Protein-Protein Docking and Brownian Dynamics Simulation of Electron Transfer Proteins

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

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

von
Dagmar Flöck
aus Neuwied (Rhein)

Frankfurt 2003

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

Dekan: Prof. Dr. Schmidt-Böcking

Gutachter: Prof. Dr. Werner Mäntele
Prof. Dr. Volkhard Helms

Datum der Disputation: 20. Oktober 2003

Zusammenfassung

Protein-Protein Wechselwirkungen sind von zentraler Bedeutung für alle lebenden Organismen. Auch in unserer post-genomischen Zeit sind noch viele Fragen auf diesem Gebiet offen. Die vielen Beiträge aus den verschiedenen Naturwissenschaften sind Beweis für die gemeinsame Anstrengung, diese Prozesse besser zu verstehen.

Im Rahmen der hier vorliegenden Arbeit wurden Modellrechnungen an Elektrontransferproteinen der Atmungskette und der bakteriellen Photosynthese durchgeführt. Diese Systeme sind besonders gut für unsere Untersuchungen geeignet. Zum einen liegen für die beteiligten Proteine Röntgenkristallstrukturen mit hoher Auflösung vor. Dadurch ist es möglich, Computersimulationen im atomaren Detail auszuführen. Experimentelle Untersuchen haben außerdem zahlreiche Erkenntnisse über die kinetischen und die Mutationseigenschaften dieser Enzyme geliefert. Die Berücksichtigung dieser Ergebnisse erlaubt es, biologisch relevante Aussagen zu treffen, und die Resultate der Computersimulationen direkt mit dem Experiment zu vergleichen.

Die hier untersuchten Elektrontransferpartner der Atmungskette sind Cytochrom c und Cytochrom c Oxidase (COX). COX ist als terminales Enzym der Atmungskette ein integrales Membranprotein der inneren mitrochondrischen Membran. Die Oxidase nimmt Elektronen auf, um einerseits Sauerstoff zu Wasser zu reduzieren, und ausserdem um einen Protonengradient über die Membran aufzubauen. Die Elektronen werden von dem globularen, löslichen Elektronentransporter Cytochrom c in Wechselwirkung mit der Oxidase abgegeben.

In ähnlicher Weise trägt Cytochrom c auch zur Funktionsweise der bakteriellen Photosynthese bei. Hier werden die Elektronen an das Reaktionszentrum (wiederum ein integrales Membranprotein) übertragen. Auch hier produziert der Elektronentransfer über die Membran einen Protonengradient, der dann seinerseits die ATP-Synthase aus ADP durch die ATPase antreibt.

Im ersten Teil dieser Arbeit sind Protein-Protein Docking-Studien dokumentiert. Bis heute konnten die meisten Protein-Komplex-Strukturen nicht experimentell aufgeklärt werden, so auch die beiden oben genannten Elektrontransfer-Komplexe.

Numerische Docking-Methoden haben das Ziel zwei Proteine so aneinander anzupassen, daß die natürliche Reaktion, in diesem Fall der Elektrontransfer, stattfinden kann. Der simulierte Docking-Prozeß kann in drei Phasen unterteilt werden: Das eigentliche Docking, die Bewertung aller sich ergebenden Protein-Komplexe und schließlich die Verfeinerung der besten Komplexe. Der erste Schritt erfolgt üblicherweise mit Hilfe einer impliziten Oberflächenrepräsentation der Proteine auf einem Gitter. Dann erfolgt die numerische Suche im kompletten Translations- und Rotationsraum nach der besten Oberflächenkomplementarität.

In der vorliegenden Docking-Studie wurde eine aktuelle Docking-Software mit einer Energieminimierung kombiniert, die zugleich die sinnvolle Bewertung der erhaltenen Komplexe erlaubt. Das Durchsuchen des Konformationsraumes erfolgte mit einer schnellen Fourier-Transformationsroutine. Elektrostatisch ungünstige Komplexe wurden sofort eliminiert. Von den verbleibenden tausenden Komplexen wurden mit Hilfe von biochemischen Informationen die besten herausgefiltert. Diese verbleibenden möglichen Komplexe wurden nun unter Anwendung eines atomistischen Kraftfeldes energieminimiert, wobei jedes Protein für sich starr gehalten wurde. Die Addition der elektrostatischen und van der Waals Energiebeiträge ergab ein gutes Maß für die Güte dieser Komplexe.

Die dargelegte Methode wurde an einem Elektrontransfer-Protein-Komplex mit vorhandener Komplexstruktur getestet: dem Cytochrom c Peroxidase:Iso-1-cytochrom c Paar. Das Resultat ist sehr vielversprechend, da die durch diese Methode identifizierte Komplexstruktur sich nur minimal von der Röntgenkristallstruktur unterscheidet: Der mittlere Abstand der C_α-Atome (der Verbindungsglieder zwischen Rumpfstruktur und den Seitenketten der Aminosäuren) beträgt nur etwa 2 Å.

Nach diesem erfolgreichen Test wurden verschiedene Cytochrom c Oxidase:Cytochrom c Paare mit der gleichen Methode gedockt: COX aus *Paracoccus denitrificans* mit Pferdeherz Cytochrom c und COX mit dem löslichen Fragment des membrangebundenen Cytochrom c_{552} (beide aus *P. denitrificans*). Bei der letzten Docking-Studie wurden sowohl die funktionale Zweiuntereinheiten-Struktur als auch die komplette Vieruntereinheiten-Struktur der Oxidase aus *P. denitrificans* verwendet.

Das Ergebnis der ersten Studie ist einer anderen vorhergesagten Struktur eines Oxidase:Pferdeherz Cyt c Systems sehr ähnlich. Im Falle der Vorhersagen, die die physiologischen Systeme betreffen, erwarten wir Ergebnisse aus Fluoreszenz-Labeling-Experimenten. Sie könnten darüber Auskunft geben, ob diese Komplexstruktur-Vorhersagen ebenso verlässlich sind. Die Ergebnisse des COX:Cytochrom c Dockings sind den im zweiten Teil beschriebenen Brownschen Dynamik Simulationen zu Grunde gelegt.

Anhand einer weiteren Docking-Studie mit dem Reaktionzentrum aus *Rhodobacter Spaeroides* und dessen physiologischen Elektrontransferpartner cytochrom c_2 wurde die kombinierte Docking-Methode noch einmal kritisch untersucht und um einige Elemente erweitert.

Im zweiten Teil dieser Arbeit wurde die diffusive Annäherung des Cytochrom c an die Cytochrom c Oxidase mit der Brownschen Dynamik Methode simuliert. Die Diffusionsbewegung eines Brownschen Teilchens in wässriger Lösung wird durch die Langevin-Gleichung bestimmt. Der auf dieser Gleichung fußende Ermack-McCammon-Algorithmus ist Grundlage der Simulationsmethode.

Von den zu berücksichtigenden intermolekularen Kräften kommt der Elektrostatik die größte Bedeutung zu. Das elektrostatische Potential eines Makromoleküls, umgeben von wässriger ionischer Lösung, wird durch die Poisson-Boltzmann Gleichung bestimmt. Die numerische Bestimmung des Potentials erfolgt üblicherweise mit einer finiten Differenzmethode auf einem Gitter. Da die Gitterpunktabstände klein gehalten werden müssen, sind die Rechnungen zu aufwendig, um sie nach jedem Schritt der Brownschen Dynamik Simulation erneut durchzuführen. Deshalb wird hier die sogenannte Effektive-Ladungs-Methode verwendet. Das Potential der Makromoleküle im inhomogenen dielektrischen Medium wird dabei durch effektive Ladungen, die man innerhalb der Proteine plaziert, über einen Bereich eines homogenen Mediums reproduziert. Weiterhin werden während der Simulation van der Waal Abstoßung und Desolvationseffekte berücksichtigt.

Um entscheiden zu können, ob eine Brownsche Dynamik Trajektorie zu einer erfolgreichen Annäherung der beiden Proteine führt, müssen Kriterien für das Erreichen des sogenannten Encounter-Komplexes definiert werden. Ausgehend von den gedockten Komplexen wurden deshalb während der Simulation, die Abstände zwischen aus-

gewählten intermolekularen Residuenpaaren verfolgt. Wenn die Kontaktdistanz für eine bestimmte Anzahl dieser Paare auf einen bestimmten Maximalwert festgelegt ist, kann über den Anteil der erfolgreichen Trajektorien, die bimolekulare Assoziationsrate berechnet werden.

Die so ermittelten Raten für COX und Pferdeherz, sowie für COX und Cytochrom c_{552}, konnten nun mit experimentell gewonnenen Raten verglichen werden. Da die Elektrostatik für den Annäherungsprozeß dieser Proteine eine so gewichtige Rolle spielt, wirken sich Mutationen, die mit einer Ladungsänderung einhergehen, merklich aus. Dies ist vor allem dann der Fall, wenn sich die Mutation in der Nähe der Bindungsstelle befindet. Aus dem gleichen Grund ist die Assoziationsrate auch stark von der Ionenstärke der umgebenden Lösung abhängig. Steigt die Ionenkonzentration, wird die elektrostatische Komplementarität der Bindingsstellen der beiden Makromoleküle stärker abgeschirmt, und die Rate sinkt.

Diese beiden relativen Trends konnten durch die Simulationen gut reproduziert und bestätigt werden. Allerdings liegen die absoluten Resultate merklich über den experimentell gemessenen Raten. Dies kann verschiedene Gründe haben. Zum einen ist es schwierig den Encounter-Komplex exakt zu definieren, dies umso mehr da hier von numerisch gedockten Proteinkomplexen ausgegangen werden mußte. Zum anderen werden die Ratenmessungen an gelöster Oxidase durchgeführt. Bei der Lösung der Oxidase mit Detergenz, heften sich Detergenzmoleküle an den Teil der Oxidase, der vorher durch die Membran eingeschlossen war. In Lösung zeigen die hydrophilen Detergenzköpfe nach außen. Dieser Umstand, der in den Computersimulationen nicht berücksichtigt ist, könnte Einfluß auf die Elektrostatik und somit auch auf die Assoziationsraten haben. Desweiteren ist es aber auch sehr gut möglich, daß post-diffusive Effekte, die nicht in einer Brownschen Dynamik Simulation von starren Körpern berücksichtigt werden können, die Raten erniedrigen.

Um den Einfluß der Membranumgebung auf die Wechselwirkung des Elektrontransportsystems zu untersuchen, wurde eine DPPC Doppelschicht um die Oxidase modelliert und energieminimiert. Mit Poisson-Boltzmann Rechnungen wurde das elektrostatische Potential dieses Nanosystems untersucht und mit dem der einzelnen Oxidase verglichen. Durch einen modifizierten Set-up konnten dann auch für dieses Membransystem Brownsche Dynamik Simulationen durchgeführt werden. Der Vergleich mit den vorhergehenden Simulationen ohne Membran erbrachte bemerkenswerte Ergebnisse.

Während die Assoziationsraten für Pferdeherz Cytochrom c durch den Membraneinfluß erniedrigt wurden, stiegen sie im Fall des physiologischen Transferpartners c_{552}. Pferdeherz Cytochrom c weist eine positive Nettoladung und einen ausgeprägten bipolaren Charakter auf. Eine große Zahl positiv geladener Seitenketten befindet sich auf der gleichen Hemisphäre mit der Bindungsstelle. Obwohl die DPPC Lipidmoleküle neutral sind, zeigten die Elektrostatikrechnungen, daß die Membranoberfläche abstoßend auf positive Ladungen wirkt. Da sich nun die Bindungsstelle der Oxidase für Cytochrom c nur etwa 10 Å oberhalb der Membran befindet, verringert sich die Wahrscheinlichkeit der Assoziation.

Im Gegensatz dazu besitzt Cytochrom c_{552} eine negative Nettoladung und einen weniger starken bipolaren Charakter. Somit wirkt die Membran nicht inhibierend, sondern steigert die Raten sogar. Die biologische Membran in *Paracoccus denitrificans* wird demnach ähnliche elektrostatische Eigenschaften wie die hier modellierte Membran aufweisen, denn es ist anzunehmen, daß das natürliche System hoch effizient ist.

Abschließend ist zu sagen, daß die hier vorgestellten Methoden gut geeignet sind, um weitere makromolekulare Systeme zu untersuchen. Es gibt zum Beispiel experimentelle Hinweise darauf, daß die Membranproteine der Atmungskette und der Photosynthese Superkomplexe bilden. Hier wäre es nötig die Superkomplexstruktur numerisch zu docken. Nachfolgend könnte mit einer modifizierten Brownschen Dynamik Simulation die diffusive Annäherung vieler Cytochrome und deren Wechselwirkung untereinander untersucht werden.

Contents

Chapter 1

Introduction

Research in the field of proteomics has increasingly drawn attention in the post-genomic era. Contributions from different areas of natural, medical, computer science and mathematics show the common effort to get deeper insights into the field.

The study of biological systems can be as fascinating for physicists as for biologists. There are many examples of mutual stimulations in the history of sciences, one of the most famous probably being the observation of irregular movements of small particles in solvent by R. Brown, which inspired A. Einstein to develop the theory of Brownian molecular movement.

Tackling the complexity of living matter is a big challenge in modern biophysics. Even though e. g. all proteins are built out of the same building blocks they show a huge diversity and individuality in sequence, folding, and function. In studying protein-protein interaction we have to deal with a multitude of forces which govern the interplay and the flexible structure of each of the units. The development and application of physical models must be strongly correlated to the problem under study, since theoretical models are often only valid for certain time- and length-scales. The chosen model should also be as simple as possible and as accurate as necessary to achieve meaningful results.

Crystal structures of single proteins were determined as early as in the 1950s, the first one being myoglobin [1]. Crystallization of protein-protein complexes being much more demanding, the first crystal structure of protein-protein complex was determined much later. Nowadays, about 20.000 structures of individual proteins are stored in protein data base (PDB), but only about 500 of protein-protein complexes. On the

other hand, proteomics techniques have recently underlined the enormous importance of transient protein complexes as functional units in biological cells [2]. Therefore, understanding the association of multiple proteins is an enormously important problem in current life science.

In the case of computational protein-protein docking a variety of methods have been explored in the last decade. Still, a general methodology for the docking of two protein structures has not been found. The difficulty may arise from the fact that the free energy minimum of the native protein complex does not necessarily lie at the bottom of a deep energy well. In the case of electron transfer systems, as they are under study here, the complexes form only transiently. The electron transfer itself occurs at high speed once the donor:acceptor pair has assembled. It has to be assumed that this single charge transfer results in the untying of the freely moving docking partner, or in other words that the dissociation rate then will be higher than the association rate. Without doubt the full dynamical process of the system behavior can not be analyzed in quantum theoretical detail. Still, models for protein docking and protein association can be successfully explored individually.

Here, we combine standard docking methods based on surface complementarity with rigid body energy minimizations applying an atomistic force field. The association process itself is then simulated by Brownian dynamics simulations. Of special interest is the incorporation of the membrane environment.

1.1 Structure and Function of Proteins

Proteins are functional forms of polypeptides. Their structure is defined as much by the environment as by the chemical properties of the polypeptide chain. Distinctive sorts of proteins are e.g. water-soluble globular proteins, water-insoluble fibrous proteins and proteins that associate with the hydrophobic environment of membrane bilayers. Each of these is a polymer built from amino acids, characterized by a distinct composition. Common to all living organisms are 20 amino acids, where the amino and carboxylic acid groups are separated by a single C_α carbon (see Figure 1.1). In a protein the amino acids are linked by peptide bonds. Each amino acid is distinguished by the chemistry of the side chains that are attached to the C_α carbon. Still, to function proteins often require additional organic or inorganic co-factors, like the heme group of cytochromes.
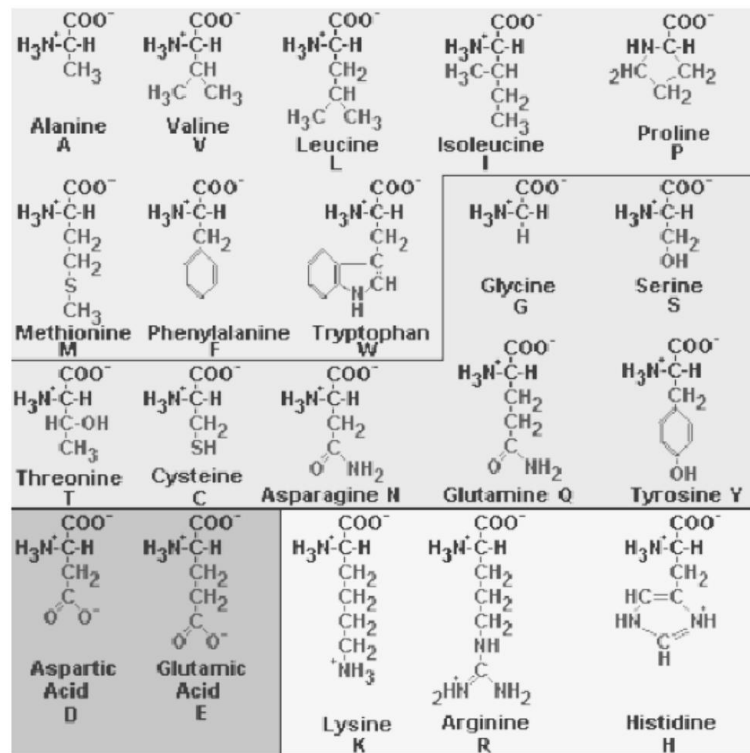
Figure 1.1: *Twenty common amino acids. The side chains can be categorized as being nonpolar (top), polar (middle), or charged at neutral pHs (bottom).*

Amino acids with polar side chains are regarded to be hydrophilic, because they prefer to be hydrated, whereas nonpolar amino acids are regarded to be hydrophobic. They tend to be part of the interior of globular proteins or reside in the membrane-spanning part of trans-membrane proteins. The resulting hydrophobic effect is a strong driving force for the folding of proteins. Hydrophobic amino acids either have alkyl or aromatic side chains. The side chains of polar amino acids may be charged under physiological pHs. The overall charge of a protein depends on the number of basic and acidic residues that are charged at a particular pH.

The regular structure of a polypeptide is called the secondary structure. Most common are the so called α-helix and the β-sheet. While the α-helix is stabilized through hydrogen bonds within the helical structure, the β-sheet has hydrogen-bonding interactions between the extended helical strands holding the sheet together. The overall three-dimensional conformation of a polypeptide chain is called the tertiary structure, which can be visualized best by a molecular graphics program, where secondary struc-

ture elements are often represented by symbols or the peptide backbone is traced by a ribbon to simplify the inspection. Finally, the non-covalent association of polypeptides to form a multimeric complex defines the quaternary structure of a protein. For instance a complex formed by two identical subunits is called a homodimer, while a complex formed by two different subunits is called a heterodimer.

1.2   Electron Transfer Systems in Photosynthesis and Respiration

A variety of important biochemical reactions occur via electron-transfer (ET) processes, among these e.g. DNA damage, photosynthesis, and much of enzymatic catalysis. Most of these processes are bimolecular and have been extensively investigated by kinetic measurements [3, 4], site directed mutagenesis [5–7], linear dichroism [8, 9], and computational docking [10, 11]. Two of these ET systems have a central role in cellular metabolism: the electron transfer from cytochrome c to cytochrome c oxidase, and also the electron transfer between cytochrome c and the photosynthetic reaction center.
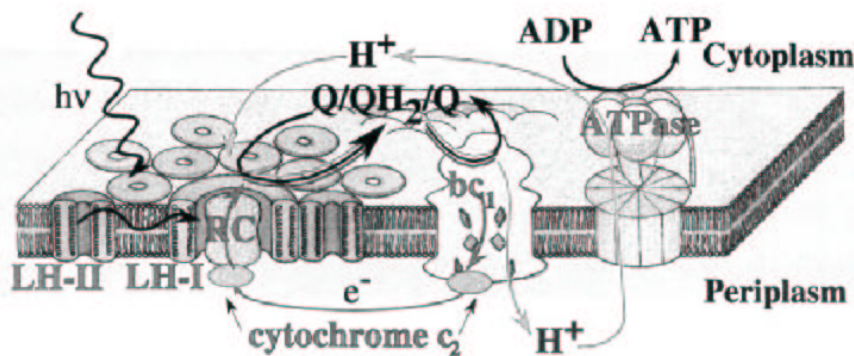
Figure 1.2: *Schematic representation of the photosynthetic apparatus in bacteria [12]. The reaction center is surrounded by multiple light-harvesting complexes (LHs). Photons are absorbed by LHs and excitation is transferred to the RC initiating a charge separation. Electrons are transfered by the cytochrome c_2 from the ubiquinone-cytochrome bc_1 complex to the RC. The electron transfer across the membrane produces a large proton gradient that drives the synthesis of ATP from ADP by the ATPase.*

The integral membrane protein photosynthetic reaction center (RC) with its electron transfer partner, cytochrome c plays a key role in the photosynthetic apparatus,

converting light energy to chemical energy (ATP synthesis). A schematic drawing of the system is shown in Figure 1.2. The reaction center was the first reaction partner where c-type cytochrome oxidation could be examined with single-turnover kinetics [13, 14]. Electron transfer occurs from the reduced ferrous heme of the cytochrome c to the light-generated, free radical of the oxidized bacteriochlorophyll dimer, P^+, in the reaction center. The ET within the bound complex has been measured to occur at a rate of $10^6 s^{-1}$ [15, 16].
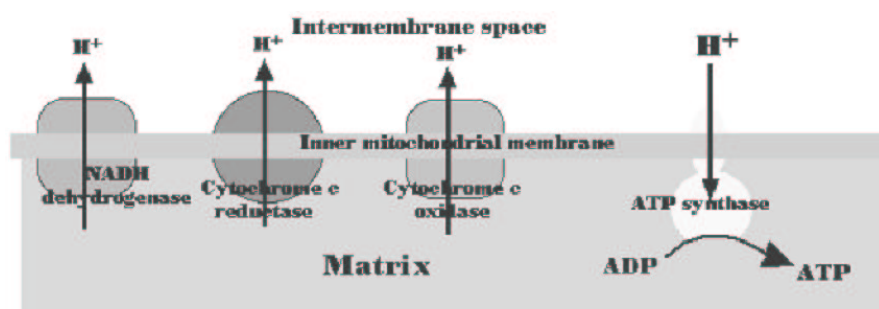
Figure 1.3: *Scheme of respiratory chain in mitochondria. Electrons are shuttled from cytochrome c reductase to cytochrome c oxidase (COX) on the matrix side. Once four electrons are transfered oxygen is reduced to water by COX and proton gradient is established over the inner-mitochondrial membrane, which is the driving force for ATP synthesis.*

Cytochrome c oxidase is the terminal enzyme of the respiratory chain (see Figure 1.3) in the inner membrane of mitochondria and the cytoplasmic membrane of many bacteria. Energy derived from the redox reaction is conserved into a proton motive force, which can be used by the cell to drive critical reactions such as synthesizing ATP from ADP.

Protein-protein ET is influenced by many structural and dynamic factors. Here we are not concerned with the details of the ET itself, but focus on the events leading to the formation of productive ET complexes, through rotational and translational diffusion, which is influenced by electrostatic and steric forces.

1.3   Protein-Protein Association Processes

The association process of a pair of proteins ($A + B \rightleftharpoons^{k_{on}}_{k_{off}} AB$) can be described as a two-step reaction:

$$A + B \rightleftharpoons^{k_1}_{k_{-1}} A:B \rightleftharpoons^{k_2}_{k_{-2}} AB, \tag{1.1}$$
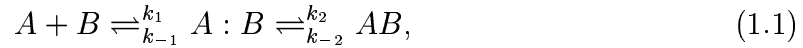
where A and B represent the freely diffusing particles, $A:B$ the encounter complex, and AB the final complex. According to this model the proteins diffuse until they reach a so called steering region where the electrostatic attraction leads them to build the encounter complex. The larger the electrostatic interaction between the proteins, the larger is the volume of the steering region. This encounter complex can be described as a precursor state before docking. In Figure 1.4 the free energy profile of the reaction pathway is shown. It has been discussed lately [17] that the encounter state could be separated from the final complex by a high energy barrier due to desolvation effects. Especially for charged residues the desolvation is thermodynamically very unfavorable. Electrostatic attractions which are favorable at intermediate separation distances, could thus be unfavorable during complex formation. Also the encounter complex might dissociate rapidly rather than forming the final complex, because of necessary structural rearrangements of the binding site and the precise structural alignment between the two proteins before final docking.
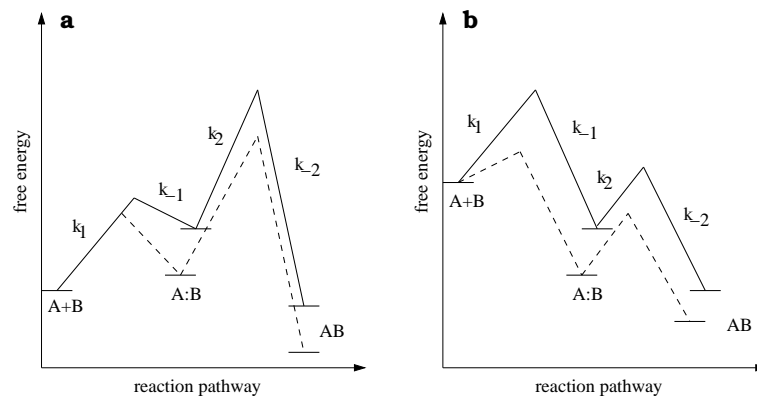
Figure 1.4: *Free energy profiles of possible reaction pathways.* **a***: Transition between encounter complex and final complex is the rate-limiting step for association.* **b***: Formation of the encounter complex is the rate-limiting step for association. Dashed lines represent the interaction of proteins in presence of favorable electrostatic forces [17].*

If there are favorable electrostatic forces, the proteins enter the steering region at

an earlier state and the random diffusion is shortened on account of increased directional diffusion. Since the ensemble of configurations defining the encounter complex on average resembles the docked complex, the encounter complex can be defined by distance constraints starting from the complex formation.

Recent investigations [18,19] identified hot spots, which caused large shifts in binding affinity upon mutation. They also have a large effect on the energy funnel [17] and are mostly located within or near the binding site to assist in steering the reaction partner. By performing Brownian dynamics simulations and computing k_{on} rates from the trajectories hot spots can be identified numerically since mutations of charged residues affect the transition state.

The association process of two proteins is often guided by long ranging electrostatic forces which result in a fast binding process. In contrast to uniformly charged colloidal particles, proteins often carry significant net charges and non-uniform charge distributions. As an example, the net charge of horse heart cytochrome c is +7e and its macromolecular dipole moment 300 Debye [20]. A direct consequence of this is a high ionic strength dependence of the association (and therefore also of the reaction) since the polar interactions of the proteins are then shielded by the ionic charges. On the other hand the desolvation forces leading to slow binding are independent of the ion concentration. Electrostatic forces also stabilize the so called encounter complex (see Figure 1.5), as well as desolvation effects, which can speed up hydrophobic paths and slow down the reaction paths involving charged residues. At this stage specific short-range interactions do not seem to play an important role [21–23].

Generally the kinetics of protein-protein interactions can be divided into diffusion controlled and reaction controlled events. Whereas the former is characterized by a very fast reaction after the time relevant step of diffusional association, the latter possesses a comparatively slow chemical reaction usually due to subsequent chemical processes. In the absence of systematic forces the upper limit for the association rate constant of the random search is $7 \times 10^9 M^{-1}s^{-1}$ - dictated by the Smoluchowski equation [24]. Of course this random collision is far from assuring complex formation, since the relative conformations of the two protein partners add additional constraints, which slow down the reaction by three to five orders of magnitude [25–27]. Association rates have been measured in the range of $10^5 - 10^9 M^{-1}s^{-1}$, where in the fast range the association is enhanced by strong, favorable electrostatic forces [28].

freely diffusing particles

bound complex

electrostatic steering

desolvation +
binding
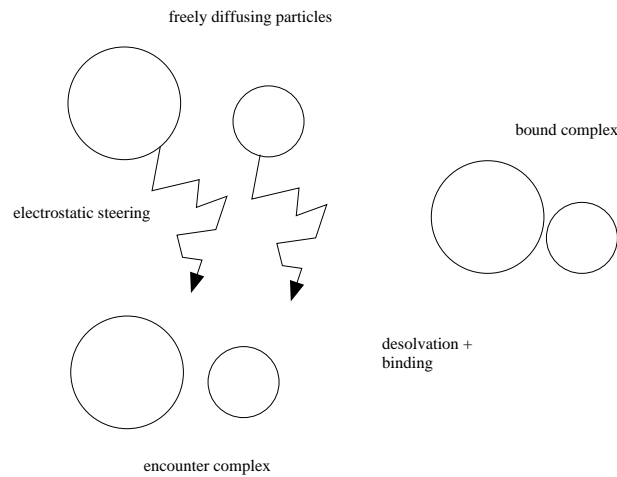
encounter complex

Figure 1.5: *Schematic representation of encounter complex and bound complex.*

The dependence of the association rate (k_{on}) on viscosity is often used to determine the control mechanism, since it has been shown that the relative association rate in the absence and presence of a viscous agent for a bimolecular interaction depends linearly on the relative viscosity of the medium for a diffusion controlled reaction [29].

The basal rate of association k_{on}^0 (in absence of electrostatic forces) is naturally related to the diffusional coefficient. Favorable electrostatic forces, as mentioned earlier, are often the driving forces for fast association. Altering the ionic strength of the solution and therefore the magnitude of the electrostatic attraction, naturally affects k_{on}. According to Debye-Hückel theory this relationship was shown to follow [23, 30]

$$\ln(k_{on}) = \ln(k_{on}^0) - \frac{U}{RT}\left(\frac{1}{1+\kappa a}\right),\tag{1.2}$$

where U is the electrostatic interaction energy, κ the inverse Debye length, and a the minimal distance of approach.

Chapter 2

Protein-Protein Docking

The prediction of protein-protein interaction sites has been the aim of many studies in recent years, since understanding the nature of protein association is of great importance for drug design as is the ability to predict potential binding sites. The physical basis for molecular encounter has been reviewed very recently [31].

Computer docking algorithms have gained significance over the past years since interest is currently shifting from the analysis of single proteins towards the understanding of protein-protein interactions, and ultimately of complex signaling pathways. The development of suitable methods has been greatly helped due to more structures of protein complexes being determined and structural data being made available through the Brookhaven Protein Data Bank (PDB) [32, 33]. Techniques for protein-protein docking have evolved following the relative success story of small ligand docking. Since the middle of the 1980s a large number of docking programs like DOCK [34] and LUDI [35] are available. The modern ones moreover include flexibility of the ligands, namely AUTODOCK [36] and DARWIN [37].

However, docking two proteins together is altogether much more demanding than docking a small ligand into the active site of a protein. The binding interface is much larger (2000 Å on average as found by Janin [38]), and the binding region may not even be known. As mentioned before common strategy to numerically determine the structural arrangement of protein-protein complexes is to use shape-complementarity, which is measured by various scoring functions (for reviews, see [38, 39]). In atomic detail, this task is computationally far too demanding at the current state. Therefore, using simplified, discretized descriptions of the protein surfaces on a grid has gained

great popularity, especially when Fast Fourier Transformation is being used to search translational and rotational space. Popular computer programs that implement this scheme are FTDOCK [40] and DOT [41].

2.1   Review of Docking methods

The term docking is used for numerical schemes which search for best matches between two molecules. Computational docking is of great current interest because it can give insights into molecular recognition, cellular pathways, and the building of macromolecular assemblies. The simulated docking process for two proteins can be divided into three parts [42]: the original docking method, a scoring procedure and finally a refinement of the highest scoring complexes. The first step is usually performed with implicit surface representations (e.g. grids) but also explicit surface representations (e.g. based on Connolly analysis [43,44]) have been explored. An effective search and matching algorithm should cover the relevant conformational space. A successful scoring function should be able to discriminate between native and non-native solutions. Mostly scoring of surface complementarity is combined with some kind of electrostatic scoring. Even more sophisticated methods implement desolvation penalties. The scoring process can be integrated in the search algorithm (e.g. in genetic algorithms) or be applied at a later stage when it will have less impact on the docking results.

In the refinement process which could be performed with rigid body minimization or minimization allowing side chain flexibility the knowledge of biochemical data can be incorporated to distinguish between good (near natural) complexes and false positives. Since the conformation of the long sidechains of lysine and arginine in fast binders might differ significantly the rigid body docking will in those cases give rise to a number of false positives. It is very helpful to diminish the number of possible complexes through this kind of filtering process as early as possible in the docking procedure. Problems arising from the search for natural binding conformations are usually connected to the fact that the original search for complexes is performed with rigid bodies and cannot account for side-chain flexibility. For the van der Waals and electrostatic contributions, soft potentials have to be used to avoid divergences. But the most challenging problem still remains to optimize the scoring function in order to eliminate all false positives from the high scoring range. It has been found that shape complementarity

based scoring works better for protein complexes with mostly nonpolar interactions. These are also mostly slow binders in contrast to the fast binders driven by polar interactions. Computational methods of docking include genetic algorithms, graph theory methods, Molecular Dynamics (MD) and Monte Carlo (MC) simulations to maintain feasible computation times. Especially for MD and MC simulations the knowledge of the binding site is essential. Several criteria for scoring functions have been applied like geometric complementarity, size of the nonpolar buried surface area, electrostatic interactions, the number of hydrogen bonds and unsatisfied buried charges, as well as specific amino acid contacts and solvation energy contributions. To reduce the number of solutions often rapid calculated parameters (geometric) are used. Docking and protein folding procedures will become even more related when the folding of multi-unit proteins will be tackled where the individual subunits will have to be docked against each other.

Particularly for electron transfer complexes the crystallization of complexes is extremely difficult because their association is weak and has to be stabilized during crystallization. It is even under discussion whether electron transfer complexes form a single productive complex or whether multiple productive conformations exist [45]. The kinetics of the interaction on the other hand is fast and makes major conformational changes very unlikely. Electron transfer (ET) systems are therefore very attractive for theoretical studies based on rigid body models. Understanding the fine details of the transport mechanism that requires structural models of the productive electron transfer complexes and the docking of two physiological electron transfer partners is the major interest in this study.

One can argue, however, that proteins may anyhow undergo small conformational changes upon binding to optimize the binding interface, and using rigid, atomistic models therefore would make it even harder to find good solutions. On the other hand, the use of atomistic force fields has the big advantage that one can more easily and reliably compare the relative energies of various docked conformations, as the task is to exactly find the lowest (free) energy conformation.

Here, we focus on the interaction of cytochrome c with its physiological redox partners cytochrome c oxidase and the bacterial reaction center. COX is the terminal oxidase in the bacterial and mitochondrial respiratory chains. As explained in the introduction during one catalytic cycle, four cytochrome c molecules bind on a surface

patch of COX and deliver a total of four electrons. Crystal structures from bovine COX and *Paracoccus denitrificans* were determined about six years ago [46–48]. No complex structures exist so far, however. Knowledge about the binding interface of COX and cytochrome *c* came from several experimental results [20, 49], and recently from site directed mutagenesis [5, 6]. The interaction between cytochrome *c* oxidase and cytochrome *c* was already subject of a previous theoretical study [11] where bovine heart cytochrome *c* oxidase and horse heart cytochrome *c* were docked with the program DOT [41]. The complex with the shortest distance from the exposed heme edge to the electron entry site Trp121 of cytochrome *c* oxidase (ranked 19th in energy) was proposed as the electron transfer configuration. However, the system studied by Roberts *et al.* is an artificial one, unlike the system of two physiological electron transfer partners studied here. Cytochrome c_{552} of *Paracoccus denitrificans* contains 72 N-terminal residues (membrane anchor) preceeding the typical cytochrome *c*-like domain. No structural information regarding these two domains is available. The first 30 of those residues are most likely forming an α-helical membrane anchor. The next 40 residues, most of them acidic, could be unstructured [50, 51]. Kinetic studies so far were performed with a soluble fragment of c_{552} where the membrane anchor was stripped off and this fragment was structurally characterized by X-ray crystallography and by NMR [52, 53]. In this study we combine the docking program FTDOCK [40], that has previously been used successfully in docking studies of enzyme/inhibitor and antibody/antigen systems with a rigid body energy minimization routine which has been implemented into the molecular simulation package CHARMM [54]. A crystal structure of the complex exists, however, for the complex of cytochrome c_2 with reaction center from *Rhodobacter sphaeroides*.

2.2   Structures

Coordinates for cytochrome *c* peroxidase (CcP) and iso-1-cytochrome *c* (cc) were taken from the crystal structure of the yeast CcP:cc complex determined at 2.3 Å resolution by Pelletier *et al.* [55] (PDB code 2pcc). All ionizable residues were assigned their usual protonation states at pH 7, so that the total charge is -13e for CcP and +6e for cc. Coordinates for the two-subunit cytochrome *c* oxidase (COX), determined by Ostermeier *et al.* [46] (PDB code 1ar1) to 2.7 Å resolution (see Figure 2.1) and for

a soluble fragment of cytochrome c_{552} (PDB code 1ql3), determined by Harrenga *et al.* [52] to 1.4 Å resolution (see Figure 2.2), both of *Paracoccus denitrificans* were also taken from the protein data bank. Additionally, a four subunit COX of *Paracoccus denitrificans* determined by Harrenga *et al.* [47] (PDB code 1qle) to 3.0 Å resolution was used to investigate the influence of subunits III and IV on the complex conformation. Of the four c_{552} molecules contained in the data file chain B was taken. COX was cut down to the non-membrane region interacting with cytochrome *c*. The selected coordinates included the "top" parts of subunit II with the copper *a* center and that of subunit I. The heme groups lie in the membrane spanning part and were not included. Each copper atom was assigned a charge of +1.5e [56,57], giving the selected subsystem a total charge of -13e. For horse heart cytochrome *c* (cyt *c*) we used the PDB structure determined by Bushnell *et al.* [58] (PDB code 1hrc) at a resolution of 1.9 Å. Using neutral histidine residues, a total charge of +7e resulted.
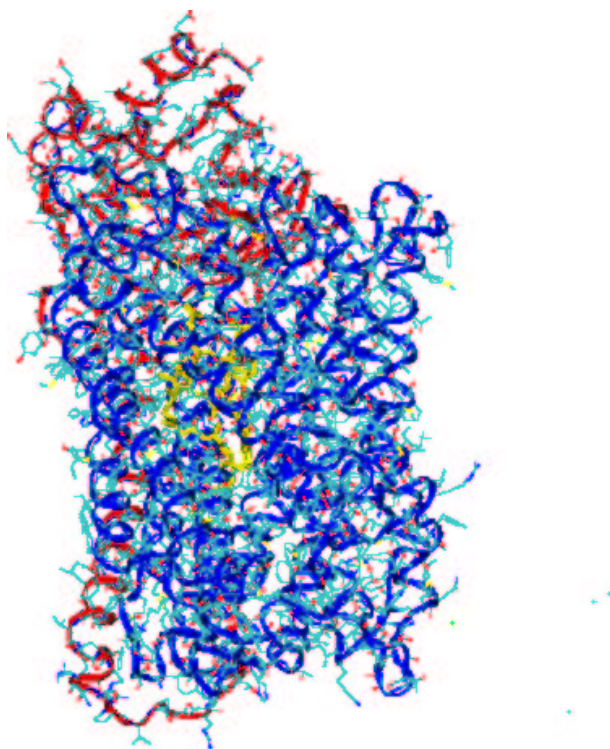
Figure 2.1: *Structure of the two-subunit cytochrome c oxidase determined to 2.7 Å resolution [46]. All-atom stick representation and ribbon drawing of chain A (blue) and B (red). The heme and the other co-factors are drawn in yellow.*
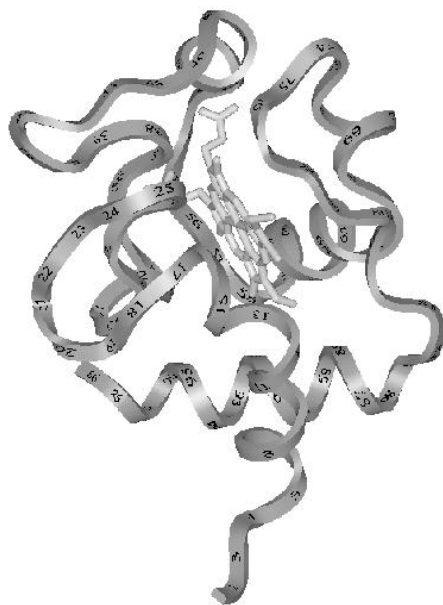
Figure 2.2: *Structure of cytochrome c_{552} oxidase determined to 1.4 Å resolution [47] in ribbon representation.*

The coordinates for the photosynthetic reaction center (RC) of *Rhodobacter sphaeroides* were taken from the crystal structure of uncomplexed RC determined by Ermler *et al.* [59] at 2.65 Å resolution (PDB entry 1PCR). The coordinates for the bound conformation of cytochrome c_2 (bcyt c_2) were taken from the crystal structure of the complex (see Figure 2.3) of cytochrome c_2 with photosynthetic reaction center from *Rhodobacter sphaeroides* determined by Axelrod *et al.* [60] at 2.40 Å resolution (PDB entry 1L9B). The coordinates for the uncomplexed form of cytochrome c_2 (cyt c_2) of *Rhodobacter sphaeroides* were taken from the crystal structure which was recently re-refined by Herb Axelrod to 1.6 Å resolution (unpublished data). For horse heart cytochrome c (cyt c), coordinates of the structure determined at 1.9 Å resolution by Bushnell [58] (PDB entry 1HRC) were used.

2.3 Docking algorithm

The docking of all protein pairs was performed with the FTDOCK program developed by Gabb *et al.* [40]. To achieve optimal shape complementarity, FTDOCK scans the full translational and rotational space by the use of a shape recognition algorithm
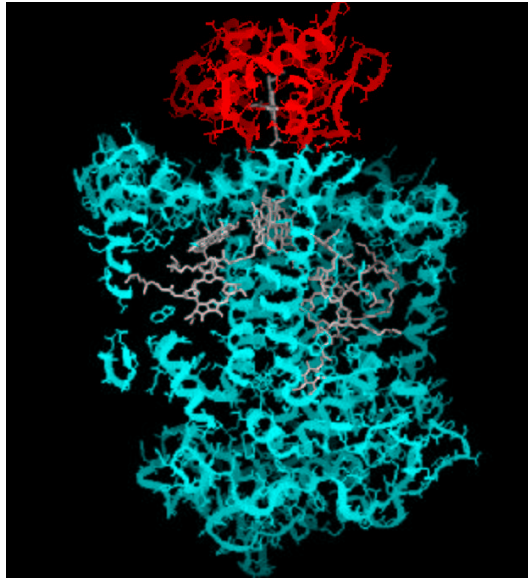
Figure 2.3: *Ribbon drawing of the complex of cytochrome c_2 (red) and the bacterial reaction center (blue) in the co-crystal [60]. The cytochrome c_2 porphyrin and the co-factors of RC are shown in grey.*

which is very similar to the one of Katchalski-Katzir *et al.* [61]. The two molecules are discretized on three-dimensional grids and the correlation of their surfaces is calculated by Fast Fourier Transformation. For the larger molecule a (COX and CcP), the shape distribution is represented by a value of 1 at each node (l,m,n), a negative value ρ in the core of the molecule and a value 0 for open space. The smaller molecule b is represented by a value of 1 at every node.

$$a_{l,m,n} = \left\{ \begin{array}{ll} 1 & \text{surface of molecule} \\ \rho & \text{core of molecule} \\ 0 & \text{outside of molecule} \end{array} \right\} \text{large molecule} \qquad (2.1)$$

$$a_{l,m,n} = \left\{ \begin{array}{ll} 1 & \text{surface of molecule} \\ 0 & \text{outside of molecule} \end{array} \right\} \text{small molecule} \qquad (2.2)$$

A correlation function is computed as the sum over all grid-points of the product of the two position scores.

$$c_{\alpha,\beta,\gamma} = \sum_{l=1}^{N} \sum_{m=1}^{N} \sum_{n=1}^{N} a_{l,m,n} b_{l+\alpha,m+\beta,n+\gamma} \qquad (2.3)$$

In this way a high correlation score denotes a complex with good surface complementarity. If the smaller molecule significantly overlaps with the larger one, the correlation is negative. A zero correlation score indicates that in most cases the molecules are not in contact. After each translational scan, the smaller molecule is rotated about its three Euler angles until the rotational space has been scanned completely. For an angular resolution of $\alpha = 10$ degree steps, for example, there exist 22,105 non-degenerate orientations. A simple Coulombic model is then used as a binary filter, leaving only those complexes with attractive electrostatic interactions. Out of these remaining complexes a large chosen number of complexes is recorded to file. To reduce the number of complexes to a manageable scope, post-processing consists of a filtering procedure, where distance constraints are applied to the residues, which are known to be part of the docking site through experimental investigations. In all our calculations $128 \times 128 \times 128$ grids are used. The grid size is the sum of the respective molecule diameters plus 1 Å and defines the grid spacing. In the following the parameters for the four systems investigated are listed in more detail:

a) CcP:cc:

With a maximum diameter of the complex of 108 Å the 128 cubic grid yields a grid spacing of 0.84 Å. A coarse scan was performed using 10 degree steps for the rotation of cyt *c* around its Euler angles. By keeping the best 5 complexes of each translational scan, the 20,000 complexes with highest correlation scores were stored. The distance constraint applied required the separation between the following residues and protein chains to be smaller than 4.5 Å (chain C:CcP, chain D:cc):

ALA193C-D   ALA194C-D   VAL197C-D   ASP34 C-D   GLU290C-D
ARG131C-D   TYR139C-D   GLU120C-D   HEM104D-C   LYS87D-C
LYS 86 D-C

This notation means that for a complex fulfilling the first filtering constraint (ALA193C-D), one atom of residue Ala 193 of CcP needs to be closer than 4.5 Å to a deliberate atom of cyt *c*. A subsequent fine docking (angular range of 9 degrees, 3 degree steps) of the 26 complexes could increase the correlation scores.

b) COX:cyt$c$:

This complex has a diameter up to 120.0 Å which gives a grid spacing of 0.94 Å. Again, the Euler angles were scanned in 10 degree steps and the best 20000 conformations were stored, keeping only those with the best score from each translational scan. In the filtering process, we used six acidic residues on subunits I and II and Trp 121 (chain A : subunit I of COX, chain B : subunit II of COX, chain C : c_{552}):

TRP121B-C   ASP178B-C   ASP257A-C   ASP156A-C   ASP135B-C
GLU126B-C   ASP159B-C

c) COX (1ar1):c_{552}:

The maximum diameter of the complex adds up to 111.1 Å yielding a 128 cubic grid with 0.87 Å spacing. By storing the 10 best of each translational scan, 10000 complexes were kept in the stack.

d) COX (1qle):c_{552}:

With a complex diameter of 123.9 Å the cubic grid yields a spacing of 0.97 Å. Again the best 10000 complexes were kept.

2.4 Combining Docking with a Force Field Minimization

Because the surface correlation scores of the FTDOCK program do not allow for a comparison of different docked complexes in terms of an interaction energy, we combined the docking procedure with an energy minimization applying an atomistic force field, namely the academic version of the molecular modeling package CHARMM [54]. Parameter set 22 was used [62]. Missing bonding parameters for heme a of COX were constructed in analogy to related fragments. The partial charges were those of a reduced heme. This is not thought to be crucial in this work because heme a is buried almost 30 Å below the binding surface of COX. As mentioned above, each of the copper

atoms of the dicopper site Cu A in subunit II of COX was assigned a partial charge of +1.5e.

Starting the energy minimization, first hydrogen positions were generated by the HBUILD command in CHARMM and were minimized in all configurations [63]. For the energy calculations we modified the standard steepest descent energy minimization routine of CHARMM to perform rigid body energy minimization including translation and rotation around the center of mass. To do so, the larger molecules (COX, CcP) were kept fixed while the smaller ones (cyt c's) were allowed to freely translate and rotate under the influence of the non-bonded forces exerted by the larger particle. The initial step size was 0.01 Å and the minimizations were carried out until a tolerance step size of 0.0001 Å was reached. The non-bonded interactions were computed using a distance dependent dielectric constant (which is a numerically efficient way to represent the solvent implicitly [64]) and an atom based cutoff of 40.0 Å for CcP and COX complexes. After energy minimizing all systems and complexes, the hydrogen positions were removed, newly generated and re-optimized with CHARMM to ensure maximally comparable configurations.

Also for the RC:bcyt c_2 pair the non-bonded interactions were computed using a distance dependent dielectric constant ($\epsilon = 5 \times r$) and an atom based cut-off of 40 Å for all complexes. After refining the relative orientation of the two proteins, the hydrogen positions of these minimized complexes were deleted, rebuilt and optimized using CHARMM. Further, 100 steps rigid body minimization with atom based cut-off of 100 Å for non-bonded interactions were carried out. A similar protocol of rigid body minimization using CHARMM was employed subsequent to MULTIDOCK [65] runs, which perform side-chain optimizations. The cofactors for RC were inserted back directly from the complexes built by FTDOCK. This is possible because during refinement using MULTIDOCK, the RC was kept static and the cofactors are not directly involved at the interface. The porphyrin was reinserted into these complexes refined by MULTIDOCK from the cytochrome (bcyt c_2 or cyt c_2 or cyt c depending on the case) that was transformed in reference to the cytochrome part of the docked complexes by superimposing the backbone atoms of the peptide chains of both. The aim was finally to compare these various refinement strategies so as to narrow down to one efficient way.

2.5   Docking of cytochrome _c_ peroxidase and cytochrome _c_

To develop and fine-tune a successful docking strategy we started working on this protein-protein complex with known three-dimensional structures. As is described above, a freely available docking program was combined with an atomistic force field. The docking strategy is shown schematically in Figure 2.4. The first method (FT-DOCK) is designed to efficiently scan the relative orientations of the molecular complex in a systematic way. The second method (CHARMM) allows to compute a realistic interaction energy for a drastically reduced number of complexes after energy minimization of their relative orientations. This additional rigid-body energy minimization is necessary because both methods use different representations of the molecular surface - a discretized surface on a grid vs. an atomistic representation.
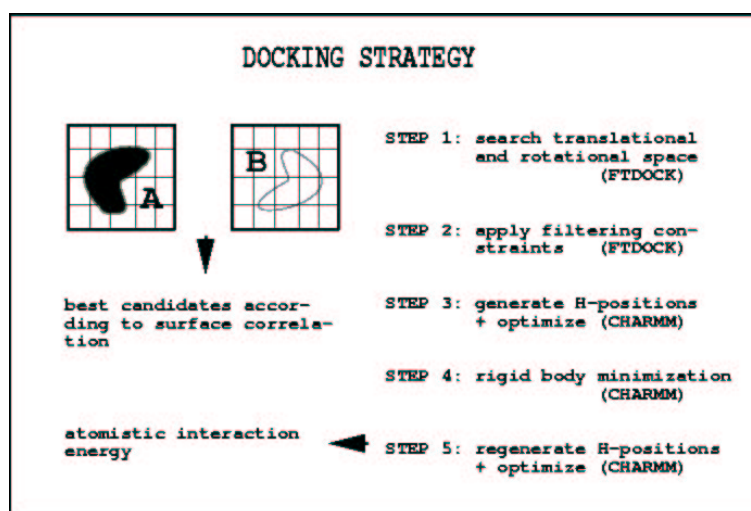
Figure 2.4: _Scheme for docking proteins A and B_

The calculated CHARMM energies given in Tables 2.1 to 2.6 refer to final values of complexes optimized individually. The results for the cytochrome _c_ peroxidase and cytochrome _c_ system are listed in Table 2.1. Successive filtering of eleven residue-chain contacts chosen from the crystal structure of the complex reduced the number from about 20,000 to 26 complexes. The distance constraints applied required that the separation between specific residues on one protein and the closest of the other protein to be smaller than 4.5 Å. For the crystal structure and the two docked complexes lowest

in energy (docking score 3271 and 10778) we tested using distance dependent dielectric constants of $\epsilon = 1 \times r$, $\epsilon = 3 \times r$ and $\epsilon = 5 \times r$, respectively.

Assuming that the crystal structure defines a natural binding site, one should expect that, with the appropriate value for the dielectric constant ϵ, the energy calculated using the atomistic force field should be minimal for this conformation. Since this was only the case for $\epsilon = 5 \times r$ this model was used for all energy minimizations. Note that a value of $\epsilon = 5 \times r$ was previously used successfully in modeling studies of protein kinases [66, 67]. Note also that the complexes were refined individually, so the van der Waals (vdW) energies differ as well. To explain the positive vdW energies it has to be mentioned that electrostatic and vdW energies also include intramolecular interactions inside the two proteins kept rigid. If we subtract the vdW energies for the isolated proteins we find that the vdW force is attractive for all given complexes. One of the energetically best ranking complexes (FTDOCK ranking 10778) is structurally very similar to the crystal complex. The RMS deviation to the energy minimized X-ray structure, which is referred to in Table 2.1, is 2.03 Å and only 1.82 Å when referring to the original complex structure. The center of mass of cytochrome c is shifted by 0.78 Å and 0.59 Å, respectively.

2.6 The complex of reaction center from *Rhodobacter sphaeroides* with cytochrome c_2 (Free and bound docking)

All calculations regarding the reaction center were carried out by V. Chandran [68], who was co-supervised by the author. Electrostatic energies of complexes docked in various conformations were computed by solving the Poisson-Boltzmann equation using the finite difference solver UHBD [69]. The interior dielectric constant within the protein was set to 4 and that outside was set to 78. An ionic strength of 50 mM and a radius of 1.5 Å were used for all calculations. Two sets of calculations were performed for RC: bcyt c_2 complexes, one with a 150 cubic grid and 1 Å grid spacing and the other with a 240 cubic grid and 0.625 Å grid spacing. This energy was then added to the

FTDOCK-score	ϵ	Electrostatic energy (kcal/mol)	vdW energy (kcal/mol)	Total energy (kcal/mol)	
X-ray	$5 \times r$	-165.8	562.0	396.2	
X-ray	$3 \times r$	-286.5	562.6	276.1	
X-ray	$1 \times r$	-987.5	575.1	-412.4	
24		-160.7	566.8		406.1
2050		-153.0	563.2		410.4
2115		-164.8	575.7		411.1
2263		-154.4	573.6		419.2
2374		-165.3	575.6		410.3
2616		-149.5	572.7		423.2
3271	$5 \times r$	-174.3	573.4	399.1	
3271	$3 \times r$	-301.8	575.9	274.1	
3271	$1 \times r$	-1040.4	593.0	-447.4	
4892		-153.9	571.7		417.2
5424		-166.4	572.4		406.0
6373		-159.9	574.1		414.2
10693		-163.1	573.6		410.5
10778	$5 \times r$	-166.2	565.5	399.3	
10778	$3 \times r$	-287.5	567.5	280.0	
10778	$1 \times r$	-991.5	583.1	-408.4	
12458		-159.3	566.6		407.3
12552		-158.8	587.8		429.0
13637		-162.4	566.4		404.0
13711		-163.2	577.0		413.8
16205		-157.0	591.3		434.3
16582		-162.5	566.9		404.4
16956		-172.1	596.8		424.7

Table 2.1: CHARMM energies for complexes of yeast cytochrome c and cytochrome c peroxidase: The 19 complexes listed are those that remained after applying the filtering procedure in FTDOCK and were then optimized in CHARMM by rigid-body minimization. The distance dependent dielectric constant was $\epsilon = 5 \times r$, except for three cases where different models ($\epsilon = 1 \times r$, $\epsilon = 3 \times r$ or $\epsilon = 5 \times r$) are compared.

interaction energies of both proteins computed with the molecular mechanics package CHARMM [54].

The structural difference between a docked complex of RC with cytochrome c_2 (bound and free conformation) and the crystal complex was determined in the following way: first, the docked structures were superimposed on the template crystal structure by the molecular graphics package InsightII [70] using the backbone atoms of RC. Then, the residual RMSD was computed between the C_α atoms of cytochrome c_2 in both structures using VMD [71]. Another measure of contact is the percentage of native contacts (PNC) which characterizes intermolecular residue-residue contacts as defined by Zhang *et al.* [72]:

$$\text{PNC} = \left(1 - \frac{\sum_i \sum_j |CN_{ij}^N - CN_{ij}^D|}{\sum_i \sum_j (CN_{ij}^N - CN_{ij}^D)} \right). \tag{2.4}$$

Here CN_{ij}^N and CN_{ij}^D represent the contact number matrix for native and docked structures, respectively, and hold detailed information about residue-pairing patterns between the two proteins in the complex. For any pair of residues i and j (i from the larger (static) molecule and j from the smaller (mobile) molecule) the contact number matrix CN_{ij} between them is computed as $\sum \sum n_{kl}$, where n_{kl} is the contact number between the kth atom of residue i and l^{th} atom of residue j. n_{kl} is defined as a function of the distance between the two atoms (r_{kl}) as follows ($r_{on} = 6$ Å and $r_{off} = 10$ Å):

$$n_{kl} = \begin{cases} 1 & : \quad r_{kl} \leq r_{on} \\ (r_{off} - r_{kl})^2 \frac{(r_{off} + 2r_{kl} - 3r_{on})}{(r_{off} - r_{on})^3} & : \quad r_{on} < r_{kl} \leq r_{off} \\ 0 & : \quad r_{kl} > r_{off} \end{cases} . \tag{2.5}$$

For evaluating PNC, it was not necessary to consider the cofactors of RC because they are located within the molecule. Hence, the peptide chains of the reaction center (L1-L281, M35-M301, H11-H250) and all atoms of cytochrome c_2 molecule (C1-C124 and heme) were used for PNC computations.

FTDOCK score	Electrostatic energy (kcal/mol)	vdW energy (kcal/mol)	Total energy (kcal/mol)	Heme-Trp121 distance (Å)
460	-153.0	6140.7	5987.7	12.4
955	-164.6	6137.4	5972.8	14.3
1473	-177.1	6137.8	5960.7	4.9
4049	-154.6	6128.9	5974.3	9.2
8672	-154.6	6141.5	5986.9	9.2

Table 2.2: CHARMM energies for the five best 1ar1-1hrc complexes: Two-subunit cytochrome *c* oxidase of *Paracoccus denitrificans* and horse heart cytochrome *c*.

2.7 Docking of cytochrome *c* oxidase from *Paracoccus denitrificans* and horse heart cytochrome *c*

The filtering constraints were motivated by the work of Witt *et al.* [5,6] who determined residues at the binding site critical for electron transport by site directed mutagenesis. Six acidic residues on subunits I (Asp256, Asp156) and II (Asp178, Asp135, Glu126, Asp159, and Trp121, which is proposed to be the electron entry site) were used in the filtering process. Successive filtering was performed for every possible combination choosing five out of these seven residues. 112 complexes fulfilled the condition that five of these residue-chain distances must not exceed 8.5 Å. In order to reduce this number to a manageable size for the next step in the docking procedure, we also requested the Heme(CBC)-Trp(CZ3) distance to be lower than 15 Å. Doing so, 23 complexes remained. The energies of the five complexes lowest in energy after CHARMM optimization are shown in Table 2.2. The complex with the lowest CHARMM energy (FTDOCK ranking 1473) also has the shortest Heme-Trp separation (4.9 Å) and the lowest electrostatic energy. This complex is shown in Figure 2.5.

Five putative salt-bridges (O-N distance $\leq$ 6.0 Å) are formed between the acidic residues of the COX interface and Lys residues on the cyt *c* side. A criterion of 6.0 Å was used instead of the contact distance of about 3 Å because the highly flexible lysine side chains could easily adapt if they were allowed to be flexible. A top view of the
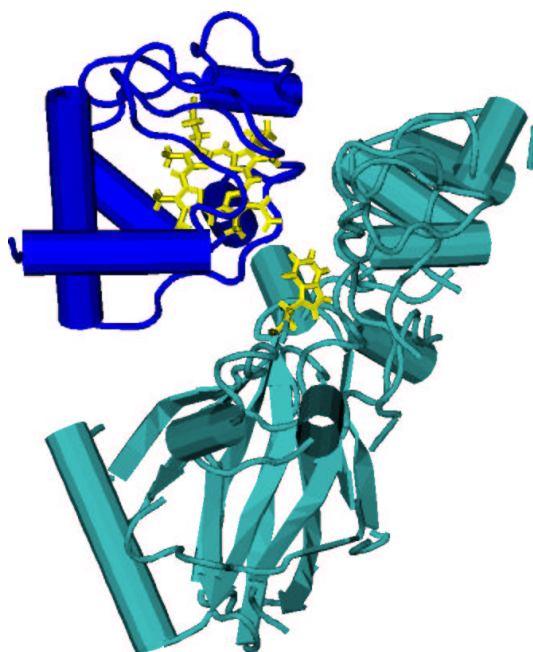
Figure 2.5: *Predicted best complex (FTDOCK ranking 1473) of horse heart cytochrome c (upper left) and the interface region of cytochrome c oxidase from Paracoccus denitrificans (lower right) in ribbon representation. The cyt c heme and Trp 121 of COX, the probable electron entry site, are colored yellow.*

interface section is shown schematically in Figure 2.6. This complex, obtained from docking the two-subunit structure of COX (*Paracoccus denitrificans*/PDB entry 1ar1) with cytochrome *c* of horse heart, has a very similar orientation to the complex found by Roberts *et al.*, where horse heart cytochrome *c* was docked against the bovine crystal structure of COX. The position and orientation of these two complexes differ only by a few Ångstrom and degrees, respectively. When superimposing the corresponding amino acids of the two COX structures the RMS deviation between the cyt *c* coordinates amounts to 3.8 Å and the center of mass is translated by 2.9 Å.
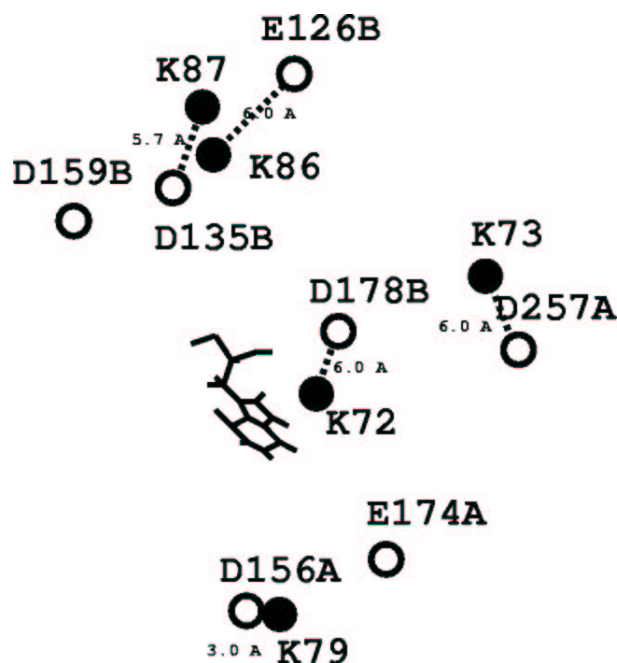
Figure 2.6: *Residue-residue contacts at the interface of the best predicted complex of horse heart cytochrome c and cytochrome c oxidase from Paracoccus denitrificans. In this orientation, the observer views through cyt c onto Trp 121 of cytochrome c oxidase. Black filled circles represent lysine residues (position of atom NZ) of cytochrome c. Empty circles are negatively charged Asp and Glu residues of COX (positions of atom CG and CD, respectively). The distances given in Å indicate putative salt-bridges and are marked by dashed lines when the circles are not positioned next to each other.*

2.8 Docking two subunit cytochrome c oxidase with cytochrome c_{552} fragment both of *Paracoccus denitrificans*

The same filtering conditions were applied as to the COX-c_{552}-complex, since we expected a similar interaction domain for all cytochrome c molecules. From the 48 complexes that remained after filtering, 15 satisfied the condition that the Heme-Trp121 distance should not exceed 15 Å. The results for docking the physiological electron transfer partner cytochrome c_{552} of *Paracoccus denitrificans* with the two subunit COX are shown in Table 2.3. In this case the fourth best complex (FTDOCK ranking 4491) in energy has the shortest Heme-Trp separation. However, as in the case of the best

FTDOCK score	Electrostatic energy (kcal/mol)	vdW energy (kcal/mol)	Total energy (kcal/mol)	Heme-Trp121 distance (Å)
788	-157.4	5454.4	5297.0	7.3
4389	-149.9	5470.4	5283.4	9.0
4491	-166.2	5454.5	5288.3	4.9
8062	-141.8	5410.5	5268.7	11.0
10763	-153.5	5438.7	5285.2	11.7
chimera	-153.9	5438.2	5284.3	4.6

Table 2.3: CHARMM energies for the five best 1ar1-1ql3 complexes (two-subunit cytochrome c oxidase and cytochrome c_{552} both from *Paracoccus denitrificans*) and of the chimera (best orientation of cyt c docked with four-subunit cytochrome c oxidase).

docked complex with horse heart cytochrome c, this complex has the lowest electrostatic energy. The interface with in this case three residue pairs forming salt-bridges is shown schematically in Figure 2.7. Comparing K77 and K70 of *Paracoccus denitrificans* with the corresponding lysines K79 and K72 of the horse heart structure shows a small counterclockwise rotation of about 6 degrees.

To investigate the role of subunit III cyt c_{552} of *Paracoccus denitrificans* was also docked against the interface region of the four subunit crystal structure of COX. The same filtering conditions were applied as for COX(1ar1)-c_{552} and COX:cyt c. From the 42 complexes obtained after filtering, eight satisfied the condition that the Heme-Trp 121 distance should not exceed 15 Å. The energies of the five best complexes are listed in Table 2.4. In this case, the third best complex (FTDOCK ranking 3757) has the shortest Heme-Trp separation of 4.2 Å and an electrostatic energy comparable to the two other complexes. Its interface region is shown in Figure 2.8. Compared to the two subunit complex (see Figure 2.7) cyt c_{552} is rotated counterclockwise by about 95 degrees.

FTDOCK score	Electrostatic energy (kcal/mol)	vdW energy (kcal/mol)	Total energy (kcal/mol)	Heme-Trp121 distance (Å)
2904	-154.8	5421.2	5266.4	12.1
3757	-154.3	5437.2	5282.9	4.2
5839	-151.2	5432.4	5281.2	7.1
7698	-138.0	5436.0	5298.0	10.1
10102	-137.5	5428.3	5290.8	6.7
chimera	-166.5	5432.2	5265.7	4.7

Table 2.4: CHARMM energies of the five best 1qle-1ql3 complexes (four-subunit cytochrome c oxidase and cytochrome c_{552} both of *Paracoccus denitrificans* and of the chimera (best orientation of cyt c Docked with two-subunit cytochrome c oxidase).

2.9 Docking of pre-organized binding partners (RC:bcyt c_2)

The bound conformation of cytochrome c_2 from the X-ray structure of the complex (Figure 2.3) was docked with the crystal structure of the reaction center in its unbound conformation using the docking program FTDOCK. Only the free form of RC was used for docking studies since there are no major conformation changes visible in the interface region of RC [60]. The resulting complexes were further optimized by the program MULTIDOCK [65] or with the molecular mechanics package CHARMM. The goal was to develop a successful docking strategy using this pair, later extending this to docking of native structures. Two sets of filtering constraints were used to reduce the total number of complexes to a manageable number and to extract the biologically relevant ones. The first set of constraints based on long-range electrostatic interactions between acidic and basic residues reduced the total number of docked complexes from 10,000 to 296. The second constraint between the cyt c porphyrin and the putative electron entry site RC:TyrL62 further reduced this number to 6. The final energies computed using CHARMM force field, the solvation energies computed using UHBD program and interaction energies calculated by the program MULTIDOCK are presented in Table 2.6. They are listed along with RMSD and PNC calculated in reference to the crystal structure. The parameters used for the computations are described above.

Complex (FTDOCK Rank)	Electrostatic energy[b] (kcal/mol)	vdW energy (kcal/mol)	Total energy CHARMM (kcal/mol)	UHBD energy[a] (kcal/mol)	CHARMM+ UHBD energy (kcal/mol)	MULTIDOCK energy (kcal/mol)	RMSD Å	PNC %	RPScore
X-ray	-176.3	3758.8	4691.5	-2261.5	2430.0	-80.2			
(a) FTDOCK									
495									-0.21
4248									-0.29
2840									-0.68
1072									-1.17
2852									-5.29
(b) FTDOCK + CHARMM[c]									
X-Ray	-178.1	3768.5	4699.3	-2256.2	2443.1		1.01	91.0	
495	-175.3	3755.8	4689.8	-2267.9	2421.9		2.07	85.0	
4248	-175.3	3747.8	4681.6	-2267.4	2414.2		2.02	85.4	
2840	-172.4	3745.1	4681.6	-2270.3	2411.3		2.13	80.6	
1072	-171.8	3747.5	4685.3	-2272.2	2413.1		3.21	76.8	
2852	-171.0	3772.6	4716.3	-2279.2	2437.1		19.31	7.4	
(c) FTDOCK + MULTIDOCK[d]									
495	-181.8	3889.1	4839.6	-2256.1	2583.5	-79.4	2.56	74.0	
4248	-179.8	3804.6	4744.3	-2263.6	2480.7	-71.8	2.52	80.0	
2840	-180.1	3765.3	4708.5	-2265.9	2442.6	-68.6	3.78	65.1	
1072	-180.1	3687.2	4628.7	-2259.4	2369.3	-74.5	3.70	73.0	
2852	-176.0	4530.1	5511.9	-2321.8	3190.1	-60.3	20.28	8.1	
(d) FTDOCK, MULTIDOCK + CHARMM[e]									
495	-188.3	3786.7	4700.5	-2260.4	2440.1		5.27	62.1	
4248	-174.8	3762.7	4693.2	-2269.4	2423.8		2.16	83.6	
2840	-177.3	3619.5	4550.8	-2273.6	2277.2		3.77	64.9	
1072	-177.0	3629.4	4561.4	-2262.4	2299.0		1.95	82.3	
2852	-171.4	3848.5	4786.2	-2273.7	2512.5		19.46	7.0	

Table 2.5: Calculated interaction energies between RC and docked cytochrome c_2 in its bound conformation

[a] computed with grid spacing 0.625 Å

[b] computed using $\epsilon = 5 \times r$

[c] CHARMM minimization of complexes (a)

[d] MULTIDOCK refinement of complexes (a)

[e] CHARMM minimization of complexes (c)

Figure 2.7: *Residue-residue contacts at the interface of the best predicted complex of cytochrome c_{552} and two-subunit cytochrome c oxidase both from Paracoccus denitrificans.*

2.10 Docking of binding partners in free conformations (RC:bcyt c_2)

The docking and filtering of free cytochrome c_2 (cyt c_2) and free reaction center was done similarly as above. The first set of constraints based on long range electrostatic interaction reduced the total number of docked complexes from 50,000 to 1175. The second constraint further reduced this number to 12. The final energies computed using the CHARMM force field and solvation energies computed using UHBD program along with RMSD and PNC calculated in reference to the crystal structure are presented for CHARMM rigid body minimized docked structures in Table 2.6. CHARMM rigid body minimization for docked complexes of RC: cyt c_2 refined using MULTIDOCK did not succeed. The interaction energies computed for docked complexes refined by MULTIDOCK and the UHBD solvation energies for these are listed below along with their RMSD and PNC in reference to the crystal structure.
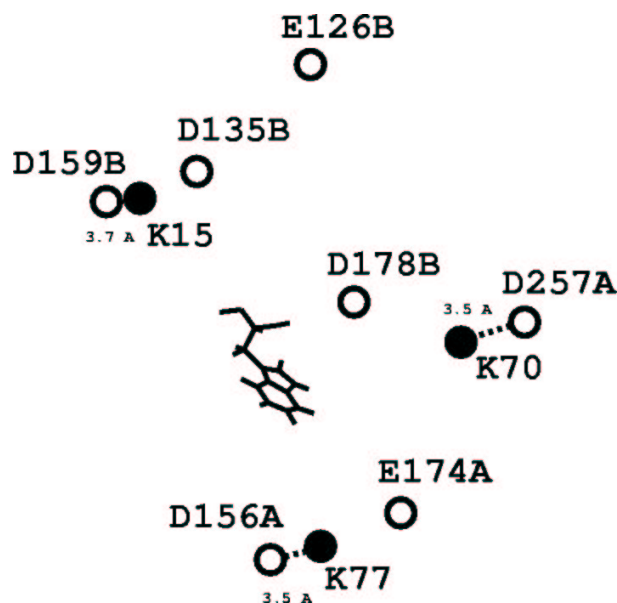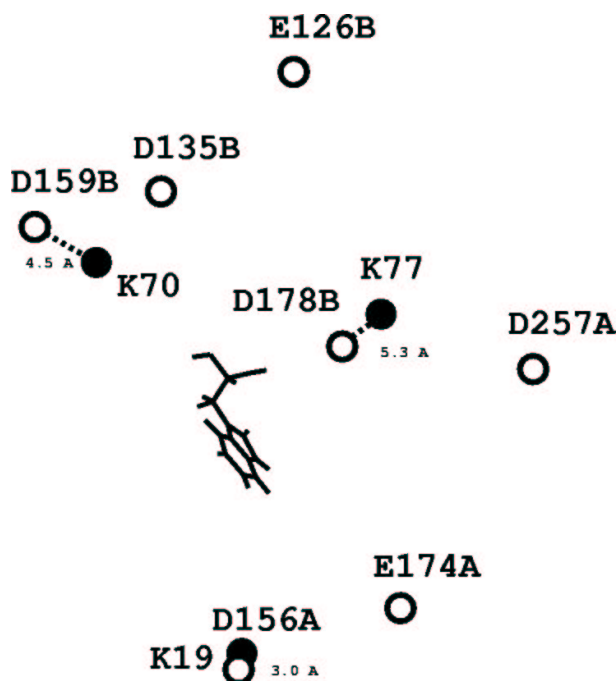
Figure 2.8: *Residue-residue contacts at the interface of the best predicted complex of cytochrome c_{552} and four-subunit cytochrome c oxidase both of Paracoccus denitrificans (1ql3-1qle). View as in Figs. 2.6 and 2.7.*

2.11 Discussion

Two different positions of cyt c_{552} with respect to COX were obtained as best complexes when COX coordinates were taken from the crystal structure of two-subunit COX or four-subunit COX. What makes the difference? Is it the tiny differences of parts of an Ångstrom between the two coordinate sets? To answer this question, chimeras were constructed, where the best cyt c_{552} position of docking was merged with the four-subunit COX coordinates, and where the cyt c_{552} position of docking was merged with the two subunit COX coordinates. Surprisingly, the chimera structures had slightly lower energies than the best original complexes (see Table 2.3 and 2.4). This indicates that small structural variations between the two- and four-subunit structures have no influence on the docking result. In another energy minimization the coordinates of subunits I and II of the four subunit structure were used for both cyt c_{552} orientations. The total CHARMM energies for these complexes with different orientations only differed by 1 kcal/mol.

Complex (FTDOCK Rank)	Electrostatic energy (kcal/mol)	vdW energy (kcal/mol)	Total energy (kcal/mol)	UHBD energy (kcal/mol)	CHARMM + UHBD energy (kcal/mol)	RPScore[a]	RMSD Å	PNC %
3808	-206.3	4757.8	5670.8	-2256.6	3414.2	-5.43	8.8	34
4613	-206.3	4785.3	5698.7	-2256.8	3441.9	-2.94	8.8	34
5308	-203.7	4780.6	5698.0	-2232.3	3465.7	-4.53	9.8	27
6083	-204.1	4784.8	5700.4	-2233.4	3467.0	-0.69	18.8	8
6900	-198.9	4772.7	5697.2	-2243.0	3454.2	0.26	3.0	69
15266	-199.0	4789.2	5713.4	-2243.0	3470.4	0.47	2.9	69
26048	-196.5	4787.6	5711.0	-2268.2	3442.8	0.42	12.6	31
29681	-209.4	4786.7	5697.4	-2257.7	3439.7	-0.51	7.0	33
31819	-197.5	4782.8	5704.3	-2256.5	3447.8	-4.00	17.1	14
40890	-200.4	4812.9	5735.1	-2236.9	3498.2	-2.20	11.0	29
46565	-201.6	4808.6	5729.0	-2249.8	3479.2	0.67	15.8	14
46612	-204.0	4805.4	5721.2	-2234.0	3487.2	-1.42	18.8	9

Table 2.6: Calculated interaction energies between RC and docked cytochrome c_2 in its free conformation.

[a] computed for docked complexes before refinement

As we started working on a complex formed by two proteins (cytochrome c peroxidase and iso-1-cytochrome c), for which a crystal structure exists, a successful docking strategy could be developed and the reliability of the method could be estimated. The results were very encouraging. First of all, the crystal structure has the lowest energy after energy minimization with CHARMM. This is a prerequisite for successful docking, because otherwise this conformation could not be located as a minimum energy conformation. 27 structures remained after docking with FTDOCK to achieve surface complementarity and successive filtering that incorporates structural biochemical information. After optimizing these in CHARMM by rigid-body energy minimization, the complex lowest in energy had a RMS deviation of less than 2.0 Å compared to the crystal coordinates. In the case of the COX(1ar1):cyt c complex, our best complex found is very similar to the best orientation found by Roberts *et al.*, who docked bovine heart COX with horse heart cyt c. This is a remarkable result, since different docking strategies were employed in both studies. The close agreement of both studies is a mutual justification of both approaches. It also indicates that horse heart cyt c is quite robust in binding in a preferential conformation. Compared to other cytochrome c's it carries a rather high positive charge and more lysines are available at the binding interface to form specific salt-bridges. The docked complexes obtained by FTDOCK can have unnaturally close interatomic distances between the two binding partners due to the discretization of the surfaces being used. It is therefore not possible to use the coordinates without modification in an energy evaluation by an atomistic force field. In our approach, a rigid body minimization was performed that carefully moves the two

proteins apart into the closest local energy minimum. When applying this procedure to the complex of Roberts *et al.* we notice that this complex is pushed apart a little bit as well. The Heme(CBC)-Trp(CZ3) distance increases from 3.3 Å to 5.5 Å and so do all other distances between interacting residue pairs. Seen in atomic detail the two proteins just sit together too closely and the van der Waals force is driving them further apart upon energy minimizing the system. This has to be kept in mind when comparing the resulting complexes of the two different approaches.

Although we found two different favorable conformations when docking c_{552} with the two-subunit or four-subunit structure, this does not necessarily indicate that subunit III of the cytochrome c oxidase plays an important role in the binding of cytochrome c. Surprisingly, the energies of the chimeras are even lower than the best complexes of the docked systems, which means that subunit III does not inhibit the favorable docking position found with the two subunit structure. Subunit III of *Paracoccus denitrificans* lacks a prominent surface loop found in subunit III of bovine COX. The presence of this loop may be of significance for the docking of cytochrome c on bovine COX. A superposition of bovine and *Paracoccus denitrificans* COX is shown in Figure 2.9 and the extra loop of bovine COX is shown in green. On the other hand, the results for the chimeras show that favorable conformations can be missed in the docking procedure. Note that translational scanning used a 0.84 Å step size and rotational scanning used a 10 degree step size. As it became obvious from further energy minimization calculations, the total energy of both complex configurations is very similar when using the crystal coordinates of the four subunit structure.

The conclusion that there exist two almost equivalent docking sites in this electron transfer complex reminds of the theoretical studies using Brownian dynamics simulations [73] and NMR data [74] which suggest the existence of alternative complex structures in the case of the CcP:cc complex. Our results suggest these two conformations as natural binding sites of the oxidase, even though early experimental results [55,75] on horse heart cytochrome c emphasize the importance of lysines 13, 72, 86, 27, and 87 (in decreasing order of importance) for the reaction with cytochrome c oxidase. In the case of the best two subunit complex (see Figure 2.7) only one of the formerly mentioned lysines (lysine 72, which refers to lysine 70 in *Paracoccus denitrificans*) forms a salt bridge with an acidic residue of COX. Lysine 13 is located at the interface but does not point to an electrostatically favorable position. This is due to the fact that the

highly flexible side chains are kept rigid here and the side chain of Lys 13 is somehow buried in the crystal structure of cyt c_{552} [5]. The second complex resulting from the docking with the four subunit structure also includes lysine 70 (72) and additionally Lys 19 which is in a similar position as the mentioned lysine 27 in horse heart cyt c when superimposing the two cyt c structures. Whether lysine 13, which seems to be of crucial importance in horse heart cyt c, should be equally important for c_{552} cannot be decided by these theoretical results alone. The position of the N-terminus seems mechanistically more important. The N-terminus of c_{552} in the first complex points away from the membrane surface. The direct distance from the membrane surface is about 50 Å. This spacing can however easily be bridged by the about 40 residue long solvent-exposed stretch that links cyt c_{552} with its N-terminal membrane anchor. On the other hand, the N-terminus of the second complex points towards the membrane. This close proximity seems unlikely. However the major aim of this work was to establish a strategy which, by incorporating biochemical data, will supply us with a number of possible active complexes. Here, two complexes were identified. We argue that the first complex is a better candidate because (a) it has a significantly better electrostatic energy, (b) its N-terminal position could have mechanistic advantages, and (c) its Lys 13 is in a more favorable orientation.

Three out of six docked complexes of bound cytochrome c_2 with Reaction Center both of *Rhodobacter sphaeroides* have a RMSD of ≤ 3 Å and a Percentage of Native Contacts (PNC) of more than 70 with reference to the crystal complex structure. This accuracy is well within the accuracy of protein docking methods. One could argue that it is not a real challenge to dock two proteins when the crystal structure of the complex is known. It is more important to predict binding conformations by using the coordinates of the unbound proteins. Here, 2 out of 12 docked complexes of free cytochrome c_2 with RC have RMSD ≤ 3 Å. Not surprisingly, docked cyt c_2:RC complexes have smaller RMSD values than cyt c_2:RC complexes, because in the former case all side chains are in their ideal conformations. Moreover, the energies for docked bcyt c_2:RC complexes are lower than for cyt c_2:RC complexes, as expected.

CHARMM rigid body minimization seemed to be a better strategy for refinement in comparison to MULTIDOCK with respect to total energies even though MULTI-DOCK improved electrostatic energies, which were calculated using CHARMM [68]. However, the CPU time required for CHARMM rigid body minimization is around

30 min on a COMPAQ Alpha 667 MHz processor whereas MULTIDOCK requires only around 3 min CPU time for refinement. The total energies computed for various docked complexes of bcyt c_2:RC are of similar magnitude as for the crystal complex suggesting multiple transient binding modes for cytochrome c_2. Northrup and co-workers studied the association process of cytochrome c with cytochrome c peroxidase by Brownian dynamics simulations [73]. In experimental conditions, the association constants are measured via the kinetics of electron transfer and thus include the possibility of transfer in multiple transiently bound conformations. Under crystallization conditions, however, one conformation may be thermodynamically preferred over the other conformations, or may be preferred in the structural context of a crystalline environment. In the case of cytochrome c_2:RC, the rate of electron transfer measured in the crystal was very similar to that measured in solution for complexes of solubilized RC and soluble cytochrome c_2. In solution, this rate is the first-order rate constant for reducing the special pair of RC by electron transfer from bound cytochrome c_2 after flashing a laser pulse to induce charge separation at the special pair. Therefore it seems valid to identify the crystal conformation of cytochrome c_2 with the binding equilibrium under excess of cytochrome c_2. Electron transfer under physiological conditions in membranes may still require only transient binding of cytochrome c_2. Crystallographic studies should therefore not be considered to be able to give a definite answer on the question whether multiple transient binding states may exist. The aim for theoretical investigations should be to identify the "binding energy funnel" by which two proteins bind as a function of (a) their distance, and (b) possibly other structural parameters such as the angle of their surface normals, or Percentage of Native Contacts. In contrast to the protein folding problem, it is not required that the bottom of the funnel only includes one structure.
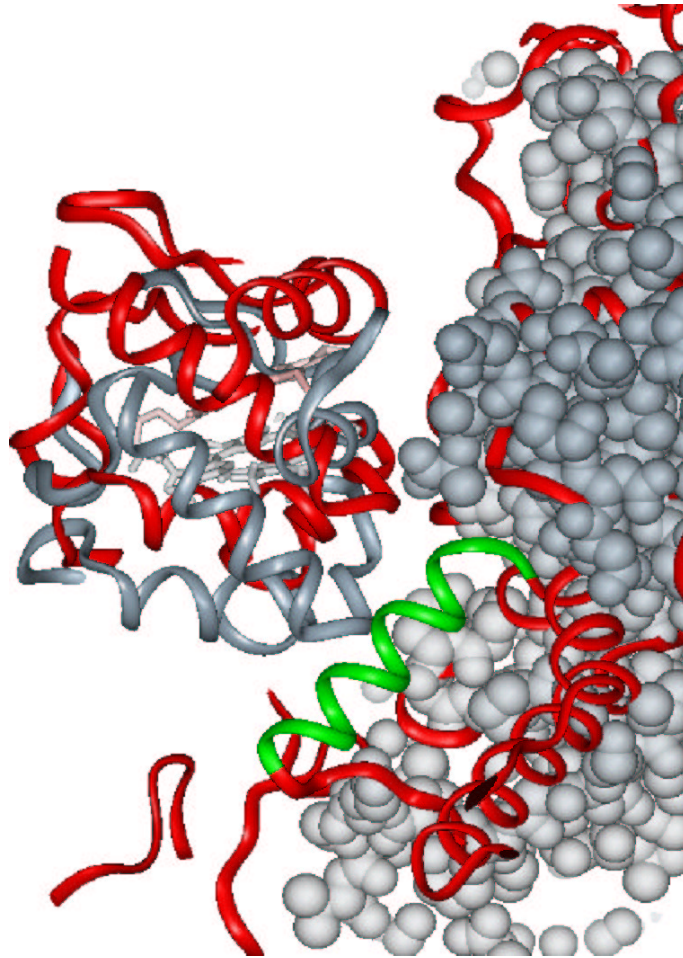
Figure 2.9: *Superposition of bovine COX (red ribbon) and COX of Paracoccus denitrif-icans (gray CPK). Horse heart cyt c as docked by Roberts et al. (red ribbon) is shown as comparison. The additional loop of bovine COX (green ribbon) collides with the position of c_{552} (gray ribbon) docked with the two-subunit structure.*

Chapter 3

Continuum Electrostatics of a Membrane System

3.1 Theory and numerical methods

Electrostatic interactions play a fundamental role for nearly all biomolecular systems and processes (for a recent review see [76]). Charged and polar groups are found ubiquitously in amino acids, nucleic acids and phospholipids. The association as well as the stability of those molecules in liquids are largely affected by these electrostatic interactions. Also it might be important for folding pathways. Early formation of secondary structure elements could prevent the polypeptide chain to explore the full energy landscape (see review [77]) and these long-ranged forces are likely candidates to orient the chain random motions towards the correctly folded structure. The importance of hydrogen bonds, largely due to electrostatics, has been established by early X-ray structures of proteins and DNA [1, 78]. Within protein-protein interfaces an average number of 10 hydrogen bonds have been found whereas one third of these involve charged residues [79]. The same study finds an average of one salt bridge per interface but stresses the large variability among complexes. Interestingly, even though direct charge-charge complementarity was not found to be significant in a data-set of protein complexes, this was the case for the complementarity of the calculated electrostatic surface potential [80]. In enzyme catalysis the stabilization of the transition state by the polar environment of proteins has been demonstrated by catalytic antibody de-

41

sign [81]. As already pointed out in the introductory part the kinetics of biomolecular association depends largely on electrostatic fields around the molecules, so that charge distribution effects should be considered. Several reviews report on macromolecular electrostatics and the interpretations of experimental results in this field [82–85]. One of the most fundamental approaches to treat these electrostatic effects in solution leads to the Poisson-Boltzmann equation. In the following this useful theory is introduced briefly.

3.1.1   Theoretical Introduction

If one treats the surrounding water molecules as a continuum medium of polar molecules, characterized by a dielectric constant ϵ, the electrostatic potential Φ of the solutes (macromolecules) is described by the Poisson equation:

$$\nabla^2 \Phi = -\frac{4\pi}{\epsilon}\rho, \tag{3.1}$$

with ρ the charge density represented in SI units. Here the charge density includes both the charge density of the fixed macromolecules as well as the charge density of the small mobile ions. For a set of point charges in a uniform dielectric the Poisson equation would reduce to Coulomb's law. The equation modifies in the general case of non-homogeneous media in order to account for polarization charges at dielectric boundaries [86]. The derivative of the space-dependent dielectric constant takes this effect into account:

$$\nabla[\epsilon(\vec{r})\nabla\Phi(\vec{r})] = -4\pi\rho(\vec{r}). \tag{3.2}$$

While the solute charges are given by the chosen molecular model, the concentration profile of the small ions is assumed to follow the Boltzmann Distribution:

$$c_i(\vec{r}) = c_0 \exp(-ez_i\Phi(\vec{r})/k_BT). \tag{3.3}$$

In this equation c_0 is the ion concentration far from the macromolecules (bulk concentration) and z_i the ion valence. The charge density of the ions is given by

$$\rho_{ions}(\vec{r}) = e \sum_{species} z_i c_i(\vec{r}). \tag{3.4}$$

Here we make the assumption that the ionic potential of mean force is equal to the average electrostatic potential multiplied by the charge of the ion. The condition of

charge neutrality requires that ρ_{ions} vanishes far from the macromolecules. Inserting 3.4 and 3.3 in 3.2 results in the second order non-linear Poisson-Boltzmann(PB) equation for the electrostatic potential with the condition that the potential vanishes at infinity:

$$\nabla\epsilon(\vec{r})\nabla\Phi(\vec{r}) = -4\pi\left(\rho(\vec{r}) + e\sum_{species} z_i c_0 \exp(\frac{-ez_i\phi(\vec{r})}{k_B T})\right). \tag{3.5}$$

In the high temperature limit where the electrostatic energy per ion is always small compared to the thermal energy and around macromolecules of low to medium charge density, we can linearize the exponential expression in 3.3 to obtain:

$$c_i(\vec{r}) \cong c_0(1 - ez_i\Phi(\vec{r})/k_B T). \tag{3.6}$$

Inserting 3.6 in 3.5 and introducing κ', which is related to the Debye-Hückel inverse length κ by

$$\kappa^2 = \frac{\kappa'^2}{\epsilon} = \frac{8\pi N_A e^2 I}{1000\epsilon k_B T}, \tag{3.7}$$

with e electronic charge, I ionic strength ($I = \sum_{species} z_i c_0$) of the solution and N_A Avogadro's number the linear PB-equation takes the form:

$$\nabla\epsilon(\vec{r})\Phi(\vec{r}) - \kappa'^2\Phi(\vec{r}) = -4\pi\rho(\vec{r}). \tag{3.8}$$

It is important to note that even though in most biologically relevant cases the potential is not very small, the solution obtained from the linear PB equation is close to the one obtained from the nonlinear PB equation. In a thorough comparison it has been shown that the onset of appreciable difference between the two treatments is linked to the charge density (electric field) at the solute-solvent interface [87].

Even the linearized equation can be solved analytically only for simple geometries. For example proteins should be modeled as spheres or ellipses (Tanford-Kirkwood theory [88, 89]), DNA as uniformly charged cylinders [90, 91] or membranes as planes (Gouy-Chapman theory [92, 93]). In the 1980s several numerical methods were developed to solve the PB equation. Nowadays more complex settings (down to atomic detail) can be solved numerically with finite element, boundary element or finite difference methods. The latter was first introduced by Warwicker and Watson [94] and became popular through software packages like DelPhi [95] and UHBD [69]. This method will be briefly described in the next section for a system of two solvated proteins.

The electrostatic free energy E of the system is given by

$$E = \frac{1}{2} \int d^3\vec{r}\, \Phi(\vec{r})\rho(\vec{r}), \tag{3.9}$$

which we will denote by $E = 1/2(\rho, \Phi)$ in the following. The electrostatic interaction free energy ΔE_{12} between two rigid solute molecules (1 and 2), where the charge density is given by atomic point charges can be decomposed as follows [97]

$$
\begin{aligned}
\Delta E_{12} &= 1/2((\rho_1, \Phi_1) - (\rho_1, \Phi_1^0)) + 1/2((\rho_2, \Phi_2) - (\rho_2, \Phi_2^0)) \\
&+ 1/2(\rho_2, \Phi_1) + 1/2(\rho_1, \Phi_2) \\
&= \Delta E_{1-c} + \Delta E_{c-2} + \Delta E_{1-2} + \Delta E_{2-1}.
\end{aligned}
$$

The first two terms (denoted with the subscript c for cavity) describe the influence of the dielectric cavity of one molecule on the potential of the other. The last two terms correspond to the interaction of each molecules charges with the potential of the other in presence of both dielectric cavities.

3.1.2 Finite difference Poisson-Boltzmann method

The solution of the linear PB-equation can be obtained by discretizing space (including the solutes and the surrounding solvent) into a cubic grid, which leads to a system of N^3 equations

$$
\begin{aligned}
\rho(i, j, k)h^2 &= \epsilon(i - \tfrac{1}{2}, j, k)(\Phi(i, j, k) - \Phi(i - 1, j, k)) \\
&+ \epsilon(i + \tfrac{1}{2}, j, k)(\Phi(i, j, k) - \Phi(i + 1, j, k)) \\
&+ \epsilon(i, j - \tfrac{1}{2}, k)(\Phi(i, j, k) - \Phi(i, j - 1, k)) \\
&+ \epsilon(i, j + \tfrac{1}{2}, k)(\Phi(i, j, k) - \Phi(i, j + 1, k)) \\
&+ \epsilon(i, j, k - \tfrac{1}{2})(\Phi(i, j, k) - \Phi(i, j, k - 1)) \\
&+ \epsilon(i, j, k + \tfrac{1}{2})(\Phi(i, j, k) - \Phi(i, j, k + 1)) \\
&+ \epsilon_s \kappa^2 \Phi(i, j, k)h^2.
\end{aligned}
$$

To each grid point values are assigned of the electrostatic potential, the charge density, the dielectric constant and the ionic strength, as is shown in Figure 3.1. The atomic

charges are allocated to the eight surrounding grid points in proportion to their separation distance from the assigned charge position. The resulting values for the potential grid can be obtained by an iterative method repeating the calculation until the values converge. Since dielectric values have to be assigned to each grid point the boundaries between the solutes and the solvent have to defined e.g. by the molecular surface or the solvent accessible surface. Then low dielectric values (usually between 2 and 4) are assigned to all points of the macromolecules and high values to all grid points outside this surface (around 80 for water). Since the dielectric properties of a material are mainly due to polarization effects and readjustment of its dipolar groups according to the outer electric field, this assignment can be justified for rigid structures. The only contribution from the conformationally stable molecule arises from the polarization effects which lead to a dielectric constant of about 2 for organic liquids [98]. The more flexible the molecule is the more the dipolar effects have to be taken into account with an increased dielectric value. Van der Waals radii and atomic charges are taken from specifically designed parameter sets or existing force fields.
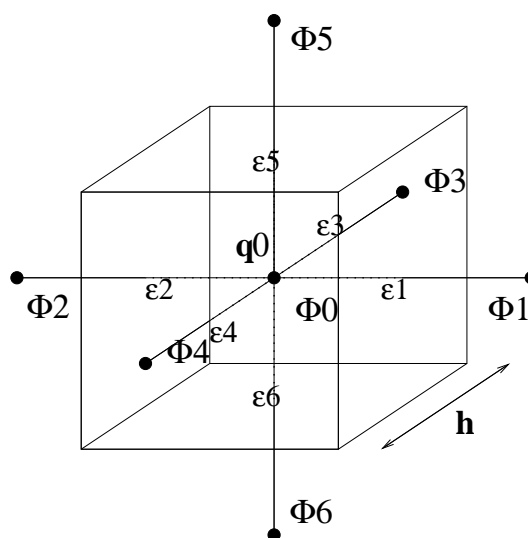
Figure 3.1: *Scheme for the finite difference Poisson-Boltzmann equation.*

The spacing of the electrostatic grid has to be chosen small enough to assure correct results (e.g. not merging opposite charges and correctly representing the solute-solvent interface). The dielectric discontinuities at the interfaces are best treated by smoothing functions to improve accuracy and convergence [96]. In order to compute free energies the grid-dependent self energy of charges must be taken into consideration. The free

energy is therefore largely dependent on the charge positions within the grid and the dimension of the grid [99]. For large systems this can get numerically very expensive and a so-called 'focusing' technique can be applied. A series of calculations starting with a coarse and ending with a fine lattice is performed, where the resulting potential values of the coarser grid are taken as boundary conditions for the subsequent refined calculation. Lately even nanosystems like microtubuli and ribosomes can be handled by the Adaptive Poisson-Boltzmann Solver (APBS) [100], which combines an adaptive finite element technique (Bank-Holst) [101] with standard focusing techniques of finite difference methods into a 'parallel focusing' method for the solution of the PB equation. If the potential serves as an input for a following simulation (e. g. a Brownian dynamics simulation), solving the full PB equation after each step is not feasible and the so-called test charge method (see e. g. [73, 102, 103]) or the effective charge method [97] are very helpful.

Applying the test charge method the potential of the protein, which is fixed during the simulation (usually the larger one), is calculated with the previously described finite difference Poisson-Boltzmann method before the simulation. Now, to calculate the inter-protein electrostatic force and torque the second protein can be treated as a test charge array moving in the field of the fixed protein. The low dielectric cavity of the moving protein is ignored and the electrolyte screening allowed throughout the protein interior: the second molecule is considered as a set of point charges immersed in the solvent with charges equal to the atomic partial charges or the net charges of atomic groups. The direct forces and torques acting on the second protein are obtained by consulting the stored grid forces and performing a summation over all test charges.

The effective charge method (ECM), which is applied for all Brownian dynamics simulations throughout this work is explained in detail in chapter 4.

The computation of the electrostatic potential at the solvent accessible surface is one of the most frequent applications of the Poisson-Boltzmann theory. It is a characteristic feature of protein-protein complexes that rather than exhibiting charge complementarity they exhibit complementarity of the electrostatic potential [3, 80]. Since the potential is greatly influenced by the charge distribution and the molecular shape, it is of special importance that the systems under investigation are carefully set up. The precise orientation of the sidechains might depend on the crystal packing, therefore one of the limitations of this approach is that the electrostatic properties are

computed on a single structure.

3.2 Set-up of membrane system

In order to study the influence of the membrane environment on the dynamics of cytochrome c and on the interaction of cytochrome c with cytochrome c oxidase a bilayer had to be constructed around the integral membrane protein. The binding patch for cytochrome c on the COX surface is about 1 nm away from the membrane surface, which is roughly 1 Debye length. Therefore it could not simply be assumed that the membrane acts as a geometrical perturbation. At such distances, atomic details may become very relevant. Therefore the membrane is represented in atomic detail as well. In this matter a protocol developed in B. Roux's group [104,105] was followed with slight modifications. The strategy consists of randomly selecting lipids from a pre-hydrated and pre-equilibrated set, scattering the lipids around the protein and reducing the number of overlaps by systematic rigid-body rotations and translations. To surround the protein by a complete lipid environment the dimension of the system in the xy-plane was set at 75 Å $\times$ 75 Å corresponding to an area of 5625 Å^2. Since the average cross-sectional area of a DMPC/DPPC molecule is 64 Å [106,107], the appropriate number of lipids can be determined to 51 for the upper bilayer and 50 for the lower bilayer. The DPPC polar heads were first represented by large spheres with the same cross-section to determine the initial position of each lipid. Therefore Molecular Dynamics (MD) and energy minimization with adequate boundary conditions are performed. The spheres were constrained at z $=+$ 19 Å and z $= -$ 19 Å for the upper and lower layers, respectively. Then these spheres were substituted by full DPPC lipids randomly chosen from a library of 2000 pre-hydrated and pre-equilibrated phospholipids [108]. Systematic rigid body rotations of the lipids around the z-axis and translations in the xy-plane were performed to remove steric clashes and unfavorable atomic contacts, which is then followed by an energy minimization to optimize the configuration. Next a pre-equilibrated water box of appropriate dimension was overlayed with the system and the resulting system was refined by energy minimization (see also Figure 3.2). The DPPC lipids exhibit a zwitter-ionic nature but are altogether neutral. The electrostatic properties of this structure are studied in the next section. For the Brownian dynamics simulations in chapter 5 the previously constructed membrane-embedded protein sys-

tem was expanded with pure lipid bilayers of the same unit length. These were placed around the central unit and since each unit is a square of length 75 Å this results in a side length of 225 Å for the total system. The structure contains over 200 000 atoms (including hydrogen atoms). Figure 3.3 shows a scheme of this larger system.
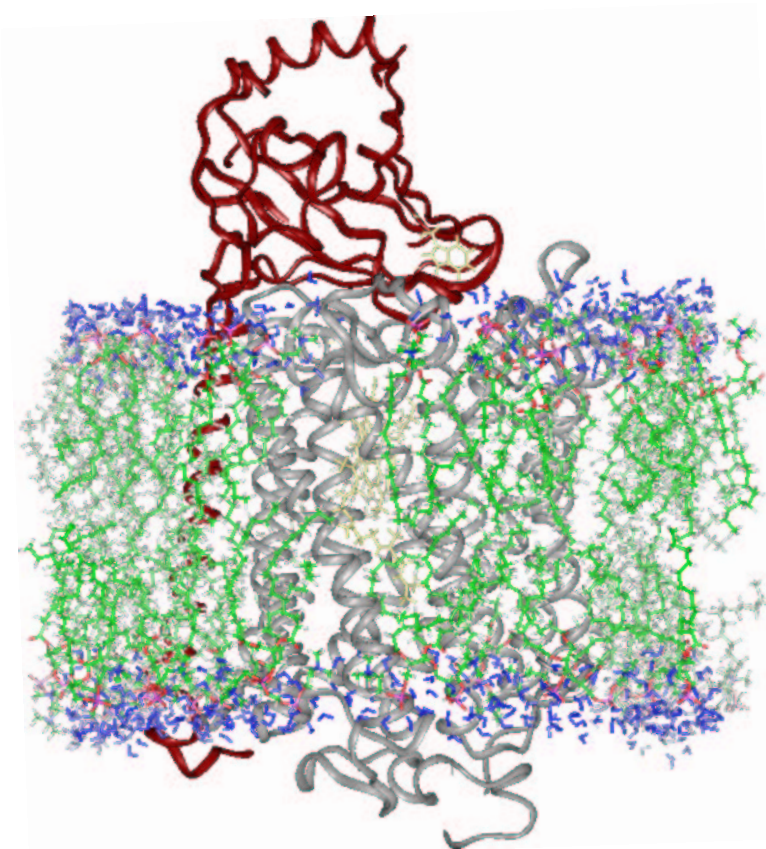
Figure 3.2: *Partly solvated lipid DPPC bilayer around cytochrome c oxidase, which corresponds to the central square of Figure 3.3*

3.3 Electrostatic properties of COX without/within the membrane environment

The structure shown in Figure 3.2 has dimensions of 7.5 nm × 7.5 nm × 7.5 nm. For the extended system (shown in Figure 3.3) the linear Poisson-Boltzmann equation was solved on a 1.55-Å mesh spacing applying a cubic 161 node grid with APBS [100]. The
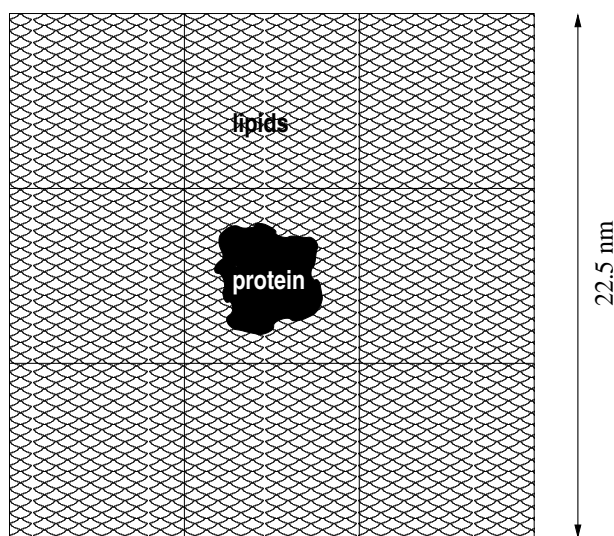
Figure 3.3: *Scheme of membrane surrounded cytochrome c oxidase (top view).*

solute dielectric constant was set to 2.0, that of the solvent to 78.5 and the calculation was performed for ionic concentrations of 50 mM, 150 mM and 250 mM for COX alone and for COX within the membrane environment. In Figure 3.4 the potential iso-contours of the central unit are shown for +1 kT/e (blue) and -1 kT/e (red).

The most obvious feature is the overall negative electrostatic potential of the oxidase as well as the positive electrostatic potential of the lipids. As expected the attraction for the positively charged lysine residues of electron transfer partner cytochrome *c* decreases as the charges are shielded by the solvent ions. Interestingly this effect is considerably decreased in the presence of the DPPC bilayer.

The calculation of the electrostatic grid potentials took about 40 min CPU on a COMPAQ Alpha 667 MHz processor with APBS and 3-4 times as long with UHBD. A comparison of grids for the same system calculated with the two different Poisson-Boltzmann solvers shows satisfactory agreement.
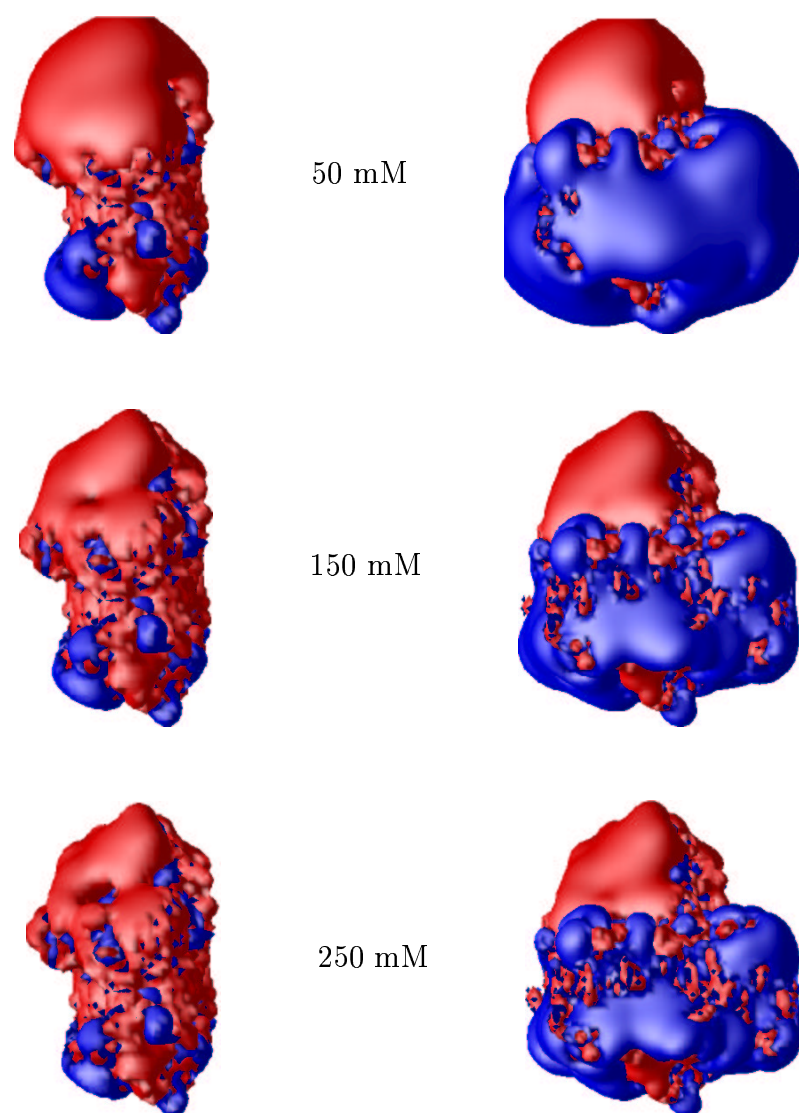
50 mM

150 mM

250 mM

Figure 3.4: *Potential iso-contours shown at $+1kT/e$ (blue) and $-1kT/e$ (red) of COX alone (left) and of the central unit of COX surrounded by lipid bilayer (right).*

Chapter 4

Brownian Dynamics Simulations

Preceding any associative protein-protein reaction an initial diffusional encounter is required, which may limit or at least partially influence the bimolecular rate constant. Therefore, the treatment of the dynamics of the associative encounters is of fundamental importance. The diffusional motion of a small particle in a solvent environment can be modeled by Brownian dynamics. Theory and simulation technique are introduced in the following.

It is of major interest to link Brownian simulations with experimental data. Therefore the scheme, which enables computation of bimolecular rate constants from trajectory statistics, is presented first.

4.1 Calculation of diffusional bimolecular rate constants

The ensemble flux through a reaction surface of particles, the motion of which is described by the steady state Smoluchowski diffusion equation, can be obtained as follows

$$\nabla \cdot D(\vec{r}) \cdot \left[\nabla - \vec{F}(\vec{r})/k_B T \right] \rho(\vec{r}) = 0 \tag{4.1}$$

or

$$\nabla \cdot \left[-\vec{J}(\vec{r}) \right] = 0, \tag{4.2}$$

Especially for MD and MC simulations the knowledge of the binding site is essential. where $\rho(\vec{r})$ is the pair probability density at configuration $\vec{r}$, $\vec{F}(\vec{r})$ is the negative gra-

dient of the intermolecular potential of mean force $U(\vec{r})$ and $D(\vec{r})$ the relative diffusion tensor.

Enforcing an outer boundary condition of constant ρ_0, the total integrated flux through the reaction surface is given as

$$j_{rx} = k\rho_0, \tag{4.3}$$

with k denoting the intrinsic reaction rate constant. The reactive flux through the spherical surface $r = b$ then is given by

$$j(b) = -4\pi b^2 r \cdot \vec{J}(\vec{r}) \mid_{r=b} = k\rho_0 . \tag{4.4}$$

The rate constant k can be decomposed into the rate k_D that any particle reaches separation b, and the probability P that it will ultimately react instead of escaping

$$k = k_D(b)P. \tag{4.5}$$

If we assume that the particles are spherical with isotropic reactivity and also react on every collision, $k_D(b)$ is given by the Smoluchowski result [109], which holds in the absence of external forces or hydrodynamic interactions for $r > b$

$$k_D^0(b) = 4\pi D_0 b. \tag{4.6}$$

Trajectories are not truncated when any given reaction criterion is met, but are continued until an outer surface of radius c is reached. After the particles pass this truncation border the run will be counted as successful or unsuccessful depending on whether the encounter definitions were fulfilled or were not fulfilled. The rate k is then given by [110]

$$k = \frac{k_D(b)\beta}{1 - (1 - \beta)k_D(b)/k_D(c)}, \tag{4.7}$$

where β is determined by the BD simulation as the fraction of the successful runs. For a scheme of the simulation setup see Figure 4.1 below. The 'start' sphere surrounding protein I typically has a radius of 5-10 Debye lengths plus the molecules' radii, whereas the 'end' sphere typically has a radius of 15-20 Debye lengths. In the ionic solutions under investigation the Debye length ranges from a few Ångstrom to a few nanometers depending on the ionic concentration (see equation 3.7).
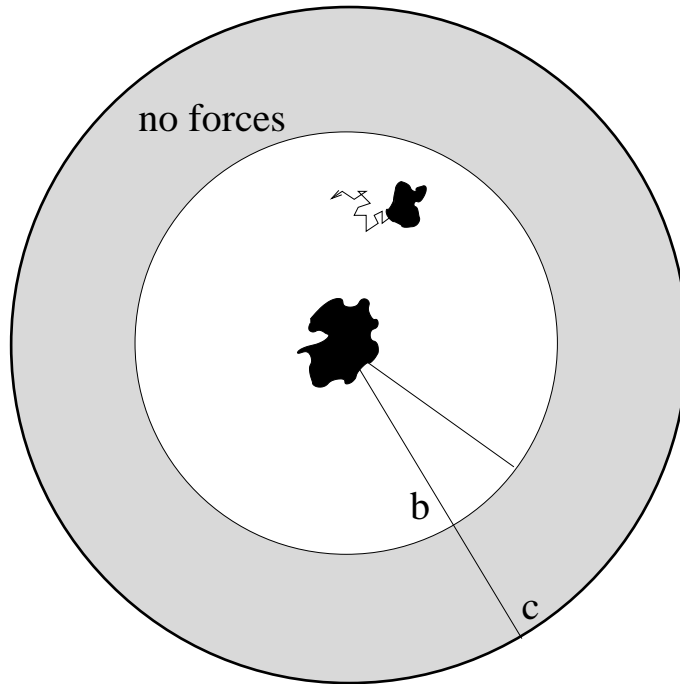
Figure 4.1: *Schematic picture of diffusional approach. Trajectories are started at random positions on a sphere with radius b and are truncated when they cross the sphere with radius c.*

4.2   Solving the diffusion equation with the Ermak-McCammon algorithm

The Langevin equations for a system of N Brownian particles are given by

$$m_i \dot{v}_i = - \sum_j \zeta_{ij} v_j + F_i + \sum_j \alpha_{ij} f_i, \qquad (4.8)$$

with $\sum_j \alpha_{ij} f_i$ the randomly fluctuating force exerted on a particle by the surrounding fluid. The moments of f_i are given by

$$\langle f_i \rangle = 0 \qquad (4.9)$$

$$\langle f_i(t) f_j(t') \rangle = 2\delta_{ij}\delta(t - t') \qquad (4.10)$$

and the coefficients are related to the hydrodynamic friction tensor

$$\zeta_{ij} = \frac{1}{kT} \sum_l \alpha_{il} \alpha_{jl}. \qquad (4.11)$$

For the diffusive time scale $\Delta t \gg \frac{m_i}{\zeta_{ij}}$ we can approximate $\dot{v}_i = 0$ and therefore equation 4.8 simplifies to

$$v_j = \frac{1}{kT} \sum_i D_{ij} F_i + \frac{1}{kT} \sum_l \sigma_{jl} f_l, \tag{4.12}$$

with the Einstein relation $\sum_j \zeta_{ij} D_{jl} = \sum_j D_{ij} \zeta_{jl} = kT\delta_{il}$ and $\sigma_{ij} = \frac{1}{kT} \sum_l D_{il} \alpha_{lj}$.

4.3   Simulation technique

For all following simulations the SDA (simulation of diffusional association of proteins) software [111] is applied to solve the equations of Brownian motions. The system of Brownian particles is placed in an initial configuration and moved along a trajectory in accordance with the configuration space distribution function $W(r_i, t)$. As the particle moves, properties like potential energy and squared displacement are calculated. At the end of the simulation the average value of each property is obtained, getting closer to the true system average for longer times t. In the computer simulation trajectories are composed of successive time steps t, each taken over short time step Δt, during which forces and torques are considered to be constant. The displacement is then given by the Ermak-McCammon algorithm [112]:

$$r_i = r_i^0 + \sum_j \frac{\partial D_{ij}^0}{\partial r_j} \Delta t + \sum_j \frac{D_{ij}^0 F_j^0}{kT} \Delta t + R_i(\Delta t). \tag{4.13}$$

$R_i(\Delta t)$ is the random displacement with Gaussian distribution function fulfilling

$$\langle R_i \rangle = 0 \tag{4.14}$$

$$\langle R_i(\Delta t) R_j(\Delta t) \rangle = 2D_{ij}^0 \Delta t. \tag{4.15}$$

Values for a set of $3N$ displacements R_i may be calculated from a weighted sum with equally distributed random deviates $[\{x_i\}; \langle x_i \rangle = 0; \langle x_i x_j \rangle = 2\delta_{ij} \Delta t]$ defining

$$R_i(\Delta t) = \sum_{j=1}^{i} \sigma_{ij} X_j \tag{4.16}$$

with

$$\sigma_{ii} = \left(D_{ii} - \sum_{k=1}^{i-1} \sigma_{ik}^2 \right)^{1/2} \tag{4.17}$$

and

$$\sigma_{ij} = \left(D_{ij} - \sum_{k=1}^{j-1} \sigma_{ik}\sigma_{jk} \right) / \sigma_{jj} \qquad (4.18)$$

for $i > j$.

To speed up the simulation hydrodynamic interactions are omitted in this work. In previous studies the computed encounter rates differed less than 20% because of this simplification [113, 114]. As will be shown, electrostatic effects of mutating single residues on either protein have far greater effects on the association rates.

4.4 Computation of forces

4.4.1 Effective charges for macromolecules in solvent

Long-range electrostatic forces, calculated using the Poisson-Boltzmann equation 3.8 are used in order to compute the diffusional trajectories of the charged proteins. This equation can be solved numerically for example through a finite difference approach via discretizing space into a cubic grid. The effective charge method (ECM) [97], then derives charges to represent the external electrostatic potential of the molecule in a uniform dielectric medium (within fitting accuracy). The effective charges ρ_i^{eff} are defined as point charges giving the same potential Φ_i^0 (see equation 3.10) outside the molecule i as the partial atomic charges in the inhomogeneous dielectric system. Then the electrostatic interaction free energy will be evaluated as the sum of the two cavity insertion terms, where as an estimation only dipole terms are taken into account, plus the interaction free energy between the effective charges of one molecule with the potential (Φ_i^0) of the other or between the effective charges of both molecules

$$\rho^{eff} = \sum_{j=1}^{n} q_j^{eff} \delta(\vec{r} - \vec{r_j}). \qquad (4.19)$$

The n effective charges are derived by minimizing a fitting functional:

$$J \;=\; \int d^3\vec{r} |\vec{\Phi}^{(0)}(\vec{r}) - \sum q_j^{eff} F_j(\vec{r})|^2 \qquad (4.20)$$

$$F_j(\vec{r}) \;=\; \frac{e^{-\kappa_s |\vec{r}-\vec{r_j}|}}{\epsilon_s |\vec{r} - \vec{r_j}|}. \qquad (4.21)$$

The integration is performed over the region relevant to intermolecular interaction. The solution is given through the following linear system of equations

$$\mathbf{A}\vec{q}^{eff} = \vec{b},\tag{4.22}$$

with

$$a_{ij} = \int d^3\vec{r}\, F_i(\vec{r})F_j(\vec{r}),\tag{4.23}$$

$$\vec{q}^{eff} = (q_1^{eff}, q_2^{eff}, ..., q_n^{eff}),\tag{4.24}$$

$$\vec{b} = \int d^3\vec{r}\, F_i(\vec{r})\Phi^{(0)}(\vec{r}).\tag{4.25}$$

To compute forces and torques acting on protein 2(1), the array of effective charges for protein 2(1) is placed on the electrostatic potential grid of protein 1(2). Molecular overlap difficulties arise because the cavity insertion terms cannot be treated as a single molecule. Therefore the described formalism requires the molecular surfaces to be separated more than twice the solvent probe radius (2×1.8 Å $= 3.6$ Å). The accuracy of the charge fitting can be measured by [97]

$$\chi = 1 - \int d^3(\vec{r})|\Phi^0(\vec{r}) - \sum q_j^{eff} F_j(\vec{r})^2|/\int d^3\vec{r}|\Phi^0(\vec{r})|^2,\tag{4.26}$$

which has its maximum value when the potential is fitted by itself and vanishes when the fitting potential is zero everywhere. The method has been successfully applied to a number of systems [111, 115, 116].

4.4.2   Van der Waals Forces

An exclusion volume prohibits van der Waals overlap. The exclusion volume is pre-calculated on a grid. If a move during the BD simulation would result in van der Waals overlap, the BD step is repeated with different random numbers until it does not overlap. For the grid calculation a spacing of 1 Å is chosen. Values of 1 are assigned to the interior of the larger protein and 0 outside. The surface-exposed atoms of the smaller protein are listed and steric overlap is defined to occur when one of the surface-exposed atoms is projected on a grid point with value 1 [118].

4.4.3 Desolvation forces

The charge desolvation penalty of protein 1 is taken as the sum of desolvation penalties of each charge of protein 1. The desolvation penalty of each charge is the sum of desolvation penalties due to the low dielectric cavity of each atom of protein 2, which gives the final expression for the desolvation penalty of protein 2 due to the presence of protein 1:

$$\Delta G_{ds} = \frac{\epsilon_s - \epsilon_p}{\epsilon_s(2\epsilon_s + \epsilon_p)} \frac{e^{2ka}}{(1 + ka)^2} \sum_{ij} (1 + kr_{ij})^2 \exp^{-2kr_{ij}} \frac{a_j^3 q_i^2}{r_{ij}^4}. \qquad (4.27)$$

ϵ_s is the dielectric constant of the solvent, ϵ_p the dielectric constant of the protein, k the Debye-Hückel parameter, q_i the value of the i-th charge on protein 2, a_j the radius of the j-th atom of protein 1 and a the sum of the average radius of the atoms of protein 1 incremented by the ionic radius of salt $\left(\sim 1.7 + 1.5 = 3.2 \text{ Å} \right)$.

Since short-range interactions, like the van der Waals attraction, are omitted, the applied force model becomes inaccurate for small protein-protein separation distances $(d \leq 4 - 5 \text{ Å})$.

4.5 Simulations of the diffusional association of cytochrome c with cytochrome c oxidase

Before presenting the results of the Brownian dynamics simulations, which were performed for this study, it should be mentioned that the reliability of the applied approach has been studied for many molecular systems before. The agreement with experimental studies, in particular, as far as comparison between different mutants of proteins and dependence on ionic strength of the solution are concerned, was found to be excellent [116]. Because the computed association rates have to be compared with experimental results, simulations were performed with the structures of the cytochrome c_{552} fragment and the soluble horse heart cytochrome c. Also, cytochrome c oxidase was not yet embedded in a membrane environment, since the stopped flow measurements for the k_{on} rates were performed with solubilized COX (subunits) [4, 117]. Membrane

Grid dimensions	161
Grid length	200 Å
Ionic strength	140 mM
Solute dielectric	78.54
Protein dielectric	2.0
Temperature	298.15 K

Table 4.1: Parameters for electrostatic calculations

proteins are isolated and solubilized by detergents which mimic the lipid bilayer environment. Hydrophobic regions of membrane proteins, normally embedded in the membrane lipid bilayer, are now surrounded by a layer of detergent molecules. The hydrophilic portions are exposed to the aqueous medium. The electrostatic potentials (solving the linear Poisson Boltzmann equation) for both proteins were computed with APBS [100] and/or UHBD [69] for comparison with the parameters given in Table 4.1.

The intermolecular electrostatic interactions were then calculated during the Brownian dynamics simulation by representing each molecule by a small number of effective charges in a uniform dielectric medium. This method is as efficient as the test charge method and at the same time provides improved accuracy. The effective charges are derived with ECM [97] by fitting them to reproduce the former computed molecular electrostatic potential calculated by the numerical solution of the finite-difference linearized Poisson-Boltzmann equation. The Brownian dynamics simulations were performed with SDA [111, 115]. The simulation parameters are given in Table 4.4. The best complex structures, which were the result of our previously performed docking study [119], see chapter 2, were taken to define the reaction criteria for the Brownian dynamics runs. The reaction patch is defined by the proximity between one or more intermolecular atom pairs (see Figure 4.2).

All inter-residue contact pairs which fulfilled the criteria of having a separation distance d_i of less than $d_c = 4.5$ Å were taken for monitoring the resulting association
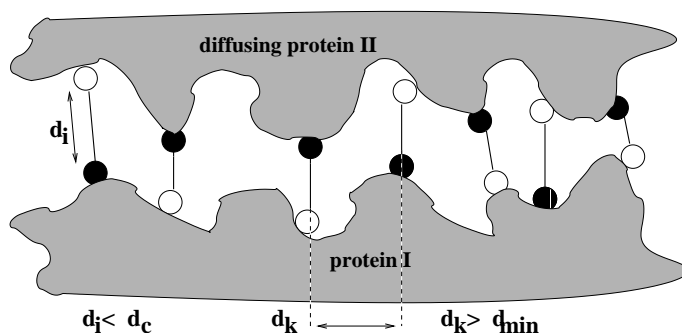
Figure 4.2: *Schematic picture of encounter definitions.*

rates. This distance assures a good definition of the protein-protein complex and is the default value of SDA [111, 115].

If a shorter separation distance is used naturally the number of contact pairs is reduced. Therefore, it becomes more and more unlikely that three to four contact pairs reach a close separation distance during the simulation of the diffusional motion. The computed association rate decreases, while the statistical error rises. Also such close inter-residue distances are rare in protein-protein complexes, especially when they are a result of a computational docking procedure with rigid body structures. On the other hand, increasing the contact distance and therefore the possible contact pairs will fail to define the docked complex and overestimate the association rates. Test simulations with different d_c showed that 4.5 Å is the fittest value to assure for reliable results.

Furthermore to assure for independent contact pairs the separation distance between the atom pairs is required to be greater than a mininimum distance $d_{min} = 7.0$ Å. For each system 4×4000 runs (with four different random number seeds) were performed. Finally the results were averaged over all simulated rates. The CPU time necessary for 16.000 runs is 32-48 h on a COMPAQ Alpha 667 MHz processor.

4.6 Comparison with experimental results

The interaction of electron transfer partners was extensively studied using Brownian dynamics simulations [73, 120, 121]. In an earlier investigation of the association between yeast iso-1-cytochrome c and cytochrome c peroxidase, the authors found that the reaction criteria giving the best match to experimental results, involved a limited heme edge distance and the relative orientation of the hemes being parallel to each other within the range of 60 degrees. Furthermore, the study provided evidence of a nonspecific Coulombic binding process between the two reaction partners. The so-called encounter complexes appeared to be stabilized by three or four ionic inter-protein contacts. Instead of a single electron transfer arrangement stabilized by strict charge complementarity, a more nonspecific association was proposed. This complex of CcP with yeast iso-1-cyt c was also the subject of a recent nuclear magnetic resonance (NMR) study [122], that investigated residues at the binding interface. In general, good agreement of the interaction in solution with the crystal conformation [55], being termed the high affinity binding site, was found. In the second Brownian dynamics study of yeast iso-1-cytochrome c and bovine cytochrome $b5$ the authors predicted a single binding domain (not a single conformation as would be observed in a crystal) contributing to the ET bimolecular dynamics. Two predominant classes of 1:1 complexes, involving two sets of interacting residues, were observed. A different study [75] used a Monte Carlo method to randomly place the two binding partners on a grid and search for low energy positions.

As it is discussed in section 1.3 the association rate gives insight into the dynamics of formation of productive electron transfer complexes. Experimentally easy to control variables are the ionic strength of the solution, which alters the strength of electrostatic interactions, and site-directed mutants where the importance of individual protein residues on the binding kinetics and thermodynamics can be investigated. Drousou *et al.* [4] extracted the association rate k_{on} for wild type and oxidase mutants from stopped-flow experiments. The observed time course was described by the sum of two exponentials. The fitted pseudo-first order rate constant was plotted against the cytochrome c concentration after mixing, and the bimolecular rate constant k_{on} calculated from the slope.

The values for wild-type, D135N and N160D mutants are shown in Table 4.2. The

Mutation	$k_{on} \times 10^6\,[M^{-1}s^{-1}]$
Wild type oxidase	3.7
D135N	0.3
N160D	2.8

Table 4.2: Oxidation of horse heart cytochrome c by wild type oxidase and subunit II mutants measured by stopped-flow at the ionic strength of 140 mM [4].

Ionic strength [mM]	$k_{on} \times 10^6\,[M^{-1}s^{-1}]$
10	4.1
35	1.5
200	0.1

Table 4.3: Oxidation of cytochrome c_{552} of *Paracoccus denitrificans* by the Cu_A fragment of wild type oxidase measured by stopped-flow at different ionic strengths [117].

mutants were considered for comparison with simulated rate constants, because D135N showed the largest effect compared to the wild type, while the association rate for N160D stayed in the same order of magnitude.

For the physiological ET partner cytochrome c_{552} k_{on} rates have been measured for different ionic strengths by Maneg *et al.* [117]. The values for 10, 35 and 200 mM are given in Table 4.3. At comparable ionic strength (200 mM vs. 140 mM) the association of cytochrome c_{552} is about 30 times slower than for horse heart cytochrome c. This can be explained by its smaller net charge and the less prominent dipolar character.

4.6.1   Wild-type COX of *Paracoccus denitrificans* and horse heart cytochrome c (140 mM)

The parameters for the Brownian dynamics runs are listed in Table 4.4. If not denoted differently they are taken for all the simulations. To assure the convergence of the

Time step $\Delta t1$	2 ps
Time step $\Delta t2$	10 ps
Relative translational diffusion constant	0.02 $\mathring{A}^2$/ps
Rotational diffusional constant	4.0×10^{-5} radian^2/ps
Radius b	115 $\mathring{A}$
Radius c	540 $\mathring{A}$

Table 4.4: **Parameters for Brownian dynamics simulations.**

simulated rates 4000 runs for 4 different random number seeds are performed for each individual setup. As can be seen from Equation 4.7 the simulated rate constant depends both on radius b and c. Still, even a large decrease of b does not change the order of magnitude of k_{on}. As mentioned earlier b should be chosen to be in the range of 5-10 Debye lengths. At distance b the electrostatic forces on the other protein should be isotropic and centrosymmetric. The trajectory truncation distance c was chosen so that making it even larger would not change the computed rates within statistical error.
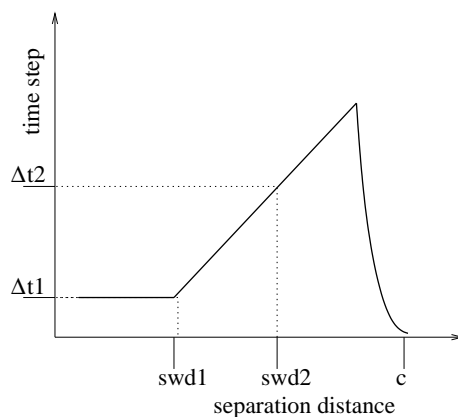

Figure 4.3: *Time step profile for Brownian dynamics simulations.*

For cytochrome c a translational diffusional constant of 0.015 $\mathring{A}^2$/ps [123] was assigned. The cytochrome c oxidase translational diffusion constant was calculated applying the Einstein-Stokes relationship. For spherical particles in aqueous solution at

room temperature, and with an effective hydration radius of approximately 50 Å for COX, this results in D = 0.005 Å^2/ps. Based on these individual diffusion constants a relative translational diffusion constant of 0.02 Å^2/ps was used. A rotational diffusion constant of 4.0×10^{-5} radian^2/ps was assigned to cytochrome c. Since the rotational diffusion constant for spheres depends on the inverse of the cube of the hydration radius the rotational diffusion constant for COX is neglectable. Test calculations supported this simplification.

A variable time step is used, that is smaller when the proteins are close to each other and also when the diffusing protein is close to surface c. The time step profile is shown in Figure 4.3. The chosen values for $\Delta t1$ and $\Delta t2$ are given in Table 4.4. For efficiency the rotations are performed less frequently than the translations. The rotations are carried out only when the accumulated random component of rotation is 3 degrees at small inter-protein separations up to 90 degrees at separations larger than 120 Å. The following inter-residue pairs served for the definition of reaction criteria, the distance between them was monitored throughout the runs:

Asp156:OD2-Lys79:NZ   Ala259:O-Lys73:NZ   Ser124B:OG-Lys86:NZ

Asp135B:OD2-Lys86:NZ   Asp135B:OD1-Lys86:NZ   Tyr122B:OH-Gly84:O

In Figure 4.4 the simulated association rates are presented for COX and horse heart cytochrome c. k_{on} rates are shown versus the RMS distance of the defined contact pairs. Naturally rates are highest for only one monitored contact pair (top line) and decrease with the increasing number of contact pairs (from top to bottom). Also, a strong dependence of the rates on the separation distance of the contact pairs can be seen: they drop markedly when the separation distance between the inter-residue pairs comes to a few Ångstroms.

For comparison the experimental measured value is introduced by a horizontal dashed-dotted line. For three contact pairs the simulated rates are closest to the experimental data. Still, even when small contact separation distances of 5.0-5.5 Å are satisfied, association rates are 10-100 fold overestimated by the simulation.

As mentioned before the experiments are conducted with solubilized COX. This might be a possible explanation for the mismatching of experimental and simulated values. A detergent belt of dodecyl maltoside molecules surrounds COX, which could change the electrostatic properties of the system and therefore also the association rate.

More likely, though, the mismatching is due to post-diffusional steps of association of this ET pair, which cannot be simulated by Brownian dynamics (see also chapter 5).
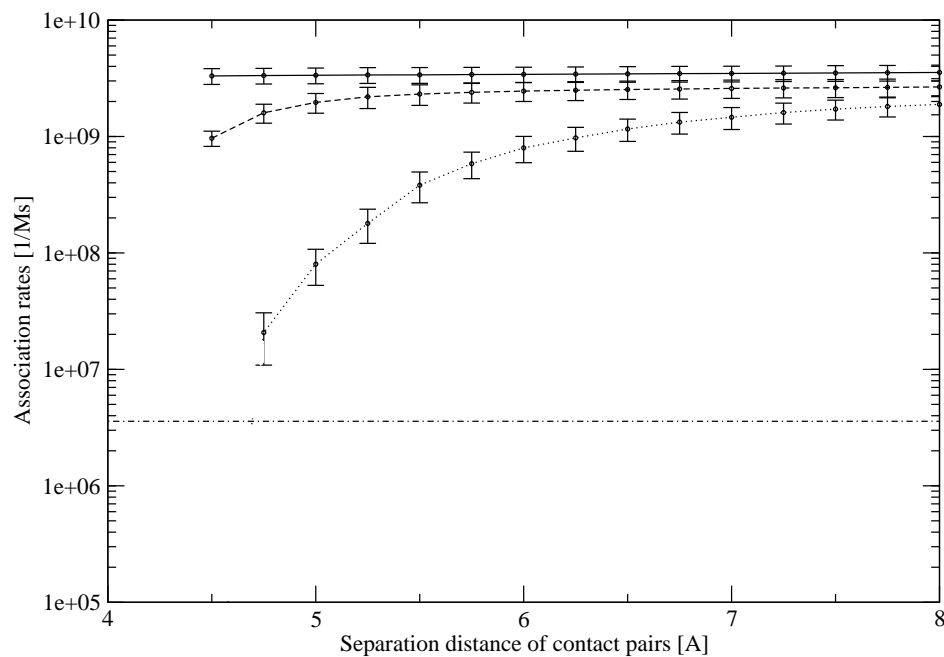

Figure 4.4: *Simulated association rates for wild-type COX and horse heart cytochrome c at 140 mM. k_{on} is shown in relationship to the reaction pair distance for one to three inter-residue pairs (from full to pointed line). The experimentally measured value is included in the graph and depicted by the horizontal dashed-dotted line. The error bars show the statistical error of the simulation.*

4.6.2 Mutation D135N of COX of *Paracoccus denitrificans* and horse heart cytochrome *c* (140 mM)

The following inter-residue pairs are monitored throughout the BD runs:

Asp156:OD2-Lys79:NZ   Ala259:O-Lys73:NZ   Ser124B:OG-Lys86:NZ
Asn135B:OD1-Lys86:NZ   Tyr122B:OH-Gly84:O

These are the same pairs as for wild-type, except that two atom pairs between Asp135-Lys86 are removed from the list. Instead the separation distance between one atom pair between Asn135-Lys86 is monitored.

The simulated association rates are presented in Figure 4.5. As for wild-type the matching of the experimental result is best realized when three contact pairs are taken to define the encounter complex and a contact separation distance of less then 5.5 Å is satisfied. Still, the experimental measured value is overestimated by about one order of magnitude.
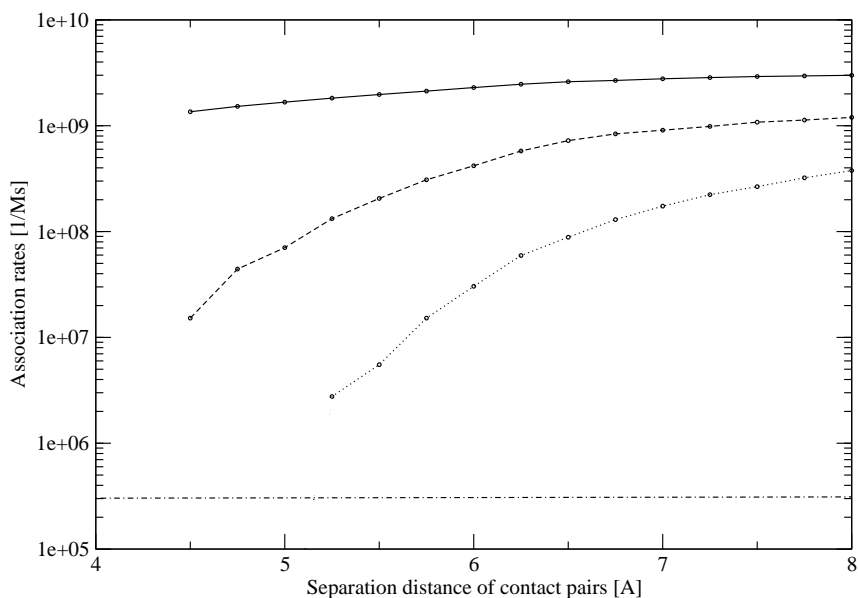

Figure 4.5: *Simulated association rates for mutant D135N COX and horse heart cytochrome c at 140 mM. k_{on} is shown in relationship to the reaction pair distance for one to three inter-residue pairs (from full to pointed line). The experimentally measured value is included in the graph and depicted by the horizontal dashed-dotted line.*

4.6.3 Mutation N160D of COX of *Paracoccus denitrificans* and horse heart cytochrome *c* (140 mM)

The following inter-residue pairs are monitored throughout the BD runs:

Asp156:OD2-Lys79:NZ	Ala259:O-Lys73:NZ	Ser124B:OG-Lys86:NZ
Asp135B:OD2-Lys86:NZ	Asp135B:OD1-Lys86:NZ	Tyr122B:OH-Gly84:O

These are the same pairs as for wild-type since they are not affected by the mutation. As for wild-type and the D135N mutant the definition of the encounter complex can best be realized by three contact pairs. The association rate for these three contact pairs and a separation distance of less than 5.5 Å up to 150 fold overestimated compared to the experimental value. Exact matching is realized if a very small contact distance of 4.5 Å is satisfied.
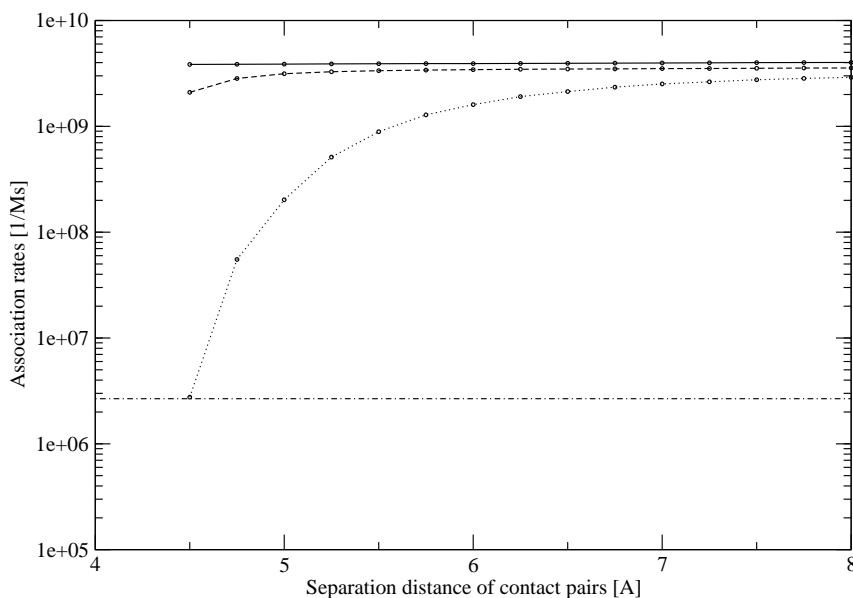

Figure 4.6: *Simulated association rates for mutant N160D COX and horse heart cytochrome c at 140 mM. k_{on} is shown in relationship to the reaction pair distance for one to three inter-residue pairs (from full to pointed line). The experimentally measured value is included in the graph and depicted by the horizontal dashed-dotted line.*

Both mutations influence the electrostatic field since a charged residue (Asp,D) is substituted by a polar, uncharged residue (Asn,N) and vice versa. The change of the electrostatic properties of COX is reflected in the observed rates, especially when the affected residues play a key role in the association process. These relative trends can
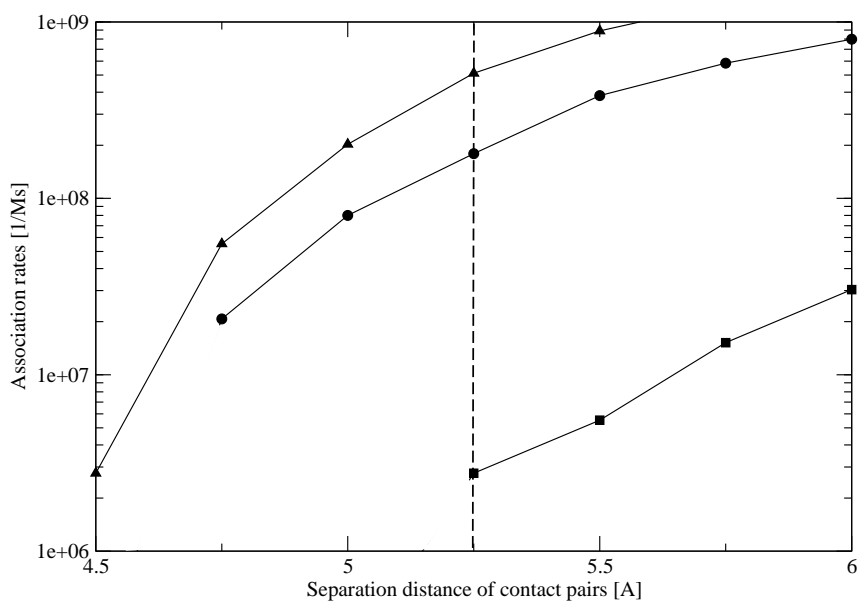
Figure 4.7: *Comparison of simulated association rates for wild-type (circles), mutant D135N (squares) and mutant N160D (triangles) COX and horse heart cytochrome c at 140 mM. k_{on} is shown in relationship to the reaction pair distance for three inter-residue pairs.*

very well be reproduced by Brownian dynamics simulations while not all overall effects can be taken into account.

Fig 4.7 shows a comparison of the results for cytochrome c association to wild type, D135N, and N160D cytochromes c oxidase for three contact pairs. In good quantitative agreement with the experimental data, N160D is close to the wild-type data whereas D135N association is markedly slower. For easier comparison a vertical line is drawn for a contact distance of 5.5 Å.

4.6.4 Wild-type COX and fragment of c_{552} of *Paracoccus denitrificans* (10, 35 and 200 mM)

For wild-type COX and its physiological ET partner cytochrome c_{552} the following inter-residue pairs are monitored throughout the BD runs:

Asp156:OD2-Asn45:ND2	Asp156:OD2-Lys77:NZ	Gln254:OE1-Asn68:ND2
Asp257:OD1-Lys70:NZ	Asp257:OD2-Lys70:NZ	Ala259:O-Asn68:ND2
Tyr122:OH-Lys13:N	Asp135:OD2-Lys15:NZ	Thr177:O-Ala81:N
ALA251:O-Lys9:NZ	Tyr122:OH-Gly12:O	Tyr122:OH-Lys13:O

The simulated association rates for 200, 35 and 10 mM ionic strength are presented in Figs. 4.8 to 4.10. As for horse heart cytochrome c simulated association rates are overestimated when compared to the experimental data. For three contact pairs and a separation distance of 5.5 Å between the inter-residue pairs simulated k_{on} are up to one order of magnitude higher. Still, like the experimental rates, the simulated k_{on} rates at comparable ionic strength (200 mM vs. 140 mM) for cyt c_{552} are much lower than for horse heart cyt c.

Figure 4.8: *Simulated association rates for COX and its physiological ET partner cytochrome c_{552} at 200 mM. k_{on} is shown in relationship to the reaction pair distance for one to three inter-residue pairs (from full to pointed line). The experimentally measured value is included in the graph and depicted by the horizontal dashed-dotted line.*
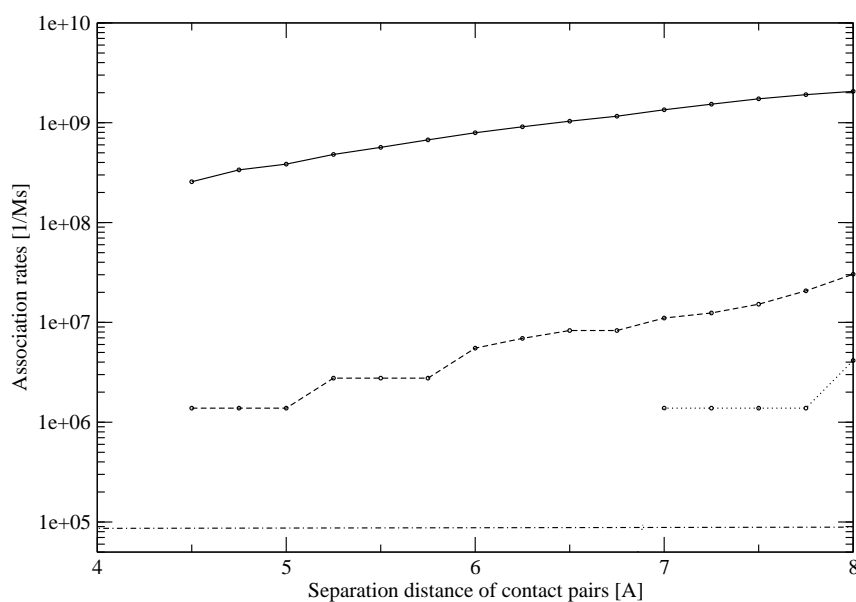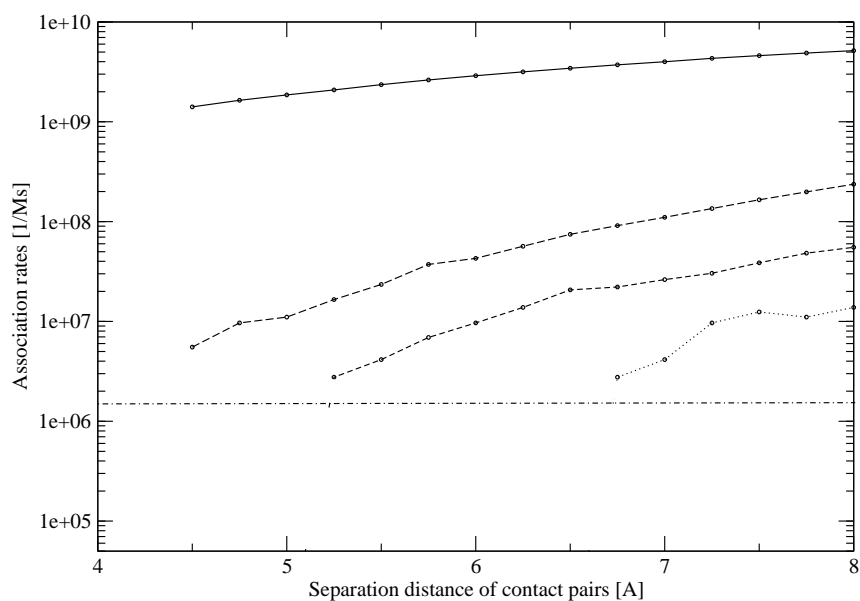
Figure 4.9: *Simulated association rates for COX and its physiological ET partner cytochrome c_{552} at 35 mM. k_{on} is shown in relationship to the reaction pair distance for one to four residue pairs (from full to pointed line). The experimentally measured value is included in the graph and depicted by the horizontal dashed-dotted line.*

Fig 4.11 shows a comparison of the results for cytochrome c_{552} association for three contact pairs. In good agreement with the experimental data the association becomes smaller as the ionic strength of the surrounding solution rises. For easier comparison a vertical line is drawn for a contact distance of 5.5 Å.
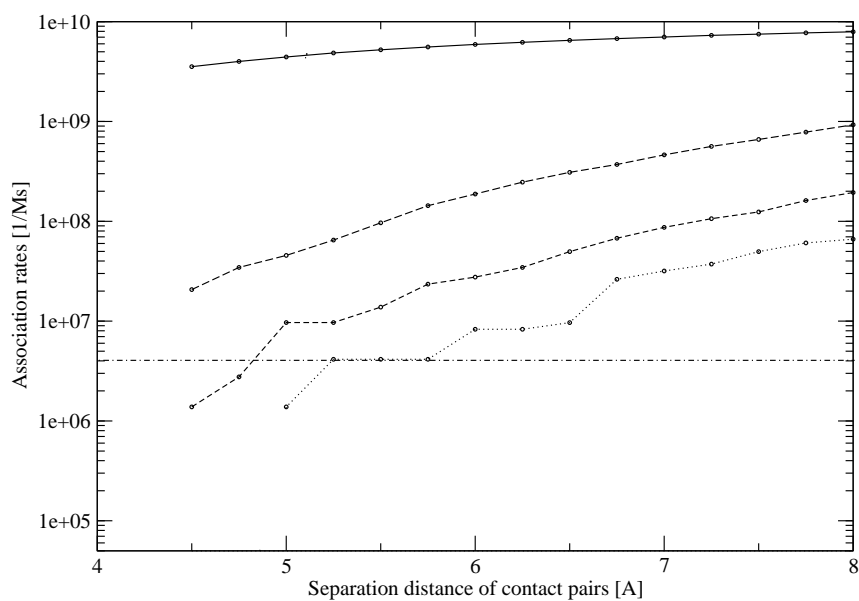
Figure 4.10: *Simulated association rates for COX and its physiological ET partner cytochrome c_{552} at 10 mM. k_{on} is shown in relationship to the reaction pair distance for one to four residue pairs (from full to pointed line). The experimentally measured value is included in the graph and depicted by the horizontal dashed-dotted line.*

Figure 4.11: *Comparison of simulated association rates for wild-type COX and cytochrome c_{552} at 10 (circles), 35 (squares) and 200 mM (triangles). k_{on} is shown in relationship to the reaction pair distance for three inter- residue pairs.*

4.7   Association rates for membrane embedded systems

The setup for the BD was modified to enable the simulations with the membrane embedded cytochrome c oxidase. Trajectories are now started only at the upper hemisphere (see Figure 4.12 for a scheme o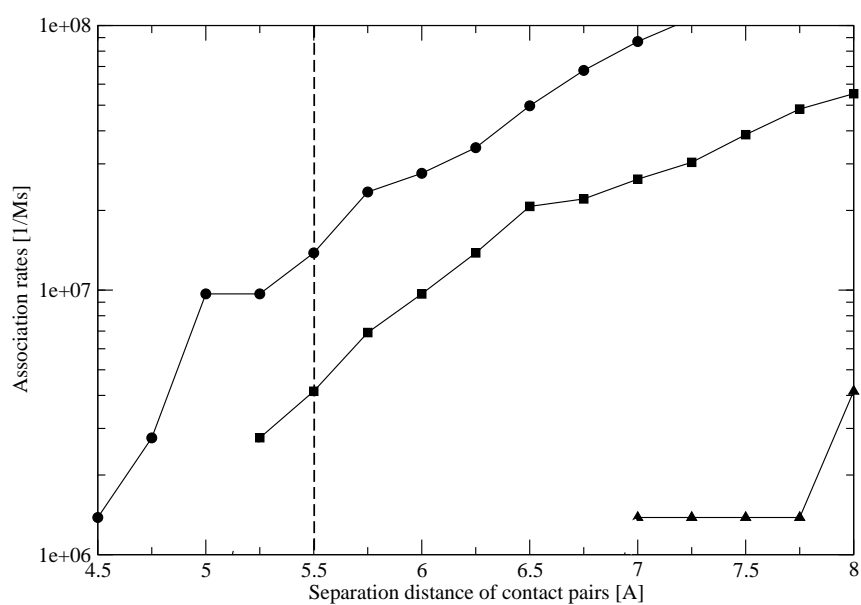f the modified setup). The algorithm was changed in the way that here the randomly chosen starting position are accepted only if they are on the protein-protein interaction side of the system. Also a minimum distance of 20 Å is assured so that the diffusional simulation of the smaller particle is not started too close to the membrane.

The new setup was tested against the standard setup with a BD simulation of cytochrome c oxidase and horse heart cytochrome c. A comparison of the simulated rates is shown in Figure 4.13 and shows excellent agreement.
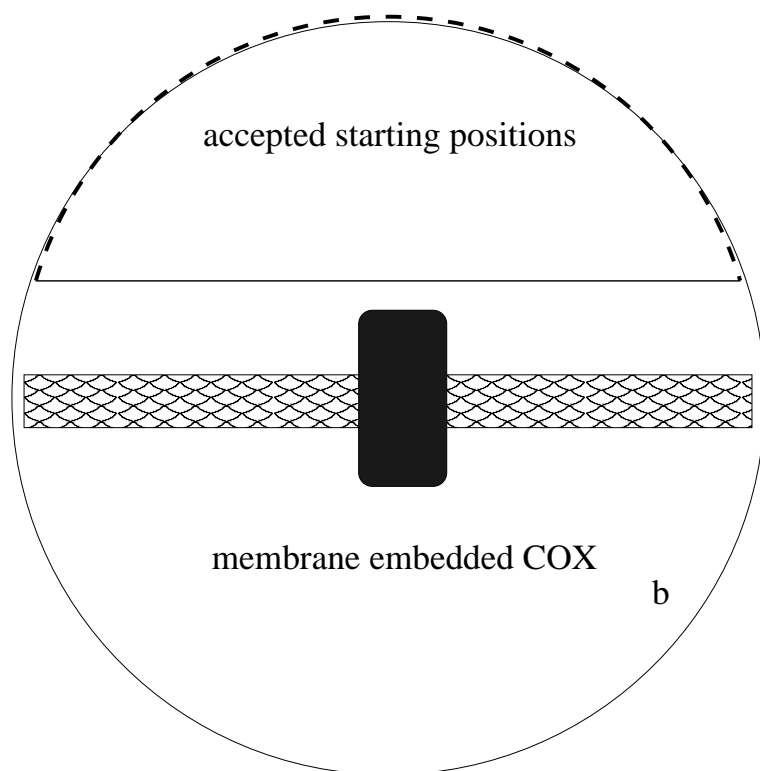

Figure 4.12: *Scheme for modified setup for Brownian dynamics simulations. Trajectories are started exclusively at the upper hemisphere.*

In Figure 4.14 the calculated association rates for membrane embedded COX and
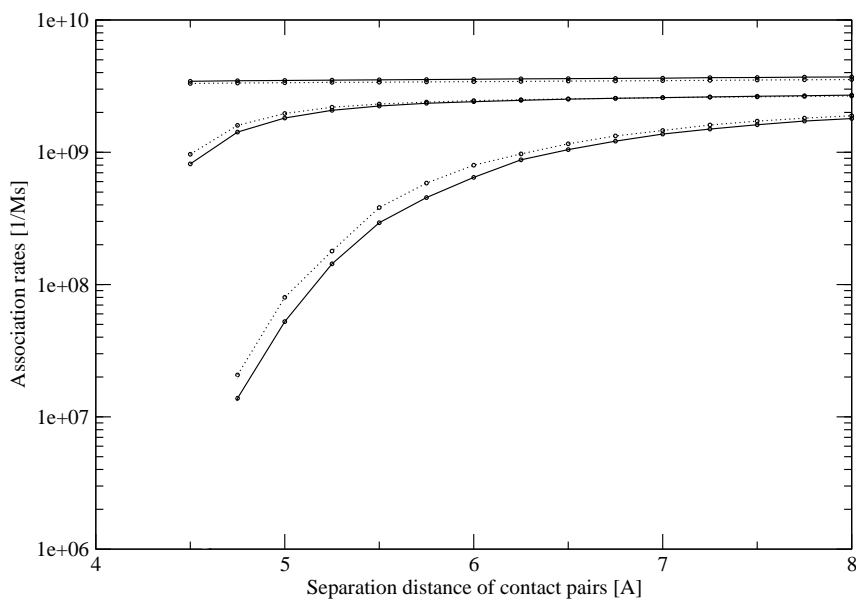
Figure 4.13: *Test simulation of COX and horse heart cytochrome c at 140 mM for original (pointed lines) and special setup of SDA, which was modified to account for the inclusion of the membrane environment.*

horse heart cytochrome c are shown. A direct comparison for three contact pairs for COX within and without the lipid bilayer is shown in Figure 4.15. The computed association rates for the membrane embedded system are two orders of magnitude lower than the computed rates for 'naked' COX. The effect can be attributed to the changed electrostatic potential (see chapter 3). The DPPC bilayer surrounding COX is repulsive for the positively charged lysines which are located close to the interaction site of horse heart cytochrome c. Therefore, the electrostatic steering effect, which is the driving force for fast association, is diminished. Also, close separation distances between the contact pairs do not occur during the BD simulation. The computed rates are now in the range of the experimentally measured data. Still, they cannot be directly compared since the stopped-flow measurements are performed with solubilized enzymes. The graph additionally shows a comparison of the calculated k_{on} rates for the membrane system simulated with UHBD and APBS calculated electrostatic potential

grids.
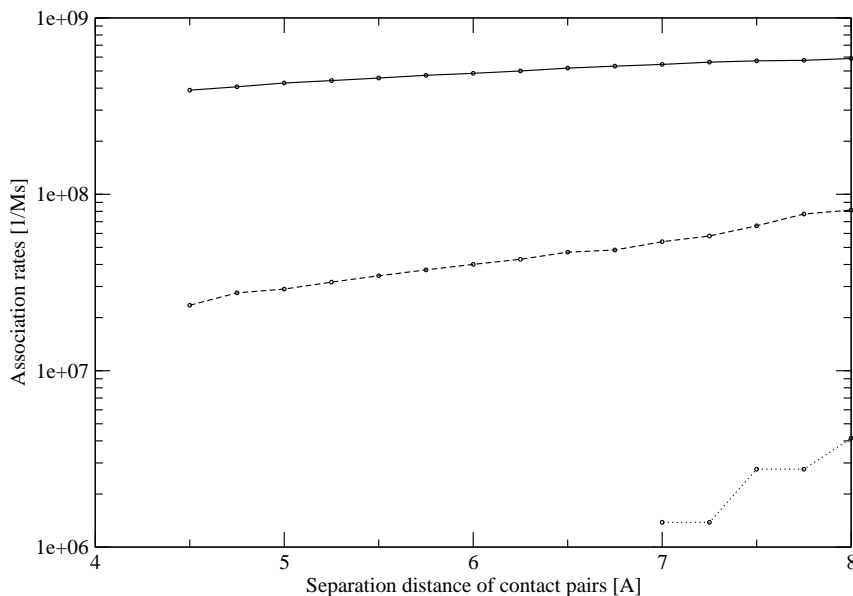
Figure 4.14: *Simulated association rates for membrane embedded COX and horse heart cytochrome c at 140 mM. k_{on} is shown in relationship to the reaction pair distance for one to three residue pairs (from full to pointed line).*

Figure 4.16 shows the calculated k_{on} rates for BD simulations of membrane embedded COX and its physiological electron transfer partner cytochrome c_{552}. A comparison with Figure 4.14 shows that the computed rates for horse heart cyt c and cyt c_{552} are now in the same range. While the association to 'naked' COX was much faster for horse heart cyt c, now the lipid environment favors the association of cyt c_{552}. In Figure 4.17 a comparison between the results of COX without and within the lipid bilayer is presented. In contrast to the previous system of COX and horse heart cytochrome c here k_{on} rates for the membrane embedded system are higher for contact distances up to 11.0 Å. This can be explained by the fact that cytochrome c_{552} has a lower net charge than horse heart cyt c. Unlike horse heart cyt c the physiological electron transfer partner shows a significantly reduced bipolar character. Fewer positively charged residues are close to the interaction site. Because the electrostatic steering effect is much smaller
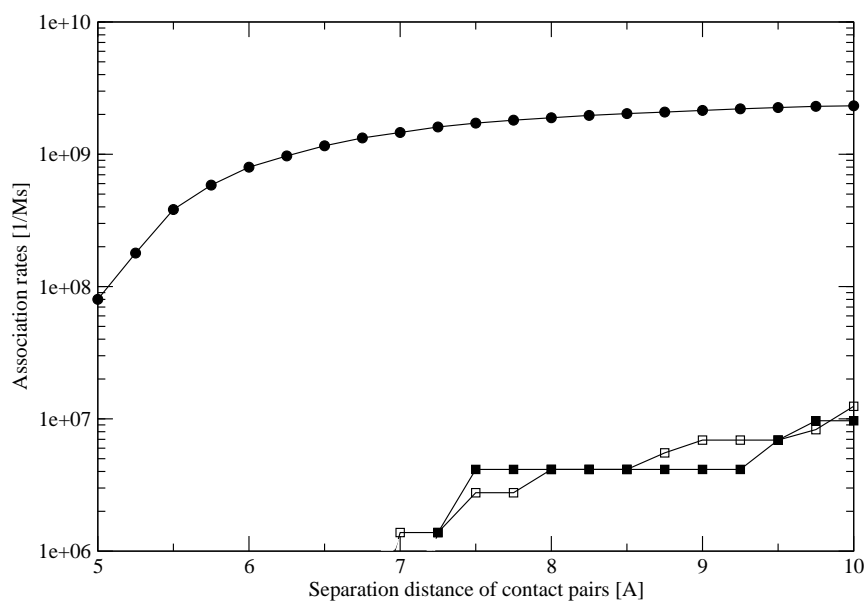
Figure 4.15: *Comparison of simulated association rates for COX without (circles) and within (squares) the membrane environment and horse heart cytochrome c at 140 mM. k_{on} is shown in relationship to the reaction pair distance for three inter-residue pairs. The simulation has been performed with UHBD (filled squares) and APBS (empty squares) calculated electrostatic grids.*

for the diffusing reacting partner, the change of the electrostatic potential introduced by the bilayer does not diminish the association rate. Possibly the effect is even turned around since cyt c_{552} carries a negative net charge. Here, the lipid bilayer surrounding, as it should be suspected in a near-natural biological environment, enhances the biological functionality of the system.
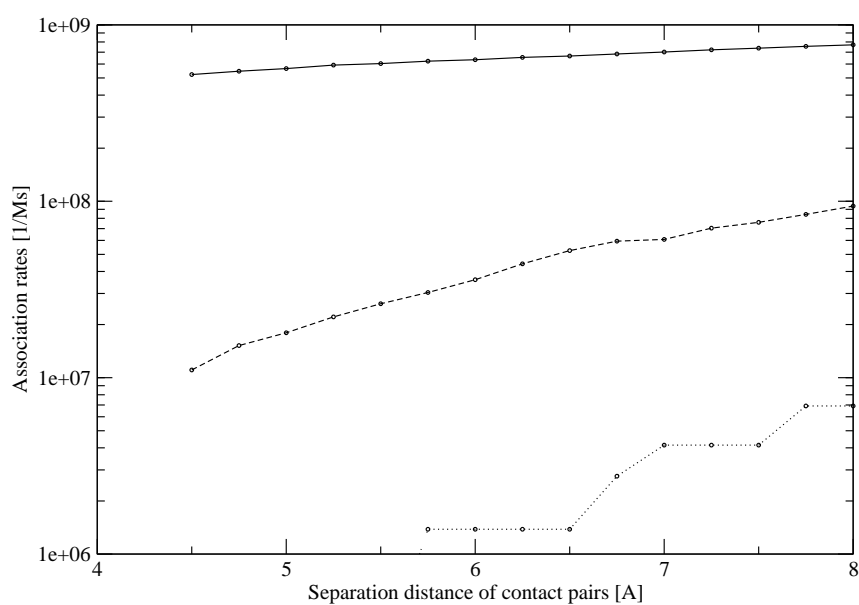
Figure 4.16: *Simulated association rates for membrane embedded COX and its physiological ET partner cytochrome c_{552} at 200 mM. k_{on} is shown in relationship to the reaction pair distance for one to three residue pairs (from full to pointed line).*

Figure 4.17: *Comparison of computed association rates for three contact pairs of membrane embedded COX (squares) and 'naked' COX (circles) with its physiological ET partner cytochrome c_{552}.*

Chapter 5

Conclusions and Outlook

5.1 Docking Studies

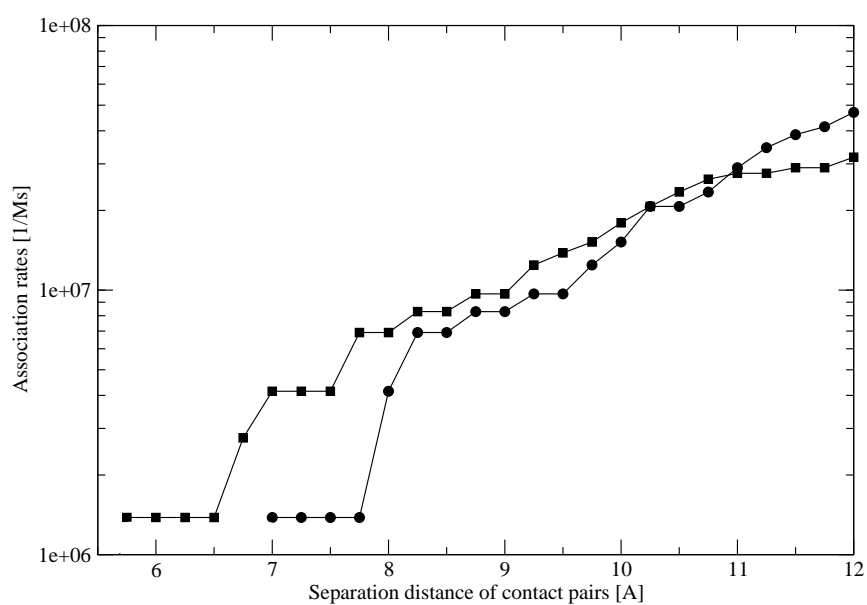A docking strategy was presented to construct realistic complexes of electron transfer proteins with a force field computation of molecular interaction energies. The results for complexes of cytochrome c with cytochrome c peroxidase and with cytochrome c oxidase are quite promising. The complexes were in one case in good agreement with a complex crystal structure (cytochrome c peroxidase) and in the other case with an independent docking study (horse heart cytochrome c with cytochrome c oxidase). We currently await experimental verification by the group of B. Ludwig. Using fluorescence labels on both binding partners to detect fluorescence resonance energy transfer will show whether the two complexes obtained for docking the physiological electron transfer partners in *Paracoccus denitrificans* cytochrome c_{552} and cytochrome c oxidase are equally valid.

In the case of the *Rhodobacter sphaeroides* complex where a well-characterized protein-protein complex was studied, the protein-protein docking techniques and energy calculations were evaluated critically. The docking tool FTDOCK can reliably generate structural models given that biochemically sensible constraints are included. If these are omitted, many 'hits' with favorable surface complementarity and energy score are generated where cytochrome c binds in the membrane region of RC. The sum of CHARMM interaction energy and UHBD solvation energy identifies the crystal structure as one among a number of favorable structures. The concept of 'induced fit'

for protein-protein association describes conformational changes at the binding interface. However, it is currently not clear when these changes occur. Most likely they occur concomitant with the association process.

5.2   Brownian dynamics simulations

Relative association rates for mutated proteins and different ionic strength could be quite accurately predicted with this model, which accounts for electrostatic and steric interactions. The quantitative prediction of real reaction rates still seems difficult to be obtained with such a simplified model. The kinetic on-rates for association to 'naked' cytochrome c oxidase are computed 10-100 times higher than measured in experiment. It is hard to exactly quantify the amount of overestimation because it is unclear where to define the encounter complex. Previous studies showed that one cannot always expect exact matching of computed and experimental rates [116]. However, the discrepancy appears a bit too high in the present case. This poses the question whether calculations and experiments are directly comparable. Experimental k_{on} rates are measured for solubilized membrane proteins. The detergent belt around cytochrome c oxidase molecules could either simply mean excluded volume or may specifically interact with the diffusing cytochrome c particles. Here, it would be helpful to know the exact protein:lipid content to set-up simulations exactly mimicking the experimental conditions.

On the other hand, the overestimation of the computed rates might very well be due to the kinetic behavior of these electron transfer systems. If the first kinetic scheme, which was presented in chapter 1 (see Figure 1.4a) is valid, post-encounter steps of association, which could not be included in the present simulations, could slow down the real association rates significantly.

One of the limitations of the applied method is that the reaction criteria should ideally be defined by the formation of native polar contacts. Since no crystal structures are available for the systems under investigation, the definition had to rely on the results of the computational docking study. Also, the association, which can here be expected to be modulated by interactions occurring after the diffusional encounter, can not be studied by rigid-body Brownian dynamics. For the same reason the inclusion of short-range interactions such as hydrogen bonds and the van der Waals attraction could

improve the simulations.

The comparison of cytochrome c binding to 'naked' cytochrome c oxidase or to membrane-embedded cytochrome c oxidase demonstrates significant differences. The k_{on} rates are 100-1000 times slower depending on the defined reaction distance when the membrane environment is included. Even though the net charge of single DPPC lipids is zero, the calculation of the electrostatic potential of the membrane-embedded protein system in chapter 4 revealed the electrostatic attraction of the membrane for negatively charged particles. Therefore the binding sites of the diffusing cytochrome c with its positively charged lysine side chains are repelled from the membrane surface. Since the binding site of COX is only about 10 Å above the bilayer, this is likely the cause for the inhibition of the association. Also, the contour plots of the electrostatic potential give evidence that the hydrophobic patch of the COX electron entry side, which is surrounded by acidic sidechains, loses its exposition when the enzyme is surrounded by a membrane.

Interestingly, the membrane environment influences the pair association of COX: horse heart cytochrome c and COX:cytochrome c_{552} differently. While the association of the natural electron transfer partner cyt c_{552} is less affected, k_{on} rates of horse heart cyt c are decreased by up to two orders of magnitude. The lipid bilayer surrounding even enhances the association of the physiological pair and is therefore favorable for the biological system, as it should be expected.

5.3   Outlook

By varying the composition of the lipid bilayer the effect of real mitochondrial membranes that are partially negatively charged can be investigated. Also, it would be highly interesting to study the effect of lipid-de-mixing, the non-homogenous membrane composition around integral membrane proteins [124].

There is experimental evidence [125,126] that the enzymes of the respiratory chain and of the photosynthetic units build super-complexes. The computational methods presented in this work are well suited to investigate these nano-systems. Once the super-complexes are computationally docked, the electrostatic properties could be studied, and Brownian dynamics simulations be performed. The composition of these systems could widely affect the electron transfer. Additionally, the diffusional motion

of many cytochromes can be incorporated in a computationally more demanding simulation. An investigation of such multi-particle systems, will deepen the understanding of such coupled systems.

Appendix A

Programs

A.1   3D-Dock

3D-Dock [40, 127] is a package of programs developed in the group of M.J.E. Stern-berg (London) designed to predict protein-protein complexes. The overall algorithm is achieved in up to 4 steps:

- a) ftdock: a global scan of translational and rotational space for best surface complementarity of the two proteins combined with an electrostatic filter

- b) rpscore: an empirical scoring of possible complexes applying a residue level pair potential

- c) filter: screening of possible complexes using biological information

- d) multidock: removal of steric clashes on the side-chains of the interface and energy minimization

A.1.1   FTDock

The program implements the algorithm of Katchalski-Katzir [61]. The two molecules are discretized onto orthogonal grids before translational and rotational space is scanned globally. The scoring is based on surface complementarity between the two grids. To speed up the calculations Fourier Transforms are used. The convolutions are replaced by multiplications in Fourier space and the number of calculations can be significantly

reduced, if the grid dimensions are integer powers of two. Specifically, direct calculation of the correlation function requires N^6 calculations, whereas Fast Fourier Transform only requires $N^3 \ln(N^3)$ calculations for every rotational angle. Additional to the surface complementarity score an electrostatic binary filter can be included, which removes complexes with unfavorable electrostatics from the stack.

The biological filter screens the possible complexes, requiring them to have a certain chain or residue of one molecule within a certain chain or residue of the other molecule, therefore forcing them to be part of the interface.

A.1.2 RPScore

Each possible complex is scored with an empirical pair potential matrix. Each potential corresponds to the likelihood of two residue types to be part of the interface (within a certain distance cut-off). The matrix is generated from 90 non-homologous interfaces found in the PDB with the aid of SCOP 1.53 (http://scop.mrc-lmb.cam.ac.uk/scop/).

A.1.3 Multidock

The program is designed to refine the interface between two proteins at the atomic level. The protein is described by electrostatic and van der Waals interactions. Multidock performs side-chain energy minimizations and removes steric clashes along the interface. The sidechains are modeled by a multiple copy representation according to a rotamer library. The refinement is achieved with a two step process:

- A probability based conformational matrix is refined iteratively. A given sidechain interacts with the fixed backbone as well as with the probability weighted average of the surrounding sidechains.

- The protein interface is relaxed by rigid-body energy minimization of protein backbone atoms and the highest probability sidechain conformations.

For further information see: http://www.bmm.icnet.ik/docking/.

A.2 CHARMM - Chemistry at HARvard Macromolecular Mechanics

CHARMM [54] is developed by the group of M. Karplus (Harvard, Strasbourg). It uses empirical energy functions to model macromolecular systems. The program can read structures, energy minimize them, perform normal mode analysis or molecular dynamics simulations. Atoms are represented as charged point masses without internal degrees of freedom. Hydrogens can be combined with neighboring heavy atoms to which they are bound, or be represented individually.

The empirical CHARMM energy function is based on internal coordinates and pairwise non-bond interaction terms:

$$E = \underbrace{E_b + E_\theta + E_\Phi + E_\omega}_{Internal\ energies} + \underbrace{E_{vdW} + E_{el}}_{Nb\ interactions} + E_{hb} + E_{c_r} + E_{c_\Phi}. \tag{A.1}$$

- Internal energy terms

  The bond potential $E_b = \sum k_b(r - r_0)^2$ and the bond angle potential $E_\theta = \sum k_\theta(\theta - \theta_0)^2$ account for bond and angle deformations, which are in most cases small enough for the harmonic approximation to apply.

  The torsion or dihedral angle term $E - \Phi = \sum |k_\Phi| - k_\Phi \cos(n\Phi)$, where $n = 1, 2, 3, 4, 6$, is a four-atom term based on the dihedral angle about an axis defined by the middle pair of atoms.

  The improper torsion term $E_\omega = \sum k_\omega(\omega - \omega_0)^2$ maintains both planarity about certain planar atoms and chirality about a tetrahedral extended heavy atom.

  Parameter tables based on the atom types of the atoms involved are read into CHARMM and provide force constants k_x and geometric constants r_0, θ_0, n and ω_0. Geometric constants are mostly derived from crystallographic data. Force constants can be obtained by fitting to vibrational data. In this work the parameter set CHARMM22 developed by MacKerell *et al.* [62] has been used.

- Non-bonded (Nb) interactions

  For typical proteins the non-bonded interactions represent the bulk of the energy evaluation time. Therefore the van der Waals term $E_{vdW} = \sum \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)$ is switched off after some distance (usually 7.5 Å) for larger systems.

To account for the electrostatic potential E_{el} one of the terms listed below can be chosen:

- Constant dielectric:

$$E_{el} = \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{A.2}$$

- Distance-dependent dielectric:

$$E_{el} = \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}^2} \, sw(r_{ij}^2, r_{on}^2, r_{off}^2), \tag{A.3}$$

  with sw switching function.

- Shifted dielectric:

$$E_{el} = \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \left(1 - \frac{2r_{ij}^2}{r_{cut}^2} + \frac{r_{ij}^4}{r_{cut}^4} \right) \tag{A.4}$$

- Hydrogen bonding:

$$
\begin{aligned}
E_{hb} &= \sum \left(\frac{A'}{r_{AD}^i} - \frac{B'}{r_{AD}^j} \right) \cos^m(\theta_{A-H-D}) \\
&\times \cos^n(\theta_{AA-A-H}) \times sw\left(r_{AD}^2, r_{h_{on}}^2, r_{h_{off}}^2 \right) \\
&\times sw\left[\cos^2(\theta_{A-H-D}), cos^2(\theta_{h_{on}}), cos^2(\theta_{h_{off}}) \right],
\end{aligned}
\tag{A.5}
$$

where i and j are positive integers, $m = (0, 2, 4)$ and $n = (0, 2)$. Involved atoms are the acceptor antecedent (AA), the acceptor (A), the hydrogen (H) and the donor (D). The hydrogen bonding term is based on the A-D distance and the A-H-D angle. The exponents n and m are determined by the type of the acceptor and the type of donor, respectively. The values for all parameters are extracted from the topology and parameter files.

This energy term is not usually used anymore in minimization or dynamics because modern parameter sets compute hydrogen bond contributions as a balance between electrostatic attraction and van der Waals repulsion terms.

- Water-water interactions:
  In the 1980s the ST2 potential [128] was usually used in CHARMM applications. Nowadays the TIP3P model [129] is widely used.

- Constraints:

  Different types of constraints can be included to restrain changes that occur in a structure when it is manipulated by minimization or dynamics. Rigid distance constraints ($\delta_{\vec{r}_{ij}} = 0$) are enforced by the so-called SHAKE algorithm [130]. Atom harmonic constraints ($E_{c_r} = \sum K_i(\vec{r}_i - \vec{r}_{i_0})^2$) are used to avoid large displacements of atoms when minimizing. Dihedral constraints $E_{c_\Phi} = \sum K_i(\Phi_i - \Phi_{i_0})^2$ are used to maintain certain local conformations.

A.3   UHBD - University of Houston Brownian Dynamics program

UHBD [69,131] has been developed in the group of J.A. McCammon (San Diego). The program is designed to perform electrostatic calculations and the simulation of diffusional motion of molecules in solution. As additional features an energy minimization routine and the calculation of protein pKa are implemented. Theory and algorithms of the routines designed for electrostatic calculations have been reported in detail in chapter 3.

For further information see : http://mccammon.ucsd.edu/uhbd.html

A.4   APBS - Adaptive Poisson-Boltzmann Solver

APBS [100] was mainly written by N. Baker during his graduate work with J.A. McCammon and M. Holst. Compared with other Poisson-Boltzmann solvers (like UHBD) it enables the parallel solution for supramolecular systems that are orders of magnitude larger in size. Combining the Bank-Holst algorithm [101], an adaptive finite element technique, with other techniques the algorithm was extended to a finite difference solver. Electrostatic focusing techniques calculate a low-accuracy solution on a coarse finite difference mesh spanning the entire problem domain. This coarse solution is then used to define the boundary conditions on a finer discretization of the desired sub-domain. Here standard focusing techniques and the Bank-Holst algorithm have been combined into a parallel focusing method for the solution of the Poisson-Boltzmann equation. The excellent parallel complexity permits the treatment of very large biomolecular sys-

tems (e.g. a million atom microtubule fragment) on parallel computational platforms. For further information see: http://agave.wustl.edu/apbs/

A.5   ECM - Effective (potential derived) Charges for Macromolecules in solvent

ECM [97] has been developed in the group of R.C. Wade (Heidelberg). The program was designed to enable the calculation of electrostatic forces and interaction free energies for two macromolecules at different mutual separations and orientations. Accurate evaluations of the electrostatic forces of these systems are only feasible for a limited number of positions. Instead of solving the PB equation for every single position, the molecules are represented by a small number of effective charges in a uniform dielectric. In this manner the electrostatic interactions can be calculated fast but still with high accuracy. The effective charges are derived by fitting them to reproduce the molecules' electrostatic potential in a heterogeneous dielectric. Thus, the PB equation has to be solved only once for each biomolecule. The fitting is performed over a layer of reasonable thickness.
For further information see: http://www.embl-heidelberg.de/~gabdoull/ecm/ecm.html

A.6   SDA - Simulation of Diffusional Association of proteins

SDA [111] was mainly developed by R.R. Gabdoulline in cooperation with R.C. Wade (Heidelberg). The program was designed to simulate the diffusional association of proteins. Theory and algorithm are explained in detail in chapter 4. SDA computes the rates of diffusional association of two proteins given the atomic structure of the bound complex of the proteins. It uses electrostatic potentials computed by UHBD or APBS and ECM generated effective charges of the biomolecules to account for the electrostatic systematic force in the Brownian motion.
For further information see: http://www.embl-heidelberg.de/~gabdoull/sda/sda.html

Bibliography

[1] Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 1958;181:662-666.

[2] Drewes G, Bouwmeester T. Global approaches to protein-protein interactions. Curr Opin Cell Biol 2003;15:199-205.

[3] Tiede DM, Vashista AC, Gunner MR. Electron-Transfer kinetics and electrostatic properties of the *Rhodobacter sphaeroides* reaction center and soluble *c*-cytochromes. Biochemistry 1993;32:4515-4531.

[4] Drosou V, Malatesta F, Ludwig B. Mutations in the docking site for cytochrome *c* on the *Paracoccus* heme aa_3 oxidase: electron entry and kinetic phases of the reaction. Eur J Biochem 2002;269:2980-2988.

[5] Witt H, Malatesta F, Nicoletti F, Brunori M, Ludwig B. Cytochrome-c-binding site on cytochrome oxidase in Paracoccus denitrificans. Eur J Biochem 1998;251:367-373.

[6] Witt H, Malatesta F, Nicoletti F, Brunori M, Ludwig B. Tryptophan 121 of subunit II is the electron entry site to cytochrome-c oxidase in Paracoccus denitrificans. J Biol Chem 1998;273: 5132-5136.

[7] Tetreault M, Rongey SH, Feher G, Okamura MY. Interactions between cytochrome c(2) and the photosynthetic reaction center from Rhodobacter sphaeroides: Effects of charge-modifying mutations on binding and electron transfer. Biochemistry 2001;40:8452-8462.

[8] Tiede DM. c-Cytochrome Orientation in Electron Transfer Complexes Formed with Photosynthetic Reaction Centers of Rhodopseudomonas sphaeroides, and When Bound to the Surface of Negatively Charged Membranes: Characterization by Optical Linear Dichroism. Biochemistry 1987;26:397-410.

[9] Drepper F, Mathis P. Structure and function of cytochrome c(2) in electron transfer complexes with the photosynthetic reaction center of Rhodobacter sphaeroides: Optical linear dichroism and EPR. Biochemistry 1997;36:1428-1440.

[10] Adir N, Axelrod HL, Beroza P, Isaacson RA, Rongey SH, Okamura MY, Feher G. Co-crystallization and characterization of the photosynthetic reaction center-cytochrome c(2) complex from Rhodobacter sphaeroides. Biochemistry 1995;35:2535-2547.

[11] Roberts VA, Pique ME. Definition of the interaction domain for cytochrome c on cytochrome c oxidase. J Biol Chem 1999;274:38051-38060.

[12] Hu X, Ritz T, Damjanović, Autenrieth F, Schulten K. Photosynthetic apparatus of purple bacteria. Q Rev Biophys 2002;35:1-62.

[13] Dutton PL, Petty KM, Bonner HS, Morse SD. [Abstract] Cytochrome c2 and reaction center of Rhodospeudomonas sphaeroides Ga. membranes. Extinction coefficients, content, half-reduction potentials, kinetics and electric field alterations. BBA 1975;387:536-556.

[14] Moser CC, Dutton PL. Cytochrome c and c2 binding dynamics and electron transfer with photosynthetic reaction center protein and other integral membrane redox proteins. Biochemistry 1988;27:2450-2461.

[15] Overfield RE, Wraight CA, Devault D. Microsecond photo-oxidation kinetics of cytochrome-c2 from Rhodopseudomonas-Sphaeroides - Invivo and solution studies. FEBS Lett 1979;105:137-142.

[16] Rosen D, Okamura MY, Abresch EC, Valkirs GE, Feher G. Interaction of cytochrome-c with reaction centers of from Rhodopseudomonas-Sphaeroides R-26 - Localization of the binding-site by chemical cross-linking and immunochemical studies. Biochemistry 1983;22:335-341.

[17] Selzer T, Schreiber G. New insight into the mechanism of protein-protein association. Proteins 2001;45:190-198.

[18] Selzer T, Albeck S, Schreiber G. Rational design of faster associating and tighter binding protein complexes. Nat Struct Biol 2000;7:537-541.

[19] Spoerner M, Herrmann C, Vetter IR, Kalbitzer HR, Wittinghofer A. Dynamic properties of the Ras switch I region and its importance for binding to effectors. Proc Natl Acad Sci USA 2001;98:4944-4949.

[20] Koppenol WH, Margoliash E. The asymmetric distribution of charges on the surface of horse cytochrome c. J Biol Chem 1982;257:4426-4437.

[21] Gabdoulline RR, Wade RC. On the protein-protein diffusional encounter complex. J Mol Recognit 1999;28:153-161.

[22] Janin J. The kinetics of protein-protein recognition. Proteins 1997;28:153-161.

[23] Selzer T, Schreiber G. Predicting the rate enhancement of protein complex formation from the electrostatic energy of interaction. J Mol Biol 1999;287:153-161.

[24] Smoluchowski MV. Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen. Z Phys Chem 1918;92:129.

[25] Berg OG, von Hippel PH. Diffusion-controlled macromolecular interactions. Annu Rev Biophys Chem 1985;14:131-160.

[26] Northrup SH, Erickson HP. Kinetics of protein-protein association explained by Brownian dynamic computer simulation. Proc Natl Acad Sci USA 1992;89:3338-3342.

[27] Camacho CJ, Weng Z, Vajda S, DeLisi C. Free energy landscapes of encounter complexes in protein-protein association. Biophys J 1999;76:1166-1178.

[28] Sheinerman FB, Norel R, Honig B. Electrostatic aspects of protein protein interactions. Curr Opin Struc Biol 2000;10:153-159.

[29] Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 1940;7:284-304.

[30] Vijayakumar M, Wong KY, Schreiber G, Fersht AR, Szabo A, Zhou HZ. Electrostatic enhancement of diffusion-controlled protein-protein association: Comparison of theory and experiment on barnase and bastar. J Mol Biol 1998;278:1015-1024.

[31] Elcock AH, Sept D, McCammon JA. Computer Simulation of Protein-Protein Interactions. J Phys Chem B 2001;105:1504-1518.

[32] Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanovivhi T, Tasumi M. The protein data bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535-542.

[33] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Research 2000;28:235-242.

[34] Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A Geometric Approach to Macromolecule-Ligand Interactions. J Mol Biol 1982;161:269-288.

[35] Böhm HJ. LUDI - Rule-Based Automatic Design of New Substituents for Enzyme-Inhibitor Leads. J Comp Aid Mol Des 1992;1:301-319.

[36] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart,WE, Belew RK, Olson AJ. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. J Comp Chem 1998;19:1639-1662.

[37] Taylor JS, Burnett RM. DARWIN: A program for Docking Flexible Molecules. Proteins 2000;41:173-191.

[38] Janin J. Protein-protein recognition. Prog Biophys Mol Biol 1995;64:145-166.

[39] Walls PH, Sternberg MJE. New algorithm to model protein-protein recognition based on surface complementarity. J Mol Biol 1992;228:277-297.

[40] Gabb HA, Jackson RM, Sternberg JE. Modeling protein docking using shape complementarity, electrostatics and biochemical information. J Mol Biol 1997;272:106-120.

[41] Ten Eyck LF, Mandell J, Roberts VA , Pique ME . Proceedings of the 1995 ACM/IEEE Supercomputing Conference .In: Hayes A, Simmons M, editors. Los Alamitos: IEEE Computer Society Press; 1995.

[42] Halperin I, Ma B, Wolfson H and Nussinov R. Principles of Docking: An Overview of Search Algorithms and a Guide to Scoring Functions. Proteins 2002;47:409-443.

[43] Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. Science 1983;21709-713.

[44] Connolly M. Analytical molecular surface calculation. J Appl Cryst 1983;16:548-558.

[45] Mathews FS, Mauk AG, Moore GR. Protein-protein complexes formed by electron transfer proteins. In: Kleanthous C, editor. Protein-Protein Recognition. New York: Oxford University Press Inc; 2000. p. 60-101.

[46] Ostermeier C, Harrenga A, Ermler U, Michel H. Structure at 2.7 angstrom resolution of the Paracoccus denitrificans two-subunit cytochrome c oxidase complexed with an antibody F-V fragment. Proc Natl Acad Sci USA 1997;94:10547-10553.

[47] Harrenga A, Michel H. The cytochrome c oxidase from Paracoccus denitrificans does not change the metal center ligation upon reduction. J Biol Chem 1999;274, 33296-33299.

[48] Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, ShinzawaItoh K, Nakashima R, Yaono R, Yoshikawa S. Science 1996;272: 1136-1144.

[49] Margoliash E, Bosshard HR. Guided by electrostatics, a textbook protein comes of age. TIBS - Sept 1983:316-320.

[50] Reincke B, Thöny-Meyer L, Dannehl C, Odenwald A, Aidim M, Witt H, Rüterjans H, Ludwig B. Heterologous expression of soluble fragments of cytochrome c552 as electron donor to the Paracoccus denitrificans cytochrome c oxidase. BBA 1999;1411:114-120.

[51] Turba A, Jetzek M, Ludwig B. Purification of Paracoccus denitrificans Cytochrome c552 and sequence analysis of the gene. Eur J Biochem 1995;231:259-265.

[52] Harrenga A, Reincke B, Ruterjans H, Ludwig B, Michel H. Structure of the soluble domain c(552) from Paracoccus denitrificans in the oxidized and reduced states. J Mol Biol 2000;295:667-678.

[53] Pristovsek P, Lücke C, Reincke B, Ludwig B, Rüterjans H. Solution structure of the functional domain of Paracoccus denitrificans cytochrome c552 in the reduced state . Eur J Biochem 2000; 267:4205-4212.

[54] Brooks BR, Roberts EB, Barry DO, David JS, Swaminathan S, Karplus M. Charmm: A program for macromolecular energy, minimization, and dynamics calculations. J Comp Chem 1983; 4:187-217.

[55] Pelletier H, Kraut J. Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c. Science 1992;258:1748-1755.

[56] Beinert H. Copper A of cytochrome c oxidase, a novel long-embattled, biological electron-transfer site. Eur J Biochem 1997;245:521-532.

[57] Kannt A, Lancaster CR, Michel H. The coupling of electron transfer and proton translocation: Electrostatic calculations on *Paracoccus denitrificans* cytochrome *c* oxidase. Biophys J 1998;74:708-721.

[58] Bushnell GW, Louie GV, Brayer, GD. High-resolution three-dimensional structure of horse heart cytochrome c. J Mol Biol 1990;214:585-595.

[59] Ermler U, Fritzsch G, Buchanan SK, Michel H: Structure of the photosynthetic reaction center from Rhodobacter sphaeroides at 2.65 A resolution:cofactors and protein-cofactor interactions. Structure 1994;2:925-936.

[60] Axelrod HL, Abresch EC, Okamura MY, Yeh AP, Rees DC, Feher G: X-Ray Structure Determination of the Cytochrome-C(2):Reaction Center Electron Transfer Complex from Rhodobacter Sphaeroides. J Mol Biol 2002; 319:501-515.

[61] Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 1992;89:2195-2199.

[62] MacKerell Jr AD, Bashford D, Bellott M, Dunbrack Jr RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera, K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher III WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem 1998;102:3586-3616.

[63] Brünger AT, Karplus M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. Proteins 1988;4,148-156.

[64] McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature 1977;267:585-590.

[65] Jackson RM, Gabb HA and Sternberg MJE. Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J Mol Biol 1998;276:265-285.

[66] Helms V, McCammon JA. Kinase conformations: A computational study of the effect of ligand binding. Prot Sci 1997;6, 2336-2343.

[67] Hünenberger PH, Helms V, Narayana N, Taylor SS, Mc Cammon JA. Determinants of ligand binding to cAMP-dependent protein kinase. Biochemistry 1999;38, 2358-2366.

[68] Chandran V, Flöck D, Axelrod H, Helms V. Critical Assessment of Protein-Protein Docking: The Complex of Reaction Center from *Rhodobacter sphaeroides* with cytochrome c_2. Submitted to J Mol Recognit.

[69] Davis ME, Madura JD, Luty BA, McCammon JA. Electrostatic and Diffusion of Molecules in Solution - Simulations with the University-of-Houston-Brownian Dynamics Program. Comp Phys Com 1991;62:187-197.

[70] Insight II User Guide, version 2.1.0, San Diego: Biosym Technologies, 1992.

[71] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33-38, 27-28.

[72] Zhang C, Chen J, DeLisi C. Protein-protein recognition: Exploring the energy tunnels near the binding sites. Proteins 1999;34:255-267.

[73] Northrup SH, Boles JO, Reynolds JC. Brownian dynamics of cytochrome c and cytochrome c peroxidase association. Science 1988;241:67-70.

[74] Moore GR, Cox MC, Crowe D, Osborne MJ, Rosell FI, Bujons J, Barker PD, Mauk MR, Mauk AG. N-epsilon, N-epsilon-dimethyl-lysine cytochrome c as an NMR probe for lysine involvement in protein-protein complex formation. Biochem J 1998;332: 439-449.

[75] Ullmann GM, Knapp EW, Kostic NM. Computational Simulation and Analysis of Dynamic Association between Plastocyanin and Cytochrome f. Consequences for the Electron-Transfer Reaction. J Am Chem Soc 1997;119:42-52.

[76] Fogolari F, Brigo A, Molinari H. The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. J Mol Recognit 2002;15:377-392.

[77] Mirny L, Shakhnovich E. Protein folding theory: from lattice to all-atom models. Annu Rev Biophys Biomol Struct 2001;30:361-396.

[78] Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE. Crystal structure analysis of a complete turn of B-DNA. Nature 1980;287:755-758.

[79] Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999;285:2177-2198.

[80] McCoy AJ, Chandana Epa V, Colman PM. Electrostatic complementarity at protein/protein interfaces. J Mol Biol 1997;268:570-584.

[81] Schultz PG. The interplay between chemistry and biology in the design by enzymatic catalysts. Science 1988;240:426-433.

[82] Davis ME, McCammon JA. Electrostatics in biomolecular structure and dynamics. Chem Rev 1990;90:509-521.

[83] Honig B, Nicolls A. Classical electrostatics in biology and chemistry. Science 1995;268:1144-1149.

[84] Warshel A, Papazyan A. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. Curr Opin Struct Biol 1998;8:211-217.

[85] Simonson T. Macromolecular electrostatics: continuum models and their growing pains. Curr Opin Struct Biol 2001;11:243-252.

[86] Jackson JD. Classical Electrodynamics. John Wiley: New York 1962.

[87] Fogolari F, Zuccato P, Esposito G, Viglino P. Biomolecular electrostatics with the linearized Poisson-Boltzmann equation. Biophys J 1999;76:1-16.

[88] Kirkwood JG. Theory of solutions of molecules containing widely separated charges with special applications to zwitterions. J Chem Phys 1934;7:351-361.

[89] Nozaki Y, Tanford C. Examination of titration behavior. Meth Enz 1967;11:715-734.

[90] Lifson S, Katchalski A. The electrostatic free energy of polyelectrolyte solutions II. Fully stretched macromolecules. J Polym Sci 1954;13:43-55.

[91] Manning GS. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. Q Rev Biophys 1978;11:179-246.

[92] Gouy M. Sur la constitution de la charge électrique a la surface d'un électrolyte. J Phys 1910:9:457-468.

[93] Chapman DL. A contribution to the theory of electrocapillarity. Phil Mag 1913;25:475-481.

[94] Warwicker J, Watson HC. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. J Mol Biol 1982;157:671-679.

[95] Nicholls A, Honig B. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J Comput Chem 1991;12:435-445.

[96] Davis ME, Madura JA. Dielectric boundary smoothing in finite difference solutions of the Poisson equation: an approach to improve accuracy and convergence. J Comput Chem 1991;12:909-912.

[97] Gabdoulline RR, Wade RC. Effective charges for macromolecules in solvent. J Phys Chem 1996;100:3868-3878.

[98] Leach AR. Molecular Modeling, Principles and Applications. Addison Wesley Longman Limited, Essex, 1996.

[99] Madura JD, Briggs JM, Wade R, Luty BA, Ilin A, Antosiewicz J, Gilson MK, Bagheri B, Ridgway Scott L, McCammon JA. Biological applications of electrostatics calculations and Brownian dynamics simulations. Rev Comp Chem 1994;5:229-267.

[100] Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 2001;98:10037-10041.

[101] Bank R, Holst M. A new paradigm for parallel adaptive meshing algorithms. Soc Ind Appl Math J Sci Comput 2000;22:1411-1443.

[102] Luty BA, Wade RC, Madura JD, Davis ME, Briggs JM, McCammon JA. Brownian Dynamics Simulations of Diffusional Encounters between Triose Phosphate Isomerase and Glyceraldehyde Phosphate: Electrostatic Steering of Glyceraldehyde Phosphate. J Phys Chem 1993;97:233-237.

[103] Kozack RE, d'Mello MJ, Subramaniam S. Computer modeling of electrostatic steering and orientational effects in antibody-antigen association. Biophys J 1995;68:807-814.

[104] Woolf TB, Roux B. Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer. Proc Natl Acad Sci USA 1994;91:11631-11635.

[105] Woolf TB, Roux B. Structure, energetics, and dynamics of lipid-protein inter-
      actions: a molecular dynamics study of the gramicidin A channel in a DMPC
      bilayer. Proteins 1996;24:92-114.

[106] Gennis RB. Biomembranes Molecular Structure and Function. New York:
      Springer-Verlag; 1989.

[107] Nagle JF. Area/lipid of bilayers from NMR. Biophys J 1993;64:1476-1481.

[108] Venable RM, Zhang Y, Hardy BJ and Pastor RW. Molecular dynamics simula-
      tions of a lipid bilayer and of hexadecane: an investigation of membrane fluidity.
      Science 1993;262:223-226.

[109] Smoluchowski MV. Drei Vorträge über Diffusion, Brownsche Bewegung und
      Koagulation von Kolloidteilchen. Phys Z 1916;17:557-585.

[110] Northrup SH, Allison SA, McCammon JA. Brownian dynamics simulation of
      diffusion-influenced bimolecular reactions. J Chem Phys 1983;80:1517-1524.

[111] Gabdoulline RR, Wade RC. Simulation of diffusional association of barnase and
      bastar. Biophys J 1997;72:1917-1929.

[112] Ermak DL, McCammon JA. Brownian dynamics with hydrodynamic interactions.
      J Chem Phys 1978;69:1352-1360.

[113] Antosiewicz J, McCammon JA. Electrostatic and hydrodynamic orientational
      steering effects in enzyme-substrate association. Biophys J 1995;69:57-65.

[114] Antosiewicz J, Briggs JM, McCammon JA. Orientational Steering in enzyme-
      substrate association: ionic strength dependence of hydrodynamic torque effects.
      Eur Biophys J 1996;24:137-141.

[115] Gabdoulline RR, Wade RC. Brownian dynamics simulation of protein-protein
      diffusional encounter. Methods 1998;14:329-341.

[116] Gabdoulline RR, Wade RC. Protein-Protein Association: Investigation of fac-
      tors influencing association rates by Brownian Dynamics Simulation. J Mol Biol
      2001;306:1139-1155.

[117] Maneg O, Ludwig B, Malatesta R. Different interaction modes of two cytochrome *c* oxidase soluble CuA fragments with their subtrates. Submitted to JBC.

[118] Northrup SH, Boles JO, Reynolds JCL. Electrostatic effects in the Brownian dynamics of association and orientation of heme proteins. J Phys Chem 1987;91:5991-5998.

[119] Flöck D and Helms V. Protein-Protein Docking of Electron Transfer Complexes: Cytochrome c Oxidase and Cytochrome c;Proteins 2002;47:75-85.

[120] Northrup SH, Thomasson KA, Miller MM. Effects of Charged Amino Acid Mutations on the Bimolecular Kinetics of Reduction of Yeast Iso-1-ferricytochrome c by Bovine Ferrocytochrome b5. Biochemistry 1993;32:6613-6623.

[121] Castro G, Boswell CA, Northrup SH. Dynamics of protein-protein docking: cytochrome c and cytochrome c peroxidase revisited. J Biomol Struct Dynam 1998;16:413-424.

[122] Worrall, JAR, Kolczak U, Canters GW, Ubbink M. Interaction of Yeast Iso-1-cytochrome c with Cytochrome c Peroxidase Investigated by [15N, 1H] heteronuclear NMR Spectroscopy. Biochemistry 2001;40:7069-7076.

[123] Eltis LD, Herbert RE, Barker PD, Mauk AG, Northrup SH. Reduction of horse heart ferricytochrome c by bovine liver ferrocytochrome b5. Biochemistry 1991;30:3663-3674.

[124] Heimburg T, Angerstein B, and Marsh D. Binding of peripheral proteins to mixed lipid membranes: The effect of local demixing upon binding. Biophys J 1999;76:2575-2586.

[125] Schägger H, Pfeiffer K. The ratio of oxidative phosphorylation complexes I-V in bovine heart mitochondria and the composition of respiratory chain supercomplexes. J Biol Chem 2001;276:37861-37867.

[126] Schägger H. Respiratory chain supercomplexes of mitochondria and bacteria. BBA 2002;1555:154-159.

[127] Moont G, Gabb HA, Sternberg MJE. Use of pair potentials across protein interfaces in screening predicted docked complexes. Proteins 1999;35:364-373.

[128] Stillinger FH, Rahman A. Improved simulation of liquid water by molecular dynamics. J Chem Phys 1974;60:1545-1557.

[129] Jorgensen WL, Chandrasekhar J, Madura JD, R. Impey W, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 1983;79:926-935.

[130] van Gunsteren WF, Berendsen HJC. Algorithms for macromolecular dynamics and constraint dynamics. Mol Phys 1977;34:1311-1327.

[131] Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosicwicz J, Gilson MK, Bagheri B, Scott LR, McCammon JA. Electrostatics and diffusion in solution: Simulations with the University of Houston Brownian Dynamics Program. Comp Phys Comm 1995;91:57-95.

Danksagung

Mein ganz besonderer Dank gilt Prof. V. Helms für die freundliche Aufnahme an das Max-Planck-Institut für Biophysik, die intensive Betreuung dieser Arbeit und die Bereitschaft, auftretende Fragen und Probleme jederzeit zu diskutieren.

Ganz herzlich danke ich auch Prof. W. Mäntele für die Betreuung der Arbeit an der Johann-Wolfgang-Goethe-Universität Frankfurt, sein Interesse und seine Unterstützung.

Vielen Dank an Dr. V. Chandran und E. Olkhova für die hervorragende Zusammenarbeit. Auch danke ich Dr. M. Lill und Dr. M. Hutter, die mir besonders bei der Einarbeitung wertvolle Hilfestellung leisteten. C. Gorba danke ich für seine Diskussionsbereitschaft, seinen gern gewährten Rat und die Hilfe bei allen Computerfragen. Herzlichen Dank auch allen anderen Mitgliedern und ehemaligen Mitgliedern der Gruppe Helms für Teamgeist und die stimulierende Arbeitsatmosphäre.

Weiterhin danke ich Prof. B. Ludwig und seiner Arbeitsgruppe, Dr. A. Harrenga und Dr. H.L. Axelrod, die mir teilweise noch unveröffentlichte Messdaten zur Verfügung stellten. Herzlichen Dank auch an Dr. R. Gabdoulline, Dr. R. Wade und Dr. N. Baker für ihren begleitenden Rat hinsichtlich spezieller Softwarefragen.

Abschliessend danke ich noch allen, die das Manuskript korrekturgelesen haben, sowie meinen Freunden und meinen Eltern, die mich während dieser Arbeit begleitet und ermutigt haben.

Lebenslauf

02.04.1971	Geboren in Neuwied (Rhein)
1981 - 1987	Görres-Gymnasium Koblenz
1987 - 1988	Ausable Valley Highschool, NY, USA
1988 - 1991	Görres-Gymnasium Koblenz
SS 1992 - WS 1998/99	Diplom in Physik, Freie Universität Berlin
Apr. 1997 - Okt. 1998	Diplomarbeit, Fritz-Haber-Institut der MPG, Berlin Prof. Dr. Gerhard Ertl Thema der Diplomarbeit: Modellierung der Muster- bildung bei der CO-Oxidation an Platin(111): Einfluß von Subsurface-Sauerstoff
seit Aug. 1999	Doktorandin am Max-Planck-Institut für Biophysik, Frankfurt Arbeitsgruppe für theoretische Biophysik, Prof. Dr. Volkhard Helms Betreuer an der Johann-Wolfgang-Goethe Universität: Prof. Dr. Werner Mäntele