Mixed quantum and classical simulation techniques for mapping electron transfer in proteins

Frank Wallrapp

DOCTORAL THESIS

THESIS DIRECTOR
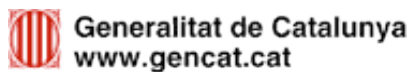
Dr. Victor Guallar

UNIVERSITAT
POMPEU FABRA

Department of Experimental and Health Services

Barcelona, 2011

Für Anna, Heidi und Oskar

Acknowledgements

First of all I would like to thank my thesis director Victor Guallar. Victor has always supported me throughout the whole PhD time. He taught me how to design and approach scientific questions, how to analyse results, how to present the research at international conferences and finally how to publish the work in peer-reviewed journals. Thank you very much Victor, you have been a great mentor to me.

Great thanks go to the University Pompeu Fabra and in particular my thesis tutor Jordi Villà i Freixa who made the shared PhD between the University Pompeu Fabra and the Barcelona Supercomputing Center possible.

I also would like to thank Alexander Voityuk, with whom I had a very fruitful collaboration and who introduced me to the exciting field of electron transfer.

A fellowship by the Generalitat de Catalunya (Ajut FI_B) as well as support by the Barcelona Supercomputing Center is acknowledged.

I also would like to thank my colleagues, who always supported me not only scientifically but also as friends. Special thanks goes to Fatima who was always open for any tricky question regarding quantum mechanics. I also want to thank Ben to whom it was always fun to talk about science, soccer or any other topic, cheers mate! Moreover, thanks goes to Israel, Ali, Suwipa, Ken, Diego, Max, Marcelo, Ryoji, Lil´ Victor, Manuel, Armin and all the other members of the BSC Life Sciences and Secretary. Winter retreats in Vall de Nuria and summer Beach Days have always been a lot of fun. Laia, Dimitry, Alexis and Romina, I enjoyed the constant battling with you guys about the air-con control. I know you guys are from Spain but 28 degrees in the office is just about too much :-)

I also would like to thank my friends back in Munich, Flo, Julia, Ben, Denise, Boris, Michi, Melli, Gerdi, Claudia, Alex, Matthias, Ruppi, Felix, Benni, Markus, Fabian, Alex, Doris, Matze, Flo. Thanks for all the escapes from PhD daily life to Ibiza, Cannes, Skiing in the Alps, Beergardening in Munich, Oktoberfest, etc. Furthermore, I would like to thank my friends here in Barcelona, Ben, Rebecca, Angi, Sonja, Michi, Bernd, James and all the others, for BBQ-parties, beach days, football and many other great memories.

Very special thanks go to my whole family, Corinna, Uli, Heidi and Oskar. Corinna and Uli, you are great as sister and brother. Dad, thank you for your tremendous support and infinite trust in me. Mom, thanks for your calming influence at home, your love and support.

Finally, I would like to thank the most important person in my life, Anna! Thank you for your help, support, patience and actually mere presence. You always believed in me and encouraged me to get over any disappointing stage within the last 3 years. Thanks to you, this time was not only PhD but also life and joy in Barcelona. I love you!

Abstract

The focus of this PhD thesis lies on electron transfer (ET) processes, belonging to the simplest but most crucial reactions in biochemistry. Getting direct information of the forces driving the process and the actual electron pathway is not a trivial task. Such atomic and electronic detailed information, however, is very valuable in terms of a better understanding of the enzymatic cycle, which might lead, for example, to more efficient protein inhibitor design.

The main objective of this thesis was the development of a methodology for the quantitative study of ET in biological systems. In this regard, we developed a novel approach to map long-range electron transfer pathways, called QM/MM e-Pathway. The method is based on a successive search for important ET residues in terms of modifying the QM region following the evolution of the spin density of the electron (hole) within a given transfer region. We proved the usefulness and applicability of the algorithm on the P450cam/Pdx complex, indicating the key role of Arg112 of P450cam and Asp48 of Pdx for its ET pathway, both being known to be important from the literature. Besides only identifying the ET pathways, we further quantified their importance in terms of electronic coupling of donor and acceptor incorporating the particular pathway residues. Within this regard, we performed two systematic evaluations of the underlying reasons for the influence of solvent and temperature onto electronic coupling in oligopeptide model systems. Both studies revealed that electronic coupling values strongly fluctuate throughout the molecular dynamics trajectories obtained, and the mechanism of electron transfer is affected by the conformational space the system is able to occupy. Combining both ET mapping and electronic coupling calculations, we finally investigated the electron transfer in the CcP/Cytc complex. Our findings indicate the key role of Trp191 being the bridge-localized state of the ET as well as the main pathway consisting of Ala194, Ala193, Gly192 and Trp191 between CcP and Cytc. Both findings were confirmed through the literature. Moreover, our calculations on several snapshots state a nongated ET mechanism in this protein complex.

The methodology developed along this thesis, mapping ET pathways together with their evaluation through electronic coupling calculations, suggests a straightforward and promising approach to investigate long-range ET in proteins.

Resumen

El objetivo de esta tesis se centra en el estudio de la transferencia de electrones (ET), una de las reacciones más simples y cruciales en bioquímica. Para dichos procesos, obtener información directa de los factores que lo promueves, asi como del camino de transferencia electronica, no es una tarea trivial. Dicha información a un nivel de conocimiento detallado atómico y electrónico, sin embargo, es muy valiosa en términos de una mejor comprensión del ciclo enzimático, que podría conducir, por ejemplo, a un diseño más eficaz de inhibidores.

El objetivo principal de esta tesis es el desarrollo de una metodología para el estudio cuantitativo de la ET en los sistemas biológicos. En este sentido, hemos desarrollado un nuevo método para obtener el camino de transferencia electrónico, llamado QM/MM e-Pathway, que se puede aplicar en sistemas complejos con ET de largo alcance. El método se basa en una búsqueda sucesiva de residuos importantes para la ET, utilizando la modificación de la región quantica en métodos mixtos QM/MM, y siguiendo la evolución de la densidad de espín dentro de la zona de transferencia. Hemos demostrado la utilidad y la aplicabilidad del algoritmo en el complejo P450cam/Pdx, identificando el papel clave de la Arg112 (en P450cam) y del Asp48 (en Pdx), ambos conocidos en la literatura. Además de obtener caminos de ET, hemos cuantificado su importancia en términos del acoplamiento electrónico entre el dador y aceptor para los diferentes caminos. En este sentido, se realizaron dos estudios de la influencia del solvente y de la temperatura en el acoplamiento electrónico para sistemas modelo oligopéptidos. Ambos estudios revelaron que los valores del acoplamiento electrónico fluctúan fuertemente a lo largo de las trayectorias de dinámica molecular obtenidas, y el mecanismo de transferencia de electrones se ve ampliamente afectado por el espacio conformacional del sistema.

La combinación del QM/MM e-pathway y de los cálculos de acoplamiento electronico fueron utilizados finalmente para investigar la ET en el complejo CCP/Cytc. Nuestros hallazgos indican el papel fundamental del Trp191 en localizar un estadio intermedio para la transferencia electronica, así como el camino ET principal que incluye Ala194, Ala193, Gly192 y Trp191. Ambos hallazgos fueron confirmados a través de la literatura.

Los resultados obtenidos para el muestro de manios de ET, junto con su evaluación a través de cálculos de acoplamiento electrónico, sugieren un enfoque sencillo y prometedor para investigar ET de largo alcance en proteínas.

Preface

It is in the human's nature to want to know and understand life in its very detail. Consequently, researchers have made precise observation and more and more complex investigations on biological and biochemical problems since the very beginning. Derived from observations, theoreticians have come up with rules and laws underlying the mechanisms of these processes and started to construct and train theoretical approaches in order to predict a priori their experimental outcomes. In the last few years, not only the interest for but also the application of these theoretical approaches in biological and biochemical sciences has increased significantly. Reasons for this are manifold. On the one hand, nowadays there exist tremendous amounts of experimental data gathered for many different systems and under different experimental conditions. Furthermore, the data is much more rigorous now, resulting from experimental techniques being directly related to theory (i.e. single molecule experiments). On the other hand, computational power has become very potent, as computers have approximately doubled their power every 2 years following Moore's law since its very beginning in the 1970ies. Theoretical investigations do not only provide a rigorous test on the knowledge of the underlying mechanisms or help experimentalists to interpret their results but also to design new experiments. Computational approaches are also much cheaper, faster and in the majority of the cases less intricate than *in-vitro* and especially *in-vivo* experiments. Furthermore, they are able to give chronological and atomistic resolution often unreached by experimental trials. Of course, computational approaches also have a negative side. Most often, the complexity of the system, respectively the possible lack of the complete knowledge only allows an estimation of the process under investigation. Thus, development and training of theoretical methodologies being capable to predict a priori experimental outcomes will keep a big challenge for the future.

The particular focus of this PhD thesis lies on electron transfer processes, being one of the simplest but most important reactions in biochemistry. The exploration of these ubiquitous processes constitutes an interdisciplinary research area, incorporating concepts and experimental techniques from a wide variety of fields. Here, important recent advances are based on concurrent progress in experiment and theory. Referring to the experimental side, strong progress was made, for example, with the introduction of femtosecond lasers allowing the study of electron transfer dynamics on the time scale of nuclear motion. Theoretical advances include, among others, the groundbreaking introduction of Marcus theory incorporating electronic and nuclear quantum effects. Throughout this PhD thesis, we purely focused on the theoretical assessment of electron transfer processes in protein through application of well-established methods and development of novel approaches.

The first chapter of this dissertation will give an introduction to important as well as recent achievements in theoretical electron transfer chemistry. All

computational methods applied within the underlying studies, especially our newly developed electron transfer mapping technique, will be described. Furthermore, we depict the systems on which we conducted our experiments. The second chapter will give the objectives that were addressed by this dissertation, and the third chapter presents the actual research carried out throughout this thesis. A critical discussion of all results is given in the fourth chapter, while the complete list of all conclusions inferred from the results is given in the fifth chapter. Finally, the list of publications resulting from the research done during this PhD thesis is provided.

Table of contents

1. INTRODUCTION

INTRODUCTION

1.1.   Electron transfer in biochemistry

Electron transfer (ET) processes belong to the simplest but most crucial reactions in biochemistry. These processes, also called redox processes, are involved in many bio-energetic reactions (respiration, photosynthesis, nitrate respiration), in the fixation of molecular nitrogen, in the repair of damaged DNA, in many detoxification processes, and in plenty other metabolic processes. In a nutshell, ET processes are involved in almost all enzymatic cycles (Marcus and Sutin 1985; Newton 1991; Beratan, Onuchic et al. 1992; Gray and Winkler 1996; Balzani, Piotrowiak et al. 2001). While the phenomenon of ET between small molecules in solution is reasonably well described (Marcus and Sutin 1985), the factors that influence protein ET reactions are less understood [Davidson 2000]. Besides fundamental research, investigation of electron transfer in proteins has also impact on medical science (Shao, O'Neill et al. 2004; Reisner, Arion et al. 2008), on improvement of agricultural technology (Dayan, Vincent et al. 2000; Barbosa, Rocha et al. 2007; Moraes, Chaves et al. 2010), on development of molecular electronic devices (Whalley, Steigerwald et al. 2007; Feldman, Steigerwald et al. 2008), and on the usage of alternative energy sources (Lovley 2006; Liu, Wang et al. 2008), among others. Consequently, much effort is spent to understand the detailed mechanism underlying these reactions in and between proteins, including the research done throughout this PhD thesis.

1.1.1. Theory

First modern attempts of electron transfer investigation were done with Fe^{3+*} + $Fe^{2+} \rightarrow Fe^{2+*} + Fe^{3+}$ in solution after the World War II. As the standard free energy ΔG^0 vanishes for self-exchange reactions, researchers were led to study other factors influencing the electron transfer rate. Consequently, starting from 1952, when Libby firstly applied the Franck-Condon principle to explain his data (Libby 1952), new formulations of electron transfer theory emerged, ranging from classical (Marcus, Zwolinski et al. 1954), over quantum (Levich and Dogonadze 1960), to current mixed classical-quantum (Kestner, Logan et al. 1974; Marcus and Sutin 1985) formulations, where classical denotes for low-frequency and quantum denotes for high-frequency motion.

From a general point of view, an electron transfer reaction occurs to be simple: An electron moves from the donor (D) to the acceptor (A) site, without any breaking or formation of chemical bonds:

$$D^- + A \xrightarrow{\;k_{ET}\;} D + A^- , \tag{1}$$

where, k_{ET} denotes the rate constant of the ET reaction, quantifying the speed of the reaction to occur. In the classical model, ET occurs at the intersection of the potential energy surfaces of the initial state (electron is localized in the donor) and final state (electron has been transferred to the acceptor). For simplicity, these multidimensional energy surfaces are typically represented as parabolas described

by the free energy (y-axis) and reaction coordinate (x-axis), see Figure 1. The reorganization energy λ is the energy difference between the initial and final states at the equilibrium nuclear configuration of the initial state (i.e. its minimum). When $\Delta G°$, the overall Gibbs free energy change of the electron transfer reaction, is equal to zero the activation energy becomes $\lambda/4$ in the parabolic case. V_{DA} is the electronic coupling matrix element that represents the extent to which the wave functions of the initial and final states overlap. The splitting at the intersection point is equal to $2V_{DA}$. This describes the degree on non-adiabicity (i.e. probability of the reaction to occur when the activation energy has been achieved). Thus, when V_{DA} is zero (a diabatic system), there is no chance of the reaction occurring regardless of the energy. At the other extreme, namely the adiabatic system, the value of V_{DA} is so large that the crossover is certain to happen when the activation energy is achieved.

In biochemistry many ET reactions only have weak electronic coupling between donor and acceptor sites, thus are non-adiabatic. In electron transfer theory an estimate of the rate constant k_{ET} for non-adiabatic transfer of an electron from a donor to acceptor can be expressed as (Marcus and Sutin 1985):

$$k_{ET} = \frac{2\pi}{\hbar}\left|V_{DA}\right|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(-\frac{\left(\Delta G° + \lambda\right)^2}{4\lambda k_B T}\right), \tag{2}$$

where $\hbar$ is Planck´s constant, k_B is Boltzmann´s constant and T is the temperature.
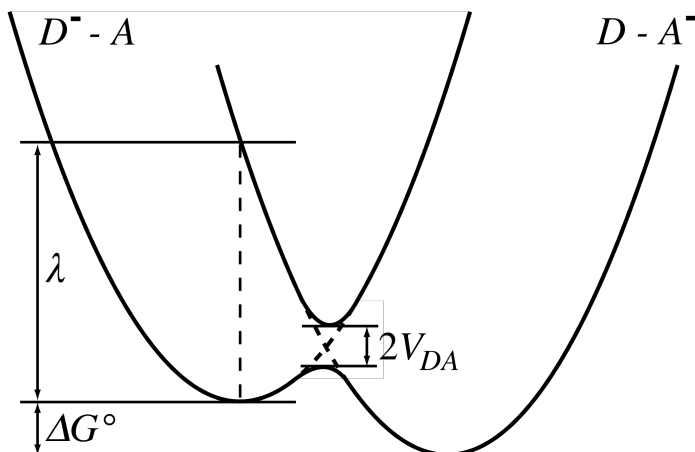
Figure 1: Diagram of key parameters of electron transfer in Marcus theory. During the reaction, the system transits from the higher potential energy surface of the initial state to the lower potential energy surface of the final state. The rate of the reaction is influenced by the reorganization energy λ and Gibbs free energy $\Delta G°$ as well as by the electronic coupling V_{DA}.

It has been shown that the parameters in equation (2) are accessible to theoretical calculations. While the reaction free energy of chemical events can usually be

obtained by umbrella sampling (Torrie and Valleau 1977), thermodynamic integration (Carter, Ciccotti et al. 1989; Den Otter and Briels 2000; Darve, Wilson et al. 2002; Fleurat-Lessard and Ziegler 2005), metadynamics (Laio and Parrinello 2002; Iannuzzi, Laio et al. 2003), adiabatic molecular dynamics (Rosso, Mináry et al. 2002), transition path sampling (Dellago, Bolhuis et al. 2003) or transition interface sampling (van Erp, Moroni et al. 2003), these approaches bare some difficulties to apply for ET reactions as very difficult to move electrons along a possible reaction coordinate respectively apply quantum mechanical molecular dynamics. A possible workaround is the approach to estimate $\Delta G°$ through ΔE by neglecting the entropic effects. Here, ΔE is given by the energy difference of the electronic states derived from quantum chemical minimizations (Mouesca, Chen et al. 1994), for example. This scheme can be extended by incorporating harmonic entropic corrections in terms of solving the Hessian in each of the minima. Several approaches exist to access the reorganization energy λ of ET reactions, including molecular dynamics simulations (Miyashita and Go 2000) and density matrix methods (Okada, Chernyak et al. 1998).

The theoretical estimation of the electron coupling V_{DA} turns out to be a challenge on its own. Within the past 20 years, many researchers have spent great efforts to come up with approaches ranging from simple pathway models (Onuchic, Beratan et al. 1992), over more sophisticated methods applying Green's function (Skourtis, Beratan et al. 1993; Skourtis and Onuchic 1993) and extended Hückel theory (Larsson 1981; Siddarth and Marcus 1993; Stuchebrukhov and Marcus 1995) to more recent Generalized Mulliken-Hush (Cave and Newton 1996; Cave and Newton 1997) and Fragment Charge Difference (Voityuk and Rösch 2002) and also strict quantum mechanical methods (Crystal, Zhang et al. 2002; Improta, Barone et al. 2006). Furthermore, there are studies applying quantum mechanical dynamics on photochemical processes in small molecules (Moret, Tapavicza et al. 2005; Moret, Tavernelli et al. 2010), but application to full biologically relevant proteins is still missing.

Within this thesis the focus lies onto electronic coupling calculations based on the well-established Generalized Mulliken-Hush and the Fragment Charge Difference methods, which are going to be explained in detail in section 1.2.2.

1.1.2. Electron transfer mechanisms

Regarding electron transfer mechanisms we have to consider two possible scenarios. The donor and acceptor sites can either be localized next to each other, for example when a substrate is bound in close vicinity of an acceptor group, or they can be at much larger distance when they are part of two separate proteins. Thus, in the latter case the electron transfer involves a molecular bridge (B) through which the electron has to travel on its way from the donor to the acceptor. Throughout this thesis we only focus on ET in proteins, generally having donor and acceptor distances of less than 30 Å. Nevertheless, in the particular case of double-helical DNA systems, ET has already been observed over 100 Å (Takada, Kawai et al. 2004).

INTRODUCTION

Extensive research on intra-molecular electron transfer between donor and acceptor through bridges of different medium and structure have disclosed two distinct mechanisms of the process. The first mechanism is the bridge-mediated superexchange (Jortner and Bixon 1999; Kuznetsov and Ulstrup 1999), which occurs if the bridge is much higher in energy as compared to the donor and acceptor. Here, the electron transfer rate, k_{ET}, decreases with the distance between the donor and acceptor, R, as

$$k_{ET} = k_0 \exp(-\beta R),$$ (3)

where k_0 is the pre-exponential factor and β is the falloff parameter. The value of parameter β strongly depends on the structure and medium of the molecular bridge. For example, highly conjugated organic bridges have the smallest falloff parameters ranging from 0.2 up to 0.6 Å^{-1} (Helms, Heiler et al. 1992; Sachs, Dudek et al. 1997; Davis, Svec et al. 1998), while free space is characterized by the largest value β of about 2 Å^{-1} (Newton 1991). In between are many synthetic as well as natural patterns including DNA, cytochromes, docked proteins and saturated organic molecules. At this point it is worth to mention that in the case of bridge-mediating superexchange electron transfer the bridge remains physically unpopulated during the entire process.

The situation changes when the bridge B becomes comparable in energy with the donor and acceptor. Here, the electron can be injected from the donor to the bridge and hence, the intermediate state where the electron localized into the bridge becomes physically occupied. Now, the overall electron transfer occurs through a process of incoherent hopping between localized electronic states within the bridge (Berlin and Ratner 2005). As a consequence, the distance dependence of the ET rate for bridge-mediated superexchange as formulated in equation (3) does not count any more. Theoretical calculations performed on long bridges with a large number N of energetically identical hopping sites separated by the distance a have shown that in this particular case the dependence k_{ET} vs. R can be approximated by

$$k_{ET} \propto N^{-\eta} = (R/a)^{-\eta},$$ (4)

with the numerical parameter η ranging from 1 (irreversible hopping along bridge) to 2 (unbiased hopping between bridge states) (Berlin, Burin et al. 2000; Wang and Nau 2001; Petrov, Shevchenko et al. 2003). These findings were also confirmed by experiments (Davis, Svec et al. 1998; Giese, Wessely et al. 1999; Malak, Gao et al. 2004). The respective contribution of the two mechanisms is dependent on the energy difference between the donor- and bridge-localized electronic states. The total rate constant of the system is then the composite of both bridge-mediated superexchange and sequential hopping mechanism. Experimental work on oligoproline spacers between ruthenium based donors and acceptors has shown that electron transfer can also change from a predominantly electron superexchange to a predominantly electron hopping mechanism when the peptide spacer distance exceeds about 20 Å (Malak, Gao et al. 2004). On the contrary, a more recent

experimental study on ET through conformationally constrained oligopeptides showed strong evidence against the hopping mechanism in this system (Polo, Antonello et al. 2005). Evidently, the actual regime of ET (superexchange versus hopping) depends on the distance as well as the nature of the bridge, respectively the energy of the bridge states. At this point, theoretical calculations seem to be a promising tool to investigate the particular ET mechanism of a given system.

Except within the case of direct electron transfer, there exist two alternatives of the electron travelling from the donor to the acceptor. Firstly, the excess electron transfer (EET), where the electron travels from the donor to the acceptor through the lowest unoccupied molecular orbitals (LUMOs) of possible bridge sites. Secondly, the hole transfer (HT), where the acceptor pulls one electron from the donor through the highest occupied molecular orbital (HOMO) of possible bridge sites. A scheme illustrating both ET mechanisms is shown in Figure 2.
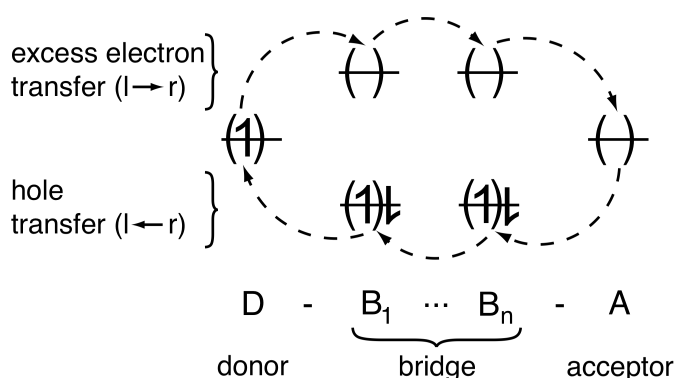
Figure 2: Schematic orbital representation of bridge-mediated superexchange excess electron transfer (top, left to right) and hole transfer (bottom, right to left). Figure adapted from (Newton 1991).

1.1.3. Conformational dynamics

One of the main parameter influencing the electronic coupling between the donor and acceptor is the mere distance between them. The reason for this is that the electronic coupling is proportional to the overlap of the two diabatic states, which itself has a roughly exponential decay with increasing distance (Ungar, Newton et al. 1999). Variations of this distance dependence have been the subject of a dynamic discussion between experimentalists (Isied, Ogawa et al. 1992; Ogawa, Wishart et al. 1993; Bixon, Giese et al. 1999; Malak, Gao et al. 2004) and theoreticians (Felts, Pollard et al. 1995; Petrov and May 2001) in the last years and is still ongoing. Many of these studies have focused on test systems of single-chained short oligopeptides to get an easier understanding of the underlying ET mechanism.

Next to the distance, dynamical flexibility of protein and solvent has a strong impact on the electronic coupling between donor and acceptor. Recently there have been several studies describing how conformational dynamics of proteins strongly

influence the donor-acceptor coupling (Wolfgang, Risser et al. 1997; Miller, Wander et al. 1999; Ungar, Newton et al. 1999; Xie, Archontis et al. 1999; Balabin and Onuchic 2000; Castner, Kennedy et al. 2000; Newton 2000; Skourtis, Archontis et al. 2001; Kawatsu, Kakitani et al. 2002; Balabin, Beratan et al. 2008). These studies show that the observed fluctuations can be large and sometimes only within the femtosecond time scale. Trying to identify the exact influences of fluctuation on the electronic coupling, a recent study on through-bond electron transfer in Ruthenium-modified azurins emphasized the central role of valence angle fluctuations in coupling dephasing (Skourtis, Balabin et al. 2005).

Very recently the group of Ursula Rothlisberger, among others, started to apply quantum mechanical dynamics to photochemical processes (Moret, Tapavicza et al. 2005; Moret, Tavernelli et al. 2010). Within their research they also focused on the dynamical interaction between the electronic and nuclear motion beyond the Born-Oppenheimer approximation (see section 1.2.1.3) (Tavernelli, Tapavicza et al. 2009). They found out that this interaction can be crucial for reactions that occur at the intersection of two or more potential energy surfaces, which is the case for ET reactions. Although their approach is computationally very expensive and hence not applicable to biologically relevant proteins, it is very worthwhile to be kept in mind for future studies.

Next to the above presented trials on dynamical flexibility there have been published experimental as well as theoretical studies on the temperature dependence of ET rates (Read, Napper et al. 1999; Davis, Ratner et al. 2001; Huppman, Arlt et al. 2002; Napper, Read et al. 2002; Jean and Krueger 2006; Eng, Martensson et al. 2008; Eng and Albinsson 2009) and electronic conductance through molecular wires (Selzer, Cabassi et al. 2004; Haiss, Zalinge et al. 2006; Poot, Osorio et al. 2006). Eng et al. showed that the electronic coupling between bridge-mediated donor and acceptor depends on the temperature due to the different conformations available at different temperatures (Eng, Martensson et al. 2008; Eng and Albinsson 2009). If donor and acceptor have a strong coupling in the energetically most favorable conformations, an increase of temperature will result in a decrease of the average coupling as the system will explore conformations with lower coupling values. On the other hand, if in the energetically most favorable conformations the couplings are weak, a temperature increase will populate conformations of higher as well as lower coupling values. Such systems are predicted to be less sensitive to temperature. Oligo p-phenyleneethynylene and oligofluorene bridges exemplify the first and the second regimes, respectively (Eng and Albinsson 2009). There also exists a third regime, where the most favorable conformation has a weak coupling. In this case the average coupling will increase with the temperature as demonstrated by fluorescence measurements on deoxytrinucleotides done by Jean and Krueger (Jean and Krueger 2006). Additionally there exist systems where the electronic coupling shows more complex temperature dependence, which is the case when the energy minimum does not correspond to a minimum (respectively maximum) of the electronic coupling. Examples are theoretical electron transfer calculations on oligo thiophene having maximal electronic coupling at about 25 K done by Eng and Albinsson (Eng and Albinsson 2009) and experimental charge separation studies on

2-phenylenevinylenes with a maximal rate constant at about 210 K done by Davis et al. (Davis, Ratner et al. 2001). In general we can say that electronic coupling becomes temperature dependent when structural dynamics is incorporated (Davis, Ratner et al. 2001) and the dependence regime strongly relies on the system itself.

Closely related to the influence of dynamic flexibility on the electronic coupling is the question if electron transfer occurs through direct coupling, mediated coupling through the molecular bridge or coupling through intermediary solvent in the case of crooked systems. There exist several studies on these so-called U-shape electron transfer systems, where a highly curved bridge imparts a vacant cleft along the line-of-sight between the electron donor and acceptor (Zimmt and Waldeck 2003; Troisi, Ratner et al. 2004; Lu, Li et al. 2005). Semiempirical calculations show that electronic coupling between donor and acceptor within these model systems results mainly from direct coupling or coupling through bridging solvent molecules. This feature differs strongly from craned model systems in which the coupling is mainly due to through-bond contributions resulting in a linear decrease of direct coupling with increasing donor-acceptor distance (Cave and Newton 1996; Cave and Newton 1997; Shin, Newton et al. 2003).

Most of the ET studies have in common that they are either based on model systems, which are rigid or on single conformations of the systems. Hence, they only represent the characteristics of natural protein to some extent as it is known from the literature that electronic coupling can strongly depend on the conformation of the system (Hoffman and Ratner 1987). In the so-called conformationally gated electron transfer the charge is not gradually transferred from the donor to the acceptor, instead there is a sudden jump enhanced by favorable conformations with high electronic coupling between the donor and acceptor. For example, a very recent study done by Hartings et al. (Hartings, Kurnikov et al. 2010) investigated long-range ET through sensitizer wire bound in the active-site channel of cytochrome P450cam. They found out that the electron tunneling rates are conformationally gated and strongly depend on single wire conformations with high electronic coupling. Consequently, despite recent efforts, there still is a need of research on ET regimes in flexible model systems, which are able to reproduce the dynamical behavior of proteins. In order to do so, it has been shown that it is necessary to calculate the average of the electronic coupling over many snapshots of a long molecular dynamics trajectory to cover full conformational space of flexible systems (Wolfgang, Risser et al. 1997; Hartings, Kurnikov et al. 2010).

1.1.4. Electron transfer pathways

In the previous section we already mentioned the question if electrons prefer to travel through the molecular bridge or through the solvent in case of U-shaped systems. Within this section we will discuss the influence of the three-dimensional scaffold of protein on the electronic coupling and its prediction with computational approaches. Within this context, we define the electron tunneling path as the trace of connected covalent bonds, hydrogen bonds, and van der Waals contacts (interactions through space), that links the donor with the acceptor. The usual experimental procedure to determine the electron transfer pathway is by means of

direct protein mutation. While this technique can probe the importance of some residues in the electron migration, it furthermore has the great disadvantage of possible secondary effects caused by the mutation, such as structural changes or significant perturbation of the protein electrostatics, which might mask the real cause for the loss of the electron conductivity. Consequently, much effort is spent to identify ET pathways by means of theoretical approaches.

Summarized in a comprehensive review, Onuchic et al. published in 1992 one of the first methods to computationally identify pathways of protein-mediated ET, referred to as the *Pathways* method (Beratan, Onuchic et al. 1990; Beratan, Onuchic et al. 1992; Onuchic, Beratan et al. 1992). The underlying idea is that electronic coupling between donor and acceptor decays with increasing distance depending on the actual ET pathway between them. The parameters for the decay of the electronic coupling via the trace of bonds forming the ET pathway are calculated separately incorporating the distance between the interacting atoms. Coupling within aromatic rings and within the guanidinium group of arginine is defined as none-decaying. The relative electronic coupling for a given pathway is then the product of all single coupling decay parameters along the path. The path providing the strongest coupling between donor and acceptor is searched by a graph theoretical algorithm called depth-first (Betts, Beratan et al. 1992). The authors applied the method on Ruthenium-modified proteins showing that the approach is able to map ET pathways in protein and also allows computing global coupling maps from all atoms on the protein surface to the redox centre. Such maps reveal important characteristic effects of primary, secondary, tertiary, and quaternary folded structure on the nature of the coupling to a specific patch of the protein as well as possible docking sites in case of protein-protein ET (Nocek, Zhou et al. 1996). A version of the *Pathways* method is implemented into a molecular simulation package called HARLEM by Igor Kurnikov (Kurnikov 1999).

The *Pathways* method has the disadvantage that it has to calculate the decay parameter for every possible bond in the protein before computing the best ET pathway between the donor and acceptor. A possible workaround was introduced by Siddarth and Marcus in 1993, where important residues for the ET are selected by and artificial intelligence (AI) approach (Siddarth and Marcus 1993). Here, a first tier of residues (say a set of nodes i) is selected by the product of both approximated electronic couplings between donor and node i and between node i and acceptor, respectively. The coupling between donor and node i is calculated iteratively by means of the product of the overlap integrals of all adjacent atoms along the path. The coupling between node i and the acceptor is approximated by means of an exponential decay of a predefined constant dependent on the distance between node i and the acceptor. With a second step, only pathways between donor and acceptor having high approximate electronic coupling are fully calculated applying the product of the overlap integrals and finally the residues along the pathway with the maximum electronic coupling are returned. Once the important residues are selected by the AI search, they collectively are considered to be the bridge for mediating ET between the donor and acceptor. The approach then diagonalizes the bridge orbitals and applies second-order perturbation theory to

calculate the electronic coupling matrix element according to the superexchange mechanism. Calculations on Ruthenium-modified cytochrome c derivates identified important residues for mediating ET and showed that the calculated electronic coupling values were consistent within one order of magnitude to experimental data.

Although the above explained methods are capable of suggesting the ET pathway and approximating the electronic coupling between the donor and acceptor to some extent, the question arises if it was possible to apply more accurate methods nowadays. Current increase in computational power together with the recent development of mixed quantum mechanics/molecular mechanics algorithms allow us to say that we actually can.

1.2. Computational methods

Computational modeling has developed considerably since the early beginning in the 1970s and 1980s in both scopes, software (methods) as well as hardware (computer power). Computational approaches are becoming significantly faster and more accurate in predicting enzymatic mechanisms and rates, binding energies, docked structures, complex network pathways, gene annotations, etc. As a consequence, the level of confidence towards computational modeling is growing among scientists, and nowadays many experimental laboratories work in close collaboration with computational modelers. Moreover, some experimentalists themselves routinely apply computational modeling tools to analyze their results and to obtain a detailed view of the mechanism in atomic resolution unreached by experimental trials.

The following section gives brief introductions to molecular modeling in terms of molecular mechanics and quantum mechanics. Furthermore, it introduces state of the art methods to compute the electronic coupling as the precursor of the ET rate constants of biological systems. Finally we depict the data analysis tools we applied through out this thesis.

1.2.1. Molecular Modeling

The currently accepted definition of molecular modeling can be stated as this: "Molecular modeling is anything that requires the use of a computer to paint, describe or evaluate any aspect of the properties of the structure of a molecule" (Pensak 1989). Methods used in the molecular modeling area regard automatic structure generation, analysis of 3D databases, construction of protein models by techniques based on sequence homology, diversity analysis and docking of ligands as well as proteins, for example. Consequently, today molecular modeling can be seen as a field concerned with the use of all sorts of different strategies to model and to deduce information of a system at the atomic level. Furthermore, this discipline includes all methodologies used in biophysics, like computation of the energy of a molecular system, energy minimization, side chain or loop predictions, or molecular dynamics.

There are mainly two different approaches to compute properties of a molecule. First, by classical molecular mechanics, a procedure based on Newton's law describing the system by only the set of nuclei and their pair wise interactions. Second, by quantum mechanics, a procedure based on first principles where nuclei are arranged in the space and the corresponding electrons are spread all over the system in a continuous electronic density and computed by solving the Schrödinger equation. Both of them are going to be briefly explained in the following.

1.2.1.1. Classical Molecular Mechanics

The "mechanical" molecular model was developed out of a need to describe molecular structures and properties in a manner as practical as possible. The range of applicability of molecular mechanics nowadays includes molecules containing

thousands and even millions of atoms (i.e. organics, oligonucleotides, peptides and saccharides or metallo-organics and inorganics in some cases). Here, one example the larger systems is the all atom molecular dynamics simulation of the complete satellite tobacco mosaic virus with up to 1 million atoms (Freddolino, Arkhipov et al. 2006). Simulations can be done in vacuum, implicit or explicit solvent environments. By molecular mechanics, we are able to compute how the structure of our system changes in time, as well as time dependent properties like diffusion or thermodynamic properties like energy, heat capacity or pressure for example. Furthermore, the great computational speed of molecular mechanics allows its use in procedures requiring large numbers of energy evaluations such as conformational energy search or docking for example.

Molecular mechanics methods are based on the following principles. First, nuclei and electrons are merged into atom-like particles, which are spherical, with radii and net charges obtained from measurements or theory. Second, interactions are based on springs and classical potentials being pre-assigned to specific sets of atoms. These interactions then determine the spatial distribution of atom-like particles and their energies. The collection of constants defining the spheres and interaction potentials is called force field.

Force field

The force field allows us to calculate the potential energy E of the system and is generally based on a pair wise sum over all the different sphere pairs present in the system. For example, the functional form of the widely used Assisted Model Building with Energy Refinement (AMBER) force field has the following expression:

$$
\begin{aligned}
E = &\sum_{bonds} \frac{1}{2} k_{r_0} \left(r - r_0 \right)^2 + \sum_{angles} \frac{1}{2} k_{\theta_0} \left(\theta - \theta_0 \right)^2 + \sum_{dihedrals} \frac{1}{2} V_n \left[1 + \cos\left(n\phi - \gamma \right) \right] \\
&+ \sum_{\substack{non-bonded \\ pairs\ i,j}} \varepsilon_{i,j} \left\{ \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}
\end{aligned}
\qquad , \qquad (5)
$$

where the first term defines the potential for bonded atom pairs, the second the potential for bonded angles, the third the potential for bonded dihedral angles and the final term the potential for non-bonded atom pairs, made up of a diffusive interaction and a Coulombic term (see Figure 3 for illustration). The equilibrium values of the bond lengths r_0 and bond angles θ_0 as well as the corresponding force constants used in the potential energy function are defined in the force field and are generally derived from averaged observations (i.e. crystal structures) respectively quantum chemical calculations. Consequently, there exist specific force fields parameters for specific systems, i.e. protein or DNA. The total energy is a measure of intramolecular strain relative to the hypothetical molecule with an ideal geometry of equilibrium. By itself, the total energy has no strict physical meaning, but

differences in total energy between two different conformations of the same molecule can be compared.
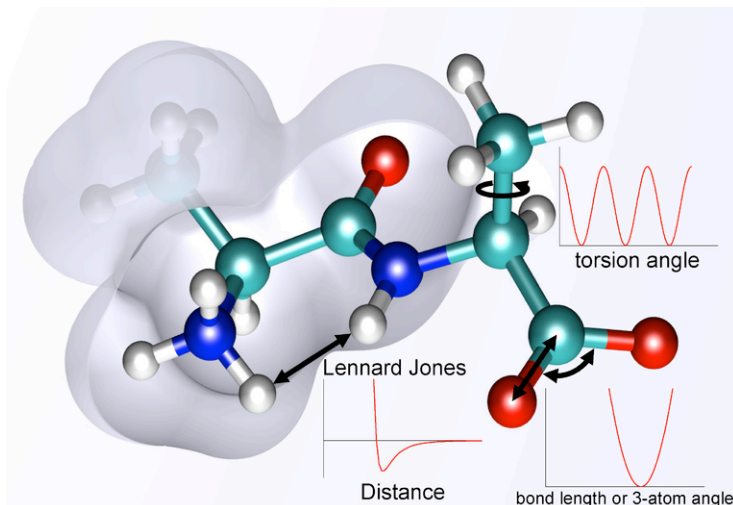

Figure 3: Schematic view of the molecular mechanics potential energy function. Figure adapted from (www.wikipedia.org).

Molecular dynamics

In general it can be said that molecular dynamics (MD) is concerned with time dependent processes in molecular systems. Here, each dynamic process (i.e. motion) has a characteristic time-scale, amplitude and energy range. Time-scales range between femtoseconds and picoseconds for local motions, like atomic fluctuations or side chain motions, and can go up to milliseconds or even hours for global motions like helix-coil transitions of protein folding/unfolding.

Within classical molecular mechanics approaches the well-known Newton's equation (*force = mass × acceleration*) is used in the molecular dynamics formalism to simulate atomic motion:

$$F_i = m_i a_i = -\nabla_i E(q), \tag{6}$$

where the force F_i is given as the derivate of the potential E. The rate and direction of motion (velocity) are determined by the forces that the atoms of the system exert on each other as described by Newton's equation. In practice, the atoms are assigned initial velocities that conform the total kinetic energy of the system, which in turn, is dictated by the desired simulation temperature. This is carried out by slowly heating the system from initially 0 K and then allowing the energy to equilibrate among the constituent atoms. The general concept of molecular dynamics is the calculation of the force on each atom, and from that information, the position of each atom after a specified period of time (typically on the order of 1 to 2 femtoseconds).

The force on an atom can be calculated from the change in energy between its current position and its position a small distance away, recognized as the derivate of the energy as defined in equation (6), while most MD codes calculate the force on an atom through the analytical derivatives of the potential, due to computational reasons. Knowledge of the atomic forces and masses then can be used to calculate the position of each atom along a series of very small time steps (in the order of femtoseconds). This small step size is necessary as constant forces are assumed through out each single step. The resulting series of snapshots of the system over time is called a trajectory. In practice, MD trajectories are not directly obtained from Newton's equation due to the lack of an analytical solution. Here approximations come into account, where the most common ones are the Verlet and Leap-Frog integrators.

Current state of the art programs for MD simulations are NAMD (Phillips, Braun et al. 2005), GROMACS (Hess, Kutzner et al. 2008) and Desmond(Bowers, Chow et al. 2006) being able to produce continuous MD trajectories of up to 10 µs on current supercomputers (Freddolino, Liu et al. 2008). In order to access even larger time scales of 10 microseconds to 1 millisecond, D. E. Shaw designed and constructed a special purpose machine called Anton that greatly accelerates the execution of MD simulations (Shaw, Dror et al. 2009) and found application in simulations of folding of an WW protein domain (Shaw, Maragakis et al. 2010), for example.

1.2.1.2. Monte Carlo methods

Next to the previously described deterministic molecular dynamics method, there also exist non-deterministic respectively statistical mechanics approaches to simulate biological systems. Among them is a whole class of algorithms called Monte Carlo methods.

In general, Monte Carlo methods rely on repeated random sampling to compute their results and are hence preferably used when it is unfeasible or impossible to compute an exact result with a deterministic algorithm. In the particular case of Monte Carlo simulations, the underlying algorithm is based on exploring the energy surface by randomly probing the geometry of the molecular system, or to speak in the statistical mechanics language, its configuration space.

The most popular realization of the Monte Carlo method is the Metropolis method, which is essentially composed of the following steps (Binder and Heermann 1992):

1 . Specify the initial atom coordinates

2 . Select some atom (side chain, loop) i randomly and move it by random displacement: ΔX_i, ΔY_i, and ΔZ_i.

3 . Calculate the change of potential energy ΔV corresponding to this displacement.

4 . If $\Delta V < 0$ accept the new coordinates and go to step 2.

5 . Otherwise, if $\Delta V \geq 0$, select a random number R in the range [0,1] and:

A: if $e^{-\Delta V/kT} < R$ accept the new coordinates and go to step 2,

B: if $e^{-\Delta V/kT} \geq R$ keep the original coordinates and go to step 2.

As a consequence of the random steps involved in the algorithm, the method does not provide time evolution of the system (in contrast to molecular dynamics). Nevertheless, the ability to jump barriers on the potential energy landscape as well as the speed reveal strong potential for its application for particular problems, such as side chain sampling or loop prediction, for example.

Protein energy landscape exploration

In order to map large conformational rearrangements due to protein-ligand and protein-protein interaction Borelli *et al.* developed a technique called protein energy landscape exploration (PELE) (Borrelli, Vitalis et al. 2005). The method, belonging to the class of Monte Carlo methods, is based on three main steps: random move, side chain sampling and minimization, and acceptance or rejection of the resulting conformation by a Metropolis acceptance test (see Figure 4).
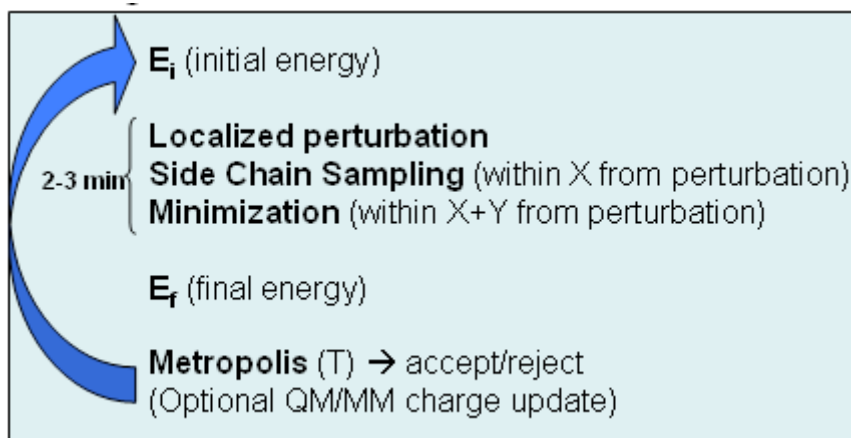
Figure 4: Sampling procedure in PELE.

The first step describes the localized perturbation. Depending on the system the region being allowed to move includes the backbone of the protein or only the backbone surrounding the center of interest (i.e. interface of the protein-protein complex). The alpha carbons are forced (by means of a harmonic constrain) to a new position resulting from a small displacement in a low frequency normal mode. As second step the algorithm replaces all specified side chains through optimized conformations derived from a rotamer library. The final step involves the minimization of a region including, at least, all residues local to the atoms involved in both previous steps. These three steps compose a move, which is accepted (if defining a new energy minima) or rejected based on a Metropolis acceptance test for a given temperature.

While the method has shown its capabilities to describe ligand escape/entry events in the micro- to millisecond time scale at a detailed atomic level (Borrelli, Vitalis et al. 2005; Guallar, Lu et al. 2009), it was mainly used throughout this

thesis to rapidly generate diverging conformations of a loosely bound protein-protein complex.

1.2.1.3. Quantum mechanics

In the early 20th century it was found out that classical mechanics had deficiencies to correctly describe some physical phenomena i.e. thermal radiation. Starting from then, researchers as Niels Bohr, Albert Einstein, Werner Heisenberg, Max Planck, Erwin Schrödinger, and many others have discovered a set of laws called *quantum mechanics*, being able to correctly describe the behavior of very small particles such as electrons or nuclei of atoms and molecules.

Quantum mechanics finds application for a multitude of problems in chemistry. It can be used to determine molecular properties, for example bond lengths and bond angles, atomic charges and dipole moments, barriers to internal rotation and energy differences between conformational isomers. Also it is applied to compute properties of transition states in chemical reactions, to estimate the relative stability of molecules or to calculate properties of reaction intermediates, for example. The following paragraphs will only give a short description of the underlying theory of quantum mechanics. For a more detailed description we refer the reader to excellent books explaining quantum mechanics as well as its application in chemistry (Pauling and Wilson 1985; Szabó and Ostlund 1996; Levine 2000).

Quantum chemistry

As mentioned before, in quantum chemistry the system is described by a set of nuclei coordinated in space with their corresponding electrons spread all over the system in a continuous electronic density. The equation describing the quantum state (wave function) of the system and its change in time is given by the Schrödinger equation:

$$\hat{H}\Psi = i\hbar \frac{\partial}{\partial t}\Psi \, , \tag{7}$$

where Ψ is the wave function expressing the probability amplitude for different configurations of the system at different times t, $i\hbar \frac{\partial}{\partial t}$ gives the time evolution of the system and $\hat{H}$ is the Hamiltonian operator. For systems in a stationary state, the time-dependent Schrödinger equation simplifies to the time-independent Schrödinger:

$$\hat{H}\Psi = E\Psi \tag{8}$$

In principle, it is possible to solve the Schrödinger equation in either its time-dependent or time-independent form, as appropriate for the underlying problem. Nevertheless, in practice we face the problem that this is not possible for many-electron atoms of molecules due to interaction between the particles, which is called

the many-body problem. As a consequence the Born-Oppenheimer approximation has to be applied, which assumes electronic and nuclear motions are independent. In practice this means all nuclei are treated frozen through the solution of the Schrödinger equation. Nowadays, there exist a wide range of methods to solve the Schrödinger equation, usually with no relativistic terms included and by incorporating the Born-Oppenheimer approximation. In doing so, the total energy of a system is determined by the sum of the electronic energy at fixed nuclei positions and the repulsion energy of the nuclei.

Quantum chemical methods that do not include any empirical or semi-empirical parameters in the equations are called *ab initio* methods. Still, these methods do not give an exact solution of the Schrödinger equation, but are approximations rigorously defined on first principles and then solved with an error margin that is qualitatively known beforehand. In order to describe the molecular orbitals (MOs) of the system, linear combinations of a set of functions (basis set) are applied with weights or coefficients to be determined throughout the calculations. Usually these functions are atomic orbitals being centered on atoms. The simplest type of *ab initio* electronic structure calculation is the Hartree-Fock (HF) method, where each electron is under the influence of the mean field created by all other electrons. Consequently, electron correlation is only introduced for those electrons in a given spin-orbital (Pauli exclusion theorem). Within the HF method, the wave function is optimized by means of an iterative algorithm following the variational principle (for a given basis set, the best wave function gives the lowest energy).

Also belonging to the *ab initio* scheme are the methods based on the density functional theory (DFT), even though many of the most common functionals use parameters derived from empirical data, or from more complex calculations. In DFT, the total energy is expressed in terms of the total one-electron density rather than the wave function. Here, an approximate Hamiltonian and an approximate expression for the total electron density are used. The major advantage of DFT methods are their little computational cost taking into account the electron correlation. Extensions of the DFT scheme, known as hybrid functional methods, are methods combining the density functional exchange functional with the Hartree–Fock exchange term.

In contrast to *ab initio* methods, semi-empirical quantum chemistry methods make use of parameters obtained from empirical data. Nevertheless, they are based on the Hartree–Fock formalism. Among them are methods where the two-electron part of the Hamiltonian is not explicitly included into the calculations, except for π-electron MOs in the case of the Hückel method (Coulson, O'Leary et al. 1978) or for valence orbitals in the case of the extended Hückel method (Hoffmann 1963). Nowadays, there exists a wide range of semi-empirical methods based on partially neglecting the differential overlap for the two-electron integrals, i.e. INDO/S (Ridley and Zerner 1973), MNDO (Bingham, Dewar et al. 1975), AM1 (Dewar, Zoebisch E.G et al. 1985) or PM3 (Stewart 1989). Despite their empirical nature, these methods are very important in computational chemistry for treating large molecules where the full Hartree–Fock method without the approximations is too expensive. Hence, large efforts are spend for proper parameterization of these

methods for all sorts of biochemical molecules, including proteins, DNA or ligands, for example.

So far we have only described the time-independent solution of the Schrödinger equation. As mentioned above it is also possible to solve the Schrödinger equation in its time-dependent form in order to investigate properties and dynamics of quantum mechanical systems in the presence of time-dependent potentials, such as electric or magnetic fields. A possible approach is time-dependent DFT (TDDFT) being an extension of the above explained DFT method. Here, the time-dependent effective potential at any given step depends on the values of the density at all previous times. Consequently, TDDFT is very expensive in terms of computational costs as well as memory and hence only finds application in small systems.

Mixed quantum mechanics / molecular mechanics

Combining the strength of both, QM (accuracy) and MM (speed) calculations, Warshel and Bromberg introduced the hybrid quantum mechanics/molecular mechanics (QM/MM) approach (Warshel and Bromberg 1970) which was later refined to the modern coupled QM/MM approach (Warshel and Levitt 1976). Here, the region of interest (i.e. the active site) can be studied with a robust *ab initio* QM methodology, employing a QM region of sufficient size to encompass any important electronic structure effects. The remainder of the protein can be modeled at the much cheaper MM theoretical level, providing the appropriate structural constraints, and the electrostatic and van der Waals non bonding interactions with the core reactive region. Figure 5 shows a schematic view of QM/MM applied to substrate (treated by QM) bound to protein (treated by MM) in a sphere of solvent.

Substrate, catalytic residues treated by QM

Protein, solvent treated MM;interacts with QM

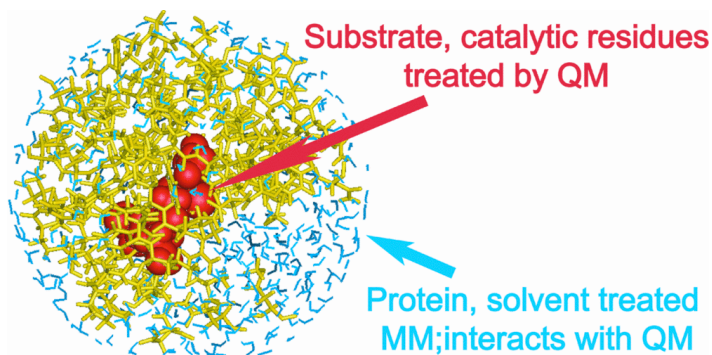**Figure 5:** Schematic view of QM/MM applied to substrate bound to protein. Figure taken from (Mulholland 2008).

The Hamiltonian $\hat{H}$ describing the complete QM/MM system is given as follows:

$$\hat{H} = \hat{H}_{QM} + \hat{H}_{QM/MM} + \hat{H}_{MM}. \tag{9}$$

While truncation of the QM region is straightforward for systems where the QM and MM regions are completely separate, special treatment is needed if the

QM/MM cut divides covalent bonds. There exist several solutions, i.e. localized-orbital methods (missing valence of the QM interface atom is filled with a frozen orbital), or link-atom approaches (missing valence of the QM interface atom is filled with dummy hydrogen atoms only visible to the QM part of the calculations) (Reuter, Dejaegere et al. 2000). A common derivate from the QM/MM setup explained here is the ONIOM approach as developed by Morokuma and co-workers (Svensson, Humbel et al. 1996). Here the energy of a system is described by:

$$\hat{H} = \hat{H}_{\text{model},QM} + \hat{H}_{\text{real},MM} - \hat{H}_{\text{model},MM}, \tag{10}$$

where *real* and *model* refer to the full system and the QM region, respectively.

The computational cost of QM/MM techniques is similar to that of QM methods alone with the advantage of having the influence of the complete MM region on the QM region. A further great advantage is its modular conception giving the ability to incorporate any kind of QM as well as MM description, based on the actual need. Thus, nowadays QM/MM techniques enjoy great popularity in the scientific community, especially for descriptions of entire enzymatic systems (Gao and Truhlar 2002; Sherwood, de Vries et al. 2003; Friesner and Guallar 2005).

1.2.2. Electronic coupling

The electronic coupling of donor and acceptor is the key characteristic that controls the rate of ET. It determines the distance dependence of the rate constant and its sensitivity to the relative orientation of donor and acceptor. Already published in the 1990s, there are a number of excellent reviews describing the quantum chemical treatment of electron transfer including electronic coupling calculations (Newton 1991; Mirkin and Ratner 1992; Newton and Cave 1997). A further review summarizes nicely available procedures to derive the coupling magnitude from experiments (Kumar, Kurnikov et al. 1998).

In general, several procedures have been proven useful for calculating the electronic coupling matrix elements, most of them applying the two-state model. In accord with the Franck-Condon principle, the electron transfer occurs at the crossing between two potential curves of the reactant and product states. The electronic coupling of the two states then determines the probability of the crossover at the crossing point. It can be determined either in terms of diabatic states (Farazdel, Dupuis et al. 1990) or adiabatic properties, for example as the minimum splitting of electronic states (Newton 1991). In the following section we will give a brief introduction to a variety of approaches to calculate the electronic coupling. Furthermore we explain in detail two very recent methods, namely generalized Mulliken-Hush and fragment charge difference, based on adiabatic state information.

1.2.2.1. Orbital diagonalization method

In 1990 Siddarth and Marcus introduced an approach applying diagonalization of the bridge orbitals and second-order perturbation theory to calculate the

electronic coupling matrix element (Siddarth and Marcus 1990) for bridge mediated systems. Based on the approximation that the transition state of ET occurs at the intersection of the energy surfaces for the reactants with that for the products, i.e. the energy of donor orbital E_χ equals that of the acceptor orbital E_μ, they calculated the electronic coupling according to the superexchange mechanism in a one-electron description.

For the case where the bridge B is connected to the donor D (acceptor A) by only one atomic orbital, the electronic coupling V_{DA} is given by (Larsson 1981)

$$V_{DA} = T_D T_A \sum_\alpha \frac{C_{D\alpha} C_{A\alpha}}{E_\chi - E_\alpha}, \tag{11}$$

where T_D (T_A) is the matrix element for the interaction between D (A) and the adjacent atomic orbital of B, $C_{D\alpha}$ ($C_{A\alpha}$) is the coefficient of the αth bridge orbital at the point of contact of B with D (A), E_α is the energy of the αth molecular orbital of B. In the case when D and A are connected to B by a number of atomic orbitals, the electron-transfer matrix element is given by

$$V_{DA} = \sum_\alpha \sum_{m_D, m_B, m_A} \frac{\left(C^*_{\chi m_D} T_{m_D m_B} C_{m_B \alpha}\right)\left(C^*_{\alpha m_B} T_{m_B m_A} C_{m_A \mu}\right)}{E_\chi - E_\alpha}, \tag{12}$$

where χ and μ denote the molecular orbitals on D and A that are involved in the electron transfer, the α's denote the MO's of B, and m_D, m_B and m_A are the atomic orbitals of D, B and A, respectively. T_{mDmB} is the interaction matrix of the donor and bridge orbitals, and T_{mBmA} is the interaction matrix of the bridge and acceptor orbitals. Within the original publication the authors applied extended Hückel theory to calculate E_α, while nowadays more accurate method can be used as well.

1.2.2.2. Energy splitting method

Like for the orbital diagonalization approach, the energy splitting method assumes ET occurring in an inactivated state with degenerate donor and acceptor localized electronic states (φ_D and φ_A) (Onuchic, Beratan et al. 1986; Jordan and Paddon-Row 1992; Liang and Newton 1992). At this point, φ_D and φ_A are mixed via bridge states (superexchange). The unmixed system is quasi-degenerate with approximate eigenstates $\varphi_0 = (1/\sqrt{2}) (\varphi_D + \varphi_A)$ and $\varphi_1 = (1/\sqrt{2}) (\varphi_D - \varphi_A)$. V_{DA} is then one-half of the energy splitting ($|E_1 - E_0|$) between these eigenstates.

As donor and acceptor sites generally differ in biological systems, it is usually not easy to derive the inactivated state of the system. A possible approach to force degeneration is to apply an electric field along the donor-acceptor axis as done by Prytkova et al. (Prytkova, Kurnikov et al. 2005). Here, the authors varied the electric field strength until degeneracy of the localized states and then computed V_{DA} as one-half of the minimum energy splitting.

1.2.2.3. Direct *ab initio* quantum chemical method

Ab initio quantum chemical methods represent in principle the most rigorous way to determine the ET matrix elements. Already in 1990, Broo and Larsson carried out calculations on small ET systems of aliphatic chains (Broo and Larsson 1990), followed by studies on moderate donor-acceptor systems (Braga and Larsson 1993; Paulson, Curtiss et al. 1996) and reaching sophisticated electronic coupling calculations between two bacteriochlorophyll molecules in 1997 (Zhang, Friesner et al. 1997). Due to the increase of computational power and development of better algorithms, not only the applied systems sizes but also the quantum chemical precision (number of basis sets) have increased steadily.

The approach described by Zhang et al. is based upon direct evaluation of the diabatic charge localized wave functions, followed by the determination of the off-diagonal coupling matrix element of the Hamiltonian between these wave functions (Zhang, Friesner et al. 1997).

Within the two-state model, the electronic coupling between donor and acceptor is given by

$$V_{DA} = \frac{H_{fi} - S_{if} H_{ii}}{1 - S_{if}^2},$$

(13)

where

$$H_{fi} = \left\langle \Phi_f \middle| H \middle| \Phi_i \right\rangle,$$

(14)

$$H_{ii} = \left\langle \Phi_i \middle| H \middle| \Phi_i \right\rangle,$$

(15)

$$S_{if} = \left\langle \Phi_i \middle| \Phi_f \right\rangle,$$

(16)

and H is the full Hamiltonian of the system. Φ_i and Φ_j are the wave functions of the initial electron transfer state (electron still in donor) and final electron transfer state (electron in acceptor), respectively. The authors distinguish intermolecular ET and intramolecular ET due to different procedures to determine Φ_i and Φ_j. In the first case, the charge localized donor and acceptor electronic states can easily be calculated by placing the charge in the respective sites. By doing so one gets Φ_{D^-} (an extra electron is in the donor), Φ_A (neutral acceptor), Φ_D (neutral donor) and Φ_{A^-} (an extra electron is in the acceptor). Thus, the total wave functions for the initial and final states are

$$\Phi_i = \hat{A} \left\{ \Phi_{D^-} \Phi_A \right\},$$

(17)

$$\Phi_f = \hat{A}\{\Phi_D \Phi_{A-}\}, \tag{18}$$

where $\hat{A}$ denotes that the wave function of the hole system is antisymmetric and hence has to be orthonormalized by Schmidt orthogonalization.

For a single donor-acceptor moiety connected by a molecular bridge (intramolecular ET), determination of localized donor and acceptor states requires introduction and fitting of an external point charge to generate an appropriate initial guess for the desired electronic state. By success, the wave functions are already the charge localized wave functions and need no further operations and the electronic coupling can be determined by equation (13).

1.2.2.4. Generalized Mulliken-Hush method

Based on a transformation from adiabatic properties to diabatic states, Cave and Newton proposed the generalized Mulliken-Hush method (GMH) to calculate the electronic coupling between a donor D and acceptor A (Cave and Newton 1996; Cave and Newton 1997). Being applicable to both, ground and excited state, this approach transforms the adiabatic to the diabatic states using the matrix that diagonalizes the adiabatic dipole moment matrix. For the two-state model in weakly coupled systems ($V_{DA} < k_B T$), the electronic coupling V_{DA} can be expressed by the following equation:

$$V_{DA} = \frac{(E_2 - E_1)|\mu_{12}|}{|\mu_D - \mu_A|}, \tag{19}$$

where $E_2 - E_1$ is the vertical excitation energy of the two relevant adiabatic states. μ_{12} is the transition dipole moment and $|\mu_D - \mu_A|$ is the difference of the diabatic dipole moments. Here, we can estimate $|\mu_D - \mu_A|$ as eR_{DA} or as $((\mu_1 - \mu_2)^2 + 4\mu_{12}^2)^{1/2}$. The transition dipole connecting localized states should be zero due to the locality of the dipole moment operator and the exponential decay of the localized states. Consequently, it is important to use only the projection of the transition and dipole moments onto the axis between the donor and acceptor rather than the length of the vectors.

One-electron approximation

Within the hole transfer process where an electron deficit is transferred from the acceptor to the donor, the electronic coupling can be calculated applying the one-electron of Koopmans' theorem approximation (Onuchic, Beratan et al. 1986; Liang and Newton 1992). Here, the properties of the adiabatic states for a radical cation can be approximated through one-electron energies and occupied molecular orbitals of the corresponding neutral system:

$$E_2 - E_1 = E_2^{N-1} - E_1^{N-1} = \left(E_2^{N-1} - E_0^{N-1}\right) - \left(E_1^{N-1} - E_0^{N-1}\right) \approx \varepsilon_{HOMO-1}^N - \varepsilon_{HOMO}^N, \tag{20}$$

where the superscripts N and $N\text{-}1$ refer to the neutral and radical cation systems, respectively. E_0^N is the ground-state energy of the neutral system and the energies ε_k (k = HOMO, HOMO-1) refer to the corresponding orbitals of that system (Rösch and Voityuk 2004).

To apply GMH, equation (19), within the Koopmans' approximation, one can calculate the difference $\mu_1 - \mu_2$ of the adiabatic dipole moments and the transition moment μ_{12} as follows:

$$\mu_1 - \mu_2 = \sum_{i,j=1}^{M} \left(C_{i,HOMO} C_{j,HOMO} - C_{i,HOMO-1} C_{j,HOMO-1} \right) d_{ij} \;. \tag{21}$$

$$\mu_{12} = \sum_{i,j=1}^{M} C_{i,HOMO} C_{j,HOMO-1} d_{ij} \;. \tag{22}$$

Superexchange correction

We already mentioned the particular case when the energies of the bridge states are close to the ones of donor and acceptor in section 1.1.2. In this case, the electron is able to physically occupy the intermediate state where the electron partly localizes into the bridge. As a consequence, the diabatic states are often represented by combinations of three or more adiabatic states. Here, the two-state model becomes inaccurate and a multistate treatment must be employed as shown in the case of hole transfer in DNA π-stacks by A. Voityuk (Voityuk 2005). The author showed that, for the case of GMH, a superexchange correction V^{SE} properly accounts for the bridge state effects and has to be added to the two-state electronic coupling V_{DA} derived from equation (19):

$$V_{DA}^{MS} = V_{DA} + V^{SE} = V_{DA} + \sum_{j=3}^{n} \left(E_j - \frac{E_1 + E_2}{2} \right) \frac{\mu_{1j} \mu_{2j}}{\left(\mu_j - \mu_D \right) \left(\mu_A - \mu_j \right)}, \tag{23}$$

where states $j \geq 3$ are the bridge states. Introducing this scheme, Cave and Newton suggested that the number of states n involved in a multistate treatment should be kept as small as possible (Cave and Newton 1997). Obviously, the number of bridge states depends on the structure of the system and hence should be determined for each type separately (Rust, Lappe et al. 2002; Voityuk 2006).

We would like to mention here that in the case of ET where the electron, respectively hole is able to fully occupy a bridge site, the actual ET system is split into two separate parts: ET between the original donor and the bridge, and ET between the bridge and the final acceptor. Both single ETs are treated separately resulting into two electronic coupling values and therefore two distinct rate constants. The final rate constant of the full ET between the original donor and final acceptor is then the smaller rate constant of both due to its rate limiting nature.

Advantages and disadvantages

The great advantage of GMH against direct *ab initio* quantum chemical electronic coupling calculation or energy splitting methods is its ability to derive the parameters needed for the calculations from pure adiabatic state properties. Thus, there is no requirement for external treatments (i.e. electric field) to force the system to localize into its diabatic states. This advantage is its greatest disadvantage at the same time. Diabatic quantities are only estimated from purely adiabatic state properties. Further advantage of the GMH method is that it is able to deal with systems where the diabatic states are described by more than two adiabatic states (see previous section). Here, the two-state model can fail to calculate the actual electronic coupling as investigated in several studies (Cave and Newton 1996; Cave and Newton 1997; Voityuk and Rösch 2002; Voityuk 2005). Finally, by applying the one-electron approximation together with GMH (see above), it is not necessary to compute the excited and ground states of radical systems but to simply derive the adiabatic state information from the HOMOs, respectively LUMOs, of the corresponding neutral systems.

Consequently, GMH has become de facto the standard method to evaluate electronic coupling on the bases of quantum chemical calculations of ET (Voityuk 2008), finding large application in the ET research field (Cave and Newton 1996; Cave and Newton 1997; Ungar, Newton et al. 1999; Castner, Kennedy et al. 2000; Rust, Lappe et al. 2002; Lambert, Amthor et al. 2004; Zheng, Kang et al. 2005; Blancafort and Voityuk 2006).

1.2.2.5. Fragment charge difference method

Like the GMH method, the fragment charge difference method (FCD) is based on the adiabatic-to-diabatic-state transformation. Here, the adiabatic states are rotated to diabatic states maximizing the charge transferred between the redox sites (Voityuk and Rösch 2002; Rösch and Voityuk 2004). In order to apply FCD it is mandatory to explicitly define D and A sites involved in the ET. Within the two-state model, the bridge-mediated electronic coupling can be calculated as:

$$V_{DA} = \frac{(E_2 - E_1)\left|\Delta q_{12}\right|}{\sqrt{(\Delta q_1 - \Delta q_2)^2 + 4\Delta q_{12}^2}} ,$$ (24)

where, Δq_1 and Δq_2 are the donor-acceptor charges difference in the two adiabatic states with energies E_1 and E_2, respectively, and Δq_{12} is the corresponding off-diagonal term. Applying the one-electron approximation explained above, the charge of fragment F (donor or acceptor) of the first adiabatic state can be estimated using the corresponding Mulliken populations of the HOMO of the neutral system:

$$q_1\left(F\right) = \sum_{i \in F} C_{i,HOMO} \sum_{j=1}^{M} C_{j,HOMO} S_{ij} .$$ (25)

Here, S_{ij} is the overlap of the atomic orbitals (AOs) i and j, where i runs only over AOs associated with the selected fragment F while j runs over all AOs. The fragment charges of the second adiabatic state are estimated analogously applying coefficients of molecular orbital HOMO-1 instead of HOMO in equation (25). The general quantity $q_{xy}(F)$ can be defined by

$$q_{xy}\left(F\right) = \frac{\sum_{i\in F} C_{i,HOMO+1-x} \sum_{j=1}^{M} C_{j,HOMO+1-y} S_{ij} + \sum_{i\in F} C_{i,HOMO+1-y} \sum_{j=1}^{M} C_{j,HOMO+1-x} S_{ij}}{2} . \tag{26}$$

The matrix elements Δq applied in equation (24) are found as

$$\Delta q_{xy} = q_{xy}\left(D\right) - q_{xy}\left(A\right). \tag{27}$$

The method provides accurate results when the adiabatic states are properly defined and comparison of GMH and FCD methods have shown that in most cases both methods give very similar results (Voityuk and Rösch 2002; Rösch and Voityuk 2004; Subotnik, Cave et al. 2009).

Furthermore, the FCD scheme has been recently extended for treatment of electronic coupling for excitation energy transfer (Hsu, You et al. 2008).

Multi-state coupling

Reliable estimation of electronic coupling is a great challenge, particularly in systems having nearly-degenerate electronic states. Here, minimal structural fluctuations of the redox centers or their environment can provoke reordering of these states. Thus, the pure two-state model taking into account only one state of each redox site, may not appropriately describe the electronic coupling of these systems. As the reordering of the donor and acceptor levels occurs at the pico-seconds time scale and thus is much faster than ET, averaging over several states should be used to approximate the electronic coupling (Troisi, Nitzan et al. 2003). The root mean square coupling $rms V_{DA}$ and the effective coupling squared V^2 are defined as following (Voityuk 2010):

$$rms V_{DA} = \sqrt{\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} V_{ij}^2} . \tag{28}$$

$$V_{DA}^2 = \frac{1}{N_1 N_2} rms V_{DA}^2 . \tag{29}$$

N_1 and N_2 denote the number of (nearly) degenerate diabatic states localized on donor and acceptor sites, respectively. The use of $rms V_{DA}$ rather than $<V_{DA}>$ is in

line with the expression for the non-adiabatic ET (Marcus and Sutin 1985; Newton 1991).

Advantages and disadvantages

Like GMH, the FCD method has the advantage/disadvantage of applying purely adiabatic state properties for its electronic coupling calculations. Clearly, the one-electron approximation can be applied as well, simplifying the excited and ground state calculations of radical systems to ground state calculations of the neutral systems. An advantage of FCD against GMH is the use of charge separation between the donor and acceptor instead of dipole moments vectors. Thus, no directions have to be applied, which constricts the application of possible bridge states to only those being in line with the donor and acceptor. A disadvantage of FCD is that the user has to define the donor and acceptor a priori as the charge difference is then calculated on these fragments.

All in all, the FCD method is an established and easy-to-apply electronic coupling method, which has been carefully evaluated against GMH as well as direct coupling methods in several studies (Voityuk and Rösch 2002; Rösch and Voityuk 2004; Hsu, You et al. 2008; Hsu 2009; Subotnik, Cave et al. 2009).

1.2.3. Data analysis tools

In order to keep track of and analyze the vast amount of data produced by computational methods or experiments, data analysis tools are needed. There exist thousands of different methods and approaches ranging from very general ones, like calculating the plain mean over all observations, to more sophisticated ones being suitable for only very specific problems. A comprehensive overview of this own field of research is given in several excellent books (Han and Kamber 2006; Hair, Black et al. 2009). Throughout this thesis we applied k-medoids clustering of conformational snapshots derived from molecular dynamics simulations as well as partial least squares regression in order to identify structural parameters being important for high coupling in oligopeptide model systems. Both methods are going to be briefly presented in the following paragraphs.

The k-medoids clustering

The k-medoids algorithm is a clustering algorithm based on the k-means algorithm and the medoidshift algorithm (Theodoridis and Koutroumbas 2006). The algorithm is partitioning the data while trying to minimize the error, the sum of all distances between the data points labeled to be in a cluster and a point designated as the center of that cluster. In clear contrast to the k-means algorithm, k-medoids chooses data points as centers, the so-called "medoids". The algorithm works iterative with a predefined number of clusters k:

1. Randomly choose k medoids m_j from all N data points.
2. Assign every data point x_i to its nearest medoid.
3. Calculate error sum

$$E = \sum_{i=1}^{N} d\left(i, m_i\right), \qquad\qquad (30)$$

with m_i being the nearest medoid of data point i.

4. Swap medoid m_i with data point j if error E gets smaller, followed by renewal of the group membership of all data points.

5. Repeat step 3 and 4 until convergence.

The k-medoids algorithm is particularly interesting for the clustering of snapshots taken from MD because it only needs a measurement for the distance between the actual data points, as all cluster centers are data points themselves. Thus, in contrast to most other clustering algorithms, no average representatives of clusters have to be calculated. Nevertheless, the algorithm is computationally intensive, as the full $N \times N$ distance matrix has to be computed, whereas the actual clustering converges after a few steps. A widely used distance measure is the root mean square distance (rmsd) calculated for all C_α atoms of the snapshots.

Partial Least Squares regression

The multivariate regression method Partial Least Squares regression (PLS-R) (Wold, Sjöström et al. 2001) is used to find the fundamental relations between two matrices X and Y. The method works by finding the multidimensional direction in the X space that explains the maximum multidimensional variance direction in the Y space, describing both X and Y by a few latent variables also known as principal components (PC). PLS-R models are usually built by extracting successive PCs, each one increasing the total percentage of Y variance explained given by "cumulative Q2", until the predictive ability of the model is optimized. PLS-R is particularly suited when there is multi-colinearity among the X values, where, by contrast, standard regression methods will fail. A further advantage is the ability to extract information about both the objects and the variables. The objects can be represented graphically according to the PC values, obtaining highly informative score plots in which similar objects appear as points closely situated in the space. In addition, the original variables can also be represented according to their overall contribution to the model by coefficient plots. These plots provide information about which variables exhibit a relevant association with the dependent variable, the direction of it and if this association is statistically significant.

1.3. Applied systems

Many ET respectively electronic coupling studies have focused on artificially engineered test systems (Newton 1991; Siddarth and Marcus 1993; Zhang, Friesner et al. 1997; Shin, Newton et al. 2003; Zimmt and Waldeck 2003; Malak, Gao et al. 2004; Berlin and Ratner 2005). These systems have a clearly defined electron donor and acceptor as well as only a single possible ET pathway in between them. Furthermore, these systems are generally straightforward to engineer and apply, both from the experimental and computational point of view. Considering the experimental case, the donor is generally easy to excite by electron pulse radiolysis or photoinduction (Dmochowski, Crane et al. 1999; Malak, Gao et al. 2004) and the resulting ET to the donor then measured by stopped flow kinetics or transient absorption spectroscopy, for example. Considering the computational case, MD simulations are necessary in order to sample conformational space followed by expensive QM calculations deriving the electronic coupling on a multiplicity of snapshots become only feasible if the applied system does not exceed certain size, meaning number of atoms, which is the case in these model systems. Consequently, they contain only a few degrees of freedom in terms of protein dynamics or ET mechanism. Thus, these model systems are well suited to investigate the underlying species of ET and its dependence on varying parameters such as donor-acceptor distance or structural conformation.

Eventually, researchers want to know how real nature works and not only how small model systems behave. Thus, they started to apply ET theory on enzymes. One of the first ET proteins under investigation was ruthenium-labeled myoglobin (Axup, Albin et al. 1988). Many other studies are based on ruthenium modified azurins (Gehlen, Daizadeh et al. 1996; Gray and Winkler 1996; Kawatsu, Kakitani et al. 2002; Prytkova, Kurnikov et al. 2005; Shih, Museth et al. 2008). Finally, researchers also investigate the ET mechanisms of none-modified proteins. Among them are the bacterial reaction center II (Ceccarelli and Marchi 2003), lysozyme (Morozova, Hore et al. 2005), ribonucleotide nucleotase (Reece, Seyedsayamdost et al. 2007) and methylamine dehydrogenase (Davidson 2008), for example.

Moreover, cytochrome P450 as well as cytochrome c peroxidase received early attention in terms of ET process. Both proteins were taken under investigation within this dissertation and are thus further explained in the following sections.

1.3.1. Oligopeptides

Throughout this PhD thesis we carried out two studies, solvent and temperature effects on donor-acceptor couplings in peptides, based on short oligopeptides of tryptophan based donor and acceptor connected by a variable bridge of prolines. The systems were constructed in dependence on experimental model systems used in the group of Isied for many years (Isied, Ogawa et al. 1992; Ogawa, Wishart et al. 1993; Malak, Gao et al. 2004), with the simplification of the ruthenium-based donor and acceptor to tryptophans. In detail, we assembled oligopeptides Trp-(Pro)n-Trp, with n ranging from 1 to 6, using Schrödinger's Maestro(2010) build

utility. Figure 6 shows extended and folded conformations for Trp-(Pro)3-Trp and Trp-(Pro)6-Trp for illustration.
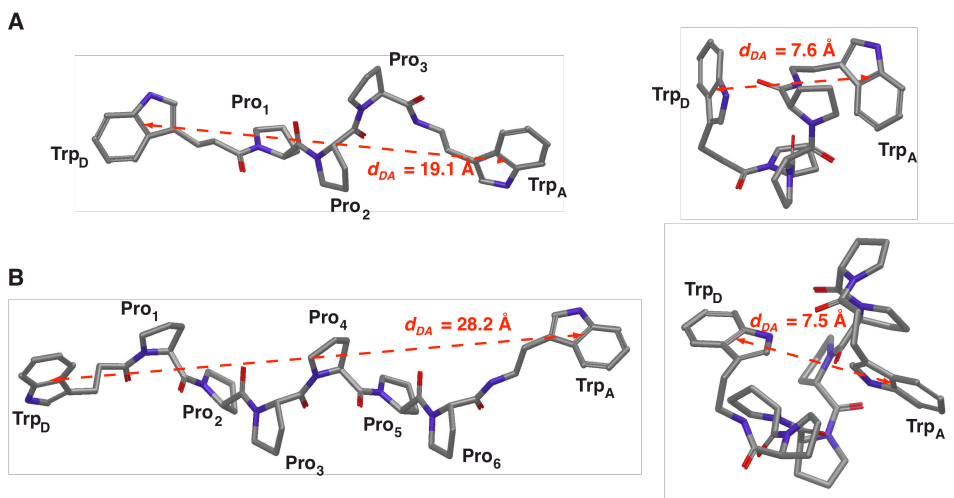

Figure 6: Oligopeptides Trp-(Pro)3-Trp and Trp-(Pro)6-Trp in panels A and B, respectively. Examples for extended and folded conformations are given for each system including their donor-acceptor distance.

Recent experimental work on the before mentioned oligoprolines between ruthenium based donors and acceptors has shown that the electron transfer mechanism can change from predominantly electron superexchange to electron hopping depending on the number of bridging prolines between the donor and acceptor (Malak, Gao et al. 2004). In detail, the study measured the rate constant of the ET between donor and acceptor in the systems having 0 to 9 prolines as spacer. In systems having 0 to 4 prolines ET was triggered by laser flash photolysis, and in systems having 3 to 9 prolines ET was triggered by pulse radiolysis. The measured maximum rate constants for all systems are given in Figure 7, clearly indicating the transition point at about 20 Å.

In our two studies based on ET model systems, we investigated the possibly underlying reasons for this change in the ET regime. In particular, we studied the dependence of conformational dynamics of the oligopeptides on electronic coupling as well as the effect of solvent. The resulting publication is given in section 3.4.1. The second study focused on the dependence of electronic coupling on the different conformational dynamics of short and long oligopeptides at different temperatures, with the publication given in section 3.4.2.
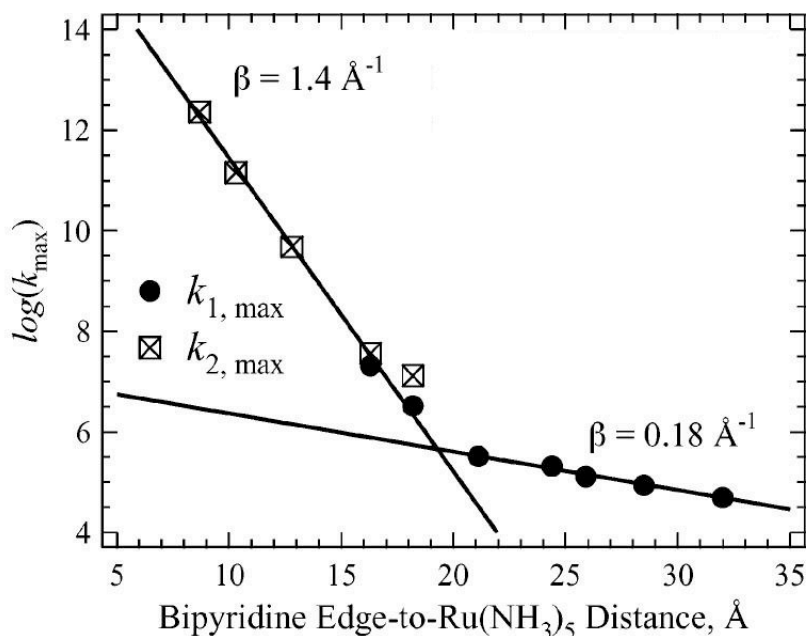
Figure 7: Plot of ET corrected rate constants $\log(k_{max})$ for radiolysis ($k_{1,max}$, circles) and photolysis ($k_{2,max}$, squares), versus the donor-acceptor distance in oligoproline systems with n = 0 to 9 prolines.

1.3.2. P450 camphor / Putidaredoxin complex

The family of cytochrome P450 is ubiquitous in human biology, playing a key role in the metabolism of pharmaceutical agents and other ingested exogenous compounds (Ortiz de Montellano 1995; Ogliaro, Harris et al. 2000). These enzymes insert an oxygen atom from O_2 into a wide variety of substrates, with substrate specificity determined by the nature of the protein active site cavity.

Among the P450 family is P450 camphor (P450cam), catalyzing the monooxygenation of *D*-camphor. In the reaction, two electrons from NADH are sequentially transferred to P450cam via two redox-linked proteins, putidaredoxin reductase and putidaredoxin (Pdx), a Fe_2S_2 protein. The first catalytic step from the low-spin resting state of P450cam involves camphor binding, which displaces the water molecule from the heme pocket (Schlichting, Berendzen et al. 2000). This binding produces a shift from low-to-high spin, accompanied by a 130-mV increase in the reduction potential which allows thermodynamic reduction by a first Pdx. In the second catalytic step, another reduced Pdx forms a binary complex with oxyferrous, *D*-camphor bound P450cam and leads to formation of the reaction products, 5-exo-hydorxycamphor, water, *D*-camphor free ferric P450 and oxidized Pdx (Gunsalus, Meeks et al. 1974; Lewis 1996; Brazeau, Wallar et al. 2003). The schematic overview of the full catalytic cycle of P450cam is illustrated in Figure 8.
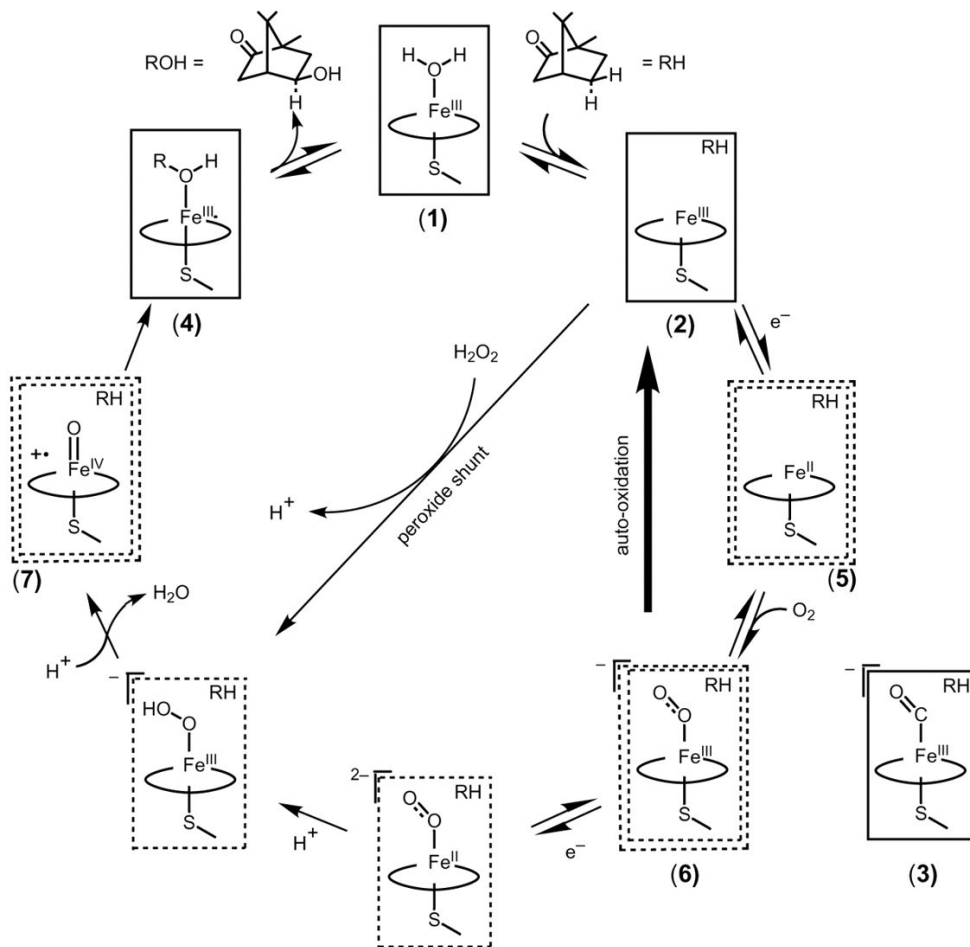
Figure 8: Catalytic cycle of P450cam. Resolved crystal structures: **(1)** P450aquo, Fe(III), **(2)** P450cam, Fe(III), **(3)** P450cam_CO, Fe(II) and **(4)** P450product, Fe(III). Intermediary states: **(5)** P450cam, Fe(II), **(6)** P540cam_O_2-, Fe(III) and **(7)** P450cpdI, Fe(IV). Figure adapted from (Schlichting, Berendzen et al. 2000).

The crystal structure of the P450cam/Pdx complex has not yet been resolved. However there are numerous studies on modeled complex structures derived from docking of the monomeric structures (Pochapsky, Lyons et al. 1996; Roitberg, Holden et al. 1998; Karyakin, Motiejunas et al. 2007). The first model, proposed by Pochapsky et al. in 1996, has a distance of 12.0Å between the heme iron of P450cam and the iron of the Fe_2S_2 cluster in Pdx closest to the surface (Pochapsky, Lyons et al. 1996). Furthermore it points out salt bridges from Arg109 (P450cam) to Asp34 (Pdx), Arg112 (P450cam) to Asp38 (Pdx) and Arg79 (P450cam) to the carboxyl group of C-terminal Trp106 (Pdx). The model proposed by Roitberg et al. two years later is very similar, also having the distance of 12Å between the two redox centers and salt bridges between the residue pairs mentioned above (Roitberg, Holden et al. 1998). They determined both binding sites in performing electrostatic Poisson-Boltzmann calculations resulting in a triangle of continuous negative

electrostatic potential between Asp34, Asp38 and Trp106 of Pdx and patches of positive electrostatic potential contours in the proximal face of P450cam. Recent studies on P450cam/Pdx also verify the salt bridge between Arg112 (P450cam) and Asp38 (Pdx), whereas the existence of the previously suggested salt bridge between Tyr78 (P450cam) and the carboxyl group of Trp106 (Pdx) was not confirmed (Shimada, Nagano et al. 2001; Wade, Motiejunas et al. 2005; Karyakin, Motiejunas et al. 2007). Here, the two proposed complex structures have a similar iron-iron distance of 14.7 Å and 16.6 Å, but the second has Pdx rotated about 90° around its vertical axis. Both complexes include a close contact of Arg112 (P450cam) and Asp38 (Pdx) forming the predicted salt bridge. Figure 9 displays the P450cam/Pdx complex in ribbon presentation.

Figure 9: Ribbon presentation of the P450cam/Pdx complex as modeled by Roitberg et al. Heme, camphor, the Fe_2S_2 cluster and specific residue important for binding and ET are highlighted by tube presentation.

Roitberg et al. also obtained pathway models for the electron transfer by calculating the electron couplings of the surface residues to their respective redox center (Roitberg, Holden et al. 1998). The results strongly highlight Asp38 and Ser44 in Pdx and Arg112 and Leu356 in P450cam as possible electron transfer residues (see Figure 10). Additionally to theoretical calculations, they applied site mutagenesis experiments identifying Arg112 (P450cam) and Asp38 (Pdx) as major residues responsible for binding and electron transfer, in agreement with previous mutagenesis studies (Nakamura, Horiuchi et al. 1994; Unno, Shimada et al. 1996). The authors propose an ET pathway from the Fe_2S_2 cluster via Cys39 to Asp38 of Pdx making a protein-protein jump onto Arg112 of P450 and from there a second through-space jump to the propionate group of the heme in P450cam. They further point out a less probable electron transfer pathway through Ser44 (Pdx) and Leu356 (P450cam). Recently, Harada et al. performed experiments with a "one-legged heme", where the 6-propionate side chain is replaced by a methyl group (Harada, Sakurai et al. 2007). The enzymatic activity of the reconstituted P450cam, having this mutated heme, is maintained whereas Pdx affinity is approximately 3.5 fold weaker. Their results indicate that the 6-propionate side chain of heme has mainly two roles, fixation of the Pdx-binding site by the Arg112 residue and prevention of transformation into the inactive P420 species in stabilizing the Fe-Cys357 ligation. The previous suggested ET pathway through Arg112 into the heme propionates (Roitberg, Holden et al. 1998) could not be confirmed.

Thus, although the mechanism of interaction and ET for this redox couple has been under investigation for over 30 years, a clear picture of the ET pathway from Pdx to P450cam is still missing. Moreover, to the best of our knowledge, theoretical ET pathway investigations incorporating the conformation dynamics of the protein complex does not exist yet. Throughout this dissertation we focused on exact these issues resulting in the publications given in section 6.
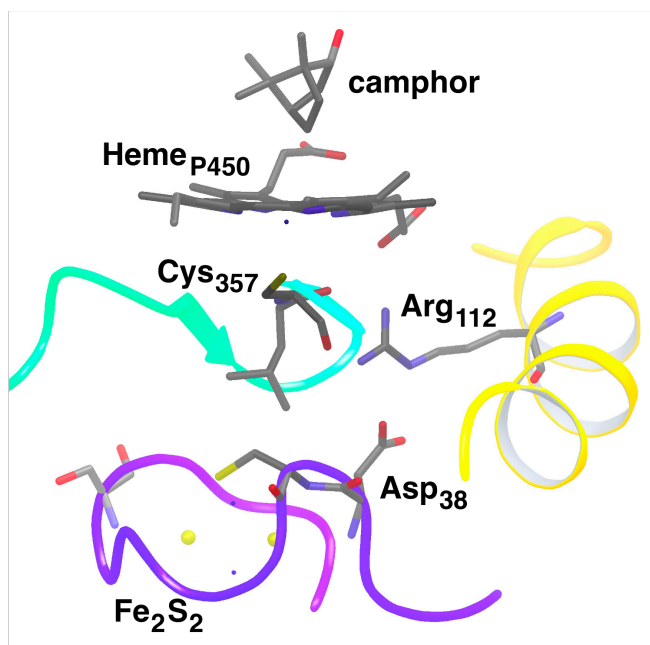
Figure 10: ET transfer region of the P450cam/Pdx complex.

1.3.3. Cytochrome c Peroxidase / Cytochrome c complex

Yeast cytochrome c peroxidase (CcP) was discovered in brewer's yeast by Altschul *et al.* in 1940 (Altschul, Abrams et al. 1940). It is localized in mitochondria of aerobically grown yeast cells (Yonetani and Ohnishi 1996) and seems to be important for the protection of the organism from high concentrations of hydrogen peroxide (Chance, Devault et al. 1967). While CcP was under investigation by many researchers trying to reveal its properties and reaction mechanism since the 1960s, its crystallographic structure was firstly discovered in 1978, providing the first three-dimensional structure of a heme enzyme (Poulos, Freer et al. 1978; Poulos, Freer et al. 1980). CcP catalyzes the reduction of hydrogen peroxide to water through the consecutive oxidation of two ferro-cytochrome c (Cytc) (see Figure 11), serving as a simple model system for the much more sophisticated electron transfer complexes of the respiratory electron transport chain.
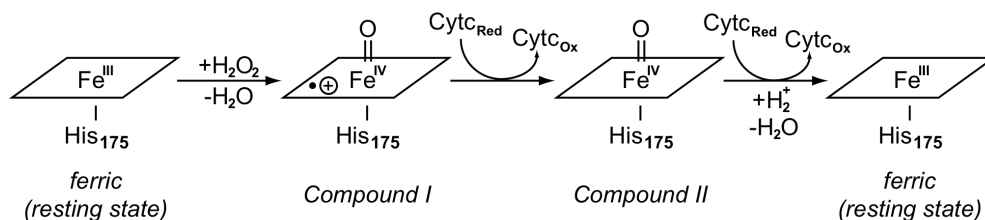

Figure 11: Catalytic cycle of CcP. Figure adapted from (Harvey, Bathelt et al. 2006).

INTRODUCTION

In the catalytic cycle, CcP undergoes a two electron oxidation by peroxide to form a common oxo-ferryl intermediate, the so-called compound I (CpdI). The iron atom in CpdI is formally in the very high +V oxidation state. However, this state is in fact not observed as an electron is transferred to the iron centre from either a high-lying doubly occupied orbital on the porphyrin ring, the proximal ligand or another group with low ionization potential in close vicinity (Sharp, Mewies et al. 2003{Gumiero, 2010 #1978)}. Thus, the iron has a formal +IV (*d4*) oxidation state with two unpaired electrons described as being Fe-O antibonding π^*. Consequently, the oxygen is not a true oxo (O2-) ligand but more of the nature of an oxy (O-·) radical and the metal center has characteristics of a low-spin Fe(III) (Bathelt, Mulholland et al. 2005). The ferromagnetic quartet and the antiferromagnetic doublet state are quasi degenerate, as the two electrons on the ferryl moiety are ferromagnetically coupled while coupling between these electrons and the third unpaired electron is much weaker (Schöneboom, Lin et al. 2002).

It is widely accepted that CpdI extracts one electron from Trp191 to create the radical intermediated Trp191+, which finally gets its missing electron from Cytc (Sivaraja, Goodin et al. 1989; Miller, Vitello et al. 1995; Barrows, Bhaskar et al. 2004; Seifert, Pfister et al. 2005). As a consequence, its electron transfer process can be considered as hole transfer. The crystal structure of the yeast CcP/Cytc complex was determined at 2.3 Å resolution by Pelletier and Kraut in 1992 (Pelletier and Kraut 1992). Here, the distance between the iron centers of both hemes is about 27 Å and the distance between Trp191 and the iron of Cytc is about 21 Å. Furthermore, the sidechain of Trp191 is parallel to, and in van der Waals contact with, the imidazole ring of His175 being ligated to the heme. Many studies have focused on the support of the existence of the radical Trp191+. Among them are site mutagenesis experiments stating that the protein environment stabilizes Trp191+ (Musah and Goodin 1997; Musah, Jensen et al. 1997), as well as another study finding strong correlation between the experimentally determined stability of the Trp191+, the enzyme activity, and the calculated electrostatic stabilization of Trp191+ (Barrows, Bhaskar et al. 2004). Based on the crystal structure of the complex, Pelletier and Kraut proposed a σ-bond tunneling pathway from Cytc to Trp191 following the residues Ala194, Ala193 and Gly192 of CcP (shown in red in Figure 12) (Pelletier and Kraut 1992). A study based on photoinduced ET showing that the electron transfer from Cytc to Trp191+ is very rapid (Liu, Hahm et al. 1995), as well as more recent studies on the replacements of the proposed pathway by a ligand-binding channel (Rosenfeld, Hays et al. 2002; Hays Putnam, Lee et al. 2009) provide support for the proposed pathway. Furthermore, it is widely accepted that CpdI extracts one electron from Trp191 to create the radical intermediated Trp191+, which finally gets its missing electron from Cytc (Sivaraja, Goodin et al. 1989; Miller, Vitello et al. 1995; Musah and Goodin 1997; Musah, Jensen et al. 1997; Barrows, Bhaskar et al. 2004; Seifert, Pfister et al. 2005). Another possible pathway was proposed by Siddarth (Siddarth 1994), including residues Leu177, Gly178, Asn195, Asn196 and Val197 of CcP, and Met80, Ala81 and Phe82 of Cytc (shown in blue in Figure 12). Previous doubts on a single docking site of Cytc on CcP (Pappa, Tajbaksh et al. 1996; Leesch, Bujons et al. 2000; Nocek, Leesch et al.

2000) were dispelled by Guo et al. who designed a covalently cross-linked complex closely resembling the non-covalent complex determined by Pelletier and Kraut (Guo, Bhaskar et al. 2004). Stopped-flow kinetics of this covalent complex showed that it closely mimics the physiological electron transfer complex, indicating that there only exists a single docking site of Cytc relevant for its electron transfer.
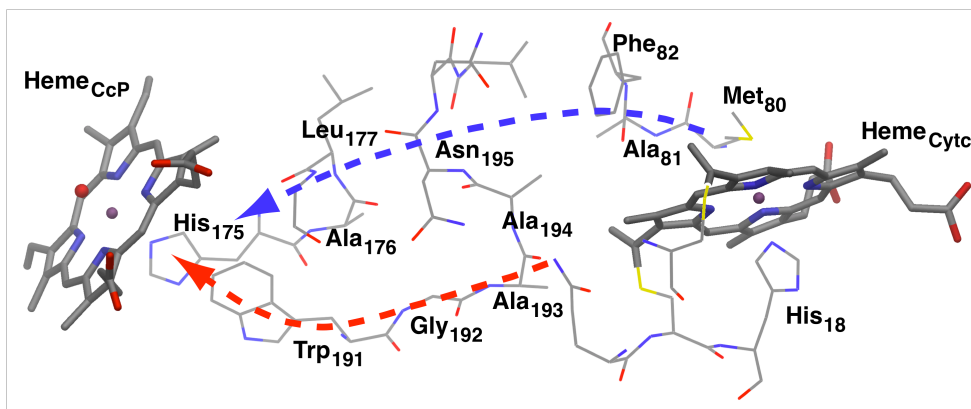

Figure 12: Electron transfer region of the CcP/Cytc complex. ET pathway proposed by Pelletier and Kraut is shown in red, ET pathway proposed by Siddarth is shown in blue.

Much experimental work was also done on detecting the rate of the CcP catalyzed reaction. The measured rates strongly depend on the ionic strength used in the experiments, resulting in rate constants ranging from about 250 s^{-1} (low ionic strength) to 2120 s^{-1} (high ionic strength) for yeast Cytc to CpdI using both stopped-flow (Summers and Erman 1988; Corin, Hake et al. 1993; Matthis and Erman 1995; Wang and Margoliash 1995; Miller 1996; Pappa, Tajbaksh et al. 1996; Guo, Bhaskar et al. 2004) and laser flash techniques (Hazzard, Poulos et al. 1987; Hazzard, McLendon et al. 1988; Hazzard, McLendon et al. 1988; Geren, Hahm et al. 1991) for both the covalent (Hazzard, Moench et al. 1988; Pappa, Tajbaksh et al. 1996) and noncovalent complexes (Hazzard, Poulos et al. 1987; Miller 1996; Guo, Bhaskar et al. 2004). The measurement of the actual rate of the pure ET process within this reaction still bears some difficulties as it is too fast for stopped-flow spectroscopy, thus being estimated to be greater than 5.0×10^4 s^{-1} (Geren, Hahm et al. 1991; Guo, Bhaskar et al. 2004).

In order to summarize, the pool of information available on this system as well as the confirmation of the true nature of Trp191 as localized bridge in the literature makes the CcP/Cytc complex to a good test system for the application of novel methodologies. Thus, in order to testify our main objective of developing a methodology for a quantitative study of ET in biological systems, we applied our novel QM/MM ET mapping approach together with electronic coupling calculations on derived pathways on this protein-protein complex and compared it against well-known information from the literature. The corresponding publication is given in section 3.5.

INTRODUCTION

2. OBJECTIVES

OBJECTIVES

The previous section has extensively discussed the most important achievements and difficulties in electron transfer calculations. Approaches to derive the parameters for theoretical rate constant, in particular the electronic coupling between the donor and acceptor, were depicted. The need for fundamental research on model systems in order to study these mechanisms and application to protein was discussed and some approaches were presented. Against this background, this section discloses the main objectives of this PhD thesis and points out in which particular publications they have been addressed.

The main objective of this PhD project was the development of a methodology for the quantitative study of ET in biological systems. From this we derived several sub-objectives which were addressed gradually throughout the thesis and can be stated as follows:

1 . Systematic review of the state-of-the-art QM/MM methods and their application to protein.

2 . Development and application of a novel method for electron transfer pathway mapping.
In particular, characterization of possible electron transfer pathways for a given system snapshot focusing on atomic detail, easy application and *ab initio* QM/MM level of theory.

3 . Electronic coupling calculations in model systems.
In particular, investigation of the influence of solvent and temperature effects on electronic coupling.

4 . Application of the newly developed electron transfer pathway mapping method in conjunction with calculations of electronic coupling and rate constants in real protein.
In particular, incorporation of conformational dynamics of the system into the study as well as calculation of kinetical and thermodynamical factors for the different electron transfer pathways obtained in order to compute the ET rate constant.

2.1. Systematic review of the state-of-the-art QM/MM methods and their application to protein.

This thesis is focused on studying electron transfer in proteins using mixed quantum and classical simulation techniques. Hence, the first objective of this thesis was to obtain an overview of the state-of-the-art QM/MM methods and their application to protein. In this context, we performed an extensive literature search of current methods and applications resulting in two reviews, presented in section 3.1.1 and 3.1.2.

2.2. Objective 1: Development and application of a novel method for electron transfer pathway mapping.

As described in the introduction, electron transfer (ET) processes belong to the simplest but most crucial reactions in biochemistry, being one of the main step in almost all enzymatic cycles. It might involve a relative short pathway, for example, from a substrate or cofactor directly bound in the vicinity of the acceptor group, or rather large electron transfer pathways, for example across protein-protein complexes, where the donor and acceptor might be considerably apart from each other. In any of these cases getting direct information of the electron pathway is not a trivial task. Such atomic and electronic detailed information, however, is very valuable in terms of a better understanding of the enzymatic cycle, which might lead, for example, into more efficient protein inhibitor design.

Hence, the second objective of this thesis was the development and application of a novel method for electron transfer pathway mapping. In this regard, we developed the QM/MM e-Pathway algorithm, based on *ab initio* QM/MM technology, described in section 3.2.

We proved the usefulness and applicability of our QM/MM e-Pathway approach through identification of important residues for ET in the P450cam/Pdx complex, being confirmed through the literature (see section 3.3). This work was furthermore presented as oral and poster communications at international conferences (see publications 9 and 11 in section 6)

2.3. Objective 2: Electronic coupling calculations in model systems.

Model systems, generally straightforward to engineer and apply, have been used for many ET, respectively electronic coupling studies. These systems have a clearly defined electron donor and acceptor as well as only a single possible ET pathway in between them.

For several systems it has been found that the electronic coupling depends on conformational dynamics. Hence, the third objective of this thesis was the study and test of applicability of electronic coupling on model systems for protein, considering the conformational dynamics of the system induced by temperature or solvent effects. In this regard, we present a combined quantum chemical/molecular

dynamics study on electronic coupling between tryptophan-based donor and acceptor in oligopeptides of variable length. We studied the dependence of electronic coupling values on conformational parameters and solvent effects. In particular, we calculated coupling values on a wealth of snapshots derived from MD calculations by applying semiempirical INDO/S as well as *ab initio* Hartree-Fock together with the generalized Mulliken-Hush approach (see section 3.4.1).

Moreover, we report a quantum chemistry and molecular dynamics study on the temperature dependence of electronic coupling in two short model oligopeptides. Here, the electron transfer parameters were estimated by using the semiempirical INDO/S method together with the charge fragment difference scheme computed on temperature-based trajectories derived from replica exchange MD calculations (see section 3.4.2).

Furthermore, we presented our findings as an oral presentation at an international conference (see publication 8 in section 6).

2.4. Objective 3: Application of the newly developed electron transfer pathway mapping method in conjunction with calculations of electronic coupling and rate constants in real protein.

Clearly, model systems are well suited to investigate the underlying species of ET and its dependence on varying parameters such as donor-acceptor distance or structural conformation. Nevertheless, at the end of the day, we are interested in understanding how real nature works. Consequently, the last and major objective of this dissertation was the application and evaluation of the newly developed electron transfer pathway mapping method together with electronic coupling and rate constants calculations in real protein.

This objective was addressed in section 3.5, where we report on the findings of the ET investigation on the CcP/Cytc complex. This protein couple provides a good test case for our purposes as it contains large complexity, i.e. large donor-acceptor distance, several proposed ET pathways and a bridge-localized state, all being well-characterized in the literature. In particular, we studied if our methodology, based on first principles and no parameters, is able to identify the true nature of Trp191 as localized bridge in the ET through electronic coupling calculations on pathways identified by our QM/MM e-Pathway approach. Moreover, we give estimated rate constants of both ET mechanisms, bridge-mediated and bridge-localized, respectively, in good agreement to experimental studies.

OBJECTIVES

3. THESIS PUBLICATIONS

THESIS PUBLICATIONS

3.1. QM/MM methods and applications to protein

THESIS PUBLICATIONS

3.1.1. Mixed quantum mechanism and molecular mechanics methods: Looking inside enzymes

Wallrapp F. H. and Guallar V.

THESIS PUBLICATIONS

Mixed quantum mechanics and molecular mechanics methods: Looking inside enzymes

Frank H. Wallrapp and Victor Guallar*

Mixed quantum mechanics and molecular mechanics (QM/MM) techniques, where the active site is treated with a robust ab initio quantum mechanics methodology and the remainder of the protein is modeled at the molecular mechanics level, offer a valuable tool for looking inside enzymes. We present here a short summary of recent QM/MM applications. We discuss the capabilities of this technology focusing in the main three steps when modeling enzymes: building a model, conformational sampling and mapping the chemical process. ©2010 JohnWiley & Sons, Ltd. WIREs Comput Mol Sci 2011 00 1–8 DOI: 10.1002/wcms.27
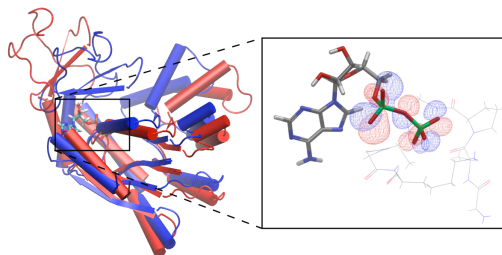
*Correspondence to: victor.guallar@bsc.es
Barcelona Supercomputing Center, Barcelona, Spain

Introduction

Understanding the mechanism of enzymes at atomic and electronic detail is of crucial importance in industrial catalysis and biomedical research. Computational modeling offers a very valuable approach to address this issue allowing electronic resolution by means of solving the Schrödinger equation. The modeling process mainly involves three steps: 1) building a model, 2) conformational sampling and 3) mapping the enzymatic chemical process. The first step opens the question on the choice of computational models: a quantum mechanics (QM) model of reduced size with only dozens of atoms held together by harmonic constraints, or a model of the entire protein (and surrounding solvent) using mixed quantum mechanics and classical mechanics (QM/MM) techniques. In the later, the reactive region of the active site is treated with a robust *ab initio* QM methodology, whereas the remainder of the protein is modeled at the MM level, by means of classical force fields.[1,2] All together, the computational cost is similar to that of QM methods alone. Furthermore, by using the correct protein constraints, QM/MM methods might reduce significantly the number of theoretical assays. Once the model is decided we need to perform conformational sampling due to inaccuracies of the initial coordinates (from

crystallography, NMR, homology modeling, etc.), or due to the possible coupling of the chemical event with low protein frequency modes. To account for this conformational sampling effectively one must dismiss the electronic degrees of freedom and perform classical force field (short runs might also include electronic degrees of freedom) molecular dynamics[3-6], Monte Carlo[7-9] techniques, proteins structure prediction algorithms[10,11], etc.

The last step is to map the enzymatic chemical process, for example bond breaking, electron transfer or electronic excitation events. Besides performing single point energy calculations or reaction coordinate analysis, we might also want to involve short (up to few nanoseconds) conformational sampling with the QM/MM Hamiltonian. The methodology used in the analysis of the chemical process will depend on the QM Hamiltonian chosen to describe the process. Less expensive QM methods will allow to run short molecular dynamics (MD) trajectories and to build potential of mean force (PMF) and free energy profiles (FEP) for the chemical process.[12,13] If a more expensive Hamiltonian is used, the study will usually narrow down to single point energies, geometry minimization and potential energy reactions coordinates.

In this review we summarize state of the art QM/MM applications on enzymatic systems. We will describe QM/MM studies addressing the three main steps introduced above: 1) choosing a model, 2) Conformational sampling and 3) mapping the chemical process, focusing on papers published in recent years.

Step 1: Choosing a model

In this section, we address the necessity of using QM/MM methods when dealing with enzymatic mechanisms. In QM/MM methods, most of the computational time corresponds to the evaluation of the QM Hamiltonian. Thus, as mentioned earlier, the computational cost of QM/MM methods is similar to that of QM methods alone. Additionally, modeling an active site with a reduced number of QM atoms in gas phase or a continuum medium, introduces several questions about the electrostatic screening (or influence) and flexibility of the system. Charge transfer, such as proton or electron transfer, will be badly modeled with a small QM model in gas phase. Adding a continuum dielectric will possibly improve the results from a gas phase environment but will not map all specific enzymatic electrostatic interactions present in the protein. Obviously, including more and more resi-

dues in the QM model will improve this description. In the limit of the QM region being the complete enzymatic system (or a large enough region) both methods, QM and QM/MM, should converge. Within this context, two recent studies have demonstrated that the convergence to the "full QM limit" is much faster when using combined QM/MM techniques than QM methods alone.[14,15] These studies apply increasing QM regions up to 1600 atoms and compare the results of QM/MM with those of only QM of the same region. Besides pointing out the significant faster convergence of QM/MM methods, both studies also agree in the necessity of applying large QM regions for a quantitative accurate result. Qualitative results, however, might be obtained at the QM/MM level of theory (not necessarily when only applying a QM model) for smaller regions.

The requisite of applying large QM regions is quite intriguing. The authors denote the need of including a minimum of 300 atoms in the QM region for accuracies of 2 kJ/mol. Of course this accuracy refers to comparison of a QM/MM model to a full QM study (of the entire system combining the MM and QM regions). This error should be added to the intrinsic QM error from the method and basis set. There are currently very few studies which treat quantum mechanically such a large number of atoms. Furthermore, the convergence is shown to be independent of the chosen method and basis set, although they do not perform an extensive study. Furthermore, the work by Fuxreiter et al.[14] evaluates the force error at the center of the QM region, a magnitude highly relevant for all QM/MM molecular dynamics studies. The authors conclude that forces may not fall below acceptable limits even for a quantum region with a radius of 9.0 Å. Thus, when computing free energies with QM/MM methods we should consider two types of errors, the errors on the potential surface and on its derivatives, which might vary differently with the QM region size.

While the necessity of using QM/MM molecular dynamics will be addressed below, we would like to introduce here another limitation of full QM models of reduced size. The lack of the enzymatic environment forces us to use harmonic constraints to hold together the system. These imposed constraints, together with the small nature of the system, makes it difficult to reproduce the enzymatic dynamics that might promote catalysis.[16] All these limitations in using reduced QM models is transforming the computational enzymology field, imposing QM/MM methodologies as the primary tool. Reduced QM models, however, might still be

2

       QM/MM methods: looking inside enzymes

quite useful for comparison analysis of the effect of the protein environment on the active site.

One of the limitations of QM/MM methods is the requirement of X-ray or NMR 3-dimensional coordinates (although producing an accurate QM reduced model does require such information as well). It is also possible to use closely related structures to produce an enzymatic model using comparative techniques, generally called homology modeling. However, when studying quantum mechanically a chemical process, small differences in the active site model might produce significant changes in the chemical energy profile. Thus, homology modeling in biochemistry should be restricted to systems with a very high level of sequence identity in the active site. Currently, most of the studies that combine homology modeling and QM/MM methods involve docking and binding energy refinements.[17-19]

Once the model is chosen, it is mandatory to perform an exhaustive inspection of the coordinates. Crystal contacts might largely affect an exposed active site. Protonation states and side chain dihedrals might be not clear from the experimental initial coordinates. Furthermore, the inclusion/deletion of residues in the quantum region might largely affect the point of view of some particular protonation state.[20]

Step 2: Conformational sampling

After building the model, it might be necessary to perform conformational sampling. We might need to eliminate crystal contact artifacts or sample missing residues and loops. Additionally, the coordinates might describe an enzymatic intermediate of no specific interest for the researcher. Therefore, it might be necessary to add some cofactor, protons or electrons to model the active state properly. As a consequence, these additions might result in conformational changes that need to be described. Furthermore, it is also possible that the enzymatic mechanism is enhanced by means of some low frequency collective motion of the system. If we listen to the latest talks by David E Shaw[4,21] or Klaus Schulten[22], mapping this collective motion is not a simple process on the order of tenths of nanoseconds. Ideally one would like to perform protein dynamics at the microsecond or even millisecond time scale.[23] While running hundreds of nanoseconds of (classical force field!) MD might be a trivial task with recent advances in molecular dynamics and computer power[4,6,22], deriving microseconds or millisec-

onds of MD will not be available as a routine job in the near future.

Of course we are not implying here that all enzymes require such an extensive conformational search. We believe that careful examination, however, should be produced prior to an expensive QM/MM study. Such a preexamination, probing the enzyme plasticity, might be produced by several alternative techniques to MD: normal modes, loop prediction techniques, etc.

Inspecting the current QM/M studies, we find very little conformational sampling coupled (or prior) to the chemical event analysis. The longest exploration times we find are on the order of tenths of nanoseconds. Most of the studies perform some equilibration and then choose the last structure, or some "suitable" structure, to run the chemical process analysis. However there are solid evidences that the potential energy from a reaction coordinate is significantly different when choosing several snapshots from a short conformational sampling. Early work by McCammon et al.[24] already established a range from 10-16 kcal/mol in the energy barrier when inspecting 10 different snapshots along a 1 ns MM trajectory in acetylcholinesterase. In this short simulation time, however, the role of the catalytic triad was very consistent. Similar results have been obtained by Warshel et al.[25] on the Ras-Gap complex. Mulholland et al. recently performed 1 ns classical MD on the fatty acid amide hydrolase followed by a clustering analysis to obtain different snapshots with energy barrier differences of 15 kcal/mol.[26] These studies do not involve long conformational sampling and already find significant differences. We find some longer MD simulations when the crystal structure is missing some loop or side chains. Studies by Garcia-Viloca et al.[27], for example, have run 7 ns of MD in Glutamate Racemase to sample a missing loop before performing the actual QM/MM study. Zhang et al performed 25 ns of classical MD on Sir2 but used only a single structure (the snapshot at 4 ns) for the QM/MM analysis.[28]

Recent extensive molecular dynamics studies by Carloni et al. on HIV-1 protease[29] and on the anthrax lethal factor[30] indicate that not every active site is being affected by large-scale protein motions. In HIV-1 protease the authors observe substrate-protein interactions, mechanically assisted by concerted movements in the complex. In the anthrax lethal factor, however, the authors do not observe large-scale protein motions affecting the active site within a 50 ns MD simulation. Still, are 50 ns enough? It seems difficult to address

these differences without performing the actual calculations. Alternative methods to produce large scale sampling coupled to QM/MM techniques include several Monte Carlo techniques.[31-34] But in principle any all-atom conformational sampling technique, such as robotics algorithms[35], normal modes[36], etc could be used to perform a pre-QM/MM analysis. As these methods become more user friendly and computationally inexpensive, we expect to see better conformational sampling prior to the analysis of the chemical process.

Step 3: Mapping chemical process

Once the model is set up, with or without foregoing conformational sampling, we are able to map the chemical process. As mentioned before, nowadays it is impossible to map slow large-scale motions with the QM/MM hamiltonian; the level of theory of the QM region will determine to which extent we can couple faster nuclear motion with the electronic degrees of freedom. With the current available computational power, we are able to apply two main distinguishable categories. First, perform stationary point searches or multiple single point calculations usually associated to expensive QM calculations (often of higher accuracy) and second, run short molecular dynamics with a less expensive QM description.

The first category aims to reduce the number of QM evaluations, mostly performing energy minimizations, transition state searches or single point energy, for a total of few hundreds energy and gradient calculations. However, due to the miss of possibly important nuclear degrees of freedom and entropic effects, this energy might be not a good estimation to the actual free energy.[16,26] However, in many systems a simple minimization reaction coordinate might give quite good results.[37] Recent applications are studies on excitation energies of rhodopsin[38] or investigations on the reactive geometry of the HIV-1 protease.[39] Thiel et al. applied QM/MM methods on the DFT level of theory in studies on the stability of redox electronmers[40] as well as in studies on NMR chemical shifts.[41] Furthermore there exist approaches of applying divide-and-conquer algorithms firstly introduced by Yang et al.[42], which are able to track large QM regions by means of consecutively calculating only fragments until the total QM region converges.[43-46] Application of these methods, in the context of higher order QM, are HF calculations on 11 proteins of only a few hundred atoms[46], DFT calculations for structural refinement of the pho-

tosystem II[45] or MP2 calculations of a synthetic protein[47], for example. An alternative fragment based approach has recently being applied to the bovine pancreatic trypsin inhibitor in salvation by Gao et al.[48] Very recently, Guallar and Wallrapp introduced the QM/MM e-Pathway approach.[49] They applied extensive HF calculations on large electron transfer regions in the cytochrome P450cam/Putidaredoxin and cytochrome c peroxidase/cytochrome c complexes to identify key amino acids within the respective electron transfer pathways, see Fig. 1.[49,50]

Within the second category of QM calculations are those which perform ab initio molecular dynamics, usually by means of semiempirical methods. Short QM/MM MD productions allow for a better coupling of nuclear and electronic degrees of freedom, mapping high frequency motions capable of promoting the chemical event.[16] Again, there exist several approaches to map the chemical processes with QM/MM MD, from which we depict few. Recent studies used transition path sampling techniques aiming to optimize the true barrier passage.[51-54] Extensive sampling is also done by Rothlisberger et al. performing several statistically independent QM/MM MD trajectories of DNA photolyase bound to double stranded DNA[55] or in a microscopic pKa analysis of cytochrome c oxidase done by Cui et al., where they combine QM/MM with a thermodynamic integration protocol.[56] Further reaction coordinate search have been introduced by Liu et al. and Field et al. applying modifications of the nudged elastic band (NEB) methods together with semiempirical methods.[57-59] We find also several potential of mean force studies, such as proton transfer analysis in Carbonic Anhydrase by Cui et al..[60,61], and in Phosphotriesteras and Bacteriorhodopsin by Gao et al..[62,63] Initial efforts applying replica exchange MD have also been applied to small peptides.[64] Important to mention is also the vast field of Car-Parrinello MD (CPMD)[65] with recent applications in studies on CphA[66] or cyclin-dependent kinase[67], for example. Furthermore, there exist several studies combining CPMD with metadynamics, such as in the reaction mechanism of catalase[68] and orotidine-5'-monophosphate decarboxylase.[69] Finally we should mention recent work by Jorgensen et al. using Monte Carlo techniques instead of molecular dynamics based techniques.[33] Due to its efficiency in exploring large conformational space, one would expect the expansion of these Monte Carlo techniques in QM/MM studies in the near future.
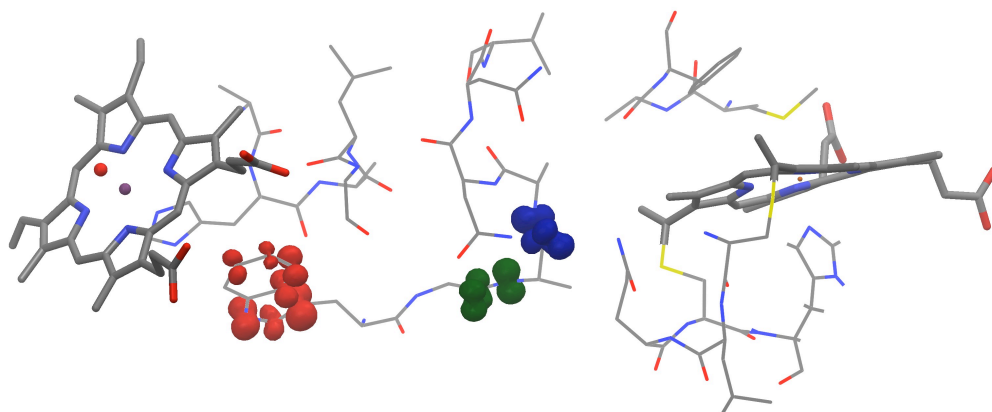
4

Figure 1. Electron transfer pathway obtained by the QM/MM e-pathway procedure in CcP/cyt c complex.

Conclusion

QM/MM methods are gaining real importance in mapping biochemical processes. They allow for an atomic and electronic detailed analysis of the mechanism. We have not addressed here the technical aspects of the methodology, dealing with the QM to MM boundary, etc. Good news is that the different codes and approximations seem to have similar level of accuracy.[70,71] However, recent studies seem to indicate the necessity of using significant larger QM regions in order to reach an accuracy of 2-10 kJ/mol (an error that could be partially cancelled when studying energy differences along a reaction coordinate). This error needs to be added on top of the error of the *ab initio* method used, usually another 5-10 kJ/mol. Thus, it does not seem reasonable to focus in reproducing rate constants and chemical properties below these limits.

Moreover most QM/MM applications nowadays are missing conformational sampling. Recent advances have made real progress in coupling fast nuclear motion with electronic degrees of freedom. The next step should involve a better description of slow large-scale protein dynamics. In these regards the field is probably focusing too much in molecular dynamics and should explore alternative methods such Monte Carlo or protein structure prediction techniques. An alternative could be to perform a chemical analysis in a slightly equilibrated structure (the actual normal procedure). From this analysis we can obtain the modes that could promote the chemical reaction, those modes highly coupled in the transition state. Then we could perform some normal mode analysis, for example using Gaussian network modeling[72] and check for possible collective protein motions affecting these modes. Such a quick test could inform us about the possible influence of the enzyme in our initial energy profile.

Overall QM/MM methods seem to be the choice for theoretical modelling of biochemical events. The numerous applications summarized here indicate that quantum chemistry has reached enough maturity to describe most electronic processes. Once we achieve a better level of large scale conformational sampling they will offer an excellent tool for *looking inside enzymes*.

Further Reading

1. Gao, J. L. & Truhlar, D. G. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **53**, 467-505 (2002).
2. Senn, H. M. & Thiel, W. QM/MM studies of enzymes. *Curr. Opin. Chem. Biol.* **11**, 182-187 (2007).
3. Senn, H. M. & Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed. Eng.* **48**, 1198-1229 (2009).

References

[1] Warshel, A. & Levitt, M. A. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **103**, 227-249 (1976).

2   Warshel, A. & Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* **94**, 5612-5625, doi:10.1021/ja00771a014 (1972).

3   Brooks, B. & Karplus, M. Harmonic Dynamics of Proteins: Normal Modes and Fluctuations in Bovine Pancreatic Trypsin Inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 6571-6575 (1983).

4   Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O. & Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struc. Biol.* **19**, 120-127, doi:10.1016/j.sbi.2009.03.004 (2009).

5   Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L. & Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-1802 (2005).

6   Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435-447, doi:10.1021/ct700301q (2008).

7   Li, Z. Q. & Scheraga, H. A. Monte Carlo -minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 2633-2636 (1987).

8   Ulmschneider, J. P. & Jorgensen, W. L. Monte Carlo backbone sampling for polypeptides with variable bond angles and dihedral angles using concerted rotations and a Gaussian bias. *J. Chem. Phys.* **118**, 4261-4271, doi:10.1063/1.1542611 (2003).

9   Evans, D. A. & Wales, D. J. Free energy landscapes of model peptides and proteins. *J. Chem. Phys.* **118**, 3891-3897 (2003).

10   Borrelli, K. W., Vitalis, A., Alcantara, R. & Guallar, V. PELE: Protein energy landscape exploration. A novel Monte Carlo based technique. *J. Chem. Theory Comput.* **1**, 1304-1311 (2005).

11   Jacobson, M. P., Pincus, D. L., Rapp, C. S., Tyler, J. F., Honig, B., Shaw, D. E. & Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**, 351-367 (2004).

12   Ruiz-Pernia, J. J., Silla, E., Tunon, I., Marti, S. & Moliner, V. Hybrid QM/MM potentials of mean force with interpolated corrections. *J. Phys. Chem. B* **108**, 8427-8433, doi:10.1021/jp049633g (2004).

13   Hu, H., Lu, Z. Y. & Yang, W. T. QM/MM minimum free-energy path: Methodology and application to triosephosphate isomerase. *J. Chem. Theory Comput.* **3**, 390-406, doi:10.1021/ct600240y (2007).

14   Solt, I., Kulhanek, P., Simon, I., Winfield, S., Payne, M. C., Csanyi, G. & Fuxreiter, M. Evaluating Boundary Dependent Errors in QM/MM Simulations. *J. Phys. Chem. B* **113**, 5728-5735, doi:10.1021/jp807277r (2009).

15   Sumowski, C. V. & Ochsenfeld, C. A Convergence Study of QM/MM Isomerization Energies with the Selected Size of the QM Region for Peptidic Systems. *J. Phys. Chem. A* **113**, 11734-11741 (2009).

16   Schwartz, S. D. & Schramm, V. L. Enzymatic transition states and dynamic motion in barrier crossing. *Nat. Chem. Biol.* **5**, 551-558 (2009).

17   Sgrignani, J., Bonaccini, C., Grazioso, G., Chioccioli, M., Cavalli, A. & Gratteri, P. Insights into Docking and Scoring Neuronal alpha 4 beta 2 Nicotinic Receptor Agonists Using Molecular Dynamics Simulations and QM/MM Calculations. *J. Comput. Chem.* **30**, 2443-2454, doi:10.1002/jcc.21251 (2009).

18   Mukherjee, P., Desai, P. V., Srivastava, A., Tekwani, B. L. & Avery, M. A. Probing the structures of leishmanial farnesyl pyrophosphate synthases: Homology modeling and docking studies. *J. Chem. Info. Model.* **48**, 1026-1040, doi:10.1021/ci700355z (2008).

19   Abdel-Halim, H., Hanrahan, J. R., Hibbs, D. E., Johnston, G. A. R. & Chebib, M. A molecular basis for agonist and antagonist actions at GABA(C) receptors. *Chem. Biol. Drug Des.* **71**, 306-327, doi:10.1111/j.1747.0285.2008.00642.x (2008).

20   Guallar, V. & Olsen, B. The Role of the Heme Propionates in Heme Biochemistry. *J. Inorg. Biochem.* **100**, 755-760 (2006).

21   Maragakis, P., Lindorff-Larsen, K., Eastwood, M. P., Dror, R. O., Klepeis, J. L., Arkin, I. T., Jensen, M. O., Xu, H. F., Trbovic, N., Friesner, R. A. *et al.* Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. *J. Phys. Chem. B* **112**, 6155-6158, doi:10.1021/jp077018h (2008).

22   Lee, E. H., Hsin, J., Sotomayor, M., Comellas, G. & Schulten, K. Discovery Through the Computational Microscope. *Structure* **17**, 1295-1306, doi:10.1016/j.str.2009.09.001 (2009).

23   Karplus, M. & Kuriyan, J. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 6679-6685, doi:10.1073/pnas.0408930102 (2005).

24   Zhang, Y. K., Kua, J. & McCammon, J. A. Influence of structural fluctuation on enzyme reaction energy barriers in combined quantum mechanical/molecular mechanical studies. *J. Phys. Chem. B* **107**, 4459-4463 (2003).

25   Klahn, M., Braun-Sand, S., Rosta, E. & Warshel, A. On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. *J. Phys. Chem. B* **109**, 15645-15650, doi:10.1021/jp0521757 (2005).

26   Lodola, A., Mor, M., Zurek, J., Tarzia, G., Piomelli, D., Harvey, J. N. & Mulholland, A. J. Conformational effects in enzyme catalysis: Reaction via a high energy conformation in fatty acid amide hydrolase. *Biophys. J.* **92**, 20-22, doi:10.1529/biophysj.106.098434 (2007).

27   Puig, E., Mixcoha, E., Garcia-Viloca, M., Gonzalez-Lafont, A. & Lluch, J. M. How the Substrate D-Glutamate Drives the Catalytic Action of Bacillus

subtilis Glutamate Racemase. *J. Am. Chem. Soc.* **131**, 3509-3521, doi:10.1021/ja806012h (2009).

28 Hu, P., Wang, S. & Zhang, Y. Highly Dissociative and Concerted Mechanism for the Nicotinamide Cleavage Reaction in Sir2Tm Enzyme Suggested by Ab Initio QM/MM Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **130**, 16721-16728 (2008).

29 Carnevale, V., Raugei, S., Micheletti, C. & Carloni, P. Large-scale motions and electrostatic properties of furin and HIV-1 protease. *J. Phys. Chem. A* **111**, 12327-12332, doi:10.1021/jp0751716 (2007).

30 Hong, R., Magistrato, A. & Carloni, P. Anthrax Lethal Factor Investigated by Molecular Simulations. *J. Chem. Theory Comput.* **4**, 1745-1756, doi:10.1021/ct8001877 (2008).

31 Guallar, V., Lu, C., Borrelli, K., Egawa, T. & Yeh, S.-R. Ligand Migration in the Truncated Hemoglobin-II from Mycobacterium tuberculosis: The Role of G8 Tryptophan. *J. Biol. Chem.* **284**, 3106-3116, doi:10.1074/jbc.M806183200 (2009).

32 Alcantara, R. E., Xu, C., Spiro, T. G. & Guallar, V. A quantum-chemical picture of hemoglobin affinity. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18451 (2007).

33 Alexandrova, A. N., Ro¨thlisberger, D., Baker, D. & Jorgensen, W. L. Catalytic Mechanism and Performance of Computationally Designed Enzymes for Kemp Elimination. *J. Am. Chem. Soc.* **130**, 15907-15915 (2008).

34 Tubert-Brohman, I., Acevedo, O. & Jorgensen, W. L. Elucidation of hydrolysis mechanisms for fatty acid amide hydrolase and its Lys142Ala variant via QM/MM simulations. *J. Am. Chem. Soc.* **128**, 16904-16913, doi:10.1021/ja065863s (2006).

35 Verbitsky, G., Nussinov, R. & Wolfson, H. Flexible structural comparison allowing hinge-bending, swiveling motions. *Proteins* **34**, 232-254 (1999).

36 Rueda, M., Bottegoni, G. & Abagyan, R. Consistent Improvement of Cross-Docking Results Using Binding Site Ensembles Generated with Elastic Network Normal Modes. *J. Chem. Info. Model.* **49**, 716-725, doi:doi:10.1021/ci8003732 (2009).

37 Friesner, R. A. & Guallar, V. Ab Initio Quantum Chemical and Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis. *Annu. Rev. Phys. Chem.* **56**, 389-427 (2005).

38 Wanko, M., Hoffmann, M., Frauenheim, T. & Elstner, M. Computational photochemistry of retinal proteins. *J. Comput.-Aid. Mol. Design* **20**, 511-518 (2006).

39 Carnevale, V., Raugei, S., Neri, M., Pantano, S., Micheletti, C. & Carloni, P. Multi-scale modeling of HIV-1 proteins. *J. Mol. Struct. (Theochem)* **898**, 97-105 (2009).

40 Schoneboom, J. C., Cohen, S., Lin, H., Shaik, S. & Thiel, W. Quantum mechanical/molecular mechanical investigation of the mechanism of C-H hydroxylation of camphor by cytochrome P450(cam): Theory supports a

two-state rebound mechanism. *J. Am. Chem. Soc.* **126**, 4017-4034 (2004).

41 Geethalakshmi, K. R., Waller, M. P., Thiel, W. & Bu¨hl, M. 51V NMR Chemical Shifts Calculated from QM/MM Models of Peroxo Forms of Vanadium Haloperoxidases. *J. Phys. Chem. B* **113**, 4456-4465 (2009).

42 Yang, W. Direct calculation of electron density in density-functional theory. *Phys. Rev. Lett.* **66**, 1438 (1991).

43 Kitaura, K., Ikeo, E., Asada, T., Nakano, T. & Uebayasi, M. Fragment molecular orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.* **313**, 701-706 (1999).

44 Kiyota, Y., Hasegawa, J.-Y., Fujimoto, K., Swerts, B. & Nakatsuji, H. A multicore QM/MM approach for the geometry optimization of chromophore aggregate in protein. *J. Comput. Chem.* **30**, 1351-1359 (2009).

45 Sproviero, E., Newcomer, M., Gascón, J., Batista, E., Brudvig, G. & Batista, V. The MoD-QM/MM methodology for structural refinement of photosystem II and other biological macromolecules. *Photosynth. Res.* **102**, 455-470 (2009).

46 He, X. & Merz, K. M. Divide and Conquer Hartree−Fock Calculations on Proteins. *J. Chem. Theory Comput.* **6**, 405-411, doi:10.1021/ct9006635 (2010).

47 Fedorov, D. G., Ishimura, K., Ishida, T., Kitaura, K., Pulay, P. & Nagase, S. Accuracy of the three-body fragment molecular orbital method applied to Moeller-Plesset perturbation theory. *J. Comput. Chem.* **28**, 1476-1484 (2007).

48 Xie, W., Orozco, M., Truhlar, D. G. & Gao, J. X-Pol Potential: An Electronic Structure-Based Force Field for Molecular Dynamics Simulation of a Solvated Protein in Water. *J. Chem. Theory Comput.* **5**, 459-467, doi:10.1021/ct800239q (2009).

49 Wallrapp, F., Masone, D. & Guallar, V. Electron Transfer in the P450cam/PDX Complex. The QM/MM e-Pathway. *J. Phys. Chem. A* **112**, 12989-12994 (2008).

50 Guallar, V. & Wallrapp, F. Mapping protein electron transfer pathways with QM/MM methods. *J. R. Soc. Interface* **5**, 233-239 (2008).

51 Ke, Z., Wang, S., Xie, D. & Zhang, Y. Born−Oppenheimer ab Initio QM/MM Molecular Dynamics Simulations of the Hydrolysis Reaction Catalyzed by Protein Arginine Deiminase 4. *J. Phys. Chem. B* **113**, 16705-16710 (2009).

52 Liu, J., Zhang, Y. & Zhan, C.-G. Reaction Pathway and Free-Energy Barrier for Reactivation of Dimethylphosphoryl-Inhibited Human Acetylcholinesterase. *J. Phys. Chem. B* **113**, 16226-16236 (2009).

53 Crehuet, R. & Field, M. J. A Transition Path Sampling Study of the Reaction Catalyzed by the Enzyme Chorismate Mutase. *J. Phys. Chem. B* **111**, 5708-5718 (2007).

54 Quaytman, S. L. & Schwartz, S. D. Comparison Studies of the Human Heart and Bacillus stearothermophilus Lactate Dehydrogrease by Transition Path Sampling. *J. Phys. Chem. A* **113**, 1892-1897 (2009).

55 Masson, F., Laino, T., Rothlisberger, U. & Hutter, J. r. A QM/MM Investigation of Thymine Dimer Radical Anion Splitting Catalyzed by DNA Photolyase13. *ChemPhysChem* **10**, 400-410 (2009).

56 Ghosh, N., Prat-Resina, X., Gunner, M. R. & Cui, Q. Microscopic pKa Analysis of Glu286 in Cytochrome c Oxidase (Rhodobacter sphaeroides): Toward a Calibrated Molecular Model. *Biochem.* **48**, 2468-2485 (2009).

57 Xie, L., Liu, H. & Yang, W. Adapting the nudged elastic band method for determining minimum-energy paths of chemical reactions in enzymes. *J. Chem. Phys.* **120**, 8039-8052 (2004).

58 Galvan, I. F. & Field, M. J. Improving the efficiency of the NEB reaction path finding algorithm. *J. Comput. Chem.* **29**, 139-143 (2008).

59 Lelimousin, M., Adam, V., Nienhaus, G. U., Bourgeois, D. & Field, M. J. Photoconversion of the Fluorescent Protein EosFP: A Hybrid Potential Simulation Study Reveals Intersystem Crossings. *J. Am. Chem. Soc.* **131**, 16814-16823 (2009).

60 Riccardi, D., Konig, P., Guo, H. & Cui, Q. Proton Transfer in Carbonic Anhydrase Is Controlled by Electrostatics Rather than the Orientation of the Acceptor. *Biochem.* **47**, 2369-2378 (2008).

61 Riccardi, D., Schaefer, P., Yang, Y., Yu, H., Ghosh, N., Prat-Resina, X., Konig, P., Li, G., Xu, D., Guo, H. *et al.* Development of Effective Quantum Mechanical/Molecular Mechanical (QM/MM) Methods for Complex Biological Processes. *J. Phys. Chem. B* **110**, 6458-6469 (2006).

62 Rajamani, R. & Gao, J. Combined QM/MM study of the opsin shift in bacteriorhodopsin. *J. Comput. Chem.* **23**, 96-105, doi:10.1002/jcc.1159 (2002).

63 Wong, K.-Y. & Gao, J. The Reaction Mechanism of Paraoxon Hydrolysis by Phosphotriesterase from Combined QM/MM Simulations *Biochem.* **46**, 13352-13369, doi:10.1021/bi700460c (2007).

64 Seabra, G. M., Walker, R. C. & Roitberg, A. E. in *Solvation Effects on Molecules and Biomolecules* Vol. 6 *Challenges and Advances in Computational Chemistry and Physics*   1-12 (Springer Netherlands, Dordrecht, 2008).

65 Car, R. & Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **55**, 2471 (1985).

66 Simona, F., Magistrato, A., Dal Peraro, M., Cavalli, A., Vila, A. J. & Carloni, P. Common Mechanistic Features among Metallo-β-lactamases. *J. Biol. Chem.* **284**, 28164-28171 (2009).

67 De Vivo, M., Cavalli, A., Carloni, P. & Recanatini, M. Computational Study of the Phosphoryl Transfer Catalyzed by a Cyclin-Dependent Kinase. *Chem. Eur. J.* **13**, 8437-8444 (2007).

68 Alfonso-Prieto, M., Biarne´s, X., Vidossich, P. & Rovira, C. The Molecular Mechanism of the Catalase Reaction. *J. Am. Chem. Soc.* **131**, 11751-11761 (2009).

69 Stanton, C. L., Kuo, I. F. W., Mundy, C. J., Laino, T. & Houk, K. N. QM/MM Metadynamics Study of the Direct Decarboxylation Mechanism for Orotidine-5'-monophosphate Decarboxylase Using Two Different QM Regions: Acceleration Too Small To Explain Rate of Enzyme Catalysis. *J. Phys. Chem. B* **111**, 12573-12581 (2007).

70 Amara, P. & Field, M. J. Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor. Chem. Acc.* **109**, 43-52 (2003).

71 Altun, A., Guallar, V., Friesner, R. A., Shaik, S. & Thiel, W. The effect of heme environment on the hydrogen abstraction reaction of camphor in P450(cam) catalysis: A QM/MM study. *J. Am. Chem. Soc.* **128**, 3924-3925 (2006).

72 Bahar, I., Atilgan, A. R. & Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* **2**, 173-181 (1997).

8

3.1.2. QM/MM methods: Looking inside heme proteins biochemistry

Guallar V. and Wallrapp F. H.

THESIS PUBLICATIONS

Guallar V, Wallrapp FH. QM/MM methods: looking inside heme proteins biochemistry. Biophys Chem. 2010; 149(1-2): 1-11.

THESIS PUBLICATIONS

3.2. Mapping protein electron transfer pathways with QM/MM methods.

Guallar, V. and Wallrapp F.

Guallar V, Wallrapp F. Mapping protein electron transfer pathways with QM/MM methods. J R Soc Interface. 2008; 5(Suppl 3): S233-9.

3.3. Electron Transfer in the P450cam/PDX Complex. The QM/MM e-Pathway.

Wallrapp, F., Masone, D. and Guallar, V.

Wallrapp F, Masone D, Guallar V. Electron transfer in the P450cam/PDX complex. The QM/MM e-pathway. J Phys Chem A. 2008; 112(50): 12989-94.

Supplementary Material

Electron Transfer in the P450cam/PDX Complex. The QM/MM e-Pathway.

Frank Wallrapp, Diego Masone and Victor Guallar
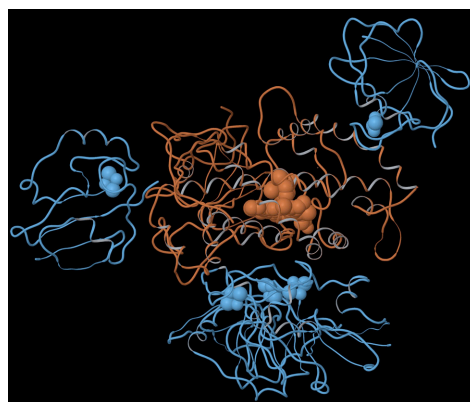
Figure S1. Representative structures of the best 10 docking results of P450 (brown) and PDX (blue) without the explicit -2 charge in the active site of PDX. The heme group and the iron-sulphur cluster are shown in a spacefill representation.
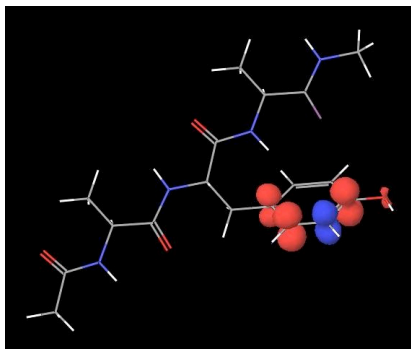
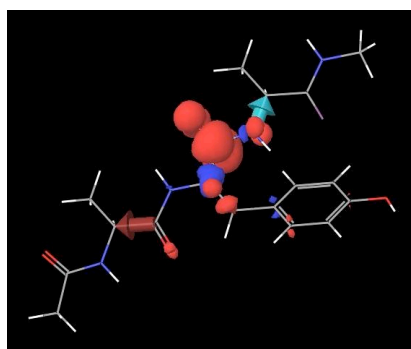**Figure S2.** Spin density plot of Ala-Tyr-Ala in vacuum having one extra electron injected, applying full QM.

Figure S3. Spin density plot of Ala-Tyr-Ala in vacuum having one extra electron injected, applying QM/MM with frozen orbital boundary type. The red and blue arrows on the backbone indicate the frozen orbitals.
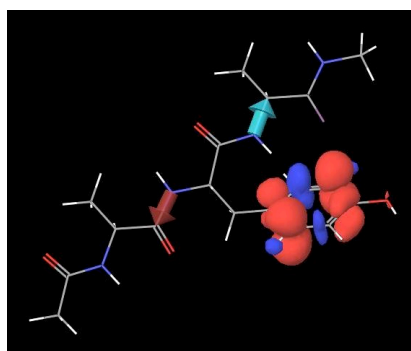

Figure S4. Spin density plot of Ala-Tyr-Ala in vacuum having one extra electron injected, applying QM/MM with H-Cap boundary type. The red and blue arrows on the backbone indicate the H-Caps.
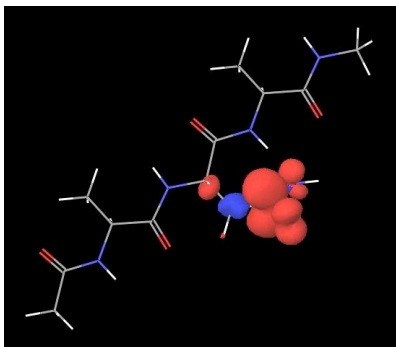
Figure S5. Spin density plot of Ala-Asn-Ala in vacuum having one extra electron injected, applying full QM.
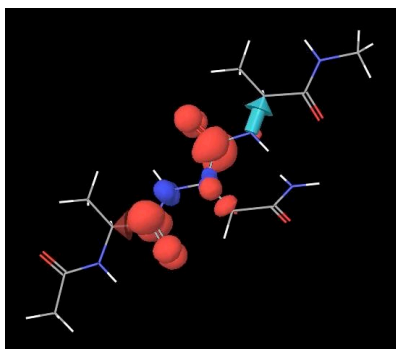

Figure S6. Spin density plot of Ala-Asn-Ala in vacuum having one extra electron injected, applying QM/MM with frozen orbital boundary type. The red and blue arrows on the backbone indicate the frozen orbitals.
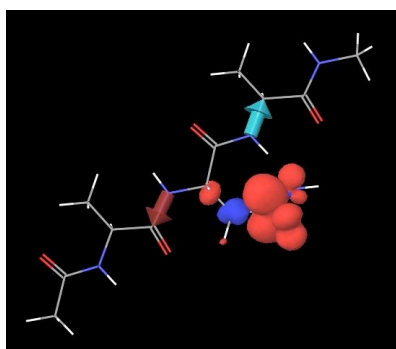

Figure S7. Spin density plot of Ala-Asn-Ala in vacuum having one extra electron injected, applying QM/MM with H-Cap boundary type. The red and blue arrows on the backbone indicate the H-Caps.

3.4.   Electronic coupling calculations in model systems

3.4.1. Solvent effects on donor-acceptor couplings in peptides. A combined QM and MD study.

Wallrapp, F., Voityuk. A. and Guallar, V.

THESIS PUBLICATIONS

Wallrapp F, Voityuk A, Guallar V. Solvent effects on donor-acceptor couplings in peptides. A combined QM and MD study. J Chem Theory Comput. 2009; 5(12): 3312-3320.

THESIS PUBLICATIONS

Supporting Information

**Solvent effects on donor-acceptor coupling in peptides.
A combind QM and MD study**

Frank Wallrapp, Alexander Voityuk, Victor Guallar

S1: Choosing step size of MD

We modeled Trp-(Pro)2-Trp in gas phase and ran MD of 10 ps with 1 fs time steps to get an overview of the behavior of the mobility of the oligopeptides. Figure S1 shows the donor acceptor distance as well as its autocorrelation (ACF) of the trajectory. The ACF oscillates about every 5 ps and has a halftime of about 1 ps indicating that the system only undergoes significant changes in respect to the donor acceptor distance within these 1 ps. Based on these results we chose a step size of 1 ps for the MD simulations carried out on the polypeptides within this study.
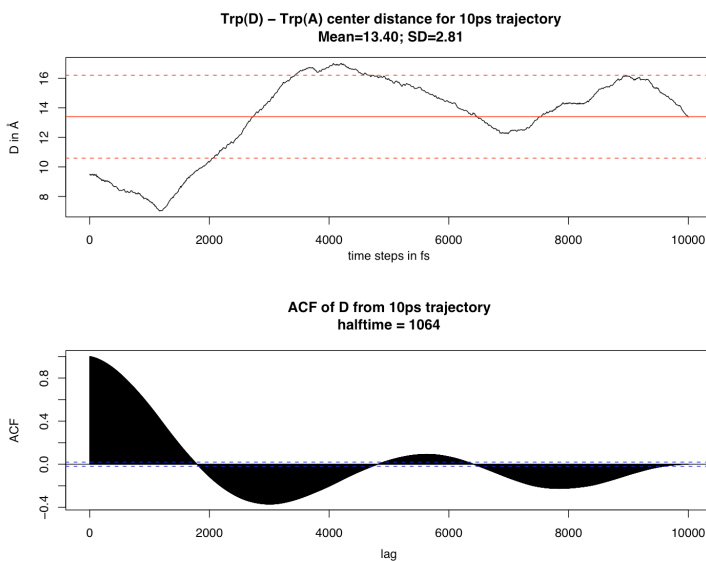
Figure S1: Donor acceptor distance and its autocorrelation in 10 ps trajectory of Trp-(Pro)2-Trp with time step of 1 fs.

S2: PLS-R analysis

We applied PLS-R on all conformations in gas phase as well as solvated-conformation-only to analyze the impact of the different conformational parameters on the coupling. Data matrix X is given by derived conformational parameters d, p and r and Y is V_{DA}^2 for each respective trajectory. For normalization we applied UV scaling on X, assuming normal distribution, and $\log_{10}$ on Y due to its large range. For each oligopeptide system we provide the model statistics including the cumulative Q2 value indicating the predictive power of the model as well as the score scatter plot of PC1 against PC2 and the loading plots of all PCs. Within the score plot green indicates conformations with low, yellow medium and red high coupling values.

Trp-(Pro)1-Trp (gas phase)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.192	0.192	0.363	0.363	0.363	0.363
2	0.102	0.294	0.117	0.48	0.182	0.478
3	0.0978	0.392	0.0189	0.499	0.0343	0.496

Score plot PC1 vs. PC2
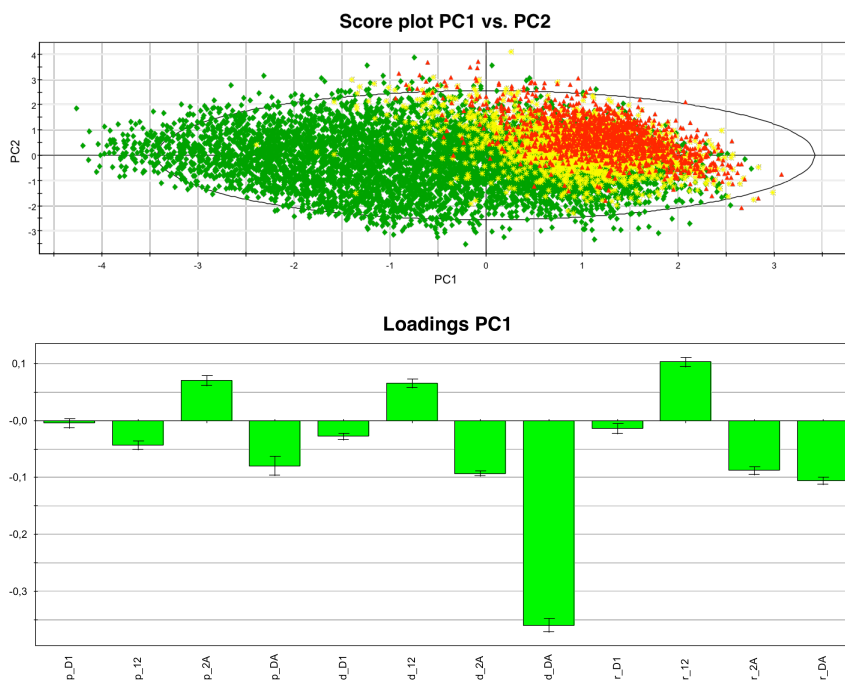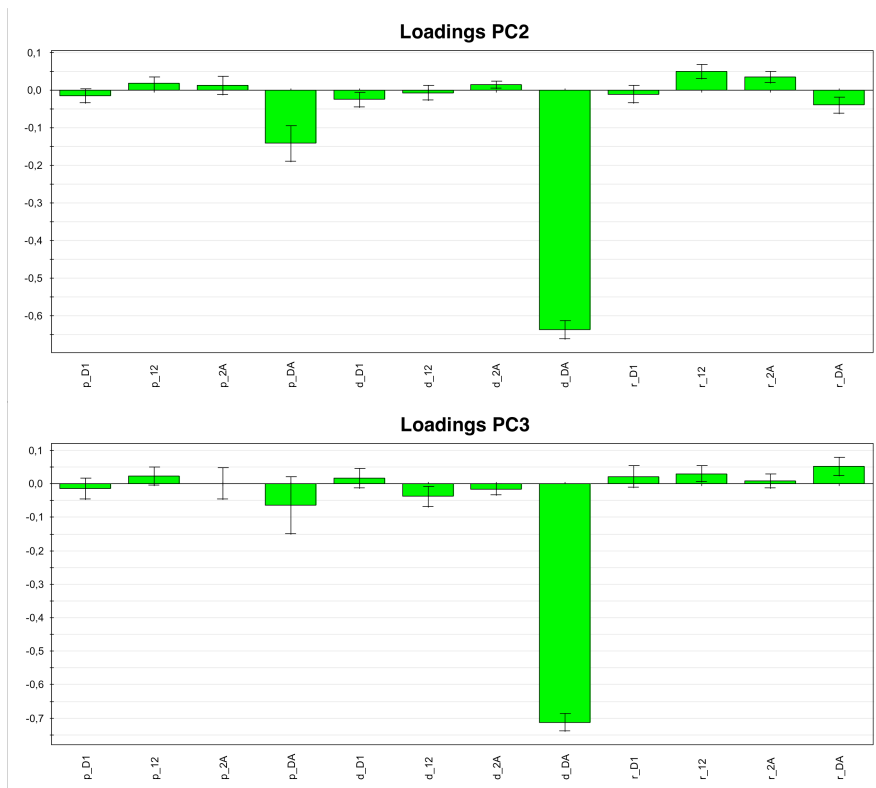
Loadings PC1

Trp-(Pro)1-Trp (solvated-conformation-only)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.238	0.238	0.355	0.355	0.354	0.354
2	0.119	0.357	0.12	0.475	0.185	0.474
3	0.0943	0.451	0.0318	0.506	0.0591	0.505

Score plot PC1 vs. PC2
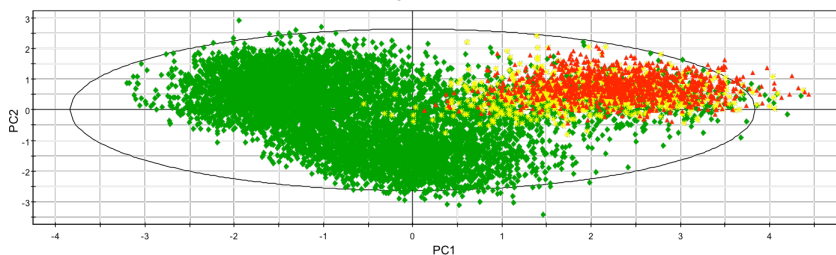
Loadings PC1
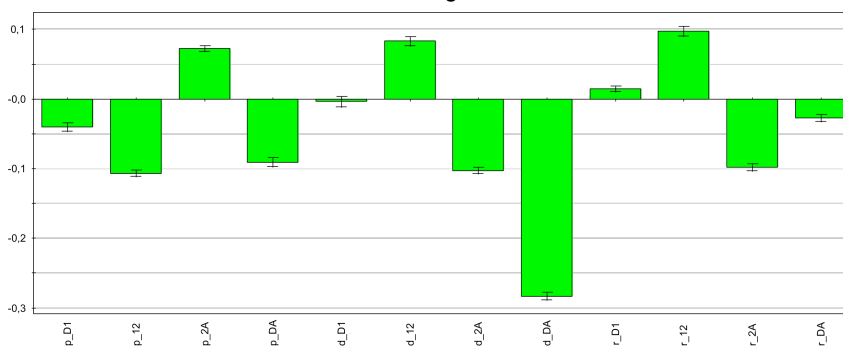
Loadings PC2
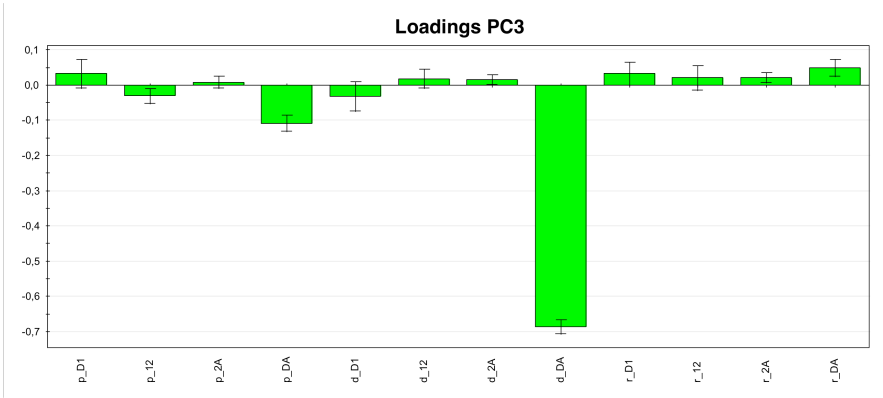
Loadings PC3

Trp-(Pro)2-Trp (gas phase)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.151	0.151	0.396	0.396	0.395	0.395
2	0.0728	0.224	0.135	0.531	0.223	0.53

Score plot PC1 vs. PC2
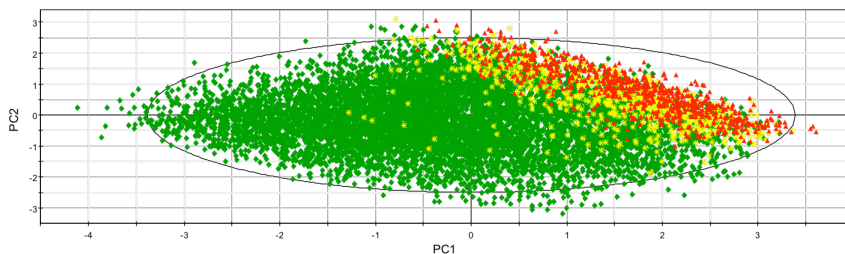
Loadings PC1
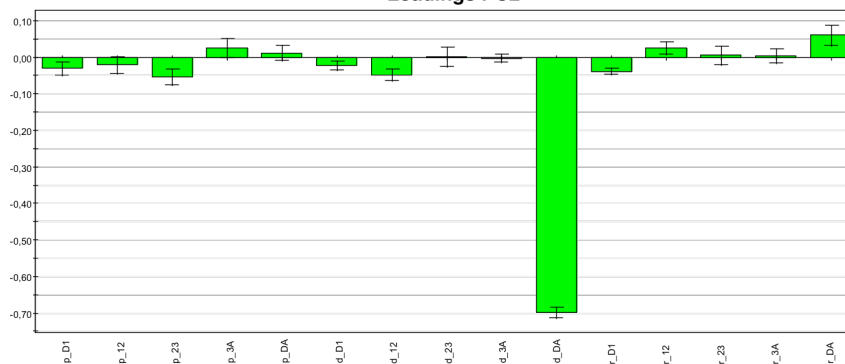
Loadings PC2

Trp-(Pro)2-Trp (solvated-conformation-only)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.153	0.153	0.186	0.186	0.185	0.185
2	0.0922	0.245	0.0143	0.2	0.0152	0.197

Score plot PC1 vs. PC2
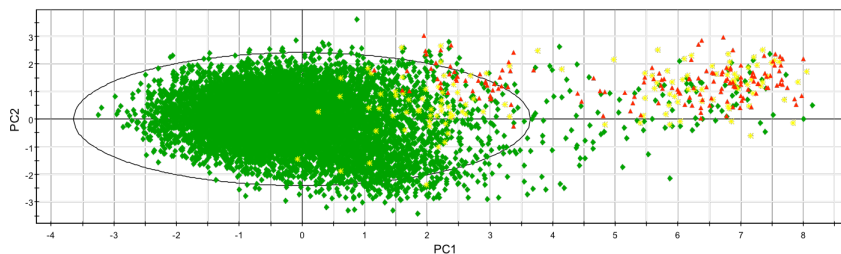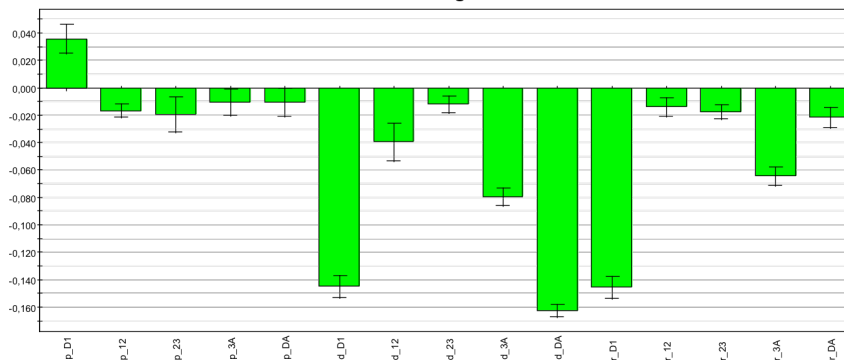
Loadings PC1
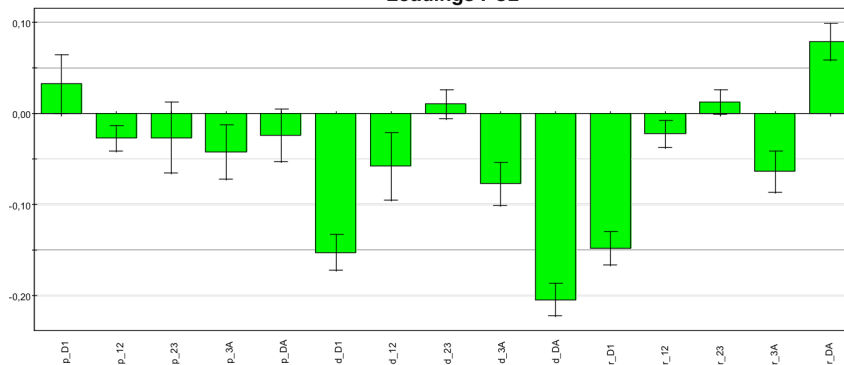
Loadings PC2

Trp-(Pro)3-Trp (gas phase)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.157	0.157	0.236	0.236	0.235	0.235
2	0.0478	0.204	0.117	0.352	0.149	0.349
3	0.091	0.295	0.0135	0.366	0.0189	0.361

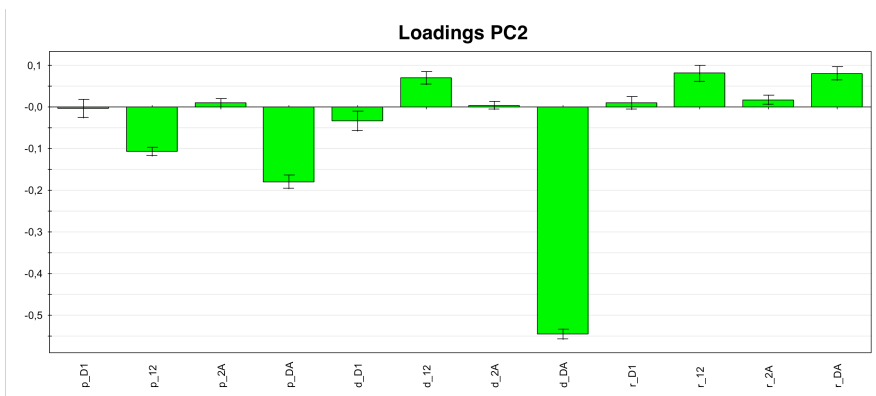Score plot PC1 vs. PC2
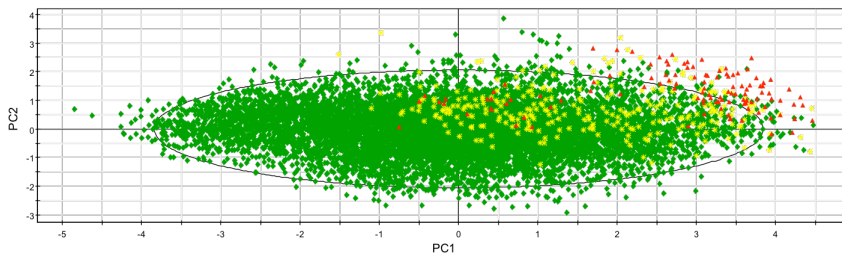
Loadings PC1
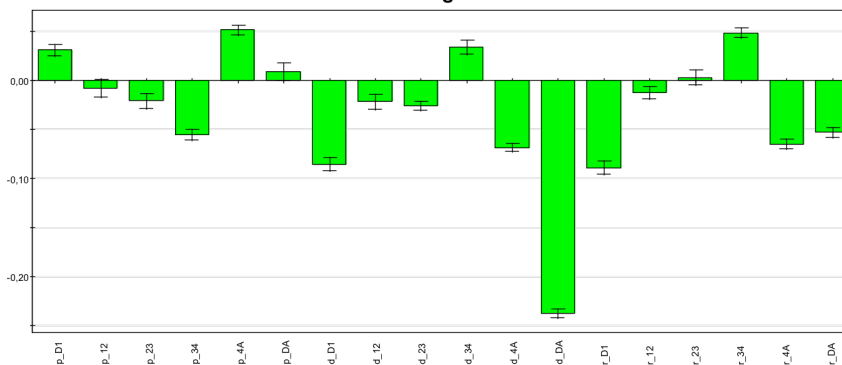
Loadings PC2

Loadings PC3
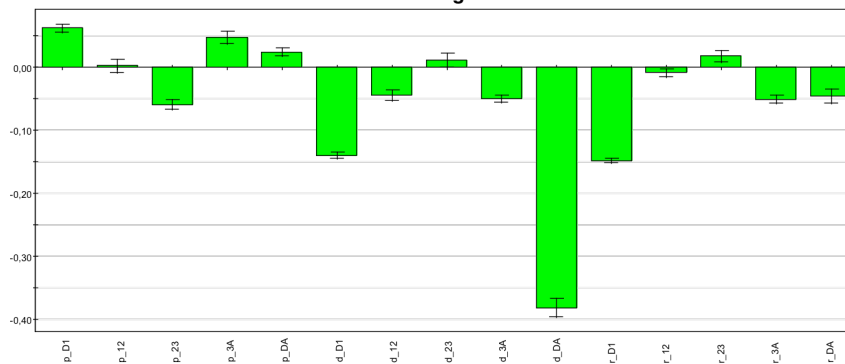
Trp-(Pro)3-Trp (solvated-conformation-only)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.177	0.177	0.184	0.184	0.183	0.183
2	0.0599	0.237	0.0305	0.214	0.0354	0.212

Score plot PC1 vs. PC2

Loadings PC1
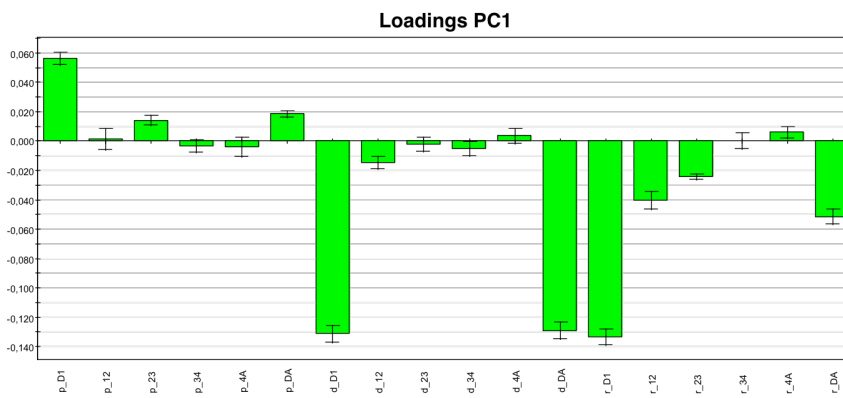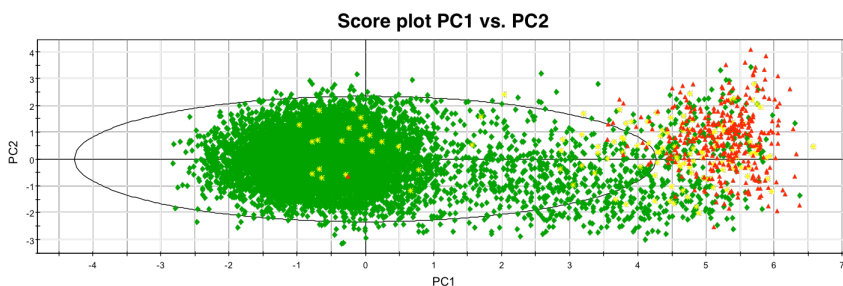
Loadings PC2

Trp-(Pro)4-Trp (gas phase)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.107	0.107	0.343	0.343	0.342	0.342
2	0.054	0.161	0.0967	0.44	0.144	0.436
3	0.061	0.222	0.0127	0.453	0.0193	0.447

Score plot PC1 vs. PC2
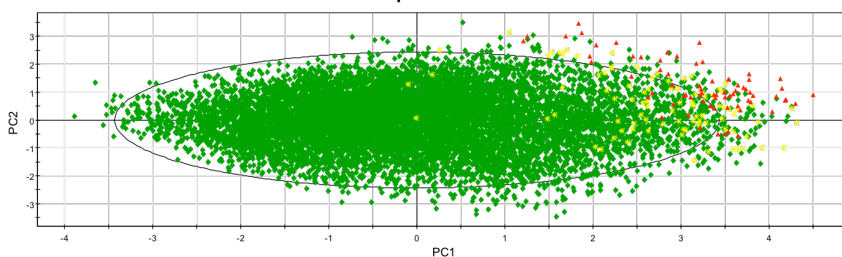
Loadings PC1
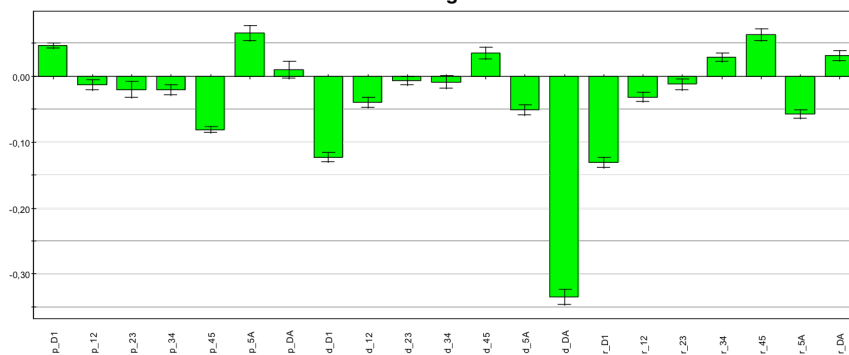
Loadings PC2
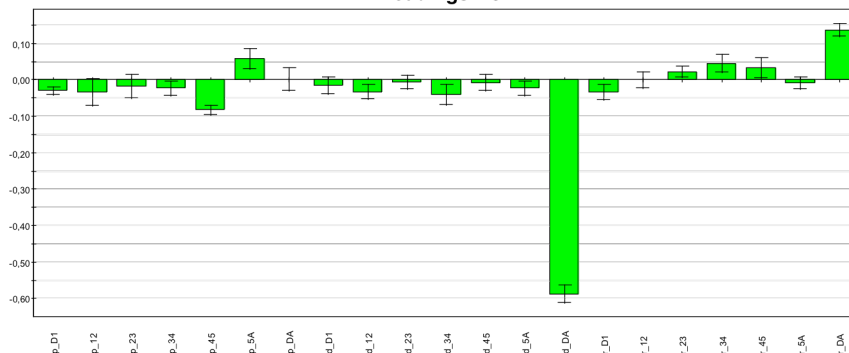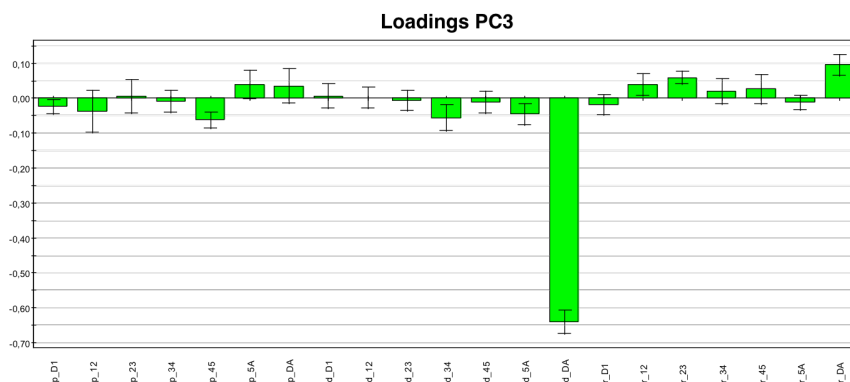
Loadings PC3

Trp-(Pro)4-Trp (solvated-conformation-only)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.0897	0.0897	0.323	0.323	0.322	0.322
2	0.0942	0.184	0.059	0.382	0.0854	0.38
3	0.0488	0.233	0.0113	0.393	0.0147	0.389

Score plot PC1 vs. PC2
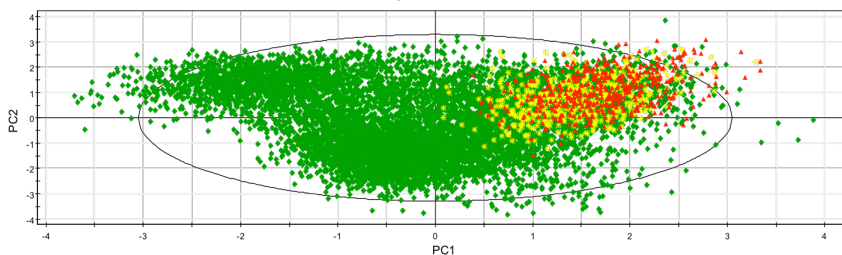
Loadings PC1
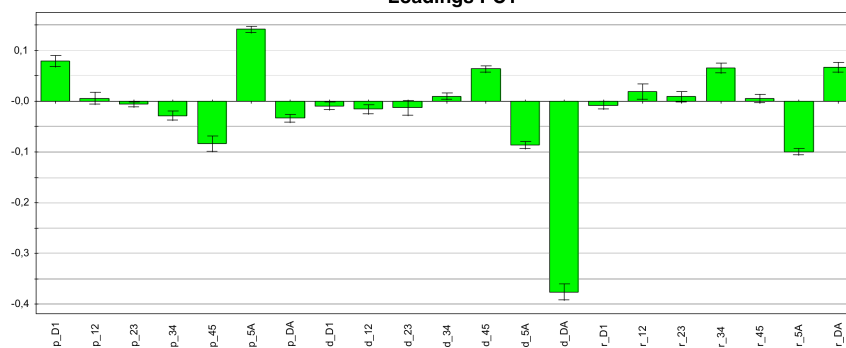
Loadings PC2
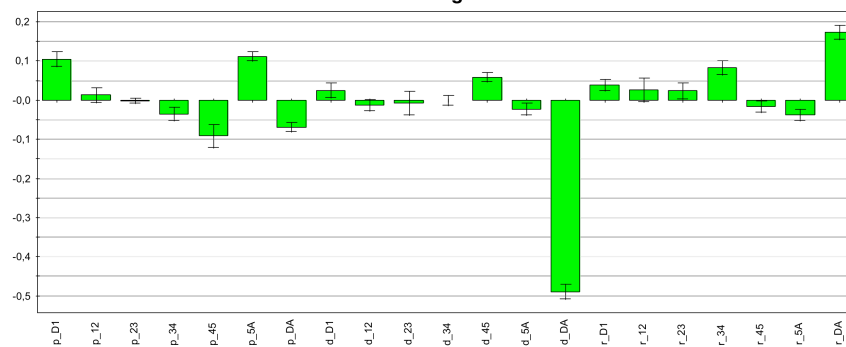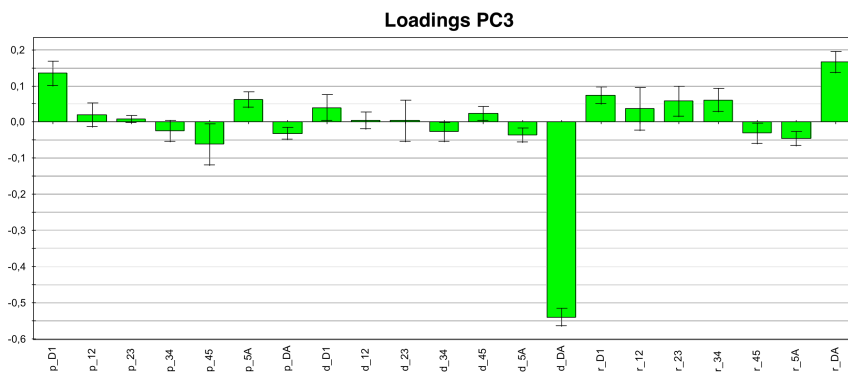
Loadings PC3

Trp-(Pro)5-Trp (gas phase)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.0971	0.0971	0.309	0.309	0.308	0.308
2	0.0499	0.147	0.0668	0.376	0.093	0.372
3	0.057	0.204	0.0109	0.387	0.0138	0.381

Score plot PC1 vs. PC2
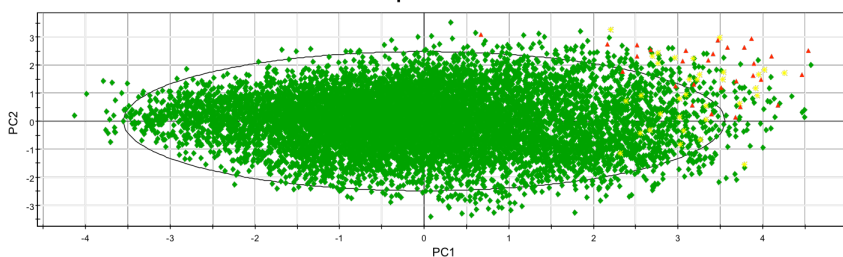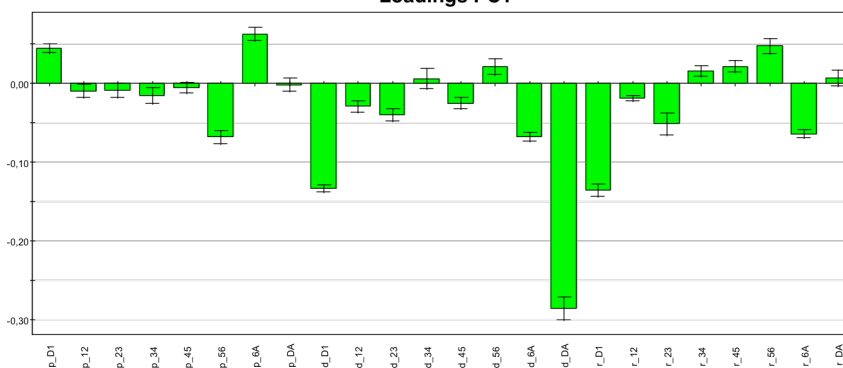
Loadings PC1
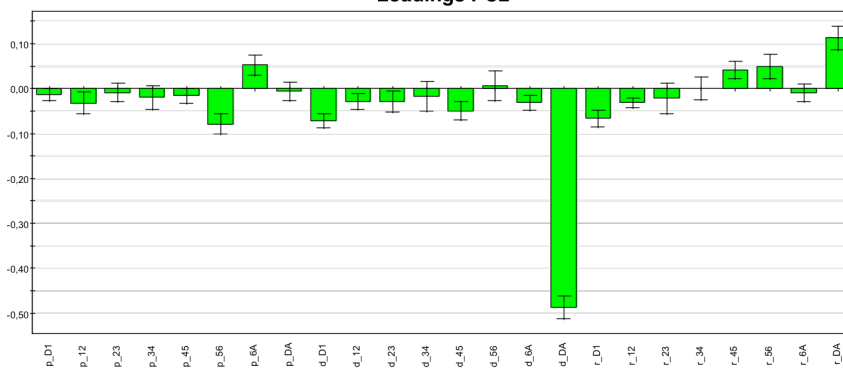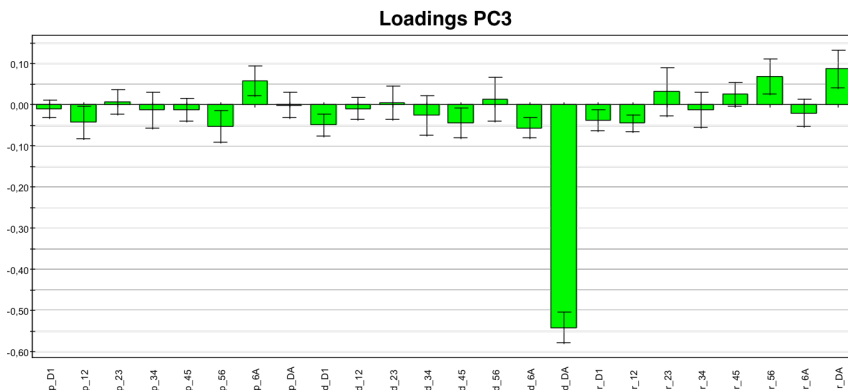
Loadings PC2

Trp-(Pro)5-Trp (solvated-conformation-only)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.0641	0.0641	0.0523	0.0523	0.0493	0.0493
2	0.0436	0.108	0.00597	0.0583	0.00148	0.0507

Score plot PC1 vs. PC2
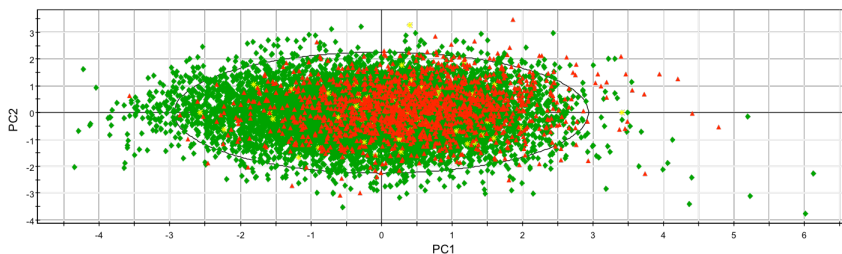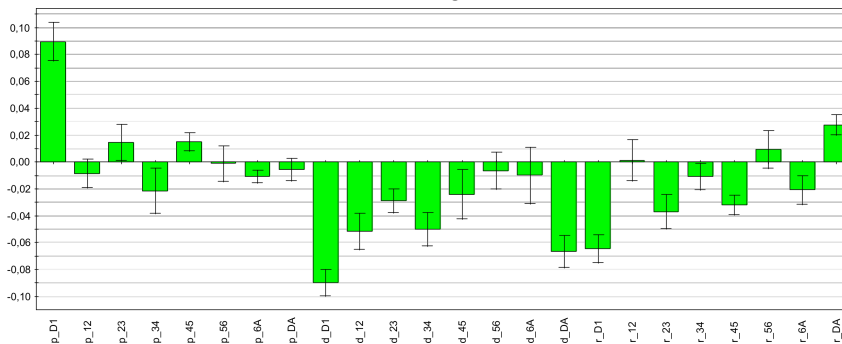
Loadings PC1
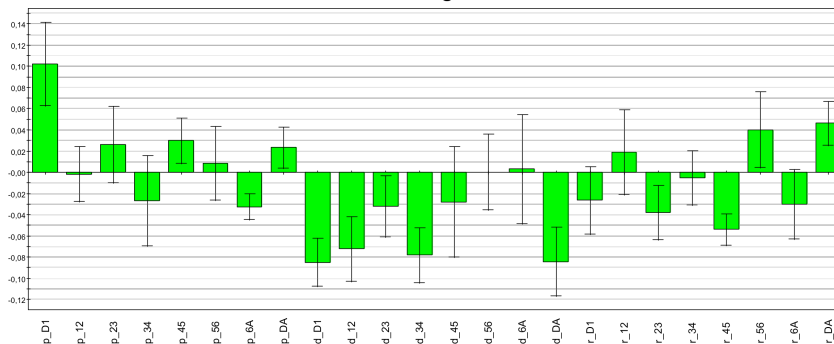
Loadings PC2

Trp-(Pro)6-Trp (gas phase)

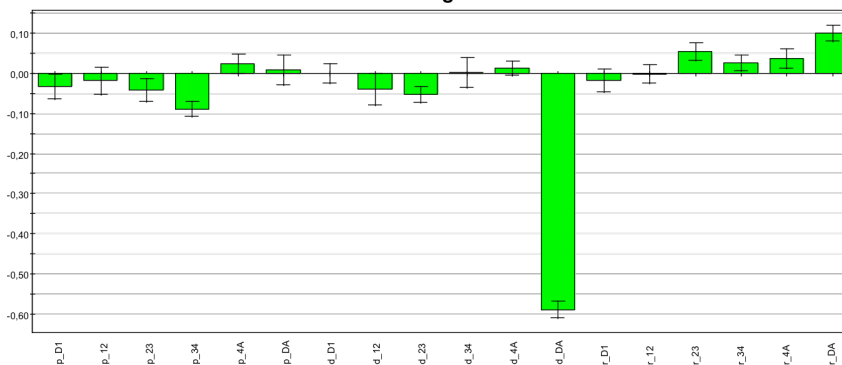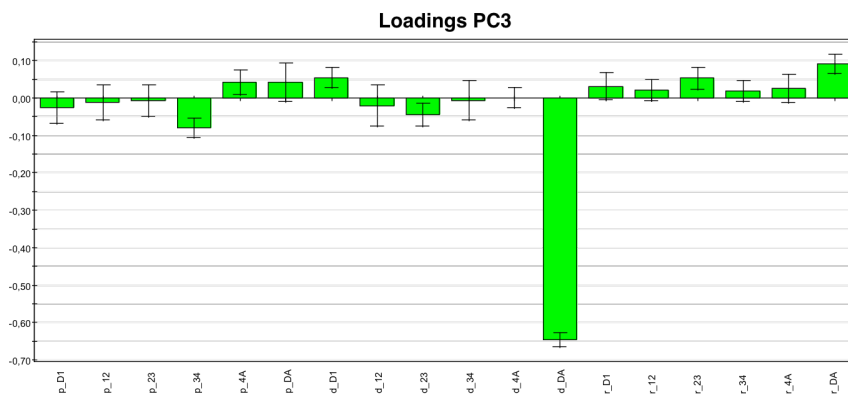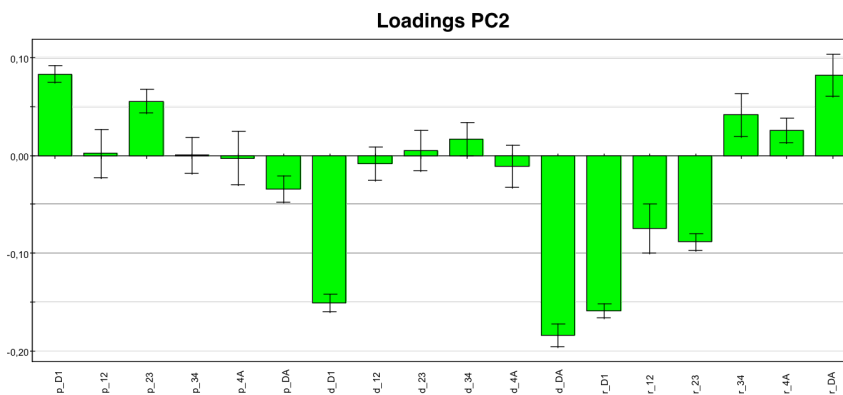| PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.0893	0.0893	0.267	0.267	0.264	0.264
2	0.0453	0.135	0.0775	0.344	0.101	0.339

Score plot PC1 vs. PC2
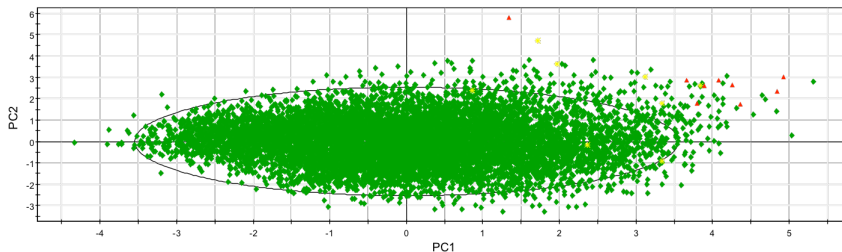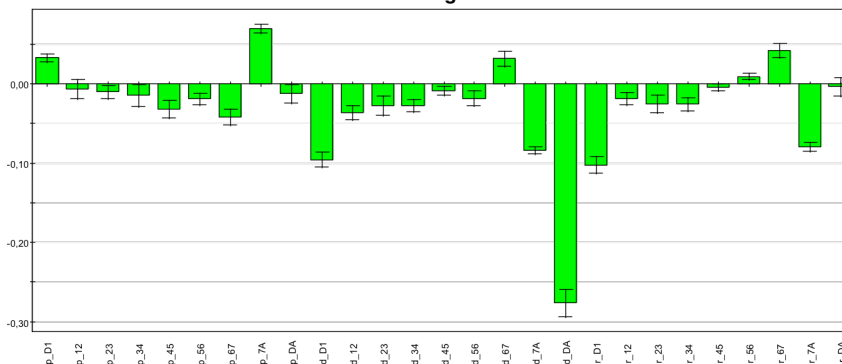
Loadings PC1
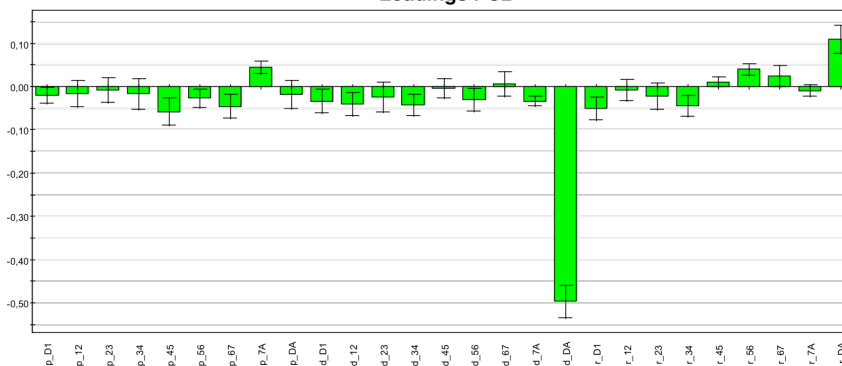
Loadings PC2

Trp-(Pro)6-Trp (solvated-conformation-only)

PC	R2X	R2X(cum)	R2Y	R2Y(cum)	Q2	Q2(cum)
1	0.0667	0.0667	0.216	0.216	0.213	0.213
2	0.0622	0.129	0.0429	0.259	0.0515	0.254
3	0.0484	0.177	0.0102	0.269	0.00971	0.261

Score plot PC1 vs. PC2
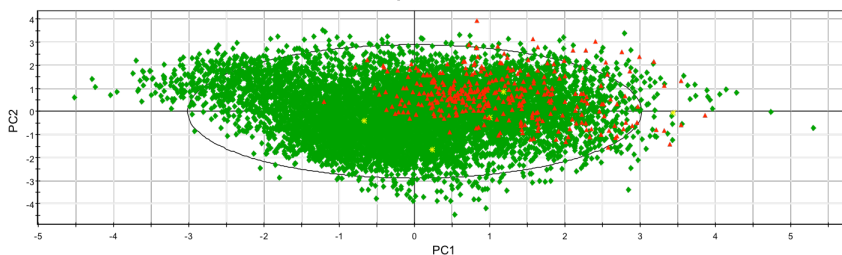
Loadings PC1
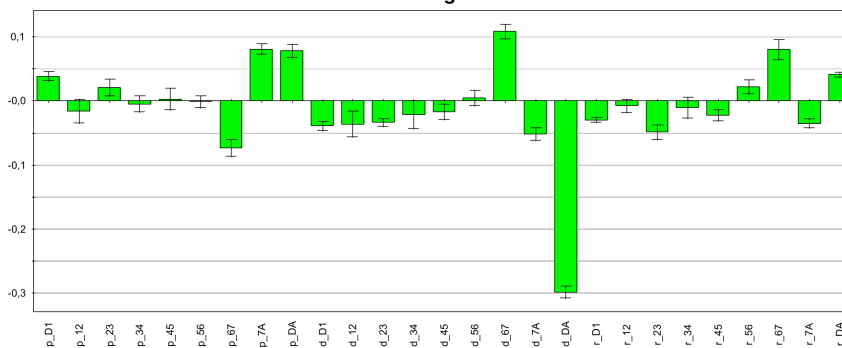
Loadings PC2

Loadings PC3
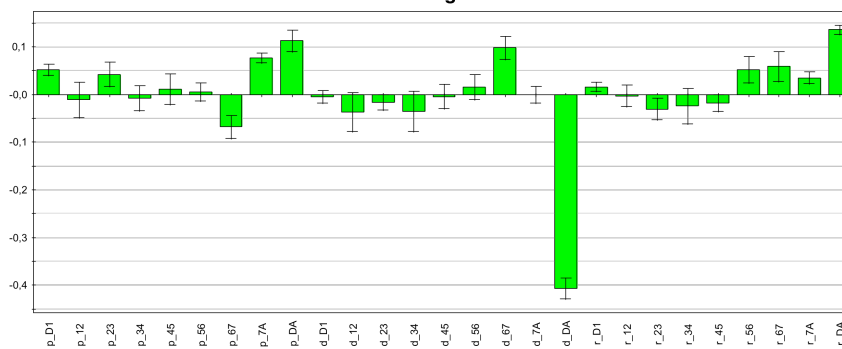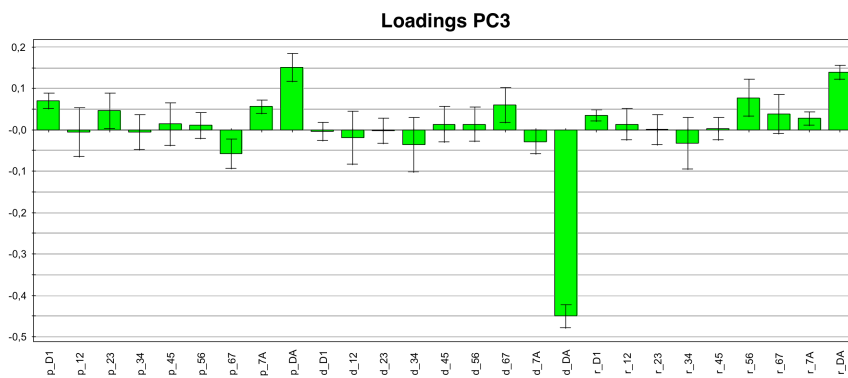
3.4.2. Temperature effects on donor-acceptor couplings in peptides. A combined QM and MD study.

Wallrapp, F. H., Voityuk. A. and Guallar, V.

Wallrapp FH, Voityuk A, Guallar V. Temperature effects on donor-acceptor couplings in peptides. A combined quantum mechanics and molecular dynamics study. J Chem Theory Comput. 2010; 6(10): 3241-3248.

Supporting Information

Temperature effects on donor-acceptor coupling in peptides. A combined QM and MD study

Frank Wallrapp, Alexander Voityuk, Victor Guallar

S1 REMD: We plotted the average potential energy against the temperature for each trajectory, shown in Figure S1. The plot clearly demonstrates a linear dependency between the potential energy and temperature indicating that our REMD sampled as expected by accepting conformations in the respective trajectories according to their energies differences.

    Furthermore, we applied REMD on Trp-(Pro)3-Trp within the temperature range of 25 to 400 K to demonstrate the ability to reach low d_{DA} at even very low temperature. The d_{DA} distribution plot of trajectories from 25 to 418 K is given in Figure S2.
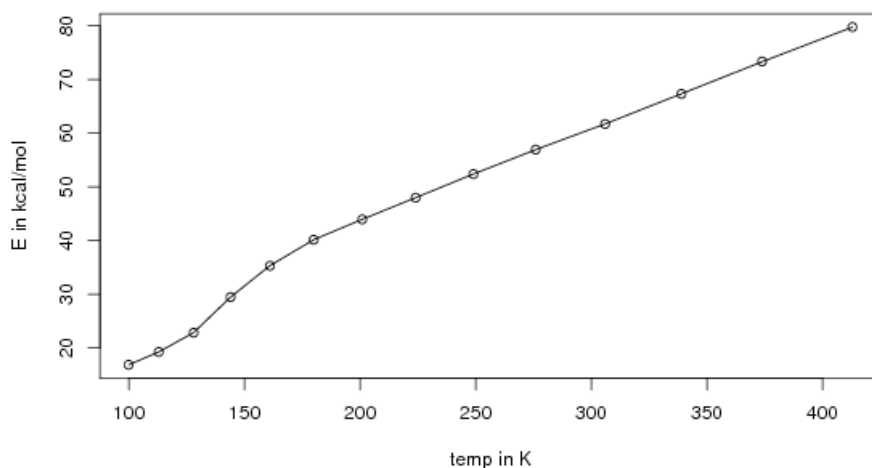
Figure S1: Average potential energy plotted against temperature of each replica exchange trajectory of Trp-(Pro)3-Trp.
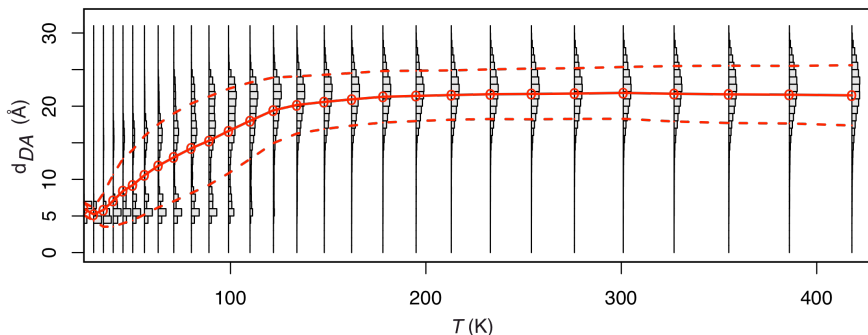
Figure S2: Distribution of the donor-acceptor distance d_{DA} of each replica exchange trajectory of Trp-(Pro)6-Trp from 25 to 122 K. Solid respectively dashed red lines show the mean values as well as the standard deviations.

Table 1: Rms V_{DA} values in eV, mean d_{DA} in Å and fluctuation in Å calculated for trajectories of Trp-(Pro)3-Trp with different temperatures T in K.

T	rms(V_{DA}) bridge mediated	rms(V_{DA}) direct	mean(d_{DA})	fluctuation
100	7.60×10^{-3}	6.78×10^{-3}	8.20	0.034
113	1.06×10^{-2}	8.68×10^{-3}	8.16	0.041
128	1.75×10^{-2}	1.72×10^{-2}	7.91	0.057
144	2.11×10^{-2}	2.08×10^{-2}	9.15	0.073
161	1.09×10^{-2}	1.07×10^{-2}	11.72	0.075
180	6.88×10^{-3}	6.88×10^{-3}	13.18	0.072
201	3.96×10^{-3}	3.99×10^{-3}	13.67	0.073
224	2.03×10^{-3}	2.06×10^{-3}	13.93	0.075
249	2.24×10^{-3}	2.24×10^{-3}	14.26	0.078
276	2.13×10^{-3}	2.15×10^{-3}	14.46	0.080
306	2.57×10^{-3}	2.53×10^{-3}	14.58	0.086
339	2.11×10^{-3}	2.10×10^{-3}	14.69	0.091
374	2.63×10^{-3}	2.56×10^{-3}	14.68	0.108
413	3.77×10^{-3}	3.57×10^{-3}	14.64	0.133

Table 2: Rms V_{DA} values in eV, mean d_{DA} in Å and fluctuation in Å calculated for trajectories of Trp-(Pro)6-Trp with different temperatures T in K.

T	rms(V_{DA}) bridge mediated	rms(V_{DA}) direct	mean(d_{DA})
100	1.40×10^{-2}	1.51×10^{-2}	10.91
113	1.36×10^{-2}	1.38×10^{-2}	11.02
128	1.46×10^{-2}	1.48×10^{-2}	12.29
144	1.06×10^{-2}	1.09×10^{-2}	14.98
161	7.75×10^{-3}	7.86×10^{-2}	17.16
180	6.43×10^{-3}	6.47×10^{-3}	18.88
201	4.16×10^{-3}	4.20×10^{-3}	20.10
224	4.06×10^{-3}	3.95×10^{-3}	20.44
249	3.36×10^{-3}	3.93×10^{-3}	20.67
276	2.30×10^{-3}	1.93×10^{-3}	20.87
306	1.63×10^{-3}	1.41×10^{-3}	20.87
339	2.69×10^{-3}	2.07×10^{-3}	20.80
374	2.69×10^{-3}	1.68×10^{-3}	20.71
413	4.12×10^{-3}	2.67×10^{-3}	20.54

THESIS PUBLICATIONS

3.5. Electron transfer in the Cytochrome c Peroxidase/ Cytochrome c complex.

Wallrapp, F. H., Voityuk, A. and Guallar, V.

THESIS PUBLICATIONS

Electron transfer in the Cytochrome c Peroxidase/Cytochrome c complex.

Frank H. Wallrapp[1], Alexander Voityuk[2,3], Victor Guallar[1,3]

[1] Barcelona Supercomputing Center, Nexus II Building, 08028, Barcelona, Spain
[2] Institute of Computational Chemistry, University of Girona, 17071 Girona, Spain
[3] Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

Abstract: Electron transfer processes are simple but crucial reactions in biochemistry, being one of the main step in almost all enzymatic cycles. Obtaining an atomic description of the transfer pathway is a difficult task, both at the experimental and theoretical level. Investigating the electron transfer pathway between Cytochrome c Peroxidase and its redox partner, Cytochrome c, we show that the combination of conformational sampling and clustering of the complex structure, followed by ET pathway mapping through our recently developed QM/MM e-Pathway approach and electronic coupling calculations on derived pathways, allows obtaining a substantial picture of the underlying ET mechanism, including the identification of the ET transfer pathway. Our results report the confirmation of residues Ala194, Ala193, Gly192 and Trp191 of CcP acting as the main ET pathway. Furthermore, low fluctuations between the electronic couplings calculated on individual complex conformations indicate a non-gated ET mechanism. We demonstrate that Trp191 appears as the most important residue in the ET region, receiving the electron hole at the first stage in almost all QM/MM e-Pathway calculations. Computed rate constant are in good agreement with experimental data, showing that the underlying ET mechanism is sequential hopping with Trp191 as localized bridge, being one order of magnitude faster than direct bridge-mediated ET between Cytochrome c and compound I of Cytochrome c Peroxidase.

Keywords: heme, CcP, Cytc, electron transfer pathway, QM/MM e-Pathway, FCD

I. Introduction

Electron transfer (ET) is one of the simplest but most crucial reactions in biochemistry.[1,2] Atomic and electronic detailed information is difficult to obtain, however, very valuable for better understanding of biological function or for the design of synthetic energy transduction systems.[3-7]

In general, ET can be considered as a transition between electronic states. Its rate is determined by the coupling between these electronic states, the reaction free energy and the reorganization energy needed by the system to adapt to its new state.

Despite bridge-mediated superexchange between donor and acceptor, electron transfer can also occur through a process of incoherent hopping between localized electronic states on the bridge.[8] The respective contributions of the two mechanisms are dependent on the energy difference between the donor- and bridge-localized electronic states. The total ET rate of the system is then the composite of both, bridge-mediated superexchange and sequential hopping mechanism. Of specific interest is not only the overall electronic coupling of donor and acceptor but also the particular pathway the electron is assumed to take. Established methods to determine the ET pathway and calculate the electronic coupling are the PATHWAYS method[9] being incorporated into the HARLEM molecular simulation package,[10] as well as an approach based on an artificial intelligence search for key residues in ET systems.[11] Very recently, Guallar et al. published another approach to map ET pathways in proteins, called QM/MM e-Pathway, which identifies residues in the order of their importance in the ET region.[12,13] There have been many investigations on ET in enzymes, [14-18] with several of them on ruthenium modified proteins.[19-23] A further subject of particular interest is ET in DNA systems.[24-33] Within this context, the widely used two-state Generalized Mulliken-Hush (GMH) method[34,35] as well as its extension to the multi-state model[36] have found much application.[30,37,38] Based upon GMH, Voityuk and Rösch developed the Fragment Charge Difference (FCD) method being able to calculate the electronic coupling through donor-acceptor charge differences in the corresponding adiabatic states of the system.[39] Recently, the FCD method was successfully applied to protein in terms of systematic investigations on solvent and temperature dependence of electronic coupling in oligopeptides.[40,41] Moreover, it found application in calculation of electronic coupling between the Fe4-S4 clusters N5 and N6a of the bacterial respiratory complex I, searching for possible hopping residues in the ET region.[42]
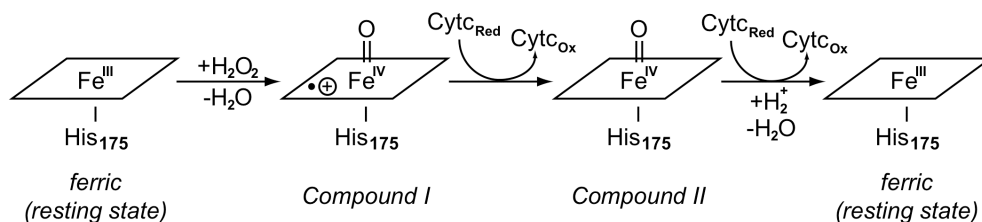
Figure 1: Catalytic cycle of CcP. Scheme adapted from Ref. [43]

Of particular interest is also the ET in the cytochrome c peroxidase/cytochrome c (CcP/Cytc) complex.[14,44-48] In its catalytic cycle, CcP undergoes a two electron oxidation by peroxide to form a common oxo-ferryl intermediate, the so-called compound I (CpdI). CpdI is then reduced by two distinct Cytc to form CpdII and finally its ferric resting state again (see Figure 1). Based on the crystal structure of the complex, Pelletier and Kraut proposed a σ-bond tunneling pathway from Cytc to Trp191 following the residues Ala194, Ala193 and Gly192 of CcP (shown in red in

Figure 2).[44] More recent studies support this thesis.[46,49,50] Furthermore, it is widely accepted that CpdI extracts one electron from Trp191 to create the radical intermediated Trp191+, which finally gets its missing electron from Cytc.[51-56] Another possible pathway, including residues Leu177, Gly178, Asn195, Asn196 and Val197 of CcP, and Met80, Ala81 and Phe82 of Cytc (shown in blue in Figure 2), was proposed in an theoretical study by Siddarth based on an artificial intelligence search to for key residues in protein ET systems.[11]
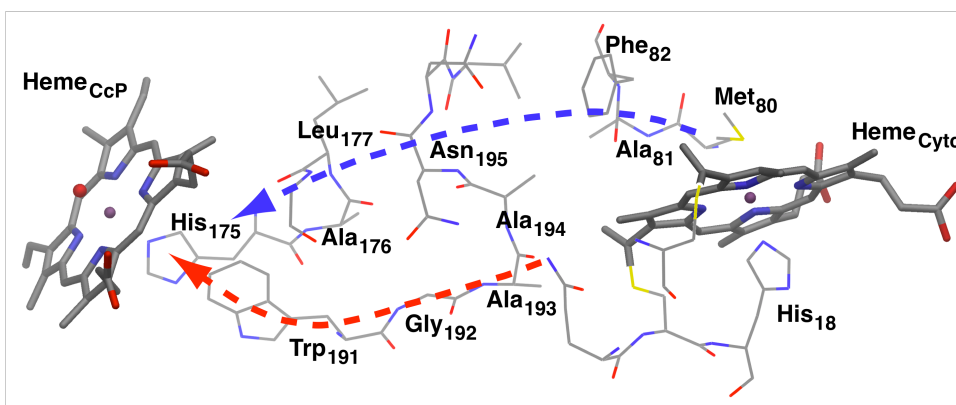
Figure 2: Electron transfer region of the CcP/Cytc complex. ET pathway proposed by Pelletier and Kraut is shown in red, ET pathway proposed by Siddarth is shown in blue.

Much experimental work was also done on detecting the rate of the CcP catalyzed reaction. The measured rates strongly depend on the ionic strength used in the experiments, resulting in rate constants ranging from about 250 s^{-1} (low ionic strength) to 2120 s^{-1} (high ionic strength) for yeast Cytc to CpdI using both stopped-flow[45,48,57-61] and laser flash techniques[62-65] for both the covalent[45,66] and noncovalent complexes[48,61,62]. The measurement of the actual rate of the pure ET process of this reaction still bears some difficulties as it is too fast for stopped-flow spectroscopy, thus being estimated to be greater than 5.0×10^4 s^{-1}.[48,65]

In this article we present a deep analysis of conformational sampling, ET pathway and electronic coupling calculations. Moreover, we introduce for the first time the approach of electronic coupling calculations applied on ET pathways derived through our recently developed QM/MM e-Pathway approach. We investigated the ET mechanism as well as pathway occurring in the reduction of CpdI by Cytc. We not only demonstrate that the most presumable ET pathway between Cytc and CcP follows residues Ala194, Ala193, Gly192 and Trp191 of CcP, but also show that our methodology is able to detect Trp191 as localized bridge of sequential hopping hole transfer from Cytc to CpdI as proposed in the literature.

2. Theoretical Methods

All QM/MM calculations were performed with the QSite program,[67] applying unrestricted M06[68] level of theory for the precursory QM/MM minimizations and electronic coupling calculations, and unrestricted Hartree-Fock for the QM/MM e-Pathway calculations. All setups included the 6-31G* basis set and a lacvp pseudo-potential for the iron center.

We chose the structure of the native CcP/Cytc complex being crystallized by Pelletier and Kraut for our calculations.[44] We applied protein preparation as well as QM/MM minimization on both hemes in their respective reduced states in order to apply the FCD method together with the Koopmans' theorem (see below). We refer to the resulting conformation as the crystal complex of CpdII/Cytc$_{RED}$ since the proteins did not undergo conformational changes throughout the equilibration process. We performed 2 ns of molecular dynamics (MD) on the complex applying GROMACS[69] together with the OPLS force field, from which we extracted snapshots every single ps. Based on the RMSD of only the ET region between the two active sites we clustered all 2000 snapshots, applying the k-medoids algorithm[70] and chose the resulting 10 clustering modes as representatives for our ET pathway mapping and electronic coupling calculations. In particular, the chosen conformations were taken at time steps 0, 162, 432, 990, 1083, 1356, 1527, 1764, 1814 and 1884. ET pathways were derived from each conformation applying our QM/MM e-Pathway approach simulating hole transfer.[12,13] Protocols of the protein preparation, MD simulation and QM/MM minimization as well as on our QM/MM e-Pathway approach are given in the supporting information and in the original publications, respectively. Furthermore, careful inspection of our snapshots discovered a twofold negative charge in about 13 Å distance of CcP due to unscreened residues Asp146 and Asp148 on the surface of CcP. As electronic coupling calculated by the FCD approach is not affected by an external electric field,[39] we neutralized this charge through careful placements of counter ions only for the Gibbs free energy and reorganization energy estimations of the ET process being described below.

Electronic coupling values were calculated between donor and acceptor applying the FCD method, where the electronic coupling is denoted as:[29,39]

$$V_{DA} = \frac{(E_2 - E_1)|\Delta q_{12}|}{\sqrt{(\Delta q_1 - \Delta q_2)^2 + 4\Delta q_{12}^2}} \, .$$

Here, Δq_1 and Δq_2 are the donor-acceptor charges difference in the two adiabatic states with their respective energies E_1 and E_2, and Δq_{12} is the corresponding off-diagonal term. Using Koopmans' theorem[71] we estimate the adiabatic splitting $E_2 - E_1$ through the one-electron energies of the highest occupied molecular orbitals being localized on the donor and acceptor sites, respectively, in the CpdII/Cytc$_{RED}$ complex, acting as the neutral system in case of hole transfer. We furthermore apply the recently defined root mean square coupling $rmsV_{DA}$ and the effective

coupling squared V^2 by averaging over several states in order to incorporate (nearly) degenerate states of the hemes:[42]

$$rms\,V_{DA} = \sqrt{\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} V_{ij}^2} \quad \text{and}$$

$$V_{DA}^2 = N_A\,rms\,V_{DA}^2,$$

where N_D and N_A denote the number of (nearly) degenerate diabatic states localized on donor and acceptor sites, respectively. The use of $rms\,V_{DA}$ rather than $<V_{DA}>$ is in line with the expression for the nonadiabatic charge transfer.[71,72] In the following we refer to direct coupling when computing V_{DA} for QM regions consisting only of the donor and acceptor; and bridge mediated coupling those including the donor, acceptor and specified bridging residues. The ET rate was estimated using Marcus expression:[72]

$$k_{ET} = \frac{2\pi}{\hbar} V_{DA}^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(-\frac{(\lambda + \Delta G^\circ)^2}{4\lambda k_B T} \right),$$

where $\hbar$ is Planck´s constant, k_B is Boltzmann´s constant, λ is the reorganization energy, T is the temperature, V_{DA} is electronic coupling, and ΔG^o is the Gibbs free energy change of the electron transfer reaction. We estimate ΔG^o through the enthalpy $\Delta E = E_{products} - E_{reactants}$, neglecting any entropic effects, which we assume to be very little in this ET reaction. In detail, we perform QM minimizations of CpdI, CpdII, Cytc_{Ox} and Cytc_{Red} separately in gas phase as well as in implicit solvent with a dielectric constant of 4.0 simulating protein environment. In addition, we perform QM/MM minimizations of the complete protein complex in the CpdI - Cytc_{Red} state (reactants) as well as the CpdII - Cytc_{Ox} state (products). In order to give a rough estimate of λ we took the wave function of the products and did a single point calculation on the minimized structure of the reactants preserving the product's oxidation state.

3. Results and Discussion

3.1 Structures. The RMSD characteristics of snapshots from our MD simulation against the QM/MM templates of heme_{CpdII} and heme_{Cytc} show low deviation and thus indicate native conformations throughout the simulation (see supporting information). The superposition of all 11 conformation applied within this study is given in Figure 3.
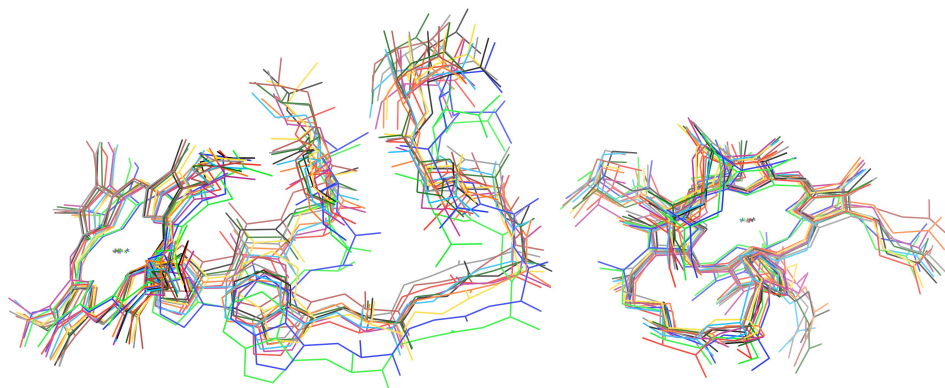
Figure 3: Superposition of all 11 conformations.

3.2 QM/MM e-Pathway. We have calculated the electron transfer pathways for the crystal as well as the 10 CcP/Cytc conformations selected by clustering of the 2 ns MD trajectory. Following the QM/MM e-Pathway procedure we determined important residues in the ET region in consecutive order. All identified residues are Ala176, Leu177, Trp191, Gly192, Ala193, Ala194, Asn195 and Asn196 of CcP as well as Ala81 and Phe82 of Cytc. We calculated a logo plot giving the order of specified residues as selected by the QM/MM e-Pathway algorithm on all conformations (see Figure 4). Here, the size of a digit specifies the number of occurrences a certain residue was selected at the corresponding position (value of the digit) within the ET pathways calculations. Occurrences less than 100% denote corresponding residues were not identified by the methodology after 7 steps. Furthermore, total occurrences of specific positions do not sum up to 1, due to the rare case of multiple residues identified by a single QM/MM e-Pathway step. The results clearly indicate the importance of Trp191 being identified as first hole acceptor in 7 of 11 conformations, followed by second and third hole acceptor in 2 conformations, respectively. Next high importance shows residue Asn195 being selected by the QM/MM e-Pathway algorithm at a very first stage in 9 conformations. Ala194 accepts the hole at a very first stage, too, if identified at all. This means if the   On the contrary low importance does show residue Asn196 being selected last and next to last in all its 6 conformations.

Figure 4: Logo plot of the order of specified residues within ET transfer region being selected by the QM/MM e-Pathway algorithm on all conformations. Size of digit specifies the number of occurrences residue was selected at corresponding position within all ET pathways calculations.

For each snapshot, we connect the identified residues to pathways in the consecutive order of the appearance. Once a connecting chain of residues between the donor and acceptor is found, we denote these residues as the main ET pathway for this snapshot. A possible alternative ET pathway is assigned if the subsequently identified residues allow a different connected pathway from the donor to the acceptor.

Our calculations come up with 2 distinct ET pathways. The first pathway consists of residues Trp191, Gly192, Ala193 and Ala194 of CcP (in the following called *path1*). The second pathway consists of residues Ala176, Leu177, Ala194 and Asn195 of CcP, and Phe82 and Ala81 of Cytc (in the following called *path2*). Path1 is found to be the only pathway in 5 out of the 11 conformations (snapshots 432, 1083, 1527, 1764 and 1814). *Path1* is found to be the main ET pathway, with *path2* as alternative, in 3 conformations (snapshots 990, 1356 and 1884). *Path2* is found to be the main ET pathway, with *path1* as alternative in the remaining 3 conformations (crystal and snapshots 0 and 162). *Path2* is never identified as the only ET pathway connecting donor and acceptor. In the very general case the first identified residues are part of the finally assigned main pathway. However, for 2 conformations (crystal and snapshot 1814), the assigned main pathway does not contain their respective firstly identified residues. Comparing our results against the literature, it can be seen that *path1* is exactly the ET pathway proposed by Pelletier and Kraut and being confirmed by several studies.[1,44,50] There is much less evidence for *path2* to be the ET pathway between Cytc and CcP, yet is was proposed by Siddarth through a different theoretical approach on the crystal structure of the CcP/Cytc complex.[11]

The QM/MM e-Pathway approach only identifies important residues in the ET region. Hence it does not quantify the probability for an electron to travel along these residues. Consequently, in order to evaluate an assigned ET pathway, it is necessary to compute the electronic coupling between the donor and acceptor incorporating the residues on the path. Besides alternative pathways, we further investigated the true nature of Trp191 as localized bridge in the ET process. Thus the following section is split into two parts: Firstly, bridge-mediated ET between Cytc and CcP and secondly, sequential hopping ET with Trp191 as localized bridge.

3.3 Single step ET

Electronic coupling. We calculated the electronic coupling on all 11 conformations based on 4 different QM setups. The first one (*direct*) consists of only $heme_{CcP}$ and $heme_{Cytc}$ in order to compute the direct electronic coupling between the donor and acceptor. The second QM setup (*all*) includes the complete ET region in addition to $heme_{CcP}$ and $heme_{Cytc}$ that to compute the full bridge-mediated electronic coupling. The third and forth setup consists of $heme_{CcP}$ and $heme_{Cytc}$ and the residues denoted by path1 and path2, respectively. We calculated the electronic coupling between donor and acceptor by applying the FCD scheme together with equation (2) (applying $N_D = N_A = 2$) in order to properly incorporate the nearly degenerate electronic states of the hemes. The *rms* V_{DA} values for all snapshot with QM setups *direct*, *path1*, *path2* and *all*, respectively, are given in Table 1. Fluctuations of the electronic coupling are depicted through the coherence factor $R_{coh} = \langle V \rangle^2 / \langle V^2 \rangle$ and given for each system.

Clearly, the electron transfer between $heme_{CcP}$ and $heme_{Cytc}$ is of bridge-mediating nature. As expected from the mere distance of donor and acceptor being about 27 Å, the direct electronic coupling lies in the very low range of 10^{-9} eV, whereas the bridge-mediated electronic coupling lies within the range of 10^{-7} to 10^{-5} eV. Here, bridge-mediating electronic coupling on pathway *path1* is as high as bridge-mediated electronic coupling incorporating all ET residues. In contrast, ET pathway *path2* mediates ET at much lower rate than *path1* or *all*. Hence, the results indicate that *path1* is the main ET pathway between Cytc and CcP, with *rms* $V_{DA} = 4.13 \times 10^{-6}$ eV. Furthermore, the results show that the electronic coupling values do not fluctuate much between the different conformations with measured coherence values R_{coh} ranging from 0.28 to 0.83. Consequently, electron transfer in the CcP/Cytc complex is not conformationally gated.

Table 1: Electronic coupling values $rms\ V_{DA}$ in eV calculated between donor and acceptor (DA) with QM setups *direct*, *path1*, *path2* and *all*, respectively. Donor-acceptor distance d_{DA} in Å, measured between both irons.

snapshot	d_{DA} (Å)	DA, *direct*	DA, *path1*	DA, *path2*	DA, *all*
crystal	26.4	5.48×10^{-9}	1.35×10^{-6}	1.17×10^{-7}	2.41×10^{-6}
0	26.7	9.19×10^{-9}	1.30×10^{-5}	1.76×10^{-7}	3.03×10^{-6}
162	27.3	4.42×10^{-9}	8.94×10^{-7}	9.99×10^{-8}	3.94×10^{-6}
432	27.3	1.26×10^{-8}	6.28×10^{-7}	2.85×10^{-7}	3.72×10^{-6}
990	26.9	4.28×10^{-9}	2.36×10^{-7}	3.65×10^{-7}	3.40×10^{-6}
1083	27.6	1.10×10^{-9}	3.33×10^{-6}	1.02×10^{-7}	4.93×10^{-6}
1356	27.2	3.10×10^{-9}	5.17×10^{-7}	3.65×10^{-7}	1.72×10^{-6}
1527	27.1	2.96×10^{-9}	1.50×10^{-6}	1.26×10^{-6}	1.32×10^{-6}
1764	27.7	5.74×10^{-9}	8.89×10^{-7}	4.90×10^{-8}	1.06×10^{-6}
1814	27.4	2.92×10^{-9}	4.89×10^{-7}	7.02×10^{-8}	9.85×10^{-7}
1884	27.3	4.05×10^{-9}	1.03×10^{-6}	3.56×10^{-7}	3.35×10^{-6}
rms	27.2	5.69×10^{-9}	4.13×10^{-6}	4.40×10^{-7}	2.99×10^{-6}
R_{coh}		0.66	0.28	0.45	0.83

Rate constant. In order to obtain the rate constant of the ET from Cytc to CcP we also performed ΔG^o as well as λ calculations. We estimated ΔG^o through the enthalpy $\Delta E = E_{products} - E_{reactants}$, neglecting any entropic effects. In detail, we performed QM minimizations of CpdI, CpdII, $Cytc_{Ox}$ and $Cytc_{Red}$ separately in gas phase as well as in implicit solvent with a dielectric constant of 4.0 simulating protein environment. From both calculations we obtained very similar $\Delta E = (E(CpdII)+E(Cytc_{Ox}))-(E(CpdI)+E(Cytc_{Red}))$ of -0.93 eV and -0.95 eV, respectively. In addition, we performed QM/MM minimizations of the complete protein complex in the CpdI - $Cytc_{Red}$ state (reactants) as well as CpdII - $Cytc_{Ox}$ state (products). In correspondence to the literature, our calculations revealed a stabilization of $heme_{CcP}$ through polarization of hydrogen atoms of Asp48 and Trp51 forming hydrogen bonds to the Oxo atom of $heme_{CcP}$.[73-75] Incorporating the counter ions for residues Asp146 and Asp148 on the surface of CcP and extending the QM region of CcP by residues Asp48 and Trp51, the QM/MM minimizations for reactants and products resulted into ΔE of -0.92 eV. This energy is in excellent agreement with energies derived from our QM model calculations as well as with ΔG^o of -0.9 ± 0.15 eV given by an experimental study on the reaction of CcP_{horse} with Cytc.[76] It further confirms that entropic effects might not play an important role in this process. It is important to say that the electronic coupling is affected neither by the presence of Asp48 and Trp51 in the QM region of $heme_{CcP}$ nor by the counter ions of Asp146

and Asp148 but mainly depends on the character of the corresponding overlapping orbitals, as mentioned in the literature before.[18,39]

In order to estimate λ we took the wave function of the products and did a single point QM/MM calculation on the minimized structure of the reactants. In doing so, we calculated λ as 0.85 eV. This value is lower than the one suggested by Conklin *et al.* but lies within the range of λ (0.8 eV to 2.0 eV) known from similar electron transfer systems.[76]

Applying equation (3) and (4) with bridge-mediated electronic coupling *rms* $V_{DA}(path1)$ = 4.13 × 10^{-6} eV between heme$_{Cytc}$ and heme$_{CcP}$ (each having 2 degenerate electronic states), reaction energy change ΔE = -0.92 eV and reorganization energy λ = 0.85 eV, the rate constant of the single step ET process was calculated as 9.35 × 10^4 s^{-1}.

From the literature we know that the ET process in the CcP/Cytc couple is expected to occur through the physical occupation of Trp191 forming a positive radical. Consequently, we split the ET process into two separate ones, ET from Cytc to Trp191 and ET from Trp191 to CcP, and calculate the electronic coupling as well as rate constants for both separately.

3.4 Two-step ET

Electronic coupling. In order to investigate the nature of Trp191 as localized bridge in the electron transfer from heme$_{CcP}$ to heme$_{Cytc}$ as proposed in the literature, we furthermore calculated the electronic coupling between heme$_{Cytc}$ and Trp191, as well as between Trp191 and heme$_{CcP}$ applying QM setups *all* and *path1*. The electronic coupling values for all snapshots are calculated in terms of the FCD scheme together with equation (2) (applying $N_A = N_D = 2$ and $N_B = 1$). The results are given in Table 2 together with their coherence factors. Again, residues on *path1* mediate the ET as good as all residues in transfer region, indicating that *path1* is the exclusive ET pathway between Cytc and Trp191 (*rms* $V_{BA}(path1)$ = 1.77 × 10^{-5} eV). Deviations between the different snapshots are small with coherence factors larger than 0.69. Comparing the values against those of the previous section we can see that electronic coupling between heme$_{Cytc}$ and Trp191 is one order of magnitude higher than those between heme$_{Cytc}$ and heme$_{CcP}$. Hence it is more likely that the hole goes into Trp191 instead of heme$_{CcP}$ once the activation energy of the reaction is established. The electronic coupling between Trp191 and heme$_{CcP}$ is very high (*rms* $V_{BA}(path)$ = 1.32 × 10^{-2} eV) as both sites are in close vicinity.

Table 2: Electronic coupling values *rms* V_{DA} in eV for bridge localized electron transfer are given for donor and bridge (DB), and bridge and acceptor (BA) applying ET pathway *all* and *path1* with Trp191 as bridge.

snapshot	d_{DB} (Å)	DB, *path1*	DB, *all*	BA, *path1*	BA, *all*
crystal	21.4	2.98×10^{-5}	1.16×10^{-5}	2.06×10^{-2}	2.46×10^{-2}
0	21.2	9.86×10^{-6}	1.01×10^{-5}	1.30×10^{-2}	1.34×10^{-2}
162	21.3	5.11×10^{-6}	2.03×10^{-5}	1.04×10^{-2}	1.01×10^{-2}
432	21.2	9.77×10^{-6}	8.34×10^{-6}	1.20×10^{-2}	1.19×10^{-2}
990	20.6	1.52×10^{-5}	1.93×10^{-5}	1.00×10^{-2}	9.99×10^{-3}
1083	21.1	3.67×10^{-5}	2.53×10^{-5}	5.83×10^{-3}	5.63×10^{-3}
1356	21.2	3.62×10^{-6}	9.35×10^{-6}	7.55×10^{-3}	7.83×10^{-3}
1527	21.4	9.78×10^{-6}	6.77×10^{-6}	2.22×10^{-2}	2.25×10^{-2}
1764	21.6	1.02×10^{-5}	1.04×10^{-5}	1.44×10^{-2}	1.45×10^{-2}
1814	21.1	2.04×10^{-5}	1.87×10^{-5}	6.07×10^{-3}	6.15×10^{-3}
1884	20.9	1.20×10^{-5}	1.12×10^{-5}	1.14×10^{-2}	1.13×10^{-2}
rms	21.2	1.77×10^{-5}	1.49×10^{-5}	1.32×10^{-2}	1.38×10^{-2}
R_{coh}		0.69	0.85	0.85	0.82

Rate constant. For the case of the two step ET process from heme$_{\text{Cytc}}$ to Trp191 and then from Trp191 to heme$_{\text{CcP}}$, where the bridge Trp191 gets physically occupied, the rate constant of the whole reaction is limited by the lower of both corresponding single rate constants. We estimated the enthalpy of the ET from heme$_{\text{Cytc}}$ to Trp191 as ΔE = -1.06 eV through QM/MM calculations explained above. Again we applied equation (3) and (4) with bridge-mediated electronic coupling *rms* $V_{DB}(path1)$ = 1.77×10^{-5} eV between heme$_{\text{Cytc}}$ and Trp191, reaction energy change ΔE = -1.06 eV and reorganization energy λ = 0.85 eV, resulting to 5.48×10^{5} s^{-1} as the rate constant of the first step of the reaction. The reaction free energy of the ET from Trp191 to CcP we derived through ΔE = -0.92 eV -(-1.06 eV) = 0.14 eV. Together with the electronic coupling *rms* $V_{BA}(path1)$ = 1.32×10^{-2} eV and the same reorganization energy of 0.85 eV, the rate constant for the second step of the bridge-localized reaction calculates to 1.35×10^{7} s^{-1}. Thus, the rate limiting step of the two-step ET between Cytc and CpdI is the hole injection into Trp191 with a rate constant of about 5.48×10^{5} s^{-1}, while ET between Trp191 and CcP is almost instantaneous.

Our calculations show that the rate constant of the two-step mechanism (k_{ET}(two-step) = 5.48×10^{5} s^{-1}) is much higher than the rate of the single step mechanism (k_{ET}(single-step) = 9.35×10^{4} s^{-1}) denoted above as well as in Figure 5, clearly stating the true nature of Trp191 as localized bridge. The results are also in good agreement with experimental values of k_{ET} of the ET process of ruthenium-modified

cytochrome c with CpdI, measured to be greater than 5.0×10^4 s^{-1},[65] as well as another study stating the ET to be faster than the dead time of the stopped-flow method ($k_{ET} \gg 2000$s^{-1})[48] The identification of the sequential hopping ET process with Trp191 as a localized bridge state is in perfect agreement with experimental data.[48]
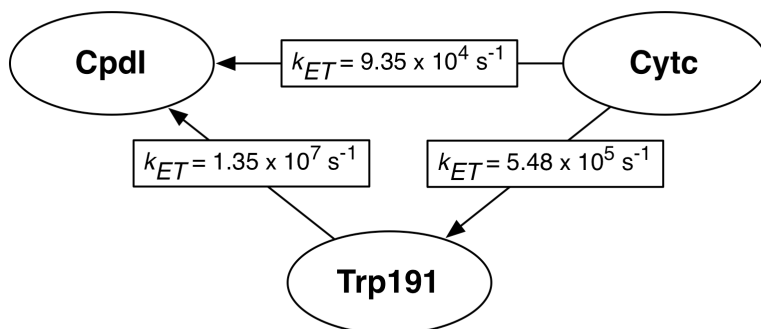
Figure 5: Calculated rate constants for single step ET process between Cytc and CpdI as well as two-step ET process between Cytc and Trp191, and Trp191 and CpdI.

4. Summary and Conclusions

We have performed a comprehensive study on the ET process in the CcP/Cytc complex. We have shown that the combination of conformational sampling and clustering of the complex structure, followed by ET pathway mapping through our recently developed QM/MM e-Pathway approach and electronic coupling calculations on derived pathways, allows obtaining a substantial picture of the underlying ET mechanism, including the identification of the ET transfer pathway as well as its quantification.

Electronic coupling calculation on distinct ET pathways derived from the QM/MM e-Pathway calculations have shown that residues Ala194, Ala193, Gly192 and Trp191 of CcP act as the main ET pathway between Cytc and CcP. Furthermore, low fluctuations between the electronic couplings calculated on individual complex conformations indicate a non-gated ET mechanism. Cleary, Trp191 appears as the most important residue in the ET region, receiving the electron hole at the first stage in almost all QM/MM e-Pathway calculations. Computed rate constant are in good agreement with experimental data, showing that the underlying ET mechanism is sequential hopping with Trp191 as localized bridge, being one order of magnitude faster than direct bridge-mediated ET between Cytc and CpdI of CcP.

Supercomputer Center, by the Spanish Ministry of Education and Science to V.G. through the project CTQ200962122 as well as to A.A.V. through the project CTQ2009-12346, and by the Catalan Agency for Management of University and Research to F.H.W. through the grant FI Mod B.

Supporting information available: Protocols of the protein preparation, MD simulation and QM/MM minimization. Short description of the QM/MM e-Pathway approach. RMSD plots of hemeCcP and hemeCytc against their QM/MM derived templates. Remarks on electronic coupling calculations on QM/MM derived properties. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1)   Beratan, D. N.; Onuchic, J. N.; Winkler, J. R.; Gray, H. B. *Science* **1992**, *258*, 1740-1741.

(2)   Langen, R.; Chang, I. J.; Germanas, J. P.; Richards, J. H.; Winkler, J. R.; Gray, H. B. *Science* **1995**, *268*, 1733-1735.

(3)   Balzani, V.; Piotrowiak, P.; Rodgers, M. A. J.; Mattay, J.; Astruc, D.; Gray, H. B.; Fukuzumi, S.; Mallouk, T. E.; Haas, Y.; de Silva, A. P.; Gould, I. R. *Electron Transfer in Chemistry*; Wiley-VCH: Weinheim, Germany, 2001; Vol. I-V.

(4)   Jortner, J.; Ratner, M. A. *Molecular Electronics*; Blackwell Scientific: Oxford, 1997.

(5)   May, V.; Kühn, O. *Charge and Energy Transfer Dynamics in Molecular Systems*; Wiley-VCH: Berlin, 2000.

(6)   Jortner, J.; Bixon, M. Electron Transfer: From Isolated Molecules to Biomolecules. In *Advances in Chemical Physics*; Wiley: New York, 1999; Vol. 106-107.

(7)   Gray, H. B.; Winkler, J. R. *Quart. Rev. Biophys.* **2003**, *36*, 341-372.

(8)   Berlin, Y. A.; Ratner, M. A. *Rad. Phys. Chem.* **2005**, *74*, 124-131.

(9)   Onuchic, J. N.; Beratan, D. N.; Winkler, J. R.; Gray, H. B. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 349-377.

(10)   Kurnikov, I. V. *Department of Chemistry, University of Pittsburgh, Pittsburgh, PA* **1999**.

(11)   Siddarth, P. *Journal of photochemistry and photobiology.A, Chemistry* **1994**, *82*, 117-121.

(12)   Guallar, V.; Wallrapp, F. *J. R. Soc. Interface* **2008**, *5*, 233-239.

(13)   Wallrapp, F.; Masone, D.; Guallar, V. *J. Phys. Chem. A* **2008**, *112*, 12989-12994.

(14)   Nocek, J. M.; Zhou, J. S.; De Forest, S.; Priyadarshy, S.; Beratan, D. N.; Onuchic, J. N.; Hoffman, B. M. *Chem. Rev.* **1996**, *96*, 2459-2490.

(15)   Ceccarelli, M.; Marchi, M. *J. Phys. Chem. B* **2003**, *107*, 5630-5641.

(16)   Morozova, O. B.; Hore, P. J.; Sagdeev, R. Z.; Yurkovskaya, A. V. *J. Phys. Chem. B* **2005**, *109*, 21971-21978.

(17)   Reece, S. Y.; Seyedsayamdost, M. R.; Stubbe, J.; Nocera, D. G. *J. Am. Chem. Soc.* **2007**, *129*, 13828-13830.

(18)   Wittekindt, C.; Schwarz, M.; Friedrich, T.; Koslowski, T. *J. Am. Chem. Soc.* **2009**, *131*, 8134-8140.

(19)   Axup, A. W.; Albin, M.; Mayo, S. L.; Crutchley, R. J.; Gray, H. B. *J. Am. Chem. Soc.* **1988**, *110*, 435-439.

(20)   Gehlen, J. N.; Daizadeh, I.; Stuchebrukhov, A. A.; Marcus, R. A. *Inorg. Chim. Acta* **1996**, *243*, 271-282.

(21)   Kawatsu, T.; Kakitani, T.; Yamato, T. *J. Phys. Chem. B* **2002**, *106*, 11356-11366.

(22)   Prytkova, T. R.; Kurnikov, I. V.; Beratan, D. N. *J. Phys. Chem. B* **2005**, *109*, 1618-1625.

(23)   Shih, C.; Museth, A. K.; Abrahamsson, M.; Blanco-Rodriguez, A. M.; Di Bilio, A. J.; Sudhamsu, J.; Crane, B. R.; Ronayne, K. L.; Towrie, M.; Vlcek, A.; Richards, J. H.; Winkler, J. R.; Gray, H. B. *Science* **2008**, *320*, 1760-1762.

(24)   Long-Range Charge Transfer in DNA (I and II). In *Topics in Current Chemistry*; Schuster, G. B., Ed.; Springer-Verlag: Berlin, New York, 2004; Vol. 236 and 237.

(25)   Arkin, M. R.; Stemp, E. D. A.; Holmlin, R. E.; Barton, J. K.; Hörmann, A.; Olson, E. J. C.; Barbara, P. F. *Science* **1996**, *273*, 475-480.

(26)   Lewis, F. D.; Liu, X.; Miller, S. E.; Wasielewski, M. R. *J. Am. Chem. Soc.* **1999**, *121*, 9746-9747.

(27)   Lewis, F. D.; Liu, X.; Liu, J.; Miller, S. E.; Hayes, R. T.; Wasielewski, M. R. *Nature* **2000**, *406*, 51-53.

(28)   Lewis, F. D.; Kalgutkar, R. S.; Wu, Y.; Liu, X.; Liu, J.; Hayes, R. T.; Miller, S. E.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2000**, *122*, 12346-12351.

(29)   Rösch, N.; Voityuk, A. A. Quantum Chemical Calculation of Donor–Acceptor Coupling for Charge Transfer in DNA. In *Long-Range Charge Transfer in DNA II*, 2004; pp 37-72.

(30)   Voityuk, A. A. *J. Phys. Chem. B* **2005**, *109*, 17917-17921.

(31)   Zheng, J.; Kang, Y. K.; Therien, M. J.; Beratan, D. N. *J. Am. Chem. Soc.* **2005**, *127*, 11303-11310.

(32)   Blancafort, L.; Voityuk, A. A. *J. Phys. Chem. A* **2006**, *110*, 6426-6432.

(33)   Félix, M.; Voityuk, A. A. *J. Phys. Chem. A* **2008**, *112*, 9043-9049.

(34)   Cave, R. J.; Newton, M. D. *Chem. Phys. Lett.* **1996**, *249*, 15-19.

(35)   Cave, R. J.; Newton, M. D. *J. Chem. Phys.* **1997**, *106*, 9213-9226.

(36)   Rust, M.; Lappe, J.; Cave, R. J. *J. Phys. Chem. A* **2002**, *106*, 3930-3940.

(37)   Voityuk, A. A. *Chem. Phys. Lett.* **2006**, *422*, 15-19.

(38)   Voityuk, A. A. *J. Chem. Phys.* **2006**, *124*, 064505-064506.

(39)   Voityuk, A. A.; Rösch, N. *J. Chem. Phys.* **2002**, *117*, 5607-5616.

(40)   Wallrapp, F.; Voityuk, A.; Guallar, V. *J. Chem. Theory Comput.* **2009**, *5*, 3312-3320.

(41)   Wallrapp, F.; Voityuk, A.; Guallar, V. *J. Chem. Theory Comput.* **2010**, *6*, 3241-3248.

(42)   Voityuk, A. A. *Chem. Phys. Lett.* **2010**, *495*, 131-134.

(43)   Harvey, J. N.; Bathelt, C. M.; Mulholland, A. J. *J. Comput. Chem.* **2006**, *27*, 1352-1362.

(44)   Pelletier, H.; Kraut, J. *Science* **1992**, *258*, 1748-1755.

(45)   Pappa, H. S.; Tajbaksh, S.; Saunders, A. J.; Pielak, G. J.; Poulos, T. L. *Biochem.* **1996**, *35*, 4837-4845.

(46)   Rosenfeld, R. J.; Hays, A.-M. A.; Musah, R. A.; Goodin, D. B. *Protein Sci.* **2002**, *11*, 1251-1259.

(47)   Erman, J. E.; Vitello, L. B. *Biochim. Biophys. Acta Prot. Struct. Mol. Enzym.* **2002**, *1597*, 193-220.

(48)   Guo, M.; Bhaskar, B.; Li, H.; Barrows, T. P.; Poulos, T. L. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5940-5945.

(49)   Liu, R.-Q.; Hahm, S.; Miller, M.; Durham, B.; Millett, F. *Biochem.* **1995**, *34*, 973-983.

(50)   Hays Putnam, A.-M. A.; Lee, Y.-T.; Goodin, D. B. *Biochem.* **2009**, *48*, 1-3.

(51)   Musah, R. A.; Goodin, D. B. *Biochem.* **1997**, *36*, 11665-11674.

(52)   Musah, R. A.; Jensen, G. M.; Rosenfeld, R. J.; McRee, D. E.; Goodin, D. B.; Bunte, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 9083-9084.

(53)   Sivaraja, M.; Goodin, D. B.; Smith, M.; Hoffman, B. M. *Science* **1989**, *245*, 738-740.

(54)   Miller, M. A.; Vitello, L.; Erman, J. E. *Biochem.* **1995**, *34*, 12048-12058.

(55)   Barrows, T. P.; Bhaskar, B.; Poulos, T. L. *Biochem.* **2004**, *43*, 8826-8834.

(56)   Seifert, J. L.; Pfister, T. D.; Nocek, J. M.; Lu, Y.; Hoffman, B. M. *J. Am. Chem. Soc.* **2005**, *127*, 5750-5751.

(57)   Summers, F. E.; Erman, J. E. *J. Biol. Chem.* **1988**, *263*, 14267-14275.

(58)   Corin, A. F.; Hake, R. A.; McLendon, G.; Hazzard, J. T.; Tollin, G. *Biochem.* **1993**, *32*, 2756-2762.

(59)   Wang, Y.; Margoliash, E. *Biochem.* **1995**, *34*, 1948-1958.

(60)   Matthis, A. L.; Erman, J. E. *Biochem.* **1995**, *34*, 9985-9990.

(61)   Miller, M. A. *Biochem.* **1996**, *35*, 15791-15799.

(62)   Hazzard, J. T.; Poulos, T. L.; Tollin, G. *Biochem.* **1987**, *26*, 2836-2848.

(63)   Hazzard, J. T.; McLendon, G.; Cusanovich, M. A.; Das, G.; Sherman, F.; Tollin, G. *Biochem.* **1988**, *27*, 4445-4451.

(64)   Hazzard, J. T.; McLendon, G.; Cusanovich, M. A.; Tollin, G. *Biochem. Biophys. Res. Commun.* **1988**, *151*, 429-434.

(65)   Geren, L.; Hahm, S.; Durham, B.; Millett, F. *Biochem.* **1991**, *30*, 9450-9457.

(66)   Hazzard, J. T.; Moench, S. J.; Erman, J. E.; Satterlee, J. D.; Tollin, G. *Biochem.* **1988**, *27*, 2002-2008.

(67)   QSite; 5.6 ed.; Schrödinger, LCC: New York, NY, 2010.

(68)   Truhlar, D. G.; Zhao. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

(69)   Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435-447.

(70)   Theodoridis, S.; Koutroumbas, K. *Pattern Recognition, Third Edition*; Academic Press, Inc.: Orlando, FL, USA, 2006.

(71)   Newton, M. D. *Chem. Rev.* **1991**, *91*, 767-792.

(72)   Marcus, R. A.; Sutin, N. *Biochimica Et Biophysica Acta* **1985**, *811*, 265-322.

(73)   Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1980**, *255*, 8199-8205.

(74)   Bonagura, C. A.; Bhaskar, B.; Shimizu, H.; Li, H.; Sundaramoorthy, M.; McRee, D. E.; Goodin, D. B.; Poulos, T. L. *Biochem.* **2003**, *42*, 5600-5608.

(75)   Bhaskar, B.; Poulos, T. L. *J. Biol. Inorg. Chem.* **2005**, *10*, 425-430-430-425-430-430.

(76)   Conklin, K. T.; McLendon, G. *J. Am. Chem. Soc.* **1988**, *110*, 3345-3350.

Supporting Information

Electron transfer in the Cytochrome c Peroxidase/Cytochrome c couple

Frank Wallrapp, Alexander Voityuk and Victor Guallar

S1: Preparation of the CpdII/CytcRED complex

We took chain A (Ccp) and B (Cytc) from the Protein Data Bank structure 2PCC for our calculations. We manually replaced the existing sulfate ligand of CcP by the oxo ligand being 6-coordinated to the Fe(IV) of CcP. The protonation states of the histidines in CcP and Cytc were assessed by visual inspection. In particular, histidines 6, 52 and 60 of CcP and 39 of Cytc were protonated on the delta position, histidines 96 of CcP and 26 of Cytc were protonated on the epsilon position and histidines 181 of CcP and 33 of Cytc were doubly protonated. The system was soaked and equilibrated using Impact[1], applying OPLS and the SPC solvent model together with 100 ps of molecular dynamics (MD) following a truncated Newton minimization. Both hemes and their ligated residues were kept frozen during the equilibration process. A layer of 10 Å of water around the proteins was kept as solvent. In order to gain the exact conformation of the system having both active sites in the reduced oxidation state, we QM/MM minimized the system in an oscillating approach. First we QM/MM minimized CcpII by having the heme, the oxo atom and the ligated His175 in the QM region (-4 charge, triplet spin state) while keeping frozen Cytc as well as all waters more than 5 Å away from CcP. The QM/MM boundary was treated by a hydrogen cut at the sidechain of His175. Within a second step we QM/MM minimized Cytc by only having the heme and its ligated residues in the QM region (-2 charge, singlet spin state) while keeping frozen CcP as well as all waters more than 5 Å away from Cytc. The QM/MM boundary was defined by an hydrogen cut at the sidechains of His18 and Met80 of Cytc as well as at the cystein bridges of Cys14 and Cys17 being linked to the hemeCytc. The procedure was repeated between CcP and Cytc until final convergence of both.

S2: MD simulation on CpdII/Cytc

In order to apply the molecular dynamics software GROMACS together with the OPLS force field we created residue templates for CpdII and CytcRED. We took the ESP charges, atom lengths and angles derived from QM/MM minimized structures applying DFT/B3LYP with the 6-31g*/lacvp basis set as implemented in the QSite program.[2] Force constants were taken from an OPLS heme-dioxy template published by Gogonea et al..[3] The templates were modified to incorporate the oxo atom and ligated His175 for CcP and ligated His18 and Met80 as well as the cystein bridges of Cys14 and Cys17 for Cytc. Force constants involving the six-coordinated irons were taken from the literature[4] respectively defined through

vibrational analysis on a QM model of CpdII in the case of the iron-oxo bond (v=954.9, mass=12.4, k=666105.5 kJ/(mol nm2)). Starting from the QM/MM minimized structure, we performed a 2 ns NVT trajectory of the system in aqueous solvent at 289.15 K, following solvation in SPC including a 0.15 M salt concentration, a truncated Newton minimization and 20 ps of equilibration. Snapshots were saved every single ps.

We plotted the RMSD of the porphyrin and the 6-coordinated atoms of $heme_{CcP}$ (panel A in Figure 1) and $heme_{Cytc}$ (panel B in Figure 1) against their respective QM/MM templates throughout the MD simulation. The RMSD of $heme_{CcP}$ has a mean value of 1.12 Å with a standard deviation of only 0.03, indicating very little fluctuations of the heme throughout the MD simulation. For hemeCytc we have a similar mean RMSD of 0.13 Å and a standard deviation of 0.02, showing even less fluctuation in the trajectory. From both results we take that $heme_{CcP}$ as well as $heme_{Cytc}$ stay in their native conformation and no QM/MM minimization is needed for the electronic coupling calculations. Preliminary test on snapshot 0 by a complete QM/MM minimization of hemeCcP and hemeCytc in their respective reduced states, have also confirmed this. The electronic coupling between both hemes is very similar before and after the minimization: *rms V_{DA}(direct*, noOpt) = 9.19×10^{-9} against *rms V_{DA}(direct*, QM/MM-Opt) = 6.38×10^{-9}, as well as *rms V_{DA}(path2*, noOpt) = 1.76×10^{-7} against *rms V_{DA}(path1*, QM/MM-Opt) = 2.56×10^{-7}.
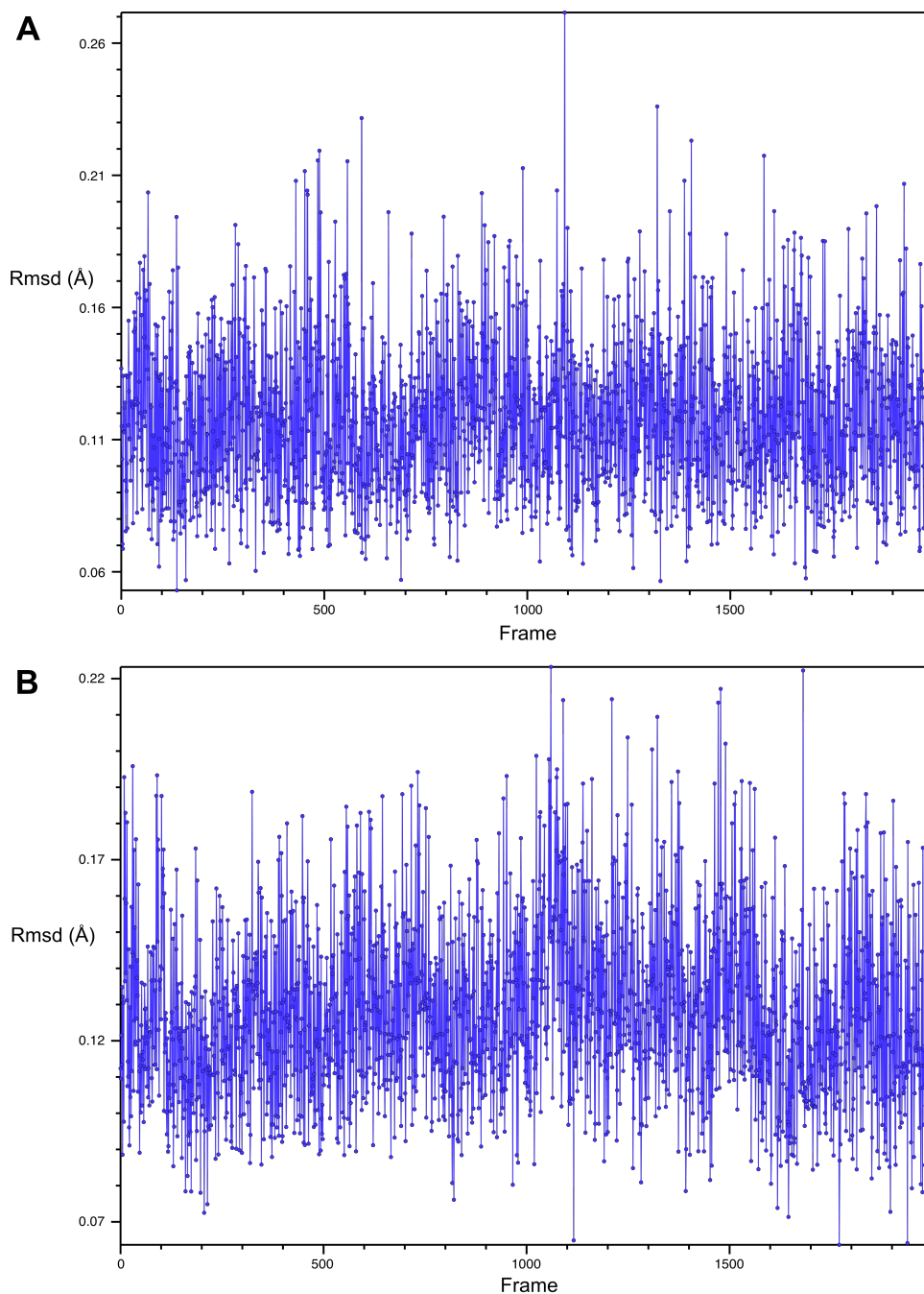
Figure 1: RMSD of hemeCcP (panel A) and hemeCytc (panel B) against their QM/MM respective templates throughout the MD simulation.

S3: QM/MM e-Pathway

In order to track electron migration, we have proposed a new algorithm called the QM/MM e-Pathway.[5] The underlying idea is that regions contained in the MM part cannot allocate electrons due to the lack of electronic description. Hence, if an unpaired electron is present in the system it must necessarily occupy the quantum region. The method's strategy is based on modifying the QM region following the evolution of the spin density. In case of an electron transfer pathway between two distant redox centers the methodology consists of the following procedure. As the initial setup, the parameters for the donor and the acceptor in the neutral state are derived, meaning both sites are reduced in our case of hole transfer. Here, the acceptor has taken itself an electron from the ET region but the donor has not given his electron away, yet. The parameterization consists of a QM/MM energy minimization of the donor and acceptor (both in their respective reduced state) from which we extract the geometry and atomic charges from the electrostatic potential. This parameterization is important because in the next step of the procedure both residues are going to be included only in the classical region, now only influencing the QM region by structural constraints as well as electrostatic and van der Waals interactions. Thus, there is going to be no electronic description of the donor and acceptor. Instead, the focus is on the "transfer region" in-between, the region that now contains the electron hole. After this setup an iterative process is started where initially the entire transfer region is included in the QM region of the QM/MM calculation. The first acceptor of the hole is found by the localization of the spin density given by a single point calculation of the system missing one electron thus having a doublet spin state. Proceeding to the next iteration, the previously identified residue is turned into a classical residue by excluding it from the quantum region. In doing so, the method does not allow for an electronic description of it and thus, the hole needs to find its next host. The QM/MM boundary is treated by hydrogen cuts placed onto the backbone of the protein residues. The procedure is repeated until the identified residues connect to a direct pathway between the donor and acceptor.

S4: Electronic coupling calculations

It is known that hydrogen cuts can disturb the partial charges of QM atoms in their vicinity.[6] Consequently, their exact placement is crucial for the correct application of FCD, being based on the charge separation of donor and acceptor sites. Preliminary tests have shown that placing the hydrogen cuts at least 3 bonds away from sites of interest do not affect their partial charges and result in correct coupling calculations. An example are hydrogen cuts from Trp191:N to Pro190:C and from Trp191:C to Gly192:N in order to incorporate the side chain of Trp191.

References:

(1)   Impact; 5.6 ed.; Schrödinger, LCC: New York, NY, 2010.

(2)   QSite; 5.6 ed.; Schrödinger, LCC: New York, NY, 2010.

(3)   Gogonea, V.; Shy, J. M.; Biswas, P. K. *J. Phys. Chem. B* **2006**, *110*, 22861-22871.

(4)   Autenrieth, F.; Tajkhorshid, E.; Baudry, J.; Luthey-Schulten, Z. *J. Comput. Chem.* **2004**, *25*, 1613-1622.

(5)   Wallrapp, F.; Masone, D.; Guallar, V. *J. Phys. Chem. A* **2008**, *112*, 12989-12994.

(6)   Ferré, N.; Olivucci, M. *J. Mol. Struct. (Theochem)* **2003**, *632*, 71-82.

THESIS PUBLICATIONS

4. DISCUSSION

DISCUSSION

The focus of this PhD thesis lies on electron transfer (ET) processes in biochemistry being involved in almost all enzymatic cycles, including respiration, photosynthesis, DNA damage repair and detoxification, among many others. Besides gathering knowledge through fundamental research on existing systems, investigation of ET in proteins has also impact on improvement of medical science and agricultural technology, on development of molecular electronic devices, and on the usage of alternative energy sources, for example. Although the phenomenon of ET between small molecules in solution is reasonably well described, we are still not able to understand, not to mention predict, the factors that influence ET reactions in proteins in satisfying detail. Consequently, much effort is spent to understand the detailed mechanism underlying ET reactions in and between proteins.

In this regard, the main objective of this PhD thesis was the development of a methodology for the quantitative study of ET in biological systems. We addressed this global aim in a gradual process through several studies carried out during this PhD thesis. While all resulting publications are represented in the previous chapter (chapter 3), the results and outcomes of this thesis will be critically discussed in the following. We start with a review on state-of-the-art QM/MM approaches and their application to protein, followed by a section about the development and application of our novel approach to map ET pathways, QM/MM e-Pathway. Furthermore, we discuss important findings in the investigation of the influence of conformational dynamics and solvent onto the electronic coupling in model systems. Finally, we critically evaluate results obtained in a quantitative study on ET in the CcP/Cytc complex. Here, we combined all beforehand gathered knowledge by the application of our novel QM/MM e-Pathway approach, electronic coupling calculations as well as calculation of kinetical and thermodynamical factors incorporating conformational dynamics, in order to compute the ET rate constant of the system.

4.1. State-of-the-art QM/MM methods and their application to protein

In the introduction of this thesis we have underlined the aim of theoretical investigation on ET processes in order to understand the underlying mechanisms and construct and train computational approaches capable of predicting a priori the experimental outcomes. We further emphasized that computational approaches are much cheaper, faster and in the majority of the cases less intricate than *in-vitro* and especially *in-vivo* experiments, and are able to give chronological and atomic resolution often unreached by experimental trials. In our particular case, the usual experimental procedure to determine the electron transfer pathway is by means of direct protein mutation, which can probe the importance of some residues in the electron migration, but has the great disadvantage of possible secondary effects. These effects, such as structural changes or significant perturbation of the protein electrostatics, might mask the real cause for the loss of the electron conductivity; hence possibly falsify the results. Here, computational approaches promise relief, as with them it is much easier to "mutate" respectively "knock-out" single residues in terms of ET without any perturbation of the protein structure. Nevertheless,

computational approaches clearly have limitations, too. They generally use simplified setups due to the mere complexity of biological systems. Also, not all approaches are based on first principles but on semiempirical or molecular mechanics force fields introducing incertitude and even errors. Thus, we should keep in mind that theoretical approaches only simulate real nature. Hence, the obtained results should always be critically assessed and compared against experimental outcome.

In order to simulate ET processes it is necessary to apply quantum chemical approaches that are able to describe the electronic state of the systems. Within this matter, QM/MM techniques offer a very valuable tool, as they treat the active site (respectively the site of interest, i.e. the ET region) with a robust *ab initio* quantum mechanics methodology and the remainder of the protein by means of a much cheaper molecular mechanics model. Within the two publications presented in sections 3.1.1 and 3.1.2 we review state-of-the-art QM/MM methods together with recent applications on protein. Here, the first review addresses not only the QM/MM methodology itself but also the complete procedure of deriving appropriate QM/MM simulations of the enzymatic systems under investigation. We especially attach importance to three steps: (1) choosing a model, (2) conformational sampling and (3) mapping the chemical process. Although most studies nowadays are based on good models and map the chemical process with the strictness needed, they are still missing sufficient conformational sampling. Consequently, we emphasize the need to investigate high frequency nuclear motion as well as large-scale protein dynamics being coupled with the enzymatic processes, and therefore propose the application of Monte Carlo or protein structure prediction techniques next to standard molecular mechanics. The second review focuses on QM/MM approaches in order to investigate heme proteins biochemistry. In presenting results on ligand migration coupled to the ligand binding event, spin density localization in compound I of cytochromes and peroxidases, mapping ET pathways, and reaction barrier calculations on Tryptophan 2,3-dioxygenase, we underline the possibilities of mixed quantum mechanics and molecular mechanics methods in addressing the enzymatic mechanisms in proteins.

4.2. Development and application of the QM/MM e-Pathway approach to map electron transfer pathways in protein

In order to quantify ET processes in protein it is necessary to know the pathway an electron is going to take on its way from the donor to the acceptor. To the best of our knowledge, there exist only the methods described in the introduction being able to search a priori for ET pathways in protein. Although recent attempts to incorporate robust *ab initio* DFT to compute the decay rates of electronic coupling, the underlying concept of multiplying predefined decay constants along the pathway have the clear disadvantage of possibly missing important factors (i.e. inference between neighbored residues) only present in the actual structure under investigation (de la Lande and Salahub 2010). Consequently, we aimed at the

development of a methodology being able to directly map ET pathways in given system snapshots of proteins, incorporating a sturdy level of theory

Applying robust QM/MM methodology, our novel method, called QM/MM e-Pathway, describes the electron transfer region at the QM level and the rest of the system (including donor and acceptor) at the MM level. The approach is based on the theory that the electron has already left the donor but has not arrived in the acceptor yet, thus, it is localized in the transfer region. Note, this only accounts for excess electron transfer; for hole transfer the donor still has its electron and the acceptor has already derived one electron from the transfer region, now hosting an electron hole. Consequently, before starting the iterative procedure, parameters for donor and acceptor are derived from QM/MM energy minimizations, both sites in their oxidized state (respectively reduced state in case of hole transfer). The iterative procedure starts by finding the first acceptor of the electron (hole) through localization of the spin density in the ET region, given by a single point calculation of the system having one extra electron (respectively one electron missing in case of hole transfer) and a doublet spin state. In the next iteration, we exclude the previously identified residue from the QM region, turning it into a classical residue. Therefore, an electronic description of this residue is no longer possible and thus, the electron (hole) needs to find its next host. The method continuous until donor and acceptor can be connected through a direct pathway built upon the identified residues. The publication describing the methodology is given in section 3.2.

In order to evaluate our QM/MM e-Pathway we chose the P450cam/Pdx complex, where Arg112 or P450cam and Asp48 of Pdx were already proposed to be crucial for the ET process. Our calculations were able to reproduce these findings and furthermore showed that all derived pathways directly connect into Cys357 being ligated to the iron of the heme, indicating that there is no involvement of the heme propionates in ET process. Again, these results are in agreement with experimental data. The corresponding publication of this study is given in section 3.3. All in all, we presented a novel approach based on *ab initio* QM/MM methodology for mapping ET pathways in protein. Moreover, we proved its usefulness as well as its applicability to identify residues important for ET in the study on the P450cam/Pdx complex.

The method finds currently application in several other studies on ET processes in proteins. Among them are the intra-complex ET in Cytochrome c Oxidase, the ET in the Adrenodoxin reductase/Adrenodoxin complex and the ET in the Ferredoxin NADP+ reductase/Ferredoxin complex.

4.3. Electronic coupling and the influence of conformational dynamics on it

In the previous section we discussed our novel approach to map long-range ET pathways in protein. In order to quantify the importance of these pathways, we compute the electronic coupling of donor and acceptor incorporating the particular pathway residues. As denoted in section 1.2.2, the electronic coupling describes the probability of the ET to occur, once the activation energy of the reaction is established. For several systems it has been found that the conformational dynamics

of the protein strongly influences the electronic coupling between the donor and acceptor. In order to study this influence as well as approve the applicability of the GMH and FCD method on protein, we performed two combined quantum chemical/molecular dynamics studies on electronic coupling between tryptophan-based donor and acceptor in oligopeptide model systems of variable length. In particular, we studied the dependence of electronic coupling values on conformational parameters and solvent (see section 3.4.1) as well as on temperature (see section 3.4.2).

In our first study on solvent effects on donor-acceptor couplings in peptides, we performed MD on Trp-(Pro)n-Trp (n = 1 to 6) molecules in gas phase and aqueous solvent. In particular, we applied both, INDO/S as well as *ab initio* Hartree-Fock, together with the GMH approach to compute the electronic coupling for thermal hole transfer. Comparing the results of both levels of theory, our findings indicate a high accuracy of the semi-empirical INDO/S method for calculating the electronic coupling in simple peptides. We show that the coupling values strongly vary throughout the molecular dynamics trajectory. Furthermore, our results indicate that the presence of solvent affects the mechanism of electron transfer through restriction of the conformational space. In contrast to solvent calculations that establish a nongated mechanism dominated by bridge-mediated coupling, gas phase simulations show gated electron transfer dominated by direct through-space coupling due to strong conformational changes bringing donor and acceptor in close vicinity. Finally in agreement with experimental data, our results point to a distance of approximately 20 Å as the possible point for the transition from superexchange to hopping ET mechanism.

In our second study on temperature effects on donor-acceptor coupling in peptides, we performed 10 ns replica exchange molecular dynamics on Trp-(Pro)n-Trp (n = 3 or n = 6) in gas phase. In particular, parameters for thermal hole transfer were calculated by using INDO/S together with the FCD scheme. We demonstrated that systems with only few degrees of conformational freedom (Trp-(Pro)3-Trp) can show unexpected temperature dependence of electronic coupling. In more detail, there is a local peak of the electronic coupling at a temperature of 144 K followed by a monotonic decrease with higher temperature. In contrast, systems with more degrees of conformational freedom (Trp-(Pro)6-Trp) do not show local peaks but purely monotonically decreasing electronic coupling. We found out that the underlying reason is the donor-acceptor distance of conformations in the trajectories at the respective temperature. In our systems, mainly short distance gives rise to high coupling. Consequently, the coupling follows closely the changes in donor-acceptor distances, being shortest for trajectories of Trp-(Pro)3-Trp at a temperature of 144 K and longer at other temperatures. In Trp-(Pro)6-Trp, short distances are already reached in conformations at very low temperatures and constantly increase from there with increasing temperature. Within both model systems, only a few conformations with strong donor-acceptor coupling are shown to be crucial of the overall ET rate, which indicates a gated ET mechanism.

In order to summarize, both studies revealed that electronic coupling values strongly fluctuate throughout the molecular dynamics trajectories obtained, and the

mechanism of ET is affected by the conformational space the system is able to occupy. Systems being capable of bringing donor and acceptor in close vicinity show gated electron transfer dominated by direct through-space coupling while systems of less flexibility establish a nongated mechanism dominated by bridge-mediated coupling. Here, our results have shown the clear importance of sufficient sampling in order to identify the underlying mechanism of ET in particular systems. Moreover, both studies stated the applicability of both methods, GMH and FCD, in electronic coupling calculations in proteins, deriving the properties of the adiabatic states for the hole transfer from the occupied molecular orbitals of the corresponding neutral system through incorporation of the one-electron approximation.

4.4. Application of QM/MM e-Pathway together with electronic coupling and rate constant calculations in the CcP/Cytc complex

The last and major objective of this dissertation was the application and evaluation of our newly developed electron transfer pathway mapping method together with electronic coupling and rate constants calculations in real protein, in order to fulfill the main goal of developing a methodology for a quantitative study of ET in biological systems.

This objective was addressed in section 3.5, where we report on the ET investigation on the CcP/Cytc complex. We explicitly chose this protein couple due to its suitability as a good test case for our purposes. It is a sophisticated example of intra-complex ET between two separate proteins, having a donor-acceptor distance of about 27 Å, for example. Nevertheless, the system is under investigation over many years now, resulting in a great pool of information available. Most importantly, the crystal-structure of the protein complex is available, allowing theoretical studies on the ET process in atomistic detail. Furthermore, there is comprehensive knowledge on Trp191 as localized bridge state of the ET between Cytc and CcP, as well as a proposed ET pathway. Additionally, a wide range of experiments was performed, giving ET rate constants to compare against. All in all, the system provides excellent data for the application as well as the evaluation of our methodology.

Our results of QM/MM e-Pathway clearly indicate the important role of Trp191 in the ET process. The approach finds Trp191 in 7 out of 11 snapshots to be the first hole acceptor in the ET region. Furthermore, we are able to reproduce the proposed ET pathway from Cytc to CcP following Ala194, Ala193, Gly192 and Trp191 of CcP. Additionally we found an alternative pathway through residues Ala81 and Phe82 of Cytc, Asn196, Asn195, Leu177 and Ala176 of CcP as proposed by Siddarth. FCD calculations on derived pathways clearly distinguished both alternatives, with the first pathway having much higher electronic coupling between the donor and acceptor. Finally, by incorporating kinetical factors such as the enthalpy as well as reorganization energy of the ET process, we were able to approximate the ET rate constants in the complex and in particularly along the

proposed pathways. The results show that ET between Cytc and CcP is in fact a two-step ET process from Cytc to Trp191 (through residues Ala194, Ala193 and Gly192 of CcP) and then from Trp191 to CcP, with the first step being the rate limiting one. Moreover, low fluctuating electronic coupling over several snapshots revealed a nongated ET mechanism in the CcP/Cytc complex.

In a nutshell, we performed a comprehensive study on the CcP/Cytc couple in which we have shown that the combination of conformational sampling and clustering of the complex structure, followed by ET pathway mapping through our recently developed QM/MM e-Pathway approach together with electronic coupling calculations, allows obtaining a substantial picture of the underlying ET mechanism, including the identification as well as the quantification of the ET pathway.

4.5.   Limitations

The research carried out in this thesis is based on theoretical and computational approaches that per se bare some drawbacks. In the following we will discuss in more detail some of the major limitations of the methodologies we developed and applied.

Due to its mere complexity it is necessary to simplify a biological system. Current computational power and implementation allows only precise QM calculations of about 400 to 500 atoms with high level of theory such as DFT within reasonable time on a normal computer cluster of 16 to 32 processors. Consequently, the remaining part of the system, being usually much larger than the QM region, has to be described by means of less accurate and thus cheaper level of theory, i.e. by MM. In order to apply the QM/MM methodology, the boundary between the QM and MM regions has to be defined, for instance through hydrogen cuts. It is already known to the computational community that the exact location of hydrogen cuts is very crucial for certain properties in their vicinity such as spin density as well as partial charges. Consequently, we placed hydrogen cuts very carefully and always with a minimum distance of three molecular bonds to sites of interest, i.e. fragments considered as donor or acceptor for electronic coupling calculations. Moreover, a vague knowledge of the underlying ET pathway(s) between the donor and acceptor is of great advantage in order to narrow down the QM region and consequently to reduce the above-mentioned computational costs of the calculations. Furthermore, we are aware that there exist more exact QM methods than the here applied DFT, though we obtained good results comparable to experimental data within reasonable computational time.

From the electronic coupling calculations in model systems we know that conformational dynamics can have strong influence on ET mechanisms. Consequently, conformational sampling is crucial in order to gather the true nature of the ET process. Clearly, compared to model systems, conformational sampling and subsequent QM/MM calculations of real protein are limited due to the size of the system. Nevertheless, since we observed less fluctuation in the protein complex, fewer samples were necessary to identify the true nature of the ET mechanism.

All in all, despite the here discussed limitations, we observed good agreement of our results with experimental outcomes. Hence, we are confident that our methods can help to get a comprehensive picture of ET mechanisms in biological systems.

4.6. Summary and outlook

In this PhD thesis we have addressed the sophisticated problem of getting direct information of the underlying mechanism and actual pathway of ET processes. Gradually following the main objective of this thesis (development of a methodology for the quantitative study of ET in biological systems) we developed a novel approach to map long-range electron transfer pathways, called QM/MM e-Pathway. The method is based on a successive search for important ET residues in terms of modifying the QM region following the evolution of the spin density of the electron (hole) within a given transfer region. We evaluated and proved the applicability of the algorithm on the P450cam/Pdx complex, confirming knowledge known from the literature. As a next step, we focused on the quantification of identified ET pathways in terms of electronic coupling calculations. In two separated studies on electronic coupling in oligopeptide model systems, we investigated how the electronic coupling and even the mechanism of ET is affected by the conformational dynamics of the system. Moreover, through both studies we proved the applicability of both methods, GMH and FCD, in electronic coupling calculations on proteins, deriving the properties of the adiabatic states for the hole transfer from the occupied molecular orbitals of the corresponding neutral system through incorporation of the one-electron approximation.

Within the final study on the ET process in the CcP/Cytc couple, we applied the combination of both, ET mapping and electronic coupling calculations. Our findings indicate that our QM/MM e-Pathway methodology is able to identify not only the key residue within the ET region but also the correct ET pathway. Furthermore, it is shown that we are able to quantify distinct ET pathways and give reliable estimates of ET rate constants in agreement to experiments through the application of electronic coupling calculations together with the computational assessment of kinetical factors. Moreover, we have depicted some critical points which have to be taken care of within the application of the methodologies, such as hydrogen cuts and maximum atoms in the QM region.

In summary, we developed a methodology to investigate long-range ET processes in protein, which we critically and successfully evaluated on sophisticated protein-protein complexes. Thus, it is greatly hoped that the work done in this PhD thesis will eventually help to improve the understanding of the underlying mechanisms of ET in biological systems. Further application in different systems, such as the intra-complex ET in Cytochrome c Oxidase, the ET in the Adrenodoxin reductase/Adrenodoxin complex and the ET in the Ferredoxin NADP+ reductase/Ferredoxin complex, is currently in progress.

DISCUSSION

5. CONCLUSIONS

CONCLUSIONS

1 . Current QM/MM methods and their application to proteins have been critically evaluated. We proposed three crucial steps in order to simulate enzymatic systems: (1) Choosing a suitable model, (2) conformational sampling of the protein and (3) careful mapping of the chemical process. In conclusion, we emphasize the great suitability of QM/MM methods when incorporating these steps and provide several example studies.

2 . A novel approach to map electron transfer pathways in proteins, called QM/MM e-Pathway, has been developed. It is based on an iterative procedure to identify important residues in the electron transfer region by applying *ab initio* QM/MM technology. The usefulness and applicability of the method was demonstrated in several examples.

3 . Application of the QM/MM e-Pathway approach on several snapshots of the P450cam/Pdx complex identified the important role of Arg112 of P450cam and Asp48 of Pdx in the electron transfer. Furthermore, identified pathways directly connect to Cys357, indicating no involvement of heme propionates in the electron transfer. These findings were confirmed through the literature.

4 . A systematic evaluation of the influence of solvent on the electronic coupling in model systems has been performed. Our findings revealed strong fluctuations throughout derived molecular dynamics trajectories as well as the mechanism of electron transfer being affected by the presence of solvent through restriction of conformational space. In particular, gas-phase calculations showed gated electron transfer dominated by direct through-space coupling due to strong conformational changes bringing donor and acceptor in close vicinity. Solvent calculations establish a nongated mechanism dominated by bridge-mediated coupling.

5 . A systematic evaluation of the influence of temperature on the electronic coupling in model systems has been performed. Conformational analysis of temperature-based trajectories derived from replica exchange molecular dynamics revealed that electronic coupling becomes temperature dependent when incorporating structural dynamics of the system. In particular, short oligopeptides, having only few degrees of freedom, show complex temperature dependence of electronic coupling while more flexible long oligopeptide show monotonic dependence.

6 . The strong influence of conformational dynamics on the electronic coupling emphasizes the clear importance of conformational sampling within ET investigations.

CONCLUSIONS

7. We approved the applicability of the GMH and FCD methods in electronic coupling calculations in proteins, deriving the properties of the adiabatic states for the hole transfer from the occupied molecular orbitals of the corresponding neutral system through incorporation of the one-electron approximation.

8. Application of QM/MM e-Pathway and FCD on the CcP/Cytc complex confirms Trp191 as localized bridge in the ET process as well as the ET pathway from Cytc through Ala194, Ala193, Gly192 and Trp191 to CcP. Computed ET rate reveal a two-step mechanism of ET from Cytc to Trp191 followed by ET from Trp191 to CcP, with rate constants in good agreement to the literature. The calculations establish a nongated ET mechanism.

9. The combination of conformational sampling and clustering of the complex structure, followed by ET pathway mapping through the QM/MM e-Pathway approach together with electronic coupling calculations, allows obtaining a substantial picture of the underlying ET mechanism, including the identification as well as the quantification of the ET pathway.

6. LIST OF PUBLICATIONS

LIST OF PUBLICATIONS

Articles

1 . **Wallrapp F.**, Voityuk, A., and Guallar, V., "Electron transfer in the Cytochrome c Peroxidase/Cytochrome c complex." in preparation (2011)

2 . **Wallrapp F.** and Guallar V., "Mixed quantum mechanism and molecular mechanics methods: Looking inside enzymes" *WIREs Computational Molecular Science*, in press (2011).

3 . Guallar, V. and **Wallrapp F.**, **"**QM/MM methods: Looking inside heme proteins biochemistry." *Biophysical Chemistry* **149** (1-2): 1-11 (2010).

4 . **Wallrapp, F.**, Voityuk, A., and Guallar, V., "Temperature Effects on Donor−Acceptor Couplings in Peptides." *Journal of Chemical Theory and Computation* **6** (10): 3241-3248 (2010).

5 . **Wallrapp, F.**, Voityuk, A., and Guallar, V., "Solvent Effects on Donor−Acceptor Couplings in Peptides. A Combined QM and MD Study." *Journal of Chemical Theory and Computation* **5** (12): 3312-3320 (2009).

6 . **Wallrapp, F.**, Masone, D., and Guallar, V., "Electron Transfer in the P450cam/PDX Complex. The QM/MM e-Pathway." *Journal of Physical Chemistry A* **112** (50): 12989-12994 (2008).

7 . Guallar, V. and **Wallrapp, F.**, "Mapping protein electron transfer pathways with QM/MM methods." *Journal Royal Society, Interface* **5** (Suppl. 3): 233-239 (2008).

Oral communications

8 . **Wallrapp F.**, Voityuk, A., and Guallar, V., "Protein-protein electron transfer pathways and coupling constants", *239th ACS National Meeting*, San Francisco, USA, March 21st – 25th, 2010.

9 . **Wallrapp F.** and Guallar, V., "QM/MM studies of electron transfer in the P450cam/PDX complex", *24th Reunion of the Theoretical and Computational Chemistry of Catalunya*, Gerona, Spain. June 26th – 27th, 2008.

10 . Guallar V. and **Wallrapp F.**, "Heme Electron Transfer: The propionate pathway", *2nd CCPB annual conference*, Bristol Zoo, UK, January 7th – 9th, 2008.

Poster communications

11 . **Wallrapp F.** and Guallar, V., "Mapping electron transfer pathways with QM/MM methods. Application on P450-PDX and CCP-cytC complexes", *Great Challenges in Computational Biology*, Barcelona, Spain, June 2nd – 4th, 2008.

7. REFERENCES

REFERENCES

"Maestro". New York, NY, Schrödinger, LCC (2010).

Altschul, A. M., R. Abrams, et al., "Cytochrome c peroxidase." *Journal of Biological Chemistry*, **136**: 777-794 (1940).

Axup, A. W., M. Albin, et al., "Distance dependence of photoinduced long-range electron transfer in zinc/ruthenium-modified myoglobins." *Journal of the American Chemical Society*, **110** (2): 435-439 (1988).

Balabin, I. A., D. N. Beratan, et al., "Persistence of Structure Over Fluctuations in Biological Electron-Transfer Reactions." *Physical Review Letters*, **101** (15): 158102-158106 (2008).

Balabin, I. A. and J. Onuchic, "Dynamically Controlled Protein Tunneling Paths in Photosynthetic Reaction Centers." *Science*, **290** (5489): 114-117 (2000).

Balzani, V., P. Piotrowiak, et al., *Electron Transfer in Chemistry*. Weinheim, Germany, Wiley-VCH (2001).

Barbosa, L. C. A., M. E. Rocha, et al., "Synthesis of 3-(4-Bromobenzyl)-5-(aryl methylene)-5H-furan-2-ones and Their Activity as Inhibitors of the Photosynthetic Electron Transport Chain." *Journal of Agricultural and Food Chemistry*, **55** (21): 8562-8569 (2007).

Barrows, T. P., B. Bhaskar, et al., "Electrostatic Control of the Tryptophan Radical in Cytochrome c Peroxidase." *Biochemistry*, **43** (27): 8826-8834 (2004).

Bathelt, C. M., A. J. Mulholland, et al., "QM/MM studies of the electronic structure of the compound I intermediate in cytochrome c peroxidase and ascorbate peroxidase." *Dalton Transactions* (21): 3470-3470 (2005).

Beratan, D. N., J. N. Onuchic, et al., "Electron tunneling pathways in ruthenated proteins." *Journal of the American Chemical Society*, **112** (22): 7915-7921 (1990).

Beratan, D. N., J. N. Onuchic, et al., "Electron-Tunneling Pathways In Proteins." *Science*, **258** (5089): 1740-1741 (1992).

Berlin, Y. A., A. L. Burin, et al., "On the Long-Range Charge Transfer in DNA." *Journal of Physical Chemistry A*, **104** (3): 443-445 (2000).

Berlin, Y. A. and M. A. Ratner, "Intra-molecular electron transfer and electric conductance via sequential hopping: Unified theoretical description." *Radiation Physics and Chemistry*, **74** (3-4): 124-131 (2005).

Betts, J. N., D. N. Beratan, et al., "Mapping electron tunneling pathways: an algorithm that finds the "minimum length"/maximum coupling pathway between electron donors and acceptors in proteins." *Journal of the American Chemical Society*, **114** (11): 4043-4046 (1992).

Binder, K. and D. M. Heermann, *Monte Carlo Simulation in Statistical Physics, An Introduction.* Berlin, Springer-Verlag (1992).

Bingham, R. C., M. J. S. Dewar, et al., "Ground states of molecules. XXV. MINDO/3. Improved version of the MINDO semiempirical SCF-MO method." *Journal of the American Chemical Society*, **97** (6): 1285-1293 (1975).

Bixon, M., B. Giese, et al., "Long-range charge hopping in DNA." *Proceedings of the National Academy of Sciences of the United States of America*, **96** (21): 11713-11716 (1999).

REFERENCES

Blancafort, L. and A. A. Voityuk, "CASSCF/CAS-PT2 Study of Hole Transfer in Stacked DNA Nucleobases." *Journal of Physical Chemistry A*, **110** (20): 6426-6432 (2006).

Borrelli, K. W., A. Vitalis, et al., "PELE: Protein energy landscape exploration. A novel Monte Carlo based technique." *Journal of Chemical Theory and Computation*, **1** (6): 1304-1311 (2005).

Bowers, K. J., E. Chow, et al., *Scalable algorithms for molecular dynamics simulations on commodity clusters*. Proceedings of the 2006 ACM/IEEE conference on Supercomputing, Tampa, Florida, ACM (2006).

Braga, M. and S. Larsson, "Electronic factor for electron transfer through cyclohexane-type spacers." *Journal of Physical Chemistry*, **97** (35): 8929-8936 (1993).

Brazeau, B. J., B. J. Wallar, et al., "Effector proteins from P450(cam) and methane monooxygenase: lessons in tuning nature's powerful reagents." *Biochemical and Biophysical Research Communications*, **312** (1): 143-8 (2003).

Broo, A. and S. Larsson, "Electron transfer due to through-bond interactions. Study of aliphatic chains." *Chemical Physics*, **148** (1): 103-115 (1990).

Carter, E. A., G. Ciccotti, et al., "Constrained reaction coordinate dynamics for the simulation of rare events." *Chemical Physics Letters*, **156** (5): 472-477 (1989).

Castner, E. W., D. Kennedy, et al., "Solvent as Electron Donor: Donor/Acceptor Electronic Coupling Is a Dynamical Variable." *Journal of Physical Chemistry A*, **104** (13): 2869-2885 (2000).

Cave, R. J. and M. D. Newton, "Generalization of the Mulliken-Hush treatment for the calculation of electron transfer matrix elements." *Chemical Physics Letters*, **249** (1-2): 15-19 (1996).

Cave, R. J. and M. D. Newton, "Calculation of electronic coupling matrix elements for ground and excited state electron transfer reactions: Comparison of the generalized Mulliken-Hush and block diagonalization methods." *Journal of Chemical Physics*, **106** (22): 9213-9226 (1997).

Ceccarelli, M. and M. Marchi, "Simulation and Modeling of the Rhodobacter sphaeroides Bacterial Reaction Center II: Primary Charge Separation." *Journal of Physical Chemistry B*, **107** (23): 5630-5641 (2003).

Chance, B., D. Devault, et al., "Kinetics of electron transfer reactions in biological systems". *Fast Reaction and Primary Processes in Chemical Kinetics*. S. Claesson. New York, Interscience**:** 437-464 (1967).

Corin, A. F., R. A. Hake, et al., "Effects of surface amino acid replacements in cytochrome c peroxidase on intracomplex electron transfer from cytochrome c." *Biochemistry*, **32** (11): 2756-2762 (1993).

Coulson, C. A., B. O'Leary, et al., *Hückel theory for organic chemists*, Academic Press (1978).

Crystal, J., L. Y. Zhang, et al., "Computational Modeling for Scanning Tunneling Microscopy of Physisorbed Molecules Via Ab Initio Quantum Chemistry." *Journal of Physical Chemistry A*, **106** (9): 1802-1814 (2002).

Darve, E., M. A. Wilson, et al., "Calculating Free Energies Using a Scaled-Force Molecular Dynamics Algorithm." *Molecular Simulation*, **28** (1): 113-144 (2002).

Davidson, V. L., "Protein Control of True, Gated, and Coupled Electron Transfer Reactions." *Accounts of Chemical Research*, **41** (6): 730-738 (2008).

Davis, W. B., M. A. Ratner, et al., "Conformational Gating of Long Distance Electron Transfer through Wire-like Bridges in Donor−Bridge−Acceptor Molecules." *Journal of the American Chemical Society*, **123** (32): 7877-7886 (2001).

Davis, W. B., W. A. Svec, et al., "Molecular-wire behaviour in p - phenylenevinylene oligomers." *Nature*, **396** (6706): 60-63 (1998).

Dayan, F. E., A. C. Vincent, et al., "Amino- and Urea-Substituted Thiazoles Inhibit Photosynthetic Electron Transfer." *Journal of Agricultural and Food Chemistry*, **48** (8): 3689-3693 (2000).

de la Lande, A. and D. R. Salahub, "Derivation of interpretative models for long range electron transfer from constrained density functional theory." *Journal of Molecular Structure: THEOCHEM*, **943** (1-3): 115-120 (2010).

Dellago, C., P. G. Bolhuis, et al., "Transition Path Sampling". *Advances in Chemical Physics*. Hoboken, NJ, USA, John Wiley & Sons, Inc. **123:** 1-78 (2003).

Den Otter, W. K. and W. J. Briels, "Free energy from molecular dynamics with multiple constraints." *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics*, **98** (12): 773-781 (2000).

Dewar, M. J. S., Zoebisch E.G, et al., "Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model." *Journal of the American Chemical Society*, **107** (13): 3902 - 3909 (1985).

Dmochowski, I. J., B. R. Crane, et al., "Optical detection of cytochrome P450 by sensitizer-linked substrates." *Proceedings of the National Academy of Sciences of the United States of America*, **96** (23): 12987-12990 (1999).

Eng, M. P. and B. Albinsson, "The dependence of the electronic coupling on energy gap and bridge conformation - Towards prediction of the distance dependence of electron transfer reactions." *Chemical Physics*, **357** (1-3): 132-139 (2009).

Eng, M. P., J. Martensson, et al., "Temperature Dependence of Electronic Coupling through Oligo-*p*-phenyleneethynylene Bridges." *Chemistry - A European Journal*, **14** (9): 2819-2826 (2008).

Farazdel, A., M. Dupuis, et al., "Electric-field induced intramolecular electron transfer in spiro .pi.-electron systems and their suitability as molecular electronic devices. A theoretical study." *Journal of the American Chemical Society*, **112** (11): 4206-4214 (1990).

Feldman, A. K., M. L. Steigerwald, et al., "Molecular Electronic Devices Based on Single-Walled Carbon Nanotube Electrodes." *Accounts of Chemical Research*, **41** (12): 1731-1741 (2008).

Felts, A. K., W. T. Pollard, et al., "Multilevel Redfield Treatment of Bridge-Mediated Long-Range Electron Transfer: A Mechanism for Anomalous Distance Dependence." *Journal of Physical Chemistry*, **99** (9): 2929-2940 (1995).

REFERENCES

Fleurat-Lessard, P. and T. Ziegler, "Tracing the minimum-energy path on the free-energy surface." *Journal of Chemical Physics*, **123** (8): 084101-17 (2005).

Freddolino, P. L., A. S. Arkhipov, et al., "Molecular Dynamics Simulations of the Complete Satellite Tobacco Mosaic Virus." *Structure*, **14** (3): 437-449 (2006).

Freddolino, P. L., F. Liu, et al., "Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain." *Biophysical Journal*, **94** (10): L75-L77 (2008).

Friesner, R. A. and V. Guallar, "Ab Initio Quantum Chemical and Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis." *Annual Review of Physical Chemistry*, **56** (1): 389-427 (2005).

Gao, J. L. and D. G. Truhlar, "Quantum mechanical methods for enzyme kinetics." *Annual Review of Physical Chemistry*, **53**: 467-505 (2002).

Gehlen, J. N., I. Daizadeh, et al., "Tunneling matrix element in Ru-modified blue copper proteins: pruning the protein in search of electron transfer pathways." *Inorganica Chimica Acta*, **243** (1-2): 271-282 (1996).

Geren, L., S. Hahm, et al., "Photoinduced electron transfer between cytochrome c peroxidase and yeast cytochrome c labeled at Cys 102 with (4-bromomethyl-4'-methylbipyridine)-[bis(bipyridine)]-ruthenium2+." *Biochemistry*, **30** (39): 9450-9457 (1991).

Giese, B., S. Wessely, et al., "On the Mechanism of Long-Range Electron Transfer through DNA." *Angewandte Chemie International Edition*, **38** (7): 996-998 (1999).

Gray, H. B. and J. R. Winkler, "Electron transfer in proteins." *Annual Review of Biochemistry*, **65**: 537-561 (1996).

Guallar, V., C. Lu, et al., "Ligand Migration in the Truncated Hemoglobin-II from Mycobacterium tuberculosis: The Role of G8 Tryptophan." *Journal of Biological Chemistry*, **284** (5): 3106-3116 (2009).

Gunsalus, I. C., J. R. Meeks, et al. *Molecular Mechanisms of Oxygen Activation*. H. O. New York, Academic Press**:** 559-613 (1974).

Guo, M., B. Bhaskar, et al., "Crystal structure and characterization of a cytochrome c peroxidasecytochrome c sitespecific crosslink." *Proceedings of the National Academy of Sciences of the United States of America*, **101** (16): 5940-5945 (2004).

Hair, J. F., W. C. Black, et al., *Multivariate Data Analysis*, Prentice Hall (2009).

Haiss, W., H. v. Zalinge, et al., "Thermal gating of the single molecule conductance of alkanedithiols." *Faraday Discussions*, **131**: 253-264 (2006).

Han, J. and M. Kamber, *Data mining: concepts and techniques*. Burlington, USA, Morgan Kaufmann (2006).

Harada, K., K. Sakurai, et al., "Evaluation of the functional Role of the Heme-6-propionate Side Chain in Cytochrome P450cam." *Journal of the American Chemical Society*, **130**: 432-433 (2008).

Hartings, M. R., I. V. Kurnikov, et al., "Electron tunneling through sensitizer wires bound to proteins." *Coordination Chemistry Reviews*, **254** (3-4): 248-253 (2010).

Harvey, J. N., C. M. Bathelt, et al., "QM/MM modeling of compound I active species in cytochrome P450, Cytochrome c Peroxidase, and Ascorbate Peroxidase." *Journal of Computational Chemistry*, **27** (12): 1352-1362 (2006).

Hays Putnam, A.-M. A., Y.-T. Lee, et al., "Replacement of an Electron Transfer Pathway in Cytochrome c Peroxidase with a Surrogate Peptide." *Biochemistry*, **48** (1): 1-3 (2009).

Hazzard, J. T., G. McLendon, et al., "Effects of amino acid replacements in yeast iso-1 cytochrome c on heme accessibility and intracomplex electron transfer in complexes with cytochrome c peroxidase." *Biochemistry*, **27** (12): 4445-4451 (1988).

Hazzard, J. T., G. McLendon, et al., "Formation of electrostatically-stabilized complex at low ionic strength interprotein electron transfer between yeast cytochrome and cytochrome peroxidase." *Biochemical and Biophysical Research Communications*, **151** (1): 429-434 (1988).

Hazzard, J. T., S. J. Moench, et al., "Kinetics of intracomplex electron transfer and of reduction of the components of covalent and noncovalent complexes of cytochrome c and cytochrome c peroxidase by free flavin semiquinones." *Biochemistry*, **27** (6): 2002-2008 (1988).

Hazzard, J. T., T. L. Poulos, et al., "Kinetics of reduction by free flavin semiquinones of the components of the cytochrome c-cytochrome c peroxidase complex and intracomplex electron transfer." *Biochemistry*, **26** (10): 2836-2848 (1987).

Helms, A., D. Heiler, et al., "Electron transfer in bis-porphyrin donor-acceptor compounds with polyphenylene spacers shows a weak distance dependence." *Journal of the American Chemical Society*, **114** (15): 6227-6238 (1992).

Hess, B., C. Kutzner, et al., "GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation." *Journal of Chemical Theory and Computation*, **4** (3): 435-447 (2008).

Hoffman, B. M. and M. A. Ratner, "Gated electron transfer: when are observed rates controlled by conformational interconversion?", *Journal of the American Chemical Society*, **109** (21): 6237-6243 (1987).

Hoffmann, R., "An Extended Hückel Theory. I. Hydrocarbons." *Journal of Chemical Physics*, **39** (6): 1397-1397 (1963).

Hsu, C.-P., "The Electronic Couplings in Electron Transfer and Excitation Energy Transfer." *Accounts of Chemical Research*, **42** (4): 509-518 (2009).

Hsu, C.-P., Z.-Q. You, et al., "Characterization of the Short-Range Couplings in Excitation Energy Transfer." *Journal of Physical Chemistry C*, **112** (4): 1204-1212 (2008).

Huppman, P., T. Arlt, et al., "Kinetics, Energetics, and Electronic Coupling of the Primary Electron Transfer Reactions in Mutated Reaction Centers of Blastochloris viridis." *Biophysical Journal*, **82** (6): 3186-3197 (2002).

Iannuzzi, M., A. Laio, et al., "Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics." *Physical Review Letters*, **90** (23): 238302-4 (2003).

Improta, R., V. Barone, et al., "A Parameter-Free Quantum-Mechanical Approach for Calculating Electron-Transfer Rates for Large Systems in Solution." *ChemPhysChem*, **7** (6): 1211-1214 (2006).

REFERENCES

Isied, S. S., M. Y. Ogawa, et al., "Peptide-mediated intramolecular electron transfer: long-range distance dependence." *Chemical Reviews*, **92** (3): 381-394 (1992).

Jean, J. M. and B. P. Krueger, "Structural Fluctuations and Excitation Transfer between Adenine and 2-Aminopurine in Single-Stranded Deoxytrinucleotides." *Journal of Physical Chemistry B*, **110** (6): 2899-2909 (2006).

Jordan, K. D. and M. N. Paddon-Row, "Long-range interactions in a series of rigid nonconjugated dienes. 1. Distance dependence of the .pi.+,.pi.- and .pi.+*,.pi.-* splittings determined from ab initio calculations." *Journal of Physical Chemistry*, **96** (3): 1188-1196 (1992).

Jortner, J. and M. Bixon, "Electron Transfer: From Isolated Molecules to Biomolecules". *Advances in Chemical Physics*. New York, Wiley. **106-107** (1999).

Karyakin, A., D. Motiejunas, et al., "FTIR studies of the redox partner interaction in cytochrome P450: The Pdx-P450cam couple." *Biochimica Et Biophysica Acta (BBA) - General Subjects*, **1770** (3): 420-431 (2007).

Kawatsu, T., T. Kakitani, et al., "Destructive Interference in the Electron Tunneling through Protein Media." *Journal of Physical Chemistry B*, **106** (43): 11356-11366 (2002).

Kestner, N. R., J. Logan, et al., "Thermal electron transfer reactions in polar solvents." *Journal of Physical Chemistry*, **78** (21): 2148-2166 (1974).

Kumar, K., I. V. Kurnikov, et al., "Use of Modern Electron Transfer Theories To Determine Electronic Coupling Matrix Elements in Intramolecular Systems." *Journal of Physical Chemistry A*, **102** (28): 5529-5541 (1998).

Kurnikov, I. V., "HARLEM Molecular Modeling Package." *Department of Chemistry, University of Pittsburgh, Pittsburgh, PA* (1999).

Kuznetsov, A. M. and J. Ulstrup, *Electron Transfer in Chemistry and Biology*. Chichester, U.K., Wiley (1999).

Laio, A. and M. Parrinello, "Escaping free-energy minima." *Proceedings of the National Academy of Sciences of the United States of America*, **99** (20): 12562-12566 (2002).

Lambert, C., S. Amthor, et al., "From Valence Trapped to Valence Delocalized by Bridge State Modification in Bis(triarylamine) Radical Cations: Evaluation of Coupling Matrix Elements in a Three-Level System." *Journal of Physical Chemistry A*, **108** (31): 6474-6486 (2004).

Larsson, S., "Electron transfer in chemical and biological systems. Orbital rules for nonadiabatic transfer." *Journal of the American Chemical Society*, **103** (14): 4034-4040 (1981).

Leesch, V. W., J. Bujons, et al., "Cytochrome c Peroxidase−Cytochrome c Complex: Locating the Second Binding Domain on Cytochrome c Peroxidase with Site-Directed Mutagenesis." *Biochemistry*, **39** (33): 10132-10139 (2000).

Levich, V. G. and R. R. Dogonadze, "An Adiabatic Theory of Electron Processes in Solution." *Doklady Akademii Nauk SSSR*, **133** (1): 158-161 (1960).

Levine, I. N., *Quantum chemistry*. Upper Saddle River, USA, Prentice Hall (2000).

Lewis, D. F. V., *Guide to Cytochrome P450: Structure and Function*. London, Taylor & Francis Ltd (1996).

Liang, C. and M. D. Newton, "Ab initio studies of electron transfer: pathway analysis of effective transfer integrals." *Journal of Physical Chemistry*, **96** (7): 2855-2866 (1992).

Libby, W. F., "Theory of Electron Exchange Reactions in Aqueous Solution." *Journal of Physical Chemistry*, **56** (7): 863-868 (1952).

Liu, R.-Q., S. Hahm, et al., "Photooxidation of Trp-191 in cytochrome c peroxidase by ruthenium-cytochrome c derivatives." *Biochemistry*, **34** (3): 973-983 (1995).

Liu, W.-z., A.-j. Wang, et al., "Electrochemically Assisted Biohydrogen Production from Acetate." *Energy & Fuels*, **22** (1): 159-163 (2008).

Lovley, D. R., "Bug juice: harvesting electricity with microorganisms." *Nature Reviews Microbiology*, **4** (7): 497-508 (2006).

Lu, S.-Z., X.-Y. Li, et al., "Electronic coupling matrix elements of U-shaped donor-bridge-acceptor molecules and influence of mediated benzene solvent." *Chemical Physics Letters*, **414** (1-3): 71-75 (2005).

Malak, R. A., Z. Gao, et al., "Long-Range Electron Transfer Across Peptide Bridges: The Transition from Electron Superexchange to Hopping." *Journal of the American Chemical Society*, **126** (43): 13888-13889 (2004).

Marcus, R. A. and N. Sutin, "Electron transfers in chemistry and biology." *Biochimica Et Biophysica Acta*, **811**: 265-322 (1985).

Marcus, R. J., B. J. Zwolinski, et al., "The Electron Tunnelling Hypothesis for Electron Ex-change Reactions." *Journal of Physical Chemistry*, **58** (5): 432-437 (1954).

Matthis, A. L. and J. E. Erman, "Cytochrome c peroxidase-catalyzed oxidation of yeast iso-1 ferrocytochrome c by hydrogen peroxide. Ionic strength dependence of the steady-state parameters." *Biochemistry*, **34** (31): 9985-9990 (1995).

Miller, M. A., "A Complete Mechanism for Steady-State Oxidation of Yeast Cytochrome c by Yeast Cytochrome c Peroxidase." *Biochemistry*, **35** (49): 15791-15799 (1996).

Miller, M. A., L. Vitello, et al., "Regulation of Interprotein Electron Transfer By Trp 191 of Cytochrome c Peroxidase." *Biochemistry*, **34** (37): 12048-12058 (1995).

Miller, N. E., M. C. Wander, et al., "A Theoretical Study of the Electronic Coupling Element for Electron Transfer in Water." *Journal of Physical Chemistry A*, **103** (8): 1084-1093 (1999).

Mirkin, C. A. and M. A. Ratner, "Molecular Electronics." *Annual Review of Physical Chemistry*, **43** (1): 719-754 (1992).

Miyashita, O. and N. Go, "Reorganization Energy of Protein Electron Transfer Reaction: Study with Structural and Frequency Signature." *Journal of Physical Chemistry B*, **104** (31): 7516-7521 (2000).

Moraes, G. A. B. K., A. R. M. Chaves, et al., "Why is it better to produce coffee seedlings in full sunlight than in the shade? A morphophysiological approach." *Photosynthetica*, **48** (2): 199-207 (2010).

Moret, M.-E., E. Tapavicza, et al., "Quantum mechanical/molecular mechanical (QM/MM) car-parrinello simulations in excited states." *Chimia*, **59** (7-8): 493-498 (2005).

Moret, M.-E., I. Tavernelli, et al., "Electron Localization Dynamics in the Triplet Excited State of [Ru(bpy)3]2+ in Aqueous Solution." *Chemistry - A European Journal*, **16** (20): 5889-5894 (2010).

Morozova, O. B., P. J. Hore, et al., "Intramolecular Electron Transfer in Lysozyme Studied by Time-Resolved Chemically Induced Dynamic Nuclear Polarization." *Journal of Physical Chemistry B*, **109** (46): 21971-21978 (2005).

Mouesca, J.-M., J. L. Chen, et al., "Density Functional/Poisson-Boltzmann Calculations of Redox Potentials for Iron-Sulfur Clusters." *Journal of the American Chemical Society*, **116** (26): 11898-11914 (1994).

Mulholland, A., "Computational enzymology: modelling the mechanisms of biological catalysts." *Biochemical Society Transactions*, **36** (1): 22-26 (2008).

Musah, R. A. and D. B. Goodin, "Introduction of Novel Substrate Oxidation into Cytochrome c Peroxidase by Cavity Complementation: Oxidation of 2-Aminothiazole and Covalent Modification of the Enzyme." *Biochemistry*, **36** (39): 11665-11674 (1997).

Musah, R. A., G. M. Jensen, et al., "Variation in Strength of an Unconventional C−H to O Hydrogen Bond in an Engineered Protein Cavity." *Journal of the American Chemical Society*, **119** (38): 9083-9084 (1997).

Nakamura, K., T. Horiuchi, et al., "Significant contribution of arginine-112 and its positive charge of Pseudomonas putida cytochrome P-450cam in the electron transport from putidaredoxin." *Biochimica Et Biophysica Acta*, **1207** (1): 40-48 (1994).

Napper, A. M., I. Read, et al., "Electron Transfer Reactions of C-shaped Molecules in Alkylated Aromatic Solvents: Evidence that the Effective Electronic Coupling Magnitude Is Temperature-Dependent." *Journal of Physical Chemistry A*, **106** (18): 4784-4793 (2002).

Newton, M. D., "Quantum chemical probes of electron-transfer kinetics: the nature of donor-acceptor interactions." *Chemical Reviews*, **91** (5): 767-792 (1991).

Newton, M. D., "Modeling donor/acceptor interactions: Combined roles of theory and computation." *International Journal of Quantum Chemistry*, **77**: 255-263 (2000).

Newton, M. D. and R. J. Cave, "Molecular control of electron and hole transfer processes: electronic structure theory and applications.". *Molecular electronics*. M. A. Ratner and J. Jortner. Oxford, Blackwell Science (1997).

Nocek, J. M., V. W. Leesch, et al., "Multidomain binding of cytochrome c peroxidase by cytochrome c: Thermodynamics vs. microscopic binding constants." *Israel Journal of Chemistry*, **40**: 35-46 (2000).

Nocek, J. M., J. S. Zhou, et al., "Theory and Practice of Electron Transfer within Protein−Protein Complexes: Application to the Multidomain Binding of Cytochrome c by Cytochrome c Peroxidase." *Chemical Reviews*, **96** (7): 2459-2490 (1996).

Ogawa, M. Y., J. F. Wishart, et al., "Distance dependence of intramolecular electron transfer across oligoprolines in [(bpy)2RuIIL.bul.-(Pro)n-CoIII(NH3)5]3+, n = 1-6: different effects for helical and nonhelical polyproline II structure." *Journal of Physical Chemistry*, **97** (44): 11456-11463 (1993).

Ogliaro, F., N. Harris, et al., "A Model "Rebound" Mechanism of Hydroxylation by Cytochrome P450: Stepwise and Effectively Concerted Pathways, and Their Reactivity Patterns." *Journal of the American Chemical Society*, **122** (37): 8977-8989 (2000).

Okada, A., V. Chernyak, et al., "Solvent Reorganization in Long-Range Electron Transfer: Density Matrix Approach." *Journal of Physical Chemistry A*, **102** (8): 1241-1251 (1998).

Onuchic, J. N., D. N. Beratan, et al., "Some aspects of electron-transfer reaction dynamics." *Journal of Physical Chemistry*, **90** (16): 3707-3721 (1986).

Onuchic, J. N., D. N. Beratan, et al., "Pathway analysis of protein electron-transfer reactions." *Annual Review of Biophysics and Biomolecular Structure*, **21**: 349-377 (1992).

Ortiz de Montellano, P., *Cytochrome P450: Structure,Mechanism and Biochemistry*. New York, Plenum (1995).

Pappa, H. S., S. Tajbaksh, et al., "Probing the Cytochrome c Peroxidase−Cytochrome c Electron Transfer Reaction Using Site Specific Cross-Linking." *Biochemistry*, **35** (15): 4837-4845 (1996).

Pauling, L. and E. B. Wilson, *Introduction to quantum mechanics*. New York, USA, Courier Dover Publications (1985).

Paulson, B. P., L. A. Curtiss, et al., "Investigation of Through-Bond Coupling Dependence on Spacer Structure." *Journal of the American Chemical Society*, **118** (2): 378-387 (1996).

Pelletier, H. and J. Kraut, "Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c." *Science*, **258** (5089): 1748-1755 (1992).

Pensak, D. A., "Molecular modelling: scientific and technological boundaries." *Pure and Applied Chemistry*, **61** (3): 601-603 (1989).

Petrov, E. G. and V. May, "A Unified Description of Superexchange and Sequential Donor−Acceptor Electron Transfer Mediated by a Molecular Bridge." *Journal of Physical Chemistry A*, **105** (45): 10176-10186 (2001).

Petrov, E. G., Y. V. Shevchenko, et al., "On the length dependence of bridge-mediated electron transfer reactions." *Chemical Physics*, **288** (2-3): 269-279 (2003).

Phillips, J. C., R. Braun, et al., "Scalable molecular dynamics with NAMD." *Journal of Computational Chemistry*, **26** (16): 1781-1802 (2005).

Pochapsky, T. C., T. A. Lyons, et al., "A structure-based model for cytochrome P450cam-putidaredoxin interactions." *Biochimie*, **78** (8-9): 723-733 (1996).

Polo, F., S. Antonello, et al., "Evidence Against the Hopping Mechanism as an Important Electron Transfer Pathway for Conformationally Constrained Oligopeptides." *Journal of the American Chemical Society*, **127** (2): 492-493 (2005).

Poot, M., E. Osorio, et al., "Temperature Dependence of Three-Terminal Molecular Junctions with Sulfur End-Functionalized Tercyclohexylidenes." *Nano Letters*, **6** (5): 1031-1035 (2006).

REFERENCES

Poulos, T. L., S. T. Freer, et al., "The crystal structure of cytochrome c peroxidase." *Journal of Biological Chemistry*, **255**: 575-580 (1980).

Poulos, T. L., S. T. Freer, et al., "Crystallographic determination of the heme orientation and location of the cyanide binding site in yeast cytochrome c peroxidase." *Journal of Biological Chemistry*, **253**: 3730-3735 (1978).

Prytkova, T. R., I. V. Kurnikov, et al., "Ab Initio Based Calculations of Electron-Transfer Rates in Metalloproteins." *Journal of Physical Chemistry B*, **109** (4): 1618-1625 (2005).

Read, I., A. Napper, et al., "Solvent-Mediated Electronic Coupling: The Role of Solvent Placement." *Journal of the American Chemical Society*, **121** (47): 10976-10986 (1999).

Reece, S. Y., M. R. Seyedsayamdost, et al., "Direct Observation of a Transient Tyrosine Radical Competent for Initiating Turnover in a Photochemical Ribonucleotide Reductase." *Journal of the American Chemical Society*, **129** (45): 13828-13830 (2007).

Reisner, E., V. B. Arion, et al., "Electron-transfer activated metal-based anticancer drugs." *Inorganica Chimica Acta*, **361** (6): 1569-1583 (2008).

Reuter, N., A. Dejaegere, et al., "Frontier bonds in QM/MM methods: A comparison of different approaches." *Journal of Physical Chemistry A*, **104** (8): 1720-1735 (2000).

Ridley, J. and M. Zerner, "An intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and the azines." *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, **32** (2): 111-134 (1973).

Roitberg, A. E., M. J. Holden, et al., "Binding and Electron Transfer Between Putidaredoxin and Cytochrome P450cam. Theory and Experiments." *Journal of the American Chemical Society*, **120** (35): 8927-8932 (1998).

Rösch, N. and A. A. Voityuk, "Quantum Chemical Calculation of Donor–Acceptor Coupling for Charge Transfer in DNA". *Long-Range Charge Transfer in DNA II*: 37-72 (2004).

Rosenfeld, R. J., A.-M. A. Hays, et al., "Excision of a proposed electron transfer pathway in cytochrome c peroxidase and its replacement by a ligand-binding channel." *Protein Science*, **11** (5): 1251-1259 (2002).

Rosso, L., P. Mináry, et al., "On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles." *Journal of Chemical Physics*, **116** (11): 4389-4389 (2002).

Rust, M., J. Lappe, et al., "Multistate Effects in Calculations of the Electronic Coupling Element for Electron Transfer Using the Generalized Mulliken−Hush Method." *Journal of Physical Chemistry A*, **106** (15): 3930-3940 (2002).

Sachs, S. B., S. P. Dudek, et al., "Rates of Interfacial Electron Transfer through π-Conjugated Spacers." *Journal of the American Chemical Society*, **119** (43): 10563-10564 (1997).

Schlichting, I., J. Berendzen, et al., "The Catalytic Pathway of Cytochrome P450cam at Atomic Resolution." *Science*, **287** (5458): 1615-1622 (2000).

Schöneboom, J. C., H. Lin, et al., "The Elusive Oxidant Species of Cytochrome P450 Enzymes: Characterization by Combined Quantum Mechanical/Molecular Mechanical (QM/MM) Calculations." *Journal of the American Chemical Society*, **124** (27): 8142-8151 (2002).

Seifert, J. L., T. D. Pfister, et al., "Hopping in the Electron-Transfer Photocycle of the 1:1 Complex of Zn−Cytochrome c Peroxidase with Cytochrome c." *Journal of the American Chemical Society*, **127** (16): 5750-5751 (2005).

Selzer, Y., M. A. Cabassi, et al., "Thermally Activated Conduction in Molecular Junctions." *Journal of the American Chemical Society*, **126** (13): 4052-4053 (2004).

Shao, F., M. A. O'Neill, et al., "Long-range oxidative damage to cytosines in duplex DNA." *Proceedings of the National Academy of Sciences of the United States of America*, **101** (52): 17914-17919 (2004).

Sharp, K. H., M. Mewies, et al., "Crystal structure of the ascorbate peroxidase-ascorbate complex." *Nature Structural Biology*, **10** (4): 303-307 (2003).

Shaw, D. E., R. O. Dror, et al., *Millisecond-scale molecular dynamics simulations on Anton*. Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, Portland, Oregon, ACM (2009).

Shaw, D. E., P. Maragakis, et al., "Atomic-Level Characterization of the Structural Dynamics of Proteins." *Science*, **330** (6002): 341-346 (2010).

Sherwood, P., A. H. de Vries, et al., "QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis." *Journal of Molecular Structure: THEOCHEM*, **632**: 1-28 (2003).

Shih, C., A. K. Museth, et al., "Tryptophan-Accelerated Electron Flow Through Proteins." *Science*, **320** (5884): 1760-1762 (2008).

Shimada, H., S. Nagano, et al., "Putidaredoxin-cytochrome P450cam interaction." *Journal of Inorganic Biochemistry*, **83** (4): 255-260 (2001).

Shin, Y.-g. K., M. D. Newton, et al., "Distance Dependence of Electron Transfer Across Peptides with Different Secondary Structures: The Role of Peptide Energetics and Electronic Coupling." *Journal of the American Chemical Society*, **125** (13): 3722-3732 (2003).

Siddarth, P., "An artificial intelligence search for key residues in protein electron transfer systems." *Journal of photochemistry and photobiology.A, Chemistry*, **82** (13): 117-121 (1994).

Siddarth, P. and R. A. Marcus, "Comparison of experimental and theoretical electronic matrix elements for long-range electron transfer." *Journal of Physical Chemistry*, **94** (7): 2985-2989 (1990).

Siddarth, P. and R. A. Marcus, "Electron-transfer reactions in proteins: an artificial intelligence approach to electronic coupling." *Journal of Physical Chemistry*, **97** (10): 2400-2405 (1993).

Sivaraja, M., D. B. Goodin, et al., "Identification by ENDOR of Trp191 as the free-radical site in cytochrome c peroxidase compound ES." *Science*, **245** (4919): 738-740 (1989).

REFERENCES

Skourtis, S. S., G. Archontis, et al., "Electron transfer through fluctuating bridges: On the validity of the superexchange mechanism and time-dependent tunneling matrix elements." *Journal of Chemical Physics*, **115** (20): 9444-9462 (2001).

Skourtis, S. S., I. A. Balabin, et al., "Protein dynamics and electron transfer: Electronic decoherence and non-Condon effects." *Proceedings of the National Academy of Sciences of the United States of America*, **102** (10): 3552-3557 (2005).

Skourtis, S. S., D. N. Beratan, et al., "The two-state reduction for electron and hole transfer in bridge-mediated electron-transfer reactions." *Chemical Physics*, **176** (2-3): 501-520 (1993).

Skourtis, S. S. and J. N. Onuchic, "Effective two-state systems in bridge-mediated electron transfer. A Green function analysis." *Chemical Physics Letters*, **209** (1-2): 171-177 (1993).

Stewart, J. J. P., "Optimization of parameters for semiempirical methods I. Method." *Journal of Computational Chemistry*, **10** (2): 209-220 (1989).

Stuchebrukhov, A. A. and R. A. Marcus, "Theoretical Study of Electron Transfer in Ferrocytochromes." *Journal of Physical Chemistry*, **99** (19): 7581-7590 (1995).

Subotnik, J. E., R. J. Cave, et al., "The initial and final states of electron and energy transfer processes: Diabatization as motivated by system-solvent interactions." *Journal of Chemical Physics*, **130** (23): 234102-234102 (2009).

Summers, F. E. and J. E. Erman, "Reduction of cytochrome c peroxidase compounds I and II by ferrocytochrome c. A stopped-flow kinetic investigation." *Journal of Biological Chemistry*, **263** (28): 14267-14275 (1988).

Svensson, M., S. Humbel, et al., "ONIOM: A multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(t-Bu)(3))(2)+H-2 oxidative addition." *Journal of Physical Chemistry*, **100** (50): 19357-19363 (1996).

Szabó, A. and N. S. Ostlund, *Modern quantum chemistry*. New York, USA, Courier Dover Publications (1996).

Takada, T., K. Kawai, et al., "Direct observation of hole transfer through double-helical DNA over 100 Å." *Proceedings of the National Academy of Sciences of the United States of America*, **101** (39): 14002-14006 (2004).

Tavernelli, I., E. Tapavicza, et al., "Nonadiabatic coupling vectors within linear response time-dependent density functional theory." *Journal of Chemical Physics*, **130** (12): 124107-124110 (2009).

Theodoridis, S. and K. Koutroumbas, *Pattern Recognition, Third Edition*. Orlando, FL, USA, Academic Press, Inc. (2006).

Torrie, G. M. and J. P. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling." *Journal of Computational Physics*, **23** (2): 187-199 (1977).

Troisi, A., A. Nitzan, et al., "A rate constant expression for charge transfer through fluctuating bridges." *Journal of Chemical Physics*, **119** (12): 5782-5782 (2003).

Troisi, A., M. A. Ratner, et al., "Dynamic Nature of the Intramolecular Electronic Coupling Mediated by a Solvent Molecule: A Computational Study." *Journal of the American Chemical Society*, **126** (7): 2215-2224 (2004).

Ungar, L. W., M. D. Newton, et al., "Classical and Quantum Simulation of Electron Transfer Through a Polypeptide." *Journal of Physical Chemistry B*, **103** (34): 7367-7382 (1999).

Unno, M., H. Shimada, et al., "Role of Arg112 of cytochrome p450cam in the electron transfer from reduced putidaredoxin. Analyses with site-directed mutants." *Journal of Biological Chemistry*, **271** (30): 17869-74 (1996).

van Erp, T. S., D. Moroni, et al., "A novel path sampling method for the calculation of rate constants." *Journal of Chemical Physics*, **118** (17): 7762-7762 (2003).

Voityuk, A. A., "Electronic Couplings in DNA π-Stacks: Multistate Effects." *Journal of Physical Chemistry B*, **109** (38): 17917-17921 (2005).

Voityuk, A. A., "Donor-acceptor electronic couplings in [pi]-stacks: How many states must be accounted for?", *Chemical Physics Letters*, **422** (1-3): 15-19 (2006).

Voityuk, A. A., "Charge-on-site scheme to estimate the electronic coupling in electron transfer systems." *Chemical Physics Letters*, **451** (1-3): 153-157 (2008).

Voityuk, A. A., "Electron transfer between [4Fe-4S] clusters." *Chemical Physics Letters*, **495** (1-3): 131-134 (2010).

Voityuk, A. A. and N. Rösch, "Fragment charge difference method for estimating donor-acceptor electronic coupling: Application to DNA pi-stacks." *Journal of Chemical Physics*, **117** (12): 5607-5616 (2002).

Wade, R. C., D. Motiejunas, et al., "Multiple molecular recognition mechanisms. Cytochrome P450 - a case study." *Biochimica Et Biophysica Acta*, **1754** (1-2): 239-244 (2005).

Wang, X. and W. M. Nau, "Kinetics of One- and Two-Directional Charge Hopping in One-Dimensional Systems: Application to DNA." *ChemPhysChem*, **2** (12): 761-766 (2001).

Wang, Y. and E. Margoliash, "Enzymic Activities of Covalent 1:1 Complexes of Cytochrome c and Cytochrome c Peroxidase." *Biochemistry*, **34** (6): 1948-1958 (1995).

Warshel, A. and A. Bromberg, "Oxidation of 4a,4b-Dihydrophenanthrenes. III. A Theoretical Study of the Large Kinetic Isotope Effect of Deuterium in the Initiation Step of the Thermal Reaction with Oxygen." *Journal of Chemical Physics*, **52** (3): 1262-1269 (1970).

Warshel, A. and M. A. Levitt, "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme." *Journal of Molecular Biology*, **103**: 227-249 (1976).

Whalley, A. C., M. L. Steigerwald, et al., "Reversible Switching in Molecular Electronic Devices." *Journal of the American Chemical Society*, **129** (42): 12590-12591 (2007).

Wold, S., M. Sjöström, et al., "PLS-regression: a basic tool of chemometrics." *Chemometrics and Intelligent Laboratory Systems*, **58** (2): 109-130 (2001).

Wolfgang, J., S. M. Risser, et al., "Secondary Structure Conformations and Long Range Electronic Interactions in Oligopeptides." *Journal of Physical Chemistry B*, **101** (15): 2986-2991 (1997).

REFERENCES

Xie, Q., G. Archontis, et al., "Protein electron transfer: a numerical study of tunneling through fluctuating bridges." *Chemical Physics Letters*, **312** (2-4): 237-246 (1999).

Yonetani, T. and T. Ohnishi, "CcP, a mitochondrial enzyme of yeast." *Journal of Biological Chemistry*, **241**: 2983–2984 (1996).

Zhang, L. Y., R. A. Friesner, et al., "Ab Initio Quantum Chemical Calculation of Electron Transfer Matrix Elements for Large Molecules." *Journal of Chemical Physics*, **107** (2): 450-459 (1997).

Zheng, J., Y. K. Kang, et al., "Generalized Mulliken−Hush Analysis of Electronic Coupling Interactions in Compressed π-Stacked Porphyrin−Bridge−Quinone Systems." *Journal of the American Chemical Society*, **127** (32): 11303-11310 (2005).

Zimmt, M. B. and D. H. Waldeck, "Exposing Solvent's Roles in Electron Transfer Reactions: Tunneling Pathway and Solvation." *Journal of Physical Chemistry A*, **107** (19): 3580-3597 (2003).