

Development of an Efficient Method to
Simulate Proton Transport (Q-HOP MD)

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Johann Wolfgang Goethe - Universität
in Frankfurt am Main

von
Markus Alexander Lill
aus Frankfurt am Main

Frankfurt 2002

vom Fachbereich Physik der
Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan: Prof. Dr. Dr. h.c. mult. Walter Greiner.

Gutachter: Prof. Dr. Werner Mäntele
Prof. Dr. Joachim Maruhn

Datum der Disputation: 12. Februar 2002

Zusammenfassung

Protonentransfer-Reaktionen sind Bestandteil vieler enzymatischer und bioenergetischer Prozesse und damit von essentieller Bedeutung für die meisten biologischen Systeme. Die Protonen bewegen sich dabei nicht als freie Teilchen, sondern sind stets an eine chemische Gruppe im Protein gebunden. Beim Transfer findet eine Säure-Base-Reaktion statt, das Proton wandert vom Donator zum Akzeptor. Bereits existierende quantenmechanische Simulationstechniken erfordern einen hohen numerischen Aufwand, weshalb heutzutage nur kleine Systeme (bis zu hundert Atome) über wenige Pikosekunden simuliert werden können. Die Zeitskala für den Protonentransport in Proteinen, die viele Tausende von Atomen enthalten, erstreckt sich jedoch vom Pikosekunden- bis in den Millisekundenbereich.

In Rahmen der hier vorliegenden Arbeit wurde deshalb eine neue Methode entwickelt, um Protonentransport in Proteinen effizient über Zeitintervalle von einigen Nanosekunden simulieren zu können. In einem vorgegebenen Protonierungszustand, in dem alle Protonen fest an chemische Gruppen des Proteins gebunden sind, wird die Dynamik des Proteins durch klassische Moleküldynamik-Simulationen (MD) beschrieben. In dieser sehr effizienten Simulationsmethode werden die klassischen Punktteilchen gemäss den Newtonschen Bewegungsgleichungen verschoben, wobei der Kraftberechnung zwischen den Teilchen ein empirisches Kraftfeld zugrundeliegt. Das Kraftfeld beinhaltet Coulomb- und Van-der-Waals-Wechselwirkungsterme sowie harmonische Potentialterme, die die kovalenten Bindungen des Moleküls beschreiben. Der Protonentransferschritt wird durch instantanes Protonhüpfen (“Proton Hopping”) beschrieben. Nach einem bestimmten Zeitintervall (üblicherweise 10 fs) wird die Transferwahrscheinlichkeit zwischen allen möglichen Donator-Akzeptor-Paaren bestimmt. Eine Zufallszahl zwischen 0 und 1 wird mit Hilfe eines Zufallszahlenalgorithmus numerisch berechnet und mit der Transferwahrscheinlichkeit verglichen. Ist die Zufallszahl kleiner als die Transferwahrscheinlichkeit, wird das Proton instantan vom Donator zum Akzeptor verschoben. In der Realität geht dieses Protonhüpfen mit dem Aufbrechen und der Neubildung einer kovalenten Bindung einher. In der klassischen Kraftfeld-Beschreibung des Proteins werden jedoch alle kovalenten Bindungen durch harmonische Potentiale beschrieben, so dass keine Bindungen aufgebrochen oder neu gebildet werden können. In der hier vorgestellten Methode wird daher die Proteintopologie, welche die Kon-

nektivität aller Atome des Proteins beschreibt, instantan verändert. Dabei wird die Donatorgruppe instantan deprotoniert, die Akzeptorgruppe protoniert, und die entsprechenden Kraftfeldparameter wie atomare Partiaalladungen angepasst.

Ein wesentlicher Teil der hier vorliegenden Arbeit bestand darin, eine sehr effiziente Vorgehensweise zu entwickeln, um während der Simulation schnell und einfach die Transferwahrscheinlichkeit zwischen allen Donator-Akzeptor-Paaren für jede beliebige Proteinkonfiguration zu berechnen. Zur Bestimmung der Protontransferwahrscheinlichkeiten wurden die eindimensionalen Energiehyperflächen als Funktion des Abstandes zwischen Proton und Donatoratom mit Hilfe zeitunabhängiger quantenmechanischer Rechnungen bestimmt. Hierbei wurden die aus Donator- und Akzeptorgruppe bestehenden Systeme um zusätzliche kleine Molekülgruppen um das Donator-Akzeptor-System ergänzt, um Umgebungseinflüsse auf den Transferprozess wiederzugeben. Die Umgebungseinflüsse sind dabei hauptsächlich elektrostatischer Natur. Es wird gezeigt, dass es zur Berechnung der eindimensionalen Energiehyperflächen des Protonentransfers ausreicht, ausschließlich Donator- und Akzeptorgruppe quantenmechanisch zu behandeln, und die umgebenden chemischen Gruppen als klassische Punktteilchen mit Partiaalladungen zu betrachten. Der Einfluss dieser umgebenden Gruppen auf Donator und Akzeptor lässt sich mit Hilfe reiner Coulomb-Wechselwirkung berechnen.

In der klassischen, auf einem Kraftfeld basierenden, Beschreibung des Proteins, werden die Atome üblicherweise durch klassische Punktteilchen mit unveränderlichen Partiaalladungen beschrieben. Des weiteren wurde ein Kraftfeld getestet, welches ausser den festen Partiaalladungen induzierbare Dipole auf den Punktteilchen enthält. Für den Protonentransfer zwischen zwei Wassermolekülen mit vielen umgebenden Wassermolekülen ergab sich jedoch kein Unterschied in der Genauigkeit der Beschreibung der Umgebungseinflüsse.

Nach Auswertung einer bedeutenden Anzahl von quantenmechanischen Rechnungen ergab sich, dass die Protonentransferenergiekurven hauptsächlich von drei Faktoren abhängig sind: einerseits vom Abstand zwischen Donator und Akzeptor, $R(DA)$, und andererseits von der Protonenaffinität von Donator und Akzeptor und der relativen elektrostatischen Stabilisierung von donator- bzw. akzeptorgebundenem Proton, welche sich zu einer Energiedifferenz, E_{12} , zwischen beiden Zuständen summieren. Sofern eine ausreichende Energiebarriere vorhanden ist, lässt sich zur Berechnung der Transferrate die "Transition State Theory" anwenden, wobei sich die Energiebarriere

re als einfache Funktion der drei genannten Faktoren ausdrücken lässt. Zuerst wurde die Energiedifferenz ohne Umgebungseinflüsse, E_{12}^0 , als einfache quadratische Funktion von $R(\text{DA})$ formuliert. Die drei Parameter, welche in die Funktion eingehen, beschreiben die unterschiedliche Protonaffinität von Akzeptor und Donator. Die Umgebungseinflüsse des umgebenden Proteins werden durch Berechnung der Coulombenergie in donierendem und akzeptierendem Zustand berücksichtigt. Durch Addition zu E_{12}^0 ergibt sich die vollständige Energiedifferenz zwischen akzeptierendem und donierendem Zustand, E_{12} . Mit Hilfe von weiteren sieben Parametern kann nun die Energiebarriere als Funktion von E_{12} und $R(\text{DA})$ bestimmt werden. Quantenmechanische Einflüsse des Protons wurden anschließend durch Nullpunktsenergien im donatorgebundenen Zustand sowie durch Tunnelkorrekturen auf semiklassischem Niveau berücksichtigt, wobei die Näherung von Wentzel-Kramers-Brillouin (WKB) verwendet wurde. Auch hierfür wurden einfache mathematische Ausdrücke entwickelt, um diese quantenmechanischen Einflüsse sehr schnell während der Simulation berechnen zu können. Der Einfluß der Tunnelkorrektur ist dabei stark temperaturabhängig.

Zum anderen kommen sowohl in Simulationen von Flüssigkeiten als auch in Proteinen häufig Situationen vor, bei denen das Proton barriereelos transferiert werden kann. Dies tritt insbesondere dann auf, wenn Donator und Akzeptor einen sehr kleinen Abstand zueinander einnehmen ($R(\text{DA}) \approx 2.4 - 2.7 \text{ \AA}$), und die Energiedifferenz E_{12} nicht zu gross ist ($E_{12} < 20 \text{ kcal/mol}$). Für diese Fälle wurde die eindimensionale, zeitabhängige Schrödingergleichung numerisch in dem den Protonentransfer beschreibenden Potential gelöst. Ein Wellenpaket, welches das zu transferierende Proton beschreibt, wird zu Beginn in das Minimum des Potentials positioniert, welches den donierenden Zustand beschreibt. In einem Zeitintervall von typischerweise 10 - 30 fs, bewegt sich das Wellenpaket zum Akzeptorzustand. Deshalb wird in der klassischen Simulation alle 10 fs abgefragt, ob ein Protonentransfer stattfindet. Der Anteil des Wellenpaketes, der sich nach einem Zeitintervall von 10 fs auf der Akzeptorseite befindet, wird als Transferrate definiert. Die so berechneten Transferraten wurden wiederum als Funktion von $R(\text{DA})$ und E_{12} ausgedrückt.

Während "transition state theory" für große Werte von $R(\text{DA})$ und E_{12} gültig ist, lassen sich die Transferraten für kleine $R(\text{DA})$ und E_{12} mit Hilfe der zeitabhängigen Schrödingergleichung bestimmen. In unserer Methode verknüpften wir beide Bereiche durch lineare Interpolation auf einer logarithmischen Skala miteinander.

Diese Trennung von eigentlicher Simulation der Protein- und Protonendynamik und der zeitaufwendigen quantenmechanischen Berechnung des Transferschrittes bestimmt die Effizienz der hier vorgestellten Methode.

Die Methode wurde zuerst auf ein typisches Modellproblem angewendet, in dem ein Wassermolekül mit zusätzlichem Proton (H_3O^+) in einer Wasserbox gelöst ist. Hier lassen sich die Resultate der hier vorgestellten Methode mit denen aufwendigerer quantenmechanischer Simulationstechniken sowie mit experimentellen Daten vergleichen. Die Übereinstimmung bezüglich Transferrate sowie Diffusionskoeffizient des Protons sind bemerkenswert. Die meisten Transferprozesse wurden bei Abständen $R(\text{DA}) \leq 2.5$ Å beobachtet. Allerdings treten viele Konfigurationen auf, bei denen sich Donator und Akzeptor näher als 2.5 Å annähern und bei denen kein Transfer stattfindet. Ob ein Transfer zwischen Donator und Akzeptor stattfindet, wird durch die umgebenden Wassermoleküle bestimmt. Das donierende Hydronium-Ion ist normalerweise von drei Wassermolekülen, mit einem Abstand der Sauerstoffatome zwischen 2.4 Å und 2.6 Å, umgeben. Es zeigt sich, dass ein Transfer nur dann stattfindet, wenn das akzeptierende Wasser näher als 2.5 Å entfernt ist, und gleichzeitig die beiden anderen wasserstoffgebundenen Wassermoleküle weiter als 2.6 Å entfernt sind. Dies führt zu einer ausreichenden Destabilisierung des Protons am Donator und somit zum Protonentransfer.

Anschließend wurde zum ersten Mal dynamisch das Protonierungsgleichgewicht zwischen einem Asparaginsäure-Molekül und den umgebenden Wassermolekülen simuliert.

Protonentransport ist durch die Rate des Auftretens reaktiver Konfigurationen bestimmt. Um längere Zeitintervalle simulieren zu können, wurde die Methode “Locally Enhanced Sampling” implementiert. Mögliche Donatoren und Akzeptoren liegen dabei mehrfach kopiert vor, wobei keine Wechselwirkungen zwischen den einzelnen Kopien wirken. Den Kopien werden dabei zu Beginn leicht unterschiedliche Geschwindigkeiten zugewiesen. Im Verlauf der Simulation können die Kopien daher unterschiedliche Konfigurationen annehmen. Dies führt zu einem besseren Sampling von Donator-Akzeptor-Konfigurationen und damit zu einem schnelleren Auffinden von Protonentransportpfaden. “Locally Enhanced Sampling” kann als eine Art “Mittleres-Feld-Näherung” betrachtet werden, bei der die Energiebarrieren zwischen verschiedenen stabilen Konfigurationen im Vergleich zur Barriere für eine vollständige Aminosäure reduziert sind. Somit werden lokale Umorientierungen von Donator und Akzeptor energetisch einfacher, die für den Protonentransport durch ein Protein der zeitlich limitierende Faktor

sein können. “Locally Enhanced Sampling” führt somit zu einer zusätzlichen Beschleunigung der Transferhäufigkeit. Die Methode wurde für ein Imidazol-Molekül, welches eine hohe Protonenaffinität besitzt, in Mitte von umgebenden Wassermolekülen getestet, und führte zu einer signifikanten Beschleunigung des ersten Transferschrittes von Imidazol zu einem benachbarten Wassermolekül.

Schließlich wurde die Methode auf das Grün fluoreszierende Protein (GFP) angewandt. Hierbei handelt es sich um die erste atomare Simulation eines mehrstufigen Protonentransferprozesses, bei dem Aminosäuren direkt beteiligt sind. Dabei konnte ein dynamisches, atomares Bild sowohl des Protonentransfers in GFP nach Lichtanregung des proteininternen Chromophors sowie des Rücktransport zum Ausgangszustand nach Photonenemission erstellt werden. Neben dem lichtabsorbierenden Chromophor-Molekül sind ein proteininternes Wassermolekül, sowie zwei unterschiedliche Aminosäuren, Serin und Glutaminsäure, am Protonentransport beteiligt. Diese vier Gruppen sind über ein Wasserstoffbrückennetzwerk miteinander verbunden. Nachdem Licht vom Chromophor absorbiert wurde, wird ein Proton des angeregten Chromophors an das benachbarte Wassermolekül abgegeben. In einer konzertierten Reaktion gibt Serin sein Proton an die Glutaminsäure ab und nimmt gleichzeitig das Proton vom Hydronium-Ion auf. Nachdem das Chromophor wieder in den Grundzustand zurückgefallen ist (dieser Prozess wurde durch eine Änderung der Partiaalladungen des Chromophors modelliert), wird die Reprotonierung des Chromophors initiiert. Dabei wird zuerst ein Proton des Wassermoleküls vom Chromophor akzeptiert. In einer konzertierten Reaktion wird dann das Hydroxyl-Proton des Serins an das Hydroxid-Ion doniert und gleichzeitig das Proton von der Glutaminsäure an das Serin abgegeben. Es wurden die Raten für die einzelnen Transferschritte bestimmt. Dabei zeigt sich, dass die Hinreaktion nach Absorption eines Lichtquants sehr schnell (30 - 100 fs) abläuft, während die Rückreaktion etliche Pikosekunden benötigt.

Durch ihre Effizienz erlaubt die vorgestellte Methode erstmalig die dynamische Simulation komplexer Protonentransportreaktionen in Flüssigkeiten und in Biomolekülen. Ein bedeutsamer Vorteil ist weiterhin das einfache Parametrisierungsschema, so dass Transferraten zwischen beliebigen organischen Molekülen leicht ergänzt werden können.

Contents

1	Introduction	1
1.1	Structure and Function of Proteins	1
1.2	Molecular Dynamics Method (MD)	5
1.2.1	Physical Description of Protein Dynamics	5
1.2.2	Molecular Dynamics (MD) Simulations of Proteins	6
1.2.3	Numerical Integration of the Equations of Motion	9
1.2.4	Boundary Conditions	9
1.2.5	NpT-Ensemble	10
1.3	Quantum Mechanical Methods	11
1.3.1	Hartree-Fock Equations	11
1.3.2	Basis Sets	13
1.3.3	Many Body Perturbation Theory	16
1.4	Proton Transfer	18
1.4.1	Biological Significance	18
1.4.2	Theoretical Approaches	19
1.4.3	New Approach: Q-HOP MD - Principal Ideas	20
2	Energy Barriers for Proton Transfer	25
2.1	General Shape of Energy Curves of Proton Transfer	25
2.2	Environmental Influence	27
2.2.1	Method of Calculation	28
2.2.2	Proton Transfer Barriers without Secondary Water Molecules	29
2.2.3	Influence of Secondary Water Molecules on the Proton Transfer Barrier	30

2.2.4	Modeling of the Secondary Waters by Point Charges	32
2.2.5	Analytical Description of the Environmental Effects on the Proton Transfer Barriers	35
2.2.6	Conclusions	40
2.2.7	Towards a Compact Parameter Set	41
2.3	Large Energy Barriers	41
2.3.1	Method of Calculation	43
2.3.2	Motivation	44
2.3.3	Relationship between E_b^{\rightarrow} and E_{12}	47
2.3.4	Relationship between E_{12}^0 and $R(\text{DA})$	52
2.3.5	Test Case	54
2.3.6	Accounting for Spatial Uncertainty of the Proton along the Transfer Reaction Coordinate	56
2.3.7	Accounting for Tunneling in Transition State Theory	58
2.3.8	Temperature Dependence of the Tunneling Contribution to the Transfer Rate	62
2.3.9	Conclusions	63
2.4	Small or no Energy Barriers	64
2.4.1	Method of Calculation	65
2.4.2	Time-independent Schrödinger Equation	67
2.4.3	Time-dependent Schrödinger Equation	68
2.5	Merging Time-dependent Schrödinger Equation with TST	73
2.5.1	Discussion	76
3	Q-HOP MD	79
3.1	Locally Enhanced Sampling	79
3.1.1	Theoretical Foundation	79
3.1.2	Implementation	83
3.2	Proton Hopping	84
3.2.1	Principal Procedure	84
3.2.2	Implementation	90
3.3	Polarizable Environment	93
3.4	How to Conserve Energy and Momentum	97

4	Test Systems	99
4.1	Excess Proton in a Water Box	99
4.1.1	Computational Details	99
4.1.2	Properties	100
4.2	Aspartic Acid in a Water Box	106
4.2.1	Computational Details	106
4.2.2	Properties	106
4.3	Imidazole Ring in a Water Box with Multi-Copy Simulation	108
4.3.1	Computational Details	108
4.3.2	Properties	108
5	Green Fluorescent Protein (GFP)	109
5.1	Introduction of the GFP Protein	109
5.2	Proton Shuttle in GFP: What is known ?	110
5.3	Proton Shuttle in GFP: Simulation with Q-HOP MD	114
5.3.1	Simulation Parameters and Starting Conditions	115
5.3.2	Proton Shuttle after Excitation	117
5.3.3	Back Shuttle to Chromophore in Ground State	118
6	Outlook	123
6.1	Summary	123
6.2	Possible Developments of Q-HOP MD	126
A	Møller-Plesset Perturbation theory	129
A.1	Single Excitations	129
A.2	Double Excitations	130
B	Danksagung	131

Chapter 1

Introduction

1.1 Structure and Function of Proteins

Proteins play a crucial role in almost all biological processes. For instance, they are essential as enzymes in catalytic processes, as molecules for transport and storage, as messengers in signal transduction pathways, as antibodies in the immune system, or in regulation of cell activities.

Proteins are linear hetero-polymers with a length ranging from some tens up to about thousand homomers. The building blocks are amino acids that consist of an amino group (NH_3^+), a carboxylic-acid group (COO^-), and a side chain 'R' that is characteristic for each type of amino acid (Figure 1.1). 20 different amino acids are used in nature (Figure 1.2).

There are nine amino acids which can function as a proton donor or acceptor. These amino acids with variable protonation states are arginine (Arg), aspartic acid (Asp), cysteine (Cys), glutamic acid (Glu), histidine (His), lysine (Lys), serine (Ser), threonine

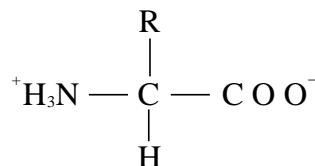


Figure 1.1: The general structural formula for amino acids. There are 20 different side groups R in the naturally occurring amino acids (Figure 1.2)

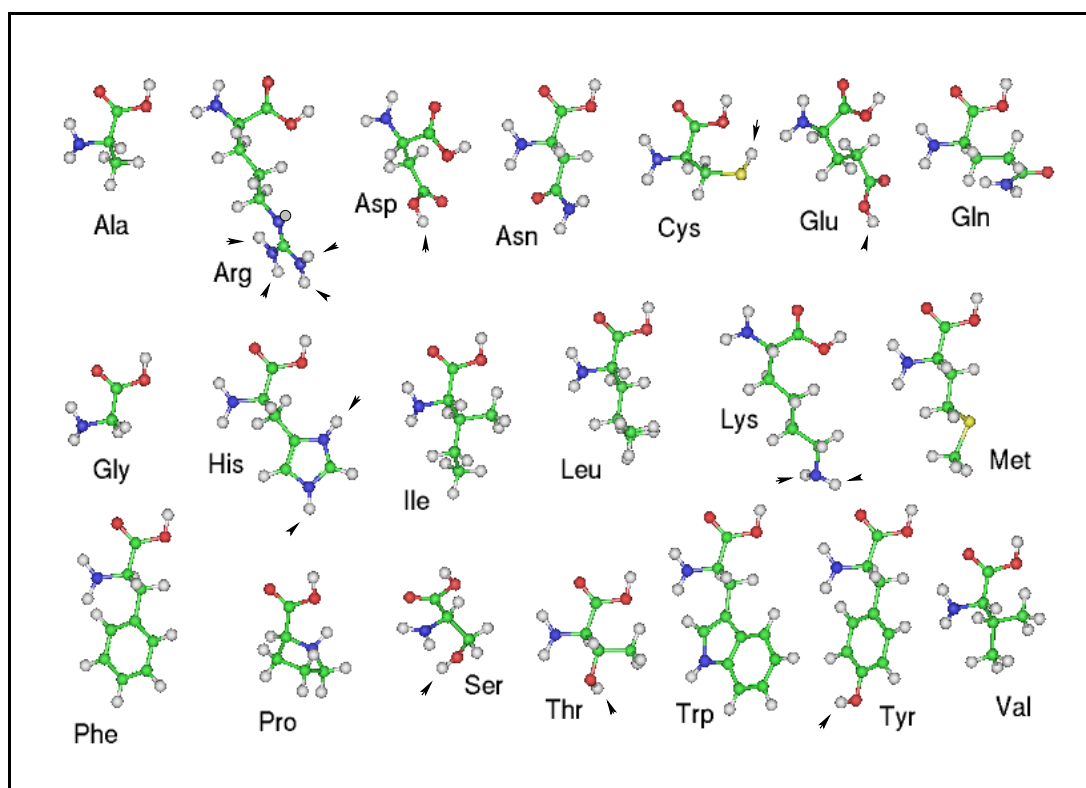


Figure 1.2: Naturally occurring amino acids. Carbon atoms are drawn green, nitrogens blue, oxygens red, sulphur yellow and hydrogens white. The possible donating protons are marked by the black sticks.

(Thr) and tyrosine (Tyr). The protonation equilibria are reflected by their pK_a -values.

The amino acids are bonded to each other via peptide bonds (Figure 1.3). The order of amino acids, which is unique for every type of protein, is called the protein sequence or primary structure of the protein. Figure 1.4 (C) and (D) show small parts of such a polypeptide. The atoms (...-NH-CH-CO-...), colored and drawn as ball-and-sticks, are found in all amino acids (except proline) and construct the backbone of the polypeptide. Atoms of the so called side-chains of each specific amino acid are represented as black lines. They are different for every type of amino acid. Certain parts of the backbone of polypeptides are locally ordered, forming the secondary structure. Two types of such ordered structure elements are found in proteins, α -helices (C) and β -sheets (D) (Figure 1.4), which are stabilized by hydrogen bonds (dashed line). These structurally ordered

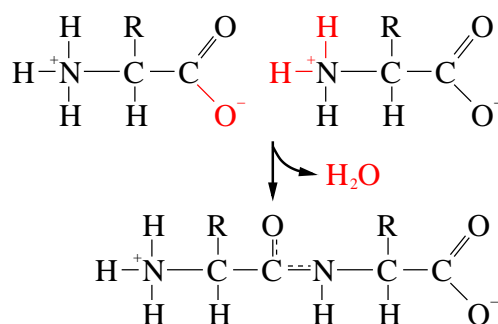


Figure 1.3: Condensation of two amino acids to form a dipeptide under elimination of a water molecule. The extra electron pair is delocalized over O-C-N-H forming a stable, planar peptide bond.

elements are often schematically represented as tubes (α -Helices) and as parallel arrows (β -sheets) (Figure 1.4 (B)). Besides these structures, there are also unordered parts of the backbone in each protein (B: gray). The side-chains are also oriented in an irregular fashion (Figure 1.4 (A), (C) and (D)).

Only a small fraction of the large number of possible protein sequences codes for foldable proteins. These proteins, however, possess a well-defined three dimensional structure (A) called tertiary structure of the protein, which is often dependent on the environmental conditions. The tertiary structure can be interpreted as three dimensional compact arrangement of ordered and unordered secondary structure elements (B).

To understand the function of a protein, one wishes to know its three-dimensional structures in atomic detail. X-ray crystallography, nuclear magnetic resonance (NMR) and electron microscopy (EM) experiments have allowed to determine the structure of about 10000 proteins up to atomic resolution [1]. This knowledge is essential to understand the function at atomic level. However these techniques provide only static protein structures.

Proteins, however, should not be understood as static assemblies, which the three dimensional X-ray crystallographic, NMR- or EM-structures of proteins could suggest. To perform their functions, proteins in general have to undergo several types of dynamic movements. These range from high frequency, thermal motions of single atoms around their average positions found in the time-averaged crystallographic structure, (10^{-14} – 10^{-13} s), over collective motions of groups of atoms (e.g. the movement of amino acid

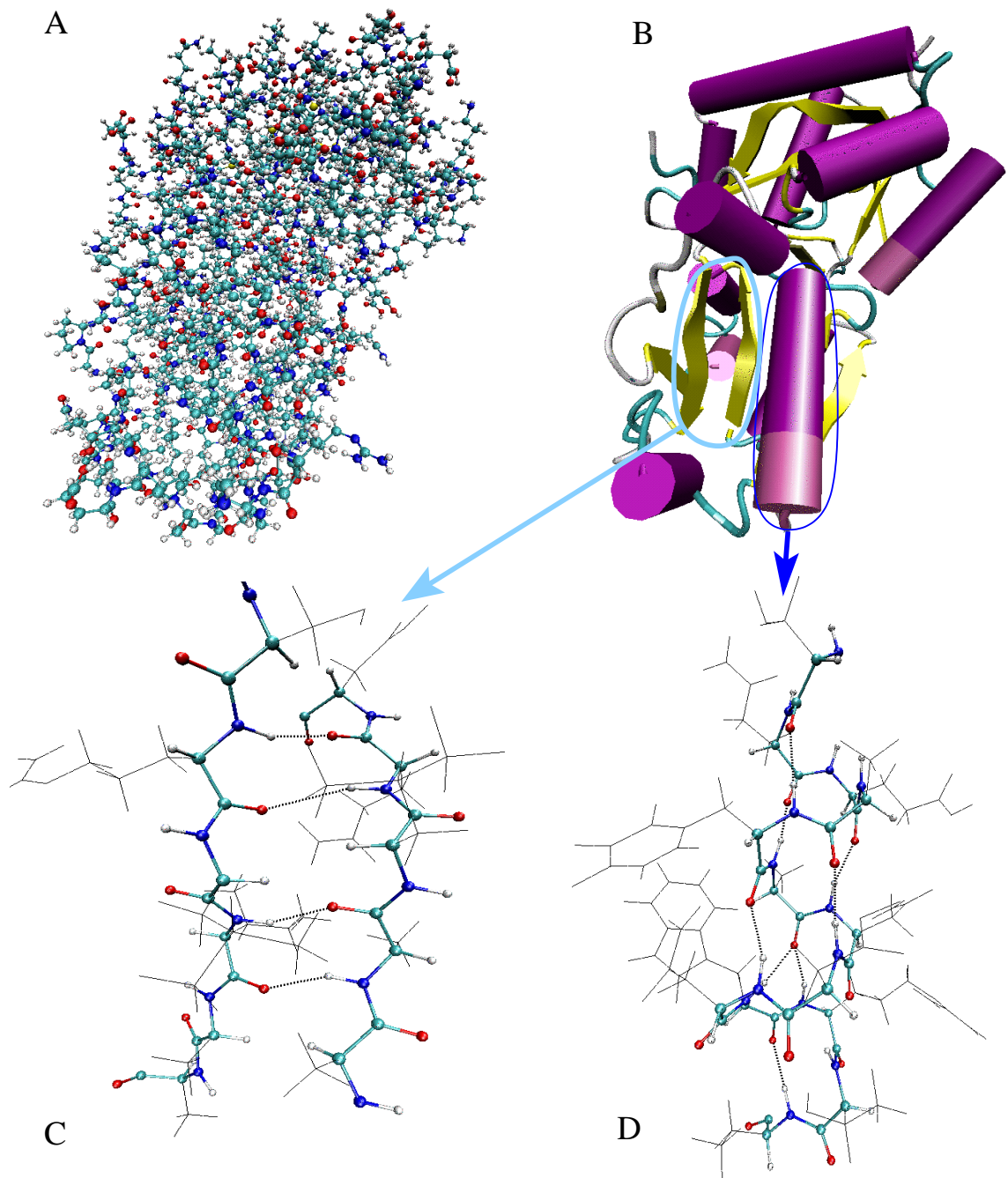


Figure 1.4: Example of the 3-D structure of a particular protein; here shown is the lysine-arginine-ornithine binding protein from *Salmonella typhimurium*. (A) The full atomistic representation, (B) a simplified representation concentrating on the secondary structure elements, (C) β -sheet and (D) α -helix. The atoms (...-NH-CH-CO-...), colored and drawn as ball-and-sticks, construct the backbone of the polypeptide.

side chains) ($10^{-12} - 10^{-10} s$), to larger conformational changes (e.g. the opening of a channel) ($10^{-9} - 10^{-3} s$). The folding of the protein itself can be as fast as microseconds, but usually takes seconds or even longer.

For example in a catalytic process, e.g. the cleavage of a bond of a ligand bound to the enzyme, structural change must take place, so that the ligand can bind to the active center, which is often located in the core of the protein. Additional conformational changes are necessary to expel the ligand after cleavage. During cleavage some amino acids in the active center change their positions and even their chemical nature (e.g. different protonation states [2], ...).

Different experimental techniques, especially spectroscopic techniques [3] have allowed to obtain valuable insights into the molecular mechanisms of protein functions. Usually these techniques can measure relative changes in the protein while it is performing its function. Although the experimental results are important to understand the function of proteins (and theoretical simulations always have to match these experimental results), they usually do not give an atomic picture of the observed relative changes in the protein. Additional analysis by theoretical simulation techniques are able to provide insight into the dynamics of the protein on an atomic scale. The most popular simulation methods that provides such information are introduced in the following section.

1.2 Molecular Dynamics Method (MD)

1.2.1 Physical Description of Protein Dynamics

The chemical heterogeneity of the amino acids results in a great diversity of possible protein sequences, and therefore protein structures. Due to this high complexity and heterogeneity of the protein with almost no symmetry, the common many particle methods known for example in solid state theory cannot be applied.

Explicit simulation techniques which treat each atom individually are the most appropriate methods. To precisely describe the atomic dynamics of a molecular system, the time-dependent Schrödinger equation has to be solved for the many-particle system

consisting of nuclei $\{\mathbf{r}_i\}$ and electrons $\{\rho_I\}$:

$$i\hbar \frac{\partial}{\partial t} \Psi(\{\mathbf{r}_i\}, \{\rho_I\}, t) = \hat{H}(\{\mathbf{r}_i\}, \{\rho_I\}, t) \Psi(\{\mathbf{r}_i\}, \{\rho_I\}, t) \quad . \quad (1.1)$$

Even for small systems of a few atoms, this can be done only numerically and becomes computationally very expensive. Thus, in order to treat larger systems such as proteins, several approximations have to be introduced.

First of all, the electrons move much faster than the heavier nuclei. Therefore, the electronic degrees of freedom can be separated from the nuclear degrees of freedom. Additionally, the nuclei are usually described classically. The Born-Oppenheimer approximation is used to separate nuclear and electronic degrees of freedom. The solutions of the electronic degrees of freedom correspond to the minimal energy states Φ_0 in the adiabatically changing nuclear configuration $\{\mathbf{r}_i(t)\}$

$$\hat{H}_e \Phi_0 = E_0 \Phi_0 \quad \text{where} \quad \hat{H}_e = - \sum_{I=1}^N \frac{\hbar^2}{2m_e} \nabla_I^2 + \hat{V}_{e-e}(\{\rho_I\}) + \hat{V}_{n-e}(\{\rho_I\}, \{\mathbf{r}_i(t)\}) + V_{n-n}(\{\mathbf{r}_i(t)\}) \quad . \quad (1.2)$$

The classical nuclei move in the effective mean-field potential energy surface, produced by the electronic states Φ_0

$$m_i \ddot{\mathbf{r}}_i = -\nabla_i \langle \Phi_0 | \hat{H}_e | \Phi_0 \rangle - \nabla_i V_{n-n}(\{\mathbf{r}_i(t)\}) \quad (1.3)$$

$$= -\nabla_i V_e^{eff}(\{\mathbf{r}_i(t)\}) \quad . \quad (1.4)$$

The time consuming part of this procedure is to determine the quasi-stationary electronic states at given nuclear configuration by solving the time-independent Schrödinger equation with one or another approximation. The degree of accuracy increases from semi-empirical methods to more time consuming 'ab-initio' methods. These methods allow in principle to dynamically simulate chemical reactions, but are currently limited to small molecules (up to about hundred atoms) and short time-intervals (from femtoseconds to some picoseconds).

1.2.2 Molecular Dynamics (MD) Simulations of Proteins

To treat large molecules like proteins consisting of thousands of atoms for time-scales of some nanoseconds (at the current state of hardware and software), the effective potential $V_e^{eff}(\mathbf{r}_1, \dots, \mathbf{r}_N)$, which is a $3N$ -dimensional function, is substituted by an

empirical molecular force-field $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$. The potential $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is usually a series of one-, two-, three- and four-particle interaction terms (Equation 1.5). The potential $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ includes Coulomb- and Van-der-Waals interactions (first two terms) and interactions considering the chemical connectivity of the molecule (3rd to 5th term) (Figure 1.5). The third sum includes a term for every covalent bond in the molecule. Each bond is treated as a simple Hooke's law spring with a characteristic force constant k_{ij}^b and equilibrium bond length l_{ij}^0 . The 4th term accounts for the deformation energy of valence angles between covalent bonds to a given atom, and the 5th term represents the intrinsic deformation energy for twisting about an axis through covalently bonded atoms. Together, these three sums account for variations in the covalent bonding energy of the molecule, and therefore are called 'bonded interactions'. Coulomb- and Van-der-Waals contributions on the other side are called 'non-bonded' interactions. The terms are all given a simple form to be able to calculate the forces in the molecule most efficiently.

$$\begin{aligned}
 V(\mathbf{r}_1, \dots, \mathbf{r}_N) = & \sum_{i < j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{i < j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \\
 & + \sum_{bonds} \frac{1}{2} k_{ij}^b (r_{ij} - l_{ij}^0)^2 \\
 & + \sum_{angles} \frac{1}{2} k_{ijk}^a (\theta_{ijk} - \theta_{ijk}^0)^2 \\
 & + \sum_{dihedrals} \frac{1}{2} k_{ijkl}^d (1 + \cos(n(\Phi_{ijkl} - \Phi_{ijkl}^0))) \quad .
 \end{aligned} \tag{1.5}$$

By using an empirical force field $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ to describe the effective potential $V_e^{eff}(\mathbf{r}_1, \dots, \mathbf{r}_N)$, the electronic degrees of freedom are no longer treated explicitly. Nevertheless, there are positive and negative charged parts of the molecule due to internal distribution of the electrons. This is modeled by assigning partial charges to each atom. The intramolecular and intermolecular electrostatic interactions are calculated using Coulomb's law between the partial atomic charges, see Equation 1.5. The parameters used for the MD simulations presented in this thesis are taken from the AMBER 95 force field [4]. The most time consuming part of the simulation is to calculate the non-bonded interactions between all atoms, because the number of interacting pairs of atoms increases with the square of the number of atoms $N(N-1)/2$. On the other hand, the magnitude of the interaction term decreases with the distance

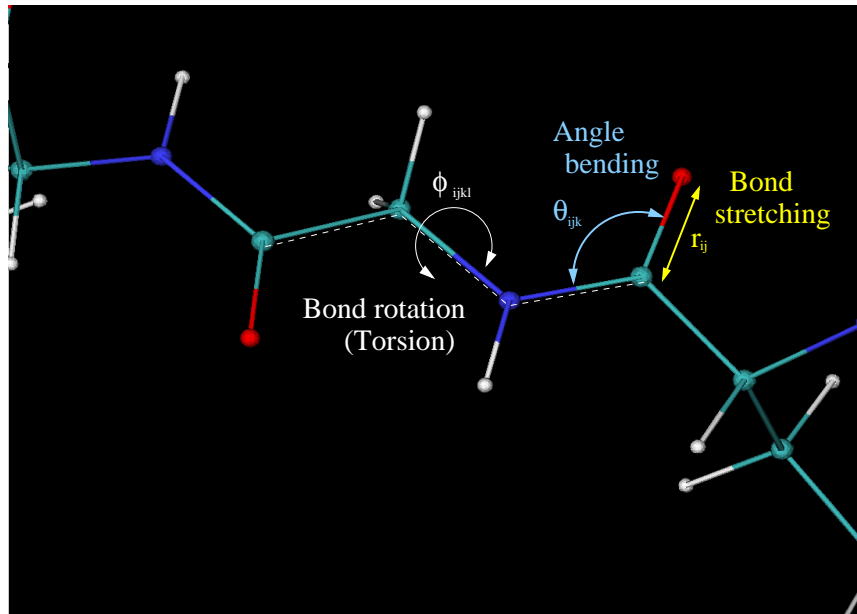


Figure 1.5: Fragment of a protein demonstrating the bonded interactions in classical force fields: bond stretching, angle bending and torsional terms.

between the interacting atoms. Therefore, usually only terms for interacting atoms are calculated where the distance is smaller than a certain value, called cut-off radius. Using this sharp cut-off for Coulomb interactions, leads to errors in size, which are not negligible. To treat the terms of interacting atoms with distances larger than the cut-off radius in an efficient way, different methods were developed [10], like fast multipole [5], particle-particle/particle-mesh [6] reaction-field [7] or Ewald-summation [8] methods. In this work the concept of charge groups was used, where small parts of the proteins are merged into neutral groups. The interaction energy between two neutral groups then decreases with r_{ij}^{-2} , like for dipole interactions.

The classical nuclei are described as classical point particles with coordinates $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ and partial charges (q_1, \dots, q_N) . Their movement is determined by Newton's equation of motion for N classical point particles in the potential $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$

$$m_i \ddot{\mathbf{r}}_i = -\frac{\partial}{\partial \mathbf{r}_i} V(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad i = 1, \dots, N \quad . \quad (1.6)$$

This simulation technique has been developed in fields like hydrodynamics, solid-state theory and astrophysics. It is now widely used in biophysics to simulate DNA-, RNA- and protein systems [9], and it is also a standard tool in refining X-ray or NMR

structures. A general overview of the MD method can be found in [9, 10, 11, 12]. The molecular dynamics package ARGOS [13] was used throughout this thesis.

However, it has also significant drawbacks. By using force fields like in Equation 1.5, covalently bonded atoms cannot change their chemical connectivity. Thus, chemical reactions cannot be studied. Therefore, it also does not allow proton transfer from donor to acceptor, which is the main focus of this thesis.

1.2.3 Numerical Integration of the Equations of Motion

In MD computer simulations, Newton's equations have to be solved numerically. The equations of motion are integrated using a finite difference method. The essential idea is that the integration is broken down into many small stages, each separated in time by a fixed time-interval Δt . The positions and velocities can be approximated as Taylor series expansion. In this work the leap-frog algorithm was used [10]. To calculate the position at time step $t + \Delta t$, in the leap-frog algorithm the velocities at half time step $v_i(t + \Delta t/2)$ are used, leading to

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \Delta t/2) \quad . \quad (1.7)$$

The velocities are updated according to

$$\mathbf{v}_i(t + \Delta t/2) = \mathbf{v}_i(t - \Delta t/2) + \Delta t \frac{\mathbf{F}_i(t)}{m_i} \quad . \quad (1.8)$$

Using this scheme, time steps of 1-2 fs in force field simulations are the upper boundary at which the energy is conserved in microcanonical simulations.

1.2.4 Boundary Conditions

To reach sufficient length of simulation time, the system size has to be limited usually to about 10 000-100 000 atoms for biological systems, and requires proper boundary conditions to minimize artifacts. Here periodic boundary conditions are used, where the cubic simulation box (protein plus solvent) is repeated in all three dimensions (Figure 1.6). A particle leaving the box during the simulation is replaced by an identical image particle that enters from the opposite side. Since the non-bonded interactions are calculated up to a cut-off radius, the minimum image convention was used. Here

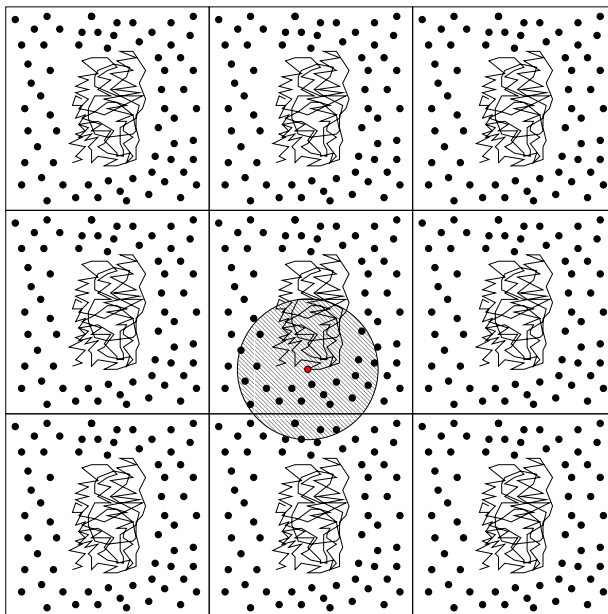


Figure 1.6: Periodic boundary conditions and minimum image convention in two dimensions.

the interactions with the neighboring partners in the periodic images are considered within the cut-off range. To avoid artificial periodicity, the cut-off radius has to be smaller than half the length of the box.

1.2.5 NpT-Ensemble

The simulation scheme presented up to now works in a microcanonical NVE ensemble. Since experiments on proteins in solvent are usually performed at constant pressure p and constant temperature T , the simulation scheme has to be modified to guarantee these conditions. Several techniques have been developed to keep the system at constant temperature and pressure. An overview can be found in [10]. In ARGOS the simulation system is coupled to a heat-bath (Berendsen thermostat [14]) keeping the temperature T at a value close to a given temperature T_0 , which is usually room temperature. After moving the system along the trajectory to the next time-step t , the actual system temperature $T(t)$ is calculated from the particle velocities. Since $T(t)$

usually deviates from T_0 , the velocities are scaled by a factor λ_T ,

$$\lambda_T = \sqrt{1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T(t)} - 1 \right)} \quad , \quad (1.9)$$

where τ_T determines the strength of the coupling. In protein-solvent-systems, different thermostats for each protein ($\tau_T = 0.1 \text{ ps}$) and solvent ($\tau_T = 0.4 \text{ ps}$) are chosen.

The same idea is used to keep the system at a constant pressure p_0 . The pressure at time t is determined using the virial theorem

$$p(t) = \frac{1}{V} \left(\frac{1}{3} \sum_{i=1}^N \frac{1}{2} m_i v_i^2 + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \mathbf{r}_{ij} \cdot \mathbf{F}_{ij} \right) \quad . \quad (1.10)$$

After each time step, the coordinates and the box dimensions are scaled by a factor λ_p ,

$$\lambda_p = \left(1 + \frac{\Delta t}{\tau_p k_B T} (p(t) - p_0) \right)^{1/3} \quad , \quad (1.11)$$

where $\tau_p = 0.5 \text{ ps}$.

1.3 Quantum Mechanical Methods

In the following chapters the energy profiles of proton transfer reactions are computed using quantum mechanical methods. Here the minimum energies are calculated for the electronic degrees of freedom with given nuclear spatial configurations. Again the Born-Oppenheimer approximation is used to separate the electronic degrees of freedom from the nuclear ones.

1.3.1 Hartree-Fock Equations

The time-independent Schrödinger equation for the N electronic degrees of freedom (see Equation (1.2))

$$\hat{H}_e \Phi_0 = E_0 \Phi_0 \quad \text{where} \quad \hat{H}_e = - \sum_{I=1}^N \frac{\hbar^2}{2m_e} \nabla_I^2 + \hat{V}_{e-e}(\{\rho_I\}) + \hat{V}_{n-e}(\{\rho_I\}, \{\mathbf{r}_i(t)\}) + V_{n-n}(\{\mathbf{r}_i(t)\}) \quad (1.12)$$

generally cannot be solved exactly and one has to rely on approximations. To generate these approximate solutions Φ , the energy functional

$$E[\Phi] = \frac{\langle \Phi | \hat{H}_e | \Phi \rangle}{\langle \Phi | \Phi \rangle} \quad (1.13)$$

has to be minimized. The approximation made here is to take a single Slater determinant as trial wavefunction

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(1) & \varphi_2(1) & \dots & \varphi_N(1) \\ \varphi_1(2) & \varphi_2(2) & \dots & \varphi_N(2) \\ \dots & \dots & \dots & \dots \\ \varphi_1(N) & \varphi_2(N) & \dots & \varphi_N(N) \end{vmatrix}, \quad \langle \varphi_I | \varphi_J \rangle = \delta_{IJ} \quad , \quad (1.14)$$

where φ_I are single electron wave functions. This approximation guarantees antisymmetry of the wavefunction by construction, but neglects electron-electron correlations. The single electron wave functions φ_I are molecular orbitals (MO), which are defined as product of a spatial orbital times a spin function, known as spin orbital. The energy can be written as

$$E = \sum_{I=1}^N h_I + \frac{1}{2} \sum_{I=1}^N \sum_{J=1}^N (J_{IJ} - K_{IJ}) + V_{nn} \quad . \quad (1.15)$$

Here h_I is the energy of a single electron I in the field of all nuclei

$$h_I = \left\langle \varphi_I(I) \left| \left(-\frac{\hbar^2}{2m_e} \nabla_I^2 - \sum_i \frac{Z_i e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \boldsymbol{\rho}_I|} \right) \right| \varphi_I(I) \right\rangle \quad , \quad (1.16)$$

J_{IJ} is the Coulomb integral representing the classical electrostatic repulsion between the charge distributions of two electrons I and J

$$J_{IJ} = \left\langle \varphi_I(I) \varphi_J(J) \left| \frac{e^2}{4\pi\epsilon_0 |\boldsymbol{\rho}_I - \boldsymbol{\rho}_J|} \right| \varphi_I(I) \varphi_J(J) \right\rangle \quad , \quad (1.17)$$

and K_{IJ} is the exchange integral

$$K_{IJ} = \left\langle \varphi_I(I) \varphi_J(J) \left| \frac{e^2}{4\pi\epsilon_0 |\boldsymbol{\rho}_I - \boldsymbol{\rho}_J|} \right| \varphi_J(I) \varphi_I(J) \right\rangle \quad . \quad (1.18)$$

Variation of the energy leads to the Hartree-Fock equations

$$(\hat{h}_I + \hat{J}_I - \hat{K}_I) \varphi_I(I) = \epsilon_I \varphi_I(I) \quad \forall I \quad (1.19)$$

with

$$\hat{h}_I = -\frac{\hbar^2}{2m_e}\nabla_I^2 - \sum_i \frac{Z_i e^2}{4\pi\epsilon_0|\mathbf{r}_i - \boldsymbol{\rho}_I|} \quad , \quad (1.20)$$

$$\hat{J}_I = \sum_J \left\langle \varphi_J(J) \left| \frac{e^2}{4\pi\epsilon_0|\boldsymbol{\rho}_I - \boldsymbol{\rho}_J|} \right| \varphi_J(J) \right\rangle \hat{\mathbf{1}} \quad , \quad (1.21)$$

$$\hat{K}_I = \sum_J \left\langle \varphi_J(J) \left| \frac{e^2}{4\pi\epsilon_0|\boldsymbol{\rho}_I - \boldsymbol{\rho}_J|} \right| \varphi_I(J) \right\rangle \hat{P}_{IJ} \quad , \quad (1.22)$$

where \hat{P}_{IJ} is the permutation operator $\hat{P}_{IJ}\varphi_I(I) = \varphi_J(I)$.

1.3.2 Basis Sets

To solve the Hartree-Fock equations, the unknown MOs φ_I are expressed in terms of a set of basis functions, called atomic orbitals (AO). Each MO is expanded by a linear combination of atomic orbitals (LCAO)

$$\varphi_I = \sum_{\alpha} c_{\alpha I} \chi_{\alpha} \quad . \quad (1.23)$$

The optimization task is now to find an optimal set of $c_{\alpha I}$. The basis functions should have a radial and angular behavior that fits to the physical problem, which means that the convergence is reasonably rapid as more basis functions are added. And the chosen basis-functions should make it easy to compute the required integrals. Mainly two types of basis functions are chosen, exponential functions

$$\chi_{\zeta,n,l,m} = N Y_{l,m} \rho^{n-1} \exp(-\zeta\rho) \quad (1.24)$$

($Y_{l,m}$ are spherical harmonics, N is a normalization constant), called Slater type orbitals (STO), that are physically based on the hydrogen atom wavefunctions, and Gaussian type functions (GTO)

$$\chi_{\zeta,l_x,l_y,l_z} = N x^{l_x} y^{l_y} z^{l_z} \exp(-\zeta\rho^2) \quad (1.25)$$

(l_x , l_y and l_z determine the type of orbital, e.g. $l_x + l_y + l_z = 1$ is a p-orbital). While STOs ensure a fairly rapid convergence with increasing number of functions, it is computationally very inefficient to calculate the integrals J_{IJ} and K_{IJ} . They are used for atomic or diatomic systems, where a high accuracy is required. GTOs have the disadvantages that they poorly describe the behavior near the nucleus, and that

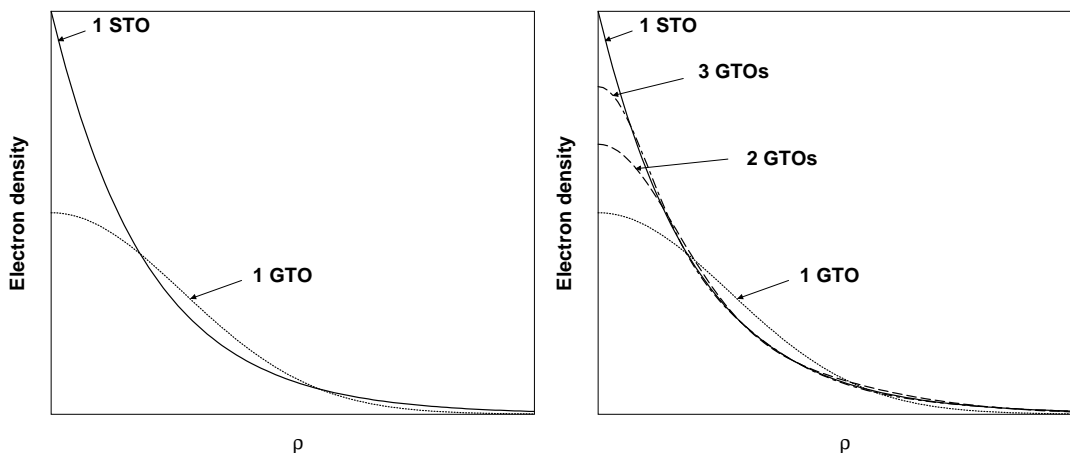


Figure 1.7: *Left:* STO and its best GTO counterpart. A single GTO poorly describes the behavior near the nucleus. *Right:* A linear combination of several GTOs can overcome this problem.

they decay too rapidly to zero with increasing ρ (see Figure 1.7, left). This problem can be overcome, if each atomic orbital is expressed as a linear combination of Gaussians, called primitive Gaussian type orbitals (PGTO) (see Figure 1.7, right)

$$\chi_{\alpha} = \sum_i d_{\alpha i} \chi_i(\text{PGTO}) \quad . \quad (1.26)$$

This increases the number of basis functions significantly and many more integrals have to be calculated. Despite this increase in number of basis functions, GTOs are still computationally more efficient due to its simpler integral calculation. Therefore GTOs are usually used as basis functions for the calculations of the electronic structure of molecules. However, with increasing number of basis functions, the number of optimization parameters increases. The optimization procedure becomes more complex.

This leads to the question how many basis functions should be used. The smallest number of basis functions would be to use as many orbitals necessary to fill in all electrons. This means that a single s-orbital is used for hydrogen (1s), and two s-orbitals and a set of p-orbitals are used for the elements in the second row of the period table (1s, 2s, 2p). A considerable drawback is, that non-spherical aspects of the distribution of the electron density cannot be described. For example, for carbon the only functions that incorporate any anisotropy are the $2p_x$, $2p_y$ and $2p_z$ functions. As the radial components of these functions are required to be the same, not a single component can differ from another. However, the types of atom that are covalently bond to a carbon

atom usually differ from each other. This leads to an anisotropic distribution of the electron density. To treat this anisotropy, usually more basis functions are chosen, e.g. $3 \times 1s$, $3 \times 2s$ and $3 \times 2p$ for elements of the second row in the periodic table. These basis functions differ in their radial distribution of the electron density.

One disadvantage of all energy optimized basis sets is the fact that they primarily depend on the wave function in the region of the inner shell electrons (especially $1s$). These wave functions account for a large part of the energy. Minimizing the energy will tend to make the basis set optimum for the core electrons, and less than optimum for the valence electrons. However, the latter ones are the electrons that take part in chemical reactions. Thus a large number of basis functions would be needed to describe the valence orbitals more accurately. A better approach is to use contracted basis sets. Consider for example the case where 10 s -basis functions on a carbon atom are used for optimization. In this case, it was observed that 6 functions describe the core orbitals, 3 the inner part of the valence orbitals and 1 the outer part of the valence orbitals. For the core orbitals the electron density close to the nuclei has to be described accurately. Since single GTOs poorly describe this part, a linear combination of most of the GTOs available are used to accurately describe the behavior close to the nuclei. Since the core orbitals are independent of different bonding situations, the exponential coefficients ζ and the expansion coefficients of the 6 PGTO are almost constant for each atom type. Therefore, to reduce the computational costs these coefficients are calculated once for each atom type and are held fixed during the optimization procedure.

Throughout this thesis Pople style basis sets are used [15, 16]. Using the 6-31G basis means that for the second period elements, 6 s -basis functions are contracted to one s -basis function ($6s \rightarrow 1s$), describing the core orbital, 3 s/p -functions are contracted to 1 s/p -function ($3s \rightarrow 1s$, $3p \rightarrow 1p$) describing the inner part of the valence orbital and 1 s/p -function remains to describe the outer valence orbitals. In summary 10 s -functions are contracted to 3 s -functions and 4 p -functions to 2 p -functions. For hydrogen, 3 s -functions are contracted to 1 s -function for the inner part of the orbital, 1 s -function remains for the outer part. A short notation is $(10s4p/4s) \rightarrow [3s2p/2s]$. In summary the MO of carbon for example can be written as a linear combination of 5 contracted sets of PGTOs, named CGTOs,

$$\varphi_I = \sum_{\alpha=1}^5 c_{\alpha I} \chi_{\alpha}(\text{CGTO}) \quad , \quad (1.27)$$

where

$$\chi_1(CGTO) = \sum_{i=1}^6 a_i \chi_i^s(PGTO) \quad (1.28)$$

$$\chi_2(CGTO) = \sum_{i=7}^9 a_i \chi_i^s(PGTO) \quad (1.29)$$

$$\chi_3(CGTO) = \chi_{10}^s(PGTO) \quad (1.30)$$

$$\chi_4(CGTO) = \sum_{i=1}^3 b_i \chi_i^p(PGTO) \quad (1.31)$$

$$\chi_5(CGTO) = \chi_4^p(PGTO) \quad (1.32)$$

The coefficients a_i and b_i , as well as all exponential coefficients ζ_i are fixed, and only the expansion coefficients $c_{\alpha I}$ are variable.

So far only s- and p-orbitals were considered for second period elements and only s-orbitals for hydrogen. It is clear that the electron density of a hydrogen atom along a bond is different from the electron density perpendicular to the bond. However, s-orbitals are not sufficient to describe an anisotropic distribution of the electron density. Therefore additional orbitals with higher angular moments, called polarization functions, are added to the basis. In the following, one type of d-orbitals is used as polarization function for second period elements, and one set of p-orbitals for hydrogen. This is denoted 6-31G(d,p) or abbreviated as 6-31G**.

1.3.3 Many Body Perturbation Theory

Up to now, electron-electron correlation was not taken into account explicitly. A widely used method to account for the main part of the correlation is Møller-Plesset perturbation theory (MP). It is based on Rayleigh-Schrödinger many body perturbation theory.

The Schrödinger equation that has to be solved for the unperturbed Hamilton operator \hat{H}_0

$$\hat{H}_0 \Phi_i = E_i \Phi_i \quad (1.33)$$

and the perturbed Hamilton operator is written as

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \quad (1.34)$$

(where λ is a parameter that represents the strength of the perturbation). To solve the

Schrödinger equation for the perturbed Hamilton operator,

$$\hat{H}\Psi_i = W_i\Psi_i \quad (1.35)$$

the energy W_i and the wavefunction Ψ_i are expressed as a Taylor series in powers of the perturbation parameter λ

$$W_i = \lambda^0 W_i^0 + \lambda^1 W_i^1 + \lambda^2 W_i^2 + \dots \quad (1.36)$$

$$\Psi_i = \lambda^0 \Psi_i^0 + \lambda^1 \Psi_i^1 + \lambda^2 \Psi_i^2 + \dots \quad (1.37)$$

where $W_i^0 = E_i$, $\Psi_i^0 = \Phi_i$, and intermediate normalization $\langle \Psi_i | \Phi_i \rangle = 1$ is claimed. Expressing the ground state Ψ_0 in first order perturbation theory Ψ_0^1 as linear combinations of the solutions Φ_i of the unperturbed Schrödinger equation

$$\Psi_0^1 = \sum_i c_i \Phi_i \quad (1.38)$$

and comparing the terms in first power of λ , one finally obtains the first order energy correction

$$W_0^1 = \langle \Phi_0 | \hat{H}' | \Phi_0 \rangle \quad \text{and} \quad c_i = \frac{\langle \Phi_i | \hat{H}' | \Phi_0 \rangle}{E_0 - E_i} \quad (1.39)$$

The second order perturbation theory is derived by expressing

$$\Psi_0^2 = \sum_i d_i \Phi_i \quad (1.40)$$

and comparing the terms in second power of λ . After some technical operations (see e.g. [12, 17, 18, 19]) the second order energy correction is given by

$$W_0^2 = \sum_{i \neq 0} \frac{\langle \Phi_0 | \hat{H}' | \Phi_i \rangle \langle \Phi_i | \hat{H}' | \Phi_0 \rangle}{E_0 - E_i} \quad (1.41)$$

In Møller-Plesset perturbation theory, the unperturbed Hamilton operator is that of the Hartree-Fock equations

$$\hat{H}_0 = \sum_I (\hat{h}_I + \hat{J}_I - \hat{K}_I) \quad (1.42)$$

Then

$$\hat{H}' = \hat{H}_e - \hat{H}_0 = \hat{V}_{ee} - \sum_I (\hat{J}_I - \hat{K}_I) + V_{nn} \quad (1.43)$$

where Equation (1.12) was used.

The sum of W_0^0 and W_0^1

$$W_0^0 + W_0^1 = E_0 + \langle \Phi_0 | \hat{H}' | \Phi_0 \rangle = \langle \Phi_0 | \hat{H}_0 | \Phi_0 \rangle + \langle \Phi_0 | \hat{H}' | \Phi_0 \rangle = \langle \Phi_0 | \hat{H}_e | \Phi_0 \rangle \quad (1.44)$$

gives exactly the Hartree-Fock energy (see Equation (1.15)). Thus correlations beyond HF level start at second order perturbation theory (MP2) with this choice of \hat{H}_0 . The calculation of W_0^2 involves matrix elements of \hat{H}' between the ground state Φ_0 and all excited states Φ_i of the unperturbed wavefunction. \hat{H}' is a two electron operator. Therefore all triple or higher excitations give zero matrix elements. As proven in Appendix A, the single excitation matrix elements are also zero. Thus, only double excitations give non-zero contributions. The result is (see Appendix A)

$$W_0^2 = \frac{\sum_I^{\text{occupied}} \sum_{J>I} \sum_A^{\text{virtual}} \sum_{B>A} \left| \langle \varphi_I(1)\varphi_J(2) | \frac{1}{\rho_{12}} | \varphi_A(1)\varphi_B(2) \rangle - \langle \varphi_I(1)\varphi_J(2) | \frac{1}{\rho_{12}} | \varphi_B(1)\varphi_A(2) \rangle \right|^2}{\varepsilon_A + \varepsilon_B - \varepsilon_I - \varepsilon_J} \quad (1.45)$$

First, the occupied and virtual MOs φ_I are determined by Hartree-Fock calculations. Then W_0^2 is evaluated by calculating the matrix elements in Equation (1.45) using these MOs. This method is used throughout the thesis to calculate the proton transfer barriers (Section 2).

An overview of the methods described in this section can be found in [12, 17, 18, 19].

1.4 Proton Transfer

1.4.1 Biological Significance

The proton, the bare hydrogen nucleus, is a nearly ideal point charge in vacuum. When solvated it forms covalent bonds with electron-rich sites of solvent molecules (one at a time). It has a high mobility in hydrophilic solvent, e.g. water, and moves among water molecules much faster than other ions.

This lead Grotthus to postulate the well known mechanism, where a proton is not transferred as a simple entity from water to water molecule. Instead, the positive charge is transferred by alternating the exchange of hydrogen and covalent bonds between hydrogen bonded water molecules. This proton jumps take place rapidly. Thus, acid-base reactions are among the fastest reactions that take place in aqueous solutions, and

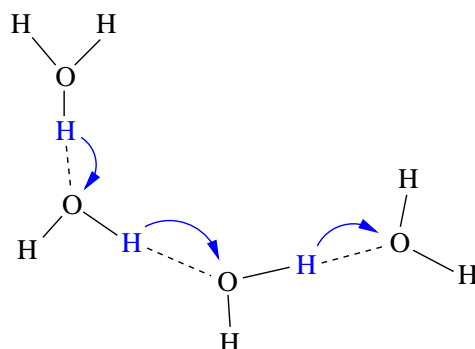


Figure 1.8: Grotthus mechanism of proton migration in aqueous solution via proton jumps. Proton jumps take place rapidly compared with direct molecular migration.

are of special importance in nature. Protons participate in the majority of enzymatic reactions as substrate, product or as an intermediate form, e.g. in serine proteases [20], in carbonic anhydrase [21], in alcohol dehydrogenase [22] or in kinases [23, 24]. Whereas in these reactions the proton is transferred only once from donor to acceptor, in some proteins the proton is transferred along a chain including several water molecules and amino acid residues, e.g. in carbonic anhydrase [21] or in green fluorescent protein [25].

In the field of bioenergetics, which deals with the mechanisms of conversion of energy provided by light or redox reactions into a form utilized by living organisms, membrane proteins are described to transport protons through the entire protein. This leads to a built up of a proton motive force (electrochemical potential difference) as an intermediate form of energy storage. For example bacteriorhodopsin [26] absorbs light to initiate proton transport. Cytochrome c oxidase [27, 28] uses the reductive power of an other molecule (cytochrome c) to initiate the proton transport against its concentration gradient. The resulting gradient is then used as driving force for another proton transporting membrane protein, the ATP synthase [29], where the proton transport initiates the torsion of a small motor which produces ATP.

1.4.2 Theoretical Approaches

A number of different theoretical approaches were previously used to dynamically simulate proton transfer in systems ranging from the simplest systems such as an excess

proton in bulk water to enzymatic reactions. The methods vary largely in complexity from classical models [30] over empirical valence bond (EVB) models [31, 32, 33] to Car-Parrinello simulations [34, 35]. One of the simplest explicit schemes is due to Schmidt & Brickmann [30], who combined classical molecular dynamics (MD) with instantaneous hopping to simulate an excess proton in bulk water. Billeter & van Gunsteren developed a specific parameterization of a protonizable water model for mixed quantum dynamics/molecular dynamics (QC/MD) simulations of proton transfer in small water clusters and liquid water [36, 37, 38]. Enzymatic proton transfer was so far mostly studied by mixed quantum mechanical/molecular mechanical methods. For example, Bala *et al.* used an approximate valence bond method and Hellmann-Feynman forces between a proton and the classically described environment to simulate proton transfer in phospholipase A2 [39]. Billeter *et al.* simulated proton and hydride transfer in liver alcohol dehydrogenase by their “molecular dynamics with quantum transition” (MDQT) surface hopping method [40]. Rothlisberger *et al.* studied enzymatic hydrogen abstraction in galactose oxidase by a mixed Car-Parrinello/classical scheme [41].

Car-Parrinello MD [34, 35] and Multi-State EVB (MS-EVB) [32, 33] clarified the coexistence of the so-called Zundel- and Eigen-cations, H_5O_2^+ and H_9O_4^+ , in bulk water. This work has recently been extended to proton wires in a model water pore [42] that were previously studied by Pomes & Roux using Feynman path integral MD simulations [43].

1.4.3 New Approach: Q-HOP MD - Principal Ideas

The theoretical methods are becoming increasingly sophisticated and are able to describe kinetic and thermodynamic aspects of individual proton transfer events in bulk solution and in macromolecules.

The question arises, why it is necessary to develop yet another technique. First of all, some of these methods such as plain Car-Parrinello MD are computationally cumbersome, and they are restricted to systems of small size covering intervals of 10-20 ps simulation time. However, transport of a proton through a biological membrane protein may take as long as some milliseconds. Therefore the new technique should be very quick, if one wants to explore proton transport paths in proteins that consist of many subsequent transfer events. Once these transport paths are known, more accurate

methods can be used to critically investigate the results. Secondly, protein residues may be essential parts of hydrogen bonded networks that transport protons. One therefore needs methods that not only work for bulk water, but that can be extended to other groups with reasonable effort. Finally, it is attractive to compute energy barriers and kinetic rates at every simulation time step, because many techniques have been developed to reduce the sampling problem in molecular simulations. Once the barriers for proton transfer are known for various possible transfer paths, these barriers could in principle be manipulated to speed up sampling.

To attain long simulation times for large systems, simulation techniques based on empirical force field methods are desirable. Commonly, molecular dynamics simulations (Section 1.2.2) of biomolecular systems are performed for a fixed protonation state. A simple way to include flexible protonation is to allow proton hopping between titratable sites. As mentioned in Section 1.4.2, such an approach was taken by Brickmann and co-workers [30] to simulate an excess proton in liquid water. In this picture, the dynamics of the system in a given protonation state is treated by classical MD simulation. Proton transfer from a proton donating amino acid or water molecule to an accepting one is treated as instantaneous proton hopping by applying a simple distance criterion [30]. If a water molecule comes closer than R to the hydronium ion (H_3O^+), the proton can hop with a probability of 50%. Choosing $R = 2.5 \text{ \AA}$ gave a diffusion coefficient of the excess proton very close to the experimental value. However, small variations in the distance criterion led to changes in the diffusion coefficient by one order of magnitude. Another major shortcoming is that this approach is applicable only in homogeneous systems. It becomes inappropriate for heterogenous transfer events, where donor and acceptor are different chemical groups, and especially for transfers in heterogenous systems, e.g. in proteins.

In this thesis it will be demonstrated that it is possible to derive simple mathematical expressions (Chapter 2) to accurately compute the proton transfer probability between arbitrary donor-acceptor pairs, depending on their distance and on environmental effects of surrounding groups. These expressions are used to compute the hopping probabilities during the MD simulations. The hopping probability is compared with a random number computed numerically, and after a hopping takes place, the simulation continues with the system in the changed protonation state.

Figure 1.9 shows an example to clarify the ideas. Starting point is a system with a

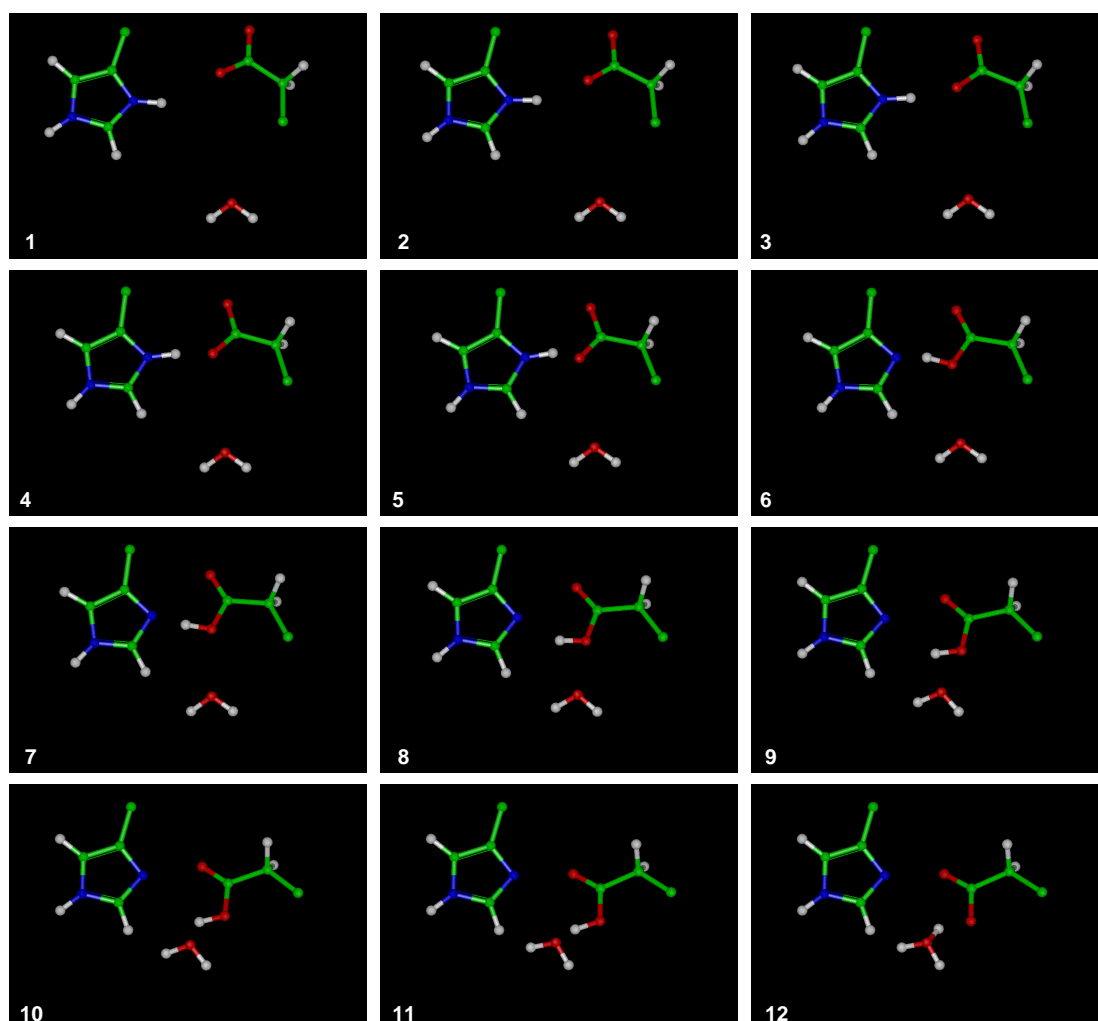


Figure 1.9: Example of a trajectory in Q-HOP MD. In the first three steps (Pictures 2-4), the probability was too small and no proton transfer has happened. After step four (Picture 5), imidazole and acetic acid are so close to each other (distance corresponds to probability, as shown in the next chapter), that the proton hops from donor imidazole to acceptor acetic acid. This is realized in the simulation by instantaneously changing the protein topology (Picture 6): the doubly protonated imidazole becomes singly protonated and the unprotonated acetic acid becomes protonated. Then, the subsequent MD steps are performed (Pictures 7-11), always comparing random number with probability. The next hopping event from acetic acid towards the water molecule is observed (Picture 12).

doubly protonated imidazole ring, an unprotonated acetic acid and a water molecule (Picture 1).

- An MD step is performed.
- At given configuration, the transfer probabilities are calculated between imidazole and the possible acceptors nearby (here acetic acid).
- A random number is computed and compared with the transfer probability. If the random number is smaller than the probability, the proton is transferred.
- The protein topology is altered in case hopping occurs. Otherwise the next MD step is performed with the same protein topology.

Dynamic conformational changes of groups involved in the proton transport might be necessary to guide the proton to the next acceptor. These movements might be the rate limiting steps. To overcome these bottlenecks for proton transport, the method of Locally Enhanced Sampling (LES) was implemented that increases conformational sampling [44](Section 3.1). Here, the groups involved in the actual transport step are represented as multiple copies of the same group with slightly different starting conditions. LES reduces the barrier heights for conformational transitions by approximately a factor $1/N$, where N is the number of copies [45, 46]. Consequently, the time to sample such transitions by MD simulations is significantly reduced, which is shown in this thesis in simulating the proton release from a protonated imidazole ring to water (Section 4.3).

In Chapter 2, the energetic properties of proton transfer will be studied, and a procedure will be presented to calculate the transfer probabilities of a proton step-by-step. This is then implemented into the classical MD program ARGOS (Chapter 3). The new method, termed Q-HOP MD, is then tested for some simple systems (Chapter 4). At the end the results from the first application on a protein, the green fluorescent protein (GFP) (Chapter 5), are presented.

Except for Sections 3.3 and 3.4, the results of Chapters 2-5 have been published [47, 48], are accepted for publication [49, 50] or are submitted for publication [51].

Chapter 2

Energy Barriers for Proton Transfer

2.1 General Shape of Energy Curves of Proton Transfer

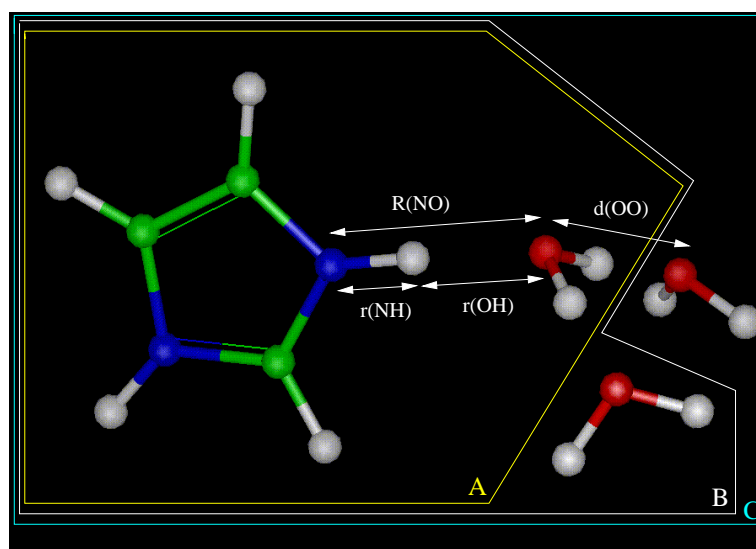


Figure 2.1: Proton transfer system consisting of imidazole, proton accepting, primary water (system A) and one (system B) or two (system C) secondary ligand waters. Four interatomic distances $r(\text{NH})$, $r(\text{OH})$, $R(\text{NO})$ and $d(\text{OO})$ are labeled, which will be used as reaction criteria in this study.

In the previous chapter, the basic ideas of Q-HOP MD were presented to simulate

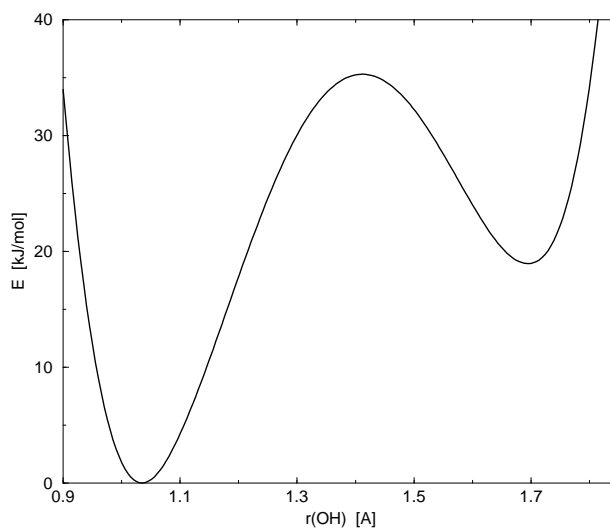


Figure 2.2: General double-well shape of the energy curve of proton transfer.

proton transfer in biological systems. But the question remains open how to calculate the transfer probabilities between all possible donor-acceptor pairs in all possible configurations of protein and solvent. To approach this question, small systems were chosen consisting of a donor and an acceptor group with some surrounding groups to model part of the environment of the donor-acceptor system. Figure 2.1 shows one example of these small systems, that has previously been studied by Lu and Voth [21]. It consists of a doubly protonated imidazole ring (the side group of histidine) as donor and a water molecule as acceptor, which is called 'primary water'. The distance between the donor and the acceptor atom is named $R(\text{NO})$. Two water molecules, called 'secondary water', are placed in a position hydrogen-bonded to the accepting water, $d(\text{OO})$ away from the oxygen of the primary water. They model part of the environment of the proton transfer system.

The advantage of concentrating on these small systems is, that they can be treated with very accurate quantum chemical calculations to get the energy curve for the proton transfer reaction, in this case transfer from imidazole to the primary water molecule. The proton is moved from the donor-bound site (here nitrogen) to the acceptor-bound site (here oxygen) step-by-step, by increasing the distance $r(\text{NH})$. The result is usually a double-well form of the energy curve of the proton transfer reaction (Fig. 2.2).

In the following, the factors are discussed which influence the energy barrier of a proton transfer (Section 2.2), and then a procedure is derived to calculate the transfer

probabilities from the energy curves (Sections 2.2.7 - 2.5).

2.2 Environmental Influence

The thermodynamic and kinetic properties of proton transfer processes in proteins are strongly influenced by the environment of the proton transferring chemical groups [20, 21, 52, 53]. Predominantly through electrostatic interactions [21], the relative energies of reactant and product states are modified, as is the energy barrier height between these two states. Environmental effects therefore have a significant impact on reaction rates and equilibria, and on local pK_a values [54] of titratable residues.

Different theoretical approaches, as described in the previous chapter, have been proposed to describe the environmental influence on proton transfer reactions. The most accurate and most costly way is to treat the complete system of proton transferring groups and their adjoining environment up to a defined distance by electronic structure methods [23]. QM/MM methods are being used to consider a larger part of the environment in an efficient manner and to simultaneously allow for a higher level of quantum mechanical description of the reactive groups. There, the active center of the reaction is described fully quantum mechanically by electronic structure methods, whereas the environmental groups are treated using force field methods [55] or are described by a polarizable continuum in a self-consistent reaction field approach [52].

In the approach presented here, an analytical description of hopping probabilities is needed for every geometry of the proton transferring and environmental groups, which would allow to determine very quickly whether a proton is transferred or not. To approach this aim, the model system, described before, for proton transfer is chosen consisting of a proton donating amino acid side group, imidazole, and an accepting group, one water molecule (Figure 2.1: system A).

The energy hypersurface, from which the transfer probabilities can be derived, is dependent on different parameters. First, the dependence of the proton transfer energy surface on the distance between the donating and the accepting atom, here $R(\text{NO})$, is analyzed. Then the influence of environmental groups, here ligand water molecules, is investigated (Figure 2.1: systems B and C). The effect of the relative positions of these secondary waters to the proton transferring imidazole-water system is analyzed. From these data an analytical approach is derived in a stepwise manner, which accounts for

the environmental effects on the proton transfer energy surface.

2.2.1 Method of Calculation

The calculations were performed with the program package NWChem 3.1 [56] on Compaq alpha workstations. To establish the computational methodology, the proton transfer in the well studied system H_5O_2^+ was chosen as test case. The results for different proton acceptor-donor distances with density functional methods (B3LYP functional) and wavefunction methods (MP2) using basis sets ranging from 3-21G to 6-31G** are shown in table 2.1, together with those from very accurate coupled cluster calculations taken from [57]. RHF results are in reasonable agreement with those found in the

Table 2.1: Energy barriers [kJ/mol] for proton transfer in H_5O_2^+ at different $R(\text{OO})$.

Method	$R(\text{OO})$ [Å]			
	2.5	2.6	2.7	2.8
MP2/6-31G* // HF/3-21G	1.3	8.4	19.7	35.2
MP2/6-31G** // HF/6-31G**	1.7	8.8	23.9	41.9
B3LYP/6-31G* // HF/3-21G	0.4	5.0	14.7	27.6
B3LYP/6-31G** // HF/6-31G**	0.4	4.2	15.9	31.0
CCSD(T)/cc-pVTZ // MP2/cc-pVTZ	1.55	8.71	20.31	35.34
QCISD(T)/cc-pVTZ // MP2/cc-pVTZ	1.59	8.63	20.18	35.17

literature [58]. As discussed in [57, 59], density functional methods underestimate the energy barriers between the two minimum energy states compared to results obtained by CCSD(T) and QCISD(T) calculations. The predictions for the energy barriers derived with MP2 are closer to those of the high level approaches. Therefore only results from MP2 calculations for the imidazole water system are shown.

The geometries of system A (Figure 2.1), consisting of imidazole and water, are optimized using HF/6-31G** and the energies are calculated at MP2/6-31G** // HF/6-31G**. To obtain the energy surface of a proton transfer reaction, the proton is moved stepwise by 0.1 Å along the $\text{N}_{\text{Im}}\text{-O}_{\text{Prim.Wat}}$ interconnecting line. At each $r(\text{NH})$ the geometry is optimized, keeping the distance $R(\text{NO})$ and the angle $\angle(\text{NHO}) = 180^\circ$ fixed. All other degrees of freedom remain free.

To study environmental effects, one (system B) or two secondary water molecules (system C) are added. The geometries are optimized using HF/3-21G to save computing time, and the energies are calculated at MP2/6-31G*//HF/3-21G. In the optimization procedure, the coordinates for the oxygens of the secondary waters are first determined at an intermediate distance $r(\text{NH}) = 1.4 \text{ \AA}$, while keeping the primary water fixed. Then, similar to system A, the energy is calculated at every $r(\text{NH})$, while constraining the coordinates of the donating atom N_{Im} , the transferred hydrogen and the accepting $\text{O}_{\text{Prim.Wat}}$, as well as the position of the oxygen of the secondary waters.

Since systems B and C contain more basis functions than system A, the importance of the basis set superposition error was checked for a correct estimation of the effect of the secondary waters. This was done by Counterpoise Correction, i.e. examining system A and ghost atoms with orbitals on the positions of the secondary waters. The energy surface compared to that of system A shows only negligible differences (results not shown).

2.2.2 Proton Transfer Barriers without Secondary Water Molecules

The energy surface of proton transfer for the imidazole-water system is shown in Figure 2.3 as function of the separation between donating atom and transferred proton, $r(\text{NH})$, at different distances between donor and acceptor, $R(\text{NO})$, ranging from 2.65 \AA to 3.45 \AA . For large separations $R(\text{NO}) \geq 2.95 \text{ \AA}$, a double-well potential is formed with a single deep minimum at about $r(\text{NH}) = 1.03 \text{ \AA}$, where the proton is bound to the donor N_{Im} , and a shallow minimum at about $r(\text{OH}) = 1.09 \text{ \AA}$ for $R(\text{NO}) = 2.95 \text{ \AA}$ to $r(\text{OH}) = 1.03 \text{ \AA}$ for $R(\text{NO}) = 3.45 \text{ \AA}$, where the proton can be assigned to the accepting $\text{O}_{\text{Prim.Wat}}$. The shape of the double-well potential reflects the stronger proton affinity of an imidazole group compared to that of a water molecule [21]. For $R(\text{NO}) \leq 2.95 \text{ \AA}$, the proton affinity of the water molecule is insufficient to form a stable hydronium ion; the second minimum disappears.

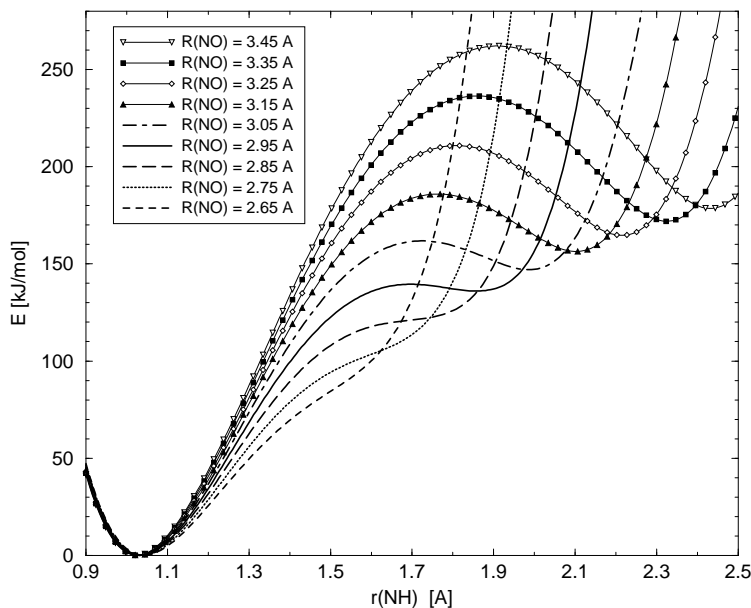


Figure 2.3: Proton transfer energy surface for system A as function of $r(\text{NH})$ at different separations between imidazole's nitrogen and primary water's oxygen, $R(\text{NO}) = 2.65 \text{ \AA} \dots 3.15 \text{ \AA}$.

2.2.3 Influence of Secondary Water Molecules on the Proton Transfer Barrier

As previously shown by Lu and Voth [21], secondary water molecules ligated to the primary, proton accepting water molecule stabilize the formation of H_3O^+ mainly by electrostatic interaction between the ligand waters and the changing charge distribution that accompanies the proton transfer process. Figure 2.4, as one example, shows the effect of one and two secondary waters at a distance of $d(\text{OO}) = 2.6 \text{ \AA}$ between primary and secondary water on the proton transfer energy surface of the imidazole-water system at $R(\text{NO}) = 3.05 \text{ \AA}$. The stabilization of the second minimum results in a decrease of the forward energy barrier by 46 kJ/mol to about 120 kJ/mol for one ligand water, and by 82 kJ/mol to about 84 kJ/mol for two ligand waters. In addition, the proton transfer barrier for the backward transfer from water to imidazole increases from about 8 kJ/mol to 33 kJ/mol for one, and to 63 kJ/mol for two secondary waters.

The effect of varying the distance between primary and secondary waters is demonstrated in Figure 2.5. As the secondary waters are shifted away, the electrostatic

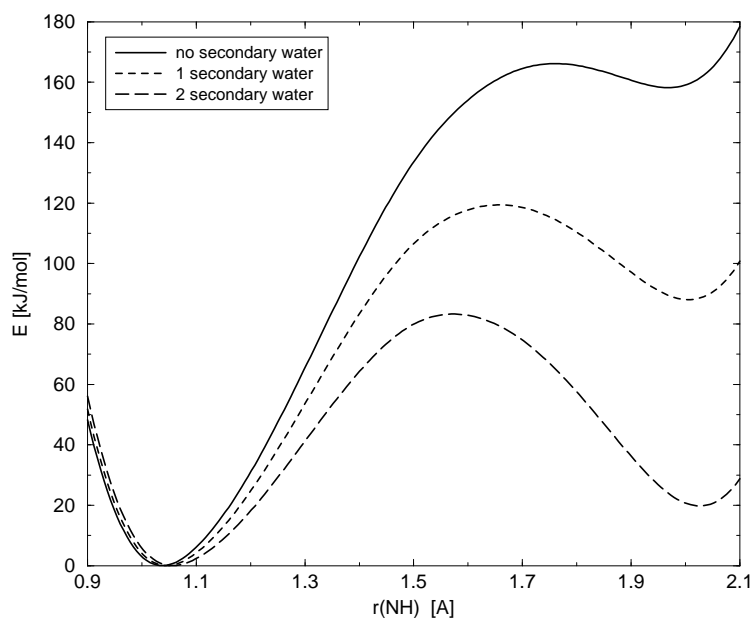


Figure 2.4: Influence of one and two secondary waters on the proton transfer energy surface for imidazole-water at $R(\text{NO}) = 3.05 \text{ \AA}$ and $d(\text{OO}) = 2.6 \text{ \AA}$.

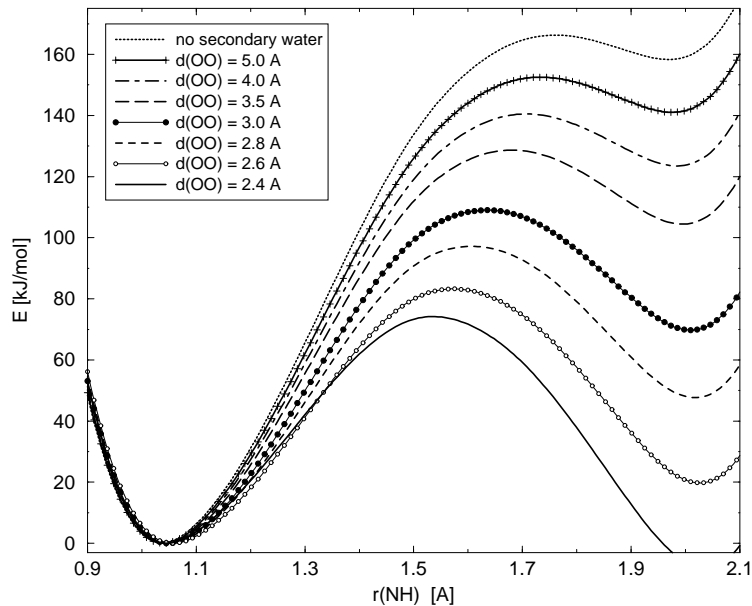


Figure 2.5: Dependence of energy barriers for system C on the position of the secondary waters relative to the primary water. The distance is varied from $d(\text{OO}) = 2.4 \text{ \AA}$ to $d(\text{OO}) = 5.0 \text{ \AA}$ and to infinity, which means no secondary water is present. These calculations were performed at $R(\text{NO}) = 3.05 \text{ \AA}$.

interaction with the imidazole water system and thus the stabilization of the second minimum diminishes. As a consequence, the barrier increases from 83 kJ/mol at $d(\text{OO}) = 2.6 \text{ \AA}$ to 152 kJ/mol at $d(\text{OO}) = 5.0 \text{ \AA}$, while the barrier for the reverse transfer decreases from 64 kJ/mol at $d(\text{OO}) = 2.6 \text{ \AA}$ to 11 kJ/mol at $d(\text{OO}) = 5.0 \text{ \AA}$.

2.2.4 Modeling of the Secondary Waters by Point Charges

The question now arises, whether the proton transfer energy surface of the complete system C can be reproduced by treating proton donor and acceptor (system A) with electronic structure methods (here: MP2) and by representing the secondary waters molecular mechanically, as three point charges at the positions of the nucleic centers. To examine this situation, the explicit ligand waters were replaced by point charges. Single point calculations were performed varying $r(\text{NH})$, where the geometry of the complete system with explicit secondary waters was optimized at each $r(\text{NH})$. The point charges of the hydrogen atoms of the waters are chosen to be identical and the total charge on each water equals zero.

Figure 2.6 shows the energy profiles for $d(\text{OO})$ ranging from 2.4 \AA to 5.0 \AA . For distances $d(\text{OO}) \geq 5.0 \text{ \AA}$, the energy profiles computed with the simple point charge water model (SPC) of Berendsen and co-workers [63] ($q_{\text{O}} = -0.82/q_{\text{H}} = 0.41$, solid line) are in excellent agreement with those for the full quantum system (results for $d(\text{OO}) = 6.0 \text{ \AA}$ are not shown here). For shorter distances, one observes increasing deviations the closer the secondary waters approach the primary water. To reproduce the quantum mechanical energy barrier, the charges on the secondary waters had to be increased up to $q_{\text{O}} = -1.20/q_{\text{H}} = 0.60$. Table 2.2 shows that by systematically increasing the point charges, energy surfaces are obtained in excellent agreement with those obtained by calculations with quantum mechanically treated ligand waters.

As the secondary waters approach the primary water, the magnitude of the point charges that give the best fit increases continuously from $q_{\text{O}} = -0.82/q_{\text{H}} = 0.41$ at $d(\text{OO}) = 5.0 \text{ \AA}$ and $d(\text{OO}) = 6.0 \text{ \AA}$ to $q_{\text{O}} = -1.20/q_{\text{H}} = 0.60$ at $d(\text{OO}) = 2.6 \text{ \AA}$. The closer a secondary water approaches the primary water, the interaction between the dipole of the latter with the electron density of the ligand water becomes stronger. Therefore the induced polarization of the secondary water and thus the absolute values of the charges on the nuclei positions increase.

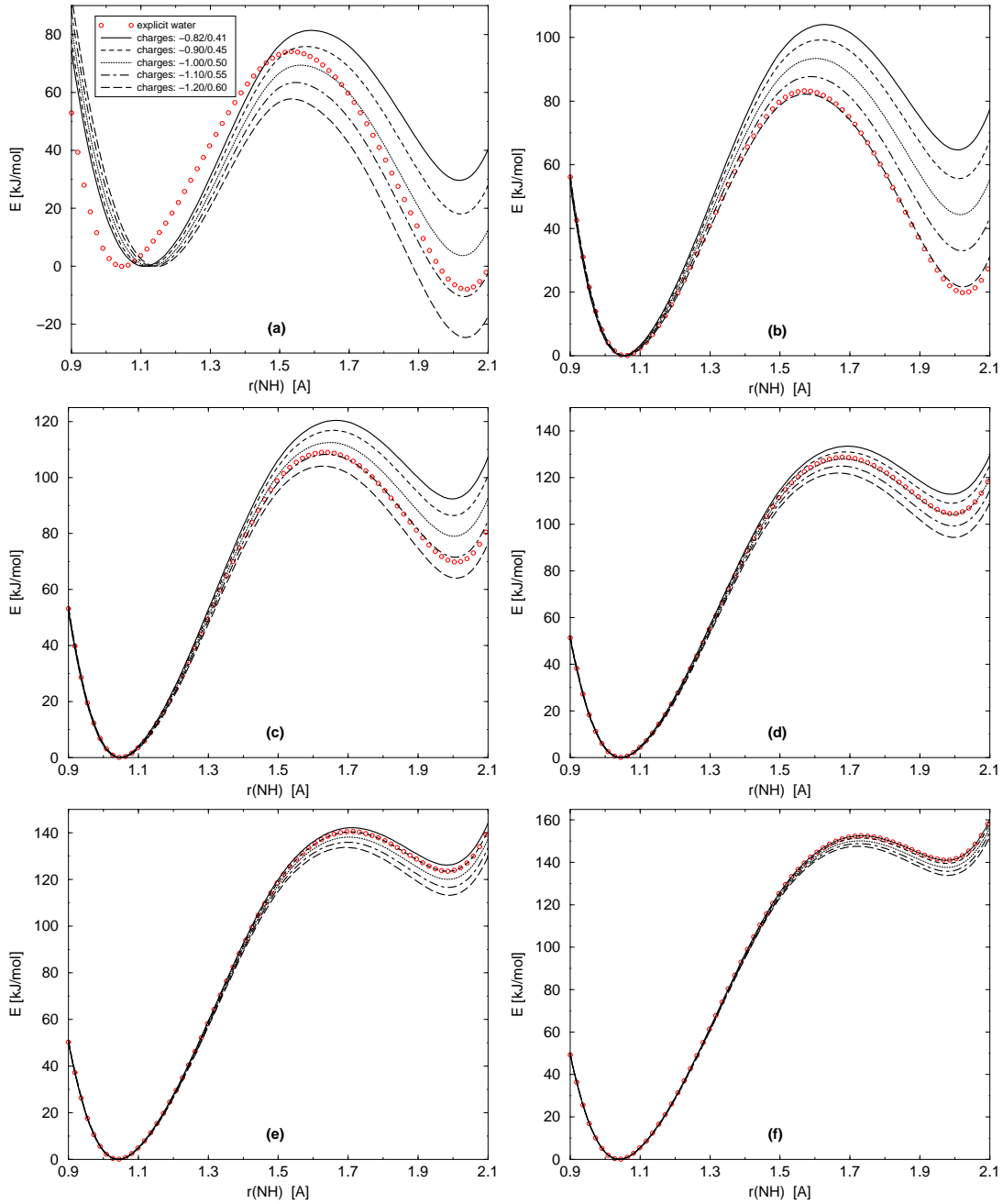


Figure 2.6: Comparison of the energy barriers for system C (where the secondary waters are part of the electronic structure calculation) with those having the ligand waters modeled by point charges. The point charges are varied between $q_{\text{O}} = -0.82/q_{\text{H}} = 0.41$ (SPC) and $q_{\text{O}} = -1.20/q_{\text{H}} = 0.60$. The distance between imidazole and primary water is $R(\text{NO}) = 3.05$ Å. Comparisons are shown for distances between primary and secondary waters ranging from (a) $d(\text{OO}) = 2.4$ Å over (b) $d(\text{OO}) = 2.6$ Å, (c) $d(\text{OO}) = 3.0$ Å, (d) $d(\text{OO}) = 3.5$ Å, (e) $d(\text{OO}) = 4.0$ Å to (f) $d(\text{OO}) = 5.0$ Å.

Table 2.2: Comparison between energy barriers of fully quantum mechanically treated system C and the energy barriers, where the secondary waters are modeled by point charges.

$d(\text{OO})$ [\AA]	point charges of best fit [e]	E_b^{\rightarrow} [kJ/mol]		E_b^{\leftarrow} [kJ/mol]	
		QM waters	point charge waters	QM waters	point charge waters
	$q_{\text{O}}/q_{\text{H}}$				
2.6	-1.20/0.60	83.3	82.1	63.6	60.3
3.0	-1.10/0.55	108.9	108.0	38.9	36.4
3.5	-1.00/0.50	128.5	127.7	24.3	23.4
4.0	-0.90/0.45	140.3	140.3	16.7	16.7
5.0	-0.82/0.41	152.4	152.4	11.3	11.3
6.0	-0.82/0.41	157.8	157.8	10.0	10.0

For $d(\text{OO}) = 2.4 \text{ \AA}$, the energy surface of the full quantum system cannot be reproduced by point charges on secondary waters anymore. This may be due to a strong polarization of the system which leads to charged secondary water molecules. We found that Mulliken charges on the atoms of a secondary water sum up to a total charge varying from $0.03 e$ at $r(\text{NH}) = 0.9 \text{ \AA}$ to $0.08 e$ at $r(\text{NH}) = 2.1 \text{ \AA}$, whereas the sum for $d(\text{OO}) = 2.6 \text{ \AA}$ varies only from $0.035 e$ at $r(\text{NH}) = 0.9 \text{ \AA}$ to $0.065 e$ at $r(\text{NH}) = 2.1 \text{ \AA}$. This may explain why a neutral secondary water with equal charges for every $r(\text{NH})$ is not able to reproduce the correct energy surface for the proton transfer process.

For distances $d(\text{OO}) \geq 2.6 \text{ \AA}$, it is feasible to calculate the proton transfer surface between an amino acid side chain, here imidazole, and water surrounded by ligand waters, treating these molecular mechanically using point charges. While for $d(\text{OO}) \geq 5.0 \text{ \AA}$ charges of SPC water can be used, the charges of the ligand waters for $2.6 \text{ \AA} \leq d(\text{OO}) \leq 5.0 \text{ \AA}$ should be enhanced by a factor, which is dependent on $d(\text{OO})$ (Figure 2.7). Very similar results for the fitted point charges were obtained at $R(\text{NO}) = 2.65 \text{ \AA}$.

A water molecule represented by point charges with $q_{\text{O}} = -1.20/q_{\text{H}} = 0.60$ has a dipole moment of approximately 3.2 D . This may appear quite large when compared to the experimental value in vacuum (approx. 1.8 D) or to the empirical SPC water (approx. 2.3 D). However, dipole moments as high as 2.9 D were previously reported in

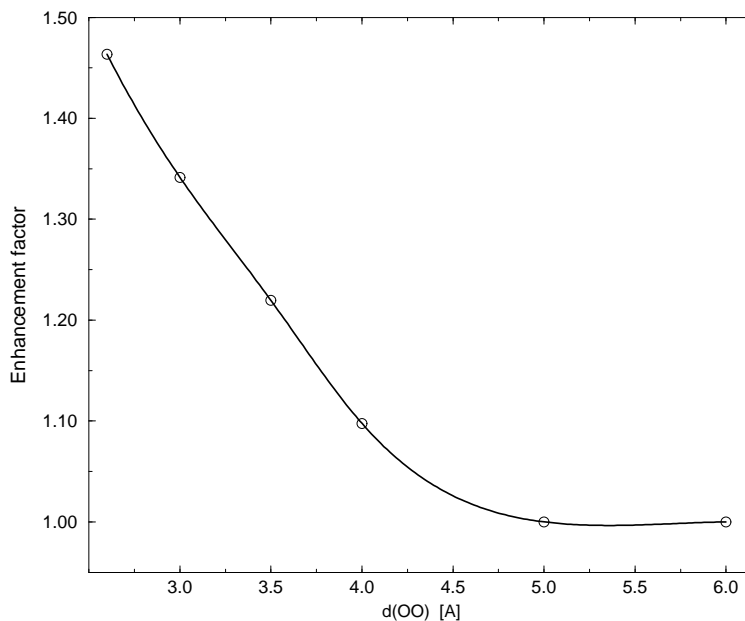


Figure 2.7: Factor by which the SPC charges of the ligand waters have to be enhanced to give the best fit for the proton transfer energy barrier at a specific $d(\text{OO})$.

molecular dynamics simulations of bulk water with explicit polarizabilities [60, 61, 62]. Even higher values should be possible for a well-coordinated water molecule for example in a protein interior.

2.2.5 Analytical Description of the Environmental Effects on the Proton Transfer Barriers

In the previous section it was shown that the environmental effect of secondary water molecules on a proton transfer reaction can be well reproduced by treating the secondary waters as point charges with distance-dependent magnitude, at least for $d(\text{OO}) \geq 2.6 \text{ \AA}$. However, it is still necessary to re-calculate the energy surface of the proton transfer between isolated donor and acceptor for every rearrangement of the surrounding waters. A significant improvement is achieved, if the proton transfer energy surface only had to be calculated once for the isolated system (system A), and if the influence of surrounding ligand waters on the energy landscape could be added as a correction later on.

Table 2.3: Energy barriers for different approaches in increasing order of simplification

			d(OO)=2.6 Å		d(OO)=2.8 Å	
Point charges on 2nd H ₂ O [e]:			$q_O = -1.20/q_H = 0.60$		$q_O = -1.20/q_H = 0.60$	
1st H ₂ O, imidazole ^a	2nd H ₂ O ^b	Energy calculation ^c	E_b^{\rightarrow} [kJ/mol]	E_b^{\leftarrow} [kJ/mol]	E_b^{\rightarrow} [kJ/mol]	E_b^{\leftarrow} [kJ/mol]
QM	QM	MP2	83.3	63.6	97.1	49.4
QM	MM	MP2	82.1	60.3	94.6	48.2
MM _C	MM	All Multipoles	84.2	62.8	97.6	47.7
MM _C	MM	Dipole	97.1	44.0	108.9	38.1
MM _A	MM	All Multipoles	90.9	45.2	102.2	38.9
MM _A	MM	Dipole	105.1	47.3	114.3	39.8

			d(OO)=3.0 Å		d(OO)=3.5 Å	
Point charges on 2nd H ₂ O [e]:			$q_O = -1.10/q_H = 0.55$		$q_O = -1.00/q_H = 0.50$	
1st H ₂ O, imidazole	2nd H ₂ O	Energy calculation	E_b^{\rightarrow} [kJ/mol]	E_b^{\leftarrow} [kJ/mol]	E_b^{\rightarrow} [kJ/mol]	E_b^{\leftarrow} [kJ/mol]
QM	QM	MP2	108.9	38.9	128.5	24.3
QM	MM	MP2	108.0	36.4	127.7	23.4
MM _C	MM	All Multipoles	113.5	36.0	133.6	23.4
MM _C	MM	Dipole	121.0	30.1	135.7	21.8
MM _A	MM	All Multipoles	114.7	30.6	133.6	21.4
MM _A	MM	Dipole	124.8	32.2	137.3	22.2

^aQM: Imidazole and primary water are treated quantum mechanically; MM_{C/A}: Geometry and Mulliken charges are taken from electronic structure calculation on system C/A.

^bLigand waters are treated quantum mechanically (QM) or are modeled by point charges (MM).

^cThe energy surface is determined by electronic structure calculations (MP2), by equation (2.2), where all multipoles are taken into account, or by equation (2.4), where just the dipole term is considered.

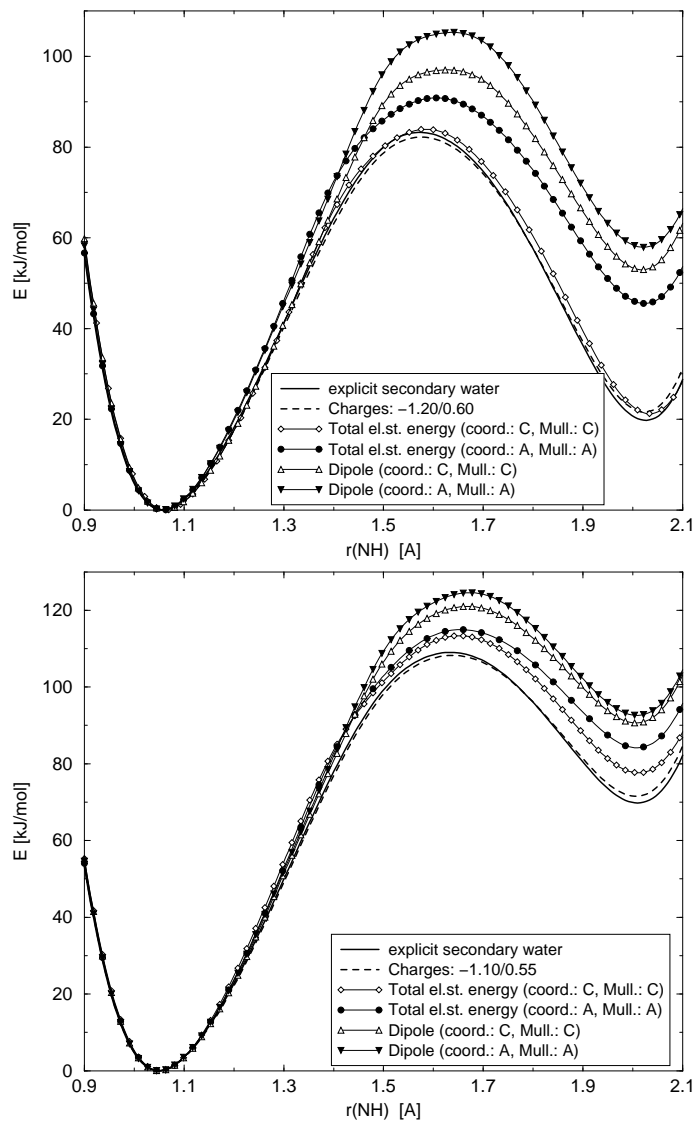


Figure 2.8: Energy surfaces, where the effect of the secondary waters is calculated with formula (2.2). Mulliken charges and geometries of imidazole and primary water are taken from electronic structure calculations on system C (solid line with open diamonds) or on system A, (solid line with filled circle), or with formula (2.4) (solid line with open triangles or filled triangles, respectively, for Mulliken charges and geometries as described before). Again $R(\text{NO}) = 3.05 \text{ \AA}$. Results are shown (a) for $d(\text{OO}) = 2.6 \text{ \AA}$, and (b) for $d(\text{OO}) = 3.0 \text{ \AA}$.

Table 2.2.5 and Figures 2.8 a,b show the results of a stepwise approach to determine the energy surface of the complete system C, by adding the electrostatic interaction energy between ligand waters and imidazole-water system A to the energy landscape of system A for the cases $d(\text{OO}) = 2.6 \text{ \AA}$ and $d(\text{OO}) = 3.0 \text{ \AA}$ at $R(\text{NO}) = 3.05 \text{ \AA}$. This approach was motivated by the work of Borgis [64], who derived a dipole approximation to estimate the effect of an external electric field \mathbf{E} on the proton transfer barrier between two water molecules.

In the first step, using the geometry and the Mulliken charges of the optimized system C, the electrostatic interaction between the Mulliken charges of donor and acceptor Q_i ($i = 1, \dots, 13$) and the electrostatic potential Φ created by the point charges of the secondary waters q_j ($j = 14, \dots, 19$) is calculated according to

$$E_{\text{el}} = \sum_{i=1}^{13} Q_i \Phi(\mathbf{r}_i) \quad (2.1)$$

$$= \sum_{i=1}^{13} Q_i \sum_{j=14}^{19} \frac{q_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \quad (2.2)$$

This electrostatic energy is added to the energy surface of the proton transfer in the isolated system A. Using charges of $q_{\text{O}} = q_{14} = q_{17} = -1.2$ and $q_{\text{H}} = q_{15} = q_{16} = q_{18} = q_{19} = 0.6$ for $d(\text{OO}) = 2.6 \text{ \AA}$, the resulting energy curve (Figure 2.8a: solid line with open diamonds) is in a very good agreement with the curve of system C (solid line). For $d(\text{OO}) = 3.0 \text{ \AA}$, using $q_{\text{O}} = -1.1$ and $q_{\text{H}} = 0.55$, the maximal error is about 5 kJ/mol (Figure 2.8b). The same remarkable agreement is also found for $d(\text{OO}) = 2.8 \text{ \AA}$ and $d(\text{OO}) = 3.5 \text{ \AA}$ (Table 2.2.5). Therefore, the interaction between the quantum mechanically treated part A and the point charges on the secondary waters can be reduced to the energy expression (2.2), in which the electronic structure part A is treated by point charges as well, here as Mulliken charges on the nucleic centers. But still the geometry and the Mulliken charges of the complete system C are being used.

The next step is now to employ the optimized geometry and the Mulliken charges of system A instead of system C. Only the positions of the secondary waters are taken from the optimized structure of system C. The electrostatic energy is calculated as described above and is added to the energy curve of system A. The result (solid line with filled circles) for $d(\text{OO}) = 2.6 \text{ \AA}$ deviates from the fully quantum mechanically calculated curve (solid line) by 7.6 kJ/mol for $E_{\text{b}}^{\rightarrow}$ and by 18.4 kJ/mol for $E_{\text{b}}^{\leftarrow}$. Compared to

the overall effect of the two secondary waters, which led to a reduction of the barrier of the isolated system by 82 kJ/mol for E_b^{\rightarrow} and an increase by 55 kJ/mol for E_b^{\leftarrow} , a deviation by 7.6 kJ/mol for E_b^{\rightarrow} is still satisfactory, whereas the discrepancy of 18.4 kJ/mol for E_b^{\leftarrow} is quite high. This tendency (the deviation for E_b^{\leftarrow} is larger than for E_b^{\rightarrow}) is also found at $d(\text{OO}) = 2.8 \text{ \AA}$ and $d(\text{OO}) = 3.0 \text{ \AA}$. It results from the strong interaction between the hydronium ion and the close secondary waters, which stabilize the ion by influencing its geometry and charge distribution. As $d(\text{OO})$ is increased, the error compared to the absolute effect of the environment reduces to about 10%, as shown in Table 2.2.5 for $d(\text{OO}) = 3.5 \text{ \AA}$.

In a final step of simplification, the calculation of the electrostatic energy is replaced by their first order Taylor expansion, the dipole approximation, as discussed by Borgis *et al.* [64] for the proton transfer between two water molecules under the influence of an external electric field \mathbf{E} . Expanding the potential Φ around a point \mathbf{r}_0 ,

$$\Phi(\mathbf{r}) = \Phi(\mathbf{r}_0) - (\mathbf{r} - \mathbf{r}_0)\mathbf{E}(\mathbf{r}_0) - \frac{1}{2} \sum_{k=x,y,z} \sum_{l=x,y,z} (r^k - r_0^k)(r^l - r_0^l) \frac{\partial E^l}{\partial r^k}(\mathbf{r}_0) + \dots \quad (2.3)$$

leads in first order to an electrostatic interaction energy

$$E_{el}^{(1)} = \sum_{i=1}^{13} Q_i \sum_{j=14}^{19} \frac{q_j}{4\pi\epsilon_0|\mathbf{r}_0 - \mathbf{r}_j|} - \sum_{i=1}^{13} Q_i(\mathbf{r}_i - \mathbf{r}_0) \sum_{j=14}^{19} \frac{q_j(\mathbf{r}_0 - \mathbf{r}_j)}{4\pi\epsilon_0|\mathbf{r}_0 - \mathbf{r}_j|^3} \quad (2.4)$$

This expression was evaluated for the optimized geometry and Mulliken charges of system C. The midpoint between the donating nitrogen atom of imidazole and accepting oxygen atom of the primary water was chosen as \mathbf{r}_0 . As can be seen in Figure 2.8 (solid line with open triangles), the deviations for this dipole approximation from the energy surface of system C, calculated by electronic structure method (solid line), are significantly larger than for the complete electrostatic interaction. The differences are $\Delta E_b^{\rightarrow} = 13.8 \text{ kJ/mol}$ and $\Delta E_b^{\leftarrow} = 19.6 \text{ kJ/mol}$. While the dipole approximation gives excellent results for system A under the influence of a homogenous electric field (results not shown), the effect of the inhomogeneous electric potential produced by secondary waters, in close distance to the donor-acceptor system, cannot be well described just using the dipole approximation. For $d(\text{OO}) \geq 3.5 \text{ \AA}$ the effect of the multipoles higher than the dipole decreases and energy barriers calculated with the electric field-dipole term approximate the effect of the complete electrostatic interaction. Thus the dipole approximation seems to be an efficient way to describe the effect of environmental

groups at distances of more than 3.5 Å from the accepting group, and higher multipoles should be used to describe the effect of secondary waters in hydrogen-bond distance.

2.2.6 Conclusions

The results presented show that the nature and the geometry of surrounding hydrogen bonded ligand waters significantly influence the energy surface of the proton transfer reaction. Thus one has to be aware of this effect when neglecting the environment of proton transferring groups and more generally of sites where chemical reactions happen. Since it is computationally cumbersome to use electronic structure methods to describe the chemically reactive groups together with a large environment, a QM/MM approach provides, in principle, an efficient alternative to treat such problems.

It is demonstrated that it is possible to determine the energy surface of a proton transfer reaction between an amino acid, here a histidine side chain, and a water molecule that is coordinated by secondary ligand waters, by calculating only the proton donating and accepting groups with electronic structure methods, while treating the environment, here the secondary waters, as point charges. The only correction necessary, due to polarization effects between the hydrogen bonded water molecules, is to enhance the charges of the point particle waters by a factor that depends on the distance between primary and secondary waters.

For environmental groups weakly hydrogen bonded to the accepting water molecule, one can go one step further. After calculating a wave function for the isolated system consisting of proton donor and acceptor, the effect of the environmental charges is added subsequently by calculating the electrostatic interaction energy between proton transferring groups and the surrounding point charges. It is thus possible to study the influence of different environments on a specific reaction in a very simple and fast manner.

Thus it is possible to compute the energy landscape only including the groups which are directly involved in the chemical reaction studied by electronic structure methods. Afterwards, the influence of a surrounding protein matrix or solvent can be added by just considering the electrostatic interaction between the reactive part and the environment. This will be used in the following section.

2.2.7 Towards a Compact Parameter Set

When the ideas of Q-HOP MD are taken into consideration, it is desirable to calculate the hopping probabilities in the framework of a classical force field approach. Having that in mind, the results presented in the last section led to the idea to calculate the vacuum energy curve of proton transfer by quantum chemical calculations before the simulation starts. This data is stored, and the environmental influence is calculated during the simulation by adding the electrostatic interactions between proton transferring system and environmental groups. The next sections (2.3 - 2.5) show that it is not necessary to know the full structure of the energy curve. A compact parameter set is derived to calculate the transfer probability very fast for every donor-acceptor system at every distance between donor and acceptor and for every configuration of the surrounding chemical groups.

First (Section 2.3) the situations with large energy barriers are considered. There, the proton transfer probability can be calculated with transition state theory (TST) including quantum corrections. For cases with no or only a tiny energy barrier, fit formulas are derived using the time dependent Schrödinger equation (Section 2.4). In Section 2.5, both approaches are merged to calculate hopping probabilities for the intermediate situations.

2.3 Large Energy Barriers

The general double-well shape of the energy curve (Figure 2.9) as a function of the donor-proton distance, $r(\text{DH})$, spurred several groups to fit these energy landscapes by analytical expressions [65]. For example, Scheiner and co-workers tested several functions to describe proton transport in H_5O_2^+ and N_2H_7^+ [58]. They found that the energy barrier at varying distance between proton donating and accepting atom can be well reproduced by Morse type potentials requiring only a few parameters for each donor-acceptor pair.

In the previous section, it was shown that the environmental groups can strongly affect the energy landscape of a proton transfer process (e.g. Figures 2.4, 2.5). The energy barrier depends strongly on the distance between primary, proton accepting water and secondary, ligand waters. Thus, for every environmental configuration one would

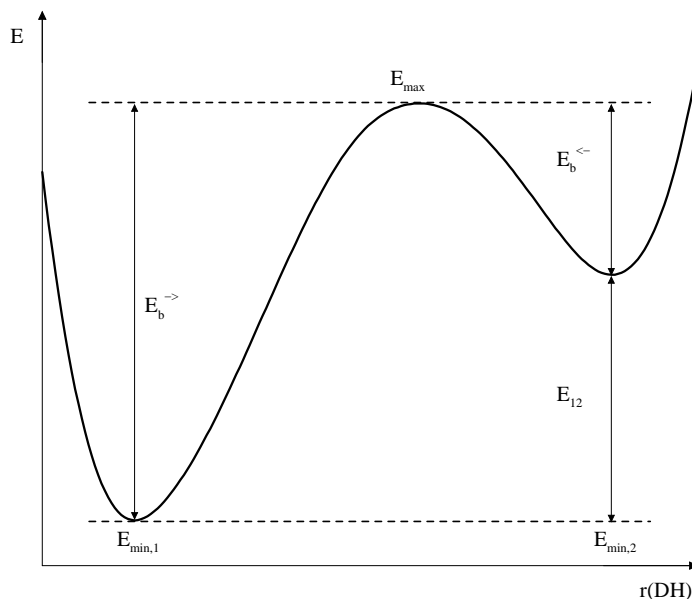


Figure 2.9: General double-well shape of the one-dimensional energy surface of proton transfer as function of the donor-proton distance, $r(\text{DH})$. $E_b^{>}$ and $E_b^{<}$ are the energy barriers in forward and backward direction. E_{12} is the relative energy of donor and acceptor bound states.

need a new set of parameters to analytically describe the one-dimensional transfer surface.

In the following, a different approach is presented. It takes into account the characteristic properties of the double-well potential. It follows the idea of proton hopping between a donating state where the proton is bound to the donating atom at about equilibrium distance, and an accepting state where it is bound to the accepting atom. For a given distance between donating and accepting atoms, $R(\text{DA})$, the energy barrier is estimated by an analytical expression using 10 fit parameters for every donor-acceptor pair and including the energetic influence of the environment on the donating and accepting state.

First it is shown that 7 parameters for every donor-acceptor pair suffice to determine the energy barrier from $R(\text{DA})$ and the energetic difference between donating and accepting state, E_{12} . Several exemplary cases show that these parameters simultaneously fit all donor-acceptor pairs that contain the same proton donating and accepting atoms. To obtain the energy difference between donating and accepting state under the

influence of the environment, the energetic difference E_{12}^0 in vacuum has to be known first. It is shown that E_{12}^0 can be expressed as a quadratic function of $R(\text{DA})$, for every donor-acceptor pair.

The quantum nature of the transferred proton is approximated at the level of an effective zero point energy along the reaction coordinate [66]. This treatment leads to a lowest vibrational energy level about 12-20 kJ/mol above the electronic potential minimum. The quantum correction can be expressed as function of E_b^{\rightarrow} using 3 adjustable parameters. Additionally, the quantum character of the proton is treated by adding tunneling corrections to transition state theory. A procedure analogous to simple transition state theory is chosen by averaging the quantum transmission coefficient $t(E)$ with the canonical equilibrium probability. $t(E)$ was numerically calculated for given energy surfaces using the WKB (Wentzel-Kramers-Brillouin) approximation method [68].

2.3.1 Method of Calculation

The ab initio calculations reported were again performed with the program package NWChem 3.1 [56] on Compaq alpha workstations at MP2/6-31G** // HF/6-31G** level. Figure 2.10 shows schematic views of the studied proton transferring systems. Donor (R_D -D-H) and acceptor (A- R_A) are $R(\text{DA})$ apart, with environmental ligand waters at a distance $d(\text{AO})$. To obtain the energy surface of a proton transfer reaction, the proton is moved stepwise by 0.1 Å along the D-A interconnecting line. At each $r(\text{DH})$, the geometry is optimized, keeping the distance $R(\text{DA})$, the angle $\angle(\text{DHA}) = 180^\circ$ and the ligand waters fixed. It is found that the preferred orientation of the primary water molecule with respect to the line, connecting donating and accepting atom, changes during the transfer step from a planar orientation as acceptor to a pyramidal orientation as donor, if the primary water is weakly bound to the secondary water molecules. Since it is expected the primary water molecule to be hydrogen-bonded to other groups most of the time, we fix its geometry to be pyramidal with respect to the donor-acceptor line. This is in contrast to the calculations presented in Section 2.2. Therefore, the energy curves vary slightly for the same $R(\text{DH})$ and $d(\text{AO})$ due to different electrostatic interactions. But this does not influence the main findings of Section 2.2. As demonstrated earlier [67], and as I confirmed for hydrogen-

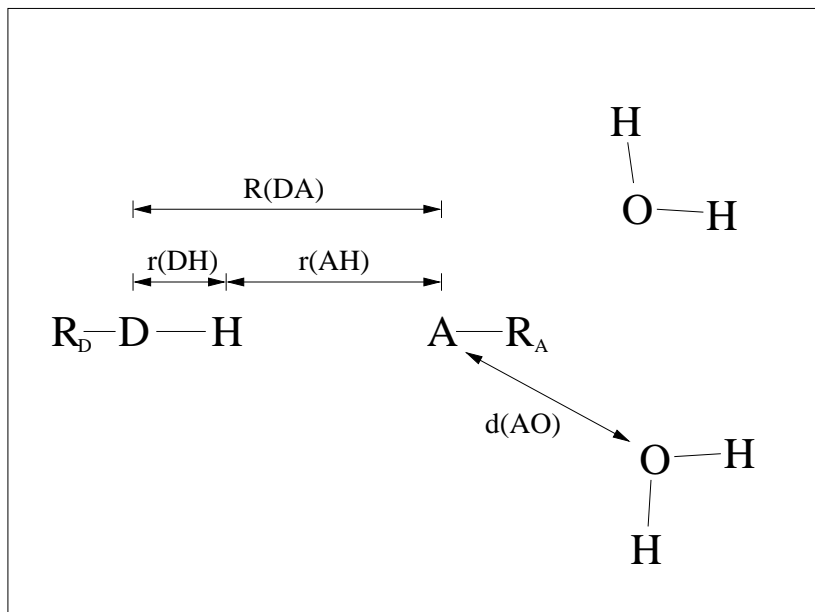


Figure 2.10: The proton transfer systems studied consist of donor R_D -D-H with proton donating atom D bound to a group R_D and transferring proton H, and of acceptor A- R_A with proton accepting atom A bound to a group R_A . To study the environmental influence on the proton transfer barrier, secondary ligand waters are present. Four interatomic distances $r(\text{AH})$, $r(\text{DH})$, $R(\text{DA})$ and $d(\text{AO})$, as labeled, are used as reaction criteria in this study. Water, ammonium, imidazole and ethanol are used as donating and accepting groups.

bonded primary waters (results not shown), this rigid-molecule assumption is in good agreement with the barriers obtained when optimizing the structure at every $r(\text{DH})$.

2.3.2 Motivation

After obtaining the energy curves, transition state theory [68, 66] can be applied, which gives for the transfer rate

$$k = \frac{k_B T}{2\pi\hbar Z_0} \exp\left(-\frac{E_b^{\rightarrow}}{k_B T}\right), \quad (2.5)$$

where Z_0 is the partition function of the educt state and E_b^{\rightarrow} is the energy barrier between educt and product state. Considering a quantum-mechanical harmonic oscillator with frequency ω_1 at the educt well minimum, then $1/Z_0 = 2\sinh(\hbar\omega_1/2k_B T)$. For

typical vibrational frequencies of bonds that contain a transferring proton (2000-3500 cm^{-1} for the studied O-H or N-H bonds), $\hbar\omega_1/2k_B T \gg 1$ and $1/Z_0 \simeq \exp(\hbar\omega_1/2k_B T)$. The rate becomes [66]:

$$k \simeq \frac{k_B T}{2\pi\hbar} \exp\left(-\frac{E_b^\rightarrow - \hbar\omega_1/2}{k_B T}\right) \quad . \quad (2.6)$$

In order to utilize this formula, the energy barriers and frequencies at the energy minimum positions have to be known. First the problem of determining the energy barrier is considered.

A simple approach to determine the energy barrier E_b^\rightarrow for a given snapshot of a simulation would be to pre-calculate and store the barrier height $E_b^{\rightarrow,0}$ in vacuum for every donor-acceptor pair at any distance $R(\text{DA})$. During the simulation, one would only need to add the influence of the environment onto the pre-calculated barrier. As previously shown (section 2.2.5)[47], this can be done by calculating the Coulombic interaction between suitable sets of point charges of the donor-acceptor pair and the environmental groups at the energy minimum $r(\text{DH})_{min}^0$, where the proton is bound to the donating atom, and at the energy maximum $r(\text{DH})_{max}^0$. However, these maximum positions $r(\text{DH})_{max}^0$, as well as the point charges at these positions would have to be stored for every donor-acceptor pair at every $R(\text{DA})$ as well.

The crux of this simple approach is demonstrated in Figure 2.11 that shows the energy surface of the system imidazole - H_2O at $R(\text{NO}) = 3.05 \text{ \AA}$ for different primary water-secondary water distances, $d(\text{OO})$. Obviously, the positions of the minima change very little with respect to $d(\text{OO})$. Namely by starting from the energy curve without secondary water and adding the influence of the secondary waters (solid arrow) in $r(\text{NH})_{min,1}^0$ and $r(\text{NH})_{min,2}^0$ gives new minima, which are in good agreement with the minima, when calculating the complete energy surface with environmental groups present (dashed arrow). However, the position of the maximum shifts significantly, from $r(\text{NH})_{max}^0 = 1.74 \text{ \AA}$ without secondary waters to $r(\text{NH})_{max} = 1.59 \text{ \AA}$ for $d(\text{OO}) = 2.6 \text{ \AA}$. Calculating the environmental effects at $r(\text{NH})_{max}^0 = 1.74 \text{ \AA}$ (solid arrow) would significantly underestimate the true barrier (dashed arrow).

A costly remedy of this problem would be to consider the environmental influence in many points along the $r(\text{NH})$ -coordinate as was discussed in Section 2.2.5. One would need to additionally store the charges of the donor-acceptor system along $r(\text{NH})$ and to calculate the Coulomb interaction energy between proton transferring system and

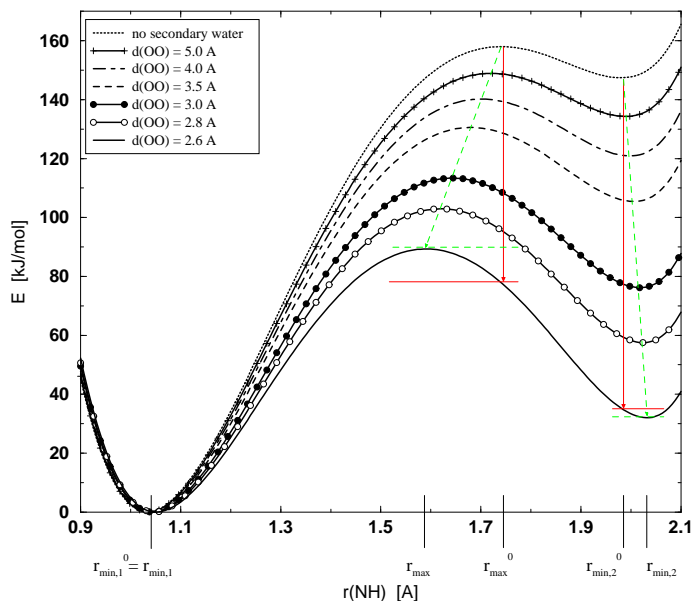


Figure 2.11: Dependence of energy barriers for the imidazole-water system on environmental effects. The distance between the two secondary, ligand waters relative to the primary water is varied from $d(\text{OO}) = 2.6 \text{ \AA}$ to $d(\text{OO}) = 5.0 \text{ \AA}$ and to infinity, which means that no secondary waters are present. All calculations were performed at fixed $R(\text{NO}) = 3.05 \text{ \AA}$. Calculating the environmental influence at the minima $r_{min,1}^0$ and $r_{min,2}^0$ (solid arrow) gives results very close to the correct energy minima at $r_{min,1}$ and $r_{min,2}$ (dashed arrow). Applying the same procedure for the maximum at r_{max}^0 strongly underestimates the barriers E_b^{\rightarrow} and E_b^{\leftarrow} .

environment along $r(\text{NH})$ during the simulation. Altogether this approach would be computationally cumbersome.

However, the ideas so far do not make use of the characteristic form of the double-well potential. As mentioned earlier, the positions of the minima on the $r(\text{NH})$ -axis change very little with shifting secondary water positions at given $R(\text{NO})$. By calculating the electrostatic interaction between donor-acceptor system and environment at given $r(\text{NH})_{min,1}^0$ and $r(\text{NH})_{min,2}^0$ and adding these energies to the energy minima $E_{min,1}^0$ and $E_{min,2}^0$ in vacuum, one arrives very close to the correct minima under environmental influence (see above). This view is in the spirit of empirical force field methods where the proton is always bound at equilibrium distance to either the donating or to the accepting atom. In dynamic simulations, the geometry of the donor-acceptor system

is known at each time step and the partial atomic charges are usually kept fixed. Both are thus directly available for the calculation of the environmental influence at $r(\text{NH})_{min,1}^0$ and $r(\text{NH})_{min,2}^0$. Thus, if the relative energies in vacuum $E_{min,1}^0$ and $E_{min,2}^0$ were known, the minima under environmental influence $E_{min,1}$ and $E_{min,2}$ could be easily determined.

Reconsidering Figure 2.11 shows that with decreasing energy difference between the minima $E_{12} = E_{min,2} - E_{min,1}$, the energy barrier $E_b^{\rightarrow} = E_{max} - E_{min,1}$ becomes smaller as well. This observation encourages to search for a relationship between E_b^{\rightarrow} and E_{12} that could be expressed analytically.

2.3.3 Relationship between E_b^{\rightarrow} and E_{12}

Figure 2.12 shows the energy barrier E_b^{\rightarrow} as a function of E_{12} at different $R(\text{DA})$ for the two systems $\text{NH}_4^+ - \text{H}_2\text{O}$ and $\text{H}_3\text{O}^+ - \text{H}_2\text{O}$ with two secondary waters for each donor and acceptor.

At each $R(\text{DA})$, the concave functional dependence can easily be fitted by a quadratic polynomial function using only three parameters S , T and V ,

$$E_b^{\rightarrow} = S + T E_{12} + V (E_{12})^2 \quad . \quad (2.7)$$

This representation has an analogy to the modified Marcus type equation [69] used by Sham *et al.* [70] to model proton translocation in proteins.

The mean errors of the fit, which are the averaged deviations between E_b^{\rightarrow} determined from quantum-chemical calculations and E_b^{\rightarrow} from the fit function, are 0.25 kJ/mol ($\text{H}_3\text{O}^+ - \text{H}_2\text{O}$), 0.29 kJ/mol ($\text{NH}_4^+ - \text{NH}_3$, data not shown) and 0.13 kJ/mol ($\text{NH}_4^+ - \text{H}_2\text{O}$). The maximum errors are 0.88 kJ/mol, 0.92 kJ/mol and 0.38 kJ/mol (in same order).

But S , T and V are different for every $R(\text{DA})$ and many parameters may be needed. Fortunately, one can observe characteristic tendencies in the functional form $E_b^{\rightarrow} = E_b^{\rightarrow}(E_{12})$: (1) the curvature decreases with increasing $R(\text{DA})$, (2) the slope is very similar for all curves, especially for high $R(\text{DA})$, and (3) the zero point value $E_b^{\rightarrow}(E_{12} = 0)$ increases continuously with $R(\text{DA})$. These observations prompted us to investigate whether S , T and V themselves can be expressed as functions of $R(\text{DA})$. The results are shown in figure 2.13 for the systems $\text{NH}_4^+ - \text{H}_2\text{O}$ (circles) and $\text{H}_3\text{O}^+ - \text{H}_2\text{O}$ (triangles).

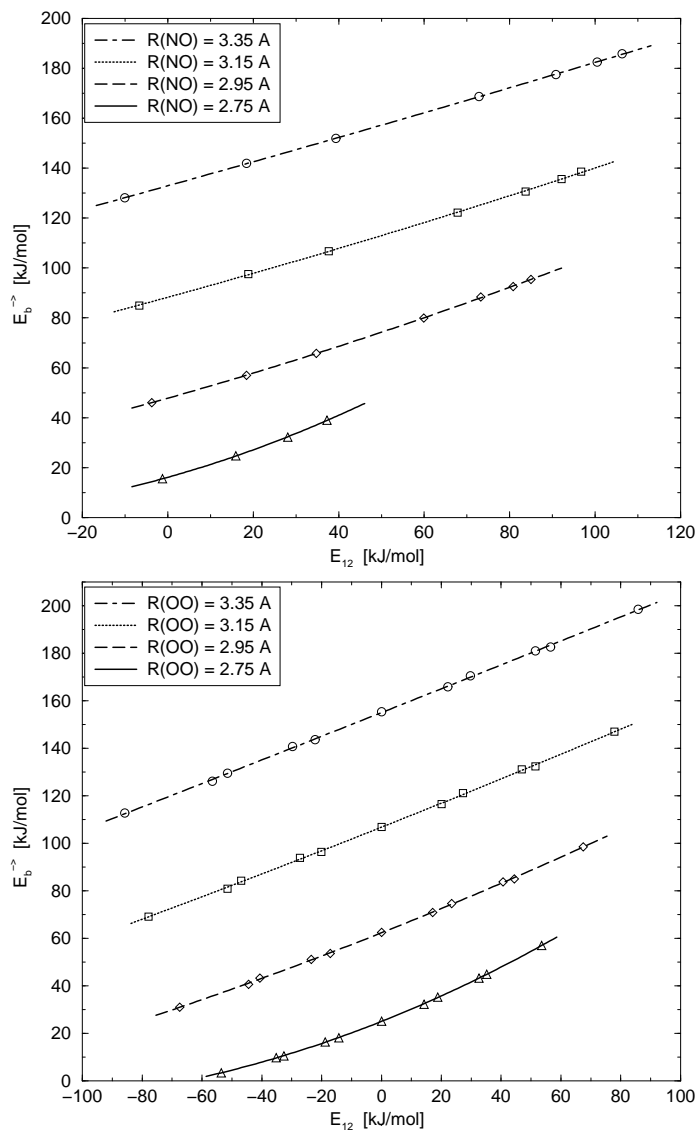


Figure 2.12: Proton transfer energy barrier E_b^{\rightarrow} as a function of the energy difference between the two energy minima, E_{12} , (a) for system $\text{NH}_4^+ - \text{H}_2\text{O}$ and (b) for system $\text{H}_3\text{O}^+ - \text{H}_2\text{O}$. In each case two secondary waters ligated to donor and two ligated to acceptor are included. The different values for E_b^{\rightarrow} are obtained by changing the distances between the secondary waters and donor as well as the distance between the secondary waters and acceptor.

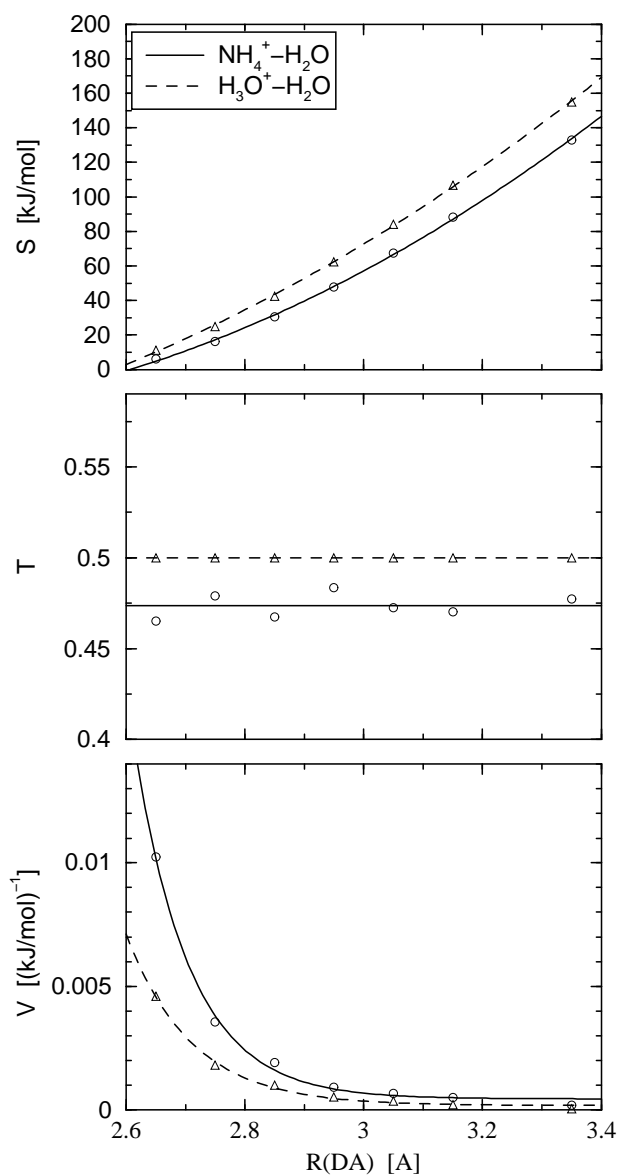


Figure 2.13: Parameters S , T and V as functions of the distance between donor and acceptor for the system $\text{NH}_4^+-\text{H}_2\text{O}$ (circles, solid line) and for the system $\text{H}_3\text{O}^+-\text{H}_2\text{O}$ (triangles, dashed line).

In both cases S shows quadratic dependence on $R(\text{DA})$, T is nearly constant and V is given in terms of $R(\text{DA})$ as an exponential function shifted along the ordinate. Thus, for a specific donor-acceptor pair, only seven parameters ($s_A, t_A, v_A, s_B, s_C, t_C, v_C$) are sufficient to calculate E_b^\rightarrow at given distance $R(\text{DA})$ and energy gap E_{12} ,

$$\begin{aligned} E_b^\rightarrow &= (s_A (R(\text{DA}) - t_A)^2 + v_A) \\ &+ s_B E_{12} \\ &+ (s_C \exp(-t_C(R(\text{DA}) - 2.0\text{\AA})) + v_C) (E_{12})^2 \quad . \end{aligned} \quad (2.8)$$

The mean errors of this fit are 0.88 kJ/mol ($\text{H}_3\text{O}^+ - \text{H}_2\text{O}$), 0.59 kJ/mol ($\text{NH}_4^+ - \text{NH}_3$) and 1.09 kJ/mol ($\text{NH}_4^+ - \text{H}_2\text{O}$). The maximum errors are 1.84 kJ/mol, 2.01 kJ/mol and 3.18 kJ/mol. The errors are obviously larger than for the first fit (Eq. 2.7), but are still considerably smaller than that of approximating the environmental influence as Coulombic interaction between classical point charges (4-17 kJ/mol) (see previous section) [47]. Also, the geometries in force-field based dynamic simulations may deviate from those in ab initio molecular dynamics simulations. Thus, the fit (Eq. 2.8) is satisfactory within a force field ansatz.

However, the number of parameters can be reduced even further. Figure 2.14 shows the energy surfaces of $\text{NH}_4^+ - \text{H}_2\text{O}$ and imidazole- H_2O under the influence of secondary waters. Both systems contain nitrogen as proton donating atom and oxygen as proton accepting atom. Yet, imidazole has a higher proton affinity than ammonia. Hence, for the same $d(\text{OO})$ the energy difference between the two minima, E_{12} , and therefore the energy barrier E_b^\rightarrow is higher for the imidazole- H_2O system than for $\text{NH}_4^+ - \text{H}_2\text{O}$. But considering energy curves with approximately the same E_{12} , for example $d(\text{OO}) = 3.0 \text{\AA}$ for $\text{NH}_4^+ - \text{H}_2\text{O}$ and $d(\text{OO}) = 2.6 \text{\AA}$ for imidazole- H_2O or $d(\text{OO}) = 4.0 \text{\AA}$ for $\text{NH}_4^+ - \text{H}_2\text{O}$ and $d(\text{OO}) = 3.0 \text{\AA}$ for imidazole- H_2O , both systems have E_b^\rightarrow values close to each other. The deviation is in the order of 4-8 kJ/mol. Thus, the same seven fit parameters ($s_A, t_A, v_A, s_B, s_C, t_C, v_C$) of system $\text{NH}_4^+ - \text{H}_2\text{O}$ are applicable to imidazole- H_2O as well.

This transfer of parameters between different donor-acceptor pairs, with the same proton binding atoms, was also verified for selected cases of imidazole- NH_3 compared to $\text{NH}_4^+ - \text{NH}_3$ and of ethanol- H_2O compared to $\text{H}_3\text{O}^+ - \text{H}_2\text{O}$. Here, ethanol serves as model of the serine side group (data not shown). The deviations lie again in the order of 4-8 kJ/mol.

Therefore, donor-acceptor pairs with the same proton donating and accepting atoms

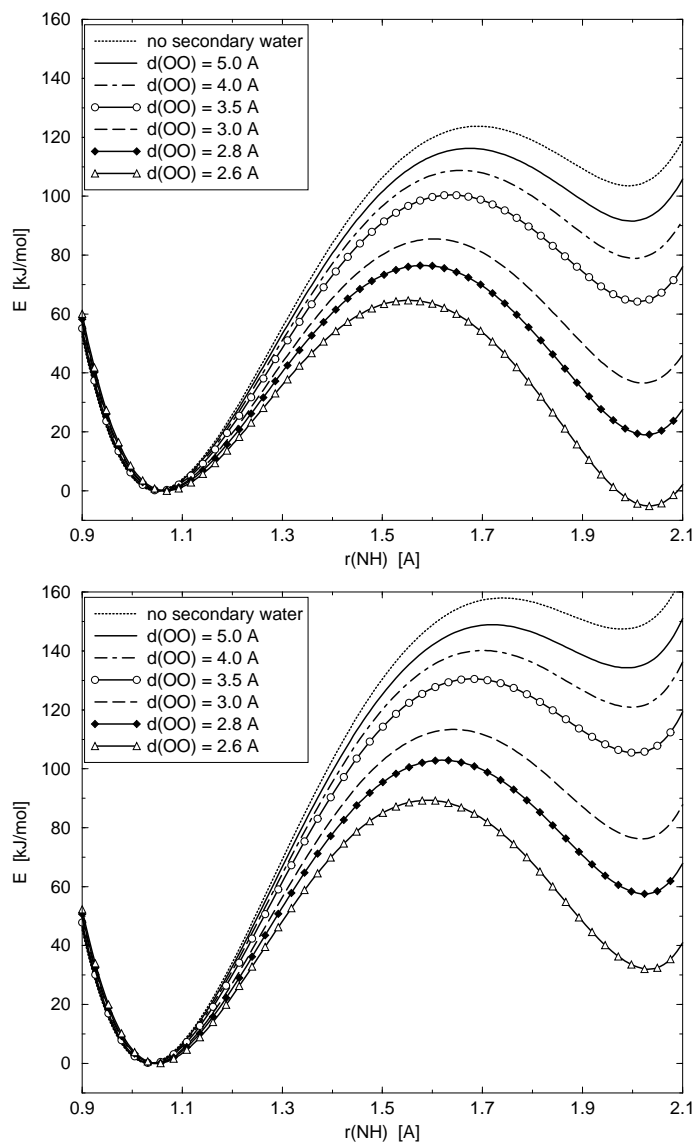


Figure 2.14: Proton transfer energy surfaces of (a) NH_4^+ - H_2O and (b) imidazole- H_2O under the influence of two ligand waters at different $d(\text{OO})$. (All curves at $R(\text{NO}) = 3.05 \text{ \AA}$.)

Table 2.4: Parameters s_A , t_A , v_A , s_B , s_C , t_C , and v_C to express E_b^{\rightarrow} as function of E_{12} and the donor-acceptor distance in equation (2.8) for different pairs of proton donating and accepting atoms. The parameters listed were obtained for the system $\text{H}_3\text{O}^+-\text{H}_2\text{O}$ (O-O), $\text{NH}_4^+-\text{H}_2\text{O}$ (O-N, N-O) and $\text{NH}_4^+-\text{NH}_3$ (N-N).

Proton donating atom	O	O	N	N
Proton accepting atom	O	N	O	N
s_A [kJ/mol \AA^{-2}]	83.1475	99.0598	99.0598	109.5189
t_A [\AA]	1.7509	2.0702	2.0702	2.2535
v_A [kJ/mol]	-57.0478	-28.5336	-28.5336	-15.0380
s_B	0.5000	0.5263	0.4737	0.4985
s_C [(kJ/mol) $^{-1}$]	1.6422	9.7837	9.7837	0.2877
t_C [\AA^{-1}]	9.1103	10.6300	10.6300	5.4425
v_C [(kJ/mol) $^{-1}$]	0.00016	0.00043	0.00043	0.00043

(e.g. Lys, Arg, His and Ser, Tyr, Asp, Glu, H_2O) seem to fall into one group. Thus, for each group, seven parameters (s_A , t_A , v_A , s_B , s_C , t_C , v_C) are sufficient to compute E_b^{\rightarrow} from given E_{12} and $R(\text{DA})$. The parameters obtained for different groups of donor-acceptor pairs are listed in Table 2.4.

From this point on, one only needs to determine $R(\text{DA})$ for a given snapshot during the simulation. Assuming that E_{12}^0 in vacuum is known at this $R(\text{DA})$, E_b^{\rightarrow} and thus the proton transfer probability can be easily derived from Equation (2.8) by including the environmental effect computed via a force field approach. The only remaining task is to express E_{12}^0 in terms of $R(\text{DA})$. That this is feasible has been previously shown by Scheiner and others [65].

2.3.4 Relationship between E_{12}^0 and $R(\text{DA})$

Figure 2.15a shows the energy surfaces of $\text{NH}_4^+-\text{H}_2\text{O}$ without secondary waters at different donor-acceptor distances $R(\text{NO})$. One minimum exists at $r(\text{NH}) = 1.05 \text{ \AA}$ and a second shallow minimum is formed for $R(\text{NO}) \geq 2.95 \text{ \AA}$. These second minima are plotted as open triangles. The resulting energy differences E_{12}^0 are drawn in Figure 2.15b

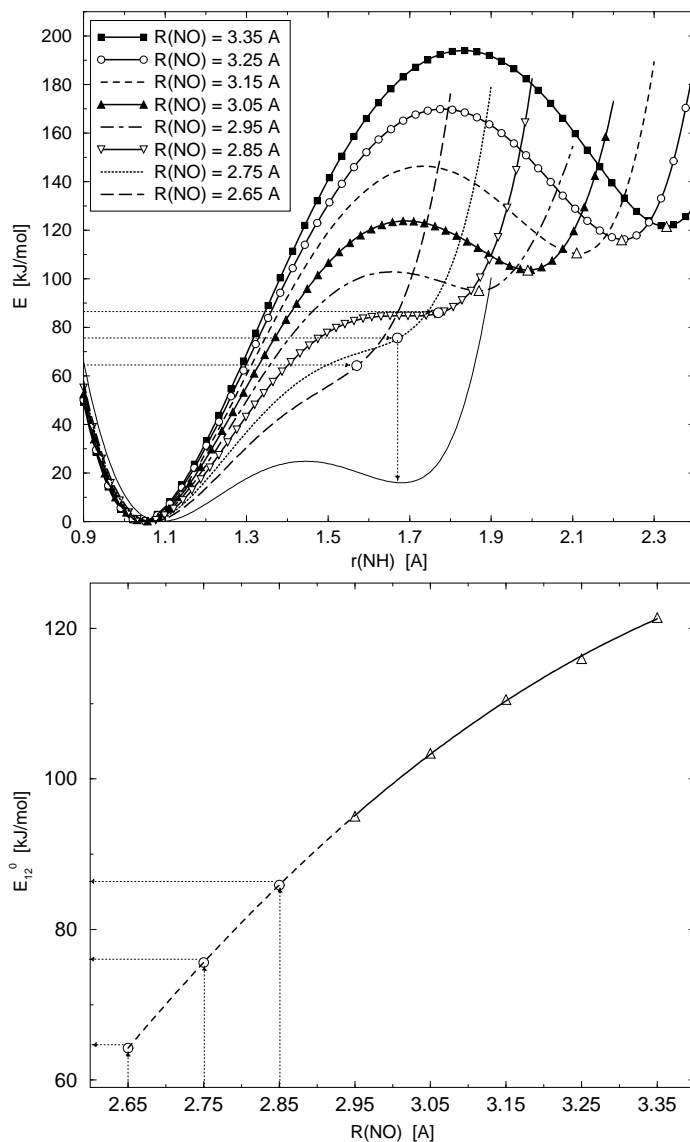


Figure 2.15: (a) Energy surface of proton transfer in vacuum for $\text{NH}_4^+ \text{-H}_2\text{O}$ at different separations between the nitrogen atom of ammonia and the oxygen atom of the primary water molecule, $R(\text{NO}) = 2.65 \text{ \AA} \dots 3.35 \text{ \AA}$. The minima $E_{min,2}^0$ are plotted as open triangles. (b) The resulting E_{12}^0 are shown as function of $R(\text{NO})$. The extrapolated (dashed line) values for $R(\text{NO}) \leq 2.85 \text{ \AA}$ (dotted arrows) are plotted in (a) as open circles. In addition, for $R(\text{NO}) = 2.75 \text{ \AA}$ the energy curve under the influence of two secondary waters at $d(\text{OO}) = 2.6 \text{ \AA}$ is shown (solid line).

as function of $R(\text{NO})$. Clearly, E_{12}^0 can be expressed as a convex, quadratic function

$$E_{12}^0 = \alpha + \beta R(\text{NO}) + \gamma R(\text{NO})^2 \quad . \quad (2.9)$$

A list of the fitted parameters α , β and γ for selected donor-acceptor pairs is found in Table 2.5. The mean error for the different systems lies in the order of 0.2 kJ/mol, and the maximum error in the order of 0.4 kJ/mol.

For $R(\text{NO}) \leq 2.85 \text{ \AA}$, no second minimum exists and therefore E_{12}^0 is not defined. But it is possible that environmental groups stabilize the product state, building a second minimum. Such a case is shown in Figure 2.15a (solid curve) for $R(\text{NO}) = 2.75 \text{ \AA}$. Two secondary waters, at 2.6 \AA distance from the primary water, stabilize the proton accepting state in such a strong way that a second minimum appears. However, one can also solve this problem with the quadratic fit function (2.9) by extrapolating $E_{12}^0 = E_{12}^0(R(\text{NO}))$ into the range $R(\text{NO}) < 2.95 \text{ \AA}$ as shown in Figure 2.15b (dashed line) by using the parameters α , β and γ . Following the dotted arrow at $R(\text{NO}) = 2.75 \text{ \AA}$ leads to $E_{12}^0 = 75.8 \text{ kJ/mol}$, a value that corresponds to a value of $r(\text{NH}) = 1.67 \text{ \AA}$ (open circle) in Figure 2.15a. This is very close to the position of the minimum of the energy surface of imidazole- H_2O with secondary waters. Thus the extrapolated energy value at $r(\text{NH}) = 1.67 \text{ \AA}$ (open circle) acts perfectly as E_{12}^0 prior to the calculation of the environmental influence.

Thus the function $E_{12}^0 = E_{12}^0(R(\text{DA}))$ with three parameters α , β and γ for every donor-acceptor pair can be used at any $R(\text{DA})$ to determine an energy value, E_{12}^0 , which is either the difference between two existing energy minima or which corresponds to the minimum position after considering product stabilization effects of the environment.

2.3.5 Test Case

In the preceding paragraphs, two fitting procedures were described to express E_b^\rightarrow as a function of E_{12} and $R(\text{DA})$, and E_{12}^0 as a function of $R(\text{DA})$. As an example of the implications of this study, the proton transfer from H_3O^+ to H_2O was chosen, now with three secondary waters coordinated to every primary water molecule. $R(\text{OO}) = 2.85 \text{ \AA}$; the three waters ligated to the donor are 2.8 \AA away from the oxygen atom of the donor, and the secondary waters of the acceptor are shifted from $d(\text{OO}) = 2.8 \text{ \AA}$ to $d(\text{OO}) = 3.5 \text{ \AA}$. The results are shown in Table 2.6. The fitted energy barriers E_b^\rightarrow deviate by

Table 2.5: Parameters α , β and γ to express E_{12}^0 as function of the donor-acceptor distance by equation (2.9) for different donor acceptor pairs. Not all possible pairs have been calculated yet. Histidine was modeled as an imidazole ring, aspartic acid and glutamic acid as acetic acid and serine as ethanol.

Donor		Acceptor						
		H ₂ O	OH ⁻	NH ₃	Ser ⁻	Asp ⁻	Glu ⁻	His
H ₃ O ⁺	α ^a	0.000	1114.600	629.549	1143.500	412.400	412.400	925.009
	β ^b	0.000	-752.800	-404.500	-774.200	-209.000	-209.000	-588.700
	γ ^c	0.000	80.683	53.832	92.198	0.000	0.000	77.757
H ₂ O	α	-1114.600	0.000		-326.209			
	β	752.800	0.000		206.838			
	γ	-80.683	0.000		-27.684			
NH ₄ ⁺	α	-629.549		0.000				220.693
	β	404.500		0.000				-137.623
	γ	-53.832		0.000				18.193
Ser	α	-1143.500	326.209		0.000	-757.000	-757.000	
	β	774.200	-206.838		0.000	502.900	502.900	
	γ	-92.198	27.684		0.000	-72.270	-72.270	
Asp-H	α	-412.400			757.000	0.000	0.000	
	β	209.000			-502.900	0.000	0.000	
	γ	0.000			72.270	0.000	0.000	
Glu-H	α	-412.400			757.000	0.000	0.000	
	β	209.000			-502.900	0.000	0.000	
	γ	0.000			72.270	0.000	0.000	
His-H ⁺	α	-925.009		-220.693				0.000
	β	588.700		137.623				0.000
	γ	-77.757		-18.193				0.000

^a[kJ/mol]

^b[kJ/molÅ⁻¹]

^c[kJ/mol Å⁻²]

about 1.2 kJ/mol from the actual barriers (calculated quantum mechanically), which is in the order of the mean error of the fitting procedure (0.88 kJ/mol). Thus, the energy barriers fitted from E_{12} with equation (2.8) yield good results for the important proton transfer system ($\text{H}_3\text{O}^+\text{-H}_2\text{O}$) with a different number of coordinating secondary waters.

Table 2.6: Comparison of the energy barriers predicted from equation (2.8) with the actual quantum chemical results for the system $\text{H}_3\text{O}^+ \text{H}_2\text{O}$ with three secondary waters at each donor and acceptor.

$d(OO_D)$ ^a [Å]	$d(OO_A)$ ^b [Å]	E_{12} [kJ/mol]	E_b^\rightarrow [kJ/mol]	$E_b^\rightarrow(\text{fit})$ [kJ/mol]
2.8	2.8	0.0	42.3	43.5
2.8	3.0	15.5	50.2	51.5
2.8	3.5	39.4	63.2	64.5

^aDistance between donor and secondary waters ligated to donating water molecule.

^bDistance between acceptor and secondary waters ligated to accepting water molecule.

2.3.6 Accounting for Spatial Uncertainty of the Proton along the Transfer Reaction Coordinate

The transferring proton as a quantum particle is spatially uncertain and the lowest energy level the proton can occupy therefore lies above the potential energy minimum. To account for this effect, the potential energy at the well is approximated to be harmonic. Consequently, the motion of the proton is that of a harmonic oscillator with frequency ω_1 (see Equation (2.6)), where $\omega_1 = \sqrt{(d^2U(r_{DH})/dr_{DH}^2)/m_H}$. The energy barrier is effectively decreased to $E_b^\rightarrow - \hbar\omega_1/2$. This harmonic approximation is expected to be valid for deep minima, but will overestimate the effect of the quantum nature of the proton for shallow minima with low energy barriers, since for these cases the energy potential becomes broader far from the potential minimum.

The frequencies at the well minima of the energy surfaces are derived from a fit of the energy curve and are typically in the range of 2000-3500 cm^{-1} (data not shown). Thus the classical energy barrier is reduced by about 12-20 kJ/mol, which means an increase in the reaction rate by a factor of about 100-3000.

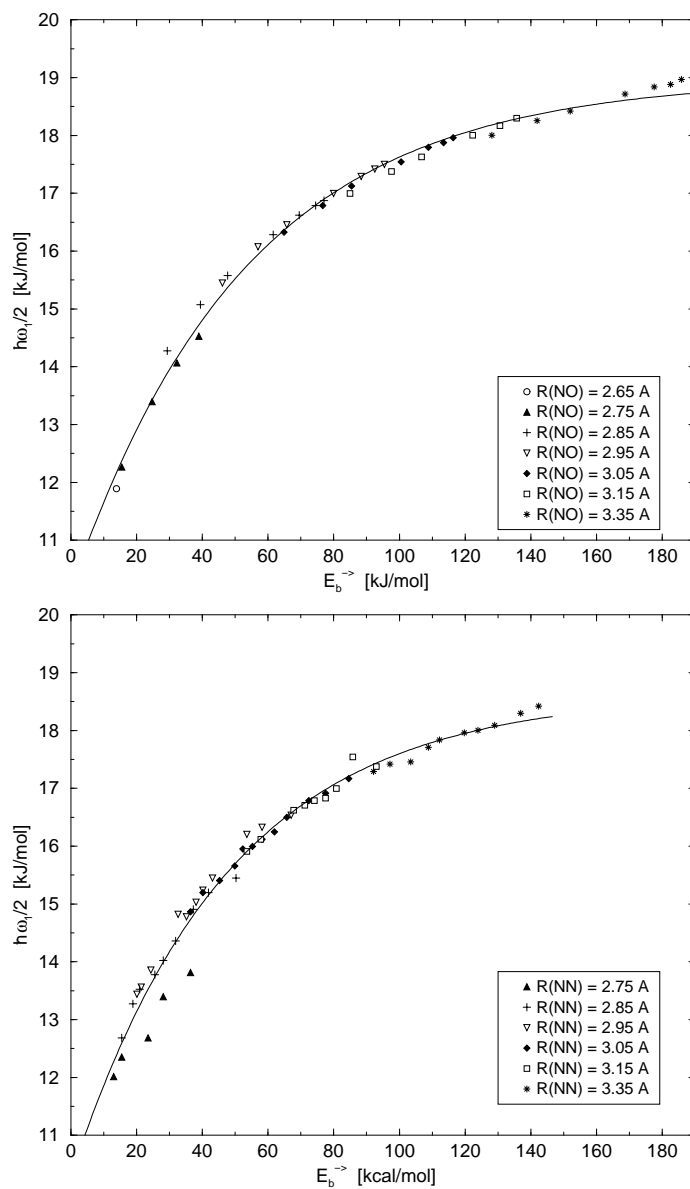


Figure 2.16: Energy $\hbar\omega_1/2$ as a function of the energy barrier between the two energy minima, E_b^{\rightarrow} , for systems (a) $\text{NH}_4^+-\text{H}_2\text{O}$ and (b) $\text{NH}_4^+-\text{NH}_3$.

Table 2.7: Parameters f , g and h in equation (2.10) to estimate a correction due to the quantum nature of the transferred proton for different proton donating atoms. The parameters are derived for the systems of Figure 2.16.

Proton donating atom	O	N
f [kJ/mol]	-9.34153	-8.34029
g [(kJ/mol) ⁻¹]	0.0184952	0.0208822
h [kJ/mol]	20.1035	18.63353

Figure 2.16 shows $\hbar\omega_1/2$ for systems $\text{NH}_4^+ \text{-H}_2\text{O}$ and $\text{NH}_4^+ \text{-NH}_3$ at different $R(\text{DA})$ as function of E_b^\ddagger . The value of $\hbar\omega_1/2$ increases with increasing E_b^\ddagger , since the curvature at the minimum position is higher for deep minima than for flat ones. The observed behavior can be fitted by an exponential function

$$\hbar\omega_1/2 = f \exp(-g E_b^\ddagger) + h \quad , \quad (2.10)$$

that only depends on the energy barrier E_b^\ddagger and that is independent of $R(\text{DA})$.

As can be seen from figure 2.16, the fitting functions of $\hbar\omega_1/2$ for the transfer from NH_4^+ to H_2O and from NH_4^+ to NH_3 are very similar. With a deviation of maximal 0.6 kJ/mol, the same three parameters can also be used to fit the transfer from imidazole to H_2O . The parameter transfer through all donating groups with the same donating atom was confirmed for the oxygen atom for the transfer from H_3O^+ to H_2O and from H_3O^+ to NH_3 . The mean deviation is about 0.17 kJ/mol, the maximal errors lie in the order of 0.63 kJ/mol. Thus, it is well justified to use the same three parameters for all donating groups with the same donating atom. The parameters for oxygen and nitrogen are found in Table 2.7.

2.3.7 Accounting for Tunneling in Transition State Theory

In the previous section, the quantum mechanical effects of the proton were considered at the level of an effective-zero point energy along the donor-acceptor interconnecting line. However, it was required that the proton has to overcome the full energy difference between ground-state (zero-point energy state) and energy barrier, and tunneling was not included in the model so far. This limitation is now going to be released and

three different tunneling models are investigated. Finally, a fit formula is derived that reproduces the results of the WKB approximation. The chosen procedure is analogous to simple transition state theory by averaging the quantum transmission coefficient $t(E)$ with the canonical equilibrium probability. $t(E)$ was numerically calculated for given energy surfaces using the WKB (Wentzel-Kramers-Brillouin) approximation method [68].

Following the idea of simple transition state theory, the transfer rate k_{qm} is obtained by averaging the quantum transmission coefficient $t(E)$ at energy E with the canonical equilibrium probability [68],

$$k_{qm} = \frac{1}{2\pi\hbar Z_0} \int_0^{\infty} t(E) \exp(-\beta E) dE \quad . \quad (2.11)$$

The transmission coefficient $t(E)$ for energy E may be calculated using the quasi-classical WKB approximation. Here $t(E)$ is given as

$$t(E) = \begin{cases} 0 & : E < \max(E_{min,1}, E_{min,2}) \\ \exp\left(-2 \int_{r_<}^{r_>} \sqrt{2m(V(r) - E)}/\hbar dr\right) & : \max(E_{min,1}, E_{min,2}) < E < E_{max} \\ 1 & : E > E_{max} \end{cases} \quad , \quad (2.12)$$

where $V(r)$ is the potential energy, and $r_<$ and $r_>$ are the left and right classical turning points ($E = V(r_{<,>})$). Substituting (2.12) into equation (2.11) gives

$$\begin{aligned} k_{qm} &\simeq \frac{k_B T}{2\pi\hbar} \exp\left(-\frac{E_b^{\rightarrow} - \hbar\omega_1/2}{k_B T}\right) \\ &+ \frac{1}{2\pi\hbar} \exp\left(\frac{\hbar\omega_1}{2k_B T}\right) \cdot \\ &\int_{\max(E_{min,1}, E_{min,2})}^{E_{max}} \exp\left(-2 \int_{r_<}^{r_>} \sqrt{2m(V(r) - E)}/\hbar dr\right) \exp\left(-\frac{E}{k_B T}\right) dE \quad , \end{aligned} \quad (2.13)$$

as the sum of the classical transfer rate including zero-point energy correction and the tunneling contribution. The latter was numerically calculated by integrating over all energies $E \in [\max(E_{min,1}, E_{min,2}), E_{max}]$ with the transmission coefficient $t(E)$ weighted by a canonical distribution. The transmission coefficient $t(E)$ at a given energy E is calculated by numerical integration from the left to the right classical turning point.

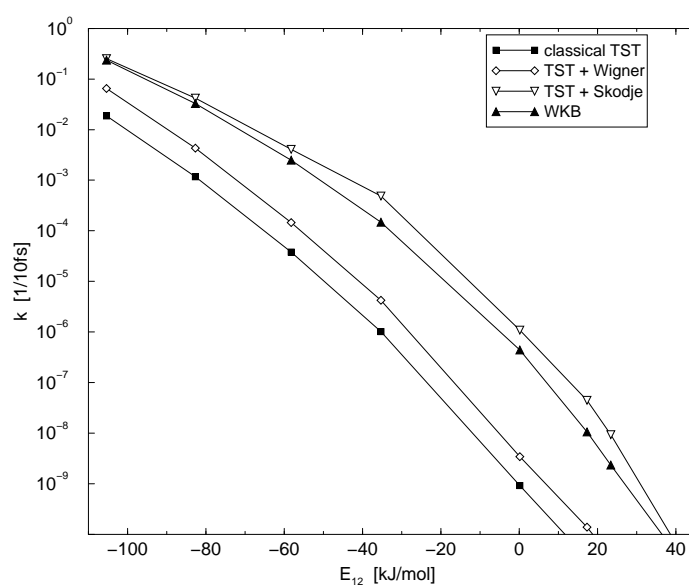


Figure 2.17: Proton transfer rate between two water molecules at $R(\text{OO}) = 2.95 \text{ \AA}$ for $T = 298 \text{ K}$ as a function of E_{12} . Together with the classical TST result (filled squares), the analytical approximations of Wigner [80] (open diamonds) and of Skodje and Truhlar [81] (open triangles down) as well as the numerical WKB based method (filled triangles up) are presented.

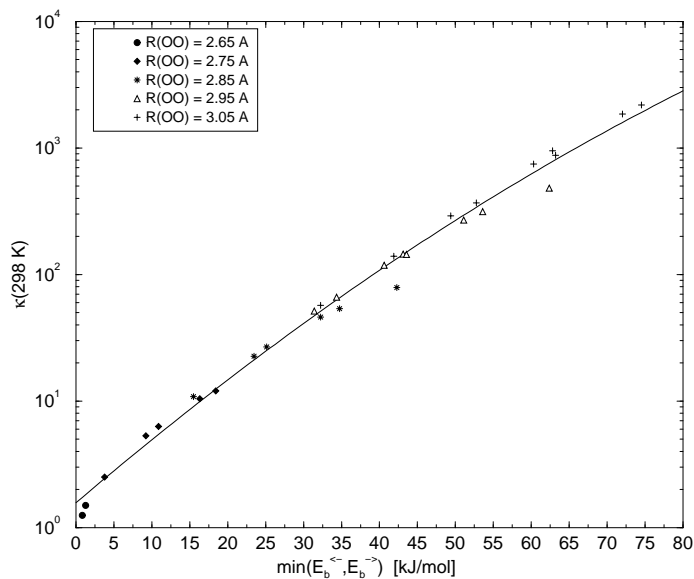


Figure 2.18: The enhancement of the classical transfer rate between two water molecules due to tunneling on WKB level at $T = 298 K$, $\kappa(298 K)$, at different separations $R(\text{OO})$, as a function of $E_M = \min(E_b^{\rightarrow}, E_b^{\leftarrow})$.

Figure 2.17 shows, as an example, the proton transfer rate between two water molecules at $R(\text{OO}) = 2.95 \text{ \AA}$ for $T = 298 K$. The transfer rate with tunneling increases compared to the classical case by about one to nearly three orders of magnitude. Thus, tunneling has a dramatic effect on the transfer rate even at this temperature. In addition to the numerical results, the analytical approximations of Wigner [80] and of Skodje and Truhlar [81] are shown. Whereas the Wigner approximation gives significantly smaller tunneling contributions to the rate, the approximation of Skodje *et al* gives results close to ours.

The classical transfer rate can easily be calculated by an approximate, analytical expression, as was shown earlier in the previous sections. The idea was now to find an expression to calculate the enhancement of the classical transfer rate due to tunneling, $\kappa(T) = k_{qm}/k_{classical}$. The results of $\kappa(298 K)$ for different $R(\text{OO})$ are shown in Figure 2.18. At 298 K tunneling increases the reaction rate by one to more than three orders of magnitude. The effect at smaller temperatures is even more dramatic (see following section). The results can be fitted (see solid line in Figure 2.18) by

$$\kappa(T) = \exp(P + QE_M + RE_M^2) \quad , \quad (2.14)$$

where $E_M = \min(E_b^{\rightarrow}, E_b^{\leftarrow})$. P , Q and R are fitting parameters that depend on the temperature.

2.3.8 Temperature Dependence of the Tunneling Contribution to the Transfer Rate

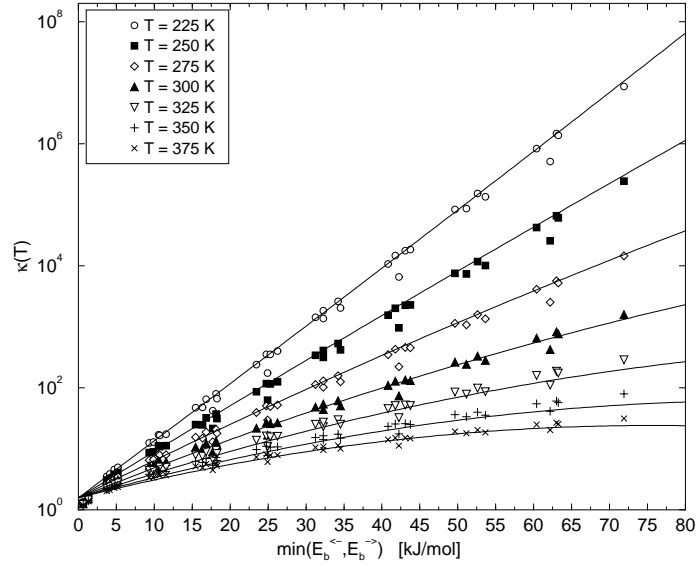


Figure 2.19: The enhancement $\kappa(T)$ of the classical transfer rate between two water molecules due to tunneling at WKB level, at different temperatures ranging from $T = 225\text{ K}$ to $T = 375\text{ K}$ as a function of $E_M = \min(E_b^{\rightarrow}, E_b^{\leftarrow})$.

Figure 2.19 shows the enhancement of the transfer rate due to tunneling, $\kappa(T)$, for the system $\text{H}_3\text{O}^+ \text{H}_2\text{O}$ at different temperatures. A strong temperature dependence is observed. Whereas $\kappa(T)$ takes on values in the order of 20 for $T = 375\text{ K}$, $\kappa(T)$ reaches 10^8 for $T = 225\text{ K}$. Each curve can be fitted by

$$\kappa(T) = \exp(P + QE_M + RE_M^2) \quad , \quad (2.15)$$

where $E_M = \min(E_b^{\rightarrow}, E_b^{\leftarrow})$. Q and R are fitting parameters which are temperature-dependent, which could be fitted by

$$P = p_1 \quad (2.16)$$

$$Q = q_1 + q_2T + q_3T^2 \quad (2.17)$$

$$R = r_1 + r_2T + r_3T^2 \quad . \quad (2.18)$$

Table 2.8: Fitted parameters p_1 , q_1 , q_2 , q_3 , r_1 , r_2 and r_3 in equations (2.16), (2.17) and (2.18) to calculate $\kappa(T)$ in equation (2.15).

Proton transfer system	O \rightarrow O	O \rightarrow N	N \rightarrow O
p_1 [1]	0.45	0.30	0.33
q_1 [(kJ/mol) $^{-1}$]	0.835	0.825	0.821
q_2 [(kJ/mol) $^{-1}$ K $^{-1}$]	$-3.848 \cdot 10^{-3}$	$-3.714 \cdot 10^{-3}$	$-3.692 \cdot 10^{-3}$
q_3 [(kJ/mol) $^{-1}$ K $^{-2}$]	$4.836 \cdot 10^{-6}$	$4.624 \cdot 10^{-6}$	$4.588 \cdot 10^{-6}$
r_1 [(kJ/mol) $^{-2}$]	$2.429 \cdot 10^{-3}$	$2.494 \cdot 10^{-3}$	$2.513 \cdot 10^{-3}$
r_2 [(kJ/mol) $^{-2}$ K $^{-1}$]	$-1.464 \cdot 10^{-5}$	$-1.794 \cdot 10^{-5}$	$-1.806 \cdot 10^{-5}$
r_3 [(kJ/mol) $^{-2}$ K $^{-2}$]	$1.836 \cdot 10^{-8}$	$2.481 \cdot 10^{-8}$	$2.517 \cdot 10^{-8}$

The parameters for p_1 , q_1 , q_2 , q_3 , r_1 , r_2 and r_3 are listed in Table 2.8.

2.3.9 Conclusions

By including environmental effects this approach significantly extends previous work to fit proton transfer barriers. In the past, Scheiner and others successfully derived formulas to express the exact shapes of the energy surface as a function of the donor-acceptor distance. Here, we focus on the relative positions of the barrier maximum and minima under the influence of environmental groups.

A fitting procedure was developed, which allows the fast calculation of energy barriers of proton transfer between all possible donor-acceptor pairs in biological systems under the influence of surrounding chemical groups. The only information required is the distance between proton donating and accepting atom, $R(\text{DA})$, and the environmental influence in the acceptor and donor bound states, $E_{el,1}$ and $E_{el,2}$.

Three parameters, α , β and γ , for every donor-acceptor pair suffice to compute the vacuum energy difference between the minima, E_{12}^0 , at a given distance between donor and acceptor, $R(\text{DA})$. The energy difference between the minima under environmental influence is then $E_{12} = E_{12}^0 + E_{el,2} - E_{el,1}$.

Finally, seven parameters (s_A , t_A , v_A , s_B , s_C , t_C , v_C) are sufficient to calculate the energy barrier E_b^\rightarrow from E_{12} and $R(\text{DA})$, and three parameters (f , g , h) are required for the calculation of a quantum correction $\hbar\omega_1/2$.

Hopping probabilities can now be efficiently computed using transition state theory with tunneling correction

$$p \simeq \kappa(t) \frac{k_B T}{2\pi\hbar} \exp\left(-\frac{E_b^{\rightarrow} - \hbar\omega_1/2}{k_B T}\right) \Delta t \quad (2.19)$$

for any given donor-acceptor distance by including the electrostatic interaction with the remaining system at both sites. This interaction can be easily computed in molecular dynamics simulations or is already known at the current time step. The hydrogen bonds are chosen to be linear and considered as optimal geometries leading to the lowest transfer barriers. As shown by Scheiner and coworkers previously [67], geometric deformation of the donating or accepting group towards the proton transfer axis leads to increasing transfer barriers. These cases are not considered explicitly here, since in dynamic simulations fluctuating geometric arrangements between donor and acceptor are expected, which will include optimal configurations with a maximal transfer rate.

On the other hand, the formalism can be combined with more elaborate Poisson-Boltzmann or Langevin Dipole calculations. It may also be useful in other contexts, e.g. to predict enzymatic rates of proton transfer from suitable crystal structures and, in particular, to efficiently estimate the effects of protein mutations.

2.4 Small or no Energy Barriers

As mentioned earlier (Section 2.2.2), the energy barrier of proton transfer is lowered with decreasing donor-acceptor distance. At some distance the barrier even disappears. This is demonstrated in Figure 2.20 for proton transfer between two water molecules at different donor-acceptor distances. In this example, the barrier and a second minimum disappear between $R(\text{OO}) = 2.55 \text{ \AA}$ and $R(\text{OO}) = 2.60 \text{ \AA}$. For these cases, simple transition state theory is clearly not applicable. Transfer at such short distances is, however, frequently observed in molecular dynamics simulations of an excess proton in a solvent box [34, 35, 31, 32, 33, 50].

In the following, such cases of very small or even absent energy barriers are studied by numerically solving the time-dependent Schrödinger equation in one dimension. Following the time evolution of the probability density of the proton between donor and acceptor, a definition is presented for the intrinsic transfer rate at the momentary

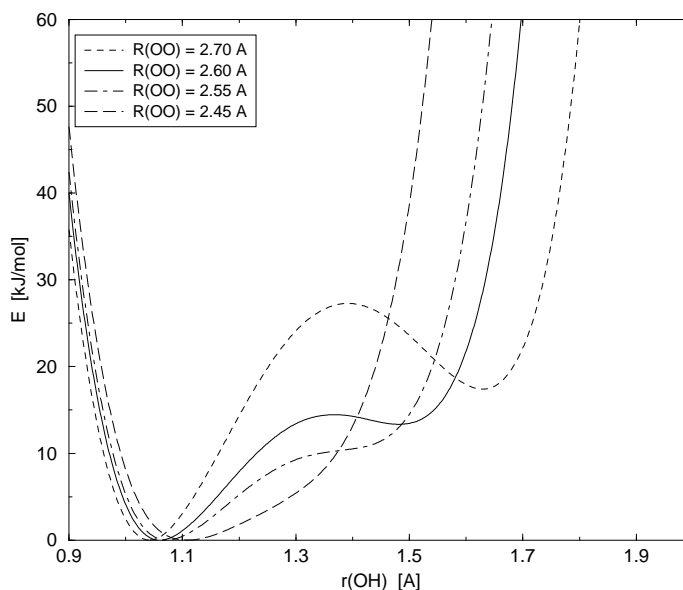


Figure 2.20: Proton transfer energy surface for system $\text{H}_3\text{O}^+ \text{H}_2\text{O}$ as a function of $r(\text{OH})$ at different separations between donating oxygen and accepting oxygen, $R(\text{OO}) = 2.45 \text{ \AA} \dots 2.70 \text{ \AA}$.

configuration of donor, acceptor and environment not accounting for solvent reorganization. An analytical fit function is derived to again quickly calculate the transfer rate as a function of E_{12} and $R(\text{DA})$ using six adjustable parameters. This approach seems to be reliable for small $R(\text{DA})$ and small E_{12} . For large $R(\text{DA})$ and large E_{12} on the other hand (for details see Figure 2.26 and corresponding text) transition state theory is valid [68, 76].

2.4.1 Method of Calculation

Energy Surfaces: The one-dimensional energy surfaces of the proton transfer reactions between various small molecules were obtained as in the previous work (Sections 2.1 - 2.3) from ab initio calculations. These were carried out with the program package NWChem 3.1 [56] on Compaq alpha workstations at MP2/6-31G**// HF/6-31G** level.

Time-dependent Schrödinger Equation: Initial wavefunctions for the time-dependent Schrödinger equation were constructed from the ground state solutions of the time-independent Schrödinger equation for the donor-bound well of the symmetric

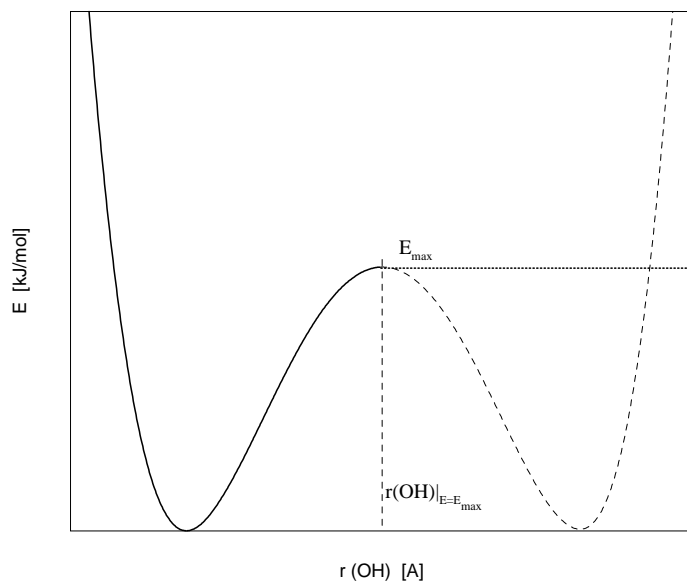


Figure 2.21: Modified proton transfer energy surface to calculate an initial wavefunction for the time-dependent Schrödinger equation. The energy curve of proton transfer was cut at $r(\text{OH})|_{E=E_{max}}$ and consecutively the right half of the energy curve was set to E_{max} (dotted line).

potential for $R(\text{OO}) = 2.75 \text{ \AA}$, $R(\text{OO}) = 2.85 \text{ \AA}$, $R(\text{OO}) = 3.05 \text{ \AA}$ and $R(\text{OO}) = 3.15 \text{ \AA}$. Precisely, the energy curve of proton transfer was cut at $r(\text{OH})|_{E=E_{max}}$ and the right half of the energy curve following $r(\text{OH})|_{E=E_{max}}$ was set to E_{max} (see Figure 2.21). For this modified energy curve, the time-independent Schrödinger equation was solved numerically. $\Psi_{2.75}^0$, $\Psi_{2.85}^0$, $\Psi_{3.05}^0$ and $\Psi_{3.15}^0$ denote the resulting initial wavefunctions. The time-dependent Schrödinger equation was then solved numerically along the proton transfer energy curve using a 0.001 fs time step. To do so, the wavefunction was represented on a grid (here 1000 grid points with a spacing of 0.0017 Å). The finite-difference representation of the evolution operator $\exp(-iHt)$ was approximated using Cayley's form leading to the implicit Crank-Nicholson method, which is unitary and second order accurate in space and time [77]. The program that is used is based on a program package written by Marten [78].

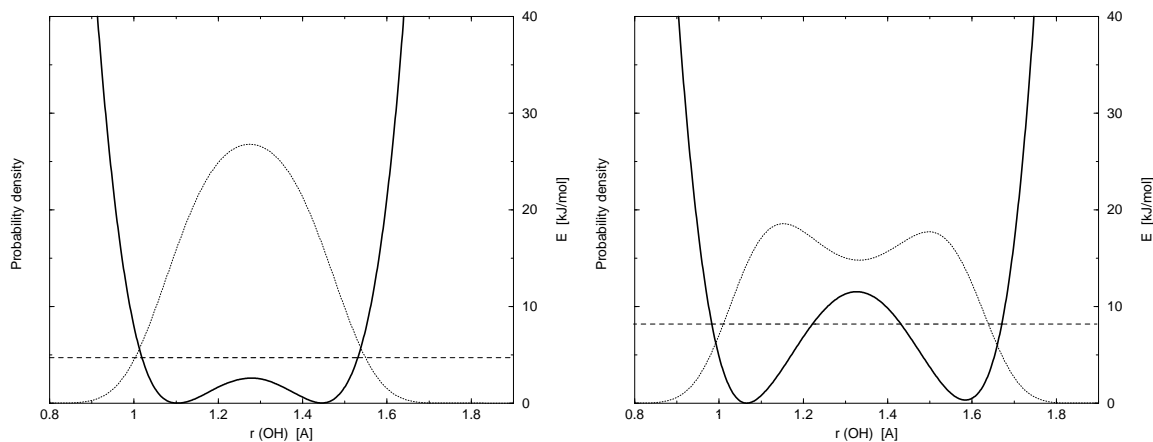


Figure 2.22: Probability density $|\Psi(r)|^2$ (dotted line) and ground state energy (dashed line) of the solution of the time-independent Schrödinger equation in the potential (solid line) of symmetric proton transfer at (a) $R(\text{OO}) = 2.55 \text{ \AA}$ and (b) $R(\text{OO}) = 2.65 \text{ \AA}$.

2.4.2 Time-independent Schrödinger Equation

In proton transfer reactions over small distances where the zero-point energy is larger than the energy barrier or where even no barrier exists, the proton can be considered de-localized over the region between donor and acceptor in a quantum mechanical treatment. On the other hand, in our new force-field based simulation method Q-HOP MD [50], the location of the proton is represented by a point in space, and one has to decide at every step whether the proton belongs to the donor or to the acceptor atom. One possible approach to determine the probability of finding the proton bound either to the donor or to the acceptor would be to solve the time-independent Schrödinger equation [30].

This approach is, however, a stationary picture, which does not give information about the time scales of the transfer process from donor to acceptor. To make this point clear, Figure 2.22 shows the situation for the transfer of an excess proton between two water molecules at $R(\text{OO}) = 2.55 \text{ \AA}$ and $R(\text{OO}) = 2.65 \text{ \AA}$. In the former case, the lowest energy state lies above the barrier, whereas in the latter, it is 2.5 kJ/mol below the barrier. Thus in the second case, the proton has to overcome an effective barrier of about thermal energy at room temperature. Therefore one would expect that the transfer rate is smaller for $R(\text{OO}) = 2.65 \text{ \AA}$ than for $R(\text{OO}) = 2.55 \text{ \AA}$. However,

solving the time-independent Schrödinger equation leads in both cases to a symmetric probability distribution with respect to donor and acceptor bound state. Consequently, for both cases the probability for a transfer of the proton would be the same, namely 0.5. Non-adiabatic effects are not included in this picture.

2.4.3 Time-dependent Schrödinger Equation

To obtain insight into the dynamics of proton transfer for small donor-acceptor distances, the time-dependent Schrödinger equation was solved numerically (see Section 2.4.1). The distance between donor and acceptor was fixed, as well as the environmental configuration. This concept of frozen solvent during proton transfer in strongly hydrogen-bonded systems is supported by a study by Tuñón *et al* [79] (Section 4.1) and the approximation of fixed donor-acceptor distance is supported by a molecular dynamics study [50] on proton transfer of an excess proton in a water box. Here, it is shown that the donor-acceptor distance changes typically by only 0.05 Å during 10 fs; and 10 fs is characteristic time-scale for proton transfer in strongly hydrogen-bonded systems, as shown in the following.

Figure 2.23 shows a typical example of the time evolution of the probability density of a proton wave packet in a double-well potential. Here, donor and acceptor atom are $R(\text{OO}) = 2.55 \text{ \AA}$ apart and the environment generates an energy difference between acceptor and donor bound state of approximately 8 kJ/mol. Two initial wavefunctions are compared, $\Psi_{2.85}^0$ (long dashed line) and $\Psi_{3.15}^0$ (dotted line). The wavepacket moves first in the direction of the acceptor and broadens. After 10 fs a second maximum is formed at the acceptor bound site. After 15 fs to 17.5 fs the maximum at the donor site disappears. The wavepacket has transferred to the acceptor site. After 20 fs the probability density starts to swap backwards to the donor site, because no damping or stabilization effects due to solvent or donor-acceptor dynamics are included in the model system. The reason behind this is that we are only interested in the time scale and probability for a first transfer of the proton from donor to acceptor. In Fig. 2.23, the wavepacket $\Psi_{3.15}$ moves faster towards the acceptor site than $\Psi_{2.85}$, because it has a higher initial kinetic energy. $\Psi_{3.05}$ shows an evolution (not shown here) which lies between $\Psi_{3.15}$ and $\Psi_{2.85}$, whereas $\Psi_{2.75}$ evolves very similar to $\Psi_{2.85}$ (not shown). The latter observation suggests to use $\Psi_{2.85}^0$ as starting wavefunction for the following

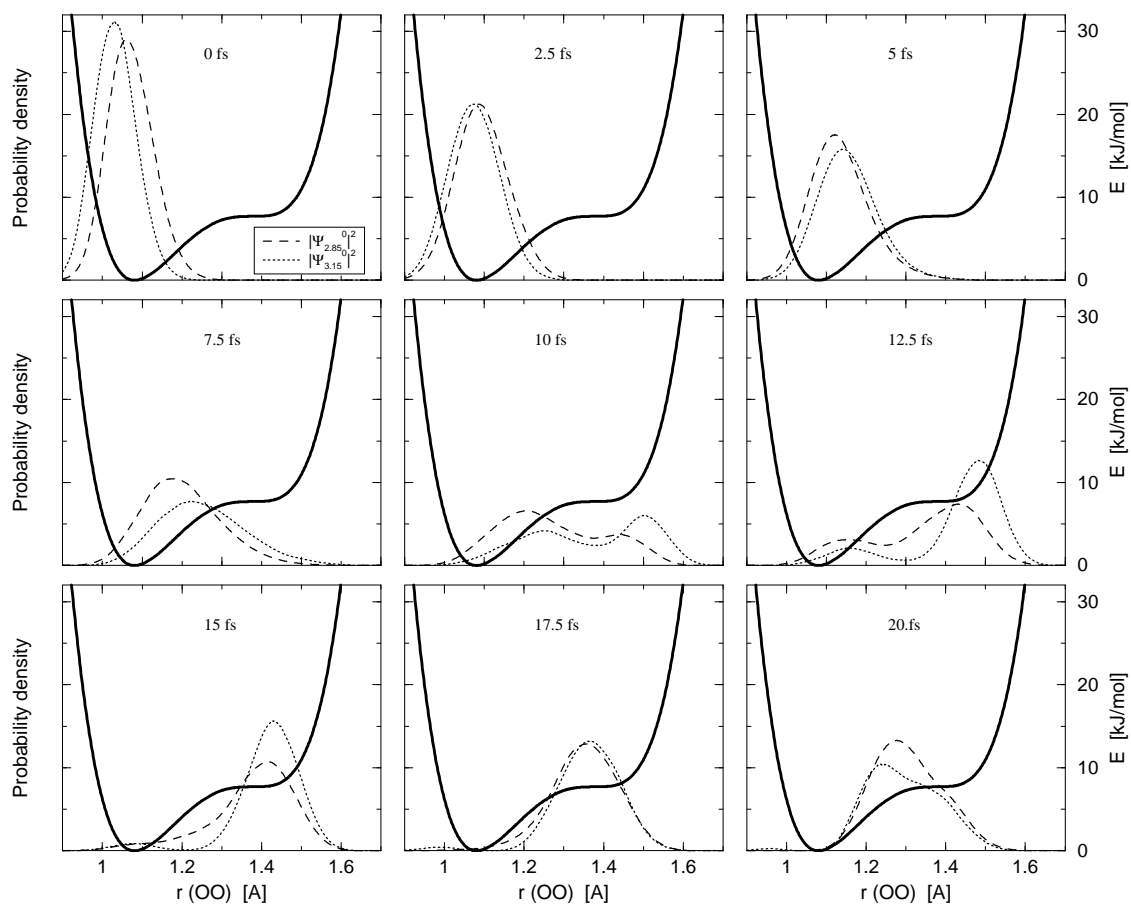


Figure 2.23: Time evolution of the probability density of a proton according to the numerical solution of the time-dependent Schrödinger equation. Donor and acceptor atoms are $R(\text{OO}) = 2.55 \text{ \AA}$ apart. The environment generates a potential energy surface (solid line) with energy difference E_{12} of approximately 8 kJ/mol. $\Psi_{2.85}^0$ (long dashed line) and $\Psi_{3.15}^0$ (short dashed line) are the initial wavefunctions. The time-interval between two consecutive snapshots is 2.5 fs.

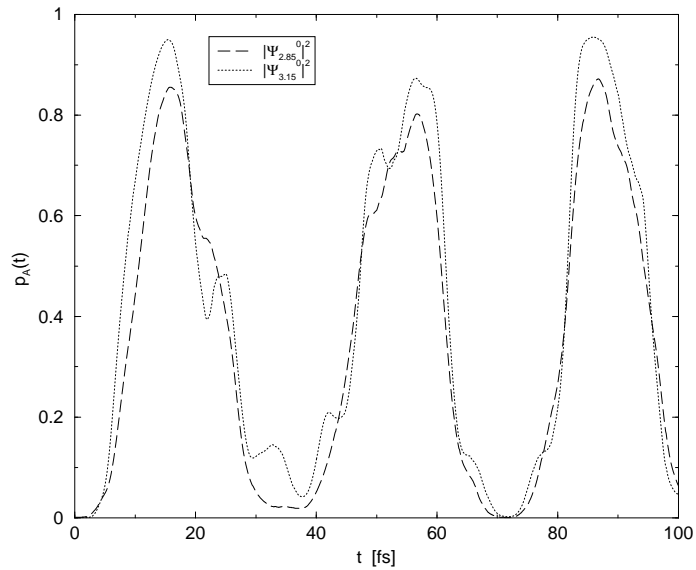


Figure 2.24: Time evolution of the probability to find the proton on the right side of the midpoint between donor and acceptor, at the acceptor bound site, for the example of Figure 5.

studies.

To determine the rate of proton transfer, the time evolution of the probability density $p_A(t)$ was examined to find the proton on the right side of the midpoint between donor and acceptor:

$$p_A(t) = \int_{R(OO)/2}^{\infty} |\Psi(r, t)|^2 dr \quad . \quad (2.20)$$

The results for the example in Figure 2.23 are shown in Figure 2.24. While the wavepacket is moving towards the acceptor, $p_A(t)$ increases to a maximum value at about 16 fs which corresponds to the situation where the wavepacket is localized at the acceptor site (compare Figure 2.23). Then, the density starts to move back to the donor site, and $p_A(t)$ decreases. Since the wavepacket $\Psi_{3.15}$ moves faster towards the acceptor site than $\Psi_{2.85}$, $p_{A,3.15}(t)$ increases faster than $p_{A,2.85}(t)$, and also the maximum value is larger for $p_{A,3.15}(t)$.

To obtain a timescale for proton transfer, Table 2.9 shows the time, $p_A(t)$ needed to reach its first maximum value for different E_{12} at $R(OO) = 2.55 \text{ \AA}$. Table 2.10 shows the same for different donor-acceptor distances at $E_{12} = 0$. In conclusion, the transfer of a proton in a strongly hydrogen bonded system happens during a time interval of 10 - 40 fs. The time interval becomes longer with increasing $R(OO)$.

Table 2.9: Time t_{max} that $p_A(t)$ needs to reach its maximum $p_{A,max}$ for different E_{12} . The values for the probability maxima $p_{A,max}$ are also printed.

E_{12} [kJ/mol]	t_{max} [fs]	$p_{A,max}$
48.0	11	0.02
31.5	12	0.19
29.5	13	0.23
21.0	15	0.48
14.5	16	0.68
10.5	16	0.86
6.5	16	0.96
0.0	17	0.98
-6.5	15	0.97
-10.5	14	0.96

Therefore, the transfer rate is defined as the probability $p_A(10\text{ fs})$ divided by 10 fs. This probability $p_A(10\text{ fs})$ as a function of E_{12} for different $R(\text{OO})$ is shown in Figure 2.25. At given $R(\text{OO})$, the transfer to the acceptor becomes more unfavorable with increasing E_{12} . Thus $p_A(10\text{ fs})$ decreases with increasing E_{12} . Also, with increasing $R(\text{OO})$ the wavepacket needs longer time to reach the accepting site, and for $E_{12} > 0$ the maximal probability $p_{A,max}$ becomes smaller. Therefore $p_A(10\text{ fs})$ decreases with increasing $R(\text{OO})$. Observing the characteristic form of $p_A(10\text{ fs})$ as a function of E_{12} and taking into account the behavior at the limits ($\lim_{E_{12} \rightarrow -\infty} p_A(10\text{ fs}) = 1$, $\lim_{E_{12} \rightarrow +\infty} p_A(10\text{ fs}) = 0$) prompted me to fit the curves by

$$p_A(10\text{ fs}) = 0.5 \tanh(-K E_{12} + M) + 0.5 \quad (2.21)$$

with fitting parameters K and M , that depend on $R(\text{DA})$. K and M as a function of $R(\text{DA})$ are expressed as

$$K = k_1 \exp(k_2 (R(\text{DA}) - 2.3\text{\AA})) + k_3 \quad (2.22)$$

$$M = m_1 \exp(m_2 (R(\text{DA}) - 2.3\text{\AA})) + m_3 \quad (2.23)$$

The fitting parameters obtained for the systems $\text{H}_3\text{O}^+ - \text{H}_2\text{O}$ ($\text{O} \rightarrow \text{O}$) and $\text{NH}_4^+ - \text{H}_2\text{O}$

Table 2.10: Time t_{max} that $p_A(t)$ needs to reach its maximum $p_{A,max}$ for different $R(OO)$; $E_{12} = 0$. The values for the probability maxima $p_{A,max}$ are also shown.

$R(OO)$ [Å]	t_{max} [fs]	$p_{A,max}$
2.45	11	0.99
2.55	17	0.98
2.60	31	0.97
2.65	44	1.00

(N→O, O→N) are listed in Table 2.11. The resulting fitting curves (shown as lines in Figure 2.25) are in good agreement with the calculated values (symbols).

Table 2.11: Fitted parameters k_1 , k_2 , k_3 , m_1 , m_2 and m_3 in equations (2.22) and (2.23) to calculate $p_A(10 \text{ fs})$ in equation (2.21) as function of E_{12} and the distance between donor and acceptor.

Proton transfer system		O → O	O → N	N → O
k_1	[(kJ/mol) ⁻¹]	0.1216	0.1972	1.7568
k_2	[Å ⁻¹]	-8.669	-9.967	-20.221
k_3	[(kJ/mol) ⁻¹]	0.031	0.038	0.047
m_1	[1]	34.36	45.4942	27.94
m_2	[Å ⁻¹]	-0.436	-0.3054	-0.5000
m_3	[1]	-30.43	-41.10	-23.54

As E_{12} increases, $p_A(10 \text{ fs})$ becomes smaller, (and $p_A(10 \text{ fs})$ becomes strongly dependent on the starting wavefunction) up to a point where only a small tail of a wavepacket reaches the acceptor bound part of the potential. Here, the approach of estimating transfer rates using the time-dependent Schrödinger equation becomes questionable. Thus, only calculation of $p_A(10 \text{ fs})$ from the time-dependent Schrödinger equation for $p_A(10 \text{ fs}) > 0.1$ is considered. This is the case for small E_{12} and small $R(DA)$. For large E_{12} and large $R(DA)$, on the other hand, the energy barriers are so high that transition state theory becomes valid. What is still missing is a connection between

2.5. MERGING TIME-DEPENDENT SCHRÖDINGER EQUATION WITH TST 73

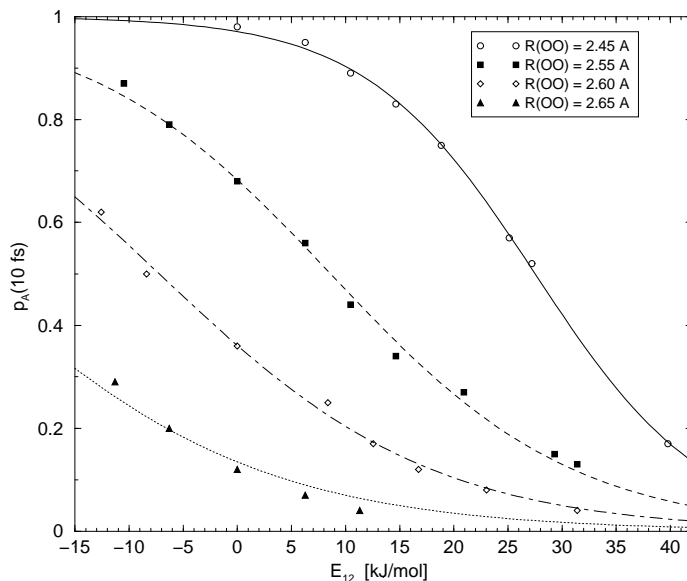


Figure 2.25: Probability to find the proton at the acceptor bound site after 10 fs, $p_A(10 \text{ fs})$, as a function of E_{12} for different $R(\text{OO})$. Symbols denote results from solutions of the time-dependent Schrödinger equation. The lines are fit functions (see text).

these two regimes. This is studied in the next section.

2.5 Merging Time-dependent Schrödinger Equation with TST

Up to now two approaches were presented which are applicable for two different regimes. The time-dependent Schrödinger equation seems to work well for small $R(\text{DA})$ and small E_{12} , while transition state theory should be used for large $R(\text{DA})$ and large E_{12} . Precisely, transition state theory is valid in a strict sense if $\exp(\beta E_b^{\ddagger}) \gg 1$, following the discussions of Hänggi [68]. As criterion of validity $\exp(\beta E_b^{\ddagger}) \geq 100$ was chosen. As criterion for the validity of the approach using the time-dependent Schrödinger equation $p_A(10 \text{ fs}) \geq 0.1$ was required, as already mentioned before. Figure 2.26 shows graphically the regions of validity for the proton transfer between two water molecules at $T = 298 \text{ K}$. In MD simulations of proton transfer in a water box (Section 4.1), $E_{12} < -40 \text{ kJ/mol}$ was never observed and usually $E_{12} > 0$. Thus, in this case

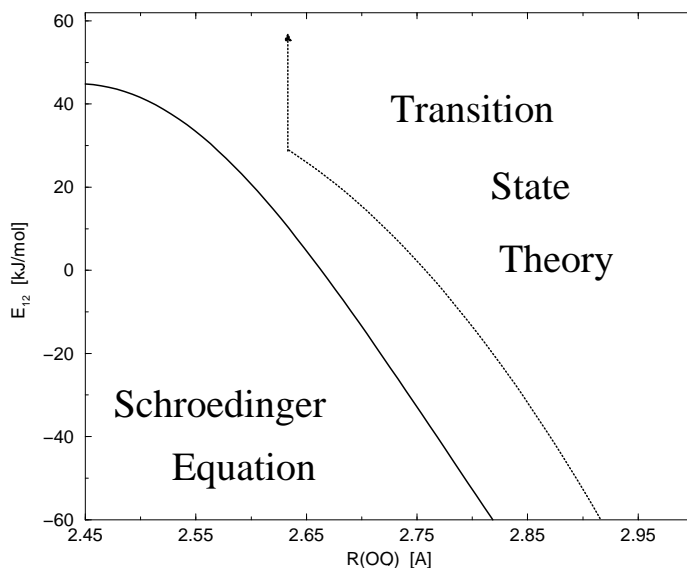


Figure 2.26: Regimes where the approach using the time-dependent Schrödinger equation is valid (small $R(\text{OO})$ and small E_{12}) and where TST is valid (large $R(\text{OO})$ and large E_{12}) for proton transfer between two water molecules at $T = 298 \text{ K}$.

transition state theory is always applicable for $R(\text{OO}) \geq 2.85 \text{ \AA}$.

Now arises the question if it is possible to bridge the gap between the solutions of the two approaches for smaller $R(\text{OO})$. Figure 2.27 shows $p_A(10\text{fs})$ at $R(\text{OO}) = 2.65 \text{ \AA}$. At small E_{12} the results from the approach using the time-dependent Schrödinger equation are drawn up to the validity limit shown as vertical line (dashed). For large E_{12} , the transition state solution including tunneling is drawn with its validity limit, again shown as vertical line (dot-dashed). The gradient of both curves at their limits differs from each other. But a linear interpolation (on a logarithmic scale) seems to be a reasonable approach to bridge the gap between the two approaches. Linearity on a logarithmic scale means

$$\log_{10} p_{\text{Gap}} = \log_{10} p_{SE}(E_{12}^L) + \frac{\log_{10} p_{TST}(E_{12}^R) - \log_{10} p_{SE}(E_{12}^L)}{E_{12}^R - E_{12}^L} (E_{12} - E_{12}^L) \quad , \quad (2.24)$$

where E_{12}^L and E_{12}^R are the validity limits of the two approaches.

An interpolation merging both extremes in a more clinging way may look better, but is as arbitrary as the approach here and can veil the fact that it is just a straightforward attempt to merge the two approaches.

2.5. MERGING TIME-DEPENDENT SCHRÖDINGER EQUATION WITH TST 75

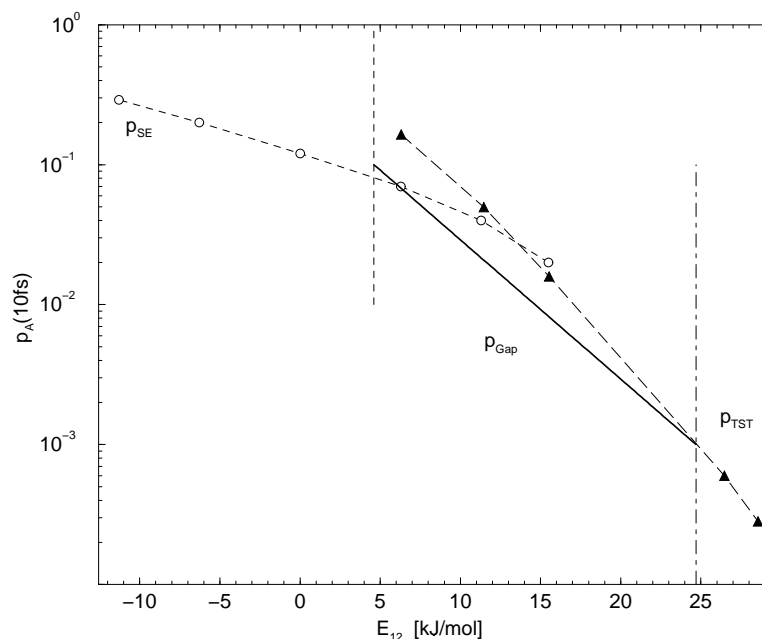


Figure 2.27: Proton transfer probability, $p_A(10\text{ fs})$, of the transfer between two water molecules at $R(\text{OO}) = 2.65\text{ \AA}$. At small E_{12} the result from the time-dependent Schrödinger equation, p_{SE} , (dashed line with open circles) up to the validity criterion ($p_A(10\text{ fs}) < 0.1$, dashed vertical line) is presented. At large E_{12} the solution of transition state theory including tunneling at WKB level, p_{TST} , (long dashed line with filled triangles) is shown with its validity criterion ($\exp(\beta E_b^{\rightarrow}) \geq 100$, dot-dashed line). Also, a linear interpolation on logarithmic scale, p_{Gap} , (solid line) is presented to bridge the gap between p_{SE} and p_{TST} .

2.5.1 Discussion

This study (Sections 2.4 and 2.5) was motivated by simulating an excess proton in a box of solvent molecules by molecular dynamics simulations. Proton transfer occurs at high frequency (experimental rate 0.63 per ps [82]) and there are frequent encounters of H_3O^+ and possibly proton accepting water molecules at distances as close as 2.5 Å. Proton transfer over larger distances is therefore quite unlikely unless a very favorable environment exists. Therefore it was necessary to supplement the previous study (see Chapter 2.3) by a special treatment for proton transfer over small energy barriers.

To solve the time dependent Schrödinger equation and to define the transition probability as the percentage of the wave package that swaps to the donor bound site after 10 fs is a compromise. On one hand, 10 fs appears as a lower limit for the time scale of proton transfer. On the other hand, classical MD simulations usually employ a 2 fs time step and proton transfer should be instantaneous according to this simulation approach. This approach was motivated by practical considerations and is clearly not a substitute for a theoretically-founded rate theory that has to be generally applicable. In that case one should certainly employ a dampening factor due to coupling with an external bath. In the present treatment, as can be seen from Fig. 2.24, the wavepacket continues to oscillate back and forth between donor and acceptor. It is, however, not clear at first sight how strong this dampening should be. Also, since barrier re-crossing is allowed in the Q-HOP MD simulation [50], the dynamics of the environment will eventually bring an end to such oscillations. The approach taken assumes a one-dimensional reaction coordinate. The time-dependency of the underlying nuclear energy surface during proton transfer is neglected.

The most ambiguous part of this study is how to connect the two seemingly opposite regimes of time-dependent Schrödinger equation and of transition state theory. Obviously, a connection must be found and a practical approach is followed. The path taken can be justified by the fact that the simulated proton transfer rates and proton diffusion coefficients in Q-HOP MD simulations (Section 4.1) of an excess proton in a box of solvent molecules are in good agreement with EVB MD simulations [31, 32, 33] and with experiment [82].

The presented approach in this section is somehow similar to the work of Azzouz and Borgis on proton transfer along asymmetrical H bonds in solution [83]. These

2.5. MERGING TIME-DEPENDENT SCHRÖDINGER EQUATION WITH TST 77

authors treated proton transfer as quantum hopping by a Landau-Zener process, and also explicitly accounted for quantum effects by a path integral formalism. Further investigation must be done to compare both approaches.

In conclusion, a parameterization scheme is presented to efficiently compute proton transfer probabilities for situations of very small or absent energy barriers.

Chapter 3

Q-HOP MD

In the last chapter a procedure was developed to calculate the hopping probabilities in a simple and fast manner for every protein configuration during the MD simulation. How proton hopping is combined with classical MD is described in detail in this chapter.

It is expected that there are local conformational reorientations between two transfer events, where for example a side-chain that has just accepted a proton has to rotate towards the next acceptor. This may be the rate limiting step along the proton transport path. To decrease the computing time necessary, the conformational sampling along the transport path is enhanced. The method used, locally enhanced sampling (LES), is described in detail in the following.

3.1 Locally Enhanced Sampling

3.1.1 Theoretical Foundation

Generally, equations of motion are derived by making the action S

$$S = \int \rho(\mathbf{X}_N, \mathbf{V}_N, t) L(\mathbf{X}_N, \mathbf{V}_N, t) d\mathbf{X}_N d\mathbf{V}_N dt \quad (3.1)$$

stationary, where the Lagrangian L is

$$L = \sum_{i=1}^N \frac{m_i}{2} v_i^2 - U(\mathbf{X}_N) \quad , \quad (3.2)$$

where $\rho(\mathbf{X}_N, \mathbf{V}_N, t)$ is the particle density and $\mathbf{X}_N = (\mathbf{x}_1, \dots, \mathbf{x}_N)$, $\mathbf{V}_N = (\mathbf{v}_1, \dots, \mathbf{v}_N)$ are the coordinates and velocities of the N particles.

By setting

$$\rho(\mathbf{X}_N, \mathbf{V}_N, t) = \prod_{i=1}^N \delta(\mathbf{x}_i - \mathbf{x}_{i,0}(t)) \delta(\mathbf{v}_i - \mathbf{v}_{i,0}(t)) \quad (3.3)$$

Newton's equation of motion for point particles is derived

$$m_i \dot{\mathbf{v}}_{i,0}(t) = - \frac{\partial U(\mathbf{X}_{N,0})}{\partial \mathbf{x}_{i,0}} \quad (3.4)$$

In Locally Enhanced Sampling (LES), those parts of the system of special interest, e.g. a relevant side chain of a protein or a surface loop etc. are replicated several times [44]. Thus the protein is split into J subsystems

$$\rho(\mathbf{X}_N, \mathbf{V}_N, t) = \prod_{j=1}^J \rho^j(\mathbf{X}^j, \mathbf{V}^j, t) \quad (3.5)$$

each represented by K_j copies

$$\rho^j(\mathbf{X}^j, \mathbf{V}^j, t) = \sum_{k_j=1}^{K_j} w_{jk_j} \delta(\mathbf{X}^j - \mathbf{X}_0^{k_j}(t)) \delta(\mathbf{V}^j - \mathbf{V}_0^{k_j}(t)) \quad (3.6)$$

where

$$\sum_{k_j=1}^{K_j} w_{jk_j} = 1 \quad \text{and e.g.} \quad \mathbf{X}_0^{k_j} = (\mathbf{x}_{s_{j-1}+1,0}^{k_j}, \dots, \mathbf{x}_{s_j,0}^{k_j}) \quad (3.7)$$

and w_{jk_j} is the weighting factor of the k_j 'th copy of subsystem j . The action can then be written as

$$S = \int (T_{eff} - U_{eff}) d\mathbf{X}_0^{k_1} \dots d\mathbf{X}_0^{k_J} d\mathbf{V}_0^{k_1} \dots d\mathbf{V}_0^{k_J} dt \quad (3.8)$$

with

$$T_{eff} = \frac{1}{2} \sum_{j=1}^J \sum_{k_j}^{K_j} \sum_{i=s_{j-1}+1}^{s_j} w_{jk_j} m_i (v_{i,0}^{k_j})^2 \quad (3.9)$$

and

$$U_{eff} = \sum_{k_1}^{K_1} \dots \sum_{k_J}^{K_J} \prod_{j=1}^J w_{jk_j} U(\mathbf{X}_0^{k_1}, \dots, \mathbf{X}_0^{k_J}) \quad (3.10)$$

In case of proton transport which is the special interest of this work, the amino acid residues and water molecules, which are involved in the current transfer step, are replicated Q times. For technical reasons (i.e. to facilitate the construction of the new

topology files in case of proton transfer) each copy k_j is present in a protonated state $p_{k_j} = 1$ and an un-protonated state $p_{k_j} = 0$. A switching function

$$\delta_{p_{k_j}} \left| \begin{array}{ll} k_j & \text{protonated} \\ k_j & \text{un-protonated} \end{array} \right. \Rightarrow \begin{array}{ll} \delta_{p_{k_j}=1} = 1 & \delta_{p_{k_j}=0} = 0 \\ \delta_{p_{k_j}=1} = 0 & \delta_{p_{k_j}=0} = 1 \end{array} \quad (3.11)$$

guarantees that only the currently active protonation state is considered in the calculation of the forces. Equations (3.6) and (3.7) thus have to be expanded, ensuring a fixed number of protonated groups N_P

$$\rho^j(\mathbf{X}^j, \mathbf{V}^j, t) = \sum_{k_j=1}^{K_j} \sum_{p_{k_j} \in \{0,1\}} w_{jk_j p_{k_j}} \delta(\mathbf{X}^j - \mathbf{X}_0^{k_j p_{k_j}}(t)) \delta(\mathbf{V}^j - \mathbf{V}_0^{k_j p_{k_j}}(t)) \quad , \quad (3.12)$$

where

$$\sum_{k_j=1}^{K_j} \sum_{p_{k_j} \in \{0,1\}} w_{jk_j p_{k_j}} = 1, \quad \sum_{j=1}^J \sum_{k_j} w_{jk_j p_{k_j}=1} = N_P \quad (3.13)$$

$$\text{and e.g. } \mathbf{X}_0^{k_j p_{k_j}} = \left(\mathbf{x}_{s_{j-1}, p_{k_j}+1, 0}^{k_j}, \dots, \mathbf{x}_{s_j, p_{k_j}, 0}^{k_j} \right) \quad .$$

The left sum guarantees that all copies of an amino acid residue or a water molecule (weighted by $w_{jk_j p_{k_j}}$) in their momentary protonation state p_{k_j} sum up to one chemical group smeared out in different configurations. The right sum over all subsystems ensures a fixed number of protonated groups N_P . The total charge of the system is therefore constrained to a fixed value. No protons are disappearing or emerging during the simulation. The implementation section will show how this constraint is automatically guaranteed during the simulation. The remaining parts of the protein are full-sized and just present as a single copy and comprise subsystem 1, i.e.

$$K_1 = 1, \quad K_{i \neq 1} = Q, \quad w_{i \neq 1 k_i \neq 1 p_{k_i \neq 1}} = \delta_{p_{k_i \neq 1}} 1/Q \quad . \quad (3.14)$$

Then

$$T_{eff} = \frac{1}{2} \sum_{i=1}^{s_1} m_i v_{i,0}^2 + \frac{1}{2Q} \sum_{j=2}^J \sum_{k_j} \sum_{p_{k_j} \in \{0,1\}} \delta_{p_{k_j}} \sum_{i=s_{j-1}, p_{k_j}+1}^{s_j, p_{k_j}} m_i \left(v_{i,0}^{k_j} \right)^2 \quad (3.15)$$

$$U_{eff} = \sum_{k_2}^Q \dots \sum_{k_J}^Q \sum_{p_{k_2} \in \{0,1\}} \dots \sum_{p_{k_J} \in \{0,1\}} \delta_{p_{k_2}} \dots \delta_{p_{k_J}} \left(\frac{1}{Q} \right)^{J-1} U \left(\mathbf{X}_0^1, \mathbf{X}_0^{k_2 p_{k_2}}, \dots, \mathbf{X}_0^{k_J p_{k_J}} \right)$$

$$\begin{aligned}
&= U_{1-1}(\mathbf{x}_{1,0}, \dots, \mathbf{x}_{s_1,0}) + \\
&\quad \frac{1}{Q} \sum_{j=2}^J \sum_{k_j=1}^Q \sum_{p_{k_j} \in \{0,1\}} \delta_{p_{k_j}} U_{1-j} \left(\mathbf{x}_{1,0}, \dots, \mathbf{x}_{s_1,0}; \mathbf{X}_0^{k_j p_{k_j}} \right) + \\
&\quad \frac{1}{Q} \sum_{j=2}^J \sum_{k_j=1}^Q \sum_{p_{k_j} \in \{0,1\}} \delta_{p_{k_j}} U_{j-j} \left(\mathbf{X}_0^{k_j p_{k_j}} \right) + \\
&\quad \frac{1}{2Q^2} \sum_{i=2}^J \sum_{\substack{j=2 \\ j \neq i}}^J \sum_{k_i=1}^Q \sum_{k_j=1}^Q \sum_{p_{k_i} \in \{0,1\}} \sum_{p_{k_j} \in \{0,1\}} \delta_{p_{k_i}} \delta_{p_{k_j}} U_{i-j} \left(\mathbf{X}_0^{k_i p_{k_i}}, \mathbf{X}_0^{k_j p_{k_j}} \right) ,
\end{aligned} \tag{3.16}$$

where U_{1-1} , U_{1-j} and U_{i-j} are the potential energies between single-copy atoms, between single-copy atoms and multi-copy atoms of subsystem j , and between multi-copy atoms of subsystem i and subsystem j . Applying the Euler-Lagrangian equations

$$\frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}_{i,0}} - \frac{\partial L}{\partial \mathbf{x}_{i,0}} = 0 \tag{3.17}$$

gives the equations of motion for non-replicated atoms

$$\begin{aligned}
m_i \dot{\mathbf{v}}_{i,0} &= - \frac{\partial U_{1-1}(\mathbf{x}_{1,0}, \dots, \mathbf{x}_{s_1,0})}{\partial \mathbf{x}_{i,0}} \\
&\quad - \frac{1}{Q} \sum_{j=2}^J \sum_{k_j=1}^Q \sum_{p_{k_j} \in \{0,1\}} \delta_{p_{k_j}} \frac{\partial U_{1-j} \left(\mathbf{x}_{1,0}, \dots, \mathbf{x}_{s_1,0}; \mathbf{X}_0^{k_j p_{k_j}} \right)}{\partial \mathbf{x}_{i,0}}
\end{aligned} \tag{3.18}$$

and for atoms represented by multiple copies

$$\begin{aligned}
m_i \dot{\mathbf{v}}_{i,0}^{k_j} &= - \frac{\partial U_{1-j} \left(\mathbf{x}_{1,0}, \dots, \mathbf{x}_{s_1,0}; \mathbf{X}_0^{k_j p_{k_j}} \right)}{\partial \mathbf{x}_{i,0}^{k_j}} \\
&\quad - \frac{\partial U_{j-j} \left(\mathbf{X}_0^{k_j p_{k_j}} \right)}{\partial \mathbf{x}_{i,0}^{k_j}} \\
&\quad - \frac{1}{Q} \sum_{l=2, l \neq j}^J \sum_{k_l=1}^Q \sum_{p_{k_l} \in \{0,1\}} \delta_{p_{k_l}} \frac{\partial U_{j-l} \left(\mathbf{X}_0^{k_j p_{k_j}}, \mathbf{X}_0^{k_l p_{k_l}} \right)}{\partial \mathbf{x}_{i,0}^{k_j}} .
\end{aligned} \tag{3.19}$$

The single-copy atoms move in the field of the other single-copy atoms and the mean field of the multi-copy atoms. Conversely, the multi-copy atoms move in the field of the single-copy atoms, in the field of the atoms of the same copy of the same

segment and in the mean field of the multi-copy atoms of different segments (at least for non-bonded interactions). The multi-copy atoms of the same amino acid residue move independently of each other.

3.1.2 Implementation

All simulations of this work were carried out with a modified version of the ARGOS simulation package [13] using the AMBER95 force field [4].

The non-bonded interactions are pair interactions that are stored in pair-lists. While interactions between single-copy atoms make full contributions to the forces, forces containing a multi-copy atom are scaled by a factor $1/Q$. This was taken into account by arrays $fofai()$ and $fofaj()$ which store these factors for both partners, atom i and j , for each interaction. An overview of possible cases is found in Table 3.1.

Table 3.1: Force factors to scale non-bonded interactions between atoms i and j in multi-copy simulations.

Atom i	Atom j		$fofai$	$fofaj$
non-copied	non-copied		1	1
non-copied	copied		$1/Q$	1
copied	non-copied		1	$1/Q$
copied	copied	same group	1	1
		different groups	$1/Q$	$1/Q$

The stretch, bending and torsional interactions are also stored in pair-, triplet- and quartet lists. Consequently, the contributions of each interaction term are also in arrays of force factors e.g. $fofashi()$, $fofashj()$ and $fofashk()$ for the bending interactions as well. Here, interactions between copies of different but consecutive groups are considered in full strength. This means that copy #1 of group i is only bound to copy #1 of group $i + 1$, copy #2 of group i is only bound to copy #2 of group $i + 1$, etc. . This leads to a higher flexibility of the poly-peptide chain, and therefore to better sampling. As an example the pre-factors for the different cases for bending interactions are presented in Table 3.2.

The same procedure is used for the energy terms according to equation (3.16).

Table 3.2: Force factors to scale bending interactions between atoms i , j and k in multi-copy simulations. ('n.c.' means non-copied.)

i	j	k	f_{ofashi}	f_{ofashj}	f_{ofashk}
n.c.	n.c.	n.c.	1	1	1
n.c.	n.c.	copied	1/Q	1/Q	1
n.c.	copied	copied	1/Q	1	1
copied	n.c.	n.c.	1	1/Q	1/Q
copied	copied	n.c.	1	1	1/Q
copied	copied	copied	1	1	1

3.2 Proton Hopping

3.2.1 Principal Procedure

There are two principal approaches implemented to treat the possible proton donors and acceptors. In one picture, the proton transport through proteins is divided into consecutive transport steps between adjacent, multi-copy donor-acceptor pairs for proton transfer (Figure 3.1). At the beginning of a transport step one proton donor is marked and multiple copies are generated. A sphere is drawn around this donor, and multiple copies are generated for all possible acceptors in this sphere (and possible additional donors as well). Their protonation state is treated as flexible during this transport step. During the transport step the protons hop between the multi-copy donors and acceptors. If more than a pre-defined percentage of the copies of an acceptor become protonated, this group is marked as the new donor for the next transport step. All non-protonated copies of this acceptor are then forced to take up a proton from one of the remaining donor copies in a given time. Then the procedure starts from the beginning. The new marked donor now becomes the center of a sphere containing the multi-copy groups for the next transport step etc.. This approach keeps the number of copied groups which can change their protonation states small, minimizing the computational overhead of the multi-copy approach. On the other hand, it does not allow for multiple proton transfer at separated positions in the protein. But it is possible that a proton transfer at one position could induce a transfer at a different

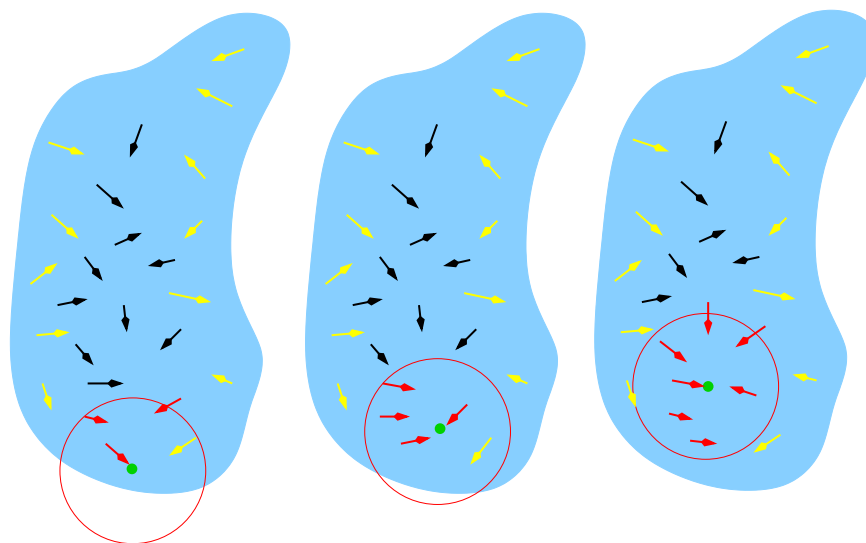


Figure 3.1: Proton transport through proteins is divided into consecutive transport steps between adjacent donor-acceptor pairs in a sphere of predefined radius. After transfer from donor to acceptor, the latter becomes the center of a new sphere for the next proton transport step. Therefore, the sphere of possible donor-acceptor pairs is slowly moving through the protein. Drawn in black are the segments which can in principle be flexibly protonated, the red ones can be flexibly protonated in the momentary transport step and the yellow ones can have only one protonation states.

protein site due to changing electrostatic interactions.

Therefore a second approach is implemented (Figure 3.2), where the segments present as multiple copies are chosen at the beginning and remain as multiple copies over the whole simulation time. The remaining steps are identical to the first implementation except that the segments which are present as multiple copies do not change. This allows parallel proton transfer at different protein sites. The disadvantage is that a larger number of multi-copied segments is present at the same time leading to a higher computational effort. Additionally, if not all segments are going to be multi-copied which could be flexibly protonated, the possible paths of proton transport through the protein should be known.

Figure 3.3 shows the procedure during a transport step in more detail. Now the quantum chemical studies on proton transfer reactions in small model systems (Sections

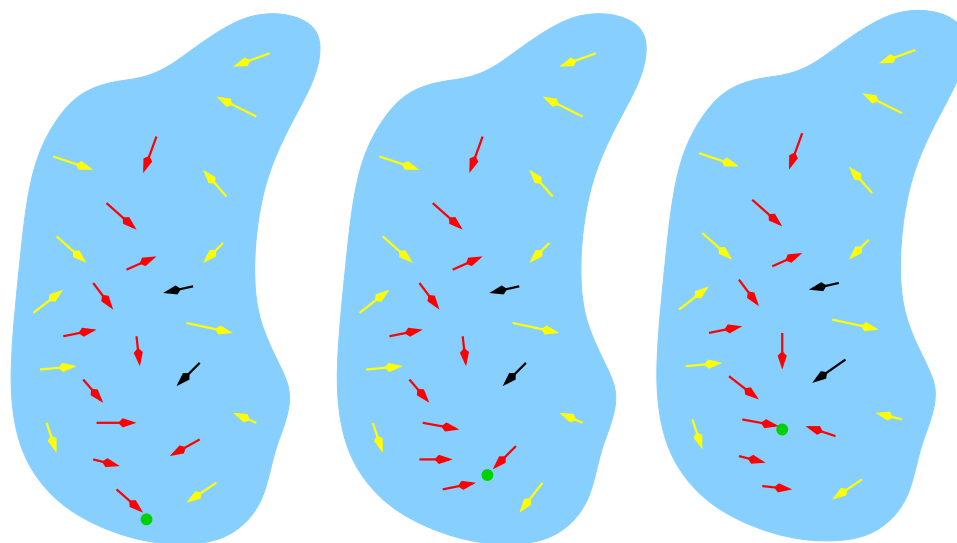


Figure 3.2: Proton transfer can happen between all segments with flexible protonation states which are selected at the beginning (red). Drawn in black are the segments which can in principle be flexibly protonated, but are not chosen at the beginning and the yellow ones can have only one protonation state.

2.3, 2.4) are used, where fit formulas were derived to calculate transfer rates under the electrostatic influence of environmental groups in an easy and fast manner. Let us consider a standard case. A classical MD simulation is performed in a starting protonation state. After a pre-defined number of MD-steps, corresponding to a time Δt , the distances $R(\text{DA})$ between all donor-acceptor pairs in the sphere are determined. Then, for each such pair, the energy difference E_{12} is calculated between the state where the proton is bound to the acceptor and the donor-bound state in vacuum. This is possible using a simple quadratic formula (see Equation (2.9))

$$E_{12}^0 = \alpha + \beta R(\text{DA}) + \gamma R(\text{DA})^2 \quad , \quad (3.20)$$

where α , β and γ are parameters tabulated for all pairs of chemical groups (see Table 2.5) present in the simulation, and $R(\text{DA})$ is the only input value measured during the simulation.

Now, the electrostatic influence of the surrounding groups on the relative energies of donor and acceptor bound state needs to be determined. A sphere of pre-defined cut-off radius is drawn around the midpoint between donating and accepting atom of

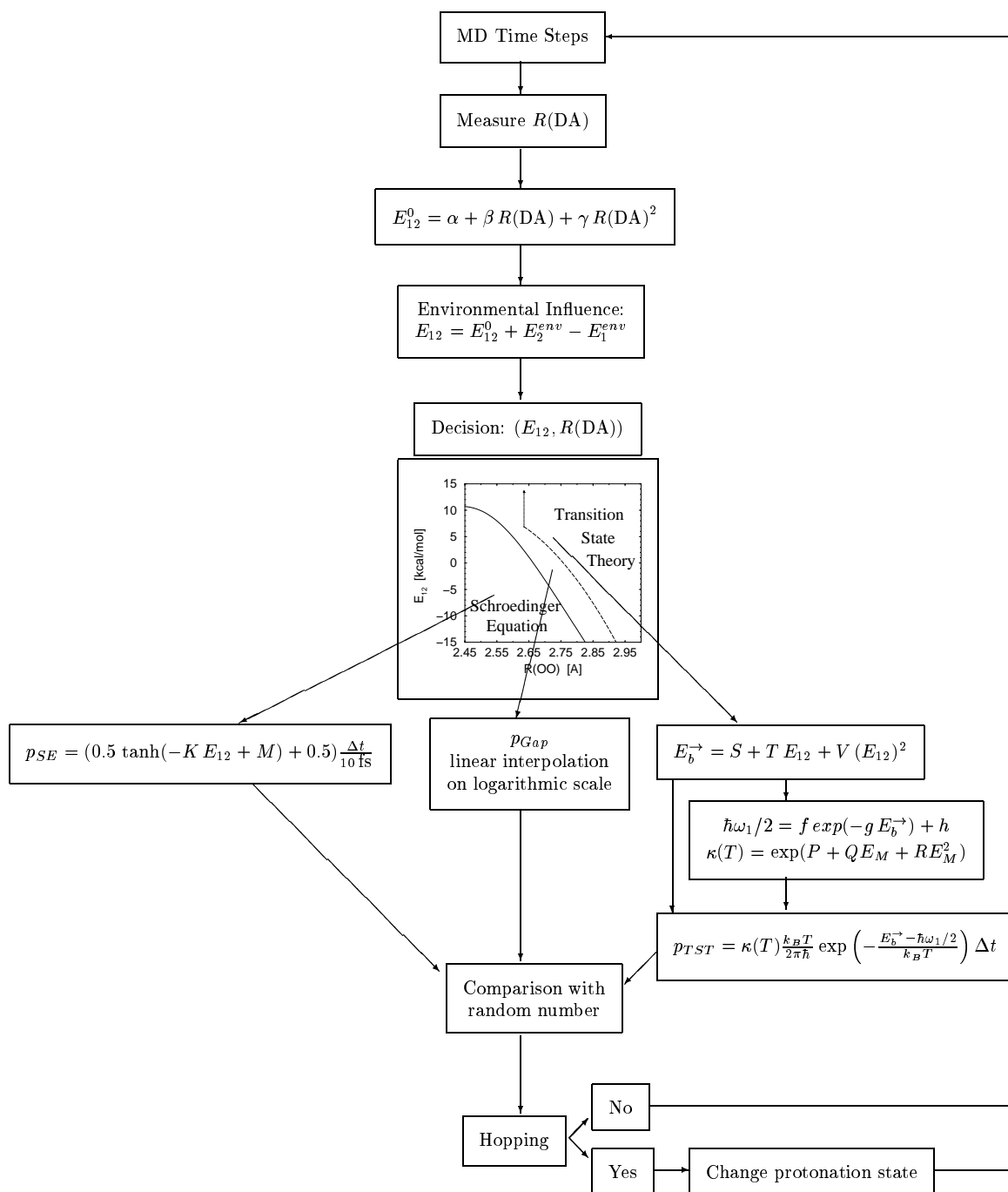


Figure 3.3: Flowchart to compute the transfer probabilities and to decide if proton hopping does occur or not during a single transport step. (For details see text.)

this donor-acceptor pair. The electrostatic interaction between donor-acceptor pair and the environmental groups in the sphere is calculated as

$$E_1^{env} = \sum_{i \in \{D-H, A\}} \sum_{j \in \{env\}} \frac{q_i q_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \quad , \quad (3.21)$$

where $q_{i/j}$ and $\mathbf{r}_{i/j}$ are charges and coordinates of the atoms i and j . E_1^{env} corresponds to the donor-bound state (protonated donor $D - H$, un-protonated acceptor A). Important to note is that the interactions between donor and acceptor are not included here, since they are already implicitly considered in equation (3.20). To calculate the environmental influence on the putative acceptor bound state a dummy situation is constructed, where the donor is de-protonated and the acceptor is protonated, respectively. The point charges are modified according to the force-field charges of de-protonated donor D and protonated acceptor group $A - H$, and E_2^{env} is calculated as

$$E_2^{env} = \sum_{i \in \{D, H-A\}} \sum_{j \in \{env\}} \frac{q_i q_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \quad . \quad (3.22)$$

So, the energy difference between donor and acceptor bound states with environmental influence E_{12} is

$$E_{12} = E_{12}^0 + E_2^{env} - E_1^{env} \quad . \quad (3.23)$$

Following previous work, three different cases have to be distinguished depending on $R(\text{DA})$ and E_{12} (see Section 2.5).

(a) For small $R(\text{DA})$ and small E_{12} , only a small or even no energy barrier E_B^{\rightarrow} between donor and acceptor is present and the probability for transfer is high. This probability can be estimated by evaluating

$$p_{SE} = (0.5 \tanh(-K E_{12} + M) + 0.5) \frac{\Delta t}{10 \text{ fs}} \quad , \quad (3.24)$$

where $K = K(R(\text{DA}))$ and $M = M(R(\text{DA}))$ are easily calculated using tabulated parameters and $R(\text{DA})$ as input value. This formula was derived by studying the time evolution of wavepackets on the one-dimensional energy surface of the proton in the donor-acceptor system using the time-dependent Schrödinger equation (see Section (2.4)).

(b) For large E_{12} and large $R(\text{DA})$, the energy barrier is so high that simple transition state theory is valid. In this case, the energy barrier E_b^{\rightarrow} is estimated by (see

Equation (2.7))

$$E_b^\rightarrow = S + T E_{12} + V (E_{12})^2 \quad . \quad (3.25)$$

$S = S(R(\text{DA}))$, $T = T(R(\text{DA}))$ and $V = V(R(\text{DA}))$ are easily calculated using tabulated parameters with $R(\text{DA})$ as input value (see Equation (2.8) and Table 2.4). The calculated energy barrier is then used to estimate the zero-point energy (see Equation (2.10))

$$\hbar\omega_1/2 = f \exp(-g E_b^\rightarrow) + h \quad (3.26)$$

(f , g and h are tabulated values (see Table 2.7)) and the enhancement of the transfer rate due to tunneling (see Equation (2.15))

$$\kappa(T) = \exp(P + Q E_M + R E_M^2) \quad (3.27)$$

($P = P(T)$, $Q = Q(T)$ and $R = R(T)$ are calculated using tabulated values for Equation (2.16) (see Table 2.8) and E_M is the minimum of forward and backward transfer barrier, $E_M = \min(E_b^\rightarrow, E_b^\leftarrow)$) to account for the quantum nature of the light proton. Introducing E_b^\rightarrow , $\hbar\omega_1/2$ and $\kappa(T)$ into simple transition state theory gives the transfer probability in this regime

$$p_{TST} = \kappa(T) \frac{k_B T}{2\pi\hbar} \exp\left(-\frac{E_b^\rightarrow - \hbar\omega_1/2}{k_B T}\right) \Delta t \quad . \quad (3.28)$$

(c) For the intermediate regime, a linear interpolation on logarithmic scale is used which bridges the gap between the limiting cases (a) and (b), (see Equation (2.24))

$$\log_{10} p_{Gap} = \log_{10} p_{SE}(E_{12}^L) + \frac{\log_{10} p_{TST}(E_{12}^R) - \log_{10} p_{SE}(E_{12}^L)}{E_{12}^R - E_{12}^L} (E_{12} - E_{12}^L) \quad , \quad (3.29)$$

where E_{12}^L and E_{12}^R are the estimated validity limits of the two approaches (a) and (b).

Having calculated the transfer probability in one of the three ways (a), (b) or (c), a random number is computed and compared with the probability of proton transfer. If the random number is larger than the probability, no hopping occurs. This comparison is done for every donor-acceptor pair. If no transfer process takes place at all, the simulation continues with the next MD-time steps in the same protonation state. If the random number is smaller than the probability, hopping takes place between donor and acceptor. The proton is instantaneously shifted from donating atom to accepting atom, and the simulation continues with MD-time steps in the altered protonation state.

3.2.2 Implementation

Although the process of proton transfer is conceptually simple, the implementation of the code is slightly tedious. Before the start of the simulation, a topology file is created which lists consecutive amino acids, water molecules and other heteroatoms with all atoms of each segment. Additionally the groups that may function as donor, $(D-H)$, or acceptor, A , in the first transport step are present as multiple copies, $(D-H)_j$ ($j = 1, \dots$, number of copies) and A_i ($i = 1, \dots$, number of copies), every copy once in the protonated and once in the un-protonated state. This means that besides the groups in their actual protonation state $((D-H)_j, A_i)$, ghost groups $(D_j, (A-H)_i)$ are listed in the topology file, which represents the groups becoming active only when they are involved in a hopping event. The topology file also contains all bonds, angles and dihedrals formed by real and ghost atoms. Since only the groups in the actual protonation state contribute to the forces computed, lists of active interaction pairs, triplets and quartets are created from the topology file which are used in the calculation of the forces and energies.

The same starting coordinates are used for the same atom of different copies. To ensure a different time evolution of the different copies of this segment, the starting velocities for corresponding atoms of different copies are chosen which are slightly different from each other.

In the simulation, a list of active donor-acceptor pairs $(D-H)_j - A_i$ is created. To minimize the additional computational effort, only pairs are selected where donating and accepting atom are closer than 3.5 Å. This is justified considering the one-dimensional energy surface of proton transfer for different systems of donor and acceptor, which shows very high energy barriers for large separations between donor and acceptor. Also, transfers were only observed for oxygen-oxygen separations much smaller than 3 Å in quantum chemical simulations for an excess proton in a water box [34, 32, 33]. In addition to the separation criterion, the angle of the hydrogen bond, the angle between donating atom, transferring proton and acceptor, has to be larger than 120°. Then for each pair the transfer probability is calculated by equations (3.24), (3.28) or (3.29) as described before (see Section 3.2.1), where the necessary fitting parameters are stored in input files which are read only at the beginning of the simulation. Then the transfer probabilities are sorted in decreasing order, so that

the pair with the highest transfer probability gets the first try to transfer its proton. A random number is computed and compared with the transfer probability p . If the random number is larger than p , no hopping between this pair happens and the next pair is checked. If the random number is smaller than p , hopping takes place between donor and acceptor of this pair. The donor $(D-H)_j$ is de-protonated and the acceptor A_i is protonated. This is done by activating the ghost atoms of the un-protonated state D_j and the protonated state $(A-H)_i$ and by deactivating the atoms of $(D-H)_j$ and A_i . The coordinates and velocities of the atoms of D_j are the same as those of the atoms of $(D-H)_j$, except for the absent proton. The coordinates and velocities of the atoms of $(A-H)_i$ are the same as those of the atoms of A_i , except for the newly appeared proton, which was not present in the protonation state A_i . The proton is positioned on the interconnecting line between former proton position and accepting atom of A_i at equilibrium distance according to the underlying force-field. Its initial velocity is set to zero. This is a simple approximation. To guarantee that the energy is conserved when a proton hops, a more elaborate approach will be presented in Section 3.4. The transfer of coordinates and velocities is insured by an array which maps the atom numbers between the different protonation states for each segment which can be in different protonation states. Thereafter the next donor-acceptor pair is checked, not containing $(D-H)_j$ or A_i . In order to decrease the computational effort, not the whole procedure is performed at every time step. The frequency at which the procedure is applied can be selected before the simulation.

When all donor-acceptor pairs have been checked and at least one transfer has happened, new lists of active interaction pairs, triplets and quartets are created. Also, the number of protonated copies of each segment is determined. The actual transport step is terminated, if more than a pre-defined percentage of a segment, A^s , which was not protonated at the beginning of the transport step, becomes protonated. If this does not happen for longer than a given time, the percentage may be decreased. The following only applies to the concept of a moving sphere. After termination of a proton transport step, the segment A^s becomes the new center of the sphere that contains the segments represented by multiple copies for the next transport step. Thus, the sphere enabling proton hopping moves slowly through the system. This means that multi-copy segments in the previous transport step may not necessarily be included in the sphere of the next transport step. This has two consequences.

First, segments have to be either fully protonated or de-protonated at the end of the transport step. While it is possible that the sphere of the terminated transport step includes more than one donating segment, the same number of donors must be present at the beginning and at the end of the step. Therefore, the segments with the largest number of protonated copies are selected to be fully protonated at the end of the transport step. Lets call them C, D, \dots whereas the other segments which have to be fully de-protonated at the end are named F, G, \dots . This is guaranteed by a number of additional transfer steps where the probability for a transfer from C, D, \dots to F, G, \dots decreases stepwise to zero, whereas the probability of a transfer from F, G, \dots to C, D, \dots increases to one.

The second consequence of the concept of a moving sphere is that the segments not included in the sphere of the next transport step have to shrink down to just one representative. This representative is selected randomly from the different copies of this segment, taking into account their Boltzmann weights, where the copy with the lowest potential energy has the highest probability. The selected copy is blown up to a complete amino acid residue in a pre-defined number of MD time-steps and the other copies are shrunk down to a dummy copy. This is done similarly to free energy perturbation techniques by introducing the Hamiltonian

$$H = \lambda H_{\text{selected copy}} + (1 - \lambda) H_{\text{not selected}} \quad , \quad (3.30)$$

where λ changes from zero to one stepwise in this time period. If two or more replicated segments A, B, \dots are connected and one, say B , has to shrink down to one representative, whereas A remains replicated, this shrinking process would lead to open ended copies of A . In this case, the coordinates of the backbone atoms of the other copies of B are approximated to the one selected representative of B . Changing the coordinates proceeds parallel to the shrinking process. A new topology file is then created and the next MD steps follow.

This procedure avoids combinatorial explosion of combining different protonation states (see equation (3.13)).

3.3 Polarizable Environment

In Section 2.2, it was shown that the energy surface of a proton transfer reaction can be determined by treating only donor and acceptor with quantum mechanical methods in vacuum, and by adding the effect of the environmental charges as the electrostatic interaction energy between proton transferring groups and the surrounding point charges. The only correction necessary is to enhance the charges of the point particles to account for polarization effects. In Sections 2.2.7 - 2.5 we showed that it is not necessary to explicitly determine the complete energy curve to calculate the transfer probabilities. Given the distance between donor and acceptor, $R(\text{DA})$, knowing the distance between the energy minima, E_{12} , is sufficient to calculate the transfer probability. While an enhancement of the point charges due to induced polarization still seems to be necessary, the point charges were not enhanced when calculating E_{12} in Section 3.2. To justify this simplification, the results for larger systems including many surrounding molecules are presented in the following .

In Section 2.2 we considered only a small isolated system as proton transferring system, composed of an amino acid residue and a water molecule, and two water molecules were included as surrounding chemical groups. A positive charge on the proton accepting water molecule will therefore have the maximum effect on the polarization of the surrounding water molecules. Additionally, the two secondary water molecules were in optimal hydrogen bond orientation, which gives maximally induced polarization. In realistic proton transfer configurations, the surrounding groups are usually positioned in non-optimal hydrogen bond arrangement, and additional surrounding groups in second or higher solvation shells around donor and acceptor interact with the groups in the first solvation shell. This will additionally influence the polarization. This causes doubts whether the magnitude of the polarization effect in Section 2.2 introduced by enhanced charges will be overemphasized when applied to real solutions or proteins.

To clarify this discussion, the proton transfer between two water molecules in a box of 216 water molecules is studied using results of a simulation, described in more detail in Section 4.1. Reactive configurations are extracted from the simulation trajectory, meaning the coordinates before and after an instantaneous proton hop. Since it is too time consuming to treat the whole system quantum mechanically, a QM/MM approach is used (see Figure 3.4). All water molecules in a sphere of radius 3.5 Å around the

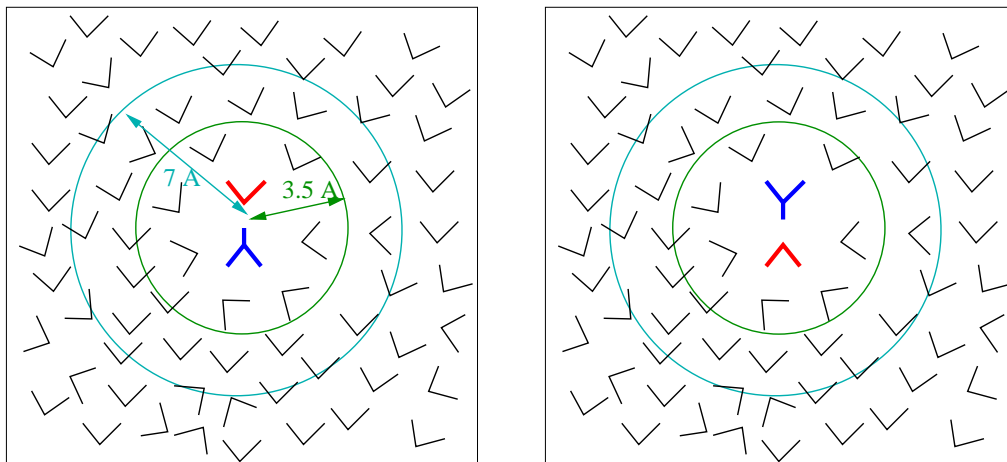


Figure 3.4: QM/MM approach to compute E_{12} for the proton transfer of an additional proton in a water box for comparison of quantum mechanical calculations with classical electrostatic or polarizable force field approach.

midpoint between donating and accepting atom are treated quantum mechanically (5-12 water molecules). Those in a spherical ring between 3.5 Å and 7 Å (44-50 water molecules) are considered as classical SPC point charges with $q_O = -0.82e$ and $q_H = 0.41e$ as in the classical MD simulations. By computing the energies in donating state, $E_1^{qm,tot}$, and in accepting state, $E_2^{qm,tot}$, at MP2/6-31G** level, $E_{12}^{qm,tot} = E_2^{qm,tot} - E_1^{qm,tot}$ is compared with the classical force field values for E_{12} for the same division of the system (sphere with radius 7 Å).

In Q-HOP MD, first the energy difference of the donating system in vacuum, E_{12}^0 is determined. For the proton transfer between two water molecules, a symmetric situation is assumed, leading to $E_{12}^0 = 0$. Since the proton hops instantaneously to the acceptor, it will be coordinated to a position leading to a hydronium ion with internal stress. Thus the internal energy of accepting state is expected to be usually higher than in the donating state, leading to $E_{12}^{0,qm} \neq 0$. To compare quantum mechanically calculated E_{12}^{qm} with classical electrostatic calculations of E_{12} , donor-acceptor system is optimized in donating and accepting state, and the energies for these configurations are calculated. In addition, the energies of donor-acceptor system in vacuum in both states are calculated and subtracted from $E_{12}^{qm,tot}$, leading to an energy difference named E_{12}^{qm} , which can be compared with the classical E_{12} used in the simulation.

Table 3.3 shows the results of this comparison for 10 transfer events. The mean

Table 3.3: Comparison of the quantum mechanically calculated energy difference between donating and accepting state, E_{12}^{qm} , with those calculated electrostatically (E_{12}) and by Tholes approach of a polarizable force field (E_{12}^{pol}).

transfer event	E_{12}^{qm} [kJ/mol]	E_{12} [kJ/mol]	$ E_{12} - E_{12}^{qm} $ [kJ/mol]	E_{12}^{pol} [kJ/mol]	$ E_{12}^{pol} - E_{12}^{qm} $ [kJ/mol]
1	20.1	19.7	0.4	18.0	2.1
2	8.4	25.5	17.1	2.9	5.4
3	41.0	47.3	6.3	47.7	6.7
4	17.6	29.3	11.7	35.2	17.6
5	-5.4	-0.8	4.6	-19.7	14.3
6	23.5	16.7	6.8	29.3	5.8
7	15.1	9.6	5.4	16.3	1.3
8	2.1	10.0	7.9	12.6	10.5
9	6.7	20.5	13.8	26.8	20.1
10	17.2	39.4	22.2	31.0	13.8

deviation is 9.6 kJ/mol. The mean donor-acceptor distance in these 10 transfer events is 2.5 Å. Thus a mean deviation of 9.6 kJ/mol corresponds to a mean deviation in transfer probability of around 10 – 20%. In non reactive configurations (not shown here, see Figure 4.4), E_{12} is typically in the order of 40–150 kJ/mol for the closest donor-acceptor pair. The electrostatic calculation of E_{12} can therefore be considered accurate enough to select non reactive configurations from reactive ones. Additionally, it is always predicted correctly, whether the reaction is energetically favorable or unfavorable.

Explicit atomic polarizabilities using the dipole model, which was developed by Thole [85, 86], where isotropic polarizabilities are assigned to each atom, are included to try to improve the prediction of E_{12} . The molecule is then considered as an arrangement of N point particles with point charges q_i and induced dipole moments $\boldsymbol{\mu}_i$. It follows the atomic dipole moments

$$\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_i \quad , \quad (3.31)$$

where α_i is the atomic polarizability and \mathbf{E}_i is the electric field at point particle i , which is the sum of the electrostatic fields from the other point charges and from the

dipole moments,

$$\mathbf{E}_i = \mathbf{E}_i^q + \mathbf{E}_i^{ind} \quad (3.32)$$

with

$$\mathbf{E}_i^q = \sum_{j \neq i} \frac{q_j}{4\pi\epsilon_0 r_{ij}^3} \mathbf{r}_{ij} \quad (3.33)$$

$$\mathbf{E}_i^{ind} = \sum_{j \neq i} \hat{\mathbf{T}}_{ij} \boldsymbol{\mu}_j \quad (3.34)$$

$\hat{\mathbf{T}}_{ij}$ is the dipole field tensor

$$\hat{\mathbf{T}}_{ij} = \frac{\hat{\mathbf{1}}r_{ij}^2 - 3\mathbf{r}_{ij}\hat{\mathbf{r}}_{ij}}{4\pi\epsilon_0 r_{ij}^5} \quad (3.35)$$

where $\hat{\cdot}$ denotes the tensor product. These equations are linearly dependent on each other and have to be solved self-consistently. The energy is calculated by

$$E = \frac{1}{2} \sum_i q_i V_i - \frac{1}{2} \sum_i \boldsymbol{\mu}_i \mathbf{E}_i \quad (3.36)$$

where V_i is the electrostatic potential at atom i .

An undesirable property of this point dipole model is that it may lead to infinite polarization by the head to tail interaction between two induced dipoles in the direction of the line connecting the two, when they are very close to each other. As a remedy Thole developed a modified dipole field tensor, where the point charge is smeared out when the dipoles come closer than $s = a(\alpha_i\alpha_j)^{1/6}$ to each other. A model for the resulting charge density was developed leading to an associated dipole field tensor [85]

$$\hat{\mathbf{T}}_{ij} = \frac{\hat{\mathbf{1}}(4v^3 - 3v^4)r_{ij}^2 - 3v^4\mathbf{r}_{ij}\hat{\mathbf{r}}_{ij}}{4\pi\epsilon_0 r_{ij}^5} \quad \text{for } r < s \quad (3.37)$$

where $v = r_{ij}/s$. The values for $a = 1.7278$, $\alpha(\text{O}) = 0.948 \text{ \AA}^3$ and $\alpha(\text{H}) = 0.519 \text{ \AA}^3$ are derived by fitting to experimental molecular polarizabilities of a learning set of 16 small molecules [86].

With this model, the energy differences between donating and accepting state including atomic polarizabilities, E_{12}^{pol} , are calculated for the 10 transfer events, where the atoms of the water molecules in the sphere of radius 3.5 \AA are polarizable, and they remain classical point charges without polarizability in the sphere ring between 3.5 \AA

and 7 Å. To give the correct gas-phase dipole moment of water [87], the partial atomic charges of the polarizable water molecules are scaled by a factor 0.82. The results (see Table 3.3) deviate on average by 9.8 kJ/mol from the quantum mechanically derived E_{12}^{qm} . Including atomic polarizabilities, thus, does not improve the calculation of E_{12} , at least not, when using Thole's model for an additional proton in a water box. Future studies should investigate, whether more complex models of polarizability (e.g. considering anisotropy [87]) can improve the calculation of E_{12} or if polarizable models at all are favorable for heterogeneous transfer systems or heterogeneous environment of a transfer system. In the following, E_{12} will be calculated only considering the Coulomb interactions between the transfer system and the point charges on the surrounding atoms.

3.4 How to Conserve Energy and Momentum

When a proton hops, it is instantaneously positioned on the interconnecting line between the former proton position and the accepting atom of A_i at equilibrium distance according to the underlying force-field. Its initial velocity is set to zero. Obviously, this violates the principle of energy and momentum conservation. The bonded interactions V^{bond} , including angle bending and dihedral torsion, and the non-bonded interactions V^{nb} , including van-der-Waals and electrostatic interactions, differ in donating and accepting states. To conserve the energy during proton hopping, the following expression has to be fulfilled

$$V_1^{bond} + V_1^{nb} + \sum_i \frac{1}{2} m_i (v_i^{(1)})^2 = V_2^{bond} + V_2^{nb} + \sum_i \frac{1}{2} m_i (v_i^{(2)})^2 \quad . \quad (3.38)$$

Among all possible parameters, only the velocities after proton hopping $\mathbf{v}_i^{(2)}$ can be adjusted to meet this condition. From a practical point of view, the question emerges which atom velocities should be changed. On one hand the approach should be simple, on the other hand it should be physically sensible. The following model is postulated. If the total energy, including V^{bond} and V^{nb} , is larger for the accepting state than for the donating state, the velocities of donating atom $\mathbf{v}_D^{(2)}$ and proton $\mathbf{v}_H^{(2)}$ are simultaneously scaled down from $\mathbf{v}_D^{(1)}$ and $\mathbf{v}_H^{(1)}$. The proton receives a 'kick' from the donor to overcome the barrier. In the opposite case, where the total energy

decreases, the velocities of accepting atom $\mathbf{v}_A^{(2)}$ and proton $\mathbf{v}_H^{(2)}$ are scaled up. To conserve momentum, the following three conditions have to be fulfilled

$$\sum_i m_i \mathbf{v}_i^{(1)} = \sum_i m_i \mathbf{v}_i^{(2)} \quad . \quad (3.39)$$

In total the six components of $\mathbf{v}_D^{(2)}$ and $\mathbf{v}_H^{(2)}$, or $\mathbf{v}_A^{(2)}$ and $\mathbf{v}_H^{(2)}$ have to fulfill four independent equations. The two remaining free parameters have to be chosen carefully, e.g. fixing the ratio $|\mathbf{v}_D^{(2)}|/|\mathbf{v}_H^{(2)}|$.

This approach is somehow artificial and in the first case it is not always fulfilled for extreme cases. However including more atoms, e.g. the atoms connected to accepting or donating atom, is only slightly more elaborate. This method is not implemented in the ARGOS code so far, but implementation is planned after further testing.

Chapter 4

Test Systems

The approach outlined in Chapters 2 and 3 uses classical point particles, while the quantum nature of proton transfer reactions is taken into account in the calculation of the transfer probabilities. To demonstrate that Q-HOP MD gives satisfying results, an excess proton in a water box was first studied. Here, the observed structural properties as well as transfer rates can be compared with extended EVB [31] and MS-EVB [33] simulations that explicitly include quantum effects in the description of the proton, as well as with experiment. The LES implementation is not used in this case, since the energy barriers for proton transfer are small enough. In fact, the time scale for proton transfer is in the order of the orientational relaxation time of water.

As an exercise for the simulation of protonation equilibria in proteins, two model systems of amino acids, aspartic acid and imidazole, were studied in water boxes. In the study of imidazole, the LES implementation was tested for its potential to speed up the sampling of donor-acceptor configurations.

4.1 Excess Proton in a Water Box

4.1.1 Computational Details

An excess proton, i.e. H_3O^+ , was simulated in a solvent box with either 211 or 771 SPC water molecules [63] and using periodic boundary conditions. The system temperature was coupled to $T = 300$ K by a Berendsen thermostat [14] and the system pressure to 10^5 Pa accordingly. The leap-frog algorithm was used with a time step of 2 fs to

numerically solve Newton’s equation of motion. The long-range Coulomb interactions were calculated up to a cut-off distance of 12 Å. Hopping is allowed every 10 fs. After a successful proton transfer, no back transfer is allowed for 20 fs. H_3O^+ is modeled with fixed bond length and flexible bond angles taken from the AMBER95 force field [4]. Partial atomic charges for H_3O^+ were derived using the “restricted electrostatic potential fit” (RESP) module in the program NWChem [56] at HF/6-31G* level according to the AMBER95 procedure [94]. The charge fitting procedure of the RESP method starts by calculating the electrostatic field at a large number of points (e.g. 10000) around the molecule from a HF wavefunction. A least squares fit algorithm is then used to derive classical atom-centered point charges which best reproduce the electrostatic potential at these points around the molecule. In our simulations, the oxygen atom of H_3O^+ carries a charge of $-0.749e$ and the three hydrogen atoms $0.583e$. The other SPC water molecules carry a charge of $-0.82e$ on the oxygen atom and $0.41e$ on the two hydrogen atoms. 16 simulations were carried out over 0.5 ns for H_3O^+ in the smaller box with 211 water molecules, and one simulation run over 1 ns in the larger box with 771 water molecules.

4.1.2 Properties

The radial distribution function $g_{O^*O}(r)$ of the hydronium oxygen O^* with any of the remaining oxygen atoms O is shown in Figure 4.1 for H_3O^+ in 771 water molecules. The radial distribution function of bulk water in experiment [84] shows a maximum at $\sim 2.8 - 2.9$ Å, corresponding to the first solvation shell. The maximum of the radial distribution function $g_{O^*O}(r)$ is shifted towards smaller oxygen-oxygen distances due to strong attractive Coulomb forces between the positive hydronium ion and the oriented dipoles of the coordinated water molecules. This leads to an enhancement of the ordered structure around the hydronium ion, which is reflected by the strongly pronounced first peak of $g_{O^*O}(r)$ at 2.58 Å. A second less pronounced peak is found at about 4.5 – 5.0 Å for $g_{O^*O}(r)$ which corresponds to the second oriented solvation shell. Integration of the radial distribution functions gives a coordination number of ~ 3 for the hydronium ion compared to ~ 4 for bulk water molecules. Whereas bulk water molecules predominantly form tetrahedral complexes with 4 hydrogen bonds of 2.8 – 2.9 Å length, solvated H_3O^+ is a nearly plane-trigonal complex forming three

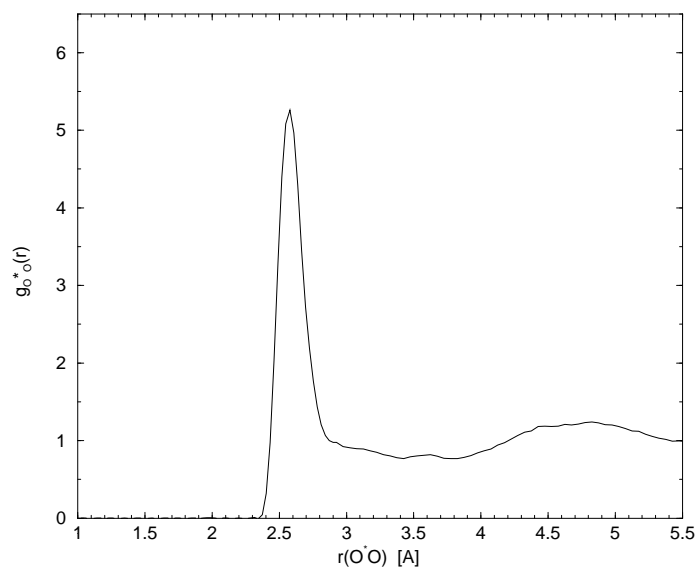


Figure 4.1: Radial distribution function $g_{O^*O}(r)$ of the hydronium oxygen O^* with any of the remaining oxygen atoms O for the simulation of an excess proton in a water box with 771 SPC molecules.

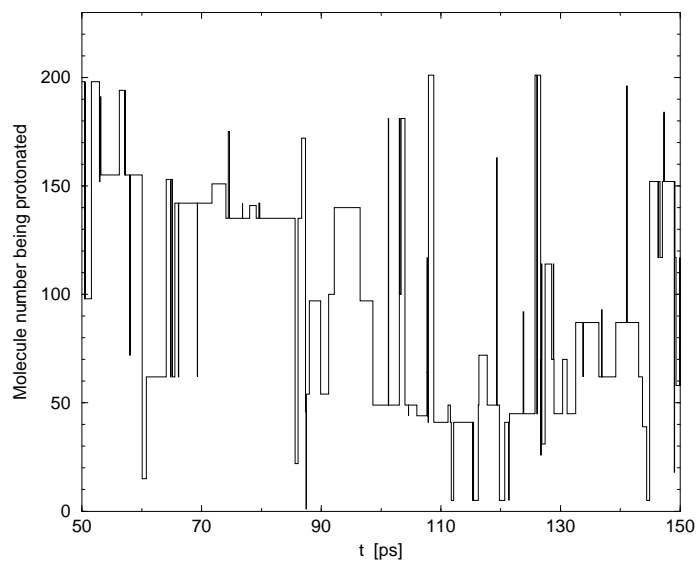


Figure 4.2: Time evolution of the molecule number of the hydronium ion for the simulation of an excess proton in a water box with 211 SPC molecules. An arbitrarily chosen window of 100 ps length is shown. A change in molecule number indicates a proton hopping event.

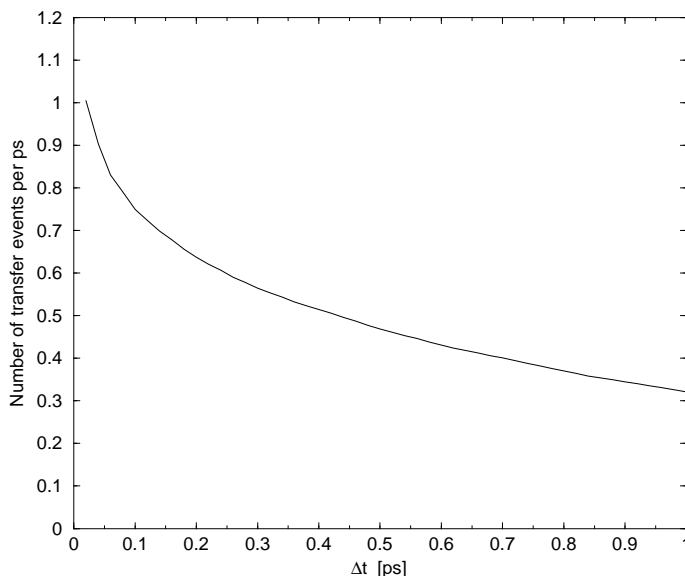


Figure 4.3: Number of transfer events per picosecond, where the proton stays longer at a molecule site than the time indicated on the ordinate before hopping to another acceptor.

hydrogen bonds to the nearest water molecules.

The dynamic properties are studied for H_3O^+ in 211 water molecules. As an example Figure 4.2 shows the time evolution of the molecule number of the hydronium ion for one of the 16 simulation runs. Compared to the results of MS-EVB simulations [33], there is no fast proton flipping between two water molecules observed. Since we are not interested in the detailed motion of the proton during the proton transfer process, we only allow consecutive hopping events every 20 fs. This leads to reduced flipping rates between two strongly hydrogen bonded water molecules sharing a proton in a quantum mechanical picture. To compute meaningful transfer rates, Vuilleumier *et al.* [31] and Schmitt *et al.* [33] proposed to consider only hopping events where the proton stays at the acceptor site for longer than 0.5 ps after transfer. By treating all nuclei classically, Vuilleumier *et al.* obtained a rate of 0.11 ps^{-1} , whereas Schmitt *et al.* observed a rate of 0.28 ps^{-1} for the classical treatment and 0.50 ps^{-1} for a centroid molecular dynamics (CMD) treatment of the proton. Here, we calculated a rate of 0.47 ps^{-1} which is very close to the CMD result [33] and is in the same order of experimental NMR results ($\sim 0.63 \text{ ps}^{-1}$) [82].

Obviously the criterion to count only transfer events which last for longer than 0.5

ps at the acceptor site after transfer is a little arbitrary, and more elaborate criteria have been used [33]. This simple measure is used here, simply to allow comparisons between our work and the EVB simulations [31, 33]. Figure 4.3 shows the number of transfer events per picosecond, where the proton stays longer at the acceptor than a particular time Δt . The number of transfer events is shown for varying Δt . Changing Δt from 0.5 ps to smaller times would lead to an increase in the computed transfer rate.

To shed light on the microscopic structural details of proton transfer, the upper panel of Figure 4.4a displays the distances between hydronium oxygen and the oxygen atoms of the three closest neighbor molecules. The energy difference between the donor and the acceptor bound state, E_{12} , is shown in the lower panel, respectively. The vertical lines mark the proton hopping events in this time window. The closest water molecule is usually at 2.4 Å to 2.6 Å distance away from the donating oxygen. Most of the transfer events observed here happen at distances smaller than 2.5 Å. But there are also many cases where the closest acceptor approaches the donor by less than 2.5 Å and does not hop (arrows). What determines hopping or non-hopping is the environmental configuration. In most hopping events shown here, the second and third closest water molecules are further away than on average. This leads to a destabilization of the hydronium ion due to weaker electrostatic interaction between the hydrogen bonded water molecules, not considering the acceptor. Thus, the transfer to the closest water molecule becomes energetically more favorable. In the cases where all three water molecules come close to the donor (arrows), such destabilization of the hydronium ion does not occur, and the transfer is energetically unfavorable. Figure 4.4b shows an example of proton hop where the second and third water molecule is only slightly further away from the donor than on average. In this case the donor stays in an energetically favorable situation for proton binding after transfer, and the proton hops back after 50 fs.

A similar mechanism concerning the distances to coordinating water molecules was suggested by Agmon [88, 89], although he mainly emphasized the dynamic event of hydrogen bond cleavage to induce proton transfer thereby changing the coordination number of O^* from 3 to 4. In the mechanism (Figure 4.5), he postulated the cleavage of a hydrogen bond in the second solvation shell around the hydronium ion (a \rightarrow b) between O_C and O_D . The hydrogen bonded water molecules readjust ultrafast to

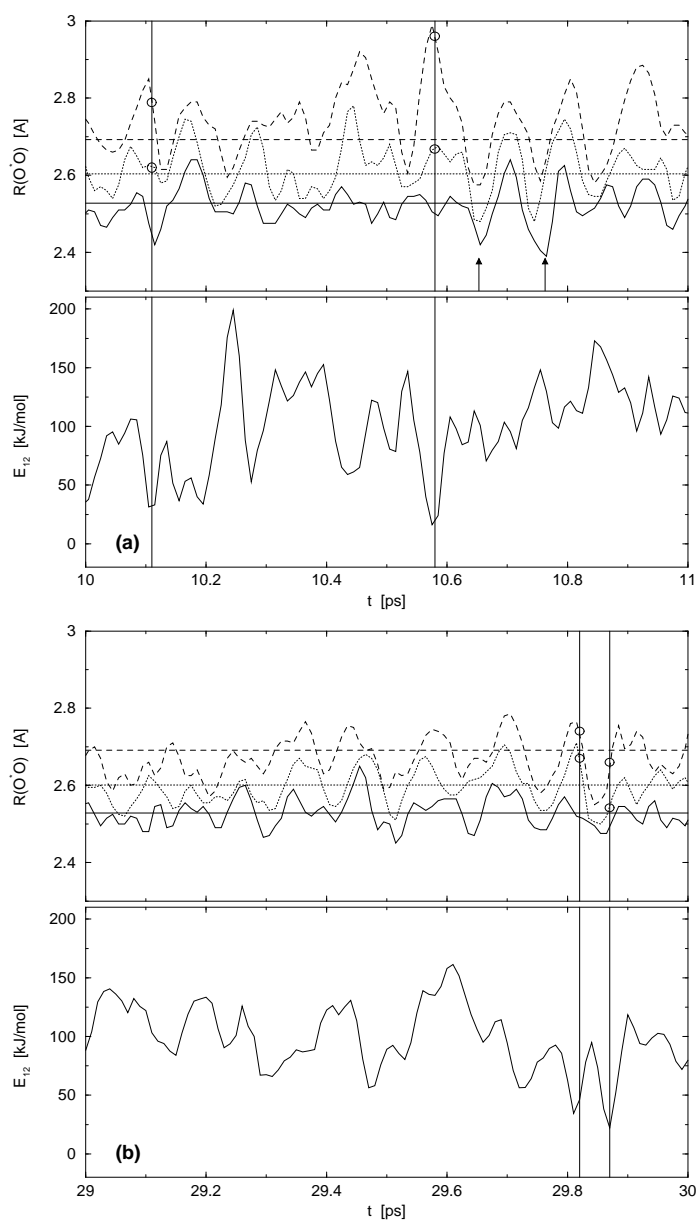


Figure 4.4: *Upper panels:* Distances between the hydronium oxygen O^* and the oxygen atoms of the three closest neighboring molecules. The mean distances of the three closest molecules during this time window are shown as horizontal lines. The vertical lines mark the successful transfer events and the small circles the actual distances between donor and second- and third-closest water molecules at this moment. *Lower panels:* The electrostatic difference E_{12} between donor and next acceptor is shown in parallel. (a) shows two separate transfer events, (b) shows a quick back-and-forth jump of the proton between two water molecules.

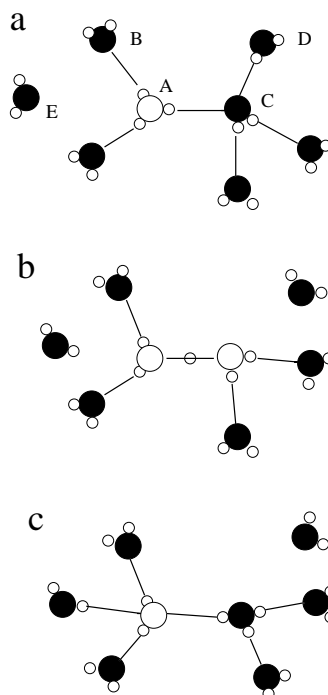


Figure 4.5: A molecular mechanism for proton transfer in water according to Agmon [88, 89].

form a H_5O_2^+ complex including O_A and O_C . The $\text{O}_A\text{-O}_C$ distance shrinks to a value around $2.4 - 2.45 \text{ \AA}$. At the same time, the distances to the two remaining closest water molecules increase to about $2.7 - 2.9 \text{ \AA}$. Fast fluctuations of surrounding water dipoles momentarily stabilize the additional proton, bonded either to O_A or to O_C . If one of these fluctuations is coupled to a reorganization of the water molecule H_2O_E towards water H_2O_A , the additional proton is localized on O_C forming a H_3O_4^+ ion. Here, the simulations demonstrated exactly the ultrafast readjustment to form a H_5O_2^+ -like configuration. Q-HOP MD cannot so far describe a proton shared between two water molecules. Nevertheless, as demonstrated in Figure 4.4, in case of hopping, the two water molecules H_2O_A and H_2O_C are around 2.45 \AA apart, whereas the distance from H_2O_A to the next water molecules is around $2.7 - 2.9 \text{ \AA}$. Whether the breaking of a hydrogen bond in the second solvation shell (between O_C and O_D) initiates the transfer could not be clarified in our simulations. Hydrogen bonds in the second solvation shell are often broken and newly formed without any proton transfer.

Another experimental observable is the diffusion coefficient of the excess proton. 16

simulation runs over 500 ps were performed for an excess proton in a box of 211 water molecules. The diffusion constant was computed using the Einstein relation

$$D = \left\langle \frac{|\mathbf{r}(t) - \mathbf{r}(0)|^2}{6t} \right\rangle, \quad (4.1)$$

where the averaging was performed over the hydronium oxygen position \mathbf{r} at various time t . The observed result of $9.0 \cdot 10^{-5} \pm 3.9 \cdot 10^{-5} \text{cm}^2 \text{s}^{-1}$ is in good agreement with the experimental value of $9.3 \cdot 10^{-5} \text{cm}^2 \text{s}^{-1}$, which is quite remarkable because the parameterization of the hopping probabilities was completely independent of any dynamic property.

4.2 Aspartic Acid in a Water Box

4.2.1 Computational Details

Keeping in mind the remarkable match of the simulation results for an additional proton in a water with experimental results and with more elaborate computational simulations, an amino acid in a water box is simulated as next step. It is the first time that the protonation equilibrium between an amino acid and a solvent box is simulated dynamically in atomic detail. First, an isolated aspartic acid with neutral end groups was simulated in a solvent box with 476 SPC water molecules. The partial charges and the bonded parameters are taken from the AMBER95 force field. All other simulation parameters and conditions (temperature, pressure, time step, cut-off, maximal hopping frequency) were the same as for the simulation of an excess proton in a water box. Five runs were simulated over 0.5 ns each.

4.2.2 Properties

Figure 4.6 shows the time evolution of the molecule number being protonated n , where $n = 1$ stands for the aspartic acid residue, for one simulation. Roughly, two different regions can be observed. One region during which the aspartic acid residue ASH is protonated most of the time, interrupted by transfers to accepting water molecules nearby. During this time, the occupancy of protonated aspartic acid was 94.7%. Sometimes the proton does not hop back to the aspartic acid residue, but is transferred to

a water molecule further away. Immediately, other water molecules between ASP and hydronium ion screen the strong Coulombic interaction between negative ASP and hydronium ion. The excess proton escapes towards the bulk water and hops through the water box until it eventually diffuses back into the attractive part of the negative ASP, where it can be caught again forming a protonated ASH. The total occupancy of protonated ASH averaged over the five runs is 18.3%. A direct comparison to the experimental pK_a -value of aspartic acid ($pK_a = 3.9$) is not possible since the simulations are not really performed at constant pH. According to this pK_a -value one would expect an occupancy of protonated ASP of 95 %. Rather, the system jumps from infinite pH (ASH, un-protonated water molecules) to $pH \sim 1$ (ASP, one hydronium ion per 476 water molecules). In order to ensure a constant pH bath, the simulation box would have to be significantly expanded, containing several hydronium ions. Also a better description of the long-range electrostatics using Ewald-summation might improve the results for this kind of simulation, where the system flips between two states, one containing two different ions and one containing only neutral molecules. For this kind of simulation, it is known that a Ewald-sum description of the long-range Coulomb interactions improves the results.

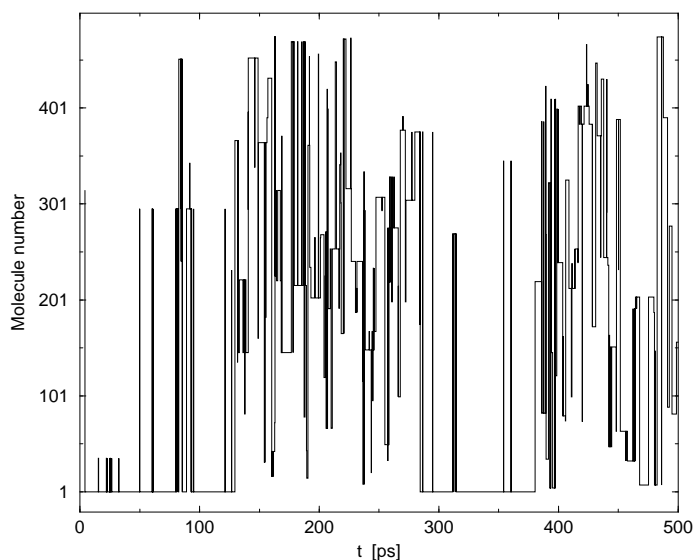


Figure 4.6: Time evolution of the molecule number being protonated for a simulation of an aspartic acid residue in a water box.

4.3 Imidazole Ring in a Water Box with Multi-Copy Simulation

4.3.1 Computational Details

In both systems described so far, transfer events happen quite often in picoseconds intervals. Therefore it was not necessary to use the LES technique to increase conformational sampling in these cases. To demonstrate the significance of LES for proton transfer reactions a system was chosen where it is expected that proton transfer events happen much more rarely. Therefore an amino acid was selected with a much higher pK_a value than aspartic acid. Nevertheless the pK_a -value should be low enough that a transfer event can be observed in reasonable computing time. Therefore an isolated imidazole ring as model for a histidine amino acid residue ($pK_a = 6.0$) is simulated in a solvent box with 207 SPC water molecules. The long-range Coulomb interactions were calculated up to a cut-off distance of 10 Å. All other simulation parameters and conditions are the same as for the two previous examples. The system was simulated five times with slightly different starting conditions over 500 ps. In the first simulations, all groups were present only as a single copy. In a second set of five simulations, the donor and the possible acceptors are represented by five copies.

4.3.2 Properties

In the single copy simulations over 500 ps, only one proton transfer was observed after 220 ps from imidazole to a water molecule in five simulations. This is in contrast to the simulations of an aspartic acid residue in a water box and reflects the much higher pK_a value of 6.0 for histidine compared to 3.9 for aspartic acid.

When representing the donating imidazole and the putative proton acceptors by multiple copies (five copies here), transfer events occur much earlier than previously in all the five simulations. The first transfer events occurred after 28 ps, 36 ps, 43 ps, 102 ps and 181 ps for the five simulation runs. The simulations were executed only until after the first proton transfer, since it was only to demonstrate the superior sampling by LES. Nevertheless this small test system demonstrates the advantage of using LES for systems where transfer events happen quite rarely.

Chapter 5

Green Fluorescent Protein (GFP)

5.1 Introduction of the GFP Protein

Green Fluorescent Protein (GFP) was discovered as a companion protein to the chemiluminescent protein aequorin in jellyfish *Aequorea*. Whereas the chemiluminescence of aequorin is blue and peaked around 398 nm, the emission spectrum of GFP is peaked at 508 nm. Therefore the GFP molecule converts the blue light emitted from aequorin to green light. This gave it the name Green Fluorescent Protein.

It is one of the currently most widely studied proteins in biophysics and biochemistry and one of the most often used protein in cell biology, biochemistry and biotechnology. Its significance is due to two effects. First its ability to emit efficiently light in the visible range, and second that it consists of only one sequence of amino acids, independent of chemical groups like chromophores, which are not amino acids and which have to be taken from the environment of the protein to built separately into the protein. The latter is the usual case for fluorescent molecules. Since GFP consists of only one DNA sequence, it can be fused to other proteins. With genetic methods one can add a DNA sequence coding the GFP after the sequence of a protein of interest, say protein A. The intact cell then reads the fused sequence and synthesizes the protein A connected with the GFP molecule which functions like an antenna. This antenna will fluoresce green when it is shined by light at proper wavelength, thus marking the location of protein A in the cell. Therefore GFP is widely used in cell biology as marker, as physiological indicator or bio-sensor.

5.2 Proton Shuttle in GFP: What is known ?

The fluorescence excitation spectrum for wild-type GFP (Figure 5.1) has a major excitation peak at 395 nm that is about three times higher in amplitude than a minor peak at 475 nm. One possibility could be that the two maxima correspond to the excitation from ground state to the first excited state and to the second excited state. But the fact that excitation at 395 nm gives an emission at 508 nm, whereas excitation at 475 nm gives an emission at 503 nm [95], favors the model that the two excitation maxima correspond to two chemically distinct species (A and B). With increasing pH the amplitude of the 475 nm excitation peak increases at the expense of the 395 nm peak [96]. A similar pH dependence is found in the unfolded protein [97]. This supports the concept of two different chemical species corresponding to the excitation peaks. Moreover, it leads to the interpretation that the two species differ in their protonation states. The simplest interpretation is that the 475 nm peak corresponds to a GFP molecule with de-protonated or anionic chromophore (state B), whereas the 395 nm peak represents a GFP molecule with protonated or neutral chromophore (state A) [95, 99].

GFP consists of 238 amino acids where residues 65-67, which are Ser-Tyr-Gly in the native sequence, form the two conjugated aromatic rings of the chromophore by an autocatalytic reaction during protein folding [100]. The structure of GFP [101, 102] is shown in Figure 5.2. Eleven β -strands form a cylinder with 42 Å in height and with a diameter of 24 Å. Many hydrogen-bonds between the backbone atoms of the different strands [103] account for its stable configuration. In the inner pore of the cylinder an α -helix is found containing the chromophore in the center. The cylinder is closed on both ends by short α -helices and turns, which shields the inner part of the cylinder from the surrounding solvent. The surrounding of the photo-active chromophore is shown in Figure 5.3.

Ionization of the chromophore after light-induced excitation of the neutral chromophore would explain the emission peak at 508 nm, which is very similar to the frequency of the emission peak of the anionic chromophore at 503 nm. The observation that phenols in the excited state become more acidic in most cases supports this model. Additional evidence for an excited state proton transfer gave time-resolved absorption and fluorescence spectroscopic studies [98, 104]. Excitation at 395 nm shifts

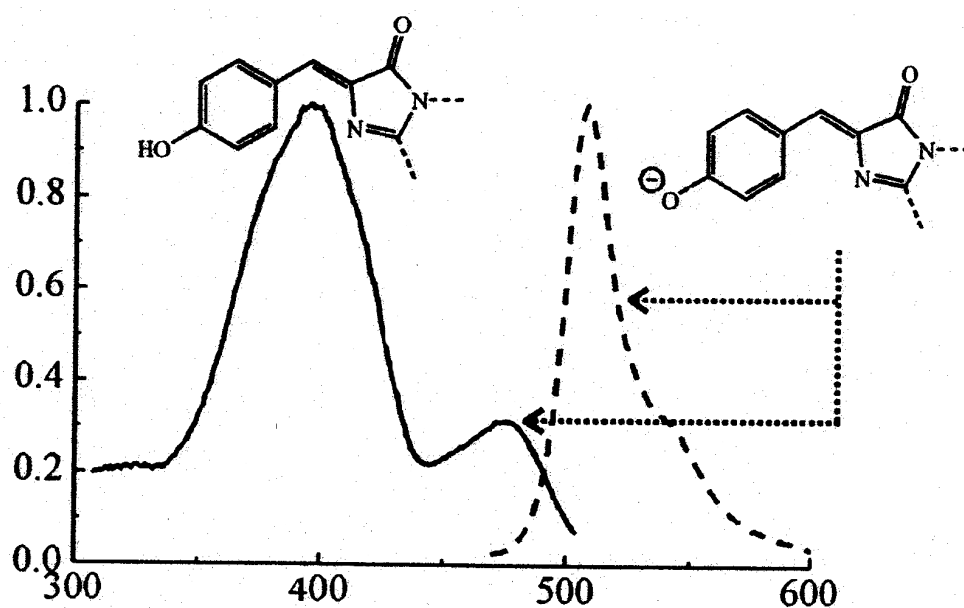


Figure 5.1: Fluorescence excitation (solid line) and emission (dashed line) spectra for wild-type GFP together with the believed protonation state responsible for the observed spectra (Taken from [25]).

the emission from a 460 nm to a 508 nm peak over about 10 ps. This is slowed down by cooling to 77 K and by deuterium substitution [98].

Two observations led to postulating an additional intermediate state I/I*. The first is that the emission peak at 508 nm after the excitation of the neutral chromophore (A-form) varies slightly from the excitation peak of the anionic chromophore (B-form) at 503 nm. The second is that during most fluorescence cycles, the proton transfer seems to eventually reverse leading to the neutral chromophore (A-form). However, in some cases the GFP molecule seems to convert into the B-form, that means that the proton seems to return not to the chromophore. This was observed by shining intense UV light for hours onto the GFP molecules. There the 395 nm absorption peak of the neutral A-form declines and the 475 nm peak of the anionic B-form increases [98, 99, 105]. Although the I-form has not been described in atomic detail so far, a structural model of the I-form was suggested on the basis of comparisons between the crystal structures of wild type (A-form) and mutant GFP proteins containing the B-form [106, 108] (see Figure 5.4). The atomistic mechanism postulated is that proton transfer occurs



Figure 5.2: Schematic model of the 3-D crystallographic structure of wild-type GFP. The chromophore is shown in atomic detail in green.

from the chromophore to Glu222 via a buried water molecule and Ser205, leading to the state I/I*. In this model Glu222 is assumed to be protonated in the I-form, thus requiring a raise of the pK_A of Glu222 by more than 4 pH-units. Stabilization energy may be provided by the extensive hydrogen bond network in the chromophore environment, and this protonation assignment is supported by electrostatic continuum calculations indicating that the protonation states of the chromophore and nearby Glu222 are tightly coupled [109]. By shifting the proton position from anti to syn position (see Figure 1), Glu222 may also play a crucial role in the transition from I to B considered to be slow [106]. Despite this discussion, currently unknown are the molecular structure of the I-form, the structural mechanism for transitions between the various forms, and the rates of interconversion.

Occasionally, Thr203 may rotate to stabilize the anionic chromophore via a hydrogen bond, forming the state B/B*. In the crystal structures of wild type GFP [106], two different conformations of Thr203 were found, about 15 % form a hydrogen bond to the phenol oxygen atom of the chromophore and 85 % with OH facing away from

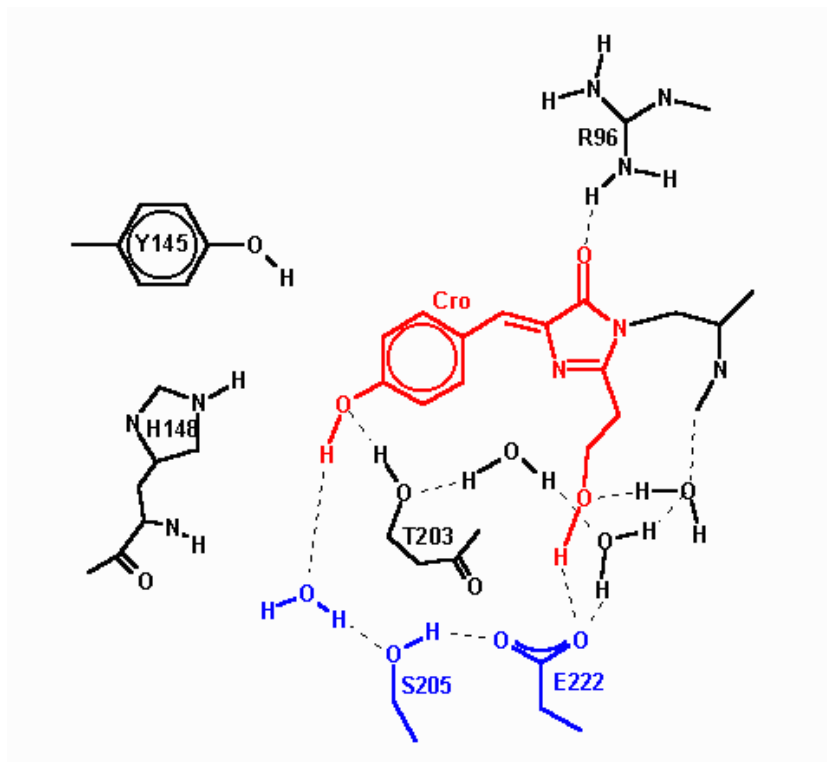


Figure 5.3: Surrounding of the chromophore (red). Drawn in blue are the chemical groups participating in the proton shuttle after excitation.

the chromophore. This is in quite good agreement with the spectroscopic ratio of anionic to neutral chromophore [98]. The postulate that there is more than one proton transfer step involved in the reaction is further settled by the time resolved absorption spectroscopic studies, where it was shown that the results for the conversion from A^* to I^* can only be described using a multi-exponential fit [98]. Newer hole-burning experiments have shown that the I-form is not just an intermediate. It is a stable state found at room temperature in the spectrum [107]. To summarize the experimental results, Figure 5.5 shows a postulated reaction scheme.

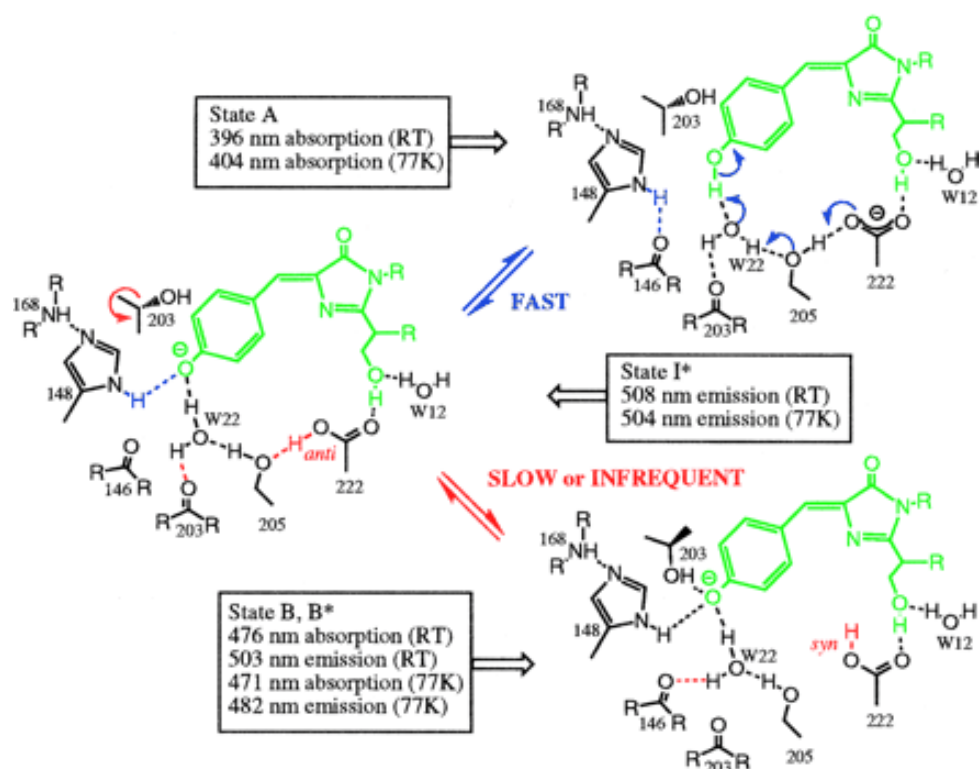


Figure 5.4: Proposed mechanism for proton shuttle after excitation ($A \rightarrow I^*$) as well as the infrequent photoisomerization to state B (Taken from [106]).

5.3 Proton Shuttle in GFP: Simulation with Q-HOP MD

The method Q-HOP MD is applied to the proton relay process in GFP. In general, the simulations agree well with the model of Brejc et al. and Palm et al. [106, 108] in that Glu222 protonated in anti-position is the I-form, and that transition from I to B may proceed via an intermediate with Glu222 protonated in syn-position. The key findings of this study are kinetic estimates for the various proton transfer steps leading from A^* to I^* and from I to A. The short life time of the fast intermediate I-form indicates that the intermediate with Glu222 protonated in anti-position cannot be the I-form observed in spectroscopic studies under equilibrium.

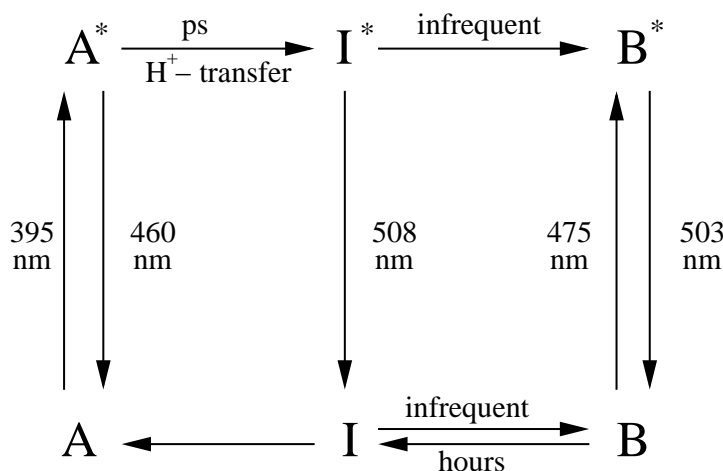


Figure 5.5: Proposed photo-cycle of wild-type GFP.

5.3.1 Simulation Parameters and Starting Conditions

The MD simulations were started from a previously equilibrated system of a single GFP molecule in a rectangular solvent box [103] using periodic boundary conditions. To save computation time, the system size was reduced to a system containing 16,993 atoms in total keeping at each side a minimum distance of 4 Å between GFP and the wall of the simulation box. 20 ps of additional equilibration were performed to adapt the system to the smaller box size. 15 ps could be simulated per day on a 1 GHz desktop PC. The system temperature was kept at 293 K and the pressure at 10^5 Pa, using Berendsen thermostats and a barostat. Missing proton transfer barriers E_{12}^0 were parameterized as described previously and are listed in Table 5.1. In the simulations, proton transfer was only allowed between chromophore, Wat22, Ser205, and Glu222. Wat22 was considered in three protonation states as H_3O^+ , H_2O , and OH^- . Partial atomic charges of neutral and anionic chromophores in their relaxed ground and first excited singlet states were previously computed by electrostatic potential fit [110].

The torsional barriers O-C-O-H and C-C-O-H for the transition between anti- and syn-positions in Glu222 were modified to barrier heights of 20 kJ/mol according to Hartree-Fock calculations of the torsional barrier.

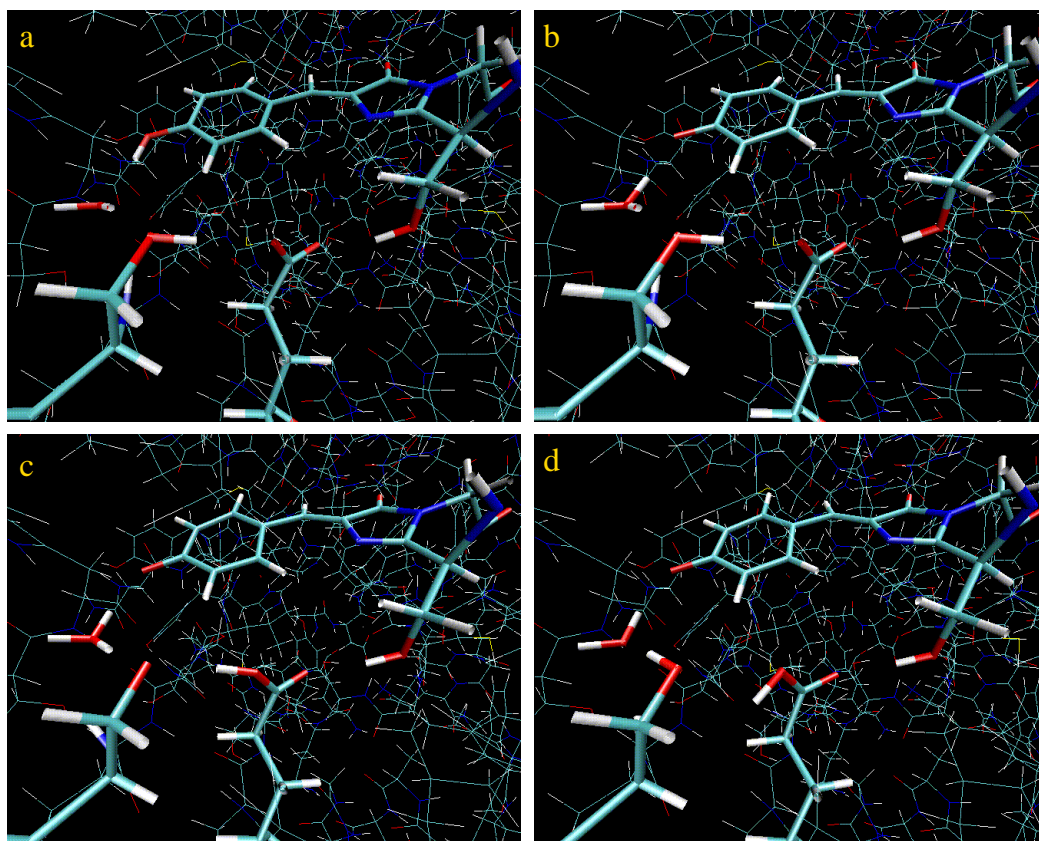


Figure 5.6: Snapshots from Q-HOP MD simulations to illustrate the proton shuttle between putative A^{*}- and I^{*}-form. Panel (a) shows the starting situation with protonated chromophore and Ser205 pointing towards Glu222. In Panel (b) a first proton is transferred from the chromophore to Wat22. The next two steps in (c) and (d) seem to occur as one concerted process: transfer from Ser205 to Glu222 and transfer from protonated Wat22 to re-protonate Ser205.

Table 5.1: Parameters α , β and γ to express E_{12}^0 as function of the donor-acceptor distance by equation (2.9) for donor acceptor pairs including chromophore in electronic groundstate.

Donor		Acceptor	
		H ₂ O	OH ⁻
Cro (groundstate)	α [kJ/mol]	-811.4	-51.0
	β [kJ/molÅ ⁻¹]	553.9	67.0
	γ [kJ/mol Å ⁻²]	-67.3	-41.9

5.3.2 Proton Shuttle after Excitation

The first proton transfer after excitation $A \rightarrow A^*$ will protonate the nearby water molecule 22, (see (a) \rightarrow (b)) in Figure 5.6. The simulations were started from a carefully equilibrated system of a GFP molecule in its electronic ground state solvated in a rectangular solvent box. To model the proton shuttle following electronic excitation of the chromophore, the partial atomic charges of the chromophore were replaced by appropriate charges for the excited state. Since it is very hard nowadays to accurately compute the energy barrier for proton release from the excited chromophore by quantum mechanical methods, the proton is simply shifted from Cro to the water molecule, and the MD simulation is started after this first step at the situation $\text{Cro}^- : \text{H}_3\text{O}^+$. To our great surprise, the next two transfer events simulated ((b) \rightarrow (c)) $\text{Ser205-OH} : \text{Glu222-COO}^- \rightarrow \text{Ser205-O}^- : \text{Glu222-COOH}$ and ((c) \rightarrow (d)) $\text{H}_3\text{O}^+ : \text{Ser205-O}^- \rightarrow \text{H}_2\text{O} : \text{Ser205-OH}$ occur within 10-80 fs and within 10 fs (see Table 5.2). 10 fs is actually the shortest transfer interval that we allow in our Q-HOP MD approach. The two transfer steps are strongly coupled to each other because the de-protonated Ser205-O^- is a very unstable intermediate in the reaction. Moreover, both steps are ultra-fast barrier-less processes leading to a much more favorable product state. The reason for this "down-hill process" seems to come from the modified charge distribution of the chromophore in its excited state because no transfer occurs from Ser205 to Glu222 when simulations are started from the $\text{Cro}^- : \text{H}_3\text{O}^+$ situation with ground state charges of the chromophore while not allowing back-transfer from H_3O^+ to the chromophore. After three subsequent transfer steps, the proton ends up at Glu222 in

anti-position. Using the re-parametrized barrier between anti- and syn-positions it will remain in anti-position during typical simulation times of 100 ps.

The second and third proton transfer events $A \rightarrow A^* \rightarrow I^*$ occur ultrafast whereas the two main decay constants for the fluorescence from A^* were measured as 3.6 picoseconds (ps) (50%) and as 12 ps (40%) [98]. Therefore these must be the decay constants for the first ESPT proton transfer from the electronically excited chromophore to the nearby water molecule and it can be expected that the chromophore will have structurally relaxed during this time. Assigning a life time of ca. 5-10 ps to the excited neutral chromophore is also in agreement with the experimentally observed weak fluorescence between 440-460 nm that originates from the neutral form.

5.3.3 Back Shuttle to Chromophore in Ground State

Following emission from I^* after 3.3 ns on average [25], the proton on Glu222 is expected to re-protonate the chromophore according to the current model [106, 108]. Consequently, a second set of simulations was started from the final configurations obtained previously, where the partial atomic charges of the chromophore were set back from electronic excited to ground state. The sequence of transfer events observed is shown in Figure 5.7. Table 5.3 shows typical transfer times. Simulations started with Glu222:H δ in syn-position did not lead to re-protonation of the chromophore because transitions between anti- and syn- conformation are very unlikely on simulation time scales of 100 ps as noted before.

Based on our observed proton transfer rates and in agreement with the crystallographic model [106, 108], the configuration with Glu222 protonated in anti-position is confirmed as the long-sought I-form when the I-form is not defined in a spectroscopic sense [107], but as the fast intermediate after ESPT from A^* [98]. Its relatively short life time (1 ps to 1 ns, compare Δt_1 in Table 5.3) clearly rules out the possibility that this form may be visible under continuous illumination by laser light [107]. On the other hand, the same molecular configuration with H δ on Glu222 relaxed into syn-position should have a considerably longer life-time and therefore be visible (Figure 5.8, lower right panel). We therefore suggest the existence of two photo-physically relevant I forms, a 'non-relaxed' I-form with H δ on Glu222 in anti position, and a 'relaxed' I-form with H δ on Glu222 in syn position. The latter form may structurally

Table 5.2: Data for proton transfer events from 14 molecular dynamics simulations of steps (2) and (3) of the proton shuttle connecting states A* and I* in wild-type GFP (see also Figure 5.8). Different snapshots from the equilibration run were used as starting configurations and different random number seeds for the generation of hopping probabilities. Δt_2 is the time interval between the start of the simulation in the excited state of the chromophore (with the first proton being transferred from Cro to Wat22) and the second proton transfer between Ser205 and Glu222. Δt_3 is the time interval between second and third proton transfer from H_3O^+ to Ser205. $R(\text{DA})$ and E_{12} are the donor-acceptor distance and the relative energy difference between donor- and acceptor-bound states at the moment of transfer, respectively.

Run #	Ser-H \longrightarrow Glu-COO ⁻			H ₃ O ⁺ \longrightarrow Ser ⁻		
	Δt_2 [fs]	$R(\text{DA})$ [Å]	E_{12} [kcal/mol]	Δt_3 [fs]	$R(\text{DA})$ [Å]	E_{12} [kcal/mol]
1	10	2.55	0.0	10	2.50	-36.3
2	10	2.59	1.8	10	2.54	-33.4
3	10	2.47	5.8	10	2.60	-42.1
4	80	2.40	1.4	10	2.57	-27.0
5	80	2.50	-0.1	10	2.59	-31.2
6	40	2.65	2.7	10	2.46	-25.4
7	10	2.54	0.0	10	2.63	-28.7
8	30	2.52	2.1	10	2.54	-27.8
9	80	2.54	3.5	10	2.50	-24.0
10	80	2.51	6.3	10	2.44	-18.8
11	20	2.51	6.5	10	2.35	-14.8
12	70	2.53	-0.8	10	2.57	-26.4
13	20	2.66	3.3	10	2.59	-34.6
14	10	2.57	2.0	10	2.64	-38.9

Table 5.3: Data from 10 molecular dynamics simulations of steps (1) to (3) of the proton back-shuttle connecting states I and A in wild-type GFP (see also Figure 5.8). In step (1), the proton may shuttle several times between H₂O and Cro before the second transfer occurs. '3 in 3.77' means that the proton hopped three times back and forth before step (2) took place.

Run #	H ₂ O → Cro ⁻			Ser-H → OH ⁻			Glu-COOH → Ser ⁻		
	Δt_1 [ps]	$R(\text{DA})$ [Å]	E_{12} [kcal/mol]	Δt_2 [fs]	$R(\text{DA})$ [Å]	E_{12} [kcal/mol]	Δt_3 [fs]	$R(\text{DA})$ [Å]	E_{12} [kcal/mol]
1	3 in 3.77	2.54	-5.5	440	2.47	9.1	10	2.59	-10.1
2	6 in 46.5	2.51	0.7	10	2.59	0.3	10	2.62	-13.3
3	1 in 0.94	2.48	2.3	330	2.50	8.4	10	2.65	-12.6
4	1 in 2.07	2.52	3.5	530	2.41	6.8	20	2.60	-6.1

be converted into the B-form by rotating Thr203 and by a side-movement of His148. Different positions of His148, either hydrogen bonding to the chromophore or not, may actually give rise to additional spectroscopic 'I-forms' e.g. observed in low-temperature studies [111].

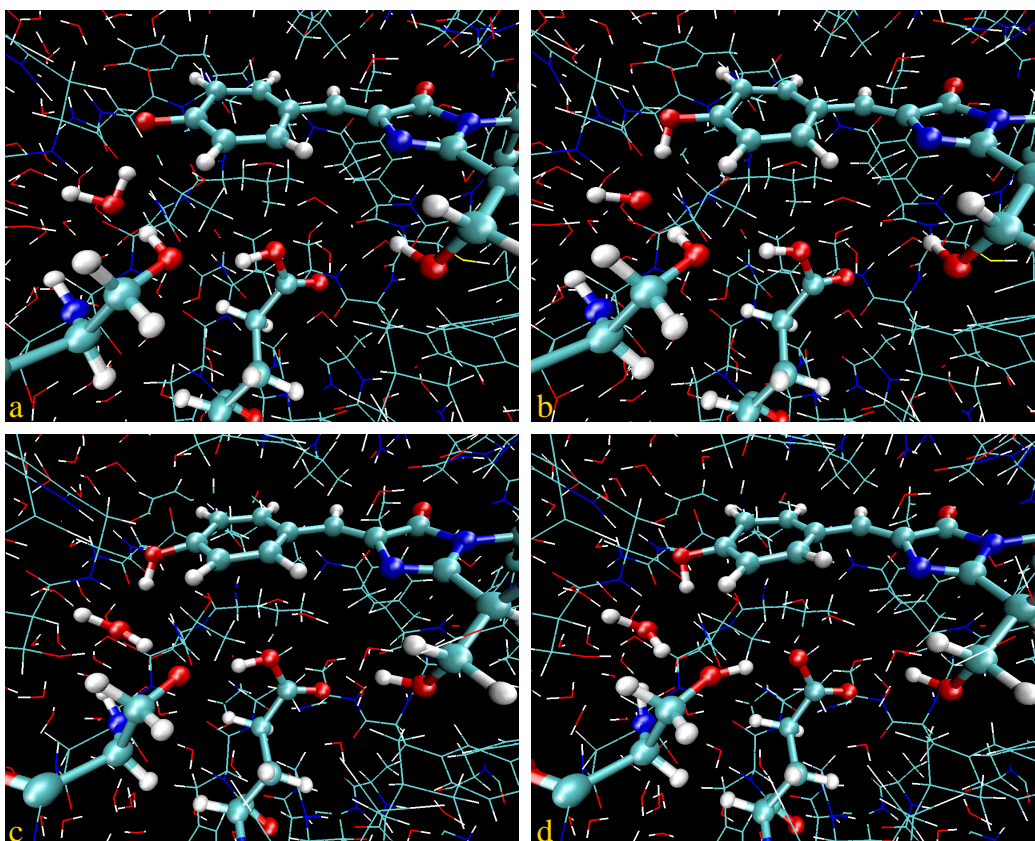


Figure 5.7: Snapshots from Q-HOP MD simulations to illustrate the proton back-shuttle between putative I- and A-form. Panel (a) shows the starting situation with protonated Glu222 and Ser205 pointing towards Wat22. In Panel (b) a first proton hopped from Wat22 to re-protonate the chromophore. The next two steps in (c) and (d) seem to occur as one concerted process: transfer from Ser205 to Wat22 and transfer from Glu222 to re-protonate Ser205.

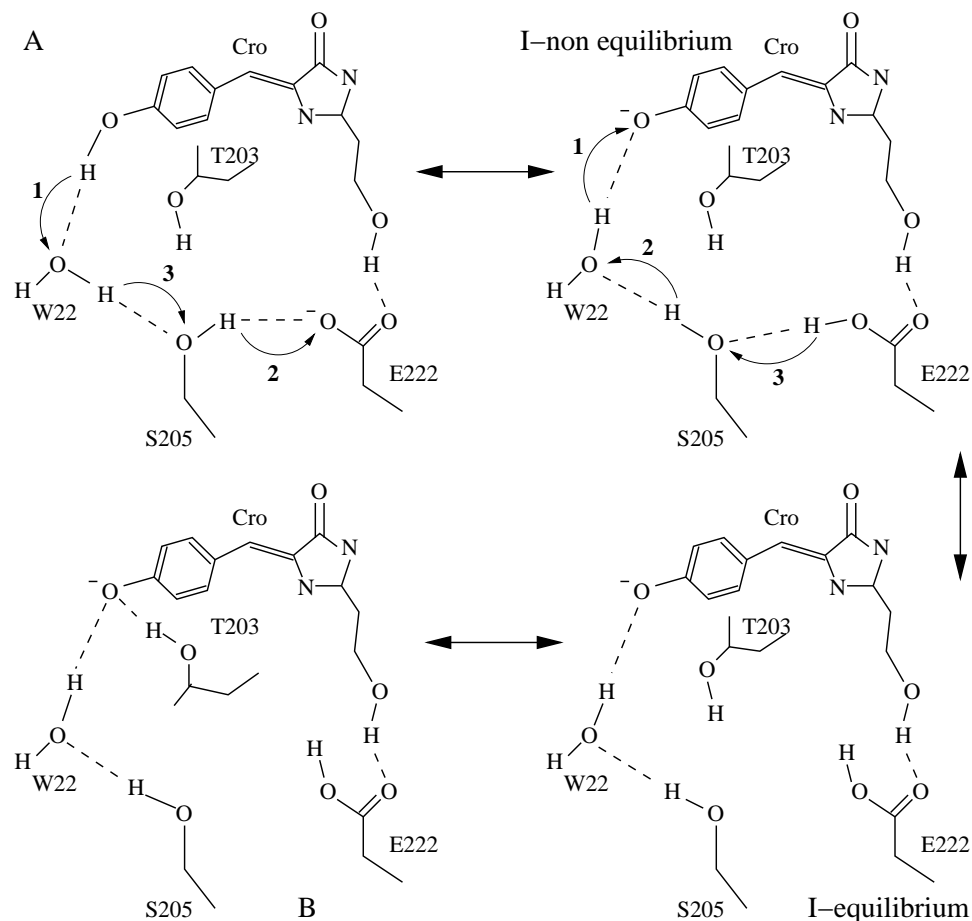


Figure 5.8: Schematic drawing of the chromophore surrounding in GFP. From upper to lower left are shown the A-form with a neutral chromophore, two putative intermediate I-forms, and the B-form with anionic chromophores. This model is closely related to the models of Brejc et al. and Palm et al. [106, 108]. The most clearly visible change between A and B is a rotation of Thr203 around its χ_1 angle to cause the hydroxyl group to either point towards the anionic chromophore or point away from the neutral chromophore. In the upper left panel, **1 - 3** indicate the three proton transfer events and the sequence in which they occur upon the transformation $A^* \rightarrow I^*$ (see also Figure 5.6). In the upper right panel, **1 - 3** denote the proton transfer events during the transformation $I \rightarrow A$ in the ground state (see also Figure 5.7).

Chapter 6

Outlook

6.1 Summary

A novel scheme was presented to simulate proton transport by classical molecular dynamics simulation coupled with quantum mechanically derived proton hopping. The main advantages of the method Q-HOP MD are high computational speed and accurate estimation of proton transfer probabilities. It is therefore ideally suited for studying proton transport in proteins.

To compute the proton transfer probabilities, the energy curve along the donor-acceptor interconnecting line was calculated in Chapter 2 using time independent quantum mechanical methods at MP2 level. Small systems consisting of donor and acceptor group were considered and some groups were added to model part of the environment of the transferring system. The energy curves are found to mainly depend on three factors, first the distance of donating and accepting atom, secondly the proton affinity of donating and accepting group and thirdly the relative stabilization of donating and accepting state due to electrostatic interactions with the environment.

TST theory can be applied, if there is a sufficient energy barrier. The energy barrier can then be computed as a simple function dependent on the three factors mentioned. The quantum nature of the transferred proton is approximated at the level of an effective zero-point energy along, and tunneling is included using the semi-classical WKB approximation.

TST is, however, not applicable for small energy barriers comparable in size to the

zero-point energy or even barrier-less configurations. These situations are common, for example, for proton transfer in bulk water. In these cases the time-dependent Schrödinger equation in one dimension was solved based on the energy curve describing the proton transfer. The timescale for proton transfer over small barriers was found to be in the order of 10-40 fs. Transition probabilities can reach values as high as 100 %, and can easily be fitted by an analytical expression. An interpolation for proton transfer rates is derived to connect the low-barrier-regime that should be treated by solving the time-dependent Schrödinger equation with the high-barrier-regime where TST applies.

Since local dynamic re-orientations of titratable groups can be the rate limiting steps in proton transport, the method of LES was implemented to increase conformational sampling. A special adaptation of LES for simulating proton transfer reactions was developed in Chapter 3.

The results presented in Chapter 4 were obtained by the simulation scheme Q-HOP MD. Proton transfer rates and the diffusional constant of the excess proton in a water box are in good agreement with the experiment. This indicates that the transfer probabilities fitted accurately to quantum mechanical calculations performed for small systems result in a realistic treatment of proton transfer in solution. On the other hand, the results for the simulations of an aspartic amino acid residue in a water box or for an imidazole ring in a water box cannot be compared with experimental values so easily. It is noteworthy that it is apparently possible to model proton exchange between an aspartic acid residue and water molecules in a small simulation box on a simulation time scale of several nanoseconds. However, the average populations of protonated and un-protonated aspartic acid residues observed in the simulations are preliminary results and should be taken carefully when comparing with experimental pK_a values. First, the pH of the simulation box is not constant as was already mentioned in the results section. And, secondly, the long range electrostatic interactions were not treated at all, while it is well-known that Ewald-type techniques should be used for proper treatment of ion pairing in solutions [90, 91]. And this also applies to the system containing an aspartic acid residue, where one is interested in the equilibrium between a contact ion pair and a situation where both ions are solvated separately. At this point I would like to mention that the implementation of Q-HOP MD in the ARGOS program [13] is a preliminary implementation. The main reason for choosing ARGOS

was the implementation of multi copy techniques for non-bonded interactions [46]. It is planned to soon implement the new methodology into the modern quantum chemistry package NWChem [56], that is specifically tailored for parallel architectures and allows to compute long-range electrostatic interactions by the Particle-Mesh-Ewald method. Its MD part is based on the sequential ARGOS program, which should considerably facilitate the implementation.

Modeling of the proton as a classical point particle does not allow to treat effects of electronic delocalization that may be important in certain transfer situations. It remains to be investigated in more detail whether Q-HOP MD can reproduce the predictions of more accurate quantum mechanical methods, for example about the co-existence of the so-called Zundel- and Eigen-cations [34, 33], H_5O_2^+ [92] and H_9O_4^+ [93]. However, in contrast to bulk solution, where a proton may be de-localized over two water molecules, we expect the proton probability density to be more localized when bound to protein residues that form deep energy minima compared to the shallow minima in bulk water. On the other hand, the simulation approach allows to employ 2 fs time steps as is commonly used in classical MD simulations, whereas 0.5 fs and 0.25 fs are typical time steps of flexible water models that are employed in the MS-EVB method [33], and in the method of Billeter & van Gunsteren [36, 37]. This gives us an automatic speed-up of 4-8. In addition, even in the preliminary implementation, Q-HOP MD shows less than 50 % computational overhead compared to standard MD.

The simulation of an imidazole ring in a water box was the first application of locally enhanced sampling (LES) for a proton transfer simulation. At each time step, the rates of proton transfer can be easily computed for the particular configuration of the system. And once the rates are known, it seems well justified to apply the well-known tricks developed for molecular simulations to speed up conformational sampling. Although we could already show the improved sampling, i.e. the facilitated proton release, when using LES in the imidazole-water simulation, the main advantage of LES may come when simulating systems such as proteins, that are not as dense as bulk solvent. Here, donor and acceptor may not regularly approach each other to 2.5 Å distance or closer like in bulk water. This is of course reflected by the much slower transfer rates in proteins. However, the LES methodology must be tested in more detail on small model systems like the ones presented, and its combination with a proper treatment of long-range electrostatics must also be tested.

Finally the method was used in Chapter 5 to simulate the proton relay in GFP. It is the first atomic simulation of a proton transfer over multiple steps where amino acids take part. The simulations were able to give a dynamic and atomistic model for the proton shuttle after excitation of the chromophore, as well as for the back shuttle to the starting ground state after the chromophore is relaxed due to light emission.

6.2 Possible Developments of Q-HOP MD

Further developments of Q-HOP MD can go into different directions. First of all, the significance of the method for simulating proton transfer reactions in solvent and in proteins has been demonstrated in this thesis. Thus one main aspect is to use the current sequential implementation and start further applications in the near future. For example, proton transfer steps in cytochrome c oxidase will be studied together with E. Olkhova and H. Michel. At the beginning, the protein will be simulated in vacuum to minimize the computational effort. Later, lipid bilayers and solvent should be included. However, to treat such protein in solvent and to reach longer time intervals, the method should be transferred to a modern MD package which is tailored for parallel architectures. This is planned to be done by T. Frigato who will continue the work presented here.

Another direction of development is to test accurately LES in combination with protein hopping for different systems. The aim is to find rules by which LES increases the transfer rate compared to simulations with donor and acceptor both single copied. This is necessary to correct the transfer rates in LES simulations to compare with experimental rates. A good test system might be the proton shuttle back to the neutral chromophore in GFP which is slow enough and seems to depend on conformational sampling. In this context, it would also be warranted to compare the approach of using the time-dependent Schrödinger equation for small energy barriers with other theoretical approaches, e.g. the work of Azzouz and Borgis [83]. This could put our approach on a stronger theoretical basis.

Additionally, polarization models going beyond the Thole's approach should be tested, whether they can compute the environmental influence on donating and accepting state more accurately than the pure electrostatic calculations based on available force fields can do. Especially for heterogeneous transfer systems or environments, this

could give a better estimate for the environmental stabilization of the proton during the transfer step.

Appendix A

Møller-Plesset Perturbation theory

The second order energy perturbation is (see Section 1.3.3)

$$W_2 = \sum_{i \neq 0} \frac{\langle \Phi_0 | \hat{H}' | \Phi_i \rangle \langle \Phi_i | \hat{H}' | \Phi_0 \rangle}{E_0 - E_i} \quad . \quad (\text{A.1})$$

where

$$\hat{H}' = \hat{H}_e - \hat{H}_0 = \hat{V}_{ee} - \sum_I (\hat{J}_I - \hat{K}_I) + V_{nn} \quad . \quad (\text{A.2})$$

Since \hat{H}' is a two electron operator, triple or higher excitations give zero matrix elements $\langle \Phi_0 | \hat{H}' | \Phi_i \rangle = 0 \quad \forall i$.

A.1 Single Excitations

In the following it will be shown that single excitations will also give zero matrix elements only.

$$\langle \Phi_0 | \hat{H}' | \Phi_i \rangle = \langle \Phi_0 | (\hat{H}_e - \hat{H}_0) | \Phi_i \rangle \quad (\text{A.3})$$

$$= \langle \Phi_0 | \hat{H}_e | \Phi_i \rangle - E_0 \underbrace{\langle \Phi_0 | \Phi_i \rangle}_{=0} \quad (\text{A.4})$$

Turning now to second quantization ($|\Phi_0\rangle \equiv |0\rangle$, $\hat{a}_a^\dagger \hat{a}_b |0\rangle = |a\bar{b}\rangle$), where

$$\hat{H}_e = \sum_{ab} h_{ab} \hat{a}_a^\dagger \hat{a}_b + \frac{1}{2} \sum_{abcd} V_{abcd} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_d \hat{a}_c \quad (\text{A.5})$$

gives for the single electron operator \hat{h}_{ab}

$$\sum_{ab} \sum_J^{\text{occ}} \sum_A^{\text{vir}} h_{ab} \langle 0 | \hat{a}_a^\dagger \underbrace{\hat{a}_b \hat{a}_A^\dagger}_{-\delta_{bA} - \hat{a}_A^\dagger \hat{a}_b} \hat{a}_J | 0 \rangle \stackrel{\langle 0 | \hat{a}_A^\dagger = 0}{=} - \sum_a \sum_J^{\text{occ}} \sum_A^{\text{vir}} h_{aA} \langle 0 | \hat{a}_a^\dagger \hat{a}_J | 0 \rangle \quad (\text{A.6})$$

$$= \sum_J^{\text{occ}} \sum_A^{\text{vir}} h_{JA} \quad (\text{A.7})$$

and for the two electron operator V_{abcd}

$$\frac{1}{2} \sum_{abcd} \sum_J^{\text{occ}} \sum_A^{\text{vir}} V_{abcd} \underbrace{\langle 0 | \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_d \hat{a}_c \hat{a}_A^\dagger \hat{a}_J | 0 \rangle}_{(\delta_{cA} \delta_{bJ} \delta_{ad} - \delta_{cA} \delta_{aJ} \delta_{bd}) - (\delta_{dA} \delta_{bJ} \delta_{ac} - \delta_{dA} \delta_{aJ} \delta_{bc})} = \sum_a \sum_J^{\text{occ}} \sum_A^{\text{vir}} (V_{JaaA} - V_{JaAa}) \quad (\text{A.8})$$

Taken together, this is exactly the Hartree-Fock Hamiltonian in the excited state

$$\langle 0 | \hat{H}_e | A\bar{J} \rangle = \langle \phi_J | \hat{H}_0 | \phi_A \rangle = 0 \quad (\text{A.9})$$

which is zero since all off-diagonal elements are zero.

A.2 Double Excitations

What remains are double excitations. Matrix elements of doubly excited states with the single electron operator \hat{h}_{ab} are obviously zero. What remains is

$$\langle 0 | \hat{H}_e | AB\bar{I}\bar{J} \rangle = \frac{1}{2} \sum_{abcd} \sum_{IJ}^{\text{occ}} \sum_{AB}^{\text{vir}} V_{abcd} \underbrace{\langle 0 | \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_d \hat{a}_c \hat{a}_A^\dagger \hat{a}_B^\dagger \hat{a}_I \hat{a}_J | 0 \rangle}_{(\delta_{dA} \delta_{cB} \delta_{bI} \delta_{aJ} - \delta_{dA} \delta_{cB} \delta_{aI} \delta_{bJ}) - (\delta_{cA} \delta_{dB} \delta_{bI} \delta_{aJ} - \delta_{cA} \delta_{dB} \delta_{aI} \delta_{bJ})} \quad (\text{A.10})$$

$$= \sum_{IJ}^{\text{occ}} \sum_{AB}^{\text{vir}} (V_{IJAB} - V_{IJBA}) \quad (\text{A.11})$$

$$= \langle IJ | \frac{1}{\rho_{12}} | AB \rangle - \langle IJ | \frac{1}{\rho_{12}} | BA \rangle \quad , \quad (\text{A.12})$$

which leads to Equation 1.45.

Appendix B

Danksagung

Die hier vorliegende Arbeit wurde am Max Planck Institut für Biophysik in Frankfurt in der Gruppe von Volkhard Helms durchgeführt. Ganz herzlich möchte ich mich bedanken bei Volkhard Helms für die Möglichkeit dieses interessante und aufregende Projekt bearbeiten zu können und die hervorragende und intensive Betreuung. Vor allem dafür dass man ihn ständig stören und nervige Fragen stellen konnte, sowie seine Geduld die abstrusesten Vorschläge und Ideen zu diskutieren. Die Ausstattung und die Räumlichkeiten in unserer Villa waren wunderbar. Auch die Gelegenheit einige nationale und auch internationale Konferenzen besuchen zu können habe ich sehr genossen.

Vielen Dank an Prof. Werner Mäntele für die begleitende Betreuung dieser Arbeit am Institut für Biophysik an der Johann Wolfgang Goethe Universität Frankfurt und seine Bereitschaft für mancherlei persönliche Diskussionen sowie die Möglichkeit die Arbeit im Rahmen eines Seminarvortrages in seiner Arbeitsgruppe präsentieren zu können.

Mein Dank gebührt auch insbesondere Michael Hutter, der mich in die Kunst der Quantenchemie eingeführt hat und auch später noch oft bei Fragen hierzu bereitwillig Rede und Antwort stand.

Herzlichen Dank auch an alle Mitglieder der Arbeitsgruppe, Dagmar Flöck, Hiqmet Kamberaj, Christian Gorba, mit dem ich viel Spass beim Aufbau unseres Linux-Clusters hatte, und Rene Staritzbichler für viele persönliche Gespräche und Diskussionen auch abseits der Forschung.

Vielen Dank auch an Tjerk Straatsma für dir Möglichkeit sein Programmpacket

ARGOS benutzen und für unsere Zwecke verändern zu können sowie für begleitenden Rat während der Implementierung.

Bedanken möchte ich mich weiterhin bei meinem Freund Jörg Marten für die zur Verfügungstellung seines Programmes zur numerischen Berechnung der zeitabhängigen Schrödinger-Gleichung.

Besonderer Dank gilt meinen Eltern und meiner Schwester, die mir auf dem Weg bis zur Dissertation mit Rat und Tat zur Seite standen.

Und ganz besonderen Dank meiner Verlobten Yoriko Iwanaga für den Rückhalt und ihre Unterstützung auch zu Zeiten, wenn es einmal nicht so gut lief und für ihr Verständnis, wenn es wieder einmal länger dauerte am Abend.

Bibliography

- [1] Protein Data Bank (<http://www.rcsb.org/pdb/>)
- [2] D. Voet and J. G. Voet, *Biochemistry* (John Wiley & Sons, Inc., New York, 1995), Chap. 14.
- [3] W. Mäntele, in *Bioanalytik* (Ed. F. Lottspeich and H. Zorbas), Chap. 7 (Spektrum, Akademischer Verlag, Heidelberg, 1998).
- [4] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [5] L. Greengard and V. Rokhlin, *J. Comp. Phys.* **73**, 325 (1987).
- [6] J. W. Eastwood and R. W. Hockney, *J. Comp. Phys.* **16**, 342 (1974).
- [7] L. Onsager, *J. Am. Chem. Soc.* **58**, 1486 (1936).
- [8] P. P. Ewald, *Ann. Phys.* **64**, 253 (1921).
- [9] J. A. McCammon and S. C. Harvey, *Dynamics of proteins and nucleic acids* (Cambridge University Press, Cambridge, 1987).
- [10] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1987).
- [11] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 1996).

- [12] A. R. Leach, *Molecular Modelling* (Addison Wesley Longman Limited, Edingburgh Fate, Harlow, 1996).
- [13] T. P. Straatsma and J. A. McCammon, *J. Comp. Chem.* **11**, 943 (1990).
- [14] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- [15] W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).
- [16] W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- [17] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover Publications, Inc., Mineola, New York, 1989).
- [18] F. Jensen, *Introduction to Computational Chemistry* (John Wiley & Sons, Chichester, 1999).
- [19] I. N. Levine, *Quantum Chemistry* (Prentice Hall, Inc., Upper Saddle River, New Jersey, 2000).
- [20] A. Warshel, G. Naray-Szabo, F. Sussman, and J.-K. Hwang, *Biochemistry* **28**, 3629 (1989).
- [21] D. Lu and G. A. Voth, *J. Am. Chem. Soc.* **120**, 4006 (1998).
- [22] A. Kohen, R. Cannio, S. Bartolucci, and J. P. Klinman, *Nature* **399**, 496 (1999).
- [23] M. C. Hutter and V. Helms, *Prot. Sci.* **8**, 2728 (1999).
- [24] M. C. Hutter and V. Helms, *Prot. Sci.* **9**, 2225 (2000).
- [25] R. Y. Tsien, *Annu. Rev. Biochemistry* **67**, 509 (1998).
- [26] H. Luecke, B. Schobert, B., H. T. Richter, J.-P. Cartailler, and J. K. Lanyi, *Science* **286**, 255 (1999).
- [27] S. Iwata, C. Ostermeier, B. Ludwig, B., and H. Michel, *Nature* **376**, 660 (1995).
- [28] H. Michel, *Biochemistry* **46**, 15129 (1999).

- [29] J. P. Abrahams, A. G. W. Leslie, R. Lutter, and J. E. Walker, *Nature* **370**, 621 (1994).
- [30] R. G. Schmidt and J. Brickmann, *Ber. Bunsenges. Phys. Chem.* **101**, 1816 (1997).
- [31] R. Vuilleumier and D. Borgis, *Chem. Phys. Lett.* **284**, 71 (1998).
- [32] U. W. Schmitt and G. A. Voth, *J. Phys. Chem. B* **102**, 5547 (1998).
- [33] U. W. Schmitt and G. A. Voth, *J. Chem. Phys.* **111**, 9361 (1999).
- [34] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, *Nature* **397**, 601 (1999).
- [35] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, *Science* **275**, 817 (1997).
- [36] S. R. Billeter and W. F. van Gunsteren, *Comput. Phys. Commun.* **107**, 61 (1997).
- [37] S. R. Billeter and W. F. van Gunsteren, *J. Phys. Chem. A* **102**, 4669 (1998).
- [38] S. R. Billeter and W. F. van Gunsteren, *J. Phys. Chem. A* **104**, 3276 (2000).
- [39] P. Bala, P. Grochowski, K. Nowinski, B. Lesyng, and J. A. McCammon, *Biophys. J.* **79**, 1253 (2000).
- [40] S. R. Billeter, S. P. Webb, I. Tzvetelin, P. K. Agarwal, and S. Hammes-Schiffer, *J. Chem. Phys.* **114**, 6925 (2001).
- [41] U. Rothlisberger, P. Carloni, K. Doclo, and M. Parrinello, *J. Biol. Inorg. Chem.* **5**, 236 (2000).
- [42] M. L. Brewer, U. W. Schmitt, and G. A. Voth, *Biophys. J.* **80**, 1691 (2001).
- [43] R. Pomes and B. Roux, *J. Phys. Chem.* **100**, 2519 (1996).
- [44] R. Elber and M. Karplus, *J. Am. Chem. Soc.* **112**, 9161 (1990).
- [45] A. Roitberg and R. Elber, *J. Chem. Phys.* **95**, 9277 (1991).
- [46] B. Das, V. Helms, V. Lonnas, and R. C. Wade, *J. Inorg. Biochem.* **81**, 121 (2000).

- [47] M. A. Lill, M. C. Hutter, and V. Helms, *J. Phys. Chem. A* **104**, 8283 (2000).
- [48] M. A. Lill and V. Helms, *J. Chem. Phys.* **114**, 1125 (2001).
- [49] M. A. Lill and V. Helms, accepted for *J. Chem. Phys.* .
- [50] M. A. Lill and V. Helms, accepted for *J. Chem. Phys.* .
- [51] M. A. Lill and V. Helms, submitted.
- [52] G.-S. Li, D. Maigret, D. Rinaldi, and F. Ruiz-López, *J. Comp. Chem.* **19**, 1675 (1998).
- [53] H. Umeyama, S. Hirono, and S. Nakagawa, *Proc. Natl. Acad. Sci. USA* **81**, 6266 (1984).
- [54] A. Kannt, C. R. D. Lancaster, and H. Michel, *Biophys. J.* **74**, 708 (1998).
- [55] M. A. Cunningham, L. L. Ho, D. T. Hguyen, R. E. Gillilan, and P. A. Bash, *Biochemistry* **36**, 4800 (1997).
- [56] High Performance Computational Chemistry Group, *NWChem, A Computational Chemistry Package for Parallel Computers, Version 3.3* (1998), Pacific Northwest National Laboratory, P.O. Box 999, Richland, Washington 99352 USA.
- [57] S. Sadhukhan, D. Muñoz, C. Adamo, and G. E. Scuseria, *Chem. Phys. Lett.* **306**, 83 (1999).
- [58] X. Duan, S. Scheiner, and R. Wang, *Int. J. Quant. Chem., Quantum Biol. Symp.* **20**, 77 (1993).
- [59] V. Barone, L. Orlandini, and C. Adamo, *Chem. Phys. Lett.* **231**, 295 (1994).
- [60] P. Ahlström, A. Wallqvist, S. Engström, and B. Jönson, *Mol. Phys.* **68**, 563 (1989).
- [61] M. Sprik, M. L. Klein, and K. Watanabe, *J. Phys. Chem.* **94**, 6483 (1990).
- [62] T. M. Nymand and P. Linse, *J. Chem. Phys.* **112**, 6386 (2000).
- [63] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and H. J. Hermans, In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981; p. 331.

- [64] D. Borgis, G. Tarjus, and H. Azzouz, *J. Chem. Phys.* **97**, 1390 (1992).
- [65] X. Duan and S. Scheiner, *Int. J. Quant. Chem., Quantum Biol. Symp.* **19**, 109 (1992) and citations [1-22] therein.
- [66] D. Antoniou and S. D. Schwartz, *Proc. Natl. Acad. Sci. USA* **94**, 12360 (1997).
- [67] S. Scheiner, *J. Am. Chem. Soc* **103**, 315 (1981).
- [68] P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).
- [69] J.-K. Hwang, G. King, S. Creighton, and A. Warshel, *J. Am. Chem. Soc.* **110**, 5297 (1988).
- [70] Y. Y. Sham, I. Muegge, and A. Warshel, *Proteins* **36**, 484 (1999).
- [71] A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York, 1991).
- [72] Y. S. Kong and A. Warshel, *J. Am. Chem. Soc.* **117**, 6234 (1995).
- [73] J.-K. Hwang and A. Warshel, *J. Am. Chem. Soc.* **118**, 11745 (1996).
- [74] M. Pavese and G. A. Voth, *Ber. Bunsenges. Phys. Chem.* **102**, 527 (1998).
- [75] M. Pavese, D. R. Berard, and G. A. Voth, *Chem. Phys. Lett.* **300**, 93 (1999).
- [76] R. I. Masel, *Chemical Kinetics and Catalysis* (Wiley, New York, 2001).
- [77] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, 1992).
- [78] J. Marten, Diplomarbeit, Friedrich-Alexander-Universität Erlangen-Nürnberg (1998).
- [79] I. Tuñón, C. Martins-Costa, C. Millot, and M. F. Ruiz-López, *J. Chem. Phys.* **106**, 3633 (1997).
- [80] E. Wigner, *Z. Phys. Chem. B* **19**, 203 (1932).
- [81] R. T. Skodje and D. G. Truhlar, *J. Phys. Chem.* **85**, 624 (1981).

- [82] Z. Luz and S. Meiboom, *J. Am. Chem. Soc.* **86**, 4768 (1964).
- [83] H. Azzouz and D. Borgis, *J. Chem. Phys.* **98**, 7361 (1993).
- [84] R. Triolo and A. H. Narten, *J. Chem. Phys.* **8**, 3624 (1975).
- [85] B. T. Thole, *Chem. Phys.* **59**, 341 (1981).
- [86] P. T. van Duijnen and M. Swart, *J. Phys. Chem.* **102**, 2399 (1998).
- [87] D. N. Bernardo, Y. Ding, K. Krogh-Jespersen, and R. M. Levy, *J. Phys. Chem.* **98**, 4180 (1994).
- [88] N. Agmon, *Chem. Phys. Lett.* **244**, 456 (1995).
- [89] N. Agmon, *Isr. J. Chem.* **39**, 493 (1999).
- [90] G. Hummer, D. M. Soumpasis and M. Neumann, *Mol. Phys.* **81**, 1155 (1994).
- [91] P. H. Hünenberger and J. A. McCammon, *J. Chem. Phys.* **110**, 1856 (1999).
- [92] G. Zundel, In *The Hydrogen Bond - Recent Developments in Theory and Experiment II*; Schuster, P., Zundel, G., Sandorfy, C., Eds.; North Holland, Amsterdam, 1976; p. 683.
- [93] M. Eigen, W. Kruse, and L. De Maeyer, *Prog. React. Kin.* **2**, 285 (1964).
- [94] C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.* **97**, 10269 (1993).
- [95] R. Heim, D. C. Prasher, and R. Y. Tsien, *Proc. Natl. Acad. Sci. USA* **91**, 12501 (1994).
- [96] W. W. Ward and S. H. Bokman, *Biochemistry* **21**, 4535 (1982).
- [97] H. Niwa, S. Inouye, T. Hirano, T. Matsuno, S. Kojima, M. Kubota, M. Ohashi, and F. I. Tsuji, *Proc. Natl. Acad. Sci. USA* **93**, 13617 (1993).
- [98] M. Chatteraj, B. A. King, G. U. Bublitz, and S. G. Boxer, *Proc. Natl. Acad. Sci.* **93**, 8362 (1996).

- [99] A. B. Cubitt, R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross, and R. Y. Tsien, *Trends Biochem. Sci.* **20**, 448 (1995).
- [100] B. G. Reid and G. C. Flynn, *Biochemistry* **36**, 6786 (1997).
- [101] F. Yang, L. G. Moss, and G. N. Phillips Jr., *Nat. Biotech.* **14**, 1246 (1996).
- [102] M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington, *Science* **273**, 1392 (1996).
- [103] V. Helms, T. P. Straatsma, and J. A. McCammon, *J. Phys. Chem. B* **103**, 3263 (1999).
- [104] H. Lossau, A. Kummer, R. Heinecke, F. Pollinger-Dammer, C. Kompa, G. Bieser, T. Jonsson, C. M. Silva, M. M. Yang, D. C. Youvan, and M. E. Michel-Beyerle, *Chem. Phys.* **213**, 1 (1996).
- [105] H. Yokoe and T. Meyer, *Nat. Biotechnol.* **14**, 1252 (1996).
- [106] K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö, and S. J. Remington, *Proc. Natl. Acad. Sci. USA* **94**, 2306 (1997).
- [107] T. M. H. Creemers, A. J. Lock, V. Subramaniam, T. M. Jovin, and S. Voelker, *Nat. Struct. Biol.* **6**, 557 (1999).
- [108] G.J. Palm, A. Zdanov, G. A. Gaitanaris, R. Stauber, G. N. Pavlakis, and A. Wlodawer, *Nat. Struct. Biol.* **4**, 361 (1997).
- [109] C. Scharnagl, R. Raupp-Kossmann, and S. F. Fischer, *Biophys. J.* **77**, 1839 (1999).
- [110] V. Helms, C. Winstead, and P. W. Langhoff, *J. Mol. Struct. (THEOCHEM)* **506**, 179 (2000).
- [111] C. Seebacher, F. W. Deeg, C. Bruchle, J. Wiehler, and B. Steipe, *J. Phys. Chem. B* **103**, 7728 (1999).

Lebenslauf



- 31.01.1974 Geboren in Frankfurt/Main
Vater: Dr. Günther Lill
Mutter: Christa Lill (geb. Schunk)
- Sep. 1980 - Juli 1984 Grundschule Herzogenaurach
- Sep. 1984 - Juli 1993 Gymnasium Herzogenaurach
- Nov. 1993 - Juli 1995 Vordiplom in Physik, Universität Erlangen-Nürnberg
- Okt. 1995 - März 1996 University of York, Großbritannien
- Mai 1996 - Juli 1998 Diplom in Physik, Universität Erlangen-Nürnberg
Lehrstuhl für Theoretische Physik II,
Prof. Dr. Christian Toepffer
Thema der Diplomarbeit: Molekulardynamische Simulationen von Fluktuationen elektrischer Felder in stark gekoppelten Plasmen
- Nov. 98 - Dez. 98 Max Planck Institut für Biochemie, Martinsried
Praktikum in Molekularbiologie und Kristallisation
- Jan. 99 - Doktorand am Max Planck Institut für Biophysik, Frankfurt
Arbeitsgruppe für theoretische Biophysik,
Dr. Volkhard Helms
Betreuer an der Johann-Wolfgang-Goethe Universität:
Prof. Dr. Werner Mäntele