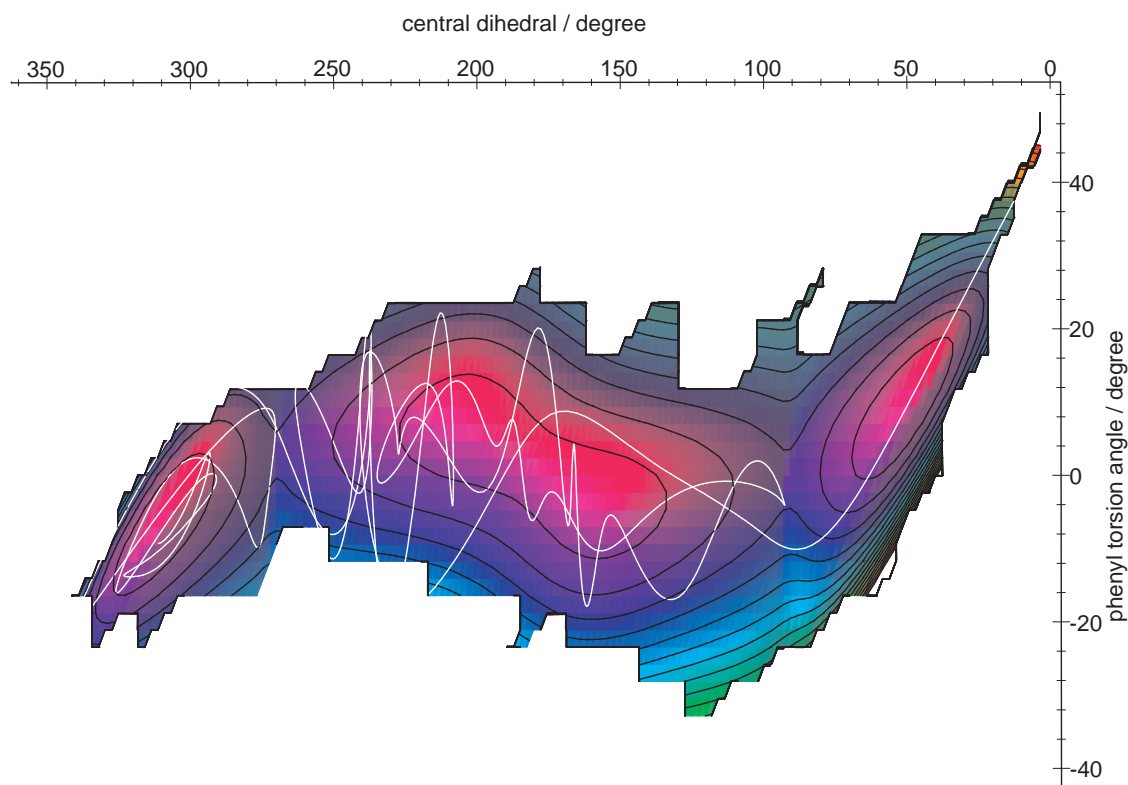


Figure 4.2: Top view of the potential energy surface of Figure 4.1. The distance between contour lines is 5 kJ/mol. A sample trajectory is drawn as a white line. It passes both barriers and ends up in the other *gauche* minimum. Note that the trajectory does not actually leave the region of the known part of the surface. This is an artifact of the representation. While the displayed surface is obtained by a symmetric cut through the real three-dimensional potential energy surface, the displayed trajectory is a projection of the real three-dimensional trajectory, i. e. one of the phenyl torsion angles is neglected.

downhill from the starting point is very narrow, which again confirms the very low variation of the reaction trajectory in the very first phase. Also, there are steep walls on the opposite side of the *gauche* minimum. This indicates that the system is able to climb high after the rapid initial downhill motion.

The potential energy landscape looks quite different compared to previous work. While the qualitative shape is similar to most suggestions, as for example by Abrash et al. [13], Repinec et al. [65], and Saltiel [66, 67], minimum and barrier are at different dihedral angles. The minimum of the excited state is near a *gauche* conformation, while in most of the previous pictures it was assumed to be at the 90° *perp* conformation. The implications on the relaxation to the ground state would be drastic, as it is generally assumed that the relaxation occurs onto the top of the barrier separating *cis*- and *trans*-stilbene in the ground state.

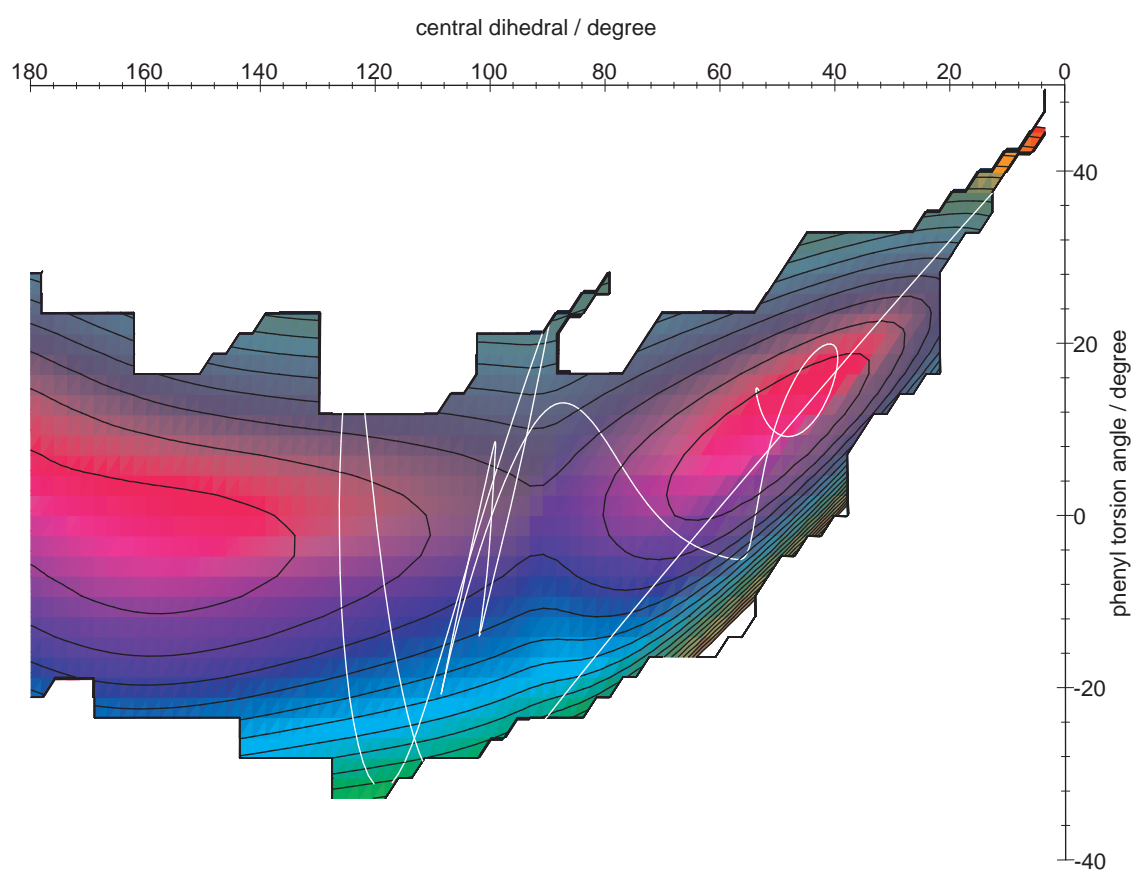


Figure 4.3: Top view of the potential energy surface of Figure 4.1. A sample trajectory is drawn as a white line. It goes actually over the barrier, but does not reach the trans minimum. It is reflected to the gauche region. See also legend for Figure 4.2.

4.4.2 Dependence on Temperature and Pressure

For the investigation of the dependence on temperature and pressure, series of simulations at different state points were performed. Each series consisted of twenty individual trajectories of 5 ps, differing in the initial configuration of the solvent. For two selected pressures, three different series of twenty trajectories each were performed to obtain more data.

The trajectories of each state point were classified into categories depending upon the behaviour of the molecular system they represent: (i) the barrier is not reached, but the system is quenched to remain in the *gauche* conformation, (ii) the system crosses the barrier (the central dihedral reaches at least 92°), but does not reach the *trans* minimum, it recrosses the barrier back to the *gauche* minimum, (iii) the system isomerises and relaxes to the *trans* conformation. The number of trajectories that belong to a certain category are given in Table 4.2 and are labeled S, R and I respectively. The boxes are labeled with numbers, the ascending integer part indicates increasing pressure. zero for simulations in vacuum will be used later. Where meaningful, a digit after the decimal point specifies the number of the simulation series. The pressures in the simulation boxes are summarised in Table 4.3. Table 4.2 allows the following conclusions. There is a strong pressure dependence in the ratio of isomerisations. Virtually the whole range from no isomerisations to all trajectories exhibiting isomerisations is encountered. In contrast, the temperature dependence is much less pronounced. For the box labeled 5.x, there is no a clear trend in the temperature dependence, while at lower pressure (box 2.x), an increase in the ratio of isomerisations is observed with increasing temperature.

If the barrier crossing events are summarized according to the time window they occur, the picture becomes clearer. Let us call transitions that occur within the first 200 fs of the simulation kinetic activations (labeled K), the ones that appear later than 500 fs are thermally activated (labeled T). Intermediate events are labeled with M, see Table 4.4. Again a strong pressure dependence is exhibited. More transitions are observed at low pressure. The temperature dependence

Box type	190 K			237 K			290 K			348 K		
	S	R	I	S	R	I	S	R	I	S	R	I
1.0	2	0	18									
2.0	6	2	12	7	0	13	4	2	14	1	2	17
2.1	7	0	13	5	1	14	5	0	15	2	0	18
2.2	3	2	15	3	1	16	7	0	13	2	1	17
3.0	5	1	14	7	2	11	7	0	13	5	1	14
4.0	10	3	7	12	3	5	7	2	11	9	1	10
5.0	14	2	4	17	1	2	13	4	3	11	3	6
5.1	14	2	4	15	2	3	17	1	2	12	3	5
5.2	15	2	3	15	2	3	14	0	6	13	3	4
6.0	20	0	0	17	2	1	20	0	0	16	1	3
7.0	20	0	0									
8.0	20	0	0							19	1	0

Table 4.2: Classification of the trajectories of state points at four temperatures and eight box sizes for different pressures according to the behaviour of the reactions. S: stayed gauche R: recrossed after a barrier crossing, I: isomerised to trans. The box types are coded as follows. The first digit indicates the pressure: 1 is very low pressure, 8 very high pressure. The digit after the decimal indicates the serial number of a series of simulations of the same box.

Box type	pressure at 190 K / bar	solution box length / nm	solvent only simulation		
			box length / nm	density / g cm ⁻³	reduced density
1	89.4	8.415			
2	257.5	6.235	6.38	0.745	0.45
3	408.9	5.815	5.92	0.932	0.56
4	1064.0	5.335	5.44	1.202	0.72
5	2432.7	5.035	5.13	1.433	0.86
6	3874.1	4.885	4.98	1.566	0.94
7	6343.1	4.735	4.72	1.840	1.10
8	10688.8	4.585			

Table 4.3: Box sizes and density in the simulations. The density of liquid argon at 87 K and ambient pressure is 1.40 g cm⁻³ (box type 5).

is different for kinetically and thermally activated events. As one expects, more thermally activated transitions are observed at higher temperature, while the kinetic activations do not depend on temperature.

By estimating the activation energy from the ratio of isomerisations (Table 4.2) using Equation 4.4, we obtain a value of $E_A = 4.1$ kJ/mol for the *cis-trans* isomerisation reaction. Although this value is likely to be very inaccurate, it is clearly lower than the barrier in our calculated potential energy surface. The energy difference from the *gauche* minimum to the *perp* saddlepoint is 14.2 kJ/mol (Table 4.1). However, as will be discussed later, most of the trajectories will cross the barrier at a higher potential energy, and not at the saddlepoint. This finding would suggest that the activation energy is even higher. As this is clearly not the case, we conclude that the assumption of a thermally activated barrier crossing process is not valid. Obviously, the process is dominated by kinetic activation from the initial motion downwards from the Franck-Condon ex-

Box type	190 K			237 K			290 K			348 K		
	K	M	T	K	M	T	K	M	T	K	M	T
1.0	18	0	1									
2.0	14	0	0	13	0	0	16	0	2	18	0	8
2.1	13	0	0	15	0	2	14	0	3	14	0	14
2.2	17	0	0	16	1	0	11	0	2	17	0	5
3.0	14	0	1	13	0	0	13	0	3	12	1	4
4.0	10	0	0	8	0	1	12	0	1	6	1	6
5.0	6	0	0	3	1	1	7	0	0	7	1	3
5.1	6	0	0	4	0	2	2	0	1	4	3	3
5.2	5	0	0	5	0	0	4	1	2	4	0	3
6.0	0	0	0	2	0	2	0	0	0	2	0	2
7.0	0	0	0									
8.0	0	0	0							0	0	1

Table 4.4: Number of barrier encounter events (the central dihedral angle reaches 92°). Kinetically activated events are shown in columns K, thermally activated events in columns T. Events that occur before 200 fs are considered to be kinetically activated, after 500 fs they contribute to columns T. The intermediate events are listed in columns M. Events may occur more than once in a single trajectory.

citation region. In fact, many trajectories traverse the *gauche* minimum in a straight line without relaxing (Figure 4.2). This explanation is in line with experimental results [14], which exhibit no temperature dependence of the reaction rate constants.

4.4.3 Viscosity Dependence

The shear viscosity of the solvent has been computed using Equations 4.5, 4.6 and 4.7. Results are listed in Table 4.5 and displayed in Figure 4.4. Errors in the average value for the diffusion coefficient, the pressure and the temperature were obtained using the procedure described by Allen and Tildesley [26]. The error of the viscosity was estimated from the spread of the

Box type	Solvent Properties					
	ratio %	k	Viscosity	Diffusion	Pressure	Temperature
190 K						
2	67	26.2±8.9	0.0048±0.0004	0.0235±0.0008	257±6	192.73±0.05
3	70	28.0±9.4	0.0079±0.0001	0.0160±0.0005	418±8	190.04±0.05
4	35	10.0±3.9	0.0150±0.0008	0.0089±0.0002	1042±15	187.64±0.03
5	18	4.7±3.0	0.0273±0.0010	0.0049±0.0001	2449±19	189.25±0.04
6	0	0.0±2.4	0.0372±0.0014	0.0031±0.0000	3834±20	187.57±0.05
7	0	0.0±2.4	0.0239±0.0009	0.0000±0.0001	6432±22	190.73±0.04
237 K						
2	72	29.9±11.1	0.0050±0.0001	0.0268±0.0008	414±5	236.16±0.04
3	55	18.6±5.9	0.0080±0.0002	0.0190±0.0005	692±9	237.10±0.04
4	25	6.7±3.3	0.0155±0.0003	0.0111±0.0003	1570±16	237.39±0.04
5	13	3.3±2.8	0.0272±0.0007	0.0065±0.0001	3222±36	235.70±0.12
6	5	1.2±2.5	0.0376±0.0006	0.0044±0.0001	4885±38	237.66±0.08
7			0.0261±0.0004	0.0000±0.0001	7755±44	237.33±0.06
290 K						
2	70	28.3±9.7	0.0055±0.0002	0.0318±0.0010	607±6	290.23±0.11
3	65	24.2±7.8	0.0083±0.0002	0.0218±0.0005	991±11	290.59±0.04
4	55	18.5±5.8	0.0153±0.0002	0.0131±0.0003	2101±34	290.40±0.08
5	18	4.8±3.0	0.0255±0.0007	0.0081±0.0001	4052±31	289.42±0.08
6	0	0.0±2.4	0.0345±0.0007	0.0058±0.0001	5884±57	289.31±0.13
7			0.0305±0.0012	0.0000±0.0001	9345±199	291.36±0.28
348 K						
2	87	47.5±22.5	0.0055±0.0001	0.0362±0.0011	821±7	350.06±0.13
3	70	28.0±9.4	0.0084±0.0002	0.0248±0.0018	1305±15	348.43±0.07
4	50	16.0±5.1	0.0156±0.0004	0.0157±0.0003	2653±53	348.50±0.28
5	25	6.6±3.3	0.0254±0.0002	0.0100±0.0002	4888±60	346.97±0.08
6	15	3.7±2.8	0.0354±0.0005	0.0073±0.0002	7001±73	351.09±0.18
7			0.0684±0.0012	0.0034±0.0001	13781±84	349.00±0.13

Table 4.5: Reaction rate constants k and some solvent properties. Ratio of isomerisations in percent; rate constant k (in ps^{-1}) estimated using Equation 4.11 and a 50° threshold angle. Viscosity (in $\text{cP} = 10^{-3} \text{ Pa s}$), diffusion coefficient (in nm^2/ps), pressure (in bar) and temperature (in K) from simulations of the solvent only.

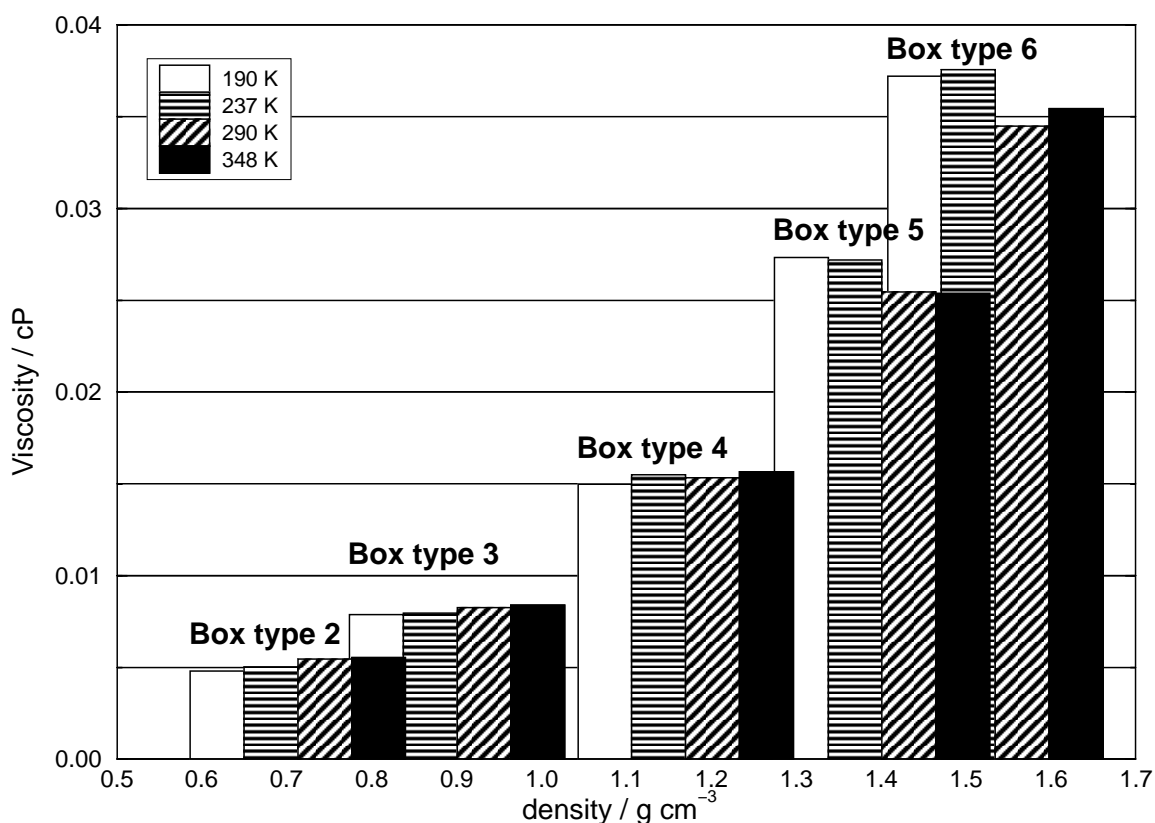


Figure 4.4: Shear viscosity as a function of temperature and box type (density). For the exact density of each box type see Table 4.3. The viscosity is virtually independent of temperature. On the other hand, the viscosity increases significantly with increasing density, i. e. pressure. The density of liquid argon at 87 K and ambient pressure is 1.40 g cm^{-3} (box type 5).

independent results from the off-diagonal elements of the pressure tensor (Equation 4.8). The calculation of the errors in the rate constant is described below for Table 4.6. Box parameters for the solvent-only simulations are listed in Table 4.3. The boxes are large enough to expect that box-size effects are absent [68], and thus the obtained viscosities really correspond to macroscopic shear viscosities.

Figure 4.4 shows that the viscosity depends strongly on the density and thus the pressure. This effect is exploited in experiment to change the viscosity of the solvent without changing the solvent itself. On the other hand, the viscosity is basically independent of temperature. This is what is expected theoretically for a liquid at constant density.

Results for the rate constant k (in ps^{-1}) calculated using Equation 4.11 for a wide range of threshold angles are given in Table 4.6. From the results of repeated series of simulations (Table 4.2), one can estimate that the error in the number of isomerisations is about two per series. From this information, error limits were derived and listed in Table 4.6. Similarly, results for the parameter A (Equation 4.13) are listed in Table 4.7. Experimental values for apolar solvents [14] at 295 K are: *n*-nonane 0.36 cP/ps, *n*-octane 0.32 cP/ps, *n*-hexane 0.23 cP/ps, *n*-pentane 0.16 cP/ps. Results within this range are written in bold in Table 4.7. Experimental and computational results match quite well. A threshold angle in the range from 40° to 60° is suggested by the results. This coincides with the *gauche* minimum, but there need not be any causal relationship.

Box type	Threshold angle / degree							
	20	30	40	50	60	70	80	90
190 K								
1	128±38	92±28	69±21	54±16	43±13	34±10	28.8±8.7	23.5±7.1
2	62±21	45±15	34±11	26.2±8.9	20.8±7.0	16.7±5.7	13.3±4.5	11.0±3.7
3	67±22	48±16	36±12	28.0±9.4	22.1±7.4	17.6±5.9	14.0±4.7	11.0±3.7
4	24.0±9.3	17.0±6.6	12.8±5.0	10.0±3.9	7.8±3.0	6.2±2.4	5.1±2.0	4.1±1.6
5	11.3±7.3	8.0±5.1	6.0±3.9	4.7±3.0	3.7±2.3	2.9±1.9	2.3±1.4	1.8±1.1
6	0.0±5.9	0.0±4.2	0.0±3.1	0.0±2.4	0.0±1.9	0.0±1.5	0.0±1.1	0.0±0.8
237 K								
2	71±27	51±19	38±14	30±11	23.8±8.9	19.0±7.1	15.0±5.6	12.2±4.5
3	44±14	31.8±10.0	23.8±7.5	18.6±5.9	14.6±4.6	11.8±3.7	9.2±2.9	7.3±2.3
4	16.0±8.0	11.4±5.7	8.6±4.3	6.7±3.3	5.2±2.6	4.1±2.0	3.3±1.6	2.8±1.4
5	8.0±6.9	5.7±4.9	4.2±3.6	3.3±2.8	2.6±2.2	2.0±1.7	1.6±1.4	1.3±1.1
6	2.9±6.2	2.0±4.4	1.5±3.3	1.2±2.5	0.9±2.0	0.7±1.5	0.6±1.2	0.4±0.9
290 K								
2	67±23	48±17	36±12	28.3±9.7	22.4±7.7	18.0±6.2	14.3±4.9	11.6±4.0
3	58±19	41±13	31.2±10.0	24.2±7.8	19.2±6.1	15.2±4.9	12.6±4.0	10.0±3.2
4	44±14	31.8±10.0	23.9±7.5	18.5±5.8	14.5±4.6	11.4±3.6	9.7±3.0	7.6±2.4
5	11.6±7.4	8.2±5.2	6.2±3.9	4.8±3.0	3.8±2.4	2.9±1.8	2.3±1.5	1.9±1.2
6	0.0±5.8	0.0±4.2	0.0±3.1	0.0±2.4	0.0±1.8	0.0±1.4	0.0±1.1	0.0±0.0
348 K								
2	113±54	81±38	61±29	47±22	38±18	30±14	25±12	19.6±9.3
3	67±22	48±16	36±12	28.0±9.4	22.2±7.5	17.5±5.9	13.9±4.7	11.8±4.0
4	39±12	27.3±8.8	20.6±6.6	16.0±5.1	12.4±4.0	9.6±3.1	7.8±2.5	5.9±1.9
5	16.2±8.0	11.4±5.6	8.6±4.2	6.6±3.3	5.1±2.5	3.9±2.0	3.3±1.7	2.6±1.3
6	9.0±7.0	6.3±4.9	4.8±3.7	3.7±2.8	2.9±2.3	2.2±1.7	1.9±1.4	1.5±1.1

Table 4.6: Rate constants k , in ps^{-1} , for several threshold angles, obtained from a series of simulations at the usual temperatures and box types. Estimated using Equation 4.11

Table 4.5 shows several macroscopic properties of the solvent at the usual state points. Some of the solvent properties correlate appreciably with the rate constant. These properties are shown in Figure 4.5. The correlation with the self diffusion coefficient is slightly better than with the inverse viscosity. This finding suggests that the former is a better measure for the reaction rate constant. This implies that the motion of the phenyl rings is more like a particle escaping from its solvation cage than displacing a continuous medium. Looking at the Lennard-Jones sizes of the moving particles, this explanation is plausible. The diameter of an argon atom is 0.34 nm, while the diameter of a phenyl ring is 0.75 nm, and its thickness is 0.24 nm. So the sizes of the moving particles are very similar. A similar conclusion has been reached for the self-diffusion of water [69]. However, the quality of the correlation coefficient is disputable. The next paragraph demonstrates that it is quite sensitive to small changes in the way the reaction rate is calculated.

The rate constant as calculated up to here comprises both kinetically and thermally activated barrier crossings. This procedure seems legitimate, as both pathways are likely to be observed in experiment. However, two objections may be raised. Firstly, the number of thermal activations

Box type	parameter $A = \eta k$ [cP/ps]								
	Threshold angle / degrees								
	20	30	40	50	60	70	80	90	
190 K									
2	0.370	0.175	0.132	0.103	0.082	0.066	0.051	0.043	0.103
3	0.526	0.374	0.283	0.220	0.174	0.139	0.110	0.087	0.168
4	0.359	0.254	0.191	0.149	0.118	0.092	0.076	0.062	0.149
5	0.247	0.175	0.132	0.101	0.079	0.063	0.048	0.038	0.101
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
237 K									
2	0.449	0.209	0.157	0.122	0.097	0.076	0.059	0.050	0.121
3	0.352	0.253	0.190	0.148	0.116	0.094	0.074	0.058	0.148
4	0.248	0.177	0.134	0.103	0.081	0.063	0.051	0.043	0.103
5	0.159	0.114	0.085	0.066	0.051	0.039	0.031	0.024	0.032
6	0.107	0.076	0.057	0.044	0.035	0.027	0.021	0.015	0.000
290 K									
2	0.364	0.228	0.171	0.133	0.105	0.083	0.066	0.055	0.101
3	0.479	0.342	0.258	0.200	0.159	0.126	0.104	0.082	0.200
4	0.680	0.487	0.366	0.284	0.223	0.175	0.148	0.116	0.247
5	0.150	0.107	0.080	0.062	0.049	0.037	0.030	0.024	0.062
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
348 K									
2	0.582	0.416	0.383	0.244	0.194	0.154	0.127	0.100	0.156
3	0.562	0.403	0.302	0.235	0.187	0.147	0.117	0.099	0.179
4	0.606	0.427	0.323	0.251	0.194	0.150	0.123	0.092	0.129
5	0.409	0.222	0.167	0.129	0.099	0.075	0.061	0.046	0.029
6	0.320	0.224	0.172	0.130	0.104	0.079	0.066	0.052	0.041

Table 4.7: *A parameters for several threshold angles, obtained from a series of simulations at the usual state points and using Equation 4.13. Bold: values within the range of experimental results (0.16 and 0.36 cP/ps). The rightmost column contains results when only the kinetically activated events are used for the calculation of the rate constant. See text.*

is determined by the lifetime of the excited state in experiment. In our study, it depends on the simulated time span, as deactivations are not considered. Secondly, the number of thermal activations depend on the temperature. When applying Equation 4.12 in analysing experimental results, these events may be erroneously attributed to the formation of DHP, so will not enter the parameter A . If this is the case, it would make sense to consider only the kinetic activations to calculate the parameter A_{kin} . Such results are given in Table 4.7 in the rightmost column and in Figure 4.5 in the lower panels. The results do not change much, as the thermal activations are rather rare. The impact on the regression in Figure 4.5 is more pronounced: The correlation coefficient with the inverse viscosity reaches the same level as with the diffusion coefficient. However, it is uncertain which type of calculation better matches the experiment.

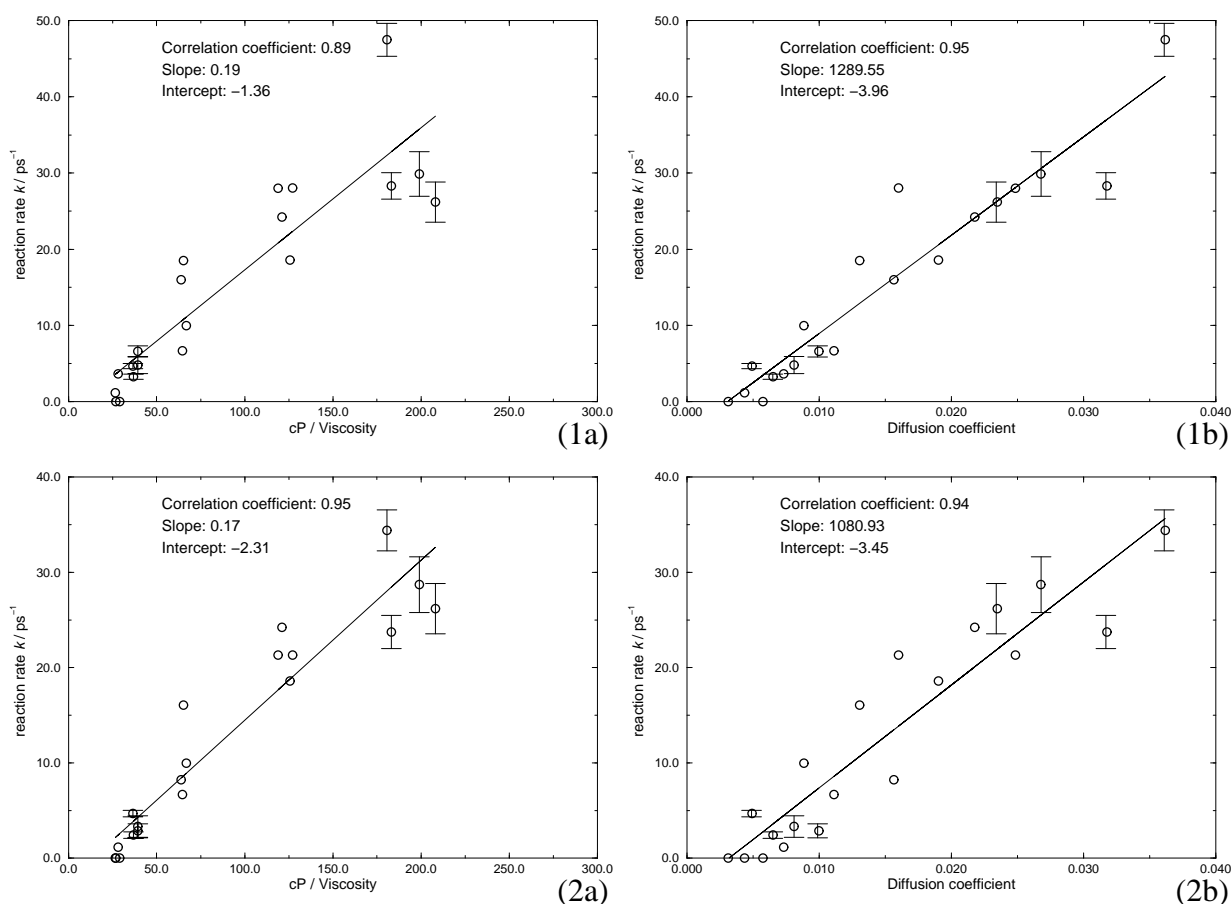


Figure 4.5: Linear regression of the inverse shear viscosity and the diffusion coefficient of the solvent with the reaction rate constant k . Error bars are shown for the points with estimated standard deviation from the three series of simulations (box types 2 and 5). (1) Normal rate constant, (2) Rate constant including kinetic activations only (see text), (a) Inverse viscosity, (b) Diffusion coefficient in nm^2/ps . Calculated using a 50° threshold angle. Results from the linear regression are shown as insets.

4.4.4 Average Trajectories

Figure 4.6 shows trajectory averages only of isomerisations to *trans* for each state point. There is no difference between the trajectories in the first 60 fs of the simulation (Phase A). After that, a pressure dependent behaviour is observed. However, the deviations are still minor in phase B, which lasts up to 120 fs. Then the motion of the molecule is more strongly quenched, strictly with increasing pressure (phase C). This finding is clear evidence for a pressure-dependent solvent friction which damps the molecule's motion more effectively with increasing pressure. In Phase D there is a clear motion towards the *trans* minimum.

The initial motion of the dihedral angles is very rapid. After about 40 fs, the *gauche* minimum is reached. This is in good agreement with Myers and Mathies [70], which concluded from resonance Raman experiments a dihedral angle change of 25° of the central ethylenic bond in 20 fs only. In our simulation, this dihedral angle change is reached after 25 fs.

The trajectories in phase C of Figure 4.6 all show an interesting feature. The slope of the trajectory ($\partial r_2/\partial t$, "speed of reaction") decreases after the barrier (at 92°) has been crossed.

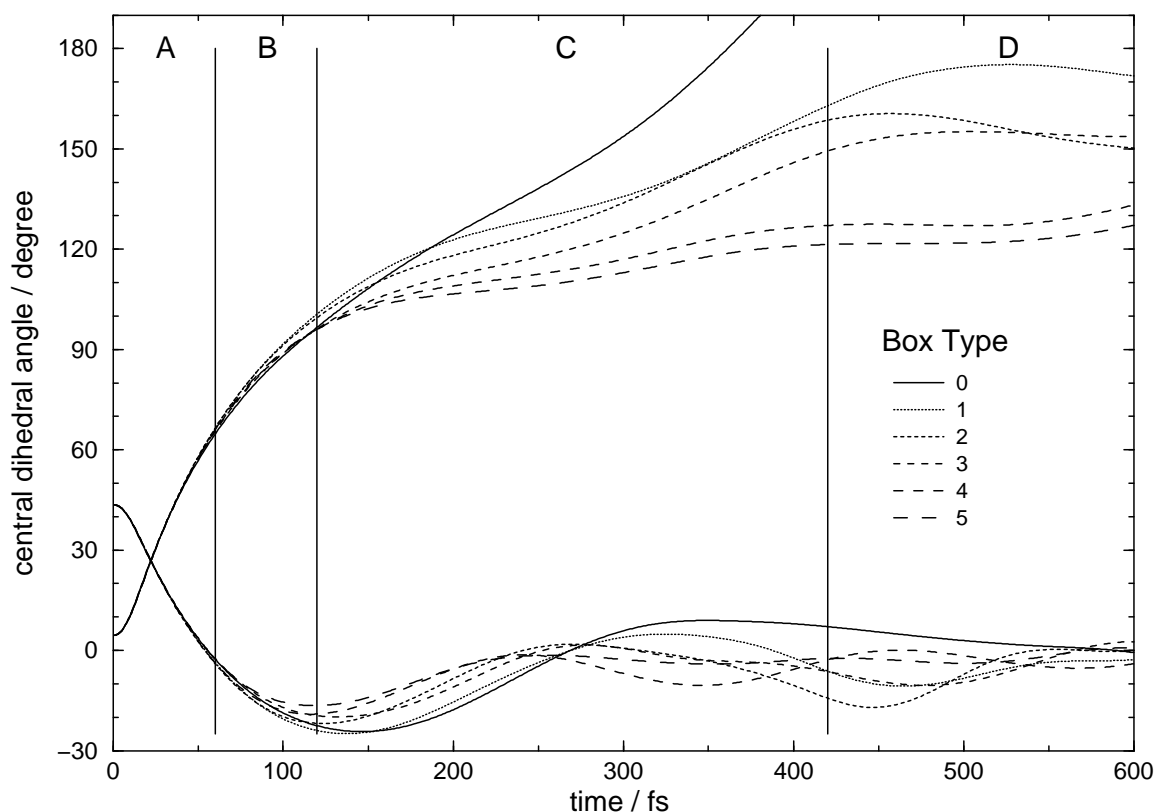


Figure 4.6: Averaged trajectories of simulations at 190 K that exhibit isomerisations to the trans minimum (class I from Table 4.2). Box types (Table 4.3) or increasing pressure is indicated by different line styles and numbers.

This is in contrast to the expectation that the molecule relaxes quickly to the minimum once the barrier has been crossed. There are two reasons for the observed behaviour: (i) the phenyl rings need to rearrange before the central dihedral is able to relax, (ii) solvent friction is particularly effective in this region. The solvent effects can be investigated by comparing to the simulation of the isomerisation in vacuo. The vacuum simulation is probably a poor representation of the gas phase reaction, in which internal vibrational energy redistribution (IVR) is likely to be an important relaxation pathway, but not allowed in the simulation. However, it is consistent with the simulations of the system in solution, in which the interaction with the solvent is assumed to be the major source of relaxation. Figure 4.7 shows the dihedral angle trajectories of the system in vacuo (dashed lines) in comparison to the system in solution (solid lines). Only the trajectories that exhibited isomerisation were averaged and are shown with standard deviations (thin lines). The trajectories originate from simulations at 190 K in a box of 5.035 nm edge length, which corresponds to the highest pressure of a system in which isomerisations still occurred. The phenyl ring torsion angles are also shown. Trajectories in both vacuum and solvent show that during the flattened phase of the central dihedral angle (between 100 and 250 fs) the motion of the phenyl torsion angles is reversed. Figure 4.9 shows a peak in the solvent-solute interaction potential energy at 250 fs, exactly the time when the dihedral angles in Figure 4.7 do not change much.

The dot-dashed line in Figure 4.7 represents the angle between the plane of the two phenyl rings. This angle may serve as a measure of impact of the reaction on the solvent. One can see that in the early phase of the reaction, up to approximately 100 fs, this angle does virtually not

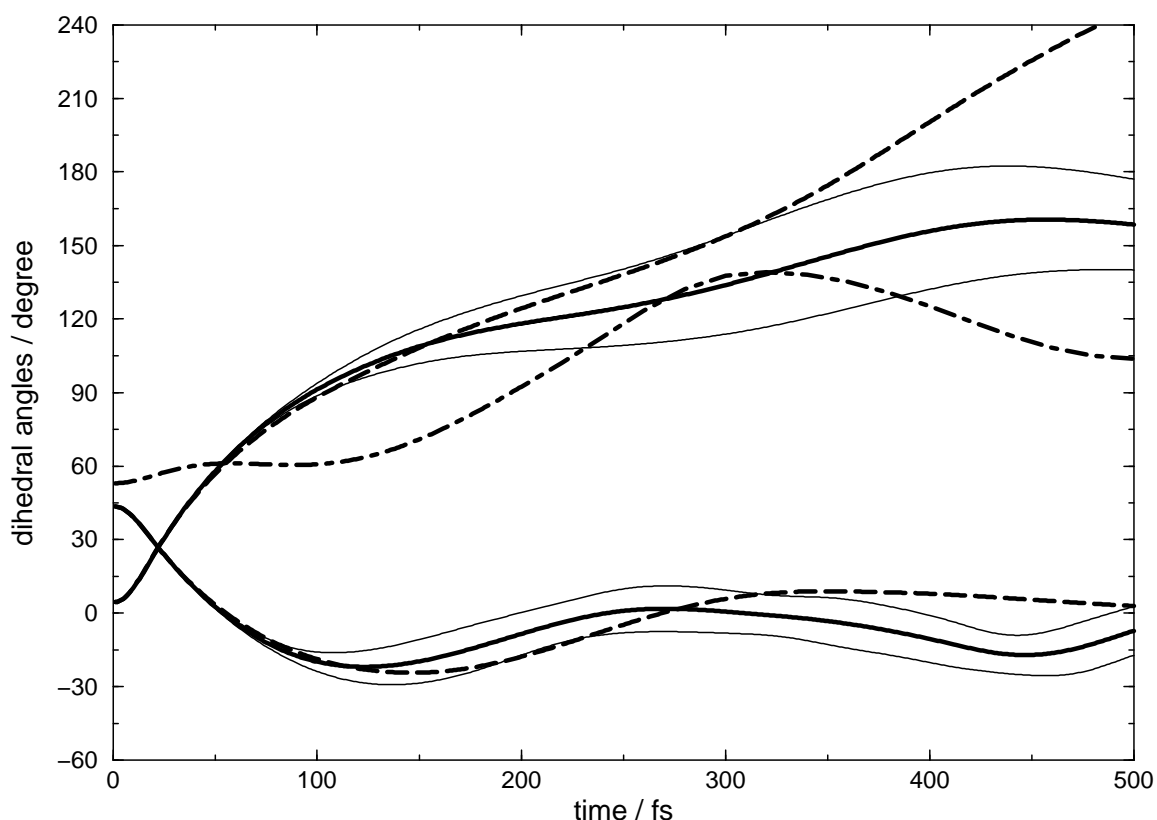


Figure 4.7: Averaged trajectories of simulations at 190 K that exhibit isomerisations to the *trans* minimum with 6.235 nm box size. Solid lines: central and one phenyl torsion dihedral angles, together with standard deviations. Both phenyl torsion angles are similar, but not identical. Dashed lines: Trajectory of a single simulation of the system in vacuo. As the system is simulated from a symmetric initial conformation, the two trajectories of the phenyl torsional angles coincide. Dot-dashed line: angle between the plane normals of the two phenyl rings.

change. Afterwards, there is a substantial change, which is nicely correlated with the flattening of the central dihedral angle trajectory (solid line) after it has crossed the barrier at 92° .

4.4.5 Some Individual Dihedral Angle Trajectories

The initial downhill motion (Figure 4.6 phase A, both central and phenyl dihedrals involved) and the barrier crossing (primarily central dihedral involved) occur in different directions in conformation space. In other words, there is a bend between the line connecting the *cis* peak to the *gauche* minimum and the line connecting the *gauche* minimum to the barrier saddlepoint in Figure 4.3. Thus, a transfer of angular momentum is required for isomerisation, although the potential energy might easily reach a value above the barrier. This observation is confirmed by looking at individual trajectories, e. g. in Figure 4.8. Both trajectories come from the same series of simulation with equal temperature and pressure. However, the solid line shows a trajectory that leads to isomerisation, while the dashed line represents a trajectory that ends up in the *gauche* minimum. Looking at the energy trajectory of the latter, it is evident that the potential energy reaches approximately 45 kJ/mol, which is considerably higher than the barrier. So, from an

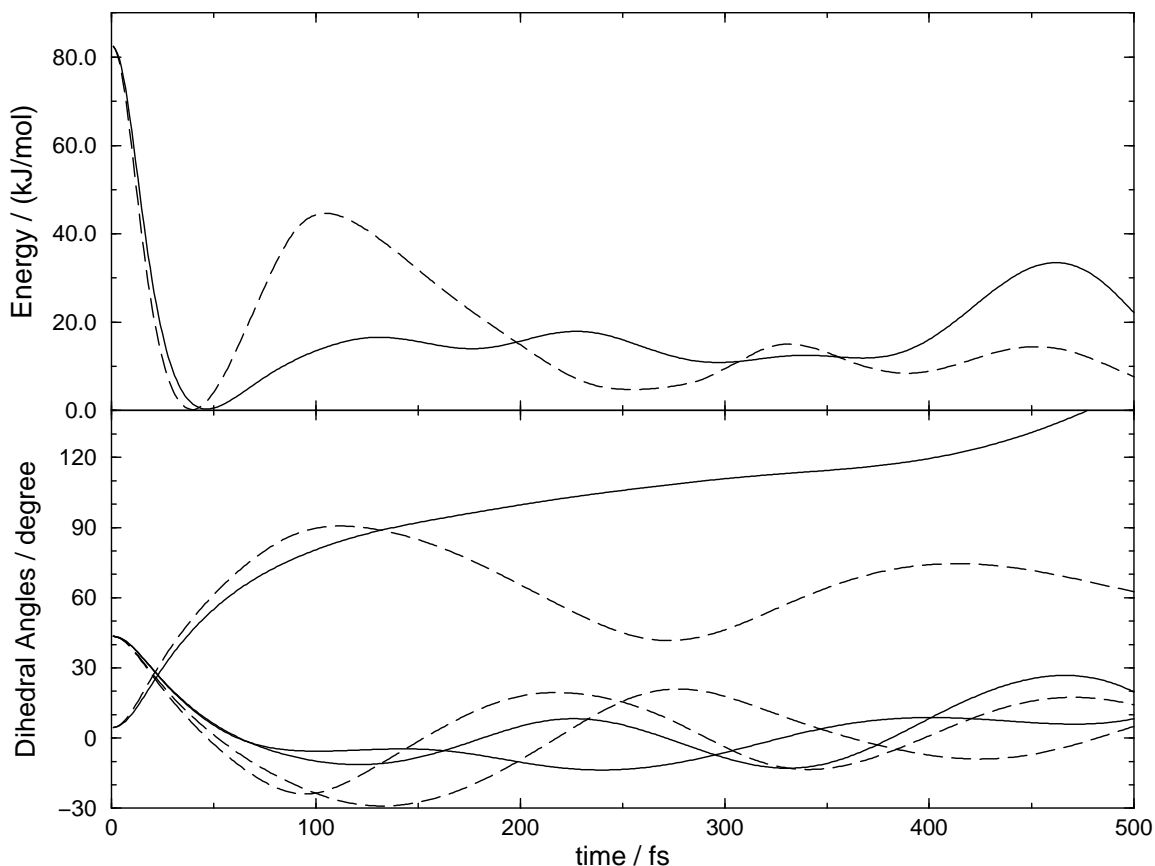


Figure 4.8: Examples of two individual trajectories at 290 K and 5.035 nm box size. The solid line depicts a trajectory that leads to isomerisation to the *trans* region, while the dashed line does not. Upper part: Potential energy trajectories; lower part: dihedral angles trajectories.

energetic point of view, a barrier crossing would easily be possible. It does not take place because the conformation is not in the vicinity of the saddlepoint.

Looking at the other trajectory (solid line), the maximum potential energy (apart from the initial part) is much lower than in the first one. Nevertheless, it is comfortably above the barrier height and exhibits an isomerisation. Interesting is again that the trajectory flattens between 90° and 120° for the central dihedral angle. The major difference between the two trajectories in the first phase is the evolution of the phenyl torsion angles. While these angles are heavily distorted to nearly -30° in the trajectory without isomerisation, they are drastically quenched in the other trajectory and hardly reach -10° . This effect directs the motion of the molecule towards the saddlepoint and over to the *trans* region. A sample trajectory is shown in Figure 4.2 as a white line. The barrier is also clearly visible and is straight on the $r_2 = 92^\circ$ line.

Looking at Figure 4.2, it is not difficult to imagine why the barrier is rarely crossed at its minimum energy point. Falling down from the initial Franck-Condon region in the upper right corner in Figure 4.2, the molecule keeps its reaction direction when climbing the wall on the other side of the *gauche* minimum. By looking carefully at the contour lines when climbing, one realises that the driving force towards the barrier is not very strong, as the contour lines are crossed nearly perpendicularly.

4.4.6 Reason for the Barrier-Recrossings

Several trajectories exhibit barrier recrossings, i. e. the barrier is actually crossed, but the *trans* minimum is not reached, and the molecule is rather pushed back to the *gauche* region, without prior relaxation. This behaviour is shown in Figures 4.3 and 4.8. Beyond the barrier, the potential energy surface does not exhibit any back-driving gradient. Therefore, the force reverting the inertial motion of the molecule must have another source. However, the gradients of the potential energy surface along the central dihedral angle are rather small, as opposed to the gradients along the phenyl torsion angles (mind the scaling of the pictures of the potential energy surface). This is likely to have two consequences: (i) The driving force to either minimum is not very strong. This is illustrated in Figure 4.6. In phase C, after the barrier has been crossed, the central dihedral angle does not move as fast as in phase A, even for the vacuum trajectory. At the same time, the phenyl torsion angles change vividly (Figure 4.3). (ii) The force required to revert the motion of the central dihedral angle does not need to be very large.

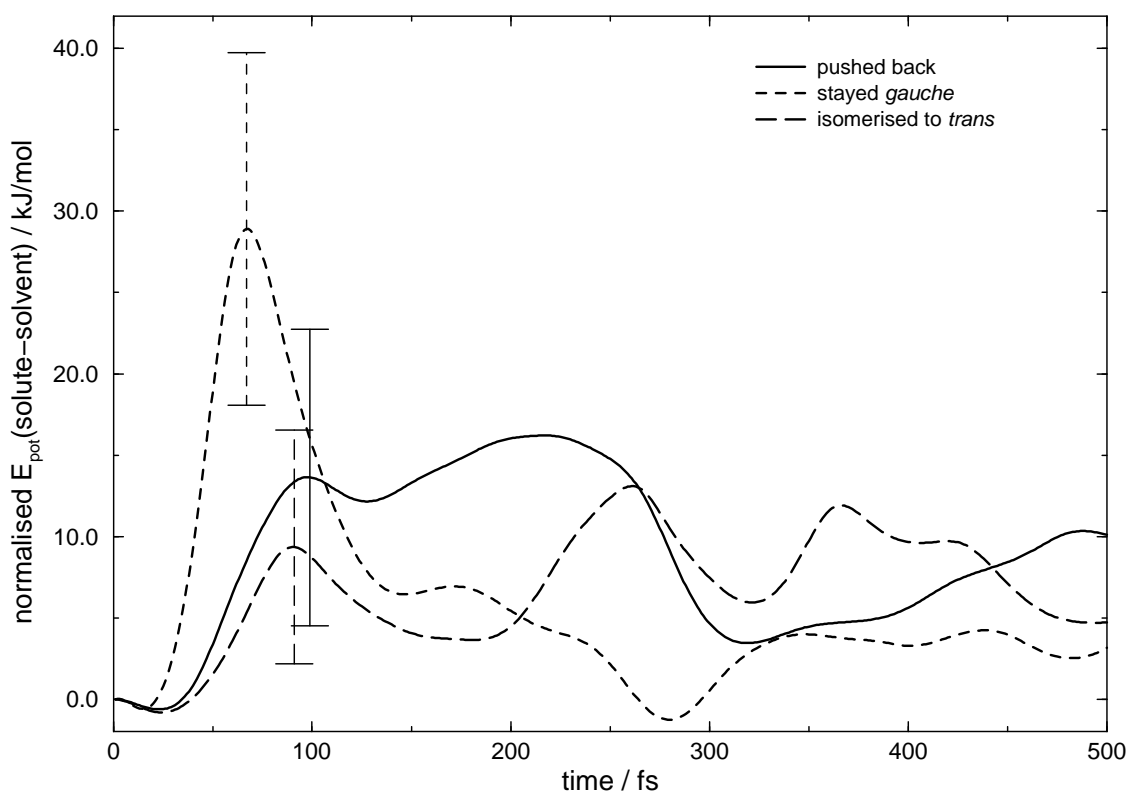


Figure 4.9: Averaged trajectories of the solvent-solute interaction potential energy of three classes of trajectories as defined in Table 4.2. The initial value of each individual trajectory has been subtracted. Average trajectories with example error bars on every first maximum are shown. An arbitrary set of 12–15 trajectories per reaction class has been averaged.

Figure 4.9 shows averaged trajectories of the solute-solvent interaction potential energy for the three classes of reactions from Table 4.2. These classes exhibit qualitatively different behaviour. (i) The reactions which are immediately quenched in the *gauche* minimum (short-dashed line) encounter a high peak of 29 kJ/mol at 70 fs. In this case, the solvent atoms form

a high energy wall which cannot be broken through. Thus isomerisations are not possible. (ii) The isomerisations (long-dashed line) feature a low peak of 9 kJ/mol at 90 fs, then a basin, and a second peak of 13 kJ/mol at 260 fs. The first frictional barrier is overcome, and the first solvation shell relaxes a bit while the isomerisation continues. The second peak is overcome when the *trans* region is reached. (iii) The barrier recrossings (solid line) are characterised by a broad lobe between 14 kJ/mol / 100 fs and 16 kJ/mol / 220 fs. These features are present in the individual trajectories in a more or less pronounced manner, and are not artifacts of the averaging. The error bars on every first maximum demonstrate that the three classes of trajectories are quite well separated.

The three classes in Table 4.2 can also be characterised by the energy fluxes between different types of energy. The first phase is equal for all three classes: the solute's potential energy is transformed into kinetic energy of the solvent. The second phase is different for the three classes. For the trajectories that remain in the *gauche* minimum, the solvent kinetic energy is transformed into potential energy of the solute-solvent interaction (see the high short-dashed peak in Figure 4.9). In the next phase, the energy moves mainly into the solvent. Afterwards, the energy fluxes become less clear.

In the case of a transition, the kinetic energy of the solute is mainly transformed back into intermolecular potential energy in the second phase, i. e. is used to climb the barrier. Only a small fraction flows into solute-solvent interaction potential energy: The long-dashed peak near 100 fs in Figure 4.9 is much smaller than the short-dashed one. After the barrier transition, there is again a peak in the solute-solvent interaction potential energy, which is overcome by slowing down the molecule's motion.

For a recrossing event, the solute's kinetic energy is distributed to all three solvent-internal, solute-solvent, and solute-internal potential energies in the second phase. Because the increase in solvent-internal and solute-solvent potential energies is slow, it is still possible for the molecule to overcome the *perp* barrier. Unlike in the other cases, the solvent-internal and solute-solvent potential energies keep increasing. These high potential energies last over a relatively long time period (see the broad solid lobe between 100 and 250 fs in Figure 4.9) and cause the inversion of the molecule's motion and eventually make it fall back to the *gauche* region.

Figure 4.10 shows the effect of the first solvation shell on the molecule. One can see that the solute-only energy trajectories (thin lines) coincide with the vacuum trajectory within the first 60 fs. For the trajectories that involve barrier transitions, the similarity to the vacuum trajectory lasts up to 170 fs, which is well after the barrier has been crossed. Thus, the trajectories of the solute plus the first solvation shell (thick lines) give an appropriate representation of the solvent effect during the reaction. At first sight, it looks like the trajectories that stay in the *gauche* region have the lowest barrier (thick short-dashed line). This is an artifact of the representation: As these trajectories do not reach the barrier, the energy remains small. It can clearly be seen that the solvent causes an increase of the barrier height, and the barrier is shifted to earlier time. To a lesser extent, the same is true for the other two reaction classes. The solute-only potential energies reach a higher level, because the barrier is indeed crossed in these cases. The solvent effects of the two classes show qualitatively different features. For the isomerisation class (long-dashed lines), the solvent effect causes an increase of the barrier by approximately 7 kJ/mol. In the recrossing class (solid lines), the barrier is increased by twice this amount and also becomes substantially broader.

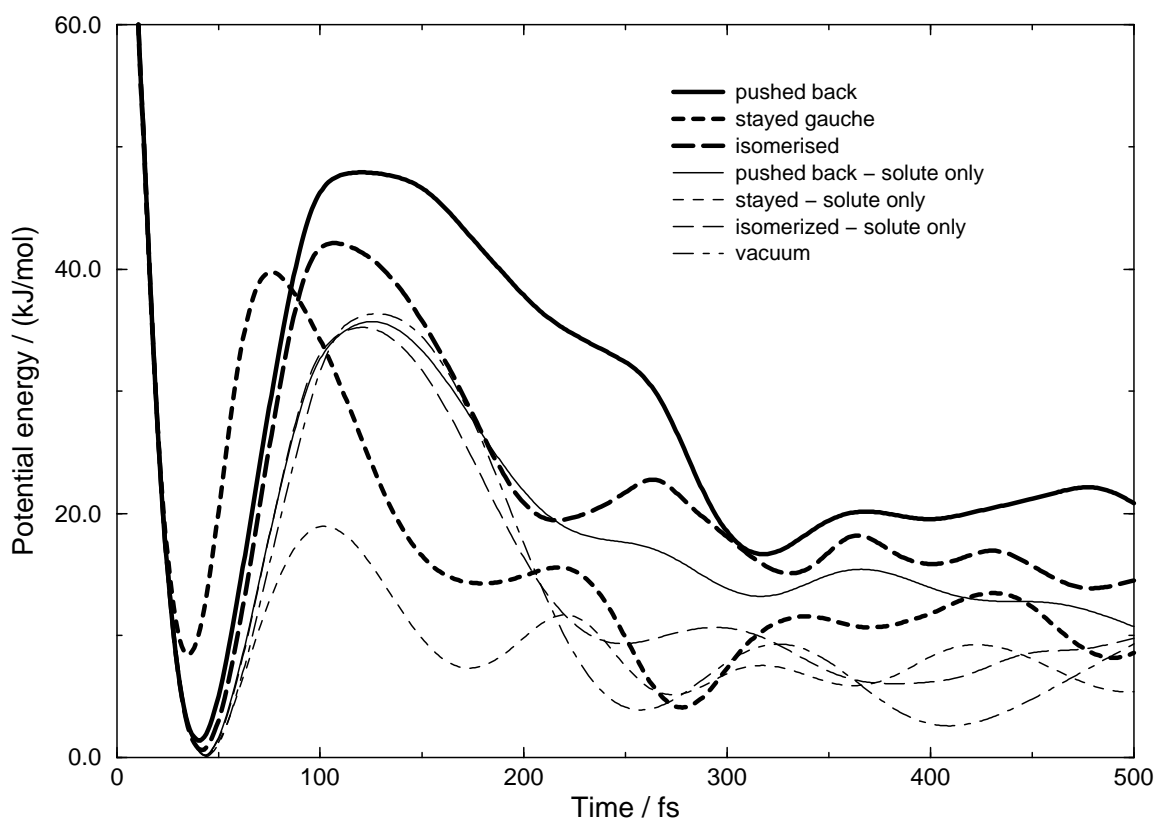


Figure 4.10: Potential energy trajectories of the solute only (thin lines) and together with the first solvation shell (thick lines). The same three reaction classes as in Figure 4.9 are shown.

4.4.7 Behaviour on the Barrier

Figure 4.11 shows at which positions the barrier is crossed. The vast majority of dots lies in the region around -20° for both phenyl ring torsion angles. At the same time, these are nearly exclusively kinetically activated events that occur before 200 fs (circles). The later events, which are thermally activated, scatter around the barrier saddlepoint at 0° for both dihedral angles (pluses and crosses). However, some transitions still occur rather far from the saddlepoint.

Table 4.8 gives averages of the barrier crossing locations and averages according to increasing time window. The vast majority of the barrier encounter events occurs in the two first time windows, before 200 fs. Their average energy is approximately 20 kJ/mol above the saddlepoint, and their location is 20° off the saddlepoint for both phenyl torsion angles. However, the later events are quite close to the saddlepoint on average and their energy is approximately 4.5 kJ/mol above its energy.

4.4.8 Barrier Close-ups

As a side product of the simulations, a detailed potential energy surface of the barrier between the *gauche* and the *trans* minimum was obtained. This barrier plays an important role in the photoisomerisation of *trans*-stilbene [12, 30, 31, 33, 36, 37, 51]. From their experimental studies, Schroeder et al. [33] draw the following conclusions: multi-dimensional barrier effects are important, and the barrier sharpens if another coordinate perpendicular to the reaction coordinate is

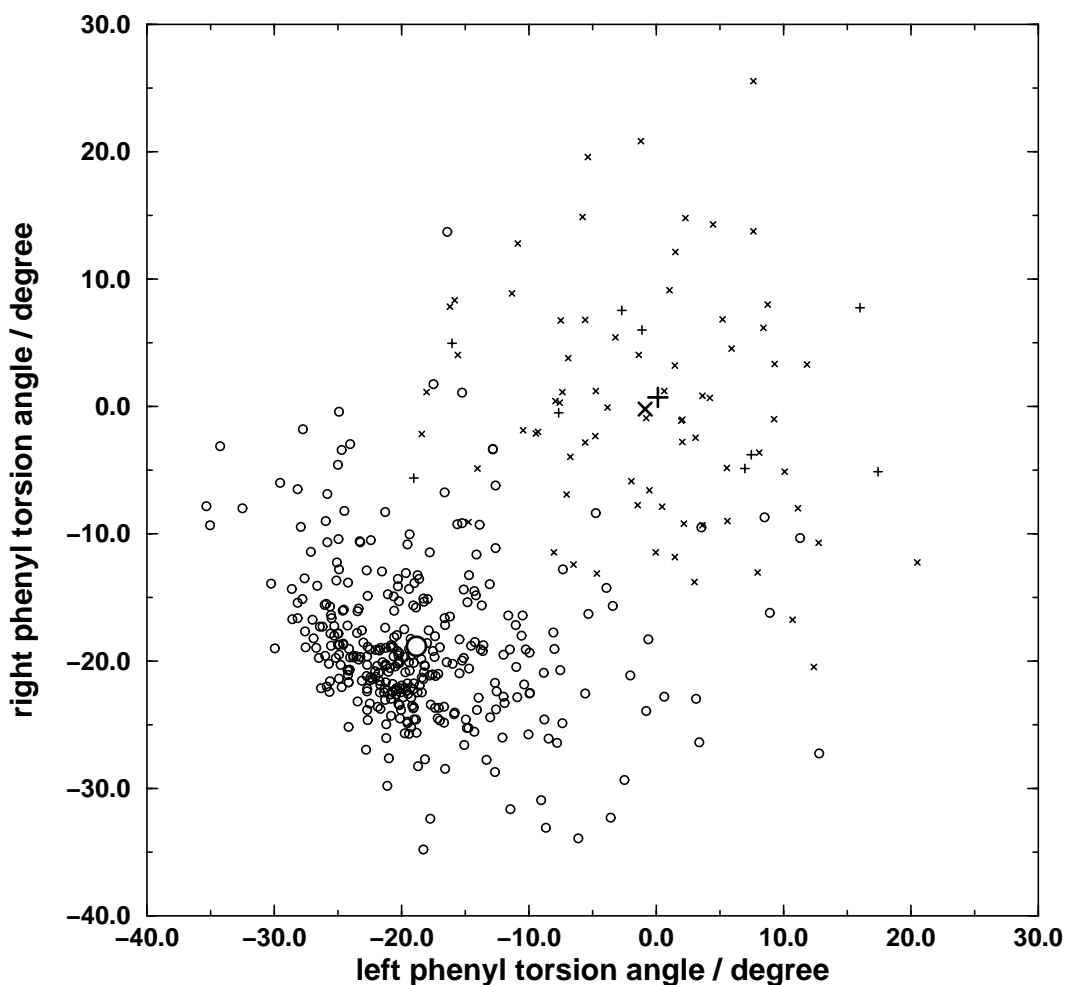


Figure 4.11: Cloud plot of the locations of all barrier crossings. The two axes represent the phenyl torsion angle values when the central dihedral angle crosses the barrier ($r_2 = 92^\circ$). Circles: early crossings (before 200 fs), +: intermediate crossings (between 200 fs and 500 fs), \times : late crossings (after 500 fs). The averages of the three sets are given by larger symbols.

time window / ps	averages				number of events
	left torsion / degree	right torsion / degree	energy / (kJ/mol)	energy above saddlepoint / (kJ/mol)	
0.0-0.1	-20.0	-20.7	34.9	20.7	102
0.1-0.2	-18.3	-18.0	32.9	18.7	243
0.2-0.4	-2.6	1.7	20.2	6.0	5
0.4-0.8	3.5	0.5	18.8	4.6	18
0.8-1.6	-1.4	0.1	18.5	4.3	27
1.6-3.2	-1.7	-1.6	17.9	3.7	16
3.2-6.4	-2.8	-0.3	18.0	3.8	16

Table 4.8: Average barrier crossing locations with corresponding average energies, depending on the time window they occur. The vast majority of the crossings are kinetically activated. They occur before 200 fs and cross the barrier far off and much above the saddlepoint. The thermally activated events pass the barrier close to the saddlepoint on average, but still not at the saddlepoint's energy.

excited. Figure 4.12 (b) shows a picture of this situation: While the barrier is relatively flat in its minimum, the curvature is stronger towards the walls of the barrier. However, in Figure 4.12(a) our calculations show a different picture: The barrier gets flatter towards its walls. This finding suggests that the postulate about the special properties of the barrier is not true. In this case, the discrepancy between the experimentally observed facts and predictions by RRKM theory is not resolved satisfactorily.

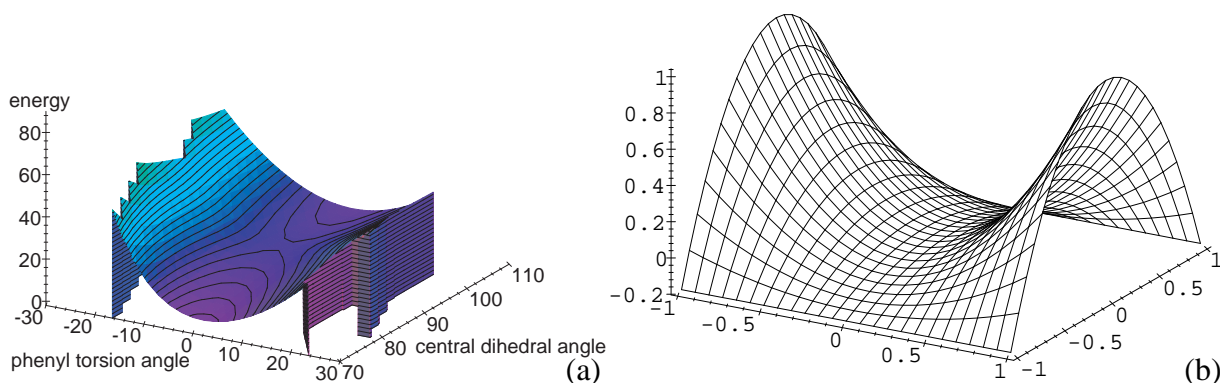


Figure 4.12: Barrier close-up views. (a): Barrier between the gauche and the trans minimum as obtained from our *ab initio* calculations. Distance between the contours: 2 kJ/mol. (b): Barrier shape as suggested by Schroeder et al. [33]. Distance between the contours: 0.1 arbitrary units.

Yet, it is likely that the multi-dimensionality of the barrier plays an important role in the photoisomerisation dynamics. This is certainly the case for *cis-trans* isomerisation, according to our simulated reaction trajectories. In the case of the *trans-cis* isomerisation, the starting point of the reaction is not in a high-energy region, but rather close to the shallow minimum. In this case, all reactions must be thermally activated. Our calculations suggest that even in this case the barrier crossings do not occur straight through the saddlepoint. In other words, modes perpendicular to the reaction coordinate are excited. Thus the multi-dimensional character of the barrier is an important aspect, as suggested by Schroeder et al. [33].

4.5 Conclusions

We have simulated the photoisomerisation of *cis*-stilbene in solution at several temperatures and pressures. The potential energy surface of the stilbene molecules is calculated by *ab initio* quantum chemistry and is represented by a finite elements grid. This representation allows a great reduction of the computational expense of the quantum chemistry. In the whole study, 4 million time steps were performed, and only 2225 explicit quantum chemical calculations were required. This gives an enhancement factor of 1800 compared to a brute force approach.

Although a rather crude model of stilbene and a low-level quantum chemical method was employed, the results are in reasonable agreement with experiment. The correlation between the reaction rate constants and the solvent shear viscosity, quantified by the parameter A in Equation 4.12, is correctly reproduced. However, in experiment the A parameter is independent of temperature and pressure, and the linear correlation is striking. In our studies, there is quite some spread in the values of the A parameters dependent on both pressure and temperature, but no trends are evident. We found that the reaction rate constant correlates with similar accuracy with

the diffusion coefficient of the solvent. This indicates that, for comparison with the reaction rate constant, a microscopic transport property is as suitable as a bulk property like the viscosity.

Since the reaction starts from a very high energy region, it is a highly non-equilibrium process. Most barrier transitions occur in one go after photoexcitation without prior relaxation to a minimum (kinetic activation), so no subsequent thermal activation is necessary. The transition energies are nearly 20 kJ/mol above the barrier saddlepoint. We also observed thermally activated barrier crossings. They average on the barrier saddlepoint, but with a considerable scatter. The picture of a minimum energy path of a reaction is inappropriate, especially for kinetically activated events.

We observed events in which the barrier was crossed, but the motion was reversed. This behaviour could be clearly attributed to a solvent effect: The solvent forms a long-lived dynamic energy barrier.

Many other authors assume a minimum on the potential energy surface of the first excited state at the 90° conformation. Our present study suggests that this state is rather at a *gauche* conformation near a 50° twist angle. This state is reached very quickly, approximately 50 fs after excitation, as suggested by Abrash et al. [13], independent of solvent friction. Similar suggestions were brought up by Myers and Mathies [70]. It is possible that the conformation probed experimentally is indeed the *gauche* conformer. This would explain the lack of spectral evolution after 100 fs. The experimentally observed exponential decay could then have a different origin than a barrier near the *cis* Franck-Condon region. It might be that the process which is experimentally monitored is the barrier crossing or other channels of disappearance from the *gauche* minimum. The former involves frictional solvent effects that are reproduced by our study in respect to the experimental work by Nikowa [14], while the effects important for the latter are ignored in our study.

Chapter 5

Simulation of the β Domain of Metallothionein

5.1 Summary

The β domain of rat liver metallothionein-2 in aqueous solution was simulated with different metal contents. The Cd_3 and the CdZn_2 variant plus the Zn_3 variant were investigated using a conventional molecular dynamics simulation, as well as a simulation with a quantum-chemical description (MNDO/d) of the metal core embedded in a classical environment. The results were compared to the corresponding experimental X-ray crystallographic and NMR solution data. The purely classical simulations were found to produce too compact a metal cluster with partly incorrect geometries, which affected the enfolding protein backbone. The inclusion of quantum chemistry for the treatment of the metal cluster improved the results to give correct cluster geometries and an overall protein structure in agreement with experiment.

5.2 Introduction

Metallothioneins are a class of small proteins with a high content of cysteines. They are capable of binding large amounts of heavy metals such as zinc, cadmium and mercury. Their primary function is believed to be detoxification of heavy metal ions. This function requires a broad but strong affinity for various toxic heavy metal ions. Rat liver metallothionein has 61 residues, of which 20 are cysteines. It binds seven heavy metal ions in two domains which are quite independent. In the β domain, consisting of residues 1–30, three metal ions are coordinated by nine cysteines. The α domain (residues 31–61) binds four metal ions to eleven cysteines. The cysteines are deprotonated and coordinate the metals in a tetrahedral fashion, similar to the structure of zincblende. The structure of rat liver metallothionein-2 has been solved by NMR [71] and X-ray crystallography [72]. A previous X-ray structure [73] was proven to be incorrect [72, 74]. The NMR structure contains seven cadmium ions, whereas the X-ray structure has four cadmium ions in the α domain, and a CdZn_2 composition in the β domain.

Simulators often hesitate to investigate proteins involving heavy metal ions. The reason is the lack of reliable force fields for these metals. The GROMOS force field [19], for example, contains parameters for zinc, but they have never been thoroughly tested. Parameters for cadmium are not generally available. The present work compares the performance of the standard

GROMOS force field, extended by estimated parameters for cadmium, to simulations with a semi-empirical treatment of the metal core embedded in a classical environment. In the case of metallothionein, primarily structural aspects are of interest. However, in many metalloproteins, the metal core displays catalytic activity (e. g. zinc in alcohol dehydrogenase), enables electron transport (e. g. iron and copper in cytochromes), or captures light (e. g. magnesium in conjunction with porphyrin in the photosystems of plants). These properties and processes are certainly not suitable for a purely classical description.

Only the β domain of rat liver metallothionein-2 is considered in the present work. Experimental structures are available for the Cd_3 and the CdZn_2 variants. Both structures are closely similar [74], having the same metal-sulphur cluster geometries and a similar polypeptide fold. A more detailed comparison is given in Section 5.4.1 below.

Figure 5.1 shows a close-up of the metal core of the β domain of metallothionein, containing two zinc ions and one cadmium ion. There are two types of sulphurs: a bridging type which is coordinated to two metal ions (from Cys7, Cys15 and Cys24), and a terminal type which is coordinated to a single metal ion. The metals and the bridging sulphurs form a twisted six-membered ring.

5.3 Methods

In order to simplify notation, let us first introduce some abbreviations. We denote a purely classical molecular dynamics simulation as MDc. A molecular dynamics simulation with a combined quantum-chemical/force-field potential energy function is denoted as MDq.

Three metal center variants were simulated: Cd_3 , CdZn_2 and Zn_3 . All three variants were simulated both fully classically (MDc) and combined with the semi-empirical method MNDO/d [24] (MDq). For the Zn_3 cluster, also MNDO without d-orbital extension was employed. The coupling scheme between the quantum-chemical core and the classical environment is described in detail in Section 1.3.3. The quantum-chemical part involves the three metal ions, the cysteinic sulphur and β carbon atoms, and the attached hydrogen atoms. Thus, in the quantum-chemical core, the cysteines are reduced in size to methylthiolates. For the MDc simulations, force-field parameters for zinc were taken from the GROMOS96 force field [19] and those for cadmium were estimated as described in Section 5.3.2 below.

The structure of the β -domain of rat liver metallothionein-2 (residues 1–30, containing the three-metal cluster) was obtained from the X-ray structure [72] (PDB entry 4MT2). The α domain was chopped off. The resulting structure of the CdZn_2 cluster was used as initial structure for the simulations of all three variants. A separate energy minimisation for each variant and force calculation scheme (MDc or MDq) was carried out prior to any dynamic simulation. The NMR structure [71] of the Cd_3 variant was used for comparison (PDB entry 2MRT).

5.3.1 Computational Details

All four lysines in the β domain were protonated and the two aspartic acids deprotonated. The protein domain was simulated in a periodic box of water with truncated-octahedral shape. A minimum protein-to-wall distance of 1.4 nm was used, giving a total of 2745 water molecules and 8446 atoms. The volume of the box was constant.

For all classical atoms, the GROMOS96 [19] force field 43A1 was employed. Water was modeled using the simple point charge model (SPC, [75]). Classical bonds were constrained

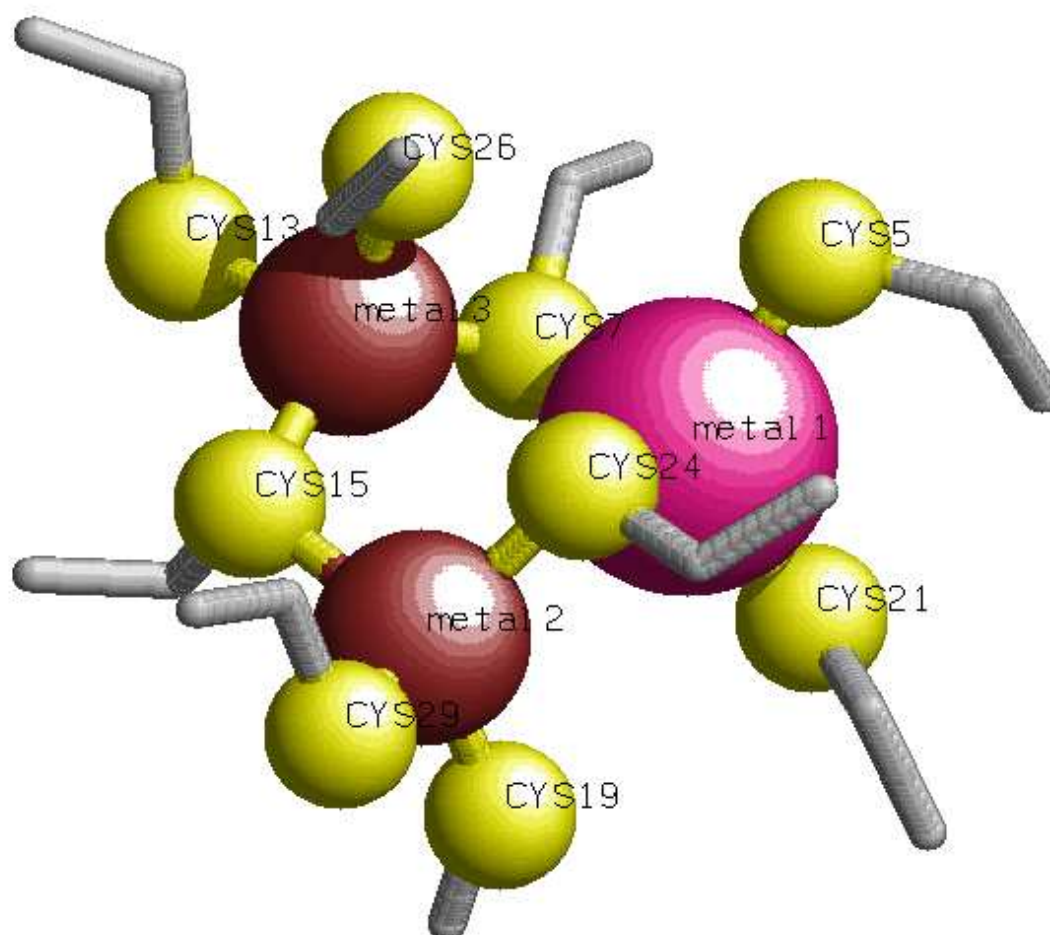


Figure 5.1: Close-up of the metal core in the X-ray structure. The large ball represents the cadmium ion, the two dark balls represent the zinc ions, and the small light balls represent the sulphur atoms. The cysteine side chains are displayed as thick sticks. The sulphur atoms of Cys7, Cys15 and Cys24 form bridges between the metals, while the other sulphur atoms are terminally coordinated.

to a relative geometric accuracy of 10^{-4} [48]. For the non-bonded forces, a twin-range cutoff of 0.8 / 1.4 nm was used with a reaction field correction [46] ($\epsilon_{RF} = 54$, as determined for SPC water [76]). The short cutoff defined at the same time the interface region of background partial charges that enter the quantum-chemical calculations. More precisely, the classical partial charges of any charge group having at least one member closer than the short-range cutoff of 0.8 nm to any quantum atom, were included in the quantum-chemical calculation. This led to an interface region usually consisting of the whole domain plus a shell of water, totally comprising about 655 atoms.

In the quantum-chemical core, the metal ions had a formal oxidation state of +2 and were coordinated to deprotonated methylthiolate. Totally, the quantum core had the composition $M_3(\text{CH}_3\text{S})_9$ and was charged $-3 e$, where M is a placeholder for any metal, Zn or Cd. One of the hydrogen atoms of each methylthiolate served as link atom. The quantum-chemical β -carbon atom was linked to the classical α -carbon atom by means of a bond-constrained link atom approach as described in Section 1.5. Hydrogen atoms were used as link atoms and the bond length ratio was 0.6948.

The protein and water were separately weakly coupled to a temperature bath of 300 K using a 0.1 ps coupling time [47]. The time step for the MDc simulations was 2 fs. In the MDq simulations, 0.5 fs was used to account for the unconstrained bonds in the quantum-chemical part. The non-bonded-interaction pair list was updated every 10 fs. Simulation of a trajectory of 10 ps on a 450 MHz dual-processor pentium-II computer took roughly $1\frac{1}{4}$ h for an MDc simulation, 14 h for an MDq simulation based on MNDO and 30 h for an MDq simulation with MNDO/d. The simulation elapsed times were about 8 ns in the MDc simulations, and about 250 ps in the MDq simulations.

5.3.2 Estimation of Van-der-Waals Interaction Parameters for Cadmium

The van-der-Waals interaction parameters for cadmium were estimated using the GROMOS96 [19] zinc parameters as a starting point. The basic structure of the metal clusters in metallothionein is equivalent to the mineral form of the metal sulfides, zincblende and cadmiumblende. Knowing the structure and the density of the latter, and the masses of the involved atoms, a metal-sulphur distance can be derived. The values shown in Table 5.1 are very close to those in the X-ray [72] and NMR [71] structures of metallothionein. However, the latter values may result from the bond restraints applied in the structure refinement process.

The zinc parameters were scaled to reflect the larger bond length to cadmium, while retaining the depth of the minimum of the van-der-Waals term in the force field. Using $r = l_{\text{Cd-S}}/l_{\text{Zn-S}}$, the ratio of metal-sulphur distances in the mineral, we obtain the scalings

$$\sqrt{C_{\text{Cd}}^{\text{six}}} = r^6 \sqrt{C_{\text{Zn}}^{\text{six}}} \quad (5.1)$$

$$\sqrt{C_{\text{Cd}}^{\text{twelve}}} = r^{12} \sqrt{C_{\text{Zn}}^{\text{twelve}}} \quad (5.2)$$

where $\sqrt{C_{\text{Zn}}^{\text{six}}}$ and $\sqrt{C_{\text{Zn}}^{\text{twelve}}}$ are the GROMOS96 van-der-Waals parameters for zinc. The results are listed in Table 5.1.

Mineral	Density g/cm ³	metal-sulphur distance $l_{\text{M-S}}$ nm	van-der-Waals parameters	
			$\sqrt{C_{\text{M}}^{\text{six}}}$ (nm ⁶ kJ/mol) ^{1/2}	$\sqrt{C_{\text{M}}^{\text{twelve}}}$ $10^{-3}(\text{nm}^{12} \text{ kJ/mol})^{1/2}$
Zincblende	4.102	0.234	0.02045	0.09716
Cadmiumblende	4.82	0.253	0.03267	0.24790

Table 5.1: Some properties of metal sulfides, and non-bonded-interaction force-field parameters for zinc and cadmium.

5.4 Results and Discussion

Both, the MDc and MDq simulations with the semi-empirical method MNDO/d were able to maintain the overall structure of the protein. However, MDq simulations using MNDO without extension to d orbitals failed in this respect. The metal cluster disintegrated after a short simulation period of 15 ps. This was not an accidentally observed unfolding event: Two more simulations with different starting velocities suffered the same fate. Therefore, only the MDq simulations with MNDO/d will be discussed.

The following subsections give more details and compare the results of the two methods against each other and against experimental data. The section is organised as follows. Tables and figures show results grouped in terms of properties such as bond lengths or NOE distances. To avoid confusion, the discussion in the text is grouped in terms of comparisons: comparisons between experimental and simulated structures, comparisons between MDc and MDq simulated results, or between Cd₃, CdZn₂ and Zn₃ variants.

5.4.1 Comparison of the CdZn₂ X-Ray Crystal Structure with the Cd₃ NMR Solution Structure

The structures of the CdZn₂ (X-ray, [72]) and the Cd₃ variant (NMR, [71]) derived from experimental data are very similar [74]. In particular, the metal cores have the same coordinative bonds and metal-sulphur cluster geometries. The polypeptide folds are closely similar. However, the polypeptide loops linking the metal-coordinated cysteines are less well defined in the NMR structure. This finding is attributed to the absence of regular secondary structure and the high degree of dynamic structural disorder.

Figure 5.2 shows the experimental structures and the structures at the end of the simulations. The top row shows the NMR structure to the left and the X-ray structure to the right. Both structures look similar indeed: The metals and sulphurs are nearly identical and the overall fold is the same, however, with quite some variation in the loops between the cysteine residues. There is a difference in the direction of the side chain of Cys13: In the X-ray structure it faces the sulphur from the top-front, while in the NMR structure from behind.

Figure 5.3 shows the backbone C_α atom distances between the X-ray and the NMR structure. The average distances of the *cysteine residues* are 0.18 nm for the C_α atoms, 0.14 nm for the C_β atoms, and 0.03 nm for both the sulphurs and the metals. The metal-sulphur configurations are very similar. The distances increase with increasing distance from the metal core, both atom-wise within the cysteines and residue-wise in the entire domain. This is not surprising because the cysteine sulphur atoms were superimposed.

The positional difference between cysteine C_α atoms is about 0.2 nm on average, which is reasonable considering the difference between metal atoms and environment. However, the average difference over all C_α atoms is about 0.3 nm which reflects a high degree of flexibility, which is most probably due to the absence of regular secondary structure and its stabilising hydrogen bonds.

There is a large variation in the C_α positions of individual residues. Not surprisingly, the cysteines mostly exhibit a very low variation. An exception is Cys13, which was previously mentioned as having a different side chain orientation in the two structures. There is a region of large deviation from Thr9 to Ser14. This loop has different conformations in the two structures, see Figure 5.2 (a) and (b), loop at the top. The same applies to Gly17, which is at the edge

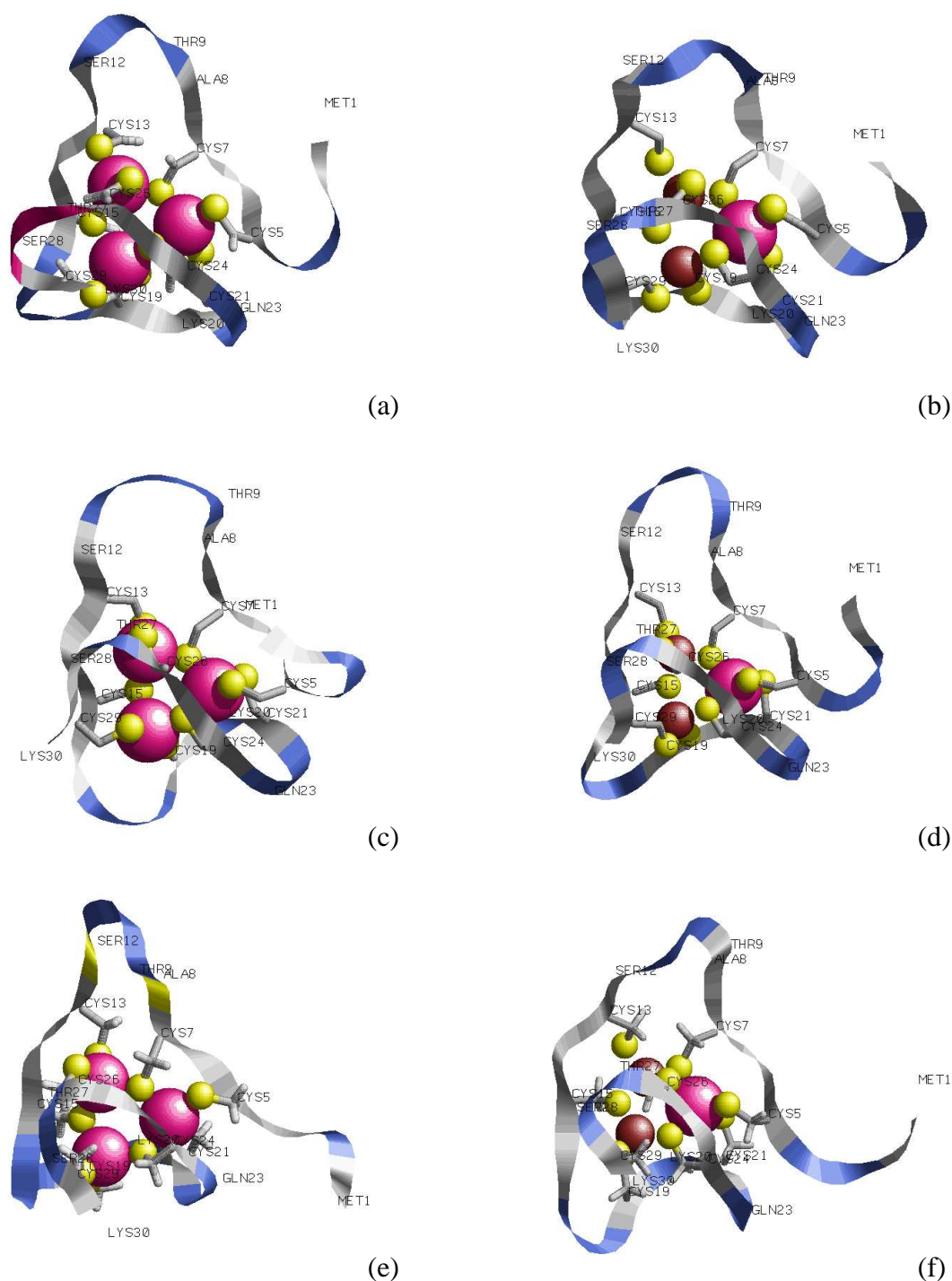


Figure 5.2: Schematic structures of the β domain of metallothionein containing three metal ions. All structures are rotated to display the same view on the metals. Cadmium ions are displayed as large balls, zinc ions as dark medium-sized balls, sulfur atoms as small light balls, the rest of the cysteine side chains as thick sticks, and the protein backbone as ribbon. Selected residues are labeled. (a) Cd_3 NMR structure [71], (b) CdZn_2 X-ray structure [72], (c) Cd_3 final MDC simulation structure, (d) CdZn_2 final MDC simulation structure, (e) Cd_3 final MDq simulation structure, (f) CdZn_2 final MDq simulation structure. The X-ray structure and the MDC simulations do not have hydrogens on the cysteine side chain, the NMR structure shows a single pseudo atom, and the MDq simulations have explicit hydrogens.

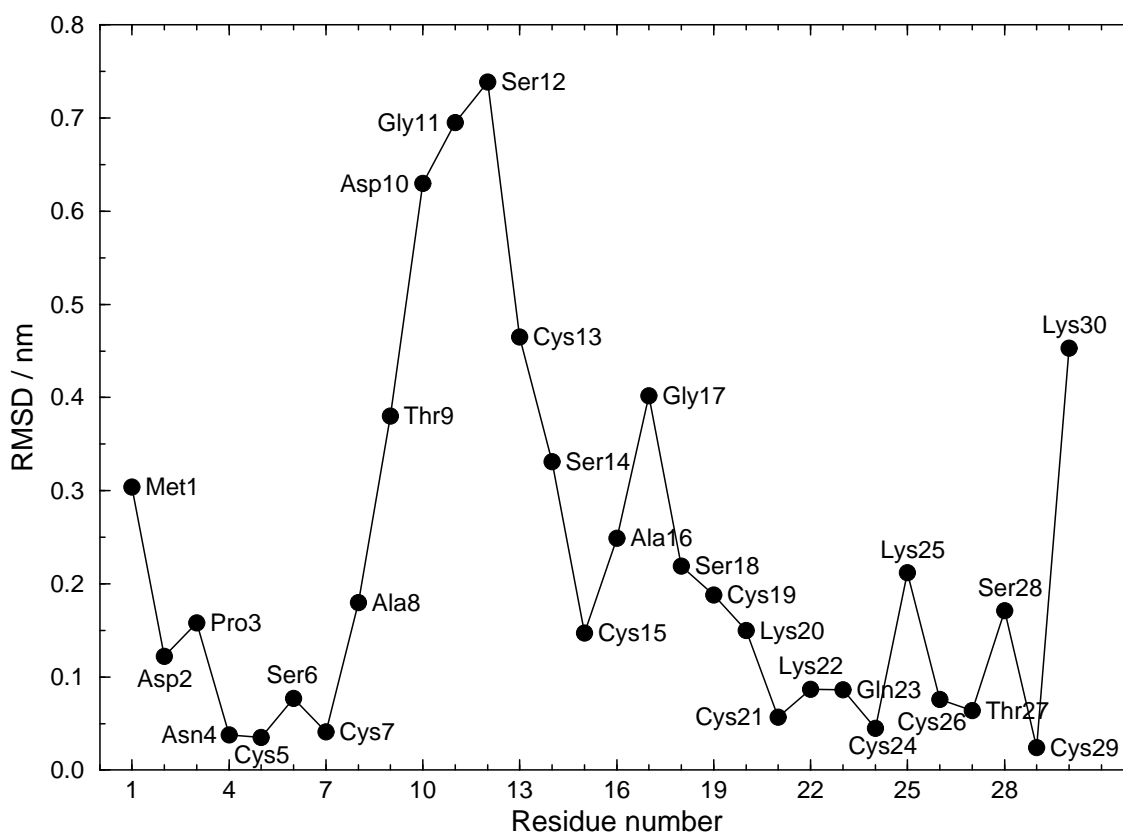


Figure 5.3: Positional distances for C_{α} -atoms between X-ray [72] and NMR structure [71]. The two structures were superimposed using a translational and rotational least-squares fit for the cysteine sulphur atoms.

of the loop to the lower back left in Figure 5.2 (a) and (b). The terminal residue Met1 and the linker to the second domain, Lys30, are expected to exhibit more structural variation. The X-ray structure only weakly defines residues Met1 and Asp2. These results suggest that the simulated root-mean-square (RMS) deviations from the experimental structures should be below 0.05 nm for the metal and cysteine sulphur atoms, below 0.15 nm for cysteine C_{β} atoms, below 0.2 nm for cysteine C_{α} atoms, and below 0.3 nm for all C_{α} atoms. These thresholds will be used as a tolerance level in analysing the simulations.

Similar regions of high structural variation are also found amongst the set of the ten best NMR structures (Figure 2 in Reference [74]). There, regions around Asp10, Ala16 and Thr27 exhibit large variations, apart from the chain ends. It is likely that the flexibility of the intercysteine loops causes these increased deviations.

Table 5.2 lists hydrogen bonds present in the experimental structures. Both experimental structures were relaxed by an energy minimisation with atom-positional restraints (force constants equal $2.5 \cdot 10^4 \text{ kJmol}^{-1}\text{nm}^{-2}$) to the original structure in order to remove stress and to adapt the structure to the GROMOS force field and its criterion of hydrogen bonds (see caption of Table 5.2). The asterisks mark hydrogen bonds already present in the original experimental structures.

There are relatively few hydrogen bonds, and the hydrogen bonding pattern differs quite much between the X-ray and the NMR structures. Two hydrogen bonds are present in both

Hydrogen bond specification		Experimental		Percentage of occurrence						
		NMR	X-ray	MDc			MDq			
				Cd ₃	CdZn ₂	Cd ₃	CdZn ₂	Zn ₃	Cd ₃	CdZn ₂
Met1-Lys25	N-H-O	—	—	21.1	—	—	—	—	—	—
Asp2-Lys25	N-H-O	—	—	31.2	—	—	—	—	—	—
Asp2-Cys5	N-H-O	—	—	—	7.7	48.7	8.9	3.2	—	—
Asn4-Asp2	ND _x -HD _{2x} -OD _x	—	—	46.1	48.9	14.2	18.0	54.7	85.0	—
Asn4-Asp2	N-H-OD _x	—	100.0*	87.9	86.4	79.1	35.9	31.7	5.0	—
Cys5-Gln23	N-H-O	—	—	—	—	—	2.3	—	41.4	—
Cys5-Asp2	N-H-OD _x	—	100.0	40.7	64.3	64.7	2.9	12.1	—	—
Cys5-Asp2	N-H-O	—	—	4.4	2.5	2.0	23.0	24.3	3.0	—
Gly11-Ala8	N-H-O	—	—	5.0	18.9	18.9	16.2	2.6	51.0	—
Ser12-Asp10	N-H-OD _x	100.0*	100.0*	3.0	25.6	3.3	40.0	12.6	4.2	—
Ser12-Asp10	OG-HG-OD _x	100.0	200.0	16.8	35.0	6.6	104.7	105.5	28.0	—
Cys15-Cys13	N-H-O	—	—	4.3	—	—	—	—	30.8	—
Cys15-Ser28	N-H-OG	—	—	—	92.9	—	—	65.1	10.2	—
Ala16-Ser28	N-H-OG	—	100.0*	—	5.1	—	—	13.4	25.4	—
Gly17-Cys29	N-H-O	—	—	—	—	—	—	—	21.2	—
Gly17-Lys30	N-H-O _x	—	—	38.1	6.6	12.8	—	—	—	—
Ser18-Cys15	N-H-O	—	100.0*	—	—	—	20.7	—	5.6	—
Cys19-Gly17	N-H-O	—	—	25.2	8.7	—	23.8	—	—	—
Lys22-Asn4	N-H-OD _x	—	—	—	—	—	22.1	—	—	—
Lys22-Asn4	N-H-O	—	100.0	8.7	11.6	27.5	58.0	4.7	5.8	—
Gln23-Gln23	NE _x -HE _{2x} -O	100.0	—	—	—	—	—	—	—	—
Gln23-Asn4	N-H-OD _x	—	—	—	4.2	22.5	47.5	—	—	—
Gln23-Asn4	N-H-O	—	100.0*	—	84.8	81.2	68.4	89.4	13.0	—
Lys25-Gln23	NZ-HZ _x -O	—	100.0*	—	—	—	—	7.9	—	—
Lys25-Cys24	NZ-HZ _x -O	—	—	—	—	—	—	2.8	43.2	—
Lys25-Asp2	NZ-HZ _x -OD _x	—	100.0*	—	17.5	2.9	—	10.3	65.8	—
Cys26-Cys29	N-H-O	100.0	—	—	11.7	—	—	—	—	—
Thr27-Lys25	N-H-O	100.0*	—	—	—	—	—	—	—	—
Ser28-Cys13	OG-HG-O	—	100.0	—	4.7	—	—	64.0	—	—
Ser28-Ser14	N-H-OG	—	—	—	—	66.0	5.9	—	3.4	—
Cys29-Cys26	N-H-O	—	100.0*	14.9	10.8	32.9	56.4	38.9	50.8	—
Lys30-Ser28	N-H-O	100.0*	—	—	10.2	3.7	9.6	11.3	2.2	—

Table 5.2: List of hydrogen bonds that occur in any experimental structure or in any simulation for more than 20 %. Values larger than 100 % result from three-center hydrogen bonds, when both single components are present at the same time. The character x denotes two equivalent atoms which can interchange. In such cases, the percentages for equivalent atoms were added. The experimental structures were relaxed by an energy minimisation to remove conformational stress. Hydrogen bonds present in the original experimental structure are marked by an asterisk. A hydrogen bond is considered to exist if the hydrogen-acceptor distance is smaller than 0.25 nm and the donor-hydrogen-acceptor angle is larger than 135° .

experimental structures, both involving Ser12 and Asp10. Interestingly, these residues are in the region of maximum distance between the two structures (Figure 5.3). It seems that a relatively stable turn is formed, whose relative position with respect to the metal core is different. The other hydrogen bonds are present in either of the structures. Moreover, there is no typical hydrogen-bonding pattern of an α helix (residue($n + 4$)-residue(n) N-H-O) or a β sheet in the entire list.

5.4.2 The Cd Zn₂ MDc Simulation Compared to the X-Ray Structure

A crucial property of the metal cluster are the metal-sulphur bond lengths. For this reason, bond lengths were averaged in time, after an equilibration phase of 50 ps, and summarised in Table 5.3. The experimental bond lengths are quite close to the ones found in the minerals (Table 5.1). Compared to the experimental X-ray structure, the MDc simulation yields bonds about 0.025 nm too short for the cadmium ion. For zinc, the difference is worse, 0.031 nm. It should be noted that the experimental values were refined against 0.24 nm for zinc and 0.25 nm for cadmium.

The structures of the entire domain were analysed by means of atom-positional root-mean-square deviations (RMSD), Figure 5.4. They were calculated by comparing structures from the trajectory with a reference structure after a translational-rotational fit over the cysteine sulphur atoms. Separate RMSD curves were calculated for the metals, for the cysteine sulphur atoms, for the cysteine β -carbon atoms, for the cysteine α -carbon atoms, and all α -carbon atoms of the domain. Not surprisingly, average RMSD values increase in that order, which reflects increasing distance from the metal core.

Figure 5.4 shows RMS deviations from the initial structures. Graph (c) for the Cd Zn₂ MDc simulation shows that the deviation of the metal and sulphur atoms is above the tolerance level of 0.05 nm as estimated in Section 5.4.1. This can be explained as a consequence of the too short bond lengths in the metal cluster. The deviations of the cysteine C $_{\alpha}$ and C $_{\beta}$ atoms are barely below 0.2 nm. The total of the C $_{\alpha}$ atoms is below 0.3 nm except for a peak shortly before 4 ns. Summarising, the MDc simulation shows too large a deviation for atoms of the metal core (metal, sulphur and cysteine C $_{\beta}$ atoms), whereas the backbone deviations are comparable to the difference between the X-ray and NMR structures in Figure 5.3. The same applies to the RMS deviations from the X-ray structure shown in Figure 5.5c. This is no surprise, as the initial structure for the simulation was obtained using energy minimisation from the X-ray structure.

The X-ray structure exhibits bond angles of the bridging sulphur in a narrow range around 104° (Table 5.4). In contrast, the MDc simulation yields average angles larger than 130°. As a consequence, the angles at the metals between bridging sulphurs are smaller in the simulation (95°) than in experiment (between 104° and 123°). These results are further discussed in Section 5.4.6.

Table 5.2 lists hydrogen bonds present in any experimental structures or in any simulation. In the Cd Zn₂ MDc simulation most hydrogen bonds of the X-ray structure are observed, Lys25-Gln23 NZ-HZ $_x$ -O and Ser18-Cys15 N-H-O are lost. Most of the frequent hydrogen bonds in the simulation are present in the crystal, however, two prominent new ones appear in the simulation: Asn4-Asp2 ND $_x$ -HD $_x$ -OD $_x$ and Cys15-Ser28 N-H-OG. Residues Lys25, Ser18, Ser28 and Cys15 have deviations larger than 0.14 nm between X-ray and NMR structures (Figure 5.3), which indicates different behaviour in the crystal and in solution. Asp2 is weakly defined in the X-ray structure. This might explain the absence of the corresponding hydrogen bond in the X-ray structure.

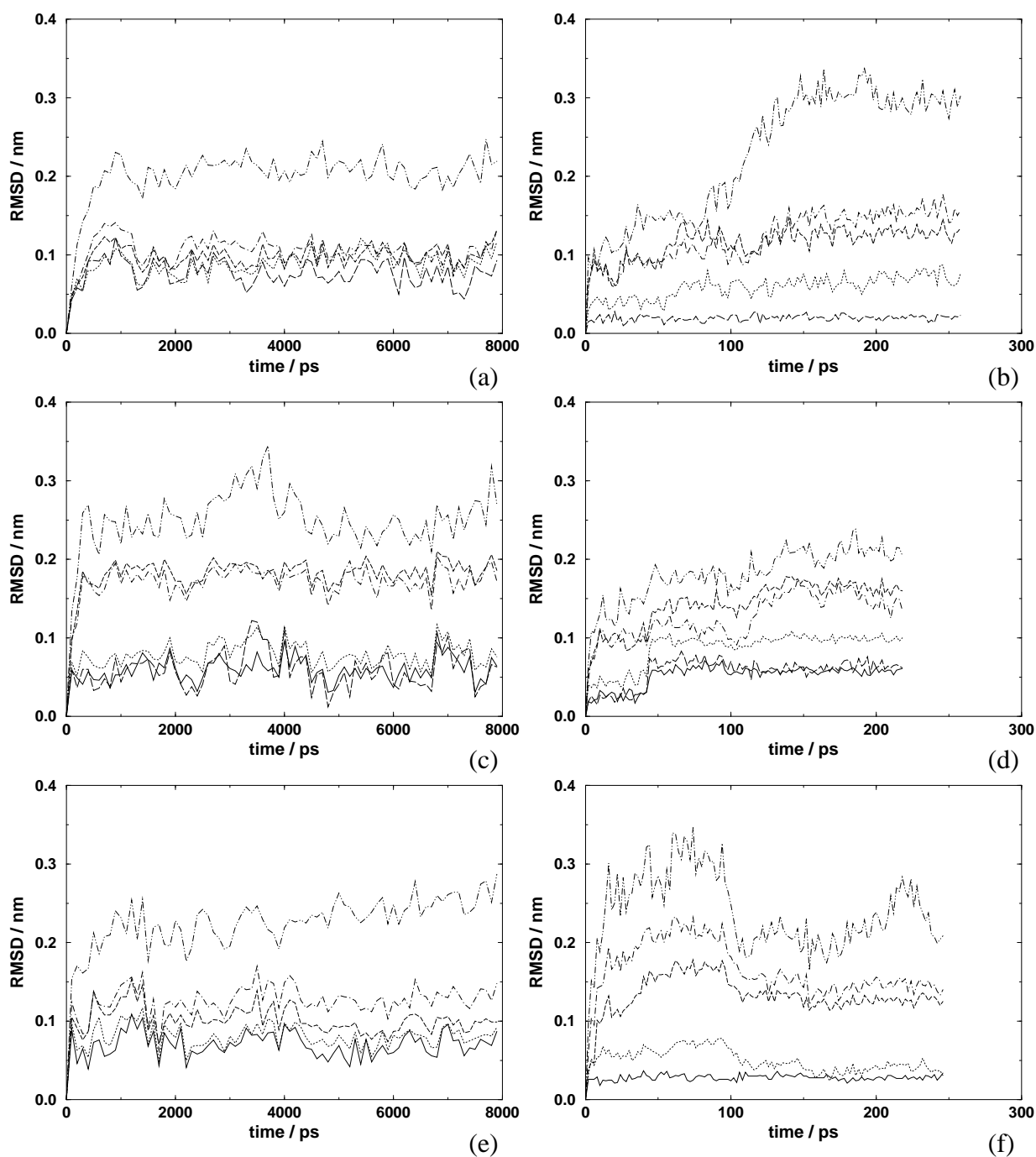


Figure 5.4: Atom-positional RMS deviations from the initial structures. The RMS deviation was calculated after a translational-rotational fit over the cysteine sulphur atoms. Left panel: MDc simulations; right panel: MDq simulations. Top row: Cd₃ variant; middle row: CdZn₂ variant; bottom row: Zn₃ variant. Solid line: zinc ions; broken line: cadmium ions; dotted line: sulphur atoms; dashed line: cysteine C_β atoms; dot-dashed line: cysteine C_α atoms; dot-dot-dashed line: all C_α atoms.

	Metal 1				Metal 2				Metal 3			
	M-S _t		M-S _b		M-S _t		M-S _b		M-S _t		M-S _b	
MDc Cd ₃												
aver	0.2229	0.2251	0.2255	0.2249	0.2293	0.2265	0.2303	0.2306	0.2246	0.2241	0.2251	0.2280
fluct	0.0017	0.0021	0.0016	0.0019	0.0027	0.0021	0.0022	0.0023	0.0016	0.0016	0.0014	0.0018
mean	0.2246				0.2292				0.2255			
MDc CdZn ₂												
aver	0.2234	0.2238	0.2273	0.2273	0.2054	0.2037	0.2110	0.2096	0.2037	0.2051	0.2083	0.2066
fluct	0.0014	0.0019	0.0020	0.0016	0.0015	0.0015	0.0021	0.0020	0.0013	0.0014	0.0014	0.0014
mean	0.2255				0.2074				0.2059			
MDc Zn ₃												
aver	0.2031	0.2032	0.2069	0.2074	0.2046	0.2024	0.2094	0.2072	0.2066	0.2115	0.2126	0.2112
fluct	0.0014	0.0013	0.0013	0.0014	0.0011	0.0011	0.0017	0.0015	0.0016	0.0023	0.0021	0.0019
mean	0.2051				0.2059				0.2105			
MDq Cd ₃												
aver	0.2441	0.2447	0.2454	0.2400	0.2429	0.2363	0.2406	0.2430	0.2447	0.2338	0.2424	0.2464
fluct	0.0015	0.0016	0.0016	0.0014	0.0028	0.0017	0.0017	0.0028	0.0017	0.0017	0.0032	0.0015
mean	0.2435				0.2407				0.2418			
MDq CdZn ₂												
aver	0.2466	0.2507	0.2517	0.2423	0.2201	0.2299	0.2136	0.2180	0.2313	0.2225	0.2164	0.2229
fluct	0.0019	0.0020	0.0038	0.0022	0.0024	0.0025	0.0026	0.0031	0.0023	0.0020	0.0031	0.0040
mean	0.2478				0.2204				0.2233			
MDq Zn ₃												
aver	0.2282	0.2285	0.2161	0.2118	0.2211	0.2325	0.2050	0.2076	0.2300	0.2292	0.2068	0.2075
fluct	0.0018	0.0018	0.0024	0.0020	0.0023	0.0029	0.0025	0.0032	0.0023	0.0021	0.0022	0.0032
mean	0.2212				0.2165				0.2184			
NMR Cd ₃												
aver	0.2509	0.2508	0.2538	0.2641	0.2539	0.2474	0.2490	0.2676	0.2508	0.2531	0.2518	0.2645
mean	0.2549				0.2545				0.2551			
X-ray CdZn ₂												
aver	0.2490	0.2490	0.2540	0.2535	0.2296	0.2405	0.2409	0.2476	0.2368	0.2370	0.2369	0.2328
mean	0.2514				0.2397				0.2359			

Table 5.3: Time-averaged bond lengths (*aver*) and fluctuations in time (*fluct*) for bonds between metal (*M*) and sulphur (*S*) atoms, in nm. The two left columns of every metal group represent the bonds to the terminal sulphur (*S_t*), the two right columns represent the bonds to the bridging sulphur (*S_b*). The means of the four bonds per metal are also given. Metal 1 is the cadmium atom in the CdZn₂ variants.

5.4.3 The CdZn₂ MDq Simulation Compared to the X-Ray Structure

The length of the cadmium-sulphur bond is well reproduced (Table 5.3). More difficulties arise with zinc, whose bond to sulphur appears about 0.016 nm too short.

The RMS deviations from the initial conformation (Figure 5.4d) for all atom types are initially below the thresholds mentioned in Section 5.4.1. However, there is a stepwise increase at 40 ps and the levels for the metals and the sulphurs rise above the thresholds. Then the deviations stay around these levels with the same small fluctuations as before. The metal core seems to undergo a transition to another stable structure. The deviations for the two C_α atom types also increase a

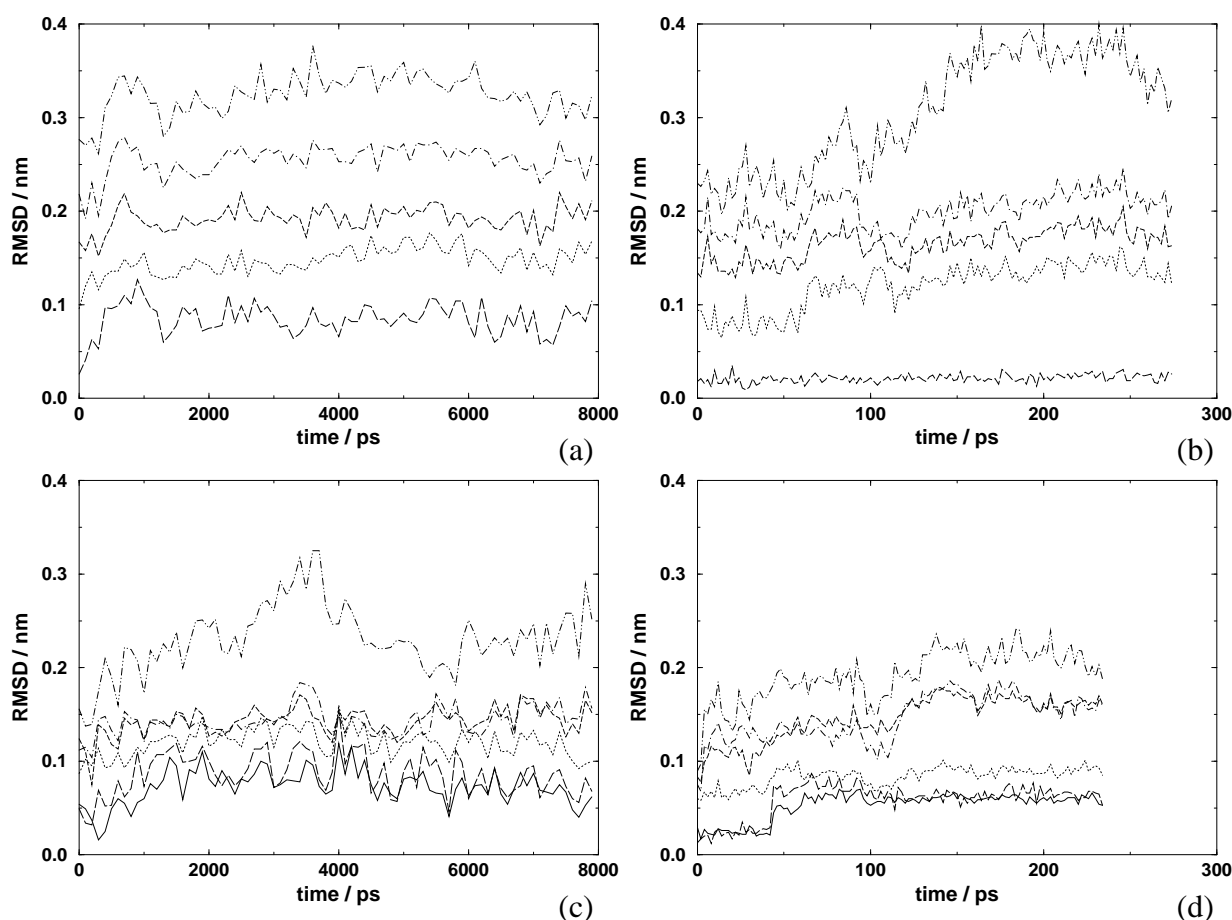


Figure 5.5: RMS deviations from the experimental structures. For the Cd_3 cluster (upper panel), this is the NMR structure [71] and for the CdZn_2 cluster (lower panel), it is the X-ray structure [72]. Left column: MDC simulations; right column: MDQ simulations. Solid line: zinc ions; broken line: cadmium ions; dotted line: sulphur atoms; dashed line: cysteine C_β atoms; dot-dashed line: cysteine C_α atoms; dot-dot-dashed line: all C_α atoms.

bit, but remain well below the tolerance threshold. Virtually the same behaviour is observed in comparison to the X-ray structure in Figure 5.5d. As the initial structure was derived from the X-ray structure, this is not surprising. However, there is a significant difference in the deviation of the sulphur atoms: it is initially already above 0.05 nm, but the increase at 40 ps is smaller, the final level lower than compared to the initial structure. Overall, the MDQ simulation gives small deviations, mostly below 0.2 nm, which is considered to be a normal structural variation for the α carbon atoms. Compared to the MDC simulations, the deviations are mostly smaller and their fluctuations narrower. However, comparing the same initial period of 240 ps, the deviations are comparable.

The bond angles of the MDQ simulations yield an irregular picture (Table 5.4). The average of the sulphur bridge angles ($\text{M}-\text{S}_b-\text{M}$) is quite close to the experimental values, but the differences between the individual angles are large, in contrast to experiment. The metal ring bridge angles ($\text{S}_b-\text{M}-\text{S}_b$) are too large, but give the correct trend in the variation of the three individual angles. Two of the terminal angles ($\text{S}_t-\text{M}-\text{S}_t$) are in good agreement with experiment, the other one deviates by 15° .

	Sulphur ring bridges M-S _b -M			Metal ring bridges S _b -M-S _b			Metal-terminal angles S _t -M-S _t		
	S _{b1}	S _{b2}	S _{b3}	M ₁	M ₂	M ₃	M ₁	M ₂	M ₃
MDc Cd ₃									
aver	132.0	133.0	147.7	101.5	87.8	92.2	96.9	156.7	125.8
fluct	2.8	3.1	2.8	3.5	2.0	1.6	5.2	14.4	4.0
mean	137.6			93.8			126.4		
MDc CdZn ₂									
aver	131.1	134.9	141.2	93.2	93.8	97.7	96.2	115.1	112.0
fluct	2.4	2.0	2.5	2.4	1.9	1.4	1.9	6.4	3.0
mean	135.7			94.9			107.8		
MDc Zn ₃									
aver	131.7	137.2	135.7	95.8	96.4	98.8	115.2	113.9	95.8
fluct	2.5	2.1	2.1	1.5	1.5	2.2	3.4	3.6	4.6
mean	134.9			97.0			108.3		
MDq Cd ₃									
aver	132.8	122.4	122.9	105.0	104.0	110.6	110.6	97.7	119.1
fluct	3.0	3.2	2.5	3.7	2.0	3.4	4.8	4.6	4.3
mean	126.1			106.5			109.1		
MDq CdZn ₂									
aver	115.9	82.8	100.7	123.3	115.0	154.2	97.0	97.3	106.8
fluct	3.7	4.8	7.3	2.7	7.1	6.0	1.7	4.5	2.4
mean	99.8			130.8			100.4		
MDq Zn ₃									
aver	112.7	111.2	109.7	100.2	108.1	128.7	94.8	92.2	92.5
fluct	6.0	5.3	8.1	2.7	4.2	8.3	2.9	3.7	1.8
mean	111.2			112.3			93.2		
NMR Cd ₃									
aver	108.5	110.0	109.4	102.9	101.6	106.5	115.9	112.3	110.1
mean	109.3			103.7			112.8		
X-ray CdZn ₂									
aver	102.3	103.3	106.2	109.0	103.7	122.5	113.1	99.1	105.6
mean	103.9			111.7			105.9		

Table 5.4: Time-averaged bond angles (*aver*) and fluctuations (*fluct*) in the metal core, in degrees. Left group: angle at bridging sulphurs between two metal ions; middle group: angle at a metal between two bridging sulphurs; right group: angle at a metal between two non-bridging sulphurs. The means of the groups are also given. The bridging sulphur atoms S_{b1}, S_{b2} and S_{b3} belong to residues Cys24, Cys7 and Cys15 respectively.

Hydrogen bonds present in the X-ray structure are mostly observed in the MDq simulation. The only hydrogen bond lost in the simulation is Ser18–Cys15 N–H–O. Prominent new hydrogen bonds established in the simulation are Asn4–Asp2 NDx–HDx–ODx and Cys15–Ser28 N–H–OG, as in the MDc simulation. Most remarkably, the three-center hydrogen bond Ser12–Asp12 OG–HG–ODx present in the X-ray structure is reproduced in the MDq simulation.

5.4.4 Comparison of the Cd₃ MDc Simulation with NMR Data

The bond lengths (Table 5.3) are systematically too short, with differences between MDc simulations and experiment of 0.028 nm. This has two reasons. Firstly, the parameters for cadmium were estimated using the zinc parameters as a basis. So the inaccuracy of the results for zinc is propagated to cadmium. Secondly, the bond lengths in the NMR structure are larger than in the X-ray structure, due to different bond-length parameters (0.26 nm) used in the structure derivation [77].

Figure 5.4a shows that the protein quickly reaches a stable conformational range without drift or long-time features. No significant conformational changes take place within 8 ns. The backbone and the cysteine C _{β} atoms exhibit small deviations well below the tolerance thresholds. However, the deviations for the metal and sulphur atoms are above the tolerance. Again, this seems to be a consequence of the too short bond lengths induced by the estimated van-der-Waals parameters for cadmium.

As the simulations were started using the X-ray structure as a basis, and the differences between the X-ray and NMR structures are partially large, it is expected that the deviations of the Cd₃ cluster against the NMR structure are quite large. As there is no significant conformational change observed, there is also no progress towards the NMR structure (Figure 5.5a).

The bond angle problems (Table 5.4) are mostly analogous to the MDc CdZn₂ case: The sulphur-bridge angles (M–S _{b} –M) are too large. The metal site 2 is heavily distorted: the ring-bridge angle is smaller than 90°, while the angle to the terminal sulphurs is larger than 150°. This raises the question whether the classical electrostatics is capable of maintaining a tetrahedral configuration around a relatively large ion.

The NMR structure exhibits a small number of hydrogen bonds (Table 5.2), of which two are not reproduced by any simulation: Gln23–Gln23 NEX–HE2x–O and Thr27–Lys25 N–H–O. The first is an intraresidual hydrogen bond which depends on the side-chain conformation. However, the latter is a backbone hydrogen bond, as are two more, Cys26–Cys29 N–H–O and Lys30–Ser28 N–H–O, which all three are not recovered in the Cd₃ MDc simulation. The two Ser12–Asp10 experimental hydrogen bonds are reproduced in the Cd₃ MDc simulations at a low percentage. These are the only hydrogen bonds in the NMR structure that are also present in the X-ray structure.

Hydrogen bonds established during the Cd₃ MDc simulation are generally similar to the CdZn₂ variant. New hydrogen bonds are formed at the beginning of the chain at Met1 and Asp2. However, it seems that the character of the initial (X-ray-based) structure is largely retained (see Figure 5.2).

The compatibility of the simulated trajectories with NOE distance limits derived from experiment [71] was investigated. Table 1 in Reference [71] groups the NOE values in three categories: (i) Sequential backbone bounds, (ii) medium-range backbone and long-range backbone bounds, (iii) interresidual bounds with side-chain protons. The NOE atom-atom distances were calculated from the trajectories using r^{-3} averaging.

Despite the large deviations to the NMR structure (Figure 5.5a), the sequential NOE bounds (top left panel in Figure 5.6) are quite well satisfied. However, there are sizeable NOE bound violations (larger than 0.1 nm) in the other two categories (middle and bottom left panel in Figure 5.6). The details of the violated NOEs are listed in Table 5.5.

The large violations of NOE bounds numbered 26 and 27 of the side-chain interresidual bounds are striking. However, these NOE bounds involve protons on Lys30 (the β -carbon protons and the amide proton). This residue is the linker to the second domain of metallothionein, which

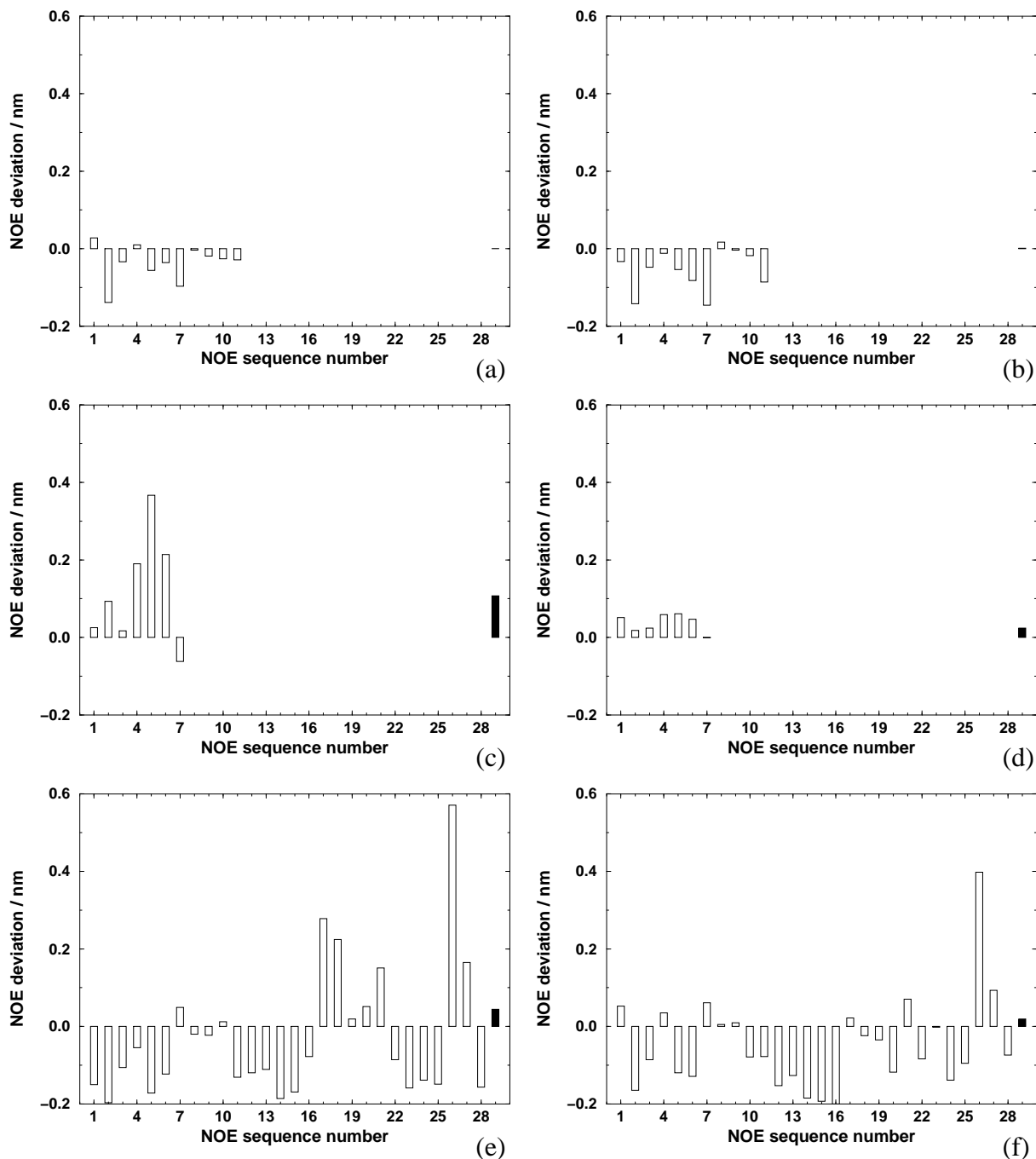


Figure 5.6: Comparison of simulated NOE atom-atom distances with experimental NOE bounds. The simulated NOE average distances are r^{-3} averages. Values greater than zero violate experimental data. Left column: MDC simulations; right column: MDq simulations. Top row: sequential backbone bounds; middle row: medium-range backbone and long-range backbone bounds; bottom row: interresidual bounds with side-chain protons. NOE values appear only once; the second entry from the NOE list (Table 1 in [71]) was omitted. The sequence of the NOEs is the same as in Table 1 of [71]. The solid black bar to the far right in each graph gives the average over all NOE violations.

NOE sequence number	residue	proton		residue	proton
backbone medium-range and long-range NOE bounds					
4	Cys5	C $_{\alpha}$	–	Cys21	C $_{\alpha}$
5	Cys5	C $_{\alpha}$	–	Gln23	Amide N
6	Cys5	C $_{\alpha}$	–	Cys24	C $_{\alpha}$
side-chain interresidual NOE bounds					
17	Ala8	C $_{\beta}$ (methyl)	–	Cys13	Amide N
18	Thr9	C $_{\gamma}$ (methyl)	–	Asp10	Amide N
21	Lys20	C $_{\beta}$ (methylene)	–	Cys21	Amide N
26	Thr27	C $_{\alpha}$	–	Lys30	C $_{\beta}$ (methylene)
27	Ser28	C $_{\beta}$ (methylene)	–	Lys30	Amide N

Table 5.5: List of violated NOE bounds. See also Figure 5.6.

was omitted in the simulations. It is likely that these violations are artifacts of the omission of the second domain: It is the last residue retained in the β domain simulations and its structure may therefore be different from the one in the complete protein.

The interresidual NOE number 18 involves Thr9 which is located in the flexible loop from Thr9 to Ser14, in which already large differences between the X-ray and the NMR structure were observed (Section 5.4.1). All other violated NOEs involve cysteine residues. The limited accuracy of the description of the metal-sulphur cluster configuration in the MDc simulation could be the origin of the NOE bound violations involving cysteine residues.

5.4.5 Comparison of the Cd₃ MDq Simulation with NMR Data

Also in the MDq Cd₃ simulation, the cadmium-sulphur bond lengths are systematically too short, on average 0.013 nm shorter than in the NMR structure (Table 5.3).

The RMS deviation from the initial structure is shown in Figure 5.4b. The deviation of the cadmium atoms is very low and the fluctuations very small. The deviation of the sulphurs drifts slowly upwards and exceeds the tolerance level of 0.05 nm. Cysteine C $_{\alpha}$ and C $_{\beta}$ atoms deviate little from the initial structure, as do the backbone C $_{\alpha}$ atoms before 100 ps. Afterwards, there is an increase up to and RMSD of 0.3 nm, seemingly stabilising there.

Compared to the NMR structure (Figure 5.5b), the results are similar, but the deviations are larger except for cadmium. Specifically, the increase of the backbone's C $_{\alpha}$ atoms in Figure 5.4b is not a progress towards the NMR structure, a similar increase is evident in Figure 5.5b, being even larger. Figure 5.2e shows that the amino terminus (Met1) turned around during the simulation, causing the increase mentioned. Figure 5.7 confirms this to be the primary cause by showing residue-wise distances for the C $_{\alpha}$ atoms from the NMR structure.

The bond angles (Table 5.4) of the sulphur bridges are considerably larger than in the NMR structure, by about 15° on average. In turn, one would expect that the metal-bridge angles would be too small. This is not the case: these angles match well, the simulation angles being on average only 3° larger than the NMR results. It should be noted that a tetrahedral symmetry on the sulphur and cadmium atoms was assumed in the structure derivation [77].

The hydrogen-bonding situation (Table 5.2) is similar to the MDc simulation. One of the experimental hydrogen bonds, Lys30–Ser28 N–H–O, is reproduced by the MDq simulation at a low percentage. Several high-percentage hydrogen bonds occur that are also present in the X-ray

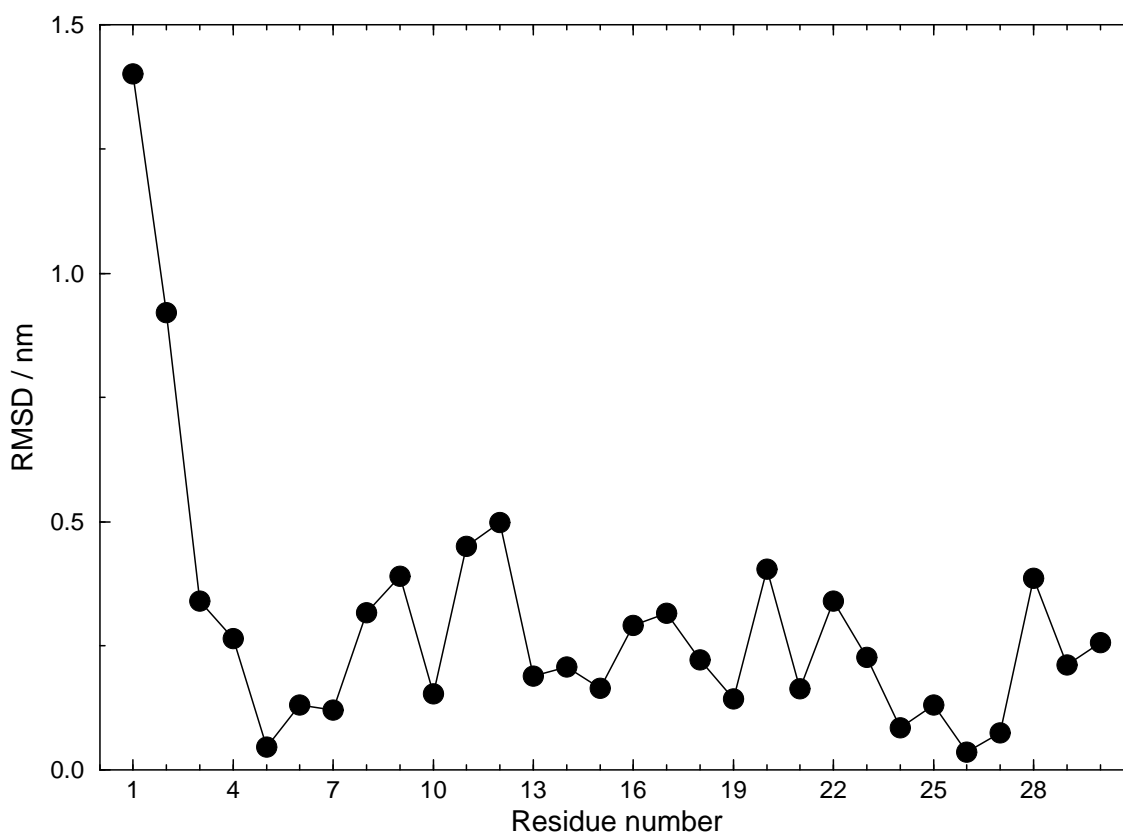


Figure 5.7: Residue-wise distances between C_{α} atoms of the final structure of the MDq Cd_3 simulation and the NMR structure.

structure, but not in the NMR structure. A prominent new hydrogen bond emerging during the simulation is Gln23–Asn4 N–H–ODx.

The sequential NOE bounds (top right panel in Figure 5.6) are quite well satisfied. The differences between the MDc and the MDq simulations are minor. The backbone medium-range and long-range NOE bounds (middle row) are better represented by the MDq simulation. The same applies to the side-chain interresidual bounds (bottom row). Only the NOE bounds involving Lys30 are still heavily violated. In total, the NOE bounds are very well satisfied using the MDq method, although the structure is quite different from the NMR structure.

5.4.6 Comparison of the Classical MDc and Quantum-Chemical MDq Simulations

In terms of the metal-sulphur bonds, the bond lengths are about 10 % too short in the MDc simulations, whereas the MDq simulations yield smaller errors, the bond lengths still being systematically too short. However, the bond lengths in the X-ray and NMR structures are largely determined by refinement parameters. It is also not clear whether the bond lengths should be the same as in the corresponding metal sulfide minerals.

A pronounced difference between the MDc and the MDq simulations is exhibited in the metal-sulphur-metal bridges. The MDc simulations yield average $M-S_b-M$ bridge angles of 136° , with all single angles larger than 130° . In contrast, the MDq simulations give an average

of 114° , with all single angles except one being below 130° . The MDq values are much closer to the ideal tetrahedral angle of 109° . This clearly points at the deficiency of classical methods in describing metal-sulphur binding. With electrostatics only, thus neglecting the lone pair, the two metals would favour a linear configuration. This leads to the elongated M–S_b–M bridge angle. Without the aid of bonded interactions, such a situation cannot properly be described by the non-bonded interaction terms of a force field (van-der-Waals and electrostatic).

Through the six-membered ring of metals and bridging sulphurs, the elongated angles at the sulphur bridges in the MDc simulations naturally lead to smaller ring bond angles at the metals (second column in every metal group in Table 5.4), which are in the range 90° to 100° . In the MDq simulations, these bond angles are larger than 100° throughout. The angles on the metal centers are more or less correctly described: The correct tetrahedral coordination is at the same time the classically favoured minimum energy structure. Especially for the angles at zinc, basically no difference between the two methods is evident. For angles at cadmium, a range from 87° to 155° is accessible in the MDc simulations, while the range narrows to 97° to 119° in the MDq simulations. However, the angles at the metal centers are irregular and disallow clear conclusions.

Comparing the three metal sites to each other, differences larger than the fluctuations are evident even when the sites contain the same metal. The short range situation is equal for all sites: a metal ion tetrahedrally coordinated by four deprotonated cysteines. So the aforementioned differences must be caused by long-range interactions such as anisotropic charge distributions or steric hindrance by the protein backbone.

When comparing structural differences of the polypeptide backbone or NOE atom-atom distances, one should bear in mind the large difference in simulation lengths between the MDc and MDq simulations: 8000 ps versus 240 ps respectively. Thus the former are much better relaxed and equilibrated than the latter.

In terms of the RMS deviations (Figures 5.4 and 5.5), the fluctuations for the metal and the sulphur atoms are significantly smaller in the MDq simulations. The quantum-chemical description seems to yield a more realistic potential-energy surface than the classical electrostatics combined with van-der-Waals terms. The crystal structure of the CdZn₂ variant is better represented by the MDq simulation, giving mostly lower deviations than the MDc simulation (Figure 5.5 lower panel).

The MDq simulation yields virtually no significant NOE bound violations, while the MDc simulation shows a handful of violations. However, this difference could be due to the different averaging times.

5.4.7 Comparison of the Simulations of the Cd₃, CdZn₂ and Zn₃ Variants

The results for the Zn₃ variant are mostly in line with what was said previously for zinc in the CdZn₂ variant. Interestingly, the atom-positional RMS deviations of the C_α atoms in the MDc simulation (Figure 5.4e) are smaller than for the CdZn₂ variant, although the initial structure is from the CdZn₂ variant, and the smaller zinc ions should induce a further contraction of the metal core. The deviations for the sulphur and zinc atoms are slightly higher, as expected.

The MDq simulation exhibits low deviations for the metals, and also the other atom types return below the tolerance thresholds after an initial disorder of about 100 ps, which is caused by motions of the N-terminus and the Thr10 loop. Changes in the metals do not induce significant changes in the overall structure.

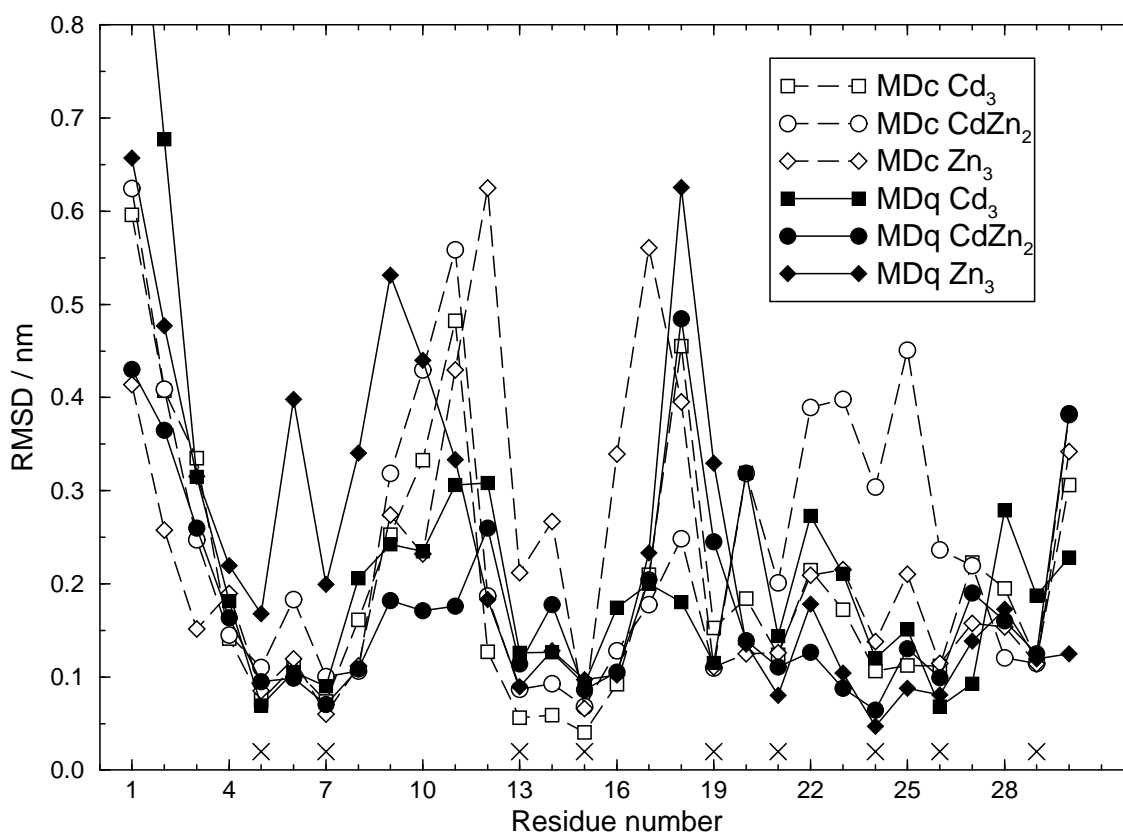


Figure 5.8: C_{α} atom-positional RMS deviations per residue between the simulated trajectories and their initial structures, averaged over the entire simulations. The \times symbols denote the cysteine residues.

Figure 5.8 shows atom-positional RMS deviations per residue from the initial structures averaged over the whole simulations. The difference between residues is remarkable. The cysteines (see the \times markers in Figure 5.8) have generally a low deviation. The chain ends deviate clearly more, as is often observed in proteins. However, the loops between the cysteines exhibit an extraordinary flexibility, especially the loops around Gly11 and around Gly17, without disrupting the geometry of the metal core. The increased flexibility is present in all simulations and all variants. This supports the hypothesis (see Reference [74] and citations therein) that the flexibility of the loops accounts for metallothionein's broad diversity in binding different metal ions.

It has recently been shown [78–84] that the GROMOS force field is capable of folding polypeptides into the correct secondary structure. The present work demonstrates that it is not a feature of the GROMOS force field to inevitably fold anything into a regular secondary structure: Metallothionein's variable loops remain irregular and flexible.

It is also appropriate to note that MNDO/d performs well on the metal cluster, despite the fact that it was not parametrised against zinc-sulphur or cadmium-sulfur bonds, neither against clusters involving several metal ions.

5.5 Conclusions

The classical MDc simulation of Cd₃ metallothionein-2 of rat liver in aqueous solution satisfies most of the NMR NOE data on this molecule; 8 atom-atom distance bounds out of a total of 46 NOE bounds are violated by more than 0.1 nm. The structure of the metal-sulphur cluster is approximately maintained.

The use of the semi-empirical method MNDO/d for the description of the metal core improved the simulated results. Specifically, bond lengths between the metal ions and the coordinated sulphur atoms are closer to experimental values. The MDc simulation yields too compact a metal core, thus affecting the whole protein. This is illustrated by the NOE bound violations. In the MDq simulation, in contrast to the MDc simulation, only one violation larger than 0.1 nm is observed.

The cost of the improved simulation is the increased computational expense. This is illustrated by Figures 5.4 and 5.5. Both the MDc and the MDq simulations ran simultaneously on the same type of computer. While the MDc simulations reached 8 ns, the MDq simulations only went shortly beyond 200 ps.

The MDc simulations indicate that the force-field parameters for zinc and cadmium would need improvement. Accidentally, the parameters derived for cadmium give bond lengths that would be appropriate for zinc. However, the current cadmium parameters are nevertheless unsuitable as improved parameters for zinc, as other zinc properties such as bond angles are not correctly reproduced by them. Moreover, a clear deficiency of a purely classical electrostatic and van-der-Waals description of the metal core could be located at the sulphur bridges, which tend to have too wide angles. Therefore, a proper description would require additional force-field terms for bond angles and maybe also bond lengths. However, it is not clear how these terms should look exactly, as the involved angles and bonds are hard to determine experimentally. The danger remains to choose “reasonable” parameters and, consequently, obtain “reasonable” results. In contrast, the treatment of the difficult core by quantum chemistry embedded in a classical environment provides an unbiased description, which is generalisable to other systems.

Chapter 6

Outlook

The work presented here has opened the doors to many new applications. A few examples are given here, as well as some ideas for the improvement of the presented methods.

6.1 Photoisomerisation of Stilbene

6.1.1 Photoisomerisation of *trans*-Stilbene

All the machinery is available as well as large portions of the potential energy surface: The simulation of the photoisomerisation of *trans*-stilbene is now straightforward. This reaction is not as fast as the photoisomerisation of *cis*-stilbene, because it starts off in a shallow minimum, and the isomerisation is expected to be a barrier crossing event with thermal activation. Nevertheless, the reaction has found an even greater echo in the literature than *cis*-stilbene, and many reaction dynamics theories, for example based on transition state theory or Rice-Ramsperger-Kassel-Marcus modeling, partly including quantum effects, have been derived and applied. It is desirable to complement the experimental and theoretical work with computational results.

6.1.2 Quantum Dynamics with Surface Hopping

The representation of a potential energy surface by finite elements is not restricted to a single surface. Two surfaces could be treated simultaneously. This would be an ideal field of application for the surface hopping method, a method for simulating quantum dynamics. It would be interesting to see to what extent the explicit inclusion of a deactivation mechanism changes the results of a classical simulation of an excited state.

6.1.3 Interpolation in More Dimensions

The interpolation method is not restricted to three dimensions as employed up to now. The scheme is generalisable to any dimensionality. Sticking to stilbene, a four-dimensional system could be obtained by unconstraining the central ethylenic bond length. Efficiency and differences to the three-dimensional treatment could be investigated.

As a technical improvement, it would be desirable to use a higher-order interpolation polynomial, which employs the full information of the energy and gradients. In the current three-dimensional implementation, the quadratic interpolation polynomial has ten coefficients, while

the information at the vertices would provide sufficient information for sixteen coefficients. A full cubic polynomial, however, requires twenty coefficients. It should be possible to construct a constrained cubic polynomial with sixteen coefficients. Such a polynomial should allow a more accurate interpolation, which in turn should allow larger elements to be used, which in turn would increase the efficiency of the method. However, stability and robustness have to be examined. First tests revealed that rotational invariance of the polynomial is important.

6.1.4 Another System

Another system of interest similar to stilbene is *azobenzene*. It has potential applications as an optical switch. For example, it can be embedded in a liquid-crystal polymer film, which enables reversible phase transitions induced by irradiating with light [85, 86].

6.2 Metallothionein

6.2.1 Other Metals in the β Domain

A next step is the investigation of the mercury variant of metallothionein. Mercury is also standardly available in MNDO/d. The complex with mercury is even more stable than with cadmium. It is interesting to see whether the conformational differences are comparable to those between the zinc and cadmium variants.

6.2.2 The α Domain

The same questions as addressed in Chapter 5 can also be investigated for the 4-metal α domain of metallothionein. The results are expected to be similar as for the β domain, however, the more complex cluster situation with two rings could raise unexpected problems.

Metallothionein is also known to bind twelve copper(I) ions. This complex is likely to be very different from the ones binding seven ions. The structure is not yet known experimentally. However, the computational prediction seems to be extremely difficult and challenging.

It would be possible to simulate homologues of rat liver metallothionein. NMR data is available for rabbit liver [77] and human metallothionein [87]. The structural and dynamic differences induced by amino acid substitution could be investigated.

6.2.3 Other Proteins

An interesting protein with catalytic activity due to a metal is liver alcohol dehydrogenase. Its active center contains a zinc ion, which serves to polarise the substrate. Besides the potential problems of the MNDO method to describe a reaction, the size of the protein (a dimer of two 374-residue chains, plus coenzyme NAD^+) raises a practical problem.

Appendix A

The fecomd Implementation

A.1 Features

The fecomd program is a simple molecular dynamics program which performs the finite element interpolation presented in Chapter 2. The name is an acronym of “finite element combined (quantum-chemical/classical) molecular dynamics”. It is programmed in C and partly C++. The mesh points are calculated by a separate quantum-chemical program, which is Gaussian 94 [45] currently, but any program could be used. Interfacing between the program is accomplished by a script which is called by the main fecomd driver. The script sets up an input file for Gaussian, starts it up and after it has finished, gathers the energy and the gradients from the output file. These results are fed back to the main fecomd program.

The fecomd program has the following features:

- Lennard-Jones interaction
- Coulomb interaction (implemented, but not actually used)
- interaction pair list
- bond constraints (SHAKE, [48])
- dihedral angle constraints [49]
- leap-frog time integration scheme [26]
- interpolation machinery (Chapter 2), any dimensionality, selectable at compile time (`#define DIMENSIO num`).
- Alternatively, a force-field-type potential energy surface can be selected by compile switches (`#define SIMPLEPOTENTIAL`). This is for testing purposes mainly.

As the fecomd is a further development from a program which used irregular elements for interpolation, it has many features in input and output which have become obsolete.

A.2 Input File

The fecomd program is controlled by a single input file. It has the following format.

Input file	Meaning
# from GROMOS96, 43A1: AR, C and HC	# introduces a comment line
atomtypes	keyword for atom definitions
Ar 39.948 0.996 0.341 0	Atom symbol, mass, LJ-epsilon, LJ-sigma, charge
C 12.011 0.40587 0.33611 0	
H 1.001 0.11838 0.23734 0	
solvent	keyword for solvent atom coordinates and velocities
Ar 1.375 1.618 0.280 0.224 0.231 0.230	Symbol, x, y, z coordinates, x, y, z velocities
...	
solute	keyword for solute atom coordinates and velocities
C -0.436 0.203 -0.717 0.023 -0.134 0.256	as above
...	
#!infile test.coo	alternatively, coordinates and velocities could be read from the named file
molecule	keyword for constraints and free dihedrals
0 1 .148	three parameters: distance constraint (shake)
12 13 .148	two atom numbers of solute, interatomic distance
...	
1 2 3 4 0.0 0.0510	six parameters: dihedral angle constraint
1 2 3 8 180.0 0.0510	four atom numbers of solute, dihedral angle, (force constant, obsolete, ignored, but must be present)
...	
2 1 0 12	four parameters: specification of free dihedrals
1 0 12 13	interpolation takes place in those dihedral space
0 12 13 14	atom numbers of solute
#!infile stilbene.topo	alternatively, the constraints can be read from the named file
shake	keyword for shake parameters
maxitera 1000	maximum of shake iterations
tolerance 1e-6	relative distance tolerance
box	keyword for box size definition
2.15 2.15 2.15	x, y, z box sizes
fes	keyword for the finite element system
maxedge 0.25	brick size, in radians
accuracy 10	for the master brick, should be high
maxvertex 50	for the master brick
maxsimplex 25	for the master brick
output	definition of output files and writing frequency
summary 10 test.sum	summary of energies, temperature, pressure etc.
solutetrajectory 1 test.itr	internal coordinate trajectory
alltraj 20 test.traj	Cartesian trajectory of all atoms
savegrid 0 test.fes	save finite-element-system at end of simulation
sim	definition of simulation parameters
timestep 0.001	timestep (picoseconds)
cutoff 0.9	interaction cutoff distance (nanometers)
tempref 348	reference temperature for coupling bath
temptau 0.1	coupling time (picoseconds)
!loadfes stilbene.fes	directive to read finite element data from named file
!startmd 5000	directive to start the MD simulations, so many steps

A.3 Output Files

Standard Output fecomd verbosely tells what it does. The standard output explains itself.

Summary File After an echo of input parameters and some internal information, the summary file contains the following data: time, total energy, total potential energy, total kinetic energy, temperature, virial and pressure.

```

*****
#*                               M D   S u m m a r y                               **
#*                               *****
#
# Initial temperature is 345.224 K
#-----
#Time ps Etot u(nm/ps)2 Epot u(nm/ps)2 Ekin u(nm/ps)2 Temp K  Virial u(nm/ps)2  Press. bar
#-----
#
  0.01      3393.53      -8507.34      11900.9      346.533      16584.7      922.276
  0.02      3395.19      -8573.37      11968.6      348.504      15811.1      907.701
  0.03      3391.45      -8630.19      12021.6      350.05      14722.8      885.273
...

```

Internal Coordinates Trajectory This file contains the time, the internal coordinates (dihedral angles in degrees), the solute's potential energy, and a word indicating whether the point has been interpolated or explicitly calculated (obsolete now, always interpolated).

```

0.001      43.4998      4.49996      43.5003      85.3759      (fepol)
0.002      43.4442      4.57468      43.4434      85.2453      (fepol)
0.003      43.3339      4.72273      43.3311      84.9866      (fepol)
...

```

Cartesian Trajectory This file has the same format as the solvent and solute entries in the input file. It is therefore suitable to be read in with the `!infile` directive.

Finite-Element-System File This file contains all available information of the finite-element grid in binary form. Explicitly quantum-chemically calculated points, their energies and gradients are reused over many trajectories via this file. See file `output.c` functions `SaveBrickSystem()` and `LoadBrickSystem()` on how this file is written and read.

A.4 Auxiliary Programs

fesconvert Reads a binary finite-element-system file and converts the finite element system to a smaller grid size. The grid is subdivided into a selectable integer number of elements in every dimension. Existing grid points are maintained. Outputs a finite-element-system file with a smaller grid size.

scanfes Reads a binary finite-element-system file, calculates a two-dimensional cut and outputs a PLOT3D file suitable for visualising the potential energy surface using Maple V.

accuracy Checks the accuracy of the interpolation when an analytic surface is used (`#define SIMPLEPOTENTIAL`). Reads in the output file of fecomd.

analysitr Analyses the internal coordinate trajectory (`*.itr`) in terms of grid point usage.

interactive-fes Reads a binary finite-element-system file and interpolates points which are interactively entered. For testing purposes mainly.

optimize-fes Reads a binary finite-element-system file and calculates the location of minima and saddlepoints.

calc-gradients-for-itr Reads a binary finite-element-system file and an internal coordinate trajectory (`*.itr`) and calculates the intramolecular forces that occurred along the trajectory.

Appendix B

The zumos Implementation

zumos is an acronym for 'Zurich molecular simulation' in honour of its two main components, GROMOS96 (Wilfred F. van Gunsteren) for the classical force field and dynamics and MNDO97 (Walter Thiel) for the semiempirical models, which were both developed in Zürich. It is an interface which combines the two programs. The combination is achieved at link level, producing a single monolithic program, and not via scripts, as many other approaches do. It establishes the combination of programs through an interface subroutine with minimum contacts to either program. Basically, it needs one subroutine call from GROMOS to the interface, which in turn calls one subroutine of MNDO. There are some additional subroutine calls at startup for initialisation.

The basic operation of the interface is:

1. it obtains atomic data from GROMOS, primarily atomic positions
2. it sets up link atoms and external charges
3. it passes control to MNDO, which calculates the energy and the gradients
4. it sums up these gradients and energies into the corresponding GROMOS arrays, and passes control back to GROMOS.

Thus the main driver is GROMOS and its input files are used. The MNDO program is controlled by a standard MNDO input file without molecular specification, which is to be placed at Fortran unit 80. See the MNDO input-file specification for details. Switches essential for zumos operation are overridden by zumos at startup.

B.1 The Quantum Topology

The quantum subsystem is specified by a quantum topology. It is usually derived from a standard GROMOS molecular topology. PROQMT is a script which does an automated conversion according to the quantum topology specifications. It performs the following steps:

- it reads a file containing sequence numbers of atoms to quantize to the specified atomic number.
- for all quantum atoms, it sets the partial charge to zero, avoiding classical Coulomb interaction with quantum atoms. The letters QQ are appended to the atom name.

- it mutually excludes all quantum atoms from the pair-list, avoiding Lennard-Jones interactions between quantum atoms.
- it sets up a charge group for every quantum atom. This feature is needed by the subroutine that dynamically finds neighbours of the quantum system. Here is an example of a changed entry within the SOLUTEATOM block. The original entry is commented out.

```
# was:      43      5 CB      11      12.01100      -0.10000      0      4      44      45      46      47
# automatically quantized atom follows
43 5 CBQQ 11 12.01100 0.00 1 39 44 45 46 47 60 61 62 63 106 107 108
109 123 124 125 126 151 152 153 154 173 174 175 176 207 208 209 210
229 230 231 232 255 256 257 258 275 276 277
```

- it removes all bonded interactions that involve quantum atoms and at most one classical atom.
- for all bonds involving both quantum and classical atoms, a link atom is set up
- it inserts a QIFACESPEC block. It has the following template structure:

Topology file entry	Meaning
QIFACESPEC	block name
39	number of quantum atoms (NRQQAT)
43 6	sequence number of quantum atom, atomic number
44 1	INQQAT[1-NRQQAT], ITQQTY[1-NRQQAT]
...	
277 30	
0	number of permanent interface atoms (NRQIAT), only
# 279	sequence numbers of permanent interface atoms
# 280	INQIAT[1-NRQIAT]
# 271	
9	number of link atoms (NRQLAT)
43 42 1 0.6948	sequence number of join atom, sequence number of connect atom,
60 59 1 0.6948	link atom type (ignored at the moment), bond length ratio
...	INQJAT[1-NRQLAT], INQCAT[1-NRQLAT],
255 254 1 0.6948	ITQLTY[1-NRQLAT], RLQLBO[1-NRQLAT]
END	end marker

Things that should be avoided:

- cut through a charge group
- cut through a very polar bond or a multiple bond
- quantum atoms on both sides of a remaining bonded interaction

B.2 Running zumos

zumos is most easily run by a driver scripts which sets up all the required files. The topology files, the initial coordinates and other GROMOS input files are linked to the fortran units as specified by GROMOS96. The MNDO input file, without molecular specification, is linked to fortran unit

80. The PROMD input is fed in as standard input to zumos. In the following example driver script, this is done via a Unix “here document”. If a density-matrix file or an eigenvalues file is to be read and/or written, these files are to be linked to fortran units 81 and 82 respectively. This is especially useful for keeping the density matrix across program restarts, as MNDO’s standard initial guess is sometimes not suitable for unusual geometries.

```
#!/bin/sh

# where the zumos executable is located
BINDIR=zumos/bin

# select the machine-specific executable
PROGRAM=${BINDIR}/zumos.`uname -i`

# specify topology label, secondary type label and serial number of simulation
TYPE='mtblq'
SECO="-s2-d"
SERIAL=2

NEXT=`expr $SERIAL + 1`

# construct file names
TOPOIN=${TYPE}.topo
COORIN=${TYPE}${SECO}-sx-${SERIAL}.coo

OUTPUT=ozd-${TYPE}${SECO}-${SERIAL}.out
TRAJOUT=${TYPE}${SECO}-rx-${SERIAL}.coo
COOROUT=${TYPE}${SECO}-sx-${NEXT}.coo

# avoid messing up fortran units when several programs run
ODIR=`pwd`
TMPDIR=/scrloc/cdb/sub-`uname -n`-$$
mkdir -p $TMPDIR
cd $TMPDIR

# link to the fortran units
rm -f fort.*
#---input units
ln -s ${ODIR}/${TOPOIN} fort.20 # molecular (quantum) topology
ln -s ${ODIR}/${COORIN} fort.21 # initial coordinates

#---output units
ln -s ${ODIR}/${TRAJOUT} fort.12 # trajectory
ln -s ${ODIR}/${COOROUT} fort.11 # final coordinates

# copy the density matrix file from the previous simulation
# to fortran unit 81. It is used as initial guess.
cp ${ODIR}/densmat-${TYPE}${SECO}-${SERIAL}.bdt fort.81

# mndo input, without molecular specification (will be added by the zumos program)
cat << ! > fort.80
  0 -10 -2 1 0 0 0 0-1 0 0 2 0 0 0 0 9 6 6 0 0 1 0 0
  0 0 066 0-1-5-5-1 2
-3 0 1 0 1 0 280-100 0 0 0 0 3-Zn-Cluster mit Methanthiol MNDO/d
!
```

```
# extract number of solvent molecules and total number of atoms from
# the PROBOX output which was used to fill the current box.
NUMSOLVMOL=`grep 'GENERATED SOLVENT MOLECULES' $ODIR/obox.out | sed 's/^.* =//g'`
NUMTOTATOM=`grep 'TOTAL NUMBER OF ATOMS' $ODIR/obox.out | sed 's/^.* =//g'`

echo "$NUMSOLVMOL solvent molecules, $NUMTOTATOM atoms in total"

#---run the program, feed promd input via here document
if time $PROGRAM << ! > $OUTPUT; then
TITLE
Metallothionein beta domain Zn3, example input
END
SYSTEM
#      NPM      NSM
#      1      $NUMSOLVMOL
END
... (rest of standard PROMD input)
!
# save the current density matrix file for future use
cp fort.81 ${ODIR}/densmat-${TYPE}${SECO}-${NEXT}.bdt
cd $ODIR
rm -rf $TMPDIR # clean up.
echo "normal termination of zumos."
else
echo "zumos did not terminate normally!"
fi
```

Bibliography

- [1] A. Warshel and M. Levitt. “Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme”. *J. Mol. Biol.*, **103**, (1976) 227–249.
- [2] M. J. Field, P. A. Bash, and M. Karplus. “A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations”. *J. Comput. Chem.*, **11**, (1990) 700–733.
- [3] D. Bakowies and W. Thiel. “Hybrid models for combined quantum mechanical and molecular mechanical approaches”. *J. Phys. Chem.*, **100**, (1996) 10580–10594.
- [4] D. Bakowies. “Hybridmodelle zur Kopplung quantenmechanischer und molekülmechanischer Verfahren”. Ph.D. thesis, Universität Zürich, (1994).
- [5] R. V. Stanton, D. S. Hartsough, and K. M. Merz Jr. “An examination of a density functional/molecular mechanical coupled potential”. *J. Comput. Chem.*, **16**, (1995) 113–128.
- [6] H. Liu, F. Müller-Plathe, and W. F. van Gunsteren. “A molecular dynamics simulation study with a combined quantum mechanical and molecular mechanical potential energy function: Solvation effect on the conformational equilibrium of dimethoxyethane”. *J. Chem. Phys.*, **102**, (1995) 1722–1730.
- [7] M. Freindorf and J. Gao. “Optimization of the Lennard-Jones parameters for a combined *ab initio* quantum mechanical and molecular mechanical potential using the 3-21G basis set”. *J. Comput. Chem.*, **17**, (1996) 386–395.
- [8] T. Helgaker, E. Uggerud, and H. J. A. Jensen. “Integration of the classical equations of motion on *ab initio* molecular potential energy surfaces using gradients and Hessians: application to translational energy release upon fragmentation”. *Chem. Phys. Lett.*, **173**, (1990) 145–150.
- [9] J. Ischtwan and M. A. Collins. “Molecular potential energy surfaces by interpolation”. *J. Chem. Phys.*, **100**, (1994) 8080–8088.
- [10] K. C. Thompson, M. J. T. Jordan, and M. A. Collins. “Molecular potential energy surfaces by interpolation in Cartesian coordinates”. *J. Chem. Phys.*, **108**, (1998) 564–578.
- [11] F. Eckert and H.-J. Werner. “Local interpolation of *ab initio* potential energy surfaces for direct dynamics studies of chemical reactions”. *Chem. Phys. Lett.*, **302**, (1999) 208–214.

- [12] D. H. Waldeck. "Photoisomerization dynamics of stilbenes". *Chem. Rev.*, **91**, (1991) 415–436.
- [13] S. Abrash, S. Repinec, and R. M. Hochstrasser. "The viscosity dependence and reaction coordinate for isomerization of *cis*-stilbene". *J. Chem. Phys.*, **93**, (1990) 1041–1053.
- [14] L. Nikowa, D. Schwarzer, and J. Troe. "Viscosity and solvent dependence of low-barrier processes: Photoisomerization of *cis*-stilbene in compressed liquid solvents". *J. Chem. Phys.*, **97**, (1992) 4827–4835.
- [15] J. K. Rice and A. P. Baronavski. "Ultrafast studies of solvent effects in the isomerization of *cis*-stilbene". *J. Phys. Chem.*, **96**, (1992) 3359–3366.
- [16] A. Z. Szarka, N. Pugliano, D. K. Palit, and R. M. Hochstrasser. "Vibrational coherence in the solution phase photoisomerization reaction of *cis*-stilbene". *Chem. Phys. Lett.*, **240**, (1995) 25–30.
- [17] D. L. Phillips, J.-M. Rodier, and A. B. Myers. "Cis-stilbene photochemistry: Direct observation of product formation and relaxation through two-color UV pump-probe Raman spectroscopy". *Chem. Phys.*, **175**, (1993) 1–12.
- [18] S. J. Lippard and J. J. Berg. *Principles of Bioinorganic Chemistry* (University Science Books, Mill Valley CA, 1994).
- [19] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. *Biomolecular Simulation: The GROMOS96 Manual and User Guide* (vdf Hochschulverlag AG an der ETH Zürich, Zürich, 1996).
- [20] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry* (Dover Publications Inc., Mineola NY, 1996), Dover ed.
- [21] J. A. Pople, R. Krishnan, H. B. Schlegel, and L. S. Binkley. "Derivative studies in Hartree-Fock and Møller-Plesset theories". *Int. J. Quant. Chem. Symp.*, **13**, (1979) 225–241.
- [22] J. B. Foresman, M. Head-Gordon, and J. A. Pople. "Toward a systematic molecular orbital theory for excited states". *J. Chem. Phys.*, **96**, (1992) 135–149.
- [23] M. J. S. Dewar and W. Thiel. "Ground states of molecules. 38. The MNDO method. Approximations and parameters". *J. Am. Chem. Soc.*, **99**, (1977) 4899–4917.
- [24] W. Thiel and A. A. Voityuk. "Extension of the MNDO formalism to *d* orbitals: Integral approximations and preliminary numerical results". *Theor. Chim. Acta.*, **81**, (1992) 391–404.
- [25] W. Thiel and A. A. Voityuk. "Extension of MNDO to *d* orbitals: Parameters and results for the second-row elements and for the zinc group". *J. Phys. Chem.*, **100**, (1996) 616–626.
- [26] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
- [27] H. R. Christen. *Grundlagen der allgemeinen und anorganischen Chemie* (Salle + Sauerländer, Frankfurt am Main, 1988), 9th ed.

- [28] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello. “A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields”. *J. Chem. Phys.*, **110**, (1999) 10452–10467.
- [29] I. S. Antes. “Combined quantum mechanical and molecular mechanical methods: From link atoms to adjusted connection atoms”. Ph.D. thesis, Universität Zürich, (1998).
- [30] J. A. Syage, W. R. Lambert, P. M. Felker, A. H. Zewail, and R. M. Hochstrasser. “Picosecond excitation and *trans-cis* isomerization of stilbene in a supersonic jet: Dynamics and spectra”. *Chem. Phys. Lett.*, **88**, (1982) 266–270.
- [31] J. A. Syage, P. M. Felker, and A. H. Zewail. “Picosecond dynamics and photoisomerization of stilbene in supersonic beams. II. Reaction rates and potential energy surface”. *J. Chem. Phys.*, **81**, (1984) 4706–4723.
- [32] B. I. Greene and R. C. Farrow. “Subpicosecond time resolved multiphoton ionization: excited state dynamics of *cis*-stilbene under collision free conditions”. *J. Chem. Phys.*, **78**, (1983) 3336–3338.
- [33] J. Schroeder, D. Schwarzer, J. Troe, and F. Voß. “Cluster and barrier effects in the temperature and pressure dependence of the photoisomerization of *trans*-stilbene”. *J. Chem. Phys.*, **93**, (1990) 2393–2404.
- [34] D. C. Todd and G. R. Fleming. “*Cis*-stilbene isomerization: Temperature dependence and the role of mechanical friction”. *J. Chem. Phys.*, **98**, (1993) 269–279.
- [35] L. Nikowa, D. Schwarzer, J. Troe, and J. Schroeder. “Photoisomerization of *cis*-stilbene in compressed solvents”. *Springer Ser. Chem. Phys.*, **55**, (1993) 603–605.
- [36] J. Schroeder, D. Schwarzer, J. Troe, and P. Vöhringer. “From barrier crossing to barrierless relaxation dynamics. Photoisomerization of *trans*-stilbene in compressed *n*-alkanols”. *Chem. Phys. Lett.*, **218**, (1994) 43–50.
- [37] J. Schroeder, J. Troe, and P. Vöhringer. “Photoisomerization of *trans*-stilbene in compressed solvents: Kramers-turnover and solvent induced barrier shift”. *Z. Phys. Chem.*, **188**, (1995) 287–306.
- [38] J. S. Baskin, L. Bañares, S. Pedersen, and A. H. Zewail. “Femtosecond real-time probing of reactions. 20. Dynamics of twisting, alignment, and IVR in the *trans*-stilbene isomerization reaction”. *J. Phys. Chem.*, **100**, (1996) 11920–11933.
- [39] V. D. Vachev, J. H. Frederick, B. A. Grishanin, V. N. Zadkov, and N. I. Koroteev. “Stilbene isomerization dynamics on multidimensional potential energy surface. Molecular dynamics simulation”. *Chem. Phys. Lett.*, **215**, (1993) 306–314.
- [40] V. D. Vachev, J. H. Frederick, B. A. Grishanin, V. N. Zadkov, and N. I. Koroteev. “Quasi-classical molecular dynamics simulations of the photoisomerization of stilbene”. *J. Phys. Chem.*, **99**, (1995) 5247–3497.
- [41] K. Bolton and S. Nordholm. “A classical molecular dynamics study of the intramolecular energy transfer of model *trans*-stilbene”. *Chem. Phys.*, **203**, (1996) 101–126.

- [42] G. Gershinsky and E. Pollak. "Theoretical study of the *trans*-stilbene isomerization reaction in ethane". *J. Chem. Phys.*, **105**, (1996) 4388–4390.
- [43] C. D. Berweger, W. F. van Gunsteren, and F. Müller-Plathe. "Finite element interpolation for combined classical / quantum-mechanical molecular dynamics simulations". *J. Comput. Chem.*, **18**, (1997) 1484–1495.
- [44] C. D. Berweger. "Molecular dynamics with regular grid finite element interpolation". Unpublished results.
- [45] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople. *Gaussian 94, Revision C.3*. Gaussian, Inc., Pittsburgh PA, (1995).
- [46] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren. "The GROMOS biomolecular simulation program package". *J. Phys. Chem. A*, **103**, (1999) 3596–3607.
- [47] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. "Molecular dynamics with coupling to an external bath". *J. Chem. Phys.*, **81**, (1984) 3684–3690.
- [48] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. "Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes". *J. Comput. Phys.*, **23**, (1977) 327–341.
- [49] D. J. Tobias and C. L. Brooks III. "Molecular dynamics with internal coordinate constraints". *J. Chem. Phys.*, **89**, (1988) 5115–5127.
- [50] M. Traetteberg and E. B. Frantsen. "A gas electron diffraction study of the molecular structure of *cis*-stilbene". *J. Mol. Struct.*, **26**, (1975) 69–76.
- [51] J. Schroeder, J. Troe, and P. Vöhringer. "Pressure dependence of solvent-induced barrier shifts in the photoisomerization of *trans*-stilbene". *Chem. Phys. Lett.*, **203**, (1993) 255–260.
- [52] H. Liu, F. Müller-Plathe, and W. F. van Gunsteren. "Molecular dynamics with a quantum-chemical potential: Solvent effects on an S_N2 reaction at nitrogen". *Chem. Eur. J.*, **2**, (1996) 191–195.
- [53] C. D. Berweger, F. Müller-Plathe, and W. F. van Gunsteren. "Molecular dynamics simulation with an *ab initio* potential energy function and finite element interpolation: The photoisomerisation of *cis*-stilbene in solution". *J. Chem. Phys.*, **108**, (1998) 8773–8781.
- [54] M. Dantus, M. J. Rosker, and A. H. Zewail. "Real-time femtosecond probing of 'transition states' in chemical reactions". *J. Chem. Phys.*, **87**, (1987) 2395.

- [55] M. J. Rosker, M. Dantus, and A. H. Zewail. “Femtosecond clocking of the chemical bond”. *Science*, **241**, (1988) 1200.
- [56] G. Gershinsky and E. Pollak. “Unimolecular reactions in the gas and liquid phase: A possible resolution to the puzzles of the *trans*-stilbene isomerization”. *J. Chem. Phys.*, **107**, (1997) 812–824.
- [57] E. Pollak, P. Talkner, and A. M. Berezhkovskii. “A theory for nonisothermal unimolecular reaction rates”. *J. Chem. Phys.*, **107**, (1997) 3542–3549.
- [58] G. Gershinsky and E. Pollak. “Isomerization of stilbene in the gas phase: Theoretical study of isotopic and clustering effects”. *J. Chem. Phys.*, **107**, (1997) 10532–10538.
- [59] G. Gershinsky and E. Pollak. “Quantum harmonic transition state theory — application to isomerization of stilbene in liquid ethane”. *J. Chem. Phys.*, **108**, (1998) 2756–2764.
- [60] G. Orlandi, P. Palmieri, and G. Poggi. “An ab initio study of the *cis-trans* photoisomerization of stilbene”. *J. Am. Chem. Soc.*, **101**, (1979) 3492–3497.
- [61] J. Troe and K.-M. Weitzel. “MNDO calculations of stilbene potential energy properties relevant for the photoisomerization dynamics”. *J. Chem. Phys.*, **88**, (1988) 7030–7039.
- [62] J. H. Frederick, Y. Fujiwara, J. H. Penn, K. Yoshihara, and H. Petek. “Models for stilbene photoisomerization: Experimental and theoretical studies of the excited-state dynamics of 1,2-diphenylcycloalkenes”. *J. Phys. Chem.*, **95**, (1991) 2845–2858.
- [63] Y. Amatatsu. “Ab initio CI study on electronically excited stilbene”. *THEOCHEM J. Mol. Struct.*, **461–462**, (1999) 311–316.
- [64] D. C. Todd, J. M. Jean, S. J. Rosenthal, A. J. Ruggiero, D. Yang, and G. R. Fleming. “Fluorescence upconversion study of *cis*-stilbene isomerization”. *J. Chem. Phys.*, **93**, (1990) 8658–8668.
- [65] S. T. Repinec, R. J. Sension, and R. M. Hochstrasser. “Femtosecond studies of the photoisomerization of *cis*-stilbene in solution”. *Ber. Bunsenges. Phys. Chem.*, **95**, (1991) 248–252.
- [66] J. Saltiel. “Perdeuteriostilbene. The role of phantom states in the *cis-trans* photoisomerization of stilbenes”. *J. Am. Chem. Soc.*, **89**, (1967) 1036–1037.
- [67] J. Saltiel. “Perdeuteriostilbene. The triplet and singlet paths for stilbene photoisomerization”. *J. Am. Chem. Soc.*, **90**, (1968) 6394–6400.
- [68] F. Müller-Plathe. “Reversing the perturbation in non-equilibrium molecular dynamics: An easy way to calculate the shear viscosity of fluids”. *Phys. Rev. E*, **59**, (1999) 4894–4899.
- [69] R. Walser, A. E. Mark, and W. F. van Gunsteren. “On the validity of Stokes’ law at the molecular level”. *Chem. Phys. Lett.*, **303**, (1999) 583–586.
- [70] A. B. Myers and R. A. Mathies. “Excited-state torsional dynamics of *cis*-stilbene from resonance Raman intensities”. *J. Chem. Phys.*, **81**, (1984) 1552–1558.

- [71] P. Schultze, E. Wörgötter, W. Braun, G. Wagner, M. Vašák, J. H. R. Kägi, and K. Wüthrich. “Conformation of [Cd₇]-metallothionein-2 from rat liver in aqueous solution determined by nuclear magnetic resonance spectroscopy”. *J. Mol. Biol.*, **203**, (1988) 251–268.
- [72] A. H. Robbins, D. E. McRee, M. Williamson, S. A. Collett, N. H. Xuong, W. F. Furey, B. C. Wang, and C. D. Stout. “Refined crystal structure of Cd, Zn metallothionein at 2.0 Å resolution”. *J. Mol. Biol.*, **211**, (1991) 1269–1293.
- [73] W. F. Furey, A. H. Robbins, L. L. Clancy, D. R. Winge, B. C. Wang, and C. D. Stout. “Crystal structure of Cd, Zn metallothionein”. *Science*, **231**, (1986) 704–710.
- [74] W. Braun, M. Vašák, A. H. Robbins, C. D. Stout, G. Wagner, J. H. R. Kägi, and K. Wüthrich. “Comparison of the NMR solution structure and the X-ray crystal structure of rat metallothionein-2”. *Proc. Natl. Acad. Sci.*, **89**, (1992) 10124–10128.
- [75] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. “Interaction models for water in relation to protein hydration”. In: “Intermolecular Forces”, (B. Pullman, ed.) (Reidel, Dordrecht, 1981) 331–342, 331–342.
- [76] P. E. Smith and W. F. van Gunsteren. “Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K”. *J. Chem. Phys.*, **100**, (1994) 3169–3174.
- [77] A. Arseniev, P. Schultze, E. Wörgötter, W. Braun, G. Wagner, M. Vašák, J. H. R. Kägi, and K. Wüthrich. “Three-dimensional structure of rabbit liver [Cd₇]metallothionein-2a in aqueous solution determined by nuclear magnetic resonance”. *J. Mol. Biol.*, **201**, (1988) 637–657.
- [78] X. Daura, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark. “Reversible peptide folding in solution by molecular dynamics simulations”. *J. Mol. Biol.*, **280**, (1998) 925–932.
- [79] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark. “Peptide folding: When simulation meets experiment”. *Angew. Chem. Int. Ed.*, **38**, (1999) 236–240.
- [80] X. Daura, W. F. van Gunsteren, and A. E. Mark. “Folding-unfolding thermodynamics of a β -heptapeptide from equilibrium simulations”. *Proteins: Structure, Function and Genetics*, **34**, (1999) 269–280.
- [81] A. M. J. J. Bonvin and W. F. van Gunsteren. “ β -hairpin stability and folding: Molecular dynamics studies of the first β -hairpin of tendamistat”. *J. Mol. Biol.*, (2000) in press.
- [82] X. Daura, A. E. Mark, and W. F. van Gunsteren. “Peptide folding simulations: No solvent required?” *Comput. Phys. Commun.*, (1999) in press.
- [83] R. Bürigi, X. Daura, A. Mark, M. Bellanda, S. Mammi, E. Peggion, and W. F. van Gunsteren. “Folding study of an Aib-rich peptide in DMSO by molecular dynamics simulations”. *J. Mol. Biol.*, (1999) submitted.
- [84] C. Peter. Personal communication, (1999).

-
- [85] T. Ikeda and O. Tsutsumi. “Optical switching and image storage by means of azobenzene liquid-crystal films”. *Science*, **268**, (1995) 1873–1875.
- [86] S. Y. Oh, S. M. Hong, and S. I. Oh. “Synthesis and properties of photoresponsive polyacetylene derivatives with an azobenzene mesogenic moiety in the side group”. *Mol. Cryst. Liq. Cryst. A*, **294**, (1997) 403–406.
- [87] B. A. Messerle, A. Schäffer, M. Vašák, J. H. R. Kägi, and K. Wüthrich. “Three-dimensional structure of human [$^{113}\text{Cd}_7$]metallothionein-2 in solution determined by nuclear magnetic resonance spectroscopy”. *J. Mol. Biol.*, **214**, (1990) 765–779.

Curriculum Vitae

Personal Information

Name Christian Daniel Berweger
Date of birth December 28, 1971
Place of birth Pfäffikon ZH, Switzerland
Citizenship Herisau, Appenzell-Ausserrhoden, Switzerland

Education

1979–1987 Primary and Secondary Schools in Russikon
1987–1991 Mathematisch-Naturwissenschaftliches Gymnasium in Wetzikon
1991 Matura Typus C
1991–1995 Study of Chemistry at ETH Zürich
1995 Dipl. Chem. ETH
1995–1999 Ph. D. thesis at the Laboratory of Physical Chemistry at ETH Zürich,
Prof. Wilfred F. van Gunsteren