
Influence of Electrostatics on
Protein-Protein Interactions of
Photosynthetic Proteins

Dissertation

der Universität Bayreuth
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
im Promotionsprogramm Fotophysik synthetischer und
biologischer multichromophorer Systeme

genehmigte Abhandlung von
Johannes Matthias Förster
geboren in Eschenbach i.d.Opf.

2018

Die vorliegende Arbeit wurde am Lehrstuhl Biopolymere an der
Universität Bayreuth unter der Betreuung von Prof. Dr. G. Matthias
Ullmann angefertigt.

Amtierender Direktor der Graduiertenschule: Prof. Dr. Dirk Schüler

Datum der Abgabe: 08.01.2018

Prüfungsausschuss:

Prof. Dr. G. Matthias Ullmann (Erstgutachter)

Prof. Dr. Birte Höcker (Zweitgutachter)

Prof. Dr. Stephan Kümmel (Vorsitzender)

Prof. Dr. Stephan Gekle

Tag des Kolloquiums: 01.02.2018

Abstract

Protein-protein interactions play a central role in many cellular processes, such as signal transduction, gene regulation and molecular bioenergetics. A wide variety of complexes is needed to fulfill the entity of interaction types ranging from strong binding complexes to weakly interacting ones like transient complexes. Transient protein complexes are particularly found in photosynthesis, facilitating electron transfer reactions. A rapid complex formation coupled to a fast reaction and protein dissociation is crucial to ensure the electron transfer not being limited in turnover. The association process of such dynamic complexes can be described by a two-step model. Initially, proteins are separated, attracting each other only by means of long-range, electrostatic interactions. Proteins approach each other to form the so-called encounter complex. In this state the interaction partners can assume different orientations within the complex prior to the formation of the active complex. The proteins then sample each others surface with the objective to form the well-defined complex, where the reaction can take place. This second step of association is dominated by short-range interactions. In order to study the influence of charge-charge interactions on complex formation, the analysis of the encounter complex is fundamental. A combination of chemical shift perturbation (CSP), paramagnetic NMR experiments, ensemble docking as well as Monte Carlo (MC) docking simulations is used to investigate and to visualize the complex orientations at the encounter state. In paramagnetic NMR experiments a protein is labeled with a spin label, that causes paramagnetic relaxation enhancements (PREs) on nuclei in its direct vicinity. This method is highly dependent on the distance and enables the visualization of lowly populated states, making it most suitable for studying the encounter complexes. Within this work the program MontyDock—a useful tool to simulate the first state of the association process—is presented. MontyDock is a rigid protein docking program, which evaluates the protein interactions solely by considering electrostatic interactions. The program is demonstrated, using the

well-established complex of cytochrome *c* and cytochrome *c* peroxidase. The analyzed complexes differ widely in the strength of electrostatic interactions. First, several plastocyanins (Pcs)—distinct in their surface charge distribution—were analyzed in complex with short, strongly charged, synthetic peptides marked with a paramagnetic spin label. The association of the peptides to the Pcs show a high dependency to electrostatic interactions. The NMR measurements of Pcs from *Populus nigra* and *Dryopteris crassirhizoma* reveal a good agreement with the docking results of MontyDock. In the visualized encounter complex the main ensembles are present at the highly charged regions of Pc, located at the eastern patch. Subsequently, the influence of a decrease in electrostatic stabilization is investigated with the complex of cytochrome *f* (Cyt*f*) and Pc from the cyanobacterium *Nostoc* sp. PCC 7119. Here the MC docking results do not fit the experimental measurements very well. Only a subordinate accordance suggests a partial reorientation of Pc to Cyt*f* based on electrostatic interactions. The dynamics of the complex and the binding orientation of Pc suggest hydrophobic interactions being the major stabilizing factor in the formation of this complex. Electrostatic interactions, however, still contribute a small part to the complex stabilization. The common description of the encounter complex by three distinct states is blurred by these findings and can be better represented by a smooth transition between the different states. This theory is emphasized by the analysis of the cross complex of *Phormidium laminosum* Pc and *Nostoc* Cyt*f*. For *Phormidium* Pc, electrostatic interactions are even of less importance than for the *Nostoc* Pc. The cross complex shows a decreased affinity and a more dynamic encounter complex compared to the *Nostoc* wild type. Nevertheless, the still-observable dependency on ionic strength is an evidence that the complex is influenced to some extent by electrostatic attraction. As a result, a model is proposed where the charges orient Pc to Cyt*f* in order to bring the hydrophobic regions in the vicinity of each other. Additionally, the cytochrome *c*₆-Cyt*f* complex from *Nostoc* was compared analytically to the Pc-Cyt*f* complex. The ensemble fitting indicates that the ensemble distribution cannot

be described by a single, well-defined complex but by a pure encounter state. The different types of complexes show that an encounter complex formation can only partially be described by electrostatic interactions alone. An interaction model is suggested where the interplay of hydrophobic interactions and electrostatic interactions regulates the dynamics and the specificity, resulting in indistinct states. This leads to an existence of various possibilities to conduct an electron transfer between transiently interacting proteins.

Keywords: Transient complex, Encounter complex, Electrostatics, Paramagnetic NMR, Protein-protein interactions, Plastocyanin, Cytochrome *f*

Zusammenfassung

Protein-Protein Interaktionen spielen in vielen zellulären Prozessen eine zentrale Rolle, zum Beispiel für Signalweiterleitung, Genregulation oder für molekulare Bioenergetik. Um die unterschiedlichen Reaktionsarten abzudecken, wird eine Vielzahl verschiedener Komplexe benötigt, welche von starken, teilweise irreversiblen Bindungen bis hin zu schwachen, kurzlebigen Interaktionen reichen können. Kurzlebige Proteinkomplexe kann man unter anderem in Photosyntheseprozessen finden, wo sie vor allem an Elektronentransferreaktionen beteiligt sind. Die Kombination aus schneller Komplexbildung und schneller Dissoziation ist entscheidend um die Geschwindigkeit der Elektronenübertragungen nicht zu limitieren. Der Assoziationsprozess solcher dynamischer Proteine wird meist durch ein Zwei-Stufen-Assoziationsmodell beschrieben. Zu Beginn der Reaktion liegen die Proteine getrennt voneinander vor und beeinflussen sich lediglich durch weitreichende, elektrostatische Wechselwirkungen. Die Proteine nähern sich einander an, um den sogenannten "Encounter Komplex" auszubilden. In diesem Schritt wird der Komplex zweier Interaktionspartner durch eine Vielzahl von Strukturen beschrieben. Hier tasten die Proteine gegenseitig ihre Oberflächen ab, um einen spezifischen Komplex auszubilden, in welchem die Reaktion stattfinden kann. Im zweiten Schritt der Komplexbildung wächst der Einfluss der kurzreichenden Wechselwirkungen. Um den Einfluss elektrostatischer Wechselwirkungen auf die Komplexbildung zu untersuchen ist die Analyse des Encounter Komplexes von zentraler Bedeutung. Hierfür wird in dieser Arbeit eine Kombination aus experimentellen Kernspinresonanz (NMR) Messungen sowie theoretischen Methoden wie Ensemble Docking und Monte Carlo Docking verwendet. Für die paramagnetischen NMR-Experimente wird ein Protein mit einer Spinsonde markiert, welche in seiner unmittelbaren Umgebung paramagnetische Relaxationseffekte (PRE) bedingt. Diese Eigenschaft prädestiniert die paramagnetische NMR für die Analyse des hochdynamischen Encounter Komplexes. In dieser Arbeit wird das Programm MontyDock vorgestellt,

welches zur Simulation des Assoziationsverhaltens zweier, voneinander getrennter Proteine, dem ersten Teilschritt einer Komplexbildung, verwendet werden kann. MontyDock behandelt Proteine als starre Körper und bewertet die Proteinwechselwirkungen ausschließlich auf elektrostatischer Basis. Das Programm wird am Beispiel des Cytochrom *c*-Cytochrom *c* Peroxidase Komplexes vorgestellt. Die analysierten Komplexe unterscheiden sich dabei stark in ihren elektrostatischen Eigenschaften. Im ersten Teil wird die Komplexbildung verschiedener Plastocyanine (Pc), die sich in ihrer Oberflächenladungsverteilung unterscheiden, mit kurzen, stark geladenen, synthetischen Peptiden untersucht. Die Peptide wurden hierfür mit einer paramagnetischen Spinsonde markiert. Die Komplexbildung der Peptide mit Pc zeigt eine hohe Abhängigkeit von elektrostatischen Interaktionen. Dabei stimmen die NMR-Messungen der Pcs, aus den Organismen *Populus nigra* und *Dryopteris crassirhizoma*, gut mit den Simulationen von MontyDock überein. Hierbei treten die Wechselwirkungen hauptsächlich in den stark geladenen Regionen von Pc auf. Des Weiteren wird am Beispiel von Cytochrom *f* (Cyt*f*) und Pc, aus dem Cyanobakterium *Nostoc* sp. PCC 7119 untersucht, wie sich die Abnahme der elektrostatischen Stabilisierung auf die Komplexbildung auswirkt. Die MC Docking Ergebnisse zeigen hier eine geringe Überstimmung zu den experimentellen Messungen. Lediglich die Übereinstimmungen einzelner geladener Reste, deutet auf eine partielle Neuausrichtung von Pc zu Cyt*f* basierend auf elektrostatischer Wechselwirkungen hin. Die Dynamik des Komplexes und die Bindungsorientierung von Pc weisen darauf hin, dass hydrophobe Wechselwirkungen den größten stabilisierenden Effekt bei der Komplexbildung haben. Durch diese Ergebnisse verwischt die bisher übliche Darstellung des Encounter Komplexes durch drei getrennte Zustände. Der Assoziationsprozess lässt sich stattdessen besser durch einen fließenden Übergang zwischen den einzelnen Zuständen beschreiben. Dieser Ansatz wird durch die Analyse des Mischkomplexes, aus *Phormidium lamniosum* Pc und *Nostoc* Cyt*f*, vertieft. Elektrostatische Wechselwirkungen spielen bei *Phormidium* Pc eine noch geringere Rolle als bei *Nostoc* Pc. Der Mischkomplex zeigt im Vergleich

zum *Nostoc* Wildtypkomplex eine geringere Affinität und einen dynamischeren Encounter Komplex. Eine Abhängigkeit der Komplexbildung von Ionenstärken zeigt jedoch, dass der Komplex noch geringfügig durch elektrostatische Interaktionen stabilisiert wird. Folglich wird ein Interaktionsmodell postuliert, bei dem Ladungswechselwirkungen Pc in Richtung Cyt*f* ausrichten, um die hydrophoben Bereiche der Proteine zueinander zu bringen. In dieser Arbeit wurde weiterhin der Cytochrom c_6 Cyt*f*-Komplex aus *Nostoc* untersucht und mit dem Pc-Cyt*f*-Komplex verglichen. Hier zeigt sich, dass die Verteilung des Komplexes nicht durch einem spezifischen Komplex ausgedrückt werden kann. Die Verteilung der Strukturen kann vielmehr durch das Auftreten eines reinen Encounter-Zustands erklärt werden. Die Ergebnisse aus den unterschiedlichen Untersuchungen zeigen, dass ein Encounter Komplex nur bedingt mittels elektrostatischer Wechselwirkungen beschrieben werden kann. Die Dynamik und Spezifität eines Encounter Komplexes kann besser durch ein Modell beschrieben werden, bei dem eine Kombination aus hydrophoben und elektrostatischen Wechselwirkungen die Bildung regulieren. Bei Elektronentransferproteinen führt dies dazu, dass man nicht mehr einen einzelnen fest definierten Zustand betrachtet, der katalytisch aktiv ist, sondern viele verschiedene, kurzlebige Strukturen, die einen Elektronentransfer ermöglichen.

Schlagwörter: Transienter Komplex, Encounter Komplex, Elektrostatik, Paramagnetische NMR, Protein-Protein Interaktionen, Plastocyanin, Cytochrom *f*

Contents

Abstract	i
Zusammenfassung	v
Contents	ix
1 Photosynthesis and its key to efficiency	1
2 Mechanism of Protein Association	5
3 Theoretical and experimental approaches for approximation of an encounter complex	9
3.1 Simulation of the free protein ensemble	9
3.2 Approximation of the encounter complex	12
3.2.1 NMR methods for analysis of the encounter complex	12
3.2.2 Ensemble docking for ensemble fitting	15
4 Characteristics of studied electron transfer proteins	17
4.1 Plastocyanin	17
4.2 Cytochrome c_6	19
4.3 Cytochrome f	20
4.4 Complex of Plastocyanin or Cytochrome c_6 with Cytochrome f	22
5 Manuscript Overview	25
5.1 Motivation and Synopsis	25
5.2 Contributions to the Joint Publications	34
6 Manuscript A	37

7 Manuscript B	69
8 Manuscript C	89
9 Manuscript D	111
10 Manuscript E	127
List of Abbreviations	139
Bibliography	141
Danksagungen/ Acknowledgements	149

Chapter 1

Photosynthesis and its key to efficiency

Photosynthesis is among the greatest influences on the evolution of life on Earth. Without photosynthesis life as we know it would not be possible. About 2.5 billion years ago, prokaryotic organisms—cyanobacteria and green sulfur bacteria—started to use sunlight as a source of energy.¹ A side product of this metabolic process was the secretion of oxygen. Over time, the highly reactive oxygen enriched in the atmosphere and caused mass mortality for organisms that could not cope with this change. The oxygenic atmosphere allowed also the evolution of a new kind of cells, namely eukaryotes, which utilize the oxygen to produce energy.² Nowadays a multitude of organisms use photosynthesis to conserve light energy like i.e. plants and bacteria. Throughout the numerous organisms using photosynthesis, various types of proteins facilitating the light harvesting and electron transfer reactions, as well as different subcellular arrangements can be found.³ Nevertheless, the fundamental mechanism for photosynthesis remains the same. Energy delivered by sun light is converted to chemical energy that is used for biomass production. Photosynthesis is a cascade of reactions, consisting of several coupled exciton, inter- and intramolecular electron transfers and proton translocation reactions.⁴ A schematic illustration of this process can be seen in Figure 1.1 for cyanobacteria. It starts with the light harvesting reaction at the

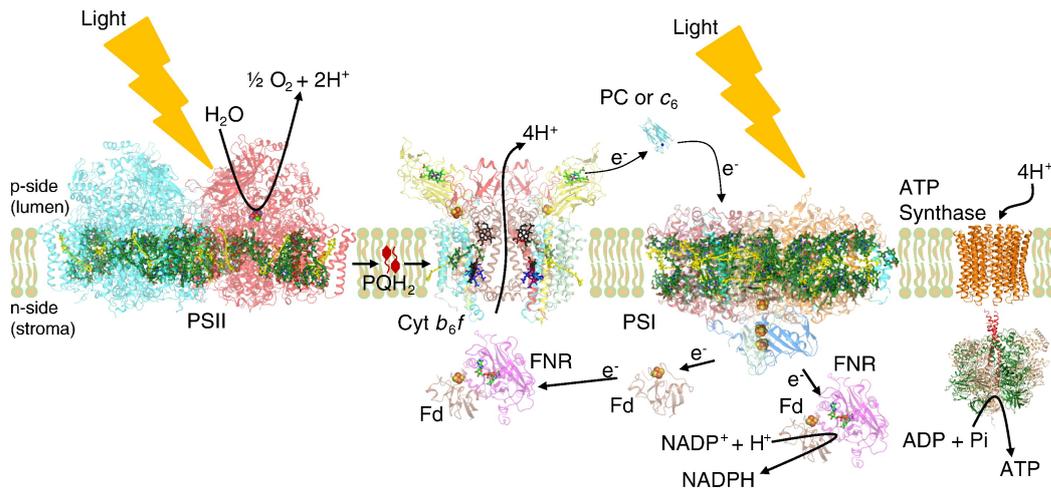


Figure 1.1. Overview of the photosynthetic process in cyanobacteria. The transmembrane proteins PSI, cytochrome b_6f , PSII and ATP-Synthase are involved in the electron transport chain of oxygenic photosynthesis to conserve energy in biomolecules. Formation of the transmembrane proton electrochemical gradient coupled to the electron transport extending from H_2O oxidation to NADP^+ reduction, in which H^+ is translocated in the protein complexes of the PSII reaction center and cytochrome b_6f ; this H^+ gradient is utilized for ATP synthesis by the ATP synthase. PDB accession for structure data: cytochrome b_6f (PDB ID: 2E74⁵), Fd (PDB ID: 1EWY⁶), ferredoxin; FNR (PDB: 1EWY⁶), ferredoxin- NADP^+ -reductase; PC (PDB: 2Q5B), plastocyanin; PSII (PDB: 3ARC⁷) and PSI (PDB: 1JB0⁸), reaction center complexes. Reprinted from Publication Hasan et al.⁹ with permission from Elsevier.

photosystem II (PSII), where a photon excites bacteriochlorophyll a molecules followed by an energy transduction towards the reaction center P680. Here the light energy is used to split H_2O in oxygen and protons, while the remaining electrons are transferred to plastoquinone (PQH_2). This molecule stabilizes the additional electron by binding a proton and transfers the electron to the cytochrome b_6f complex. From cytochrome f (Cyt f), which is part of the cytochrome b_6f complex, the electron is transferred by a mobile protein, either plastocyanin (Pc) or cytochrome c_6 (Cyt c_6), to the photosystem I (PSI). The electron is after the excitation then transferred over Ferredoxin to Ferredoxin- NADP^+ -reductase, where the reductive potential is stored at the final

acceptor NADPH. In the reaction cascade also a membrane potential is generated, where protons gather on the lumen side of the membrane, which are generated by the splitting of water or by proton translocation processes in the cytochrome b_6f complex. By translocating the superset of protons and with this degrading the membrane potential, the protein ATP-synthase utilizes the proton gradient for generating ATP, which is the universal intracellular energy carrier in nearly all living cells. Considering all steps in the energy cascade, photosynthetic efficiency regarding energy conversion is not higher than about two percent. Nature optimized each single step in this process in order to generate a downhill energy gradient to make the whole process energetically favorable. In order to ensure the energetic downhill gradient throughout the reaction cascade, each chromophore is embedded in a different, very specific chemical environment.¹⁰ Besides the tuning of intramolecular pigments, also the electron transfer reaction between the transmembrane proteins, namely PSII, cytochrome b_6f and PSI, is optimized by nature. The speed of complex formation is ensured not to be the limiting turnover factor for electron transfer reactions. Nature optimized this process to find a reasonable balance between specificity of binding and rapidity of electron transfer. This is emphasized in the interaction of Pc with the transmembrane proteins cytochrome b_6f and PSI play a crucial role for the turnover rate in photosynthesis.^{11,12} The complex formation of electron transfer proteins can not be described in the classical way. In the next chapter complexes are distinguished concerning their stability and their complex formation in terms of their transient nature, in order to give an overview of different types of protein complexes. In addition, a model to describe the dynamic character of electron transfer proteins is highlighted.

Chapter 2

Mechanism of Protein Association

Protein complexes can be described according to physical and chemical properties of the proteins, as well by means of binding affinity, complex stability, complex lifetime or by the number of proteins involved in complex formation. Nooren and Thornton¹³ provided an overview for different types of protein-protein interactions:

- (i) Homo- and hetero-oligomeric complexes
- (ii) Non-obligate and obligate complexes
- (iii) Transient and permanent complexes

This work focuses primarily on transient complexes. Permanent protein complexes, also known as static complexes, form a stable bond, while transient complexes are in equilibrium between association and dissociation. Permanent complexes consist of very strong interacting proteins. The binding interface is highly stabilized by electrostatic interactions, salt bridges, Van der Waals interactions and complementary surface structures, also described as key-lock-principle.¹⁴ Low dissociation rates (K_D) in the nM range are characteristic for this kind of complexes leading to a very tight binding, often without notable dissociation.¹⁵ A well known example is the antigen-antibody binding or the binding of an inhibitor to an enzyme. Transient complexes behave

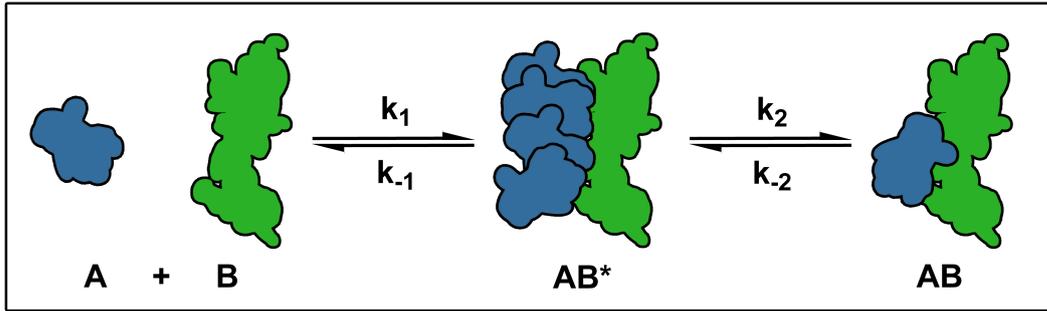


Figure 2.1. Two-step kinetic model for the description of protein-protein complex formation in the context of a productive transient complex: In the first step, free proteins form an encounter complex (AB^*) mostly driven by long-range electrostatic interactions or random collisions. In the second step, short range interactions like Van der Waals interactions, electrostatics or hydrogen bonds contribute to the formation of the well-defined complex (AB).

in the opposite way. A transient complex has a much higher dissociation rate with a K_D in μM to mM range.¹⁶ Depending on the outcome of a productive transient complex, the class can be divided in strong and weak subgroups. In very weak transient complexes one protein scans the surface of another protein and dissociates without formation of a productive complex. In strong transient complexes, the proteins also scan each others surface but form a stable, productive complex. The formation of a productive transient complex can be described by a two-step model, which is schematically shown in Figure 2.1.^{17,18} This model is also described as an encounter complex formation. This complex formation was early described as Velcro model,¹⁹ which rejects the key-lock-principle of permanent complexes. In this model the protein kinetics in equilibrium can be described by an association constant $K_A = \frac{k_{on}}{k_{off}}$ and its inverse, the dissociation constant K_D . The rate of dissociation is represented by k_{off} , which can be described with a combination of rate constants from Figure 2.1 under the assumption of AB^* being in steady state with $k_{off} = \frac{k_{-1}k_{-2}}{k_{-1} + k_2}$.²⁰ The association rate k_{on} is described as $k_{on} = \frac{k_1k_2}{k_{-1} + k_2}$, respectively. In the first step of the model, the two proteins forming a complex are separated from each

other. At this stage mainly long range interactions, like electrostatic attraction or repulsion, play a major role.²¹ Guided by charge-charge attractions, the proteins approach each other, followed by scanning each others surface for an energetically stabilized orientation or geometric complementarity. At this point, short range interactions get more important and lead to the stabilization of the complex, although they are significantly weaker compared to the forces in strong complexes. The microscopic distribution of charges is another important factor for complex stabilization. As long as the proteins are far apart, the individual proteins perceive the charges as an overall average. By approaching the surface of the binding partner, local charges become more important causing a relative rearrangement of the protein orientation. The goal of the surface scan is to get to the final complex, where the reaction between the two proteins can take place. Depending on the nature of the formed complex, the equilibrium can differ between the free proteins, the encounter state and the well-defined complex. While some proteins can form a specific complex,²² others are only present in an encounter state.²³⁻²⁵ To visualize the structure of the specific complex, crystal structures or NMR models are used as an approximation. The complexes solved by crystal structures represent densely packed proteins in energetically favorable orientation. Compared to that, for the NMR models, a combination of several measurement techniques is used to calculate an ensemble of structures with a low energy and smallest violations of the NMR data sets.

The two-step kinetic model for protein-protein complex formation is well suited to describe the interaction between mobile photosynthetic proteins and their reaction partners, the transmembrane complexes, both introduced in the previous chapter. To guarantee an efficient electron reaction, fast dissociation is needed to make room for new electron donors and acceptors whereby the flow of electrons is kept running. With the electron transfer playing a crucial role in the photosynthesis, this work attempts to analyze how the complex formation process of transient complexes can be visualized and used for a better understanding of the interplay of dynamic proteins.

Chapter 3

Theoretical and experimental approaches for approximation of an encounter complex

An encounter complex is not easy to study due to its transient character. To approximate the whole process of protein association with its individual steps (see Figure 2.1) several different computational and experimental techniques are utilized. In order to describe the free proteins and their long range attraction, Monte Carlo docking with its main focus on electrostatic interactions is used. The encounter complex is detected and visualized by a combination of several nuclear magnetic resonance (NMR) spectroscopy techniques and ensemble docking. For the specific complex either NMR structures are utilized or, if applicable, X-ray crystallography can be used.

Simulation of the free protein ensemble

In the beginning of this section, some approximations are explained, which are used for a description of molecules on atomic level. In general, all interaction of atoms rely on the interaction of electrons with protons and their distribution. On molecular basis the distribution of electrons, which interact with the atomic nucleus, can be calculated through quantum chemistry resulting in the most probable structure.

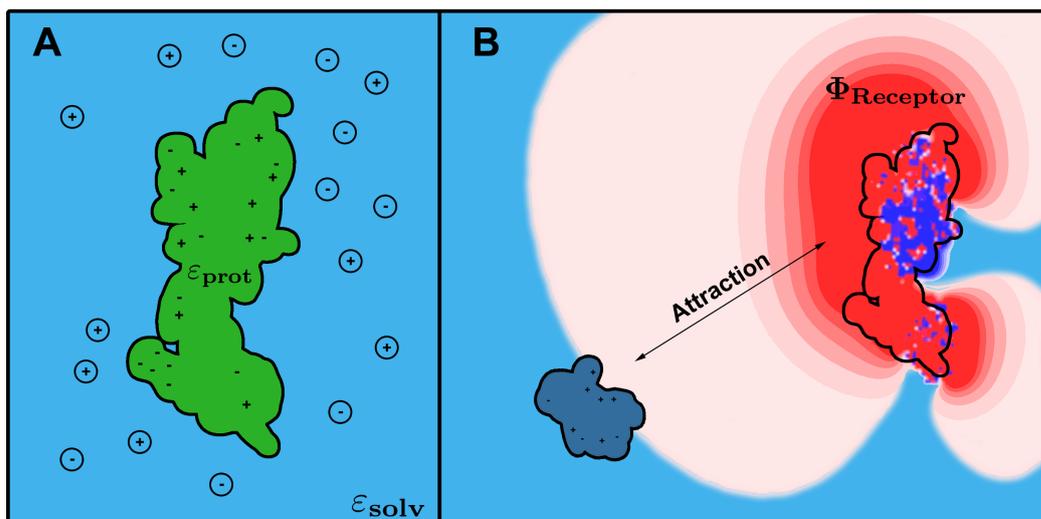


Figure 3.1. Part A: Schematic representation of spatial subdivision in the continuum electrostatic model. The protein with its partial charges is shown in green and represented with a dielectricum ϵ_{prot} , while the solvent with ions is shown in blue (ϵ_{solv}). The ions in the solvent are in equilibrium. Part B. Schematic overview of a Monte Carlo docking event. The ligand protein (blue) interacts with the electrostatic potential of the receptor (red).

This method is very limited by the size of the system because quantum chemistry calculations are computationally very expensive. Alternatively, the distribution of electrons can be approximated by a charge density or partial charges. Several force fields^{26–28} were developed to describe the charge distribution in proteins using partial charges. To account for different polarization effects, space is divided in regions with different dielectric constants,²⁹ which describe the electrostatic permittivity of every individual region (Figure 3.1A). Considering proteins as rigid with the charges fixed at their position and hence having a defined chemical environment, a low dielectric constant of $\epsilon = 4$ is used. In contrast solvent molecules, mostly water, which are highly mobile, are described a dielectric value of 80. The high value for water accounts for the high polarizability and its mobility, which results in a shielding of electric field. The partial charge description of proteins and solvents with the polarizability effects can be combined in the continuum electrostatics model.^{30,31} The Poisson-Boltzmann equa-

tion mathematically describes the continuum electrostatics model and relates the spatial charge distribution to the dielectric regions with the electrostatic potential Φ by equation 3.1:

$$\nabla[\varepsilon(\mathbf{r})\nabla\Phi(\mathbf{r})] = -4\pi\left(\rho_{\text{prot}}(\mathbf{r}) + \underbrace{\sum_{i=1}^K c_i^{\text{bulk}} Z_i e_0 \exp\left(\frac{Z_i e_0 \Phi(\mathbf{r})}{RT}\right)}_{\rho_{\text{ions}}(\mathbf{r})}\right) \quad (3.1)$$

where ε is the absolute permittivity, c_i the concentration of ionic charges, Z_i the net charge of the ion and e_0 the elementary charge with $e_0 = 1.602 \cdot 10^{-19}$ C. The potential of the ion part ρ_{ions} depends on the temperature T and the universal gas constant R . For low electrostatic potentials a linearized version of this equation can be obtained:

$$\nabla[\varepsilon(\mathbf{r})\nabla\Phi(\mathbf{r})] = -4\pi\rho_{\text{prot}}(\mathbf{r}) + \bar{\kappa}^2(\mathbf{r})\Phi(\mathbf{r}) \quad (3.2)$$

where κ is the inverse Debye length with $\bar{\kappa} = \frac{8\pi N_A e_0^2 I}{k_B T}$ and I being the sum over all mobile ions $I = \frac{1}{2} \sum_i c_i Z_i^2$. The linearized version allows the calculation of electrostatic potential for different charge distribution separately and relates them afterwards by addition, as long as the spatial distribution of dielectric permittivity does not change.³² The electrostatic description by linearized Poisson-Boltzmann equation³³ can be used to describe or predict the association of two separated proteins. The interaction between two molecules is evaluated by calculating the point charges of “molecule A” in the electrostatic potential of “molecule B”:

$$E_{\text{int}} = \sum_{n=0}^{N_a} q_{n,a} \cdot \Phi_b(r_n) \quad (3.3)$$

where n_a is the number atoms of molecule a with the respective partial charge q_a and Φ_b the electrostatic potential of molecule b at the position of the charge q_a . If the sum over all atoms is negative, the two molecules attract each other, while if the sum is positive, the two molecules repel each other in this orientation. In this way a relative position of two molecules to each other can be evaluated, whether the interaction is energetically favorable, even when the two proteins are

apart from each other. This long range attraction or repulsion can be used as an approximation for the interactions of the free proteins in an encounter complex.^{21,34} In order to generate an ensemble of energetically favorable structures describing the interacting proteins, the program MontyDock was developed. The program MontyDock is explained in detail in Manuscript A, including its algorithm. In the simulation a ligand moves randomly in the electric field of a receptor. A schematic representation of this process is shown in Figure 3.1B. For every step the electrostatic interaction energy is calculated by Equation 3.3. The energy is then either accepted or declined with a probability p , defined by the Metropolis Monte Carlo criteria, where 1 means accepted and 0 rejected.³⁵

$$p = \begin{cases} 1 & \text{if } \Delta E_s^{int} \leq 0 \quad \vee \quad (e^{\Delta E_s^{int}} < \text{Ran}() \quad \wedge \quad \Delta E_s^{int} > 0) \\ 0 & \text{if } \Delta E_s^{int} > 0 \quad \wedge \quad e^{\Delta E_s^{int}} > \text{Ran}() \end{cases} \quad (3.4)$$

The simulation result is an ensemble of structures, which are purely weighted according to their electrostatic interaction. This method has already been successfully applied in encounter complex studies before.^{22,36}

Approximation of the encounter complex

The dynamics of an encounter complex can best be described by the combination of theoretical methods together with NMR spectroscopy. NMR measurements are used for the analysis of dynamics and are therefore predestined to investigate the encounter complex. In the following section, it will be highlighted how different informations of NMR experiments are used in combination with ensemble docking to get a better insight in the dynamics of transient complexes.

NMR methods for analysis of the encounter complex

In NMR each atom of a protein gives rise to a unique signal, which is influenced by its chemical environment. One frequently used ex-

periment is the ^1H - ^{15}N hetero spin quantum coherence (HSQC) experiment, where a signal represents the spin coupling between a ^{15}N labeled amide and the bound proton. In the fast exchange regime on the NMR time scale, a peak in the HSQC represent an averaged signal of all species present in then sample, including the free or bound form of the observed protein and all possible orientations that the protein can assume within a complex. When two proteins bind to each other, the chemical environment of the amino acids in the binding site changes, which leads to a shift of the signal. The amplitude of the shift is a measurement for the binding strength and accounts for the dynamics in a complex.^{23,25} If two proteins bind very tightly, it will result in a large shift of the signal. Chemical shift perturbations (CSP) provide information on the dynamics of the complex including the calculation of K_A , the affinity constant, and K_D , the dissociation constant. This method for interaction analysis is also called chemical shift perturbation analysis. By mapping the change of chemical shifts to the amino acids in a protein, regions highly contributing to complex formations can be identified.

Another method for the analysis of encounter complexes is paramagnetic NMR spectroscopy. The paramagnetic effect has already been used in electron paramagnetic resonance (EPR) measurements for a long time. In the last few years the paramagnetic effect has also proven to be a very powerful tool in NMR analysis, especially for proteins in the fast exchange regime.³⁷⁻⁴² Battiste and Wagner³⁷ developed a site-directed spin labeling technique to derive distance restraints for structural calculations. Paramagnetic spin labels have unpaired electrons owing a large magnetic moment. By magnetic dipolar coupling of the unpaired electron with the nucleus, the nuclear transversal spin relaxation R_2 is increased, which leads to line broadening and hence to intensity decrease of the signal in the spectrum. This paramagnetic relaxation enhancement (PRE) Γ_2 can be described with equation 3.5:

$$\Gamma_2 = \frac{\gamma^2 g^2 \beta^2}{20r^6} \left(4\tau_c + \frac{3\tau_c}{1 + \omega_h^2 \tau_c^2} \right) \quad (3.5)$$

where γ and ω_h represents the nuclear gyromagnetic ratio and the

Lamor frequency; g , the electronic g-factor; β , the Bohr magneton; and r , the distance between the nucleus and the paramagnetic source. The PRE effect is inverse proportional with the sixth power to the distance of the nucleus to the spin label. This results in a high sensitivity to distance, which makes the visualization of lowly populated states up to a distance of 35 Å possible.⁴³ Due to this effect, PRE is very suitable to analyze the encounter complex described in Section 2.^{36,44} The spin labels can be separated in two groups, namely the isotropic and anisotropic spin labels. While the isotropic spin labels give rise only to PREs, anisotropic spin labels induce PREs as well as pseudocontact shifts (PCSs) and residual dipolar couplings (RDCs). The PCSs arise from the direction of the specific g-tensor of the unpaired electron, which causes dipolar couplings between the magnetic moment of the electron and the nucleus of the protein. An example for an isotropic spin label is the widely used MTSL ((1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanethiosulfonate), while for an anisotropic label caged metals can be named.^{45–47} MTSL has a stabilized nitroxid radical as a single unpaired electron.⁴⁸ As an anisotropic label also protein intrinsic metal centers, like the iron of a heme cofactor, can be used. In the experiments with PREs, two measurements need to be made, one with a paramagnetic active label (MTSL) and one with a diamagnetic control MTS ((1-acetyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanethiosulfonate). The ratio between the paramagnetic signal intensity I_{para} and the diamagnetic intensity I_{dia} shows the fraction effected by PREs. This fraction can then be related by equation 3.6 to Γ_2 :

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2d}e^{-\Gamma_2 t}}{R_{2d} + \Gamma_2} \quad (3.6)$$

The combination of equation 3.5 and 3.6 enables the calculation of distance constrains, which can be used in a further step—the ensemble docking—to visualize the encounter complex.

Ensemble docking for ensemble fitting

Interaction in protein complexes are hard to investigate with only a single method. Instead an interplay of computational and experimental methods is used for a better interpretation of a multivariate system like the formation of an encounter complex. Structural data, for example provided by NMR, X-ray crystallography or small angle X-ray scattering, can be used for structure calculations, complex modeling, docking simulations or kinetic models. In the case of NMR, experimental data sets are used as a constrain for structure calculations. Here a best fit to experimental data is obtained by getting the structure with the lowest energy and the lowest constrain violations. Referring to the encounter complex, a single structure often cannot be used to describe the extent and diversity of an encounter complex. It can better be described as an ensemble of structures, that are in equilibrium between the encounter state and the well-defined complex.⁴⁹ For this a rigid body docking method is used, where a diverging number of structures is docked to a receptor at the same time to obtain the lowest violation of experimental data. This method has already been successfully applied in different systems to visualize their encounter states.^{22,50,51} As a restrain for the ensemble docking PRE data sets are used. The procedure to get an approximation of the encounter ensemble is described and discussed in detail in Chapter 8.

Chapter 4

Characteristics and interplay of studied electron transfer proteins

Up to this point the model for the formation of an encounter complex and several methods for the analysis of such complexes have been described. To give an overview of the proteins subject of this thesis, the individual proteins will be highlighted in the following section. The detailed structure of Pc and Cyt_c₆, two mobile proteins which play an important role in the electron transfer during photosynthesis, as well as the structure of Cyt_f, one of their reaction partners, is described. Following to that, the complexes formed by these proteins are highlighted.

Plastocyanin

Pc is the main interaction partner of Cyt_f and serves as an electron carrier to the chlorophyll of PSI.⁵² The length of Pc varies from 97 to 105 amino acids, depending on the organism, and forms a secondary structure consisting of seven β -sheets. The variants with up to 105 amino acids mainly occur in cyanobacteria. Pc belongs to the type I copper protein family, containing a redox active copper, coordinated by two histidines, one methionine and one cysteine. The protein has

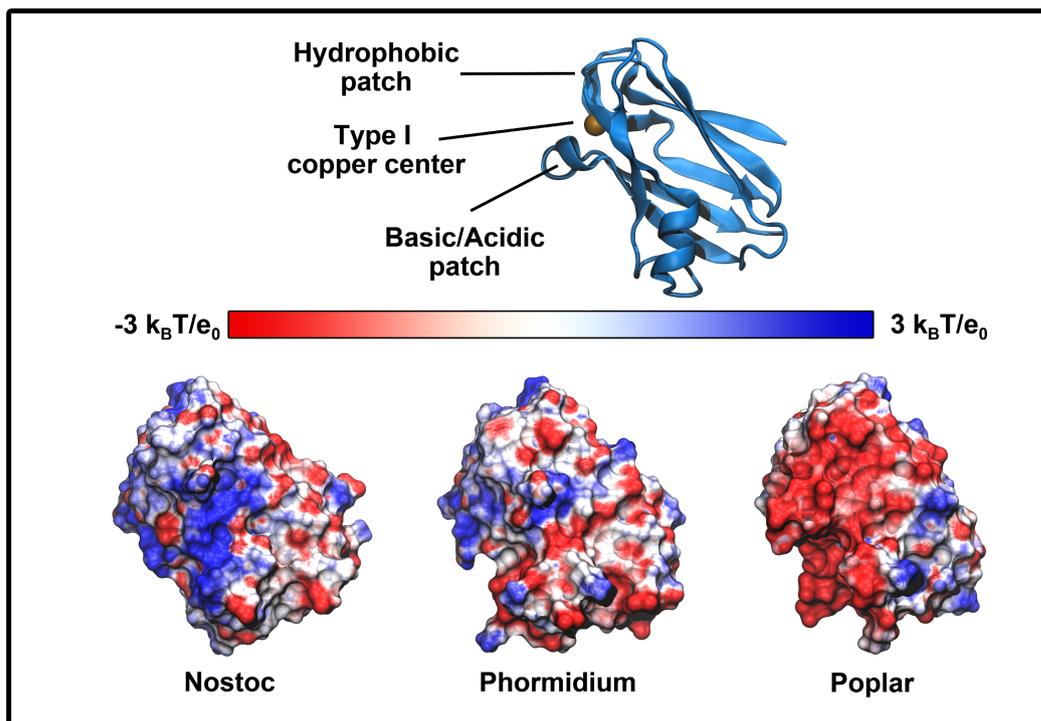


Figure 4.1. The structure of Pc consists of seven β -sheets and has a copper atom as a cofactor. Pc has two significant regions, the hydrophobic and the basic/acidic (eastern) patch, which play an important role for complex formation with Cyt*f*. The electrostatic potential for the organism *Nostoc*, *Phormidium* and *Poplar* are mapped on the surface ranging from a $-3 k_B T/e_0$ (red) to $3 k_B T/e_0$ (blue).

a characteristic blueish color at 597 nm, which is responsible for its alias 'blue copper protein'. The copper atom is buried 5 Å below the surface at a region mainly consisting of hydrophobic amino acids and therefore known as hydrophobic patch.^{53,54} In Figure 4.1 an overview for the structural regions of Pc is given. PC includes a slightly charged region at the side of a hydrophobic patch which is commonly known as eastern patch. This region plays a major role in the complex formation procedure with Cyt*f*.^{40,55} In plants and the cyanobacterium *Nostoc* sp. PCC 7119 the eastern patch is complementary charged to Cyt*f*. This regions contribute highly to the electrostatically stabilization of the complex formation and to the relative protein orientation within the complex.^{56,57} For the thermophilic *Phormidium laminosum*, where the

eastern patch is less charged, the binding orientation is different. The low presence of electrostatic interaction causes the transient complex in *Phormidium* to be much more dynamic than in plants or *Nostoc*.

Cytochrome c_6

Cytc₆ is the second electron transfer protein interacting with the cytochrome *b₆f* complex. It is suggested, that depending on the environmental factors the bacteria switches for the electron transfer reaction from Pc to Cytc₆ especially when there is an deficiency of copper.^{58,59} Cytc₆ has a similar redox potential as Pc ranging from 335-390 mV,^{60,61} and is therefore well suited to substitute Pc as an electron carrier from Cyt*f* complex to PSI.^{60,62} In higher plants a Cytc₆ analogue was found, that has a lower redox potential.⁶³ This makes it unsuited for the oxidation of Cyt*f* and is hence presumably not involved in photosynthesis.⁶⁴ The structure of Cytc₆ is well studied and several crystal structures from various organisms including cyanobacteria, algae and terrestrial plants are known.⁶⁵⁻⁶⁷ The length of Cytc₆ ranges from 83 to 90 amino acids and its secondary structure consists mainly of α -

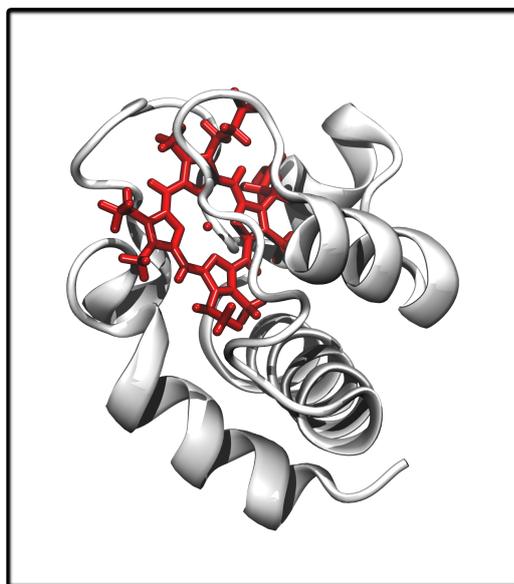


Figure 4.2. Cytochrome c_6 is a globular shaped protein consisting mainly of α -helices. The cofactor heme is shown in red, the protein backbone in gray.

helices forming a ternary structure of globular shape. An overview of the $Cytc_6$ structure is shown in Figure 4.2. $Cytc_6$ belongs to the class I cytochrome *c* family, where a single cytochrome *c* cofactor is bound to the protein by a CXXCH motive with the fifth and sixth residues coordinating the heme-iron being histidine and methionine. Considering the primary and secondary structure few parallels can be seen to Pc. In spite of the structural differences, $Cytc_6$ and Pc show a similar isoelectric point.⁶⁸ The interaction of $Cytc_6$ with *Nostoc* *Cytf* is slightly reduced, compared to the one of Pc, which can be explained by the charges being regionally not as highly concentrated as for Pc.

Cytochrome *f*

Cytf is part of the cytochrome b_6f complex, a homo-dimeric transmembrane protein, located in the thylakoid membrane in chloroplasts of plants, algae and cyanobacteria.^{69,70} *Cytf* is the soluble part of the cytochrome b_6f complex anchored to the transmembrane part of the complex by a helix of 30 hydrophobic amino acids. An overview of the cytochrome b_6f complex with *Cytf* highlighted in red can be seen in Figure 4.3. As mentioned before, *Cytf* is the direct interaction partner of Pc or $Cytc_6$, which transfers the electron yielded by *Cytf* to the PSI complex.⁵² The protein has a molecular weight of about 28 kDa and an elongated shape, which is directed along the thylakoid membrane. Macroscopically *Cytf* consists of two domains, a big and a small one.⁷¹ Just like $Cytc_6$, *Cytf* belongs to the c-type cytochrome family with the c-type heme bound in the big domain.⁷² This heme acts as the electron donor for its redox partner and thus the region surrounding the heme is the active site for an electron transfer. The cofactor is covalently bound to the protein by two thioethers and the heme iron is additionally coordinated by a histidine and the N-terminus of the protein chain.⁷³ The heme is buried in a hydrophobic region at the edge between the two domains of *Cytf*, which represents one of the binding sites for Pc and $Cytc_6$.⁷² At the upper ridge of the small domain, an acidic or basic patch is located, which is electrostatically stabilizing the

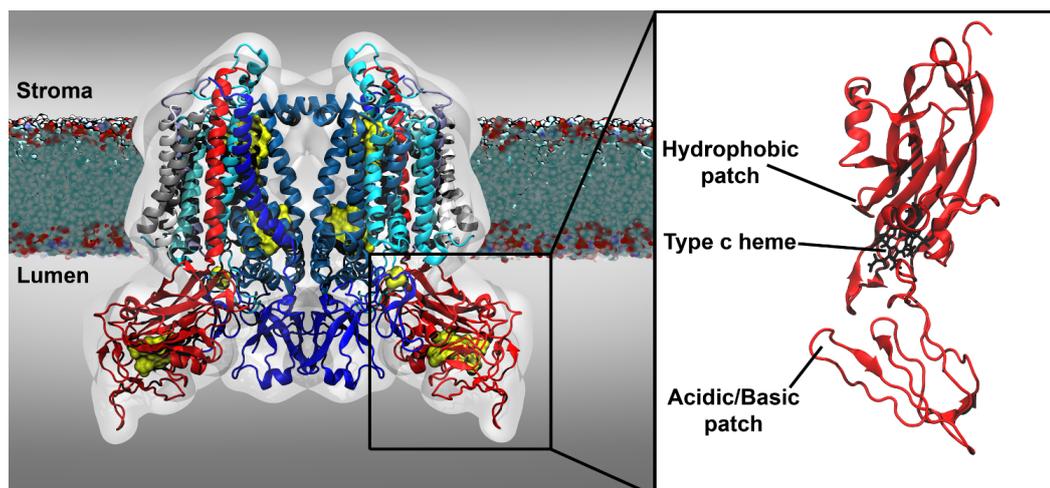


Figure 4.3. Overview of the homo-dimeric cytochrome b_6f complex anchored to the thylakoid membrane. The complex is colored according to its protein chains. The cofactors bound to cytochrome b_6f are highlighted in yellow. The soluble part Cyt f (red) is located on the thylakoid side. In Cyt f the region around the c-type heme (black) binding site is called hydrophobic patch. The small protein domain has a region, that is differently charged depending on the organism also known as acidic or basic patch.

complex formation of Pc and Cyt f . As well as in Pc the charged regions of the proteins vary depending on the organism. In detail, while the cyanobacterium *Nostoc* has a negatively charged patch in the small domain of Cyt f and a positive one in the opposing counterpart, the same region in plants is charged contrariwise. A special role has Cyt f in *Phormidium*, also a cyanobacterium, where in both proteins, Cyt f and Pc, highly charged regions are lacking.

Complex of Plastocyanin or Cytochrome c_6 with Cytochrome f

One of the first complex structures for Pc and Cyt f was solved by Ubink et al.⁴⁰ with paramagnetic NMR in 1998. Since then several structures of photosynthetic complexes from different organisms have been determined using classical NMR methods.^{38,39,74} Soon the research interest was directed towards the structural differences of this complex in various organisms. Amongst the cyanobacterial complexes, the one of *Nostoc*^{38,74} and *Phormidium*⁴¹ was solved as well as for plants the cross complexes, consisting of poplar or spinach Pc with turnip Cyt f .^{39,40} In the productive complex consisting of *Nostoc* Pc and Cyt f , Pc is orientated in a “side-on” orientation at the acidic patch of Cyt f .³⁸ An overview of the different complex orientations of *Nostoc*, plants and *Phormidium* can be seen in Figure 4.4. Several studies have shown, that the H87 is a key residue for the formation of the complex.^{40,75,76} H87 was stated to be also a key residue for the complex formation in plants.^{39,40}

Comparing the complex formation on basis of affinity studies, *Nostoc* Pc with Cyt f showed a K_A of 26 mM⁻¹ with an efficient electron transfer rate of $13.4 \cdot 10^{-3} \text{ s}^{-1}$.⁵⁶ Similar observations can be made for the plant complex composed out of turnip Cyt f and spinach Pc.⁴⁰ Compared to *Nostoc* the plant complex has oppositely charged patches and an affinity constant K_A of 6.9 mM⁻¹ in the order of magnitude with a second order rate constant of $1.76 \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$ for complex formation.⁷⁸ With the charged patch contributing highly to the binding orientation and stabilization, the complex appears in a “side-on” orientation (Figure 4.4C), similar to that observed in *Nostoc*. On the contrary in *Phormidium*, where electrostatics and eastern patch play a minor role in the specific complex formation, Pc binds in a ‘head-on’ orientation with the northern patch pointing in the direction of the Cyt f heme (Figure 4.4B). Not a highly charged region, but individual charged residues in the vicinity of the hydrophobic patch stabilize this orientation. Regarding the affinity, the *Phormidium* complex is also different from the mentioned organisms with a K_A of 0.3 mM⁻¹.^{41,79,80}

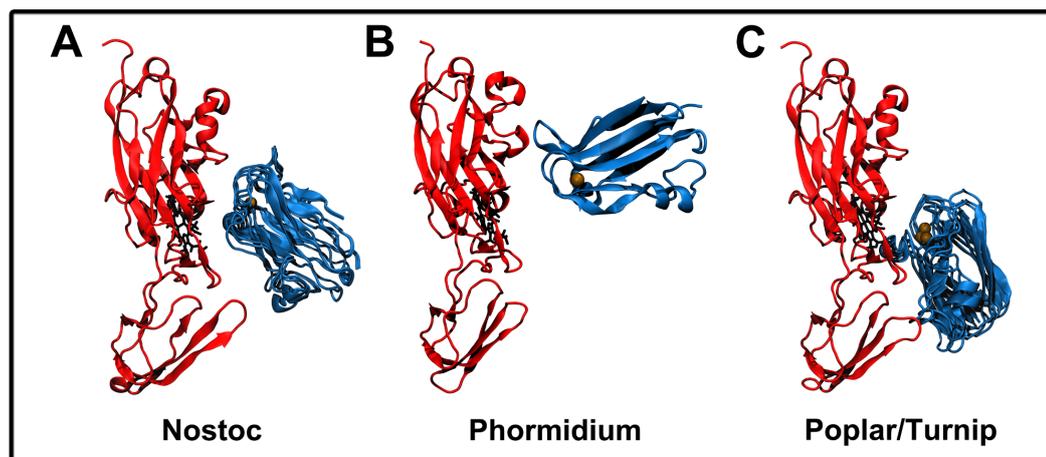


Figure 4.4. Comparison of the binding orientation of Pc (blue) to Cytf (red) for different organism. A. Specific complex of *Nostoc* (PDB: 1TU2³⁸). B. Modeled complex for *Phormidium* based on Crowley et al.⁴¹ C. Specific complex of Poplar Pc with Turnip Cytf (PDB: 1TKW³⁹). In *Nostoc* and plant, Pc is orientated to Cytf in a 'site-on' orientation, while for *Phormidium* Pc binds in a 'head-on' orientation. All complexes were determined by NMR spectroscopy.

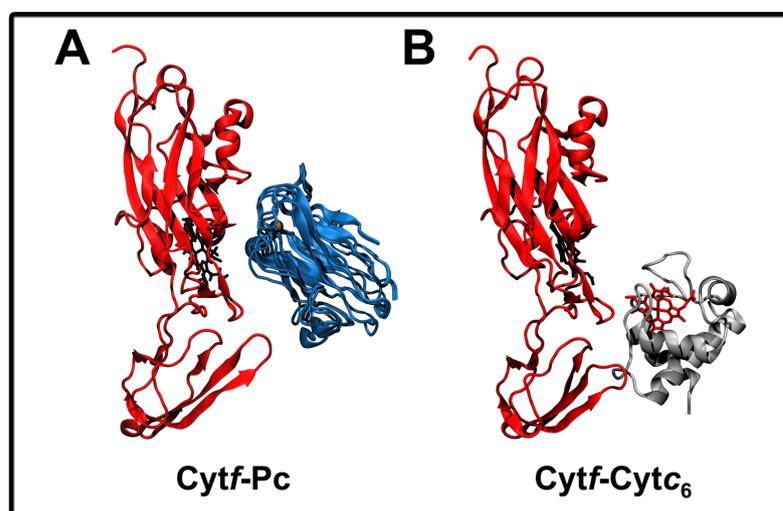


Figure 4.5. Comparison of the binding orientation for Pc-Cytf complex (A) (PDB: 1TU2³⁸) to the Cytc₆-Cytf complex (B).⁷⁷ Pc and Cytc₆ are colored blue and gray, respectively. Cytf is shown in red.

While the complex of Pc-Cytf has been subject of several studies, the interactions of cytochrome c_6 with Cytf are significantly less analyzed. Considering the differences in primary, secondary and ternary structure of Cyt c_6 compared to Pc a quite different binding and catalytic behaviour could be assumed. Crowley et al.⁸¹ found an affinity constant for *Nostoc* Cyt c_6 -Cytf of 8.0 mM^{-1} to 10.0 mM^{-1} , which is in the range of the Pc-Cytf complex, though slightly smaller. This value was affirmed by a later study.⁷⁷ A reason for the decrease in affinity might be the hydrophobic regions contributing to the encounter complex. A comparison of the specific Cyt c_6 -Cytf complex with the one of Pc-Cytf⁷⁷ for *Nostoc*, shows a similar orientational alignment of the hydrophobic part, where the heme cofactor is located (Figure 4.5). The main binding site seems also to be electrostatically stabilized by charges. This makes the complex formation comparable to the Pc-Cytf complex. With the Cyt c_6 -Cytf complex having a higher variance in complex formation for different types of organism, this gives rise of different electron transfer mechanism than in Pc-Cytf,⁸¹ though relatively little research has been conducted on this subject.

Chapter 5

Manuscript Overview

Motivation and Synopsis of the manuscripts

The formation of a close complex is necessary to achieve an effective electron transfer. Marcus^{82,83} showed, that the probability for a successful electron transfer decreases exponentially with increasing distance between the reaction centers of the electron transfer partners. Hence only few possibilities for a successful electron transfer are available. If an electron transfer process would be regulated by random collision of proteins, the probability for a successful electron transfer would be too low to maintain the physiologically necessary electron flow. Nature therefore needs a driving force to modulate the specificity and the dynamics of electron transfer complex formation. The analysis of different criteria contributing to the formation of a dynamic complex is not an easy task. The encounter complex model represents an approach that tries to summarize all occurring structures and intermediates in an ensemble of structures. The analysis of the encounter complex is important; it provides an understanding of the synergy between specific and unspecific interactions and how they contribute to protein association. Each individual step of the complex formation is therefore described; from the free proteins, to the encounter complex itself and to the specific, well-defined complex. The main focus of this work is the simulation of the initial step of the encounter complex formation, namely the long range electrostatic

interactions between free, separated proteins. The program MontyDock, described in Manuscript A, was used to calculate these interactions. MontyDock uses the electrostatic interactions of two proteins as its sole criteria for evaluating a protein complex and is therefore well suited for describing the initial phase of the encounter state. The program is demonstrated on the experimentally and theoretically well studied complex of cytochrome *c* and cytochrome *c* peroxidase. For this complex, it is well established that complex formation is highly dependent on charge-charge interactions.^{22,84,85} Several studies on a variety of electron transfer complexes and their relation to the encounter complex were performed in this work. In the following manuscripts different types of Pc or Cyt_{c6} complexes are analyzed with respect to the analogy on the encounter complex model. The influence of electrostatic interactions as well as short range interactions are tested in detail for several individual complexes. As an analysis tool the combination of paramagnetic NMR measurements and ensemble dockings are used to study the extent of the encounter complex. Manuscript B focuses on the complex formation of different Pc derivatives with highly charged peptides. This resembles an extreme case, where the complex formation is dependent almost exclusively on the charge-charge interaction of the molecules. In manuscript C and D, the focus shifts to the more dynamic complex consisting of *Nostoc* Pc and *Nostoc* Cyt_f, which is stabilized by short range interactions. In order to investigate the charge influence in more detail for this type of complex, *Nostoc* Pc was exchanged by the less charged *Phormidium* Pc (PhPc) (Manuscript D). With electrostatic interactions being nearly absent in PhPc, the manner in which individual encounter complex states change can be compared to that of the native *Nostoc* Pc-Cyt_f complex. From the electrostatic interaction point of view, another intermediate complex—Cyt_{c6} and Cyt_f from *Nostoc*—is analyzed in manuscript E. The Cyt_{c6}-Cyt_f complex is less stabilized by charges than the one with Pc as a ligand.

It is demonstrated that the free proteins can be well described by MontyDock, which depend highly on electrostatic interactions. However, with increasing importance of hydrophobic or short range in-

teractions, the encounter complex lacks the pure electrostatic influence in the first stage and cannot be separated in a two-step kinetic model anymore. The MC ensemble then only partially agrees with the NMR data and gives rise of only single residues electrostatically contributing to the complex formation. The ranged charge-charge interactions mainly affect the complex formation by reorienting the proteins, which results in the hydrophobic patches of Pc and Cyt*f* being in the vicinity of each other. For complexes, where charges play a minor role, short range charge-charge interactions still contribute to the final complex stabilization. Thus, the whole complex formation cannot be separated in distinct steps anymore, but can be seen as a fluent transition between states. The encounter complex formation, until now, is described as a two-step model, where the association step is mainly described depending on electrostatic interactions. With the increased importance of short range interactions, it is questionable whether this model is still applicable for electron transfer complexes like those in *Nostoc*, where the encounter complex only partially can be described by electrostatic interactions. For this kind of complexes the charge-charge interaction seems to play a subordinated role. This is in conformity with findings, that electron transfer reactions are possible from multiple binding configurations.⁸⁶

Manuscript A: MontyDock - A Computational Tool for Mapping Transient Protein-Protein Complexes

In manuscript A, the program MontyDock is introduced and its main feature—the analysis of transient complex interactions—is explained. In MontyDock, a ligand moves randomly in the electric field of a receptor. The driving force for protein association is limited to the electrostatic attraction or repulsion. This enables to analyze the first part of the two-step kinetic model in the encounter complex formation. Here the interaction of two proteins is meant to be dominated by long range charge-charge interactions. With the introduction of a Metropolis Monte Carlo weighting, even energetically unfavorable states can be populated, which is a good approximation for a natural distribu-

tion of complexes. A productive complex formation is unlikely, if it is only diffusion-controlled. MontyDock introduces a directed approach to simulate a protein complex formation based on electrostatic interactions. In order to cover a high amount of possible interaction orientations, all degrees of freedom are sampled for the relative protein orientation. The program with all its analysis options was exemplified on cytochrome *c* (Cc) and cytochrome *c* peroxidase (CcP). Previous studies showed, that this complex is highly controlled by electrostatic interactions and therefore well suited to provide insight on the functionality of MontyDock.^{22,84,85} By means of MontyDock, the influence of ionic strength on protein binding, different oxidation states as well as the influence of the oxidation on the binding behavior can be demonstrated. This can be of major interest especially for electron transfer proteins, where the purpose of interaction, the electron transfer, is necessarily accompanied by a change of the oxidation state. With the help of different analysis tools provided by MontyDock, lowly populated ensemble changes can be visualized and investigated. This functionality is further demonstrated on the ternary complex consisting of two Cc molecules and one CcP molecule. It can be shown, that an oxidation state change of a bound Cc has an influence on the interaction of CcP with a second, unbound Cc. Upon changes in the oxidation state, a repulsion of the two Cc molecules can be observed, which suggests a consecutive model for an electron transfer.

In order to test the two-step model of an encounter complex, the program is applied to different types of electron transfer complexes. In manuscript B a system, where the complex formation is mainly stabilized by electrostatic interactions is discussed and thereby describing one side of the story for a complex formation. In the manuscripts C, D and E, complexes with less dependency to electrostatic interactions than the one shown in manuscript B are analyzed and discussed in relation to the two-step kinetic model of an encounter complex.

Manuscript B: An ensemble of rapidly interconverting orientations in electrostatic protein-peptide complexes characterized by NMR spectroscopy

In this work, three distinct types of Pc were analyzed for the encounter complex formation with short synthetic peptides consisting of four consecutive lysine molecules and the paramagnetic label 2,2,6,6-tetramethyl-N-oxyl-4-amino-4-carboxylic acid (TOAC). The short peptides are highly charged and therefore are well suited to analyze the effect of strong electrostatic interactions on the encounter complex formation. The chosen Pcs differ in their charge distribution. The two Pcs of the plant domain, *Populus nigra* (PoPc) and *Dryopteris crassirhizoma* (DPc), have a negatively charged eastern patch. DPc represents an intermediate role, where the negative charges are not only limited to the eastern patch but are distributed over the surface. The negative charges are located in the vicinity of the hydrophobic patch. PhPc was taken as a low charge density analogue. In order to analyze the complex formation and association behavior a combination of CSP and paramagnetic NMR spectroscopy as well as Monte Carlo and ensemble docking was used. The MC docking showed a good qualitative agreement with the residues majorly affected by paramagnetic NMR as well as with the CSPs. This indicates that the same regions of the Pc surface in the MC docking were sampled by the peptides in the experimental measurements, confirming the electrostatic relevance of the complex formation. The quantitative deviation of the paramagnetic measurements can be explained with the PREs being very sensitive to the local environment and in the MC sampling, the peptides cannot get as close to the surface of the protein. Another explanation could be that though the peptides are highly charged, also lowly populated hydrogen bond formations or transient short range interactions with the TOAC molecule are present. The visualization of the encounter complex with the ensemble docking showed that the predominant cluster of the TOAC molecules is for PoPc and DPc located around the charged eastern patches. On the contrary, the ensemble of PhPc showed primarily a random character of the association. This be-

havior was expected due to for *Phormidium* electrostatic interactions were reported to play a minor role.⁴¹ The interaction hotspots identified by NMR experiments, which are distributed across the surface of PoPc and DPc, could be very well correlated to the MontyDock simulations. This study indicates that the MontyDock simulations agree very well with the experimental measurements in the context of encounter complexes and shows that MontyDock can be well applied for encounter complexes assessing strong electrostatic interactions. In the following manuscripts C, D and E, details on the change in MC docking and the complex formation upon decrease of electrostatic interactions are presented.

Manuscript C: Role of hydrophobic interactions in the encounter complex formation of the plastocyanin and cytochrome *f* complex revealed by paramagnetic NMR spectroscopy

Manuscript D: Loss of electrostatic interactions causes increase of dynamics within the plastocyanin-cytochrome *f* complex

Both manuscripts, C and D, have the complex of Pc and Cyt*f* as their topic and are therefore presented together. In the former study, the native complex consisting of *Nostoc* Pc and *Nostoc* Cyt*f* is analyzed, while in the latter, this complex is related to the cross complex of PhPc with *Nostoc* Cyt*f* and the *Phormidium* wild type complex. A previous study⁷⁴ showed that formation of the *Nostoc* Cyt*f*-Pc complex is very dynamic. Therefore the encounter complex nor can be characterized by a single structure neither can the complex formation exclusively be described by electrostatic interactions.⁷⁴ Manuscript C connects to that and investigates the encounter complex by a higher surface sampling of Cyt*f* with paramagnetic labels and by inspecting the stabilizing effects of the encounter complex. Similar to the study in manuscript B, a combination of different NMR techniques, Monte Carlo and ensemble docking is used. It becomes clear that with electrostatic interactions alone, this encounter complex cannot be described. In MC docking Pc was widely spread over the surface of Cyt*f*, while a higher

density of encounters can be seen at the charged groups surrounding the heme binding site. The PREs agreed with the CSP measurements and showed that Pc binds at several spin label positions with the hydrophobic patch facing towards Cyt*f*. Comparing the NMR experiments with the MC docking, a good fit cannot be obtained. This indicates that the charge-charge interaction plays only a minor role for this complex formation. Nevertheless, residues in the vicinity of the interaction interface are matching some in the MC docking, which is a hint for electrostatic interactions contributing to the preorientation of Pc to the surface of Cyt*f*. The diffuse distribution of negative charges on Cyt*f* causes an orientation where the hydrophobic regions are the main stabilizing factor for the interaction. The shape of the encounter complex, modeled by ensemble docking shows that Pc samples Cyt*f* over a whole region leading to a higher populated ensemble in the heme region. The gradual increase of the hydrophobic interactions smoothly leads to a productive electron transfer complex at the heme site of Cyt*f*. The formation of the encounter complex cannot be separated in individual steps but is a smooth transition that blurs the distinction between the final complex and the encounter complex.

The balance of electrostatic and hydrophobic interactions is examined in more detail in manuscript D, where the cross complex of *Nostoc* Cyt*f* with the lowly charged variant PhPc was analyzed and related to the native complexes of *Nostoc* and *Phormidium*. The decrease of electrostatic interactions in the cross complex results in an even more diffuse encounter complex than for the *Nostoc* wild type. It can be seen that an increase of ionic strength leads to a decrease of the signal for residues in the region of the hydrophobic patch. This indicates that even for the electrostatically less influenceable complex of *Phormidium*, electrostatic stabilization contributes to the complex formation. This rises the question to what extent the charge-charge interactions are involved in the formation of the cross complex. A closer analysis of the MC ensemble shows that even though the average surface charge of PhPc is almost neutral, the affected residues are mainly located in the vicinity of the hydrophobic patch. In total, half of the MC

encounter contributes to the hydrophobic patch localization. A value of around 19% for the energetic complex stabilization arises from electrostatic interactions for the cross complex. Compared to that, a contribution of about 33% comes from electrostatic stabilization for the *Nostoc* complex under low ionic strength conditions. Under physiological conditions, this contribution is probably smaller. In a nutshell, this study showed that even small electrostatic interactions affect the association by bringing the hydrophobic parts in each others vicinity, though only to a minor extent. In encounter complexes, where hydrophobic interactions are the prominent stabilizing force, a distinction in several encounter complex steps is not applicable anymore.

Manuscript E: The dynamic complex of cytochrome c_6 and cytochrome f studied with paramagnetic NMR spectroscopy

In the two studies described above, the effect of hydrophobic interactions was analyzed and how the combination of hydrophobic and electrostatic interactions influences the encounter complex formation of electron transfer proteins. In manuscript E, the model of the two-step encounter complex formation is reevaluated on the basis of the complex formation of *Nostoc* Cyt c_6 and Cyt f . In order to avoid Cyt c_6 self-reduction reactions⁸⁷ during measurements, the mutants M58C and M58H of Cyt c_6 were created. The M58C variant leads to a reduction of the midpoint redox potential by 570 mV and therefore to an inert redox state. The mutations were successfully introduced and the crystal structures were solved, revealing no major structural changes compared to the wild type. PRE experiments, in accordance with the CSP measurements, showed that the similar residues of Cyt c_6 are affected for all spin label positions of Cyt f . In detail, the hydrophobic heme region of Cyt c_6 is always rotated towards the surface of Cyt f . The visualization of the NMR data with ensemble docking showed, that the Cyt f -Cyt c_6 complex cannot be described by a single well-defined complex, but instead by a pure encounter state. The results of the MC docking simulations do not show a good agreement with the ensemble docking results, emphasizing that the complex is less influenced

by electrostatic interactions. The MC shapes roughly follow the experimental values, though quantitatively did not represent the experimental data. Merging all data reveals that Cyt_6 preorients to Cyt_f upon approaching the surface and with this rapidly the importance of hydrophobic interactions grows. Comparing the Cyt_6 - Cyt_f complex to the Pc - Cyt_f one, the complex herein seems to be even more depending on hydrophobic interactions. Both studies regarding the *Nostoc* complexes, show a high contribution of hydrophobic interactions to the encounter complex.⁸⁸ This rises the question, whether the two-step model for the encounter complex can still be applied with the first step solely represented by electrostatic interactions.

Contributions to the Joint Publications

Manuscript A

Johannes M. Foerster[†], Ina Poehner[†] and G. Matthias Ullmann (2017) MontyDock - A Computational Tool for Mapping Transient Protein-Protein Complexes. *Manuscript submitted* SID: ct-2018-00003c

[†] Contributed equally to this work

The MontyDock calculations presented in the manuscript were done by myself. The calculations concerning the structure preparation and protonation probability calculations were done by I. Poehner. The results of my calculations were analyzed by me and G. Matthias Ullmann. Most parts of the manuscript were written by me, with the help of G. Matthias Ullmann. Ina Poehner wrote the manuscript part regarding the surface mapping of the contact maps.

Manuscript B

Jia-Ying Guan, Johannes M. Foerster, Jan W. Drijfhout, Monika Timmer, Anneloes Blok, G. Matthias Ullmann and Marcellus Ubbink (2014) An ensemble of rapidly interconverting orientations in electrostatic protein-peptide complexes characterized by NMR spectroscopy. *Chem-biochem.* 15 (4), 556-566.

My work consisted of the structure preparation and Monte Carlo docking simulations. The results of my calculations were analyzed by me and G. Matthias Ullmann. The protein purification and NMR experiments, as well as its analysis was done by Jia-Ying Guan. The manuscript was prepared by Jia-Ying Guan and Marcellus Ubbink. The manuscript part of the Monte Carlo simulations was written by me and G. Matthias Ullmann.

Manuscript C

Sandra Scanu, Johannes M. Foerster, G. Matthias Ullmann and Marcellus Ubbink (2013) Role of hydrophobic interactions in the encounter

complex formation of the plastocyanin and cytochrome *f* complex revealed by paramagnetic NMR spectroscopy. *J. Am. Chem. Soc.* 135 (20), 7681-7692.

The structure preparation and Monte Carlo docking simulations were performed by me. The results of my calculations were analyzed by me and G. Matthias Ullmann, and interpreted with regard the analysis of the whole encounter complex by Sandra Scanu, Marcellus Ubbink, G. Matthias Ullmann and me. The manuscript was prepared by Sandra Scanu and Marcellus Ubbink. The manuscript part of the Monte Carlo simulations was written by me and G. Matthias Ullmann.

Manuscript D

Sandra Scanu, Johannes M. Foerster, Monika Timmer, G. Matthias Ullmann and Marcellus Ubbink (2013) Loss of electrostatic interactions causes increase of dynamics within the plastocyanin-cytochrome *f* complex. *Biochem.* 52 (38), 6615-6626.

The structure preparation and Monte Carlo docking simulations were performed by me. The results of my calculations were analyzed by me and G. Matthias Ullmann, and interpreted with regard the analysis of the whole encounter complex by Sandra Scanu, Marcellus Ubbink, G. Matthias Ullmann and me. Parts of the paramagnetic NMR measurements were done by myself. Most NMR measurements were performed by Sandra Scanu as well as the ensemble docking and the analysis for the encounter complex. The manuscript was prepared by Sandra Scanu and Marcellus Ubbink. The manuscript part of the Monte Carlo simulations was written by me and G. Matthias Ullmann.

Manuscript E

Irene Díaz-Moreno, Rinske Hulsker, Pavol Skubak, Johannes M. Foerster, Davide Cavazzini, Michelina G. Finiguerra, Antonio Díaz-Quintana, Blas Moreno-Beltrán, Gian-Luigi Rossi, G. Matthias Ullmann, Navraj S. Pannu, Miguel A. De la Rosa and Marcellus Ubbink (2014) The dynamic

complex of cytochrome c_6 and cytochrome f studied with paramagnetic NMR spectroscopy. *Biochim. Biophys. Acta.* 1837 (8), 1305-1315.

My work consisted of the structure preparation and Monte Carlo docking simulations. The results of my calculations were analyzed by me and G. Matthias Ullmann. The manuscript was prepared by Marcelus Ubbink. The manuscript part of the Monte Carlo simulations was written by me and G. Matthias Ullmann.

Chapter 6

Manuscript A

MontyDock - A Computational Tool for Mapping Transient Protein-Protein Complexes

Johannes M. Foerster[†], Ina Poehner[†], G. Matthias Ullmann

Manuscript submitted

Manuscript ID: ct-2018-00003c

[†] Contributed equally to this work

MontyDock - A Computational Tool for Mapping Transient Protein-Protein Complexes

Johannes M. Foerster,^{†,¶} Ina Poehner,^{†,‡,¶} and G. Matthias Ullmann^{*,†}

[†]*Computational Biochemistry, University of Bayreuth, 95447 Bayreuth, Germany*

[‡]*current address: Molecular and Cellular Modeling, Heidelberg Institute for Theoretical Studies,
69118 Heidelberg, Germany*

[¶]*Contributed equally to this work*

E-mail: matthias.ullmann@uni-bayreuth.de

Phone: +49 (0)921 553545. Fax: +49 (0)921 553071

Abstract

MontyDock is a docking tool particularly well suited for analyzing formation of transient macromolecular complexes. The program applies a Monte Carlo docking strategy, where the ligand moves randomly in the electrostatic field of the receptor. By applying importance sampling, the major interaction sites are mapped. The software provides a broad range of analysis options which allow to relate the simulations to experimental data and to interpret them on a structural level. The application of MontyDock is exemplified by the electron transfer complex of cytochrome c peroxidase and cytochrome c from baker's yeast. The functionality of MontyDock and the visualization of simulation data are in particular demonstrated by studying the dependence of the association on ionic strength and on the oxidation state of the binding partner. Furthermore, microscopically, a repulsion of a second ligand can be seen in the ternary complex upon change of the oxidation state of the bound cytochrome c. The software is made available as open source software together with the example and can be downloaded free of charge from <http://www.bisb.uni-bayreuth.de/index.php?page=downloads>.

Introduction

Protein-protein interactions play a central role in many cellular processes such as signal transduction, gene regulation, and molecular bioenergetics. In general, protein association processes can be separated in two or three association steps. If the two molecules are far away from each other, mainly long range interactions like electrostatics are important for attractive forces bringing them in the vicinity of each other.¹ When the proteins are in contact distance, they can scan each others surfaces.² This phase of the association is often described by an ensemble of energetically favorable orientations. During this phase, short range interactions like hydrogen bonds and hydrophobic interactions become more and more important and contribute to the complex formation and stabilization of the final complex.^{3,4}

Various experimental techniques have been developed aiming to influence the association behavior.^{5,6} One approach is to stabilize the final complex at the known binding patch with short-range interactions like hydrogen bonds. This approach requires already a more or less detailed information on the complex structure. However, to get a first idea of binding, mainly long-range interactions need to be considered. Interactions like electrostatic attraction and repulsion mainly contribute to the primal complex formation and direct the proteins towards the binding areas. In order to analyze the stability of the complex, not only the association of the proteins but also the dissociation of the complex needs to be taken into account. Protein complexes can be very stable with low dissociation rates as for instance in the case of antigen-antibody complexes, or they can be very short living, having only a transient nature, as for instance in the case of electron-transfer protein complexes. Especially for electron transfer protein complexes it is important to find a reasonable balance between directional association to ensure correct pairing of partners and the possibility of quick exchange.^{7,8} Thus, the interactions between the electron transfer proteins are often very dynamic compared to other complexes.⁹⁻¹¹

The theoretical analysis of protein-protein association is a major challenge, especially for transient protein complexes. Many different approaches are commonly used.¹² An example of a widely used tool for evaluating binding affinities to a certain extent are docking simulations. For many

docking tools it is however a major challenge to predict yet unknown binding sites. Different tools have been applied for finding and validating first guesses of docking sites, for example molecular dynamics simulations,^{13–16} Brownian dynamics^{17,18} and Monte Carlo (MC) dockings involving random movements.^{19–21}

In this paper, we present the docking software MontyDock which allows to analyze protein-ligand binding using a Metropolis MC algorithm.²² The program can be used as a tool to identify binding sites in a rigid-body docking approach, especially if the association is driven by electrostatic interactions, which is often the case for protein-protein complexes. As a ligand, a small molecule, another protein or a nucleic acid can be used. For demonstrating the usage and the capabilities of the software, we analyze the well-studied complex of Cytochrome *c* Peroxidase (CcP) and cytochrome *c* (Cc) from *Saccharomyces cerevisiae*. CcP and Cc are located in the intermembrane section of mitochondria, where the ferrous form of Cc forms a complex with CcP to reduce hydrogen peroxide to water.^{23,24} The complex formation of the two electron transfer proteins is mainly driven by electrostatics. Thus, long-range charge-charge attraction allows for preorientation of proteins and thus limits the conformational search for the final complex formation.²⁵ In the first section of the paper, we introduce the algorithm used in MontyDock and the methods for analyzing the docking results. In the second part, we illustrate the analysis of the docking simulations. In particular, we analyze the association of CcP and Cc in different oxidation states and at different ionic strength values. In previous studies, we could demonstrate that the MC simulations can be well correlated with experimental data, namely with NMR studies, in which spin-labels causing paramagnetic relaxation enhancements have been attached to the proteins.^{4,11,25–30} The software and the source code are made available free of charge and can be downloaded from <http://www.bisb.uni-bayreuth.de/index.php?page=downloads>.

Software features

Monte Carlo Docking simulation. The program MontyDock is based on a docking method, which can be used to identify binding sites of macromolecular complexes,¹⁹ both in the context of protein-protein interactions and with other types of ligands. Since the program samples all translational and rotational degrees of freedom, MontyDock can even be applied if little is known about the binding site. The protein structures of the two partners, which we from now on refer to as ligand and receptor, are kept fixed during the MC simulation while translational and rotational movements are allowed. In the MC simulation, the ligand moves in the electrostatic field of the receptor. A general overview of the workflow of MontyDock can be seen in Figure 1. For computational reasons, it is advantageous to use the smaller docking partner as the ligand. The structures have to be provided in the PQRM-format, which is a modified PDB-format containing partial charges, radii, and masses of the atoms in the last three columns after the xyz coordinates. Such a PQRM-file can for instance be obtained from CHARMM³¹ results with the help of the converter program `psfcrd2pqr`, which is provided together with MontyDock. The electrostatic potentials are provided as OpenDX-files (a format for representing volumetric data) that can for instance be calculated by the program APBS,³² which solves the Poisson-Boltzmann equation for a heterogeneous dielectric environment.³³ The electrostatic potential is provided on two grids, a large coarse grid and a small fine grid. The fine grid allows a more detailed representation of the electrostatic potential when the two proteins are in closer proximity.

The overall flowchart of the MC algorithm is shown in Figure 2. A schematic representation of a MontyDock run is depicted in Figure 3. In the beginning of a simulation, the center of mass of the ligand is randomly placed on the surface of a sphere with the radius R_{in} . The center of mass of the receptor is placed at the center of this sphere. The electrostatic potential of the receptor at this sphere should be very close to zero or at least have a spherical symmetry to a good approximation in order to avoid a bias in the calculation stemming from the initialization. The interaction energy E_{int} between ligand and receptor is obtained by multiplying the charges $q_{n,lig}$ of the ligand with

the electrostatic potential $\Phi_{rec}(r)$ of the receptor (eq 1).

$$E_{int} = \sum_{n=0}^{N_{lig}} q_{n,lig} \cdot \Phi_{rec}(r_n) \quad (1)$$

where N_{lig} is the number of atoms of the ligand. The potential Φ_{rec} at the position r_n of the charge $q_{n,lig}$ is obtained by linear interpolation of the surrounding grid points. Subsequently, the ligand is randomly translated and rotated and the energy is reevaluated using eq 1. This random move is accepted or rejected according to the Metropolis criterion.³⁴ The next steps in the simulation are performed equivalently. In order to allow for a more detailed sampling of the interaction surface in the vicinity of the proteins, the MC moves are scaled down. In each MC move, it is checked whether the ligand reaches a specified distance R_{out} originating from the center of mass of the receptor. If the ligand reaches this distance, it is being considered to be too far away from the receptor for a proper interaction and the ligand is reinitialized at the starting surface with the radius R_{in} .

In addition to the evaluation of the electrostatic energy, it has to be made sure that the two molecules do not penetrate each other. The usage of a Lennard-Jones type of energy term would be computationally too expensive, since it would require $N_{lig} \times N_{rec}$ distance evaluations. Instead we use a so-called exclusion grid, which is a regular grid describing the shape of the receptor. The exclusion grid is generated by the program `make-excl` and saved as a binary file. Each point of the exclusion grid is flagged being either in the receptor or in the solvent. If during the simulation, the distance between ligand and receptor is such that the two molecules can collide, the exclusion grid is invoked. Thereby it is tested if a surface atom of the ligand is entering a grid cell of the exclusion grid that is occupied by the receptor. If that is the case, the move is rejected. The test is computationally very inexpensive and requires only a few rounding processes to identify the appropriate grid points. Moreover, the number of operations just scales linearly with the number of surface atoms of the ligand. In case of several consecutive rejections, it is assumed that the simulation got stuck in a (local) minimum and the ligand is reinitialized at the starting radius R_{in} .

The MC simulation proceeds till the maximum number of MC runs is reached. The ensemble is stored in a so-called orientation-file which contains information on the energy and the coordinate transformation required for generating the orientation from the initial structure. Each line of this orientation-file, which is a compressed ASCII file, contains one orientation. Storing the ensemble in this form has several advantages. Besides of saving storage space compared to a storage of full structures, it is also possible to use only a subset of the original structure, for instance only the prosthetic groups, for the generation of the orientation. This method substantially speeds up the analysis. Moreover, this orientation-file can for instance be resorted by energy, facilitating only to use orientations with energies below a certain threshold for further investigation. Thus, the way how MontyDock saves its generated ensemble gives a large flexibility for the subsequent analysis.

The runtime of the docking simulation does not depend on the size of the receptor but mainly depends on the atom number of the ligand. Hence it is recommended to use the smaller molecule as a ligand to speed up the simulation. In our example, a MontyDock simulation with a ligand of 1772 atoms and 10^3 MC runs, each run consisting of 2.5×10^4 MC steps, results in about 2.6×10^4 different orientations and takes approximately 20 minutes on a single core Intel® Core™ i7-4790 with a speed of 3.6 GHz.

Ensemble Analysis and Processing. The MontyDock suite provides a couple of tools to analyze the docking ensemble. The main utility for processing the docking results is the program `print-coor`. This program can on the one hand be used for generating representations of the ensemble, which can be visualized with molecular viewers like VMD³⁵ or PyMOL.³⁶ On the other hand, `print-coor` can reduce the encounter ensemble to highlight regions of interest. In the way MontyDock stores the docking ensemble, i.e. representing each structure of the ensemble by a linear transformation of the reference orientation, the distribution of ensemble structures can not only be inspected with respect to the receptor, but also vice versa with respect to the ligand. This possibility allows for the easy identification of high affinity binding patches on the surface of either ligand or receptor.

The program `print-coor` can analyze the ensemble contained in the orientation-file of Mon-

tyDock in different modes. One possibility is to generate each individual orientation of the ligand or of the receptor, which allows to analyze individual orientations directly. Another way to analyze the docking ensemble is to generate the center of mass of either the ligand around the receptor or the receptor around the ligand. This information can also be visualized as a density (in OpenDX-format). This density gives a fast overview of the docking results and can be inspected with many molecular viewers in different representations such as iso-contour plots or volume density slices. An example for such representations can be seen in Figure 4.

The program `print-coor` can not only help to visualize the docking ensemble, but also reduce it as follows. `Print-coor` creates subensembles, which make a deeper inspection of specific regions easier. For this purpose, the program determines the distance between the ligand and the receptor and only orientations in which the distance between any given atom of the ligand and receptor is smaller than a given threshold d are taken into account for this subensemble. Another way to alter the docking ensemble is to invoke a so-called inclusion grid. The idea thereby is similar to the procedure for excluding structures in case of protein collisions as described in the section of the MontyDock algorithm. Instead of excluding a ligand orientation entering the region that is occupied by the receptor, all orientations entering the regions that are defined by an inclusion grid are saved in a new subensemble. The inclusion grid presents a more efficient way for estimating whether certain regions of ligand and receptor are in proximity of each other upon complex formation. Like the exclusion grid, the inclusion grid can be generated by the program `make-excl`; the probe sphere radius needs to be adjusted to include larger regions around the receptor. This approach is computationally efficient and particularly of interest if the receptor molecule is non-spherical.

An overview of the docking simulations can also be obtained by creating a contact map histogram of all orientations, allowing to map major interaction regions to single atoms or amino acids. If the distance of an amino acid or atom of the ligand to an amino acid or atom of the receptor is smaller than a specified threshold, the counter in the 2D histogram is increased. The resulting histogram is written as an ASCII file which can be plotted with PyCoALA (the **P**ython

Contact Area Localization and Analysis tool, distributed with MontyDock) or any plotting tool like for instance GMT³⁷ or alternatively be further processed. For relating histograms of different simulations to each other, each histogram is normalized to the total amount of orientations saved in the individual MC simulation. This normalization can be easily performed with PyCoALA. This prevents misinterpretation by comparing simulations resulting in ensembles which are differing in size. For a better representation, the histogram is then scaled such that the maximal occurrence is set to one.

Together with MontyDock, with PyCoALA, we provide a tool that supports a representation of a contact map histogram on the protein structure using two different modes, namely (i) hot spot visualization or (ii) difference mapping by histogram subtraction. For both types of modes, contact map histograms are used as input, which can either have atomic or amino acid resolution. The contact surface regions of interest are represented on the respective protein surface. For this purpose, the residues that are involved in more than a certain number of contacts are colored on the surface of receptor and ligand by a PyMOL script.³⁶ If differences between various docking runs need to be examined, a subtraction of the results can help to identify the areas where the major differences occur. Positive values of the subtracted data are visualized in blue, negative values in red. Similar as for the heatmap representation, the contact differences can be mapped back on the protein surface. An example for a visualization with the PyCoALA back-mapping strategy is described in the example below.

An Example: Docking of yeast Cytochrome c Peroxidase (CcP) and yeast Cytochrome c (Cc)

In this section, we want to demonstrate the usage of the MontyDock suite by applying the software to the complex of CcP and Cc. This complex has been analyzed in many previous experimental and also some theoretical studies.^{38–43} Moreover, structurally, this complex was characterized by X-ray crystallography⁴⁴ and NMR.^{25,28,45} The complex shows many interesting features. Most notably, its association is largely governed by electrostatic interactions. Consequently, the asso-

ciation depends strongly on the ionic strength of the solution. Moreover, since the two reaction partners perform an electron transfer reaction, their association is influenced by the redox state of the reaction partners. In the following parts, we will demonstrate how the influence of these different parameters can be simulated, analyzed and visualized using MontyDock.

Influence of Ionic Strength on the Association. The complex formation of CcP and Cc depends strongly on electrostatic interactions and is therefore well suited to be studied with MontyDock. In order to determine how protein association is influenced by the salt concentrations, we studied complex formation at 0 mM, 10 mM, 20 mM, 50 mM and 100 mM ionic strength. A change in the ionic environment can result in a relocation of binding sites. An overview of such a docking analysis for different ionic strength values is shown in Figure 5. At lower ionic strength, two binding sites can be seen, a major one at a negatively charged surface area of CcP at the crystallographic binding site and a minor one at the site distal from the primary site.^{25,45} Inspection of the density of Cc around CcP and vice versa shows that the specific interaction surface of Cc shrinks with increasing ionic strength causing random encounters being more populated at higher ionic strength.

Since CcP and Cc are electron transfer proteins, the distance between the redox cofactors, namely the hemes, is an important parameter that can be analyzed in the docking ensemble. If one plots the probability of occurrence of encounter complexes in dependence of the shortest heme-to-heme distance of CcP and Cc and the corresponding energy, two minima can be identified at low ionic strength. These two minima represent two preferred binding orientations. With increasing ionic strength, the interaction weakens, though the previously identified hot spots still are present to an extent of about 20%. To inspect the interacting regions and the binding orientations in more detail, we generate contact map histograms. Such histograms for Cc and CcP at an ionic strength of 0 mM and 100 mM are depicted in Figure 6. Comparing the contact map histogram of 0 mM with the contact map of the crystal structure, it can be seen that the contacts of the crystal structure can be well reproduced (Figure 6A and B). Some additional contacts can be identified which arise from the analysis of the whole ensemble of complex formations. These contacts are located on

the surface in the vicinity of main interacting residues. When key residues are engaged in contacts stemming from many different orientations, neighboring residues also get involved to a certain extent. These more distributed interactions underline the transient character of the encounter complex. Interestingly, for the key residues D148 and K149 of the secondary binding site contacts with a population up to 11 - 25% can also be identified at low ionic strength. At 100 mM the total docking events decrease, while the relative specificity of individual amino acids can still be sustained, as is expected with increasing ionic strength (Figure 6C). However, subtracting the normalized contact map histograms at 0 mM and 100 mM from each other allows identification of differences between the docking ensembles at different ionic strength values and highlights which residues are less involved in contacts at higher salt concentrations (Figure 6D). With the combination of contact map histograms and energy distribution plots, we can show that the experimentally characterized loss of the lowly populated binding sites and the shift of the contact surface induced by a higher salt concentration^{39,42} can be reproduced by our MC simulations. The various analysis and plotting methods provide a powerful tool to connect experimental data with structural information derived from the MC simulation.

Influence of Cc heme oxidation states on the formation of the complex. An important parameter that influences the association of electron transfer proteins is their oxidation state. Experimentally, such an influence is often hard to access, since many experimental methods are only applicable in certain redox states. Simulations are therefore a good alternative for analyzing the influence of this parameter. In MontyDock, different redox states are modeled by assigning appropriate partial charges to the redox centers. For CcP, reduced Cc (Fe^{2+} state) is the natural electron donor. In order to guarantee a high turn over, it would be biologically advantageous if oxidized Cc (Fe^{3+}) would not bind as tightly. To probe this possibility, we simulated the association of CcP with both oxidized and reduced Cc at an ionic strength of 100 mM.

At first glance, the density distributions obtained for the two oxidation states of Cc are quite similar. However, the differences can be seen when the docking density obtained for oxidized Cc is subtracted from the density obtained for reduced Cc (Figure 7A and B). The binding site

of reduced Cc is more distributed, while the oxidized form with its more positive charge is more focused on the main interaction patch (Figure 7C and D).

While the difference map gives a reasonable approximation of the extent of the encounter complexes, the surface mapping tool PyCoALA grants a more detailed overview which residues are mainly contributing to the complex formation. Utilizing PyCoALA, the difference between contact map histograms for the different oxidation states can be plotted and mapped to the protein surface. The residues which show a difference are depicted in Figure 8. It becomes clear that many interacting residues on CcP are the same for both oxidized and reduced state of Cc (shown in purple), however, the contacting areas of Cc shift slightly. With the interaction difference mapped on the surface, two distinct separated regions on Cc can be identified. While K87 and K89 exclusively show many interactions with the peroxidase for reduced Cc, the interaction profile shifts to a region located around R13 and G83 for the oxidized Cc. Even if this change in the binding area is not dramatic, a difference between the oxidized and reduced ensemble can be seen, which may also influence the binding affinity.

Ternary Complexes. For the reduction of hydrogen peroxide to water, two electrons are required. Since CcP may bind several copies of Cc,^{42,46} two different mechanistic scenarios are possible. Either, two reduced molecules of Cc bind consecutively, i.e. the first Cc dissociates after the first electron transfer, giving way to a second Cc as an electron donor. Or, two reduced Cc molecules bind to CcP at the same time and the two electron transfers occur without the necessity for dissociation.

We use our MC docking method to test if two copies of Cc could bind simultaneously to CcP. For this purpose, we use the crystal structure of the complex between CcP and Cc⁴⁴ as the receptor for two different docking simulations. In one simulation, we assume that the bound Cc is reduced, in the other simulation it is oxidized. The ligand, i.e. Cc that binds to the binary complex of Cc and CcP, is always assumed to be reduced. For the simulation in which the bound Cc is reduced (Figure 9A), we obtained two binding spots: a small one near the interface between the bound Cc and CcP and a more extended one at the position of the secondary binding site.^{25,43} In the simula-

tion with the bound oxidized Cc, the small binding spot instead virtually disappears (Figure 9B). In order to visualize the difference between the simulations, the density for the oxidized state was subtracted from the density for the reduced state (Figure 9C). The difference density confirms the lower population of the binding spot near the interface, but also shows that the outer part of the secondary binding site has a higher population in the reduced state. It seems that if a Cc molecule is already bound to the main binding site of CcP, the secondary binding site is highly promoted for interactions. The small binding spot near the interface can be explained by the influence of the remaining potential of the main binding site of CcP. The potential of CcP is not completely neutralized by the bound Cc and attracts a second Cc molecule. The changed binding behavior of the ligand to the different redox states of the CcP-bound Cc can also be seen in the population diagrams, where the energies of the different orientations are related to the distances between ligand heme and CcP heme (Figure 9D) and their probability of occurrence. The distributions look similar for both oxidation states with the energetically favored encounters being higher populated and more widely spread in the reduced form. Apparently, the oxidized form of Cc repels another Cc molecule more strongly, which can be seen by less low energy encounters and a shift of the distribution to a higher heme-to-heme distance. Considering the plot of the distances between the ligand heme and CcP heme and the one between ligand heme and bound Cc heme (Figure 9E), we can clearly identify two clusters being predominant for the reduced simulation. Upon oxidation of the bound Cc, a shift of the hot spots from a short distance to a widely spread cluster equally distant to Cc and CcP heme can be observed. These findings agree with the idea that two binding Cc molecules repel each other electrostatically.^{40,47}

Our simulation leads to a view that the oxidation state in the ternary complex influences the binding behavior of a second Cc molecule. From our analysis we would support a model in which two Cc molecules bind at the same time. The binding is electrostatically favored and even promotes the binding of Cc to the secondary binding site, while for a consecutive binding the interactions are slightly destabilized, as long as an oxidized Cc molecule is still bound to CcP.

Conclusion

In this paper, we introduced the software suite MontyDock and demonstrated the usage of the software for protein-protein docking. The software uses a rigid-body docking approach and is particularly well suited for analyzing transient protein complexes for instance involved in electron transfer processes. In MontyDock, a MC method with a ligand moving randomly in the electrostatic field of a receptor, is applied. Acceptance of MC moves and thereby weighting of the electrostatic interaction energies is done according to the Metropolis MC criterion, resulting in a more natural distribution of the energies. Moreover, this approach allows for the identification of orientations, which are energetically not favored but still have a certain likelihood of occurrence. The program comes with a variety of analysis options which we demonstrate in this paper. These analysis tools help to relate the simulation results to experimental data which can then be explained on a structural level. In particular, with PyCoALA, we present an analysis method to rapidly link 2D contact histograms to 3D structures of receptor and ligand. This allows to analyze predicted protein protein interaction hot spots for their spatial properties and their electrochemical nature. In the context of growing efforts to find drugs modulating interactions between proteins, an application of such an analysis strategy in a drug design context is conceivable. Furthermore, the possibility of mapping difference hot spots between two related simulations to the protein surfaces provides insight in microscopic factors contributing to altered system behavior and may find application not only in the context of studying the impact of redox states but also, for example post-translational modifications or mutations, on a docking ensemble. In addition to the herein demonstrated analysis of protein-protein interaction, MontyDock is well suited to analyze also heavily charged molecules, like DNA or RNA. In order to cope with the longitudinal shape and the high charge density, the boundary conditions, namely the grid size and grid spacing, need to be adjusted accordingly. MontyDock also opens the possibility to linearly scale up the docking simulation. In the way MontyDock saves orientations, all MC runs are independent of each other. By this means, MontyDock can be executed multiple times in parallel with (the same parameters but) different starting seeds in order to get one big docking ensemble. In the application to the complex

of CcP and Cc, we demonstrate how for instance the ionic strength of the solution or oxidation states of proteins can influence the binding. Moreover, our simulations suggest that a second Cc molecule can be repelled by a bound Cc in the ternary complex and vice versa. These results give insight into how the complex formation is microscopically managed and how theoretical modeling can be used as a computational microscope for a better understanding of experimental data.⁴⁸

Material and Methods

Preparation of the Protein Models. The structure of the yeast complex Cc::CcP (PDB code: 2PCC, chain A and B)⁴⁴ was used in our analysis. We used the fitting method CHELPG⁴⁹ within ORCA⁵⁰ to derive partial charges for cofactor and amino acid states, which were not part of the CHARMM force field.⁵¹ For Cc, we derived charges for a c-type heme, typically linked to the protein by two cysteines and coordinated by histidine and methionine in both states, reduced (Fe^{2+}) and oxidized (Fe^{3+}).

We added hydrogens and energy minimized their position using CHARMM³¹ with steepest descent and conjugate gradient methods in cycles of 500 steps each. During minimization, heavy atoms were kept fixed. The protonation states of all amino acids were determined with titration studies using an in-house modified version of MEAD⁵² and GMCT.⁵³ The residues E11, D224, D235 and E267 in CcP were found to be protonated at pH=7 and were treated as such in further analysis.

Calculation of the Electrostatic Potentials. Electrostatic potentials were calculated using APBS.³² The ionic strength was set to 0 mM, 10 mM, 20 mM, 50 mM and 100 mM. A relative permittivity constant of 4 and 80 was used for the protein and the solvent, respectively. For the MC simulation, the electrostatic potential was represented on cubic grids with 225 grid points in each direction and a grid spacing of 2.0 Å and 1.0 Å for the coarse and the fine grid, respectively. The temperature was set 300 K. A radius of 1.4 Å and 2.0 Å was used for the ion and the solvent, respectively.

Docking Simulations with MontyDock. For protein-protein docking, we used MontyDock,

our software for performing MC docking simulations presented herein. We performed 10^4 runs with a total of 10^5 MC steps each for 0 mM, 10 mM, 20 mM, 50 mM and 100 mM ionic strength, respectively. The simulations were initiated with a center of mass separation of $R_{in} = 130 \text{ \AA}$ between receptor and ligand. A maximum center of mass separation of $R_{out} = 180 \text{ \AA}$ was allowed. The MC run was reseted after 50 consecutively rejected MC steps. The temperature was set to 300 K. The electrostatic potential maps with a grid spacing of 1 \AA and 2 \AA were used as inner and outer potential grids, respectively. A maximal displacement of 3 \AA and a maximal rotation of 5 rad were used. Calculations using OpenDX files (such as for instance subtractions of densities) were done using the tool DXMATH which is a part of APBS.³²

Analysis of the Docking Results. For analyzing all orientations of a docking ensemble, `print-coor` was used. For generating the contact map histograms, a distance of 8.0 \AA was chosen. The heme-to-heme distances were calculated with the subprogram `min-dist`, which determines the minimal distance between two given structures or structural elements. In order to compare MontyDock runs with a different number of orientations in the encounter ensemble, the histograms were normalized to the total number of orientations. In this way the histograms can be subtracted from each other to identify differences. For a better graphical representation, the histograms were scaled by the inverse of the highest absolute value of the histogram to lead to a maximal value of one for all histograms. For visualizing residues with major differences in the contacts between two simulations, PyCoALA and PyMOL were used.³⁶

Acknowledgement

This work was supported by the German Science Foundation (DFG; GRK1640).

References

- (1) Ubbink, M. The courtship of proteins: Understanding the encounter complex. *FEBS Lett.* **2009**, 583, 1060–1066.

- (2) Northrup, S. H.; Erickson, H. P. Kinetics of Protein-Protein Association Explained by Brownian Dynamics Computer Simulation. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 3338–3342.
- (3) Grove, T. Ž.; Ullmann, G. M.; Kostić, N. M. Simultaneous true, gated, and coupled electron-transfer reactions and energetics of protein rearrangement. *J. Inorg. Biochem.* **2012**, *106*, 143–150.
- (4) Scanu, S.; Foerster, J. M.; Ullmann, G. M.; Ubbink, M. Role of hydrophobic interactions in the encounter complex formation of the plastocyanin and cytochrome f complex revealed by paramagnetic NMR spectroscopy. *J. Am. Chem. Soc.* **2013**, *135*, 7681–92.
- (5) Hamp, T.; Rost, B. Evolutionary profiles improve protein-protein interaction prediction from sequence. *Bioinformatics* **2015**, *31*, 1945–1950.
- (6) Selzer, T.; Albeck, S.; Schreiber, G. Rational design of faster associating and tighter binding protein complexes. *Nat. Struct. Biol.* **2000**, *7*, 537–541.
- (7) Kostić, N. M. *Diprotein Complexes and Their Electron Transfer Reactions.*; In: Metal Ions in Biological Systems. Vol. 27, Ed. H. Sigel, 1991; pp 97–127.
- (8) Kostić, N. M. *Dynamic Aspects of Electron-Transfer Reactions in Metalloprotein Complexes.*; In: Metal-Containing Polymeric Materials. Plenum Press New, York, 1996; pp 491–500.
- (9) Hart, S. E.; Schlarb-Ridley, B. G.; Delon, C.; Bendall, D. S.; Howe, C. J. Role of charges on cytochrome f from the cyanobacterium *Phormidium laminosum* in its interaction with plastocyanin. *Biochemistry* **2003**, *42*, 4829–36.
- (10) Schlarb-Ridley, B. G.; Bendall, D. S.; Howe, C. J. Role of Electrostatic in the Interaction between Cytochrome f and Plastocyanin of the Cyanobacterium *Phormidium laminosum*. *Biochemistry* **2002**, *41*, 3279–3285.
- (11) Scanu, S.; Foerster, J. M.; Timmer, M.; Ullmann, G. M.; Ubbink, M. Loss of electrostatic

- interactions causes increase of dynamics within the plastocyanin-cytochrome f complex. *Biochemistry* **2013**, *52*, 6615–26.
- (12) Perkins, J. R.; Diboun, I.; Dessailly, B. H.; Lees, J. G.; Orengo, C. Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. *Structure* **2010**, *18*, 1233 – 1243.
- (13) Di Nola, A.; Roccatano, D.; Berendsen, H. J. C. Molecular dynamics simulation of the docking of substrates to proteins. *Proteins: Struct., Funct., Genet.* **1994**, *19*, 174–182.
- (14) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531–568.
- (15) Pfaendtner, J.; Volkmann, N.; Hanein, D.; Dalhaimer, P.; Pollard, T. D.; Voth, G. A. Key Structural Features of the Actin Filament Arp2/3 Complex Branch Junction Revealed by Molecular Simulation. *J. Mol. Biol.* **2012**, *416*, 148–161.
- (16) Perilla, J. R.; Goh, B. C.; Cassidy, C. K.; Liu, B.; Bernardi, R. C.; Rudack, T.; Yu, H.; Wu, Z.; Schulten, K. Molecular dynamics simulations of large macromolecular complexes. *Curr. Opin. Struct. Biol.* **2015**, *31*, 64–74.
- (17) Gabdoulline, R. R.; Wade, R. C. On the Contributions of Diffusion and Thermal Activation to Electron Transfer between *Phormidium laminosum* Plastocyanin and Cytochrome f : Brownian Dynamics Simulations with Explicit Modeling of Nonpolar Desolvation Interactions and Electron Transfer Even. *J. Am. Chem. Soc.* **2009**, *131*, 9230–9238.
- (18) Gabdoulline, R. R.; Wade, R. C. Brownian Dynamics Simulation of Protein-Protein Diffusional Encounter. *Methods* **1998**, *14*, 329–341.
- (19) Ullmann, G. M.; Knapp, E.-W.; Kostić, N. M. 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.

- (20) Liu, M.; Wang, S. MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 435–451.
- (21) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293–304.
- (22) Metropolis, N.; Ulam, S. The Monte Carlo method. *J. Am. Stat. Assoc.* **1949**, *44*, 335–341.
- (23) Altschul, A. M.; Abrams, R.; Hogness, T. R. Cytochrome c Peroxidase. *J. Biol. Chem.* **1940**, *136*, 777–794.
- (24) Beetlestone, J. The Oxidation of Cytochrome c by Cytochrome c Peroxidase. *Arch. Biochem. Biophys.* **1960**, *89*, 35–40.
- (25) Bashir, Q.; Volkov, A. N.; Ullmann, G. M.; Ubbink, M. Visualization of the Encounter Ensemble of the Transient Electron Transfer Complex of Cytochrome c and Cytochrome c Peroxidase. *J. Am. Chem. Soc.* **2010**, *132*, 241–247.
- (26) Scanu, S.; Förster, J.; Finiguerra, M. G.; Shabestari, M. H.; Huber, M.; Ubbink, M. The complex of cytochrome f and plastocyanin from *Nostoc* sp. PCC 7119 is highly dynamic. *Chembiochem* **2012**, *13*, 1312–8.
- (27) Díaz-Moreno, I.; Hulsker, R.; Skubak, P.; Foerster, J. M.; Cavazzini, D.; Finiguerra, M. G.; Díaz-Quintana, A.; Moreno-Beltrán, B.; Rossi, G. L.; Ullmann, G. M.; Pannu, N. S.; De La Rosa, M. A.; Ubbink, M. The dynamic complex of cytochrome c6 and cytochrome f studied with paramagnetic NMR spectroscopy. *BBA-Bioenergetics* **2014**, *1837*, 1305–1315.
- (28) Volkov, A. N.; Worrall, J. A. R.; Holtzmann, E.; Ubbink, M. Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 18945–18950.

- (29) Iwahara, J.; Clore, G. M. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **2006**, *440*, 1227–1230.
- (30) Tang, C.; Iwahara, J.; Clore, G. M. Visualization of transient encounter complexes in protein-protein association. *Nature* **2006**, *444*, 383–386.
- (31) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (32) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *P. Natl. Acad. Sci. USA* **2001**, *98*, 10037–10041.
- (33) Ullmann, G. M.; Bombarda, E. In *Protein Modelling*; Náray-Szabó, G., Ed.; Springer International Publishing: Cham, 2014; pp 135–163.
- (34) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **1953**, *21*, 1087–1092.
- (35) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–8, 27–8.
- (36) Schrödinger, LLC,
- (37) Wessel, P.; Smith, W. H. F.; Scharroo, R.; Luis, J.; Wobbe, F. Generic Mapping Tools: Improved Version Released. *Trans., Am. Geophys. Union* **2013**, *94*, 409–410.
- (38) Volkov, A. N.; Nicholls, P.; Worrall, J. A. The complex of cytochrome c and cytochrome c peroxidase: The end of the road? *BBA-Bioenergetics* **2011**, *1807*, 1482 – 1503.
- (39) Mei, H.; Wang, K.; McKee, S.; Wang, X.; Waldner, J. L.; Pielak, G. J.; Durham, B.; Millett, F. Control of formation and dissociation of the high-affinity complex between cytochrome c and

- cytochrome c peroxidase by ionic strength and the low-affinity binding site. *Biochemistry* **1996**, *35*, 15800–15806.
- (40) Nakani, S.; Vitello, L. B.; Erman, J. E. Characterization of Four Covalently-Linked Yeast Cytochrome c /Cytochrome c Peroxidase Complexes: Evidence for Electrostatic Interaction between Bound Cytochrome c Molecules. *Biochemistry* **2006**, *45*, 14371–14378.
- (41) Kornblatt, J. A.; English, A. M. The binding of porphyrin cytochrome c to yeast cytochrome c peroxidase: A fluorescence study of the number of sites and their sensitivity to salt. *Eur. J. Biochem.* **1986**, *155*, 505–511.
- (42) Zhou, J.; Hoffman, B. Stern-volmer in reverse: 2:1 stoichiometry of the cytochrome c-cytochrome c peroxidase electron-transfer complex. *Science* **1994**, *265*, 1693–1696.
- (43) Northrup, S.; Boles, J.; Reynolds, J. Brownian dynamics of cytochrome c and cytochrome c peroxidase association. *Science* **1988**, *241*, 67–70.
- (44) Pelletier, H.; Kraut, J. Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c. *Science* **1992**, *258*, 1748–1755.
- (45) Volkov, A. N.; Ubbink, M.; Van Nuland, N. A. J. Mapping the encounter state of a transient protein complex by PRE NMR spectroscopy. *J. Biomol. NMR* **2010**, *48*, 225–236.
- (46) Stemp, E. D.; Hoffman, B. M. Cytochrome c Peroxidase Binds Two Molecules of Cytochrome c: Evidence for a Low-Affinity, Electron-Transfer-Active Site on Cytochrome c Peroxidase. *Biochemistry* **1993**, *32*, 10848–10865.
- (47) Page, T. R.; Hoffman, B. M. Control of Cyclic Photoinitiated Electron Transfer between Cytochrome c Peroxidase (W191F) and Cytochrome c by Formation of Dynamic Binary and Ternary Complexes. *Biochemistry* **2015**, *54*, 1188–1197, PMID: 25629200.
- (48) Lee, E. H.; Hsin, J.; Sotomayor, M.; Comellas, G.; Schulten, K. Discovery Through the Computational Microscope. *Structure* **2009**, *17*, 1295–1306.

- (49) Breneman, C. M.; Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **1990**, *11*, 361–373.
- (50) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (51) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (52) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environments: First International Conference, ISCOPE 97 Marina del Rey, California, USA December 8–11, 1997 Proceedings*; Ishikawa, Y., Oldehoeft, R. R., Reynders, J. V. W., Tholburn, M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1997; pp 233–240.
- (53) Ullmann, R. T.; Ullmann, G. M. GMCT : A Monte Carlo simulation package for macromolecular receptors. *J. Comput. Chem.* **2012**, *33*, 887–900.

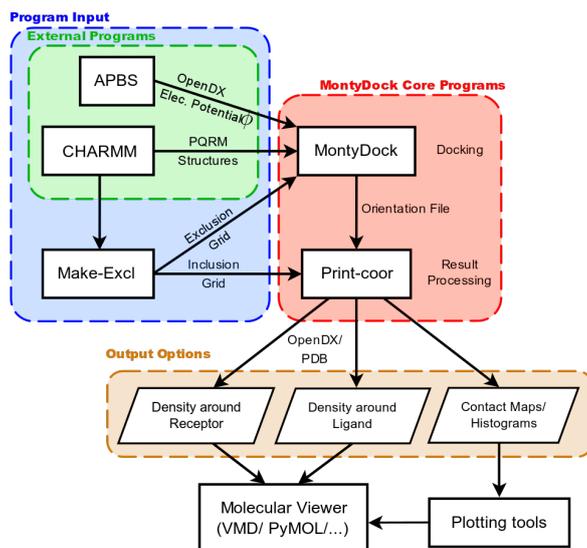


Figure 1: Overall workflow diagram for the usage of MontyDock: MontyDock needs structural information, the electrostatic potentials and an exclusion grid as a starting point. After the docking simulation, the results are processed with `print-coor`. Depending on the chosen analysis, the docking results can either be visualized with the focus on the receptor or on the ligand. Alternatively, the whole simulation can be analyzed by contact map histograms.

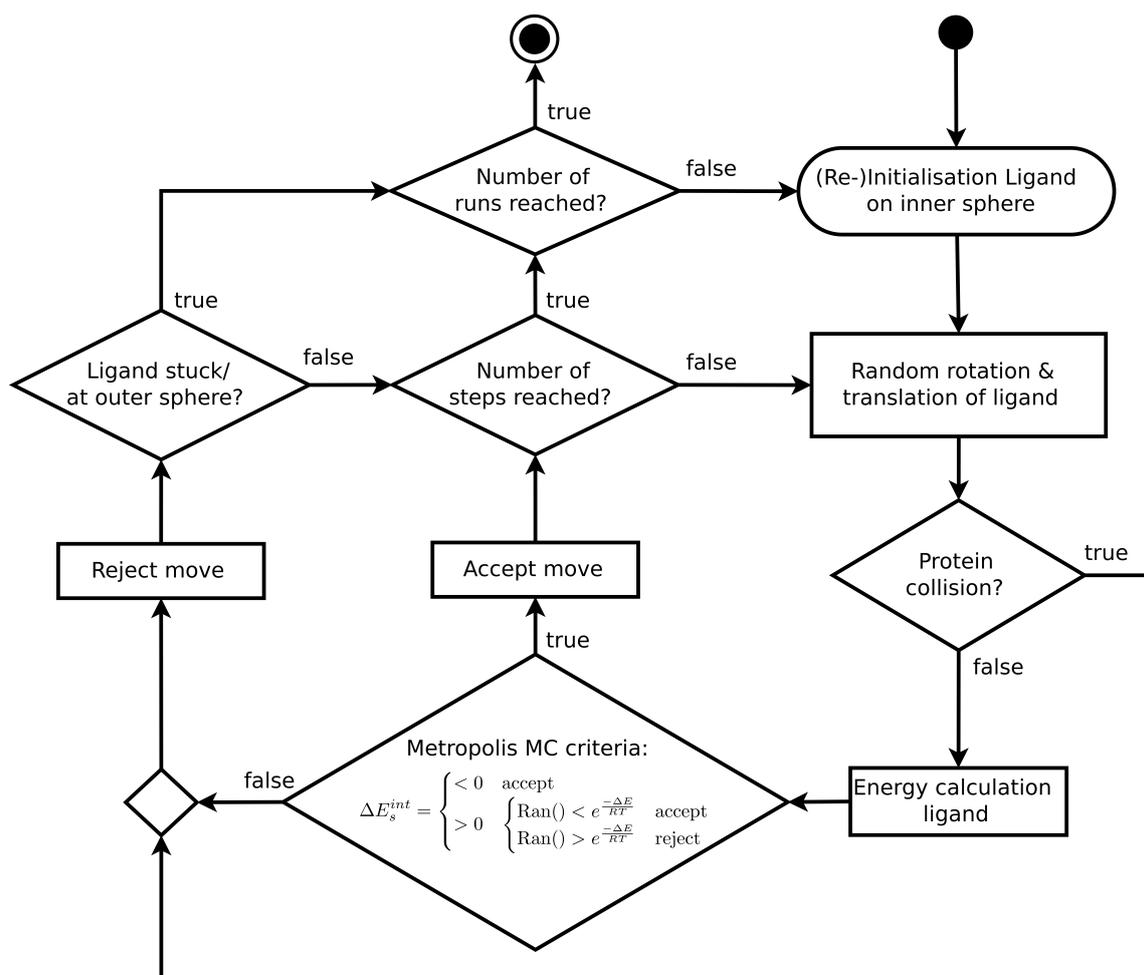


Figure 2: Flowchart of the algorithm implemented in MontyDock: The simulation starts by placing the ligand on the surface of the inner sphere (see Figure 3). Then the ligand moves randomly in space. If the proteins do not collide, the ligand energy is calculated and the step is evaluated by the Metropolis MC criterion. This cycle is repeated until the maximum number of steps is reached.

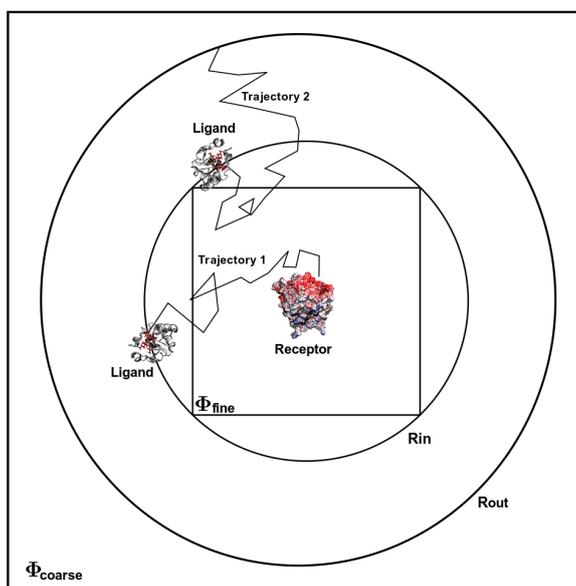


Figure 3: Schematic illustration of two exemplary MC runs. The inner and outer boxes around the receptor represent its electrostatic potential on a fine and coarse grid, respectively. In trajectory 1, the ligand starts at radius R_{in} and proceeds to the surface of the receptor with each MC step visualized by connected lines. In trajectory 2, the ligand starts at a different position on R_{in} and moves to the radius R_{out} , where the MC run is aborted.

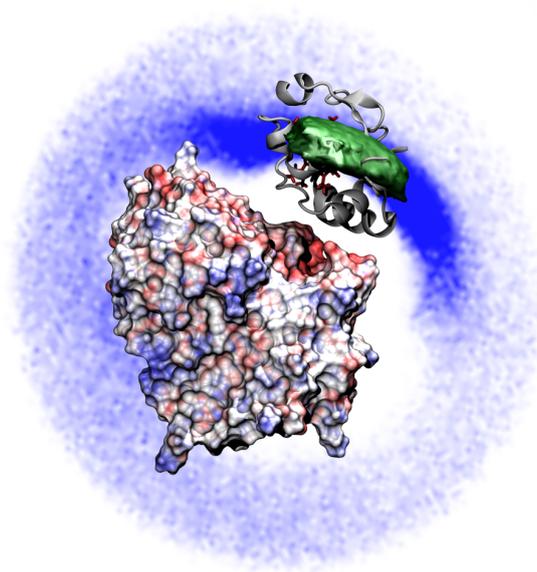


Figure 4: Overview of different representations of the center of mass density obtained from the MC docking simulations. CcP is represented with the electrostatic potential mapped on the surface ranging from -5 (red) to 5 kcal/e° (blue). Cc is shown in silver with the heme highlighted in red. The green isosurface represents the positions where the center of mass of Cc was found at least 100 times in the ensemble. The blue color in the background illustrates a slice through the volume of the docking ensemble ranging from 0 (white) to at least 40 (dark blue) orientations.

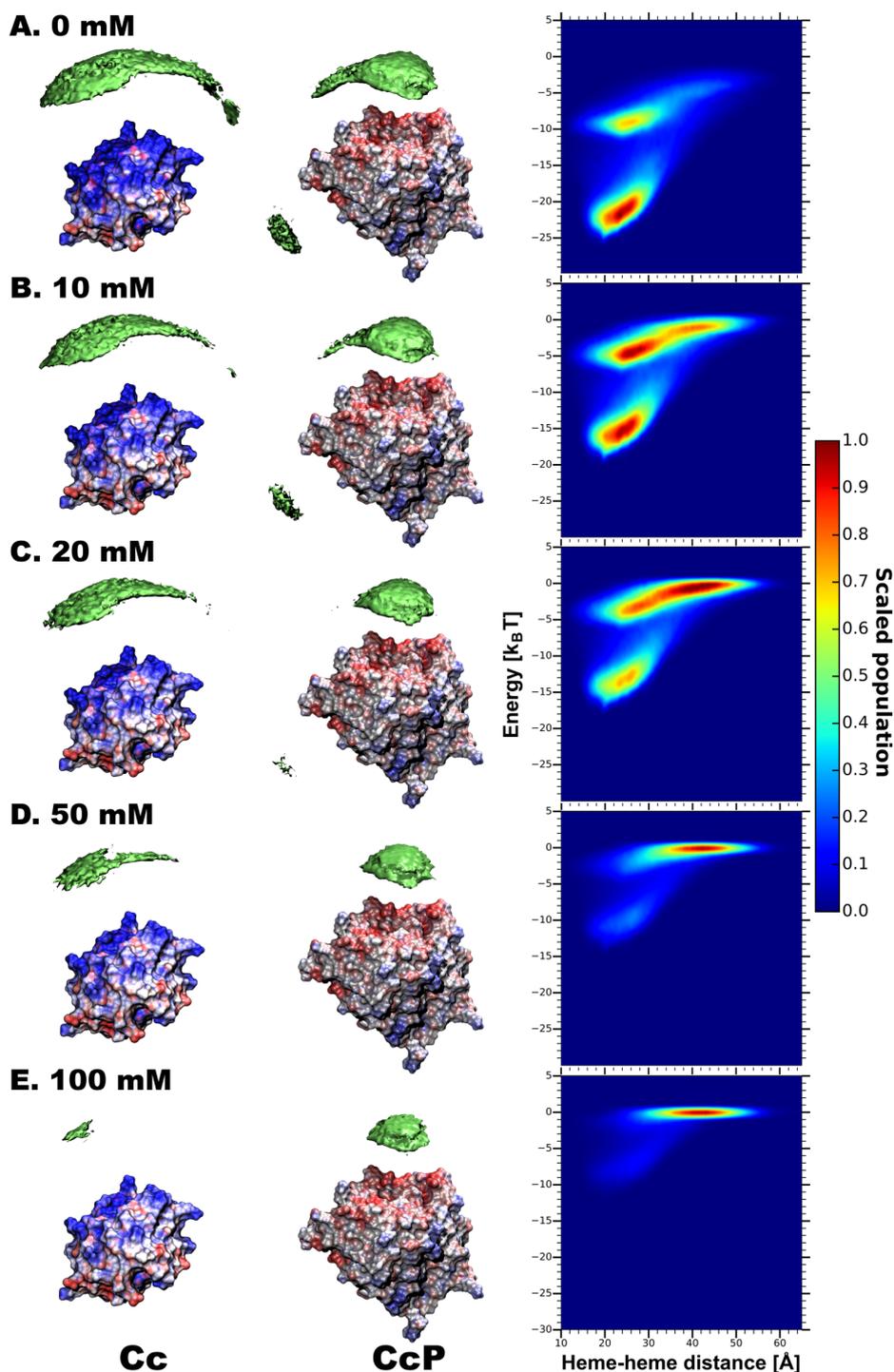


Figure 5: MC docking results for Cc (left) and CcP (middle) in dependency of ionic strength (0, 10, 20, 50, 100 mM). The green isosurfaces represent the count of the center of mass of CcP (left) and Cc (middle). The cutoff was set to 150 for the representation. The electrostatic potential is mapped on the surface of CcP and Cc ranging from -5 (red) to 5 kcal/e $^\circ$ (blue). The plots on the right show the scaled normalized population of the docking ensemble in dependency of the heme-to-heme distance of Cc and CcP and the energy of the orientation.

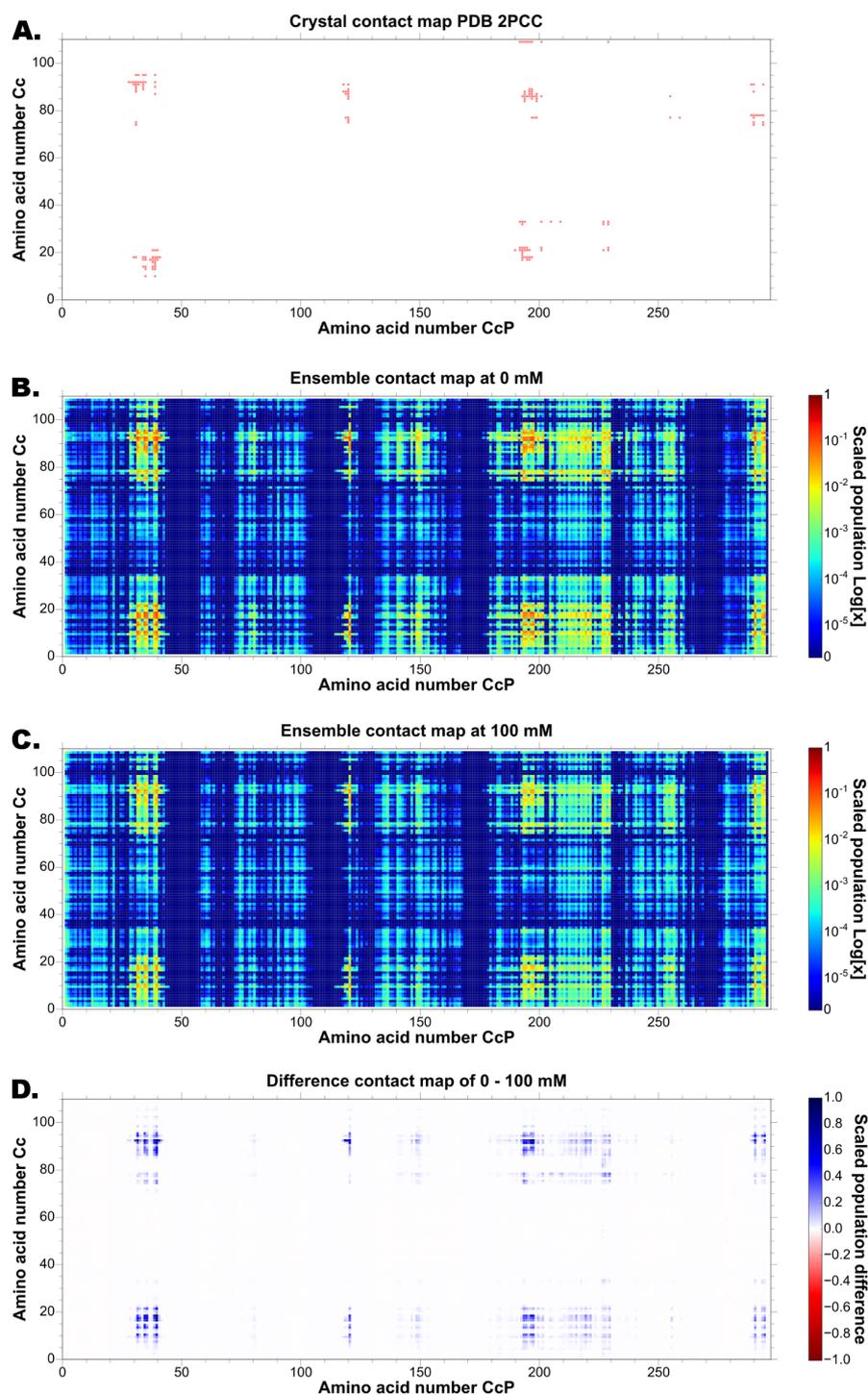


Figure 6: Contact map for the crystal structure (PDB: 2PCC) (A) and contact map histograms of the docked CcP-Cc complexes at 0 mM (B) and 100 mM (C) ionic strength. The difference between the contact map histograms at 0 mM and 100 mM can be seen in part D. The x- and y-axis correspond to the amino acids of the respective protein. Any contact within a distance of 8 Å or less is counted and counts are color coded as shown in the legends.

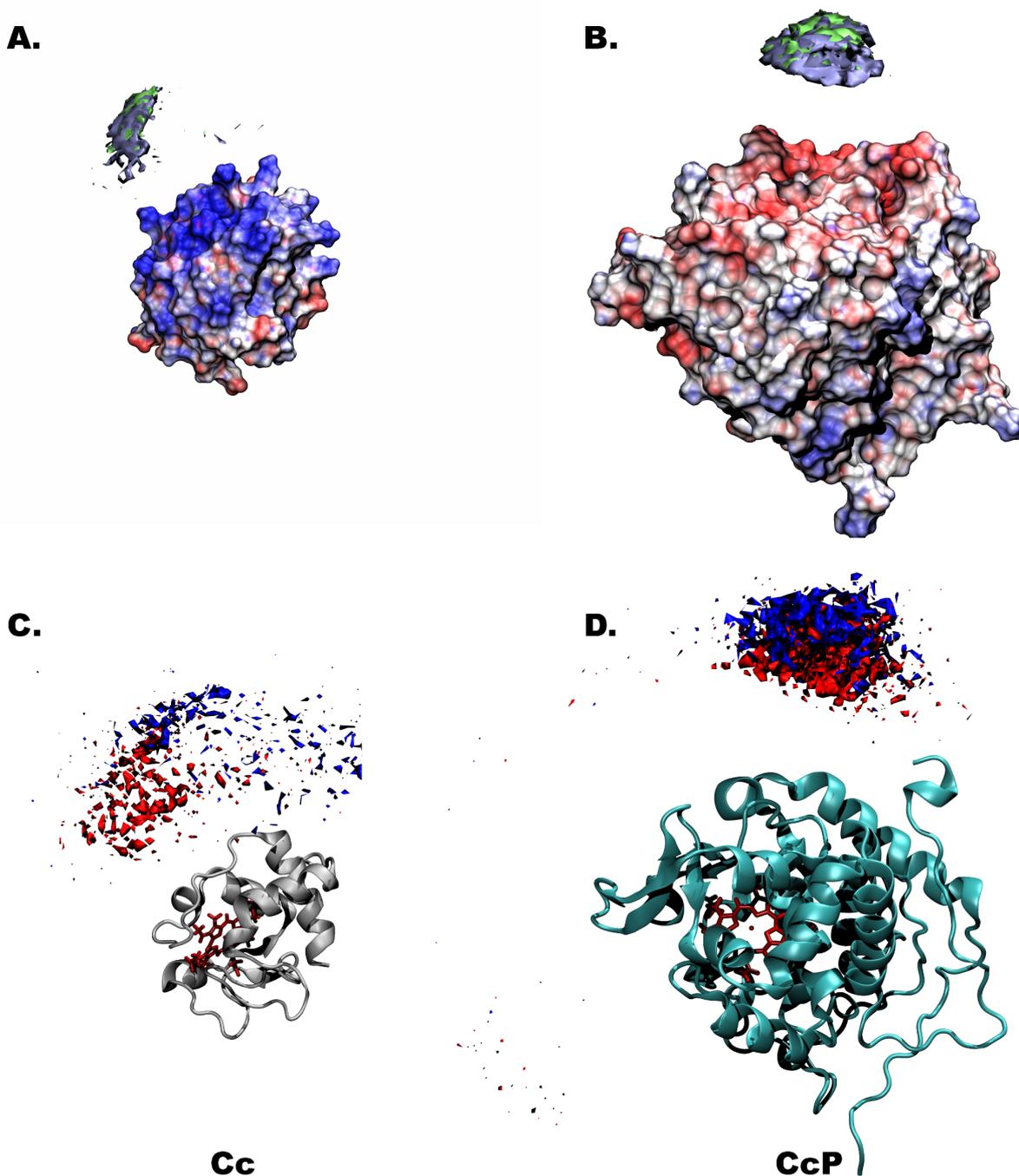


Figure 7: MC docking results for different oxidation states of Cc with CcP at an ionic strength of 100 mM. The docking ensemble is shown as original ensemble (A, B) and as a difference between the Fe^{2+} and the Fe^{3+} (C, D) ensemble: The docking density of CcP around Cc (A/C) and vice versa (B/D). The total docking density is highlighted in green for Fe^{2+} and in iceblue for Fe^{3+} . The difference density in red shows the region with a difference of at least -6 (C) or -40 (D) docking events, i. e. docking events of the Fe^{3+} oxidation state were predominant in these regions. A difference of +6 (C) or +40 (D) is shown as a blue isosurface.

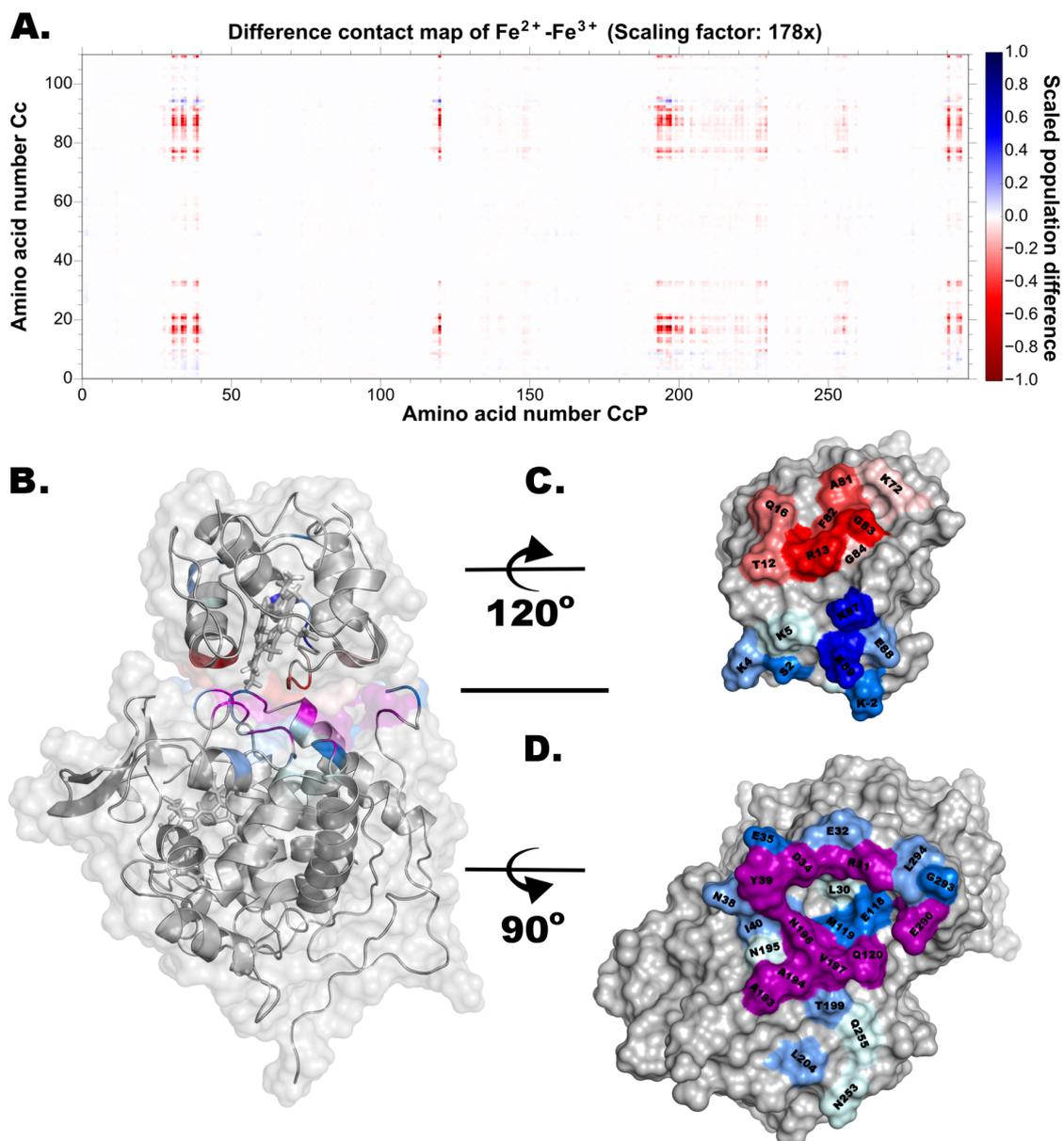


Figure 8: The differences between oxidized and reduced state of Cc visualized with the most affected residues highlighted: Part A: Difference of the contact map histogram between amino acids of CcP and Cc after subtraction of Cc Fe^{3+} oxidation state from Cc Fe^{2+} oxidation state results. Part B: Overview of the complex between Cc and CcP with the regions with the highest differences between both oxidation states color coded. Residues highlighted in red have the most negative values in the subtraction, i. e. they are more populated for Fe^{3+} oxidation state, while residues in blue show the most positive values. Residues highly populated in both oxidation states are colored purple. The major binding region is highlighted separately on the surfaces of Cc (C) and CcP (D). The amino acids are labeled like in the crystal structure.

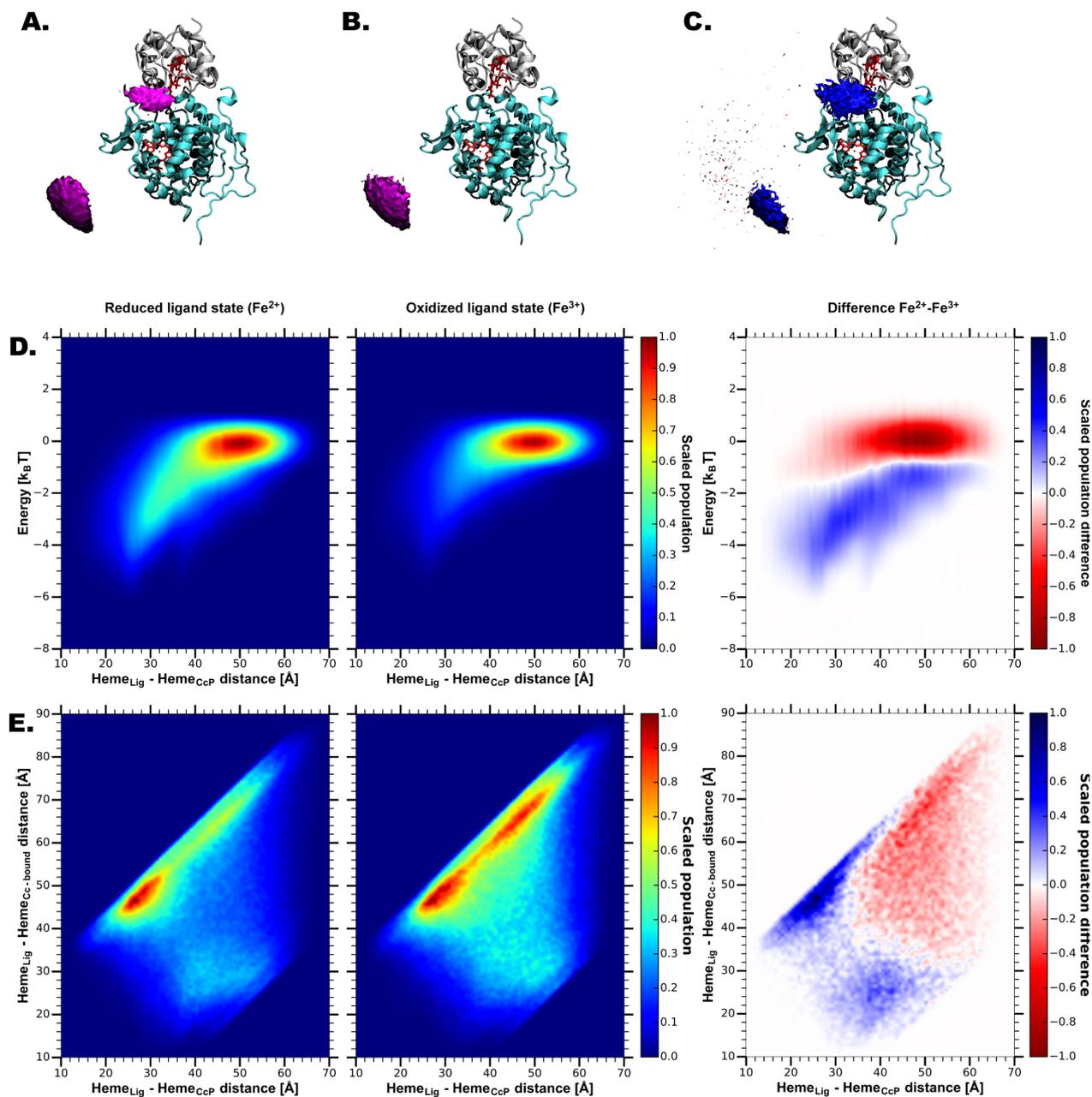


Figure 9: Docking results for the ternary complex consisting of the crystal complex CcP-Cc (PDB: 2PCC) and an additional Cc molecule as ligand. In one simulation the bound Cc is oxidized (A), in the other it is reduced (B). The isosurfaces represent half the maximal density of all orientations. The difference between the ensembles ($\text{Fe}^{2+} - \text{Fe}^{3+}$) are shown in part C, with the positive and the negative isosurface colored blue and red respectively. For each simulation the shortest heme-to-heme distance for the ligand to the CcP is plotted versus the electrostatic interaction energy (D). Another way to identify orientational changes is to analyze the population of the various heme-to-heme distances from the ligand Cc to the complex Cc as a function of the heme-to-heme distance of the ligand Cc to CcP. In all distance mappings, the ensemble distribution seems to have a clear edge at the top border. This edge is caused by the maximal separation of the ligand and the receptor considered in the analysis and is not a feature of the encounter complex.

Chapter 7

Manuscript B

An ensemble of rapidly interconverting orientations in electrostatic protein-peptide complexes characterized by NMR spectroscopy

Jia-Ying Guan, Johannes M. Foerster, Jan W. Drijfhout, Monika Timmer, Anneloes Blok, G. Matthias Ullmann and Marcellus Ubbink

ChemBioChem. 2014, 15 (4), 556-566

doi: 10.1002/cbic.201300623

Reprint with permission of *ChemBioChem*.

© 2014 Copyright by Wiley-VCH Verlag GmbH & Co. KGaA.

DOI: 10.1002/cbic.201300623

An Ensemble of Rapidly Interconverting Orientations in Electrostatic Protein–Peptide Complexes Characterized by NMR Spectroscopy

Jia-Ying Guan,^[a] Johannes M. Foerster,^[b] Jan W. Drijfhout,^[c] Monika Timmer,^[a] Anneloes Blok,^[a] G. Matthias Ullmann,^[b] and Marcellus Ubbink^{*[a]}

Protein complex formation involves an encounter state in which the proteins are associated in a nonspecific manner and often stabilized by interactions between charged surface patches. Such patches are thought to bind in many different orientations with similar affinity. To obtain experimental evidence for the dynamics in encounter complexes, a model was created using the electron transfer protein plastocyanin and short charged peptides. Three plastocyanins with distinct sur-

face charge distributions were studied. The experimental results from chemical shift perturbations, paramagnetic relaxation enhancement (PRE) NMR, and theoretical results from Monte Carlo simulations indicate the presence of multiple binding orientations that interconvert quickly and are dominated by long-range charge interactions. The PRE data also suggest the presence of highly transient orientations stabilized by short-range interactions.

Introduction

According to current models, formation of a specific protein complex is preceded by that of an encounter complex.^[1] It is believed that, in this state, the partners assume multiple orientations to enhance the probability of finding the specific binding site.^[2] Often in the encounter complex, charge–charge interactions dominate, whereas the specific (final) state is stabilized by various short-range interactions. The assumed presence of multiple orientations in the encounter state is based on the theoretical notion of charged surface patches. Like Velcro,^[3] such patches can bind in many orientations with similar energy and thus all are assumed to be populated. The presence of multiple orientations and the dynamic exchange between them in the charge-driven encounter state is, however, not easy to demonstrate experimentally.

The aim of this study was to create a pure, charge-driven encounter state and demonstrate the existence of a rapidly changing set of binding orientations. We chose to study the complex of plastocyanin (Pc) and short, charged peptides (Lys₄), assuming that the interaction would be dominated by the strong positive charges of the peptides. The peptides are

an artificial binding partner, so Pc will not have an optimized binding site, and a specific complex is unlikely to be formed.

Pc is a type I blue copper protein involved in the electron transport process in oxygenic photosynthesis, functioning as an electron carrier between cytochrome *f* (Cyt *f*) of the *b₆f* complex and P700⁺ of photosystem I (PSI). Structures are available for Pc from various plants and bacteria.^[4–11] One of the histidine residues that is a copper ligand is considered to be the electron entrance, that is, it provides a strong coupling pathway toward the copper. It is located at the so-called “northern” side of the protein, within a hydrophobic patch. Pc is acidic in higher plants^[5,9–12] and green algae,^[13–15] possessing two highly conserved negatively charged surface regions (acidic patches) formed by amino acids at positions 42–44 and 59–61 on the so-called “eastern” side. A typical example of *Populus nigra* Pc (PoPc) is shown in Figure 1A. Compared to typical plant Pcs, the structure of Pc from the fern *Dryopteris crassirhizoma* (DPc) has the same global structure (Figure 1B), but a large acidic arc extends to the northern side surface near the hydrophobic patch, resulting in distinct electrostatic properties.^[7] In cyanobacteria, Pc can also be almost neutral^[6,16] such as in *Phormidium laminosum*^[6] (Figure 1C), or basic, such as in *Nostoc* sp. PCC 7119.^[17–19]

Charged peptides have proved useful for studying interacting sites in electron transfer proteins, including Pc, Cyt *f*, and Cyt *c*.^[20–26] Experimental results showed that positively charged polylysine peptides interact with the clustered acidic residues on Pc and competitively inhibit electron transfer from Cyt *c* or Cyt *f* to Pc.^[20,22] This competitive inhibition was explained by neutralization of charges by the formation of the Pc–peptide complexes.^[20] The binding of polylysine peptides to Pc can also subtly perturb the active-site geometry and the redox potential.^[20,23] Little information, however, is available for the

[a] Dr. J.-Y. Guan, Dr. M. Timmer, A. Blok, Prof. Dr. M. Ubbink
Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University
Einsteinweg 55, 2333 CC Leiden (The Netherlands)
E-mail: m.ubbink@chem.leidenuniv.nl

[b] J. M. Foerster, Prof. Dr. G. M. Ullmann
Structural Biology/Bioinformatics, University of Bayreuth
Universitätsstrasse 30, 95447 Bayreuth (Germany)

[c] Dr. J. W. Drijfhout
Department of Immunohematology and Blood Transfusion
Leiden University Medical Center
Albinusdreef 2, 2333 ZA Leiden (The Netherlands)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201300623>.

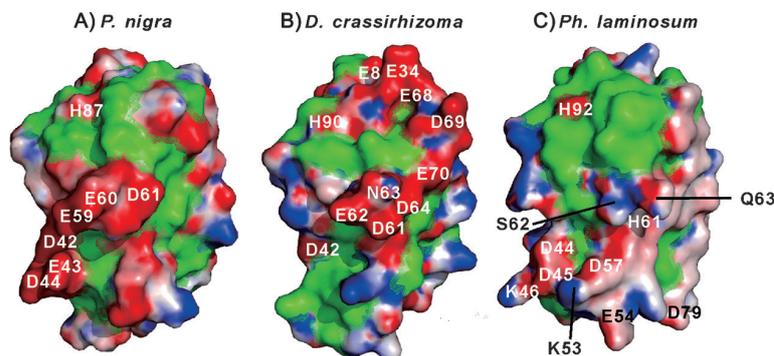


Figure 1. Electrostatic potential maps of Pc surface models (PDB IDs: 1TKW,^[32] 1KDI,^[7] and 2Q5B). The surface colors correspond to the electrostatic potential calculated by the program APBS^[63] at an ionic strength of 10 mM, pH 6.5, 300 K, to match the experimental conditions. The electrostatic potentials are colored and contoured from -8 (intense red) to $+8$ kT/e (intense blue). Hydrophobic residues (Ala, Val, Ile, Leu, Phe, Pro, Tyr, and Met) are colored in green. Several relevant residues are labeled. Pictures were generated using PyMOL.^[64]

binding interface and the underlying degree of dynamics in the interaction.

Paramagnetic relaxation enhancement (PRE) NMR spectroscopy has been used as a sensitive tool to detect lowly populated intermediates in biomolecular complexes.^[27,28] The large magnetic moment of the unpaired electron from the paramagnetic center causes relaxation of nuclear spins in the vicinity. This effect diminishes very rapidly, being proportional to the inverse sixth power of the distance between the electron and the nucleus. TOAC (2,2,6,6-tetramethyl-*N*-oxyl-4-amino-4-carboxylic acid) has been shown to be useful for PRE NMR studies of protein–peptide interactions.^[29] One of the advantages of TOAC over side chain-attached spin labels is that TOAC can be directly incorporated into the peptide backbone in automated peptide synthesis. There has been growing interest in using TOAC in peptide–protein and peptide–nucleic acid interactions and in combination with other physical techniques, such as electron paramagnetic resonance (EPR), circular dichroism (CD), fluorescence, Fourier-transform infrared (FT-IR), NMR, and X-ray crystallography, to understand molecular interactions.^[30]

In this study, the transient complexes formed by tetralysine peptides and three different Pcs were studied using chemical shift perturbation (CSP) analysis, PRE NMR spectroscopy, ensemble docking, and Monte Carlo (MC) simulations. The CSP data corresponded well with the electrostatic MC docking calculations, clearly showing that binding is dominated by charge interactions. The PRE data indicated that, within the electrostatic ensemble, the peptides assume multiple orientations in a dynamic fashion. The PRE data also provide evidence for the presence of orientations that are slightly more favored than expected from pure charge–charge interactions, perhaps due to transient hydrogen bond formation with TOAC or weak hydrophobic interactions. Overall, the experimental and simulation results provide direct evidence for dynamics in an encounter complex dominated by charge–charge interactions.

Results

Backbone assignments

To study the three Pcs by NMR, the proteins were isotopically labeled with ^{15}N for PRE measurements and $^{15}\text{N}/^{13}\text{C}$ for resonance assignments. To eliminate the paramagnetic effect of Cu^{2+} , Zn-substituted Pc was used. For DPc and PoPc, backbone amide resonances were assigned by using HNCACB experiments on $^{13}\text{C}/^{15}\text{N}$ -labeled proteins. The assignments of Cu^{I} -DPc (BMRB code 7370)^[31] and Cu^{I} -PoPc (BMRB code 4019) were used as the starting points. Data for backbone assignments (H, N, C_{α} , C_{β}) have been deposited to BMRB under codes 19236 (DPc) and 19247 (PoPc). Assignments of Zn-substituted PhPc were kindly provided by Dr. Sandra Scanu (Leiden University). For DPc, the resonance of Ser92 was not found in the spectra. For PoPc, some residues close to the N terminus have double peaks. These double resonances exist for Ile1, Asp2, Val3, Ser20, Ile21, Ser22, Pro23, Gly24, Glu25, Lys26, Ile27, Val28, Lys30, Met57, Thr69, Phe70, Glu71, Val72, Leu74, and Gly78. Similar observations were described for Cd-PoPc.^[32] The double signals were attributed to partial processing of the N-terminal methionine in the bacterial cytoplasm, as these residues are located near the N terminus in the three-dimensional structure of the protein.^[32]

Chemical shift perturbations

To study the interaction of Pc with lysine peptides, four types of peptides were used. For the PRE experiment described below, a TOAC residue (X) was introduced at the N or C terminus (X-Lys₄ and Lys₄-X). As controls for the introduction of TOAC, Ala-Lys₄ and Lys₄-Ala were also used. First, the interactions of these peptides with the three Pc variants were studied using CSP analysis.

Each ^{15}N -Pc was titrated with the four peptides individually in a low ionic strength buffer ($I=10$ mM), and $^1\text{H},^{15}\text{N}$ HSQC spectra were acquired at each titration point. For these studies, TOAC was reduced to eliminate its paramagnetic effects. Addition of the peptides gave rise to small CSPs in the $^1\text{H},^{15}\text{N}$ HSQC spectra of all Pcs, with maximum observed average shifts ($|\Delta\delta_{\text{ave}}|$) of 0.07 ppm for PoPc, 0.05 ppm for DPc, and 0.01 ppm for PhPc (Figure 2). Single, averaged resonances were observed for all amides, indicating fast exchange between the free and bound Pc on the NMR time scale. Binding maps, obtained by coloring the protein residues according to the size of CSP, show similar patterns for different peptides for the same Pc (Figure 2 for Lys₄-X and Figure S1 for the other peptides). The similar patterns observed for Lys₄-X and X-Lys₄ indicate that

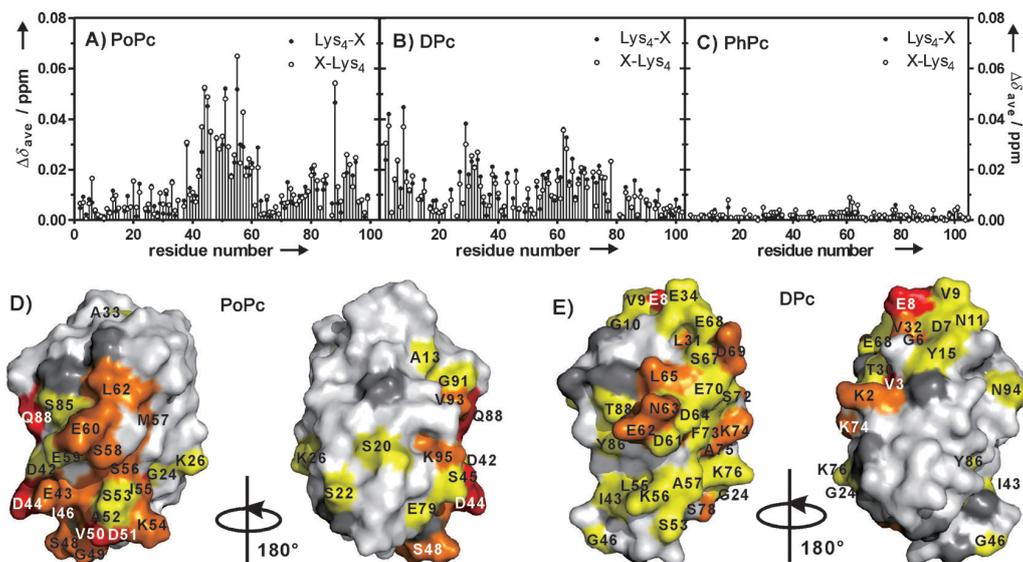


Figure 2. A)–C) Plots of NMR chemical shift perturbations measured for Pc backbone amides in the presence of TOAC-containing peptides. Extrapolated values (to 100% bound) for PoPc and DPc, and observed values for PhPc, are shown. D) and E) CSPs (extrapolated to 100% bound, see Table 1 for bound fractions) mapped onto the protein surfaces from the binding of Lys₄-X to PoPc (panel D, PDB ID: 1TKW^[32]) and DPc (panel E, PDB ID: 1KDI^[7]). Red, $\Delta\delta_{ave} \geq 0.04$ ppm; orange, $0.04 > \Delta\delta_{ave} \geq 0.02$ ppm; yellow, $0.02 > \Delta\delta_{ave} \geq 0.01$ ppm; white, $\Delta\delta_{ave} < 0.01$ ppm; Gray, no data or overlapping resonances. Binding maps for the other peptides are shown in Figure S1.

the CSPs are caused by interactions with the four lysines. The binding maps of Ala-Lys₄ and Lys₄-Ala were also similar to those of X-Lys₄ and Lys₄-X, indicating no significant effect of TOAC on peptide binding (Figure S1).

In PoPc and DPc, most CSPs occurred around the regions with the acidic patches, in agreement with the assumption that the positively charged peptides interact with the acidic residues of Pc.^[20,33] The largest CSPs for PoPc occurred for residues Asp44, Ser45, Asp51, Ile55, and Gln88. Among these residues, Asp44 belongs to the acidic patch. For DPc, the largest CSPs occurred for residues Val3 and Glu8. Glu8 is located at the acidic arc on the northern side. Although the observed CSPs are very small for PhPc, similar effects were still observed from both TOAC-containing peptides (Figure 2C). The small perturbations of the resonances of the copper ligand residues (His37, Cys84, His87, and Met92 for PoPc; His37, Cys87, His90, and Met95 for DPc; His39, Cys89, His92, and Met97 for PhPc) indicate that the copper site is not the main binding site of the peptides. Similar magnitudes of perturbations and binding maps caused by a tetralysine peptide (without an additional TOAC) were observed for Pc from the seed plant *Silene pratensis*.^[34]

Binding constants were obtained by fitting the CSP curves for the most affected residues (Figure 3, Figures S2 and S3, and Table 1). For PhPc, the magnitudes of the observed CSPs were too small ($|\Delta\delta_{H}| \leq 0.01$ ppm) to determine a dissociation constant.

The binding curves for PoPc fitted well to a single binding site model (Figure 3A). Interestingly, there were two types of dissociation constants observed in DPc titrations. The residues that are involved in stronger binding (lower K_d) were clustered on the northern side of DPc (Figure 2E and Figure S3B). This might be due to the unusual surface charge distribution of DPc compared with other plant Pcs. It is possible that there is internal competition between the two binding sites for the peptides. Clearly, a 1:1 binding model is not appropriate to explain this observation. Therefore, a two-site binding model was used to obtain the K_d values for DPc (Figure 3B and C, Figure S2B and C, and Figure S3C–F).

For most peptides, the K_d values for the same Pc are similar, indicating that the TOAC caused no significant changes in the affinity of the peptides for Pc. Only Lys₄-X has a somewhat lower K_d for PoPc than Lys₄-Ala, but the difference is within the error margins.

Table 1. Dissociation constants of the complexes formed between Pc(Zn) and tetralysine peptides and their calculated bound fractions at the end point of the titrations.

Pc	Lys ₄ -Ala		Lys ₄ -X		Ala-Lys ₄		X-Lys ₄	
	K_d [μ M]	Fraction	K_d [μ M]	Fraction	K_d [μ M]	Fraction	K_d [μ M]	Fraction
PoPc	150 ± 40	0.95	90 ± 30	0.97	110 ± 20	0.97	130 ± 40	0.96
DPc (strong)	110 ± 20	0.97	110 ± 20	0.97	110 ± 20	0.98	110 ± 20	0.96
DPc (weak)	300 ± 40	0.91	300 ± 50	0.90	340 ± 40	0.94	300 ± 100	0.94

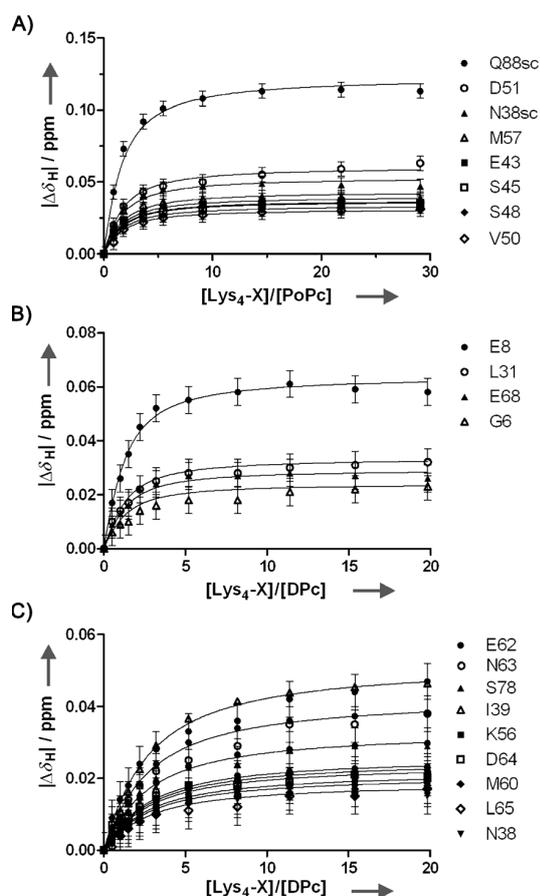


Figure 3. Chemical shift changes of selected Pcs resonances as a function of increasing [peptide]/[Pc]. The dissociation constants of the corresponding peptides (Table 1) were obtained by simultaneous fitting to a 1:1 binding model for PoPc (solid lines) and by simulation of two-site binding for DPc. Error bars represent ± 0.005 ppm. A) Lys₄-X with PoPc; B) Lys₄-X with DPc, strong binding residues; C) Lys₄-X with DPc, weak binding residues.

Paramagnetic relaxation enhancements

The paramagnetic TOAC was introduced to determine whether the bound peptide possesses a single, well-defined orientation or several orientations. If the peptide orientation is well-defined, the strong distance dependence of the PRE should result in highly localized effects. The TOAC was placed at the N or C terminus of the tetralysine peptide in order to interfere minimally with binding.^[29] The attached spin labels were thus expected to yield PRE of nuclei on nearby Pc residues. If the peptides bind in a specific orientation, the N- and C-terminal TOACS should generate different PRE patterns.

PREs were observed for some residues, as shown in Figure 4. Binding of these peptides to the three Pcs is in the fast-exchange regime, so the observed PRE is a weighted average of free Pc (no PRE) and bound Pc. By dividing the observed PRE by the fractions bound, calculated from the K_{d} , the PRE for 100% bound Pc was obtained. For DPc, the weak-binding K_{d} values were used, because most residues showed weak binding.

For PoPc binding to Lys₄-X, the resonances that were broadened beyond detection were those of Gly49, Glu59, and the side chain of Gln88. For PoPc binding with X-Lys₄, the resonance of an additional residue (Glu43) was completely broadened. These residues are located on the same side as the acidic patches, which include Glu43 and Glu59. Resonances of many residues located around the acidic patch also experienced PRE at various magnitudes. This observation indicates that the binding sites of the peptides on Pc are not restricted to the acidic patch residues only, but also extend to other polar or charged residues around this region and even to the hydrophobic patch, including some positive residues such as Lys26, Lys54, and Lys66 ($I_{\text{para}}/I_{\text{dia}}$ ratio: 0.60–0.84). This observation suggests that the peptides sample a large area of the protein surface and demonstrates the superior sensitivity of PRE for transient interactions.

For the interaction of DPc with Lys₄-X and X-Lys₄, the resonances of three residues disappeared from the spectra: Gly33, Gly36, and Glu68. Resonances of two other acidic residues (Glu34 and Asp69) were broadened but still visible in the spectra ($I_{\text{para}}/I_{\text{dia}}$ ratio: 0.59–0.82). These five residues are close together on the acidic arc at the northern side of DPc, indicating that the cluster of negative charges on the protein attracted the peptides by charge–charge interactions.

For PhPc, only one resonance (Thr75) had a clearly significant PRE ($I_{\text{para}}/I_{\text{dia}}$ ratio: 0.5) under the experimental conditions (peptide/protein ratio: 1:1). The $I_{\text{para}}/I_{\text{dia}}$ ratios of Val48, Leu55, His61 and Gln63 were 0.84, 0.83, 0.84, and 0.83, respectively. These values are close to the defined threshold for unaffected residues ($I_{\text{para}}/I_{\text{dia}}$ ratio: 0.85).

The PRE effects of tetralysine peptides on DPc are smaller than on PoPc in general. This is due to a smaller bound fraction. The nuclei that experience the largest CSP in DPc are not those that exhibit the largest PRE, probably because CSP monitors the effects from all atoms within the peptides, whereas PRE indicates the effects from the paramagnetic center only.

It is interesting to note that strongly affected residues have unaffected neighbors. One such example is Ala73 of PoPc, which is affected by PRE, while the neighboring Val72 and Leu74 are not. Similarly, Ala75 of DPc, located in between the residues with PREs (Lys74 and Lys76), remains unaffected. Another example is seen with Asp61 and Glu62 of DPc, both located on the acidic arc. Asp61 is affected, but Glu62 is unaffected. These findings suggest highly localized effects and will be discussed in more detail later.

Ensemble docking

Visualization of the encounter state on the basis of PRE data can be carried out quantitatively by using the ensemble docking approach.^[28] Calculations were performed using 1–15 copies of a pseudoatom that represents the paramagnetic center. Experimental PREs were converted into distances for ensemble docking. For DPc, the K_{d} values used here are the low affinity values, as most residues belong to the low affinity group. The high affinity residues were completely broadened; therefore, their target distance ranges are the same using

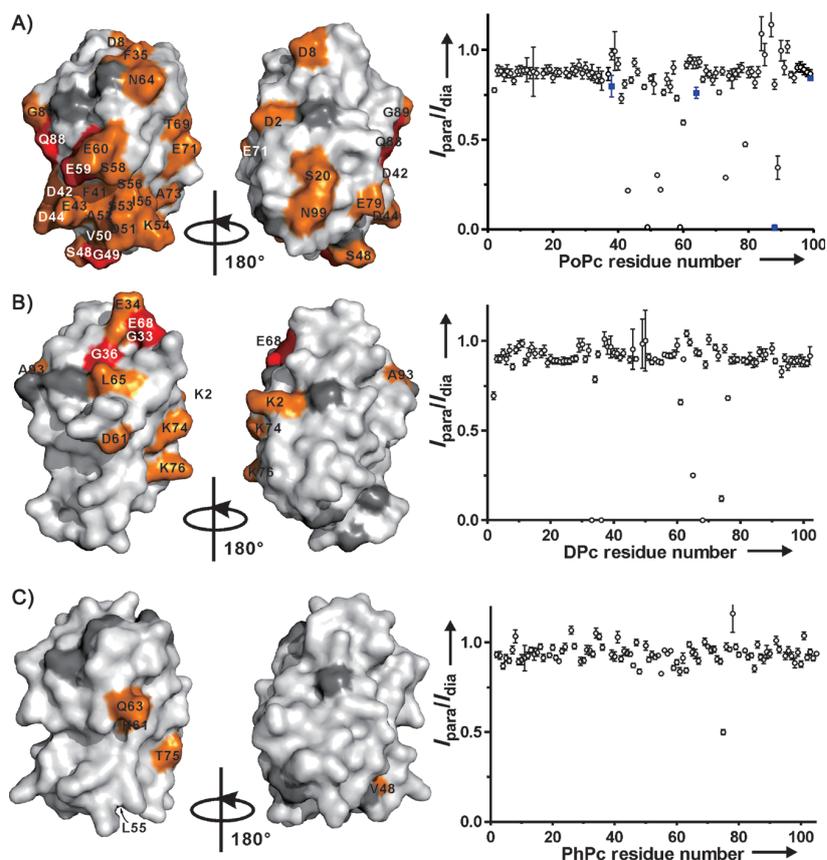


Figure 4. PRE effects in Pc-Lys₄-X complexes. The paramagnetic peptide was added to Pc at a peptide/Pc molar ratio of 0.5 for DPc and 1 for PoPc and PhPc, resulting in fractions of bound Pc of 14% for DPc and 35% for PoPc. The bound fraction for PhPc is unknown but expected to be very small. Left: A) PRE maps of PoPc (PDB ID: 1TKW^[32]), B) DPc (PDB ID: 1KD1^[7]), and C) PhPc (PDB ID: 2Q5B) bound to the Lys₄-X peptide. Surface model colors: red, $I_{para}/I_{dia} < 0.1$; orange, $0.1 \leq I_{para}/I_{dia} < 0.85$; white, $I_{para}/I_{dia} \geq 0.85$; gray, prolines, unassigned, and overlapping resonances. Right: relative ¹H,¹⁵N HSQC intensities of the backbone amide of A) PoPc (including side chains, which are shown as blue squares), B) DPc, and C) PhPc in complex with TOAC-containing peptides. Error bars denote twofold standard deviations, derived from spectral noise levels using standard error propagation procedures. For most data points, the error bars are within the symbol.

either K_d value. Violations were defined as the absolute differences between the distance back-calculated from the entire ensemble (by using r^{-6} averaging) and the experimental distance. Figure 5 shows the results of ensemble docking for Lys₄-X binding to PoPc and DPc with increasing ensemble size. Large distance violations occurred when using a single representation of the paramagnetic center (Figure 5, $N=1$), indicating that multiple orientations are required to describe the data. As a result of increasing degrees of freedom, the distance violations were reduced with increasingly larger ensembles. For PoPc (Figure 5A), no significant reduction in violation occurred at $N \geq 8$. For DPc, the violation curve flattened at $N=5$ (Figure 5B).

The resulting ensembles for $N=6$ are shown in Figure 6. Most of the paramagnetic centers are located in well-defined positions and not in a “cloud” of orientations. This correlates with the observation that some amides of Pc are strongly affected by PRE, whereas others that are nearby are not. That can be explained by assuming that the paramagnetic center

spends a short time being very close to the affected amide. Most of the affected amides have a considerable accessible surface area (ASA), which enables a close contact with the TOAC. In general, no major differences were observed for X-Lys₄ and Lys₄-X.

Monte Carlo simulations

Previous studies have shown that many encounter complexes are predominantly stabilized by electrostatic forces,^[2] although in some cases, short-range hydrophobic interactions may also contribute.^[35] Visualization of the encounter complex of Cyt c and Cyt c peroxidase was successfully achieved using PRE data and rigid-body MC simulations.^[36] The results showed that formation of this encounter complex was driven by charge-charge interactions. In MC simulations, one protein is docked to the other, guided by an electrostatic field and MC sampling.^[37] In this way, charge-charge interactions represent the only force that brings the binding partners together. Rigid-body MC docking simulations were performed for the Pc-peptide complexes, and a Boltzmann distribution of orientations of the peptide in complex with Pc was obtained. The paramagnetic centers of the peptides in this distribution are shown as green (Lys₄-X) and blue (X-Lys₄) spheres around Pc in Figure 7.

The results for PoPc (Figure 7A) and DPc (Figure 7B) show that the peptides are located close to the acidic patches. For PhPc, the population is more randomly distributed, with a relatively higher density at the side of PhPc that is farthest from the hydrophobic patch (Figure 7C).

The distances from the nitroxyl oxygen of the TOAC to the Pc amide hydrogens were measured and averaged (using r^{-6} averaging) for an ensemble existing of 2000 orientations randomly selected from the entire distribution. The distances obtained were compared with the experimental values. The violations calculated for the MC docking ensemble were 2.08, 1.70, 0.68, and 0.56 for PoPc-Lys₄-X, PoPc-X-Lys₄, DPc-Lys₄-X, and DPc-X-Lys₄, respectively. All violations are in the middle of the range of values shown in Figure 5A and B and Figure S5A and B). Thus, the MC docking ensemble does not fully agree with the PRE data. Figure 8 shows the back-calculated average dis-

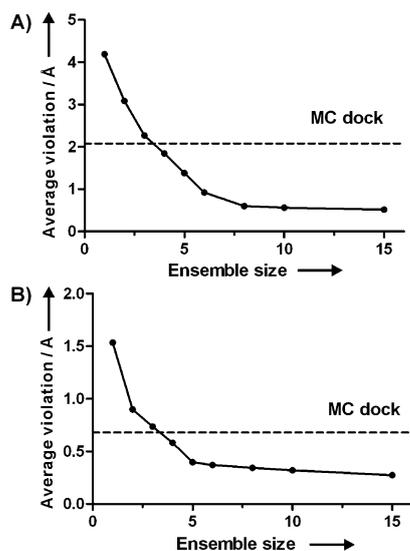


Figure 5. Averaged distance violations against a number of paramagnetic pseudoatoms ($N=1-6, 8, 10, 15$) in the ensemble docking. A) Lys₄-X-PoPc, B) Lys₄-X-DPc. The dashed horizontal lines indicate the average violations calculated from MC dock.

tances for each Pc residue in comparison with the PRE-derived distances. Although MC docking clearly places the paramagnetic center close to the affected residues, the simulation underestimates the PRE for these residues.

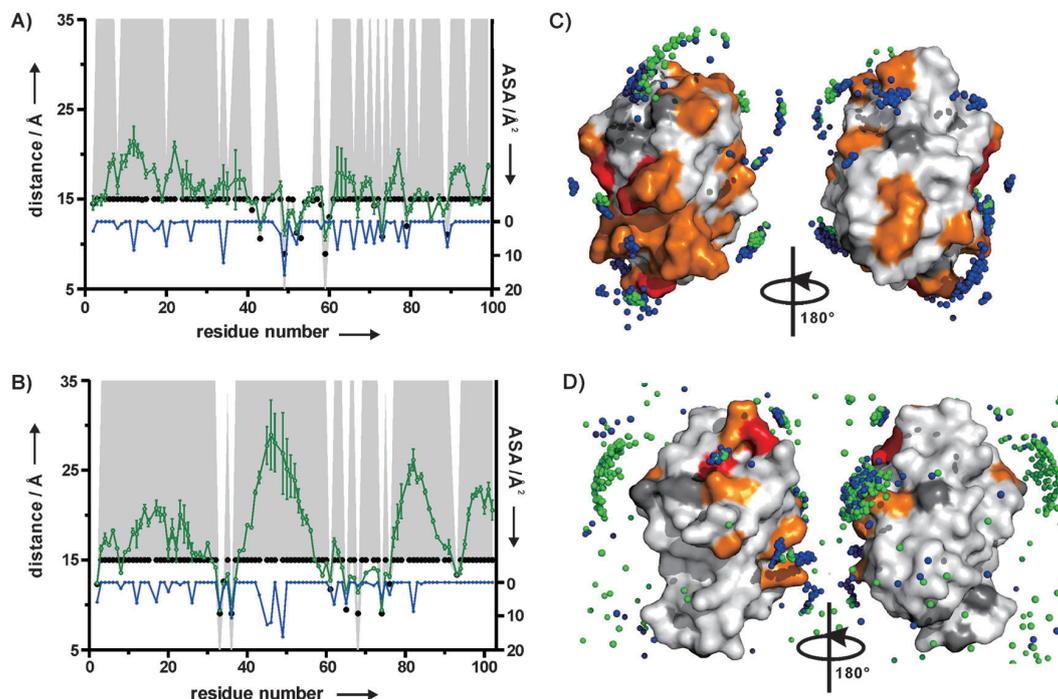


Figure 6. Ensemble docking. A) and B) Correlation of experimental distances (black dots) and back-calculated average distances (green dots with green connecting lines) from the ensemble docking ($N=6, 20$ lowest energy structures) of Lys₄-X bound to A) PoPc and B) DPc, with error bars representing the standard deviation. Right y-axes indicate the accessible surface area (ASA) of each amide proton, shown as blue dots with blue connecting lines. Gray areas indicate the error margins of the experimental distances. C) and D) PRE-based ensemble docking results ($N=6$) of C) PoPc (396 solutions for Lys₄-X and 594 for X-Lys₄) and D) DPc (630 solutions for Lys₄-X and 360 for X-Lys₄). The paramagnetic centers from TOAC are shown as spheres, with Lys₄-X in green and X-Lys₄ in blue. Protein surfaces are colored the same as in PRE maps (Figure 4).

Figure 9 shows the plots of electrostatic interaction energy distribution for the Pc-Lys₄-X complexes. PoPc (Figure 9A) and DPc (Figure 9B) have similar patterns. The highest population in DPc was at -6 kcal mol⁻¹, whereas in PoPc it was at -7 kcal mol⁻¹. For PhPc (Figure 9C), it is clear that the charge-charge interaction is much weaker (highest population at -2 kcal mol⁻¹). Histograms for the Pc-X-Lys₄ complexes are shown in Figure S5. The highest populations appeared at $-8, -7,$ and -2 kcal mol⁻¹ for PoPc, DPc, and PhPc, respectively.

Discussion

The aim of the present work was to experimentally characterize the dynamics in encounter complexes. The rationale was to create a pure encounter complex by ensuring that electrostatics dominate the interactions. For this purpose, the complexes formed by charged tetralysine peptides and three Pcs with distinct surface charge properties were studied. At pH 6.5, the net charges of PoPc, DPc, and PhPc are $-7, -5,$ and $-1,$ respectively, and the charge distributions differ markedly between these Pcs.

Previously, the interaction between the seed plant *S. pratensis* Pc and lysine peptides of varying lengths was studied using circular dichroism, UV-visible absorption, resonance Raman spectroscopy, and cyclic voltammetry. Minor changes in the geometry of the copper site were observed upon peptide binding.^[20,23] The peptides also competitively inhibited electron

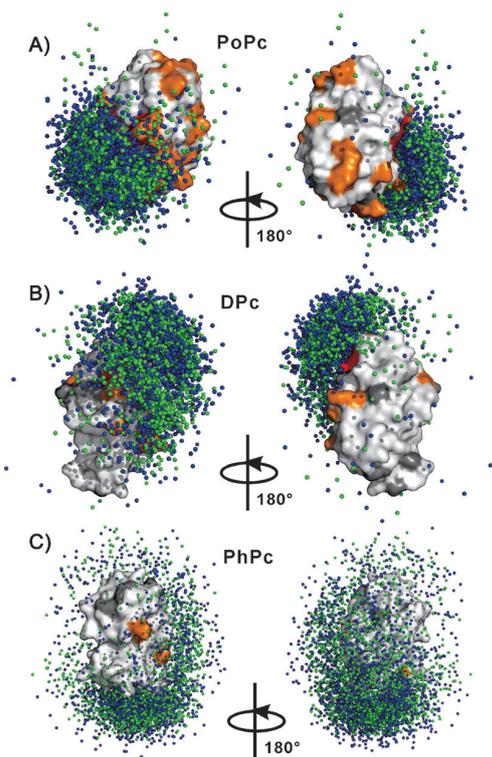


Figure 7. MC docking results showing 2000 solutions of Lys₄-X and X-Lys₄ bound to A) PoPc, B) DPc, and C) PhPc. The paramagnetic centers of the peptides are shown as green (Lys₄-X) and blue (X-Lys₄) spheres. Protein surfaces are colored according to the PRE maps (Figure 4). The 2000 orientations in each ensemble were selected randomly for the entire MC docking solution set.

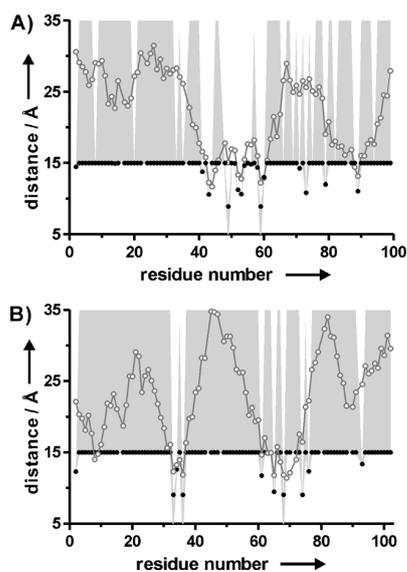


Figure 8. Comparison of experimental distances (dots) and back-calculated average distances (circles with connecting lines) between Pc amides and the TOAC nitroxyl oxygen atoms in the ensembles from MC simulations (2000 structures). A) PoPc-Lys₄-X; B) DPc-Lys₄-X.

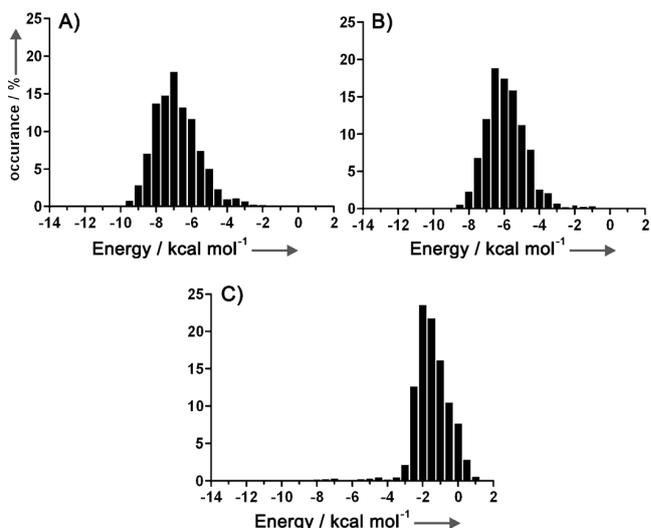


Figure 9. Histograms showing the electrostatic interaction energy distribution of 2000 structures randomly selected from the MC simulations. A) PoPc-Lys₄-X, B) DPc-Lys₄-X, C) PhPc-Lys₄-X.

transfer from Cyt *f*^[22] and Cyt *c*.^[20] Mutagenesis of Pc showed that the interaction and electron-transfer inhibition by lysine peptides decreased significantly as the net charge of the Pc negative patch decreased,^[20] showing that charge interaction contributed to the binding. The authors of this study proposed a specific and effective interaction between the positively charged peptides and the negative patches of Pc.^[20] These studies monitored spectroscopic changes caused by peptide binding but could not directly observe the binding interface and the dynamics of the interaction. Other studies of highly charged electron transfer proteins with small- or medium-sized molecules also emphasized the importance of charge-charge interactions in binding. Increasingly tight binding to Cyt *c* was observed for porphyrins with an increasing number of carboxylates, even with subnanomolar *K_d* values.^[38,39] Structural characterization of the complex of Cyt *c* with similar, though more weakly binding porphyrins (high micromolar range) by CSP analysis suggested mobility of the porphyrin on the protein surface.^[40] Cyt *c* has also been shown to interact with calixarenes. When decorated with many negative charges, these compounds compete effectively with natural protein partners for binding to Cyt *c*.^[41] A recent and elegant structural study showed that calixarenes interact with several amino groups of lysines, taking several specific conformations on the surface of the cytochrome.^[42]

Charge-charge interactions

To establish whether charge-charge interactions were the dominating interaction force in the Pc-Lys₄ complexes, the interaction surface was mapped using CSPs and compared with the results from the electrostatically driven MC simulations. In PoPc and DPc, CSPs were largest in the acidic regions. The *K_d* values were about 100 μM for PoPc and 110 μM and 300 μM for the two binding sites on DPc. For PhPc, peptide binding re-

sulted in very small CSPs, suggesting a low affinity. No dissociation constant could be determined. These results are in good agreement with the MC simulations. The electrostatic ensembles match well with the CSP-derived binding maps for PoPc and DPc. The electrostatic interaction energies indicated that these two Pcs have a strong interaction, whereas for PhPc, the affinity was quite weak. The data indicate that charge-driven binding is a good first description of the complexes.

Paramagnetic relaxation effects

To determine whether the peptides assume a single, well-defined orientation or exhibit multiple orientations, the paramagnetic amino acid TOAC was incorporated at the N or C terminus of tetralysine peptides. Control peptides with Ala instead of TOAC were used to assess the effect of TOAC incorporation on peptide binding to Pc. No significant difference between the binding affinities of TOAC- and Ala-tetralysine peptides was observed, indicating that TOAC has little influence on the thermodynamics of peptide binding.

In PoPc, the presence of TOAC caused PREs mainly in the neighborhood of the acidic patches as well as for some of the hydrophobic patch residues. Almost no CSPs were observed in the hydrophobic patch, which suggests that the PREs for those residues represent peptide orientations that are sparsely populated. The PRE is highly sensitive for minor states in which the paramagnetic center is brought within close proximity of the nucleus. Apparently, transiently peptide–protein interactions that are not dominated by electrostatic forces were present. In DPc, the area affected by PREs is smaller and more localized than in PoPc. The largest PREs were detected around the top of the acidic arc, close to the copper site. The PRE and CSP maps were similar in this case. For PhPc, few PREs and CSPs were observed, in accordance with the weak affinity for tetralysine peptides.

Dynamics in the complexes

It is believed that the overall size of the CSP is a measure for the degree of dynamics in a protein complex. Large CSPs are caused by a single, well-defined orientation in the complex, in which desolvation of the interface and multiple short-range interactions occur. Small CSPs indicate averaging of multiple orientations in the encounter state, with minimal desolvation. Small CSPs have been observed in several complexes of redox proteins that are thought to be highly dynamic, including Cyt *b*₅–myoglobin,^[48] Cyt *c*–adrenodoxin,^[49] Cyt *c*–Pc,^[50] and Cyt *c*–Cyt *b*₅.^[51] In this study, similarly small CSPs were observed in all Pc–peptide complexes. Small CSPs can be caused by a dynamic interaction or simply low affinity. In the case of PoPc and DPc, the CSPs could be extrapolated to 100% on the basis of the *K*_d value, demonstrating that the CSPs are indeed small for the fully bound Pc. For PhPc the CSPs were too small even to derive a reliable *K*_d value. To support the hypothesis that small overall CSP values correlate with dynamic interactions, we used PRE mapping. The observed PREs were scattered over the Pc surface, and both for PoPc and DPc, they could not be

satisfied by a single orientation of the peptides. Furthermore, the N- and C-terminal TOAC-containing peptides gave very similar PRE maps, which is not to be expected for peptides binding in well-defined orientations. Thus, qualitatively, the PRE results strongly support a dynamic binding model in which the peptide assumes many orientations relative to Pc and interconverts between these orientations faster than the NMR time-scale defined by the maximum CSP (exchange rate ≥ 250 s⁻¹).

Back-calculated distances using the ensemble docking approach with multiple orientations showed a good correlation with the experimental PREs for ensemble sizes much larger than 1, which is in line with dynamics within the complex. Also, the average distances between TOAC and Pc amides of the MC docking ensembles matched the experimental distances qualitatively but not quantitatively; the TOAC molecules were, on average, not close enough to the affected Pc amide groups to explain the observed PREs. This observation could be a consequence of the limitations of the docking method, such as the use of an exclusion grid to avoid steric hindrance. Alternatively, it could point toward small contributions of interactions other than electrostatics, perhaps very transient hydrogen bond formation between the exposed amide protons and the oxygen of TOAC. Evidence for the latter explanation comes from the PRE pattern. It is remarkable that the NMR resonances of several residues were broadened beyond detection due to a PRE, whereas those of neighboring amides were (almost) unaffected. The distance between neighboring amides is about 4 Å, so the PRE ratio for two amide residues is at most proportional to $r^{-6}/(r+4)^{-6}$, where *r* is the distance between the nitroxyl radical and the nearest amide proton. It can be shown that, at least for some amides, this must imply that the TOAC nitroxyl group approaches very closely, within several Ångströms for a short fraction of the time, which suggests that the sensitivity of PRE for minor states provides evidence for weak and transient short-range interactions. In physiological systems of protein–protein complexes such interactions must occur in the encounter complex next to the dominant charge–charge interactions for the complex to proceed to the final, well-defined complex.

Conclusions

The binding of tetralysine peptides to Pcs with different surface charge properties was characterized by a combination of CSP, PRE NMR, and MC simulations. The high similarity of CSP maps for the different peptides used in the study, as well as the small magnitudes of CSPs, strongly suggests a high degree of dynamics. Also, the scattered distribution of PREs indicates the presence of multiple orientations. The peculiar distribution of peptide positions obtained from ensemble docking with high densities in small areas only qualitatively matches the electrostatic docking simulations, suggesting that the PRE approach picks up very transient, short-range interactions between the peptide and the protein, in which the TOAC closely approaches specific amide protons.

Experimental Section

Peptide synthesis and preparation: Fmoc-TOAC-OH was purchased from Iris Biotech (Germany). Synthetic peptides Ala-Lys₄, Lys₄-Ala, TOAC-Lys₄ (X-Lys₄), and Lys₄-TOAC (Lys₄-X) were prepared as described,^[29] with N-terminal acetylation and C-terminal amidation. Peptide purity was verified by rpHPLC, and peptide integrity was assessed by MALDI-TOF mass spectrometry. The peptides were dissolved in 10 mM NaPi, pH 6.5. The fraction of paramagnetic peptide was checked by EPR and found to be close to 100%. The quantity of trifluoroacetic acid (TFA) in the samples was confirmed by ¹⁹F NMR with trifluorotoluene as the internal reference. A TFA/peptide molar ratio of 5:1 was used to calculate the peptide concentration.

Protein expression and purification

General procedure: ¹⁵N-enriched M9 minimal media was prepared as described previously.^[52] For PoPc and PhPc, copper was excluded during bacterial growth. For additional ¹³C labeling, the minimal medium was supplemented with 2 g L⁻¹ ¹³C-glucose. Cells were harvested by centrifugation and lysed with a French pressure cell (Stansted Fluid Power Ltd.) in the presence of 1 mg lysozyme, 3.75 mg DNase, 1 mM PMSF, and ZnCl₂ (100 μM for PoPc and DPc, 5 mM for PhPc). For PoPc and DPc, an additional 250 μM of ZnCl₂ was added after passing through the French press. Cell debris was removed by centrifugation at 7000g for 25 min, and membranes were removed by ultracentrifugation at 25000g for 1 h. All columns used for purification were purchased from GE Healthcare Biosciences. PoPc and DPc concentrations were determined using the Bradford assay (Bio-Rad) with bovine serum albumin as the standard. Pc was considered pure when the protein migrated as a single band on SDS-PAGE (15%, 200 V, 50 min).

PoPc: The PoPc gene from plasmid pETPC^[32] was subcloned into a pET28 plasmid with an additional glycine residue at the N terminus. ¹⁵N-labeled PoPc was essentially produced as described^[32] with the following modification: the protein was expressed in *E. coli* (Rosetta 2) in M9 minimal medium (1 L, 0.5 L per 2 L Erlenmeyer flask). Protein production was induced by adding IPTG to a final concentration of 0.5 mM. Incubation was continued at 16 °C overnight. The protein was purified using 3 × 5 mL HiTrap-DEAE FF ion-exchange columns in 20 mM sodium phosphate, pH 7.0. The protein was eluted with a gradient of 0–500 mM NaCl. Fractions containing PoPc were concentrated and purified by a Superdex G-75 size-exclusion column in 20 mM sodium phosphate, pH 6.8, 100 mM NaCl. The yield of pure protein was 1.5 mg L⁻¹ of culture for ¹⁵N-PoPc and 0.75 mg L⁻¹ of culture for ¹⁵N,¹³C-PoPc.

DPc: ¹⁵N- and ¹⁵N/¹³C-labeled recombinant DPc containing zinc was produced in *E. coli* BL21(DE3) and purified as described before^[31] with the following modifications: all copper salts were replaced by ZnCl₂ during purification. The protein was purified using 3 × 5 mL HiTrap-Q HP ion-exchange columns in 10 mM sodium phosphate, pH 5.8 at 4 °C. The impurities were eluted with a gradient of 0–100 mM NaCl at 4 mL min⁻¹, and the Pc protein was eluted in 100 mM NaCl at 0.5 mL min⁻¹. Then, size-exclusion chromatography with a Superdex G-75 column was performed in a buffer of 10 mM sodium phosphate, pH 6.5 and 100 mM NaCl. The yield of pure protein was 149 mg L⁻¹ of culture for ¹⁵N-DPc and 19 mg L⁻¹ of culture for ¹⁵N,¹³C-DPc.

PhPc: Uniformly ¹⁵N-enriched PhPc was produced without the leader peptide and purified as described^[53] with the following modifications: after cell lysis and ultracentrifugation, the supernatant was dialyzed against 0.5 mM ZnCl₂ and 5 mM Tris-HCl, pH 7.5

overnight at 4 °C. Pc concentrations were determined using ε₂₈₀ = 5.00 cm⁻¹ mm⁻¹ on a Cary 50 spectrophotometer (Varian). The yield of pure protein was 3.5 mg L⁻¹ of culture for ¹⁵N-PhPc.

NMR measurements: All Pcs were concentrated by ultrafiltration (Amicon, M_w-cutoff 3 kDa). The sample buffer was 10 mM sodium phosphate, pH 6.5, and 6% D₂O. For peptide titrations, the protein concentrations were 200 μM for ¹⁵N-DPc(Zn) and ¹⁵N-PhPc(Zn) and 110 μM for ¹⁵N-PoPc(Zn). The samples for fern Pc and poplar Pc backbone assignments consisted of 2.4 mM and 0.25 mM ¹³C/¹⁵N-labeled protein, respectively. Peptide solutions were prepared in 10 mM sodium phosphate, pH 6.5. All NMR spectra were recorded at 300 K on a Bruker AVIII600 spectrometer equipped with a triple-resonance TXI-Z-GRAD cryoprobe, or a Bruker 600 MHz Avance DRX spectrometer equipped with a 5 mm TCI cryoprobe. Data were processed with TopSpin (Bruker) and analyzed in SPARKY.^[55] Resonances in the HSQC spectra of DPc and PoPc were assigned using 3D HNCACB experiments. The side chain resonance assignments of PoPc were taken from PoPc(Cd).^[32] NMR assignments have been deposited to the BMRB, entry codes 19236 (DPc) and 19247 (PoPc).

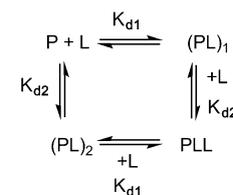
PRE analysis: The paramagnetic X-Lys₄ and Lys₄-X peptides were added into ¹⁵N-labelled Pc separately, and ¹H,¹⁵N HSQC spectra were recorded. Each paramagnetic peptide was added to Pc at a peptide/Pc molar ratio of 0.5 for DPc and 1 for PoPc and PhPc. Under these conditions, the fractions of bound Pc were 14% for DPc and 35% for PoPc. Diamagnetic spectra were recorded by reducing the peptides with sodium ascorbate (5 equiv). PREs were determined according to the procedure of Battiste and Wagner.^[54] The intensity ratio (*I*_{para}/*I*_{dia}) of the Pc resonances in the presence of X-Lys₄ or Lys₄-X was normalized by dividing by the average value of the ten largest *I*_{para}/*I*_{dia} values. The scaling factors for each Pc-peptide were 0.92, 0.87, 0.93, 1.07, 0.94, and 0.95 for PoPc-Lys₄-X, PoPc-X-Lys₄, DPc-Lys₄-X, DPc-X-Lys₄, PhPc-KKKX and PhPc-X-Lys₄, respectively.

PRE (*R*₂^{para}) values were calculated according to Equation (1):

$$\frac{I_{\text{para}}}{I_{\text{dia}}} = \frac{R_2^{\text{dia}} \exp(-R_2^{\text{para}} t)}{R_2^{\text{dia}} + R_2^{\text{para}}} \quad (1)$$

The transverse relaxation rates in the diamagnetic sample (*R*₂^{dia}) were calculated from the line width at half height obtained from a Lorentzian peak fit in the direct dimension using SPARKY. The symbol *t* denotes the time for transverse relaxation during the pulse sequence (9 ms).

Calculation of dissociation constants: Peptide binding was observed through the changes of protein resonances in the ¹H,¹⁵N HSQC spectrum upon titration with the peptide. CSP analysis was carried out as described before.^[56] The dissociation constants (*K*_d) for PoPc were determined using a two-parameter nonlinear regression curve fitting based on a one-site binding model as described previously.^[57] The fraction of bound ligand was calculated using the dissociation constant. For DPc, resonance overlap was observed for Thr79/Glu25 and Ala23/Phe12 during the titrations. These four residues were excluded from the *K*_d calculation. The peptide-DPc interaction was modeled with two independent binding sites (Scheme 1). P and L in



Scheme 1. Two-site binding model of DPc. P, protein; L, ligand; *K*_{d1} and *K*_{d2}, dissociation constants.

dicating the free protein and the free peptide, respectively. (PL)₁ and (PL)₂ are the 1:1 complexes formed by peptide binding to sites 1 and 2 on DPc, respectively. PLL is the protein with two peptides bound. K_{d1} and K_{d2} are the dissociation constants for sites 1 and 2, respectively. The binding curves were simulated numerically with varying values for K_{d1} , K_{d2} , and the $\Delta\delta$ at 100% bound Pc using Microsoft Excel.

Ensemble docking: For DPc, PhPc, and PoPc, the PDB IDs 1KDI,^[7] 2Q5B, and Pc from 1TKW model 1^[32] were used, respectively. The structure of Pc in 1TKW originated from PDB 5PCY.^[4] The RMSD of all atoms between Pc in 1TKW and 5PCY is 0.15 Å.

The PREs were converted into distances for structure calculations as described previously.^[56] τ_c was taken to be 5.54 ns for DPc, 5.14 ns for PoPc, and 5.93 ns for PhPc, on the basis of the HYDRONMR^[58] prediction of the rotational correlation time for each Pc. For each peak, R_2 was estimated from the width at half-height ($\Delta\nu_{1/2}$) of a Lorentzian fit in the proton dimension by using $R_2 = \pi\Delta\nu_{1/2}$. PRE values were calculated after normalization of the I_{para}/I_{dia} ratios and extrapolated to 100% bound by dividing the values by their bound fractions (35% for PoPc and 14% for DPc). Three classes of PRE restraints were included in the calculations:^[29] 1) For amide residues whose resonances disappear in the paramagnetic spectrum, an upper limit for I_{para} was estimated from the standard deviation of the noise level of the spectrum. The upper bound PRE (R_2^{para}) value was set to 500 s⁻¹ and the distance set to 9 Å. 2) For residues with $I_{para}/I_{dia} > 0.85$, the lower bound distance was set to 15 Å. 3) For residues with I_{para}/I_{dia} between 0.1 and 0.85, the distances (r) calculated according to a previously described equation^[56] were used, with upper and lower bounds of ($r \pm 0.1$) Å. Violations were defined as the absolute differences between the calculated distance and the experimental distance including the corresponding upper and lower bound margins for the three classes. An additional restraint ensures that the TOAC nitroxy oxygen atom and the Pc center of mass are at a distance between 10 and 30 Å. The structure calculations were done in XPLOR-NIH.^[59] The accessible surface area (ASA) of each amide proton was calculated with a Python-based implementation of the Shrake–Rupley algorithm.^[60]

Monte Carlo simulations: The peptide coordinates of X-Lys₄ and Lys₄-X were generated from the PRODRG server,^[61] and the conformations were optimized in Swiss PDB-Viewer^[62] to separate the charges as far as possible (Figure S6). For DPc, PhPc, and PoPc the PDB IDs 1KDI,^[7] 2Q5B, and 1TKW model 1^[32] were used, respectively. Structure preparation and the rigid-body MC simulation^[37] were performed as described.^[35,36] The electrostatic potential was calculated with APBS^[63] for an ionic strength of 0.01 M and a temperature of 300 K to match the experimental conditions. An ensemble of 2000 peptide orientations, randomly selected from the entire run of 2.2×10^6 saved structures, was considered for the calculations. The averaged distances were derived from the ensemble and compared to the experimental distances.

Acknowledgements

We thank Dr. Sandra Scanu for providing the resonance assignment of plastocyanin from *P. laminosum* and Martin van Son for EPR measurements of the TOAC peptides. We thank Dr. Hans Wienk and the NMR facility of the Bijvoet Center in Utrecht for use of the 600 MHz spectrometer. M.U. and M.T. received financial support from the Netherlands Organisation for Scientific Research (NWO), grant 700.58.441. G.M.U. and J.M.F. were supported by

the German Science Foundation (DFG; GRK 1640). The authors declare no competing financial interests.

Keywords: encounter complexes · molecular modeling · NMR spectroscopy · plastocyanins · tetralysine peptides

- [1] Q. Bashir, S. Scanu, M. Ubbink, *FEBS J.* **2011**, *278*, 1391–1400.
- [2] M. Ubbink, *FEBS Lett.* **2009**, *583*, 1060–1066.
- [3] G. McLendon, *Struct. Bonding (Berlin)* **1991**, *75*, 159–174.
- [4] J. M. Guss, P. R. Harrowell, M. Murata, V. A. Norris, H. C. Freeman, *J. Mol. Biol.* **1986**, *192*, 361–387.
- [5] J. M. Moore, C. A. Lepre, G. P. Gippert, W. J. Chazin, D. A. Case, P. E. Wright, *J. Mol. Biol.* **1991**, *221*, 533–555.
- [6] C. S. Bond, D. S. Bendall, H. C. Freeman, J. M. Guss, C. J. Howe, M. J. Wagner, M. C. J. Wilce, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, *55*, 414–421.
- [7] T. Kohzuma, T. Inoue, F. Yoshizaki, Y. Sasakawa, K. Onodera, S. Nagatomo, T. Kitagawa, S. Uzawa, Y. Isobe, Y. Sugimura, M. Gotowda, Y. Kai, *J. Biol. Chem.* **1999**, *274*, 11817–11823.
- [8] T. P. Garrett, D. J. Clingeffer, J. M. Guss, S. J. Rogers, H. C. Freeman, *J. Biol. Chem.* **1984**, *259*, 2822–2825.
- [9] H. Sugawara, T. Inoue, C. Li, M. Gotowda, T. Hibino, T. Takabe, Y. Kai, *J. Biochem.* **1999**, *125*, 899–903.
- [10] Y. Xue, M. Ökvist, Ö. Hansson, S. Young, *Protein Sci.* **1998**, *7*, 2099–2105.
- [11] S. Bagby, P. C. Driscoll, T. S. Harvey, H. A. O. Hill, *Biochemistry* **1994**, *33*, 6611–6622.
- [12] P. M. Colman, H. C. Freeman, J. M. Guss, M. Murata, V. A. Norris, J. A. M. Ramshaw, M. P. Venkatappa, *Nature* **1978**, *272*, 319–324.
- [13] C. A. Collyer, J. M. Guss, Y. Sugimura, F. Yoshizaki, H. C. Freeman, *J. Mol. Biol.* **1990**, *211*, 617–632.
- [14] M. R. Redinbo, D. Cascio, M. K. Choukair, D. Rice, S. Merchant, T. O. Yeates, *Biochemistry* **1993**, *32*, 10560–10567.
- [15] N. Shibata, T. Inoue, C. Nagano, N. Nishio, T. Kohzuma, K. Onodera, F. Yoshizaki, Y. Sugimura, Y. Kai, *J. Biol. Chem.* **1999**, *274*, 4225–4230.
- [16] I. Bertini, D. A. Bryant, S. Ciurli, A. Dikiy, C. O. Fernández, C. Luchinat, N. Safarov, A. J. Vila, J. Zhao, *J. Biol. Chem.* **2001**, *276*, 47217–47226.
- [17] L. Schmidt, H. E. M. Christensen, P. Harris, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62*, 1022–1029.
- [18] I. Díaz-Moreno, A. Díaz-Quintana, M. A. De La Rosa, M. Ubbink, *J. Biol. Chem.* **2005**, *280*, 18908–18915.
- [19] R. Hulsker, M. V. Baranova, G. S. Bullerjahn, M. Ubbink, *J. Am. Chem. Soc.* **2008**, *130*, 1985–1991.
- [20] S. Hirota, K. Hayamizu, M. Endo, T. Hibino, T. Takabe, T. Kohzuma, O. Yamauchi, *J. Am. Chem. Soc.* **1998**, *120*, 8177–8183.
- [21] S. Hirota, M. Endo, T. Tsukazaki, *J. Biol. Inorg. Chem.* **1998**, *3*, 563–569.
- [22] S. Hirota, M. Endo, *J. Am. Chem. Soc.* **1999**, *121*, 849–855.
- [23] S. Hirota, K. Hayamizu, T. Okuno, M. Kishi, H. Iwasaki, T. Kondo, *Biochemistry* **2000**, *39*, 6357–6364.
- [24] S. Hirota, T. Tsukazaki, O. Yamauchi, *Biochem. Biophys. Res. Commun.* **2000**, *268*, 395–397.
- [25] S. Hirota, O. Yamauchi, *Eur. J. Inorg. Chem.* **2002**, 17–25.
- [26] S. Hirota, O. Yamauchi, *Chem. Rec.* **2001**, *1*, 290–299.
- [27] J. Iwahara, G. M. Clore, *Nature* **2006**, *440*, 1227–1230.
- [28] G. M. Clore, J. Iwahara, *Chem. Rev.* **2009**, *109*, 4108–4139.
- [29] H. E. Lindfors, P. E. de Koning, J. W. Drijfhout, B. Venezia, M. Ubbink, *J. Biomol. NMR* **2008**, *41*, 157–167.
- [30] S. Schreier, J. C. Bozelli, N. Marín, R. F. F. Vieira, C. R. Nakaie, *Biophys. Rev.* **2012**, *4*, 45–66.
- [31] R. Hulsker, A. Mery, E. A. Thomassen, A. Ranieri, M. Sola, M. P. Verbeet, T. Kohzuma, M. Ubbink, *J. Am. Chem. Soc.* **2007**, *129*, 4423–4429.
- [32] C. Lange, T. Cornvik, I. Díaz-Moreno, M. Ubbink, *Biochim. Biophys. Acta Bioenerg.* **2005**, *1707*, 179–188.
- [33] S. Hirota, H. Okumura, S. Arie, K. Tanaka, M. Shionoya, T. Takabe, N. Funasaki, Y. Watanabe, *J. Inorg. Biochem.* **2004**, *98*, 849–855.
- [34] R. Hulsker, *PhD thesis*, Leiden University (The Netherlands), **2008**.
- [35] S. Scanu, J. M. Foerster, G. M. Ullmann, M. Ubbink, *J. Am. Chem. Soc.* **2013**, *135*, 7681–7692.
- [36] Q. Bashir, A. N. Volkov, G. M. Ullmann, M. Ubbink, *J. Am. Chem. Soc.* **2010**, *132*, 241–247.

- [37] G. M. Ullmann, E. Knapp, N. M. Kostić, *J. Am. Chem. Soc.* **1997**, *119*, 42–52.
- [38] R. K. Jain, A. D. Hamilton, *Org. Lett.* **2000**, *2*, 1721–1723.
- [39] T. Aya, A. Hamilton, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2651–2654.
- [40] P. B. Crowley, P. Ganji, H. Ibrahim, *ChemBioChem* **2008**, *9*, 1029–1033.
- [41] Y. Wei, G. McLendon, A. D. Hamilton, *Chem. Commun.* **2001**, 1580–1581.
- [42] R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power, P. B. Crowley, *Nat. Chem.* **2012**, *4*, 527–533.
- [43] S. Fletcher, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **2005**, *9*, 632–638.
- [44] H. Yin, A. D. Hamilton, *Angew. Chem.* **2005**, *117*, 4200–4235; *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163.
- [45] S. Fletcher, A. D. Hamilton, *J. R. Soc. Interface* **2006**, *3*, 215–233.
- [46] T. Mecca, G. Consoli, C. Geraci, F. Cunsolo, *Bioorg. Med. Chem.* **2004**, *12*, 5057–5062.
- [47] M. Blaskovich, Q. Lin, F. Delarue, J. Sun, H. S. Park, D. Coppola, A. D. Hamilton, S. M. Sebt, *Nat. Biotechnol.* **2000**, *18*, 1065–1070.
- [48] J. A. R. Worrall, Y. Liu, P. B. Crowley, J. M. Nocek, B. M. Hoffman, M. Ubbink, *Biochemistry* **2002**, *41*, 11721–11730.
- [49] X. Xu, W. Reinle, F. Hannemann, P. V. Konarev, D. I. Svergun, R. Bernhardt, M. Ubbink, *J. Am. Chem. Soc.* **2008**, *130*, 6395–6403.
- [50] M. Ubbink, D. S. Bendall, *Biochemistry* **1997**, *36*, 6326–6335.
- [51] A. N. Volkov, D. Ferrari, J. A. R. Worrall, A. M. J. J. Bonvin, M. Ubbink, *Protein Sci.* **2005**, *14*, 799–811.
- [52] M. Cai, Y. Huang, K. Sakaguchi, G. M. Clore, A. M. Gronenborn, R. Craigie, *J. Biomol. NMR* **1998**, *11*, 97–102.
- [53] P. B. Crowley, G. Otting, B. G. Schlarb-Ridley, G. W. Canters, M. Ubbink, *J. Am. Chem. Soc.* **2001**, *123*, 10444–10453.
- [54] J. L. Battiste, G. Wagner, *Biochemistry* **2000**, *39*, 5355–5365.
- [55] T. D. Goddard, D. G. Kneller, SPARKY 3.0, University of California, San Francisco, **2008**.
- [56] S. Scanu, J. M. Foerster, M. G. Finiguerra, M. H. Shabestari, M. Huber, M. Ubbink, *ChemBioChem* **2012**, *13*, 1312–1328.
- [57] A. Kannt, S. Young, D. S. Bendall, *Biochim. Biophys. Acta Bioenerg.* **1996**, *1277*, 115–126.
- [58] J. García de La Torre, M. L. Huertas, B. Carrasco, *J. Magn. Reson.* **2000**, *147*, 138–146.
- [59] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65–73.
- [60] A. Shrake, J. A. Rupley, *J. Mol. Biol.* **1973**, *79*, 351–371.
- [61] A. W. Schüttelkopf, D. M. F. van Aalten, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355–1363.
- [62] N. Guex, M. C. Peitsch, *Electrophoresis* **1997**, *18*, 2714–2723.
- [63] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10037–10041.
- [64] The PyMOL Molecular Graphics System, Version 0.99, Schrödinger, LLC.

Received: September 26, 2013

Published online on February 6, 2014

CHEMBIOCHEM

Supporting Information

© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2014

An Ensemble of Rapidly Interconverting Orientations in Electrostatic Protein–Peptide Complexes Characterized by NMR Spectroscopy

Jia-Ying Guan,^[a] Johannes M. Foerster,^[b] Jan W. Drijfhout,^[c] Monika Timmer,^[a]
Anneloes Blok,^[a] G. Matthias Ullmann,^[b] and Marcellus Ubbink^{*[a]}

cbic_201300623_sm_miscellaneous_information.pdf

Supporting Information

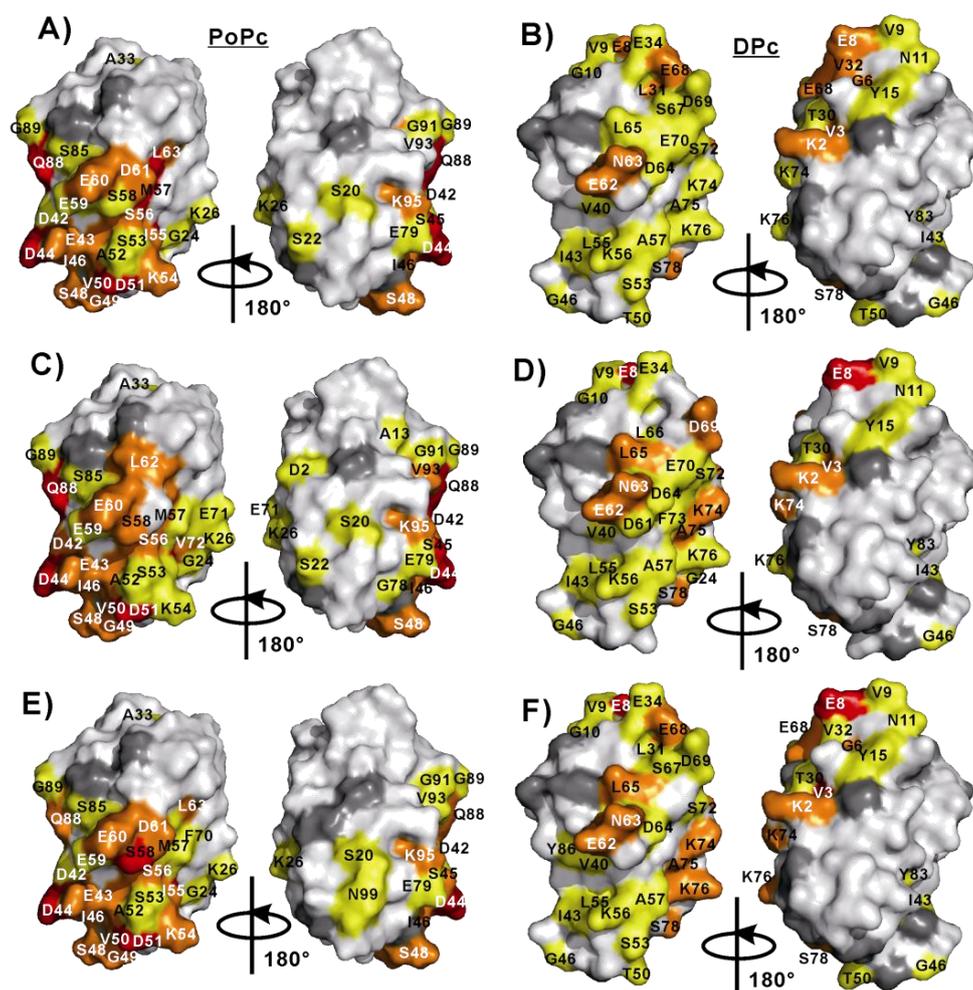


Figure S1: CSPs (extrapolated to 100% bound) mapped onto the protein surfaces from the binding of Lys₄-X (panels A and B), Lys₄-Ala (panels C and D) and Ala-Lys₄ (panels E and F) to PoPc (left, PDB entry 1TKW^[1]) and DPc (right, PDB entry 1KDI^[2]). Color representations: red, $\Delta\delta_{ave} \geq 0.04$ ppm; orange, $0.04 > \Delta\delta_{ave} \geq 0.02$ ppm; yellow, $0.02 > \Delta\delta_{ave} \geq 0.01$ ppm; white, $\Delta\delta_{ave} < 0.01$ ppm; grey, no data or overlapping resonances.

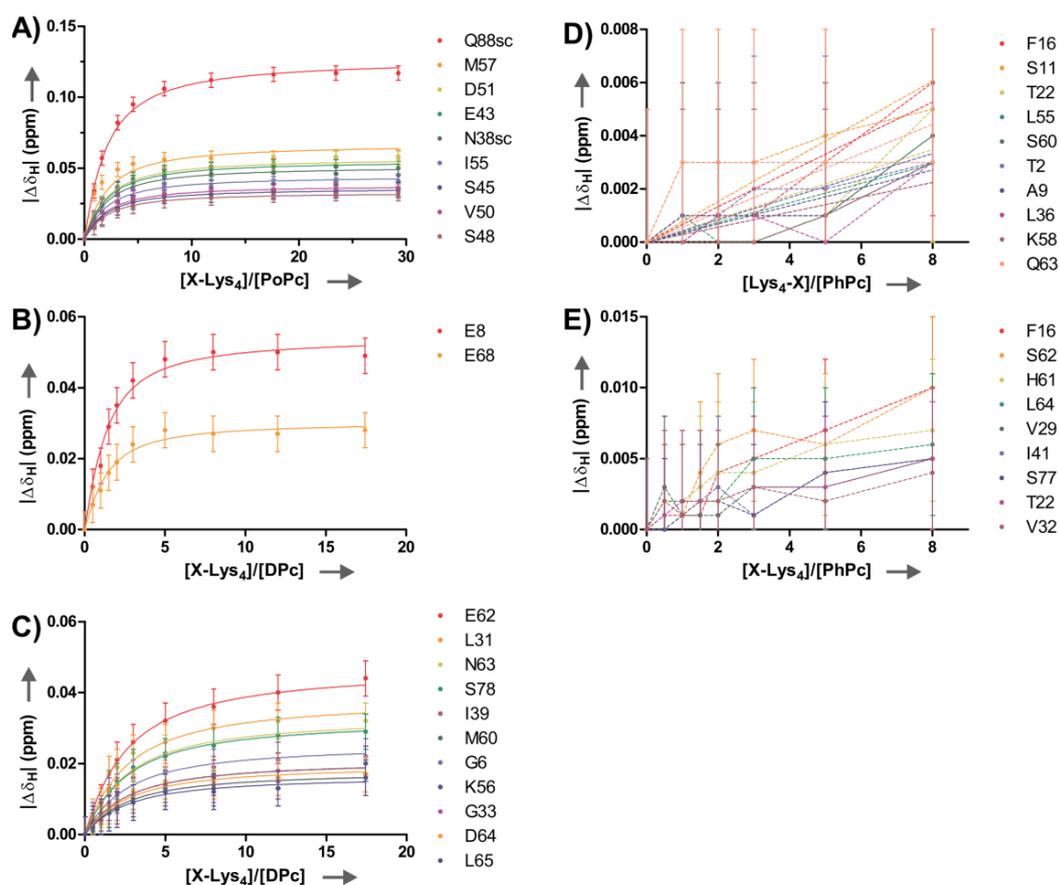


Figure S2: Chemical shift changes of Pcs resonances as a function of increasing [peptide]/[Pc] for peptides X-Lys₄ and Lys₄-X. The dissociation constants of the corresponding peptides (Table 1) were obtained by simultaneous fitting to a 1:1 binding model for PoPc (solid lines) and by simulation for 2-site binding for DPc. (A) X-Lys₄ with PoPc; (B) X-Lys₄ with DPc, strong-binding residues; (C) X-Lys₄ with DPc, weak-binding residues; (D) Lys₄-X with PhPc; (E) X-Lys₄ with PhPc. The titration points for each residue in (D) and (E) are connected with dashed lines. Error bars represent ± 0.005 ppm.

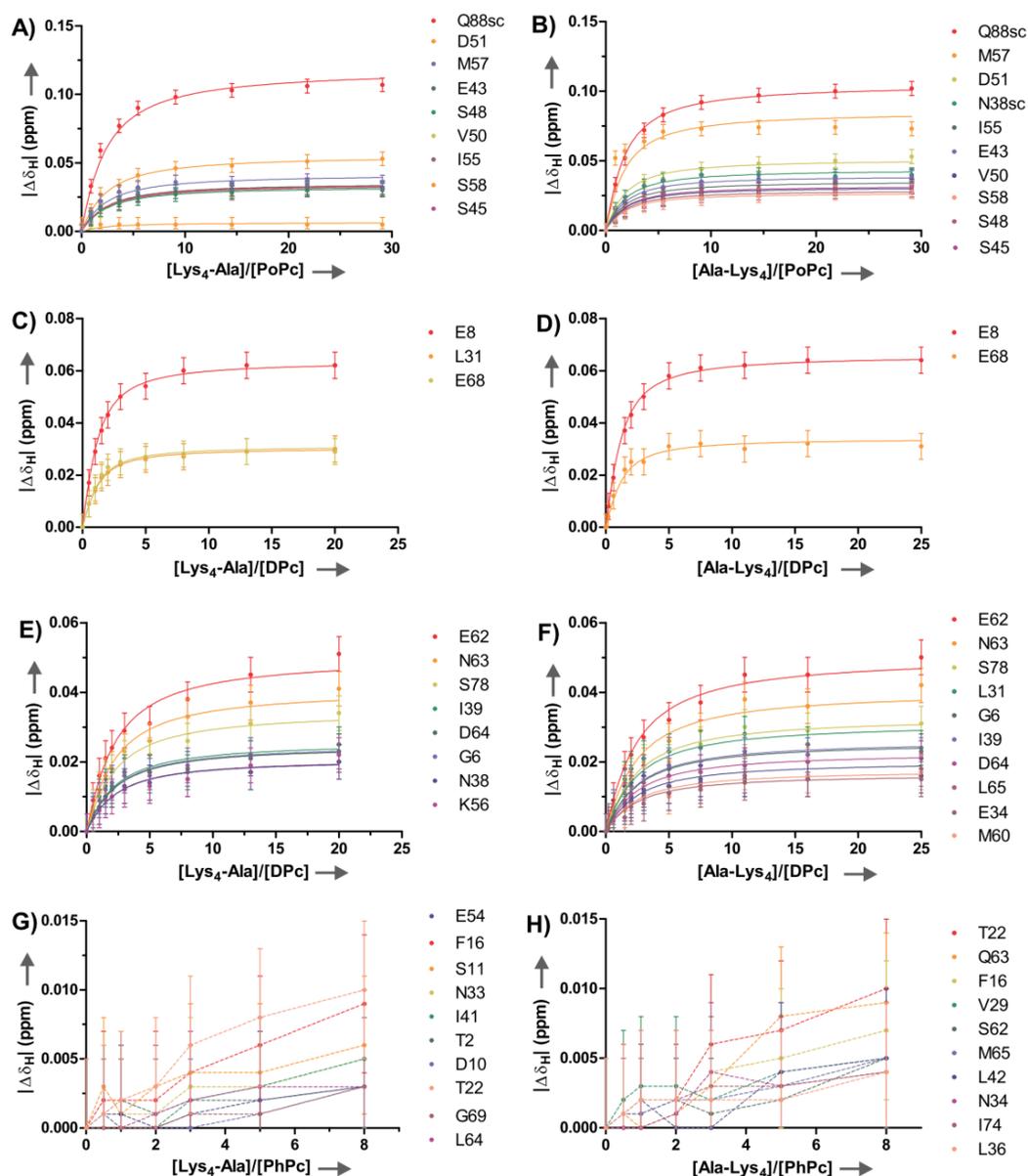


Figure S3: Chemical shift changes of Pcs resonances as a function of increasing $[\text{peptide}]/[\text{Pc}]$ for peptides Lys₄-Ala and Ala-Lys₄. The residues which showed largest perturbations are shown. The dissociation constants of the corresponding peptides (Table 1) were obtained by simultaneous fitting to a 1:1 binding model for PoPc (solid lines) and by simulation for 2-site binding for DPc. (A) Lys₄-Ala with PoPc; (B) Ala-Lys₄ with PoPc; (C) Lys₄-Ala with DPc, strong-binding residues; (D) Ala-Lys₄ with DPc, strong-binding residues; (E) Lys₄-Ala with DPc, weak-binding residues; (F) Ala-Lys₄ with DPc, weak-binding residues; (G) Lys₄-Ala with PhPc; (H) Ala-Lys₄

with PhPc. The titration points for each residue in (G) and (H) are connected with dashed lines. Error bars represent ± 0.005 ppm.

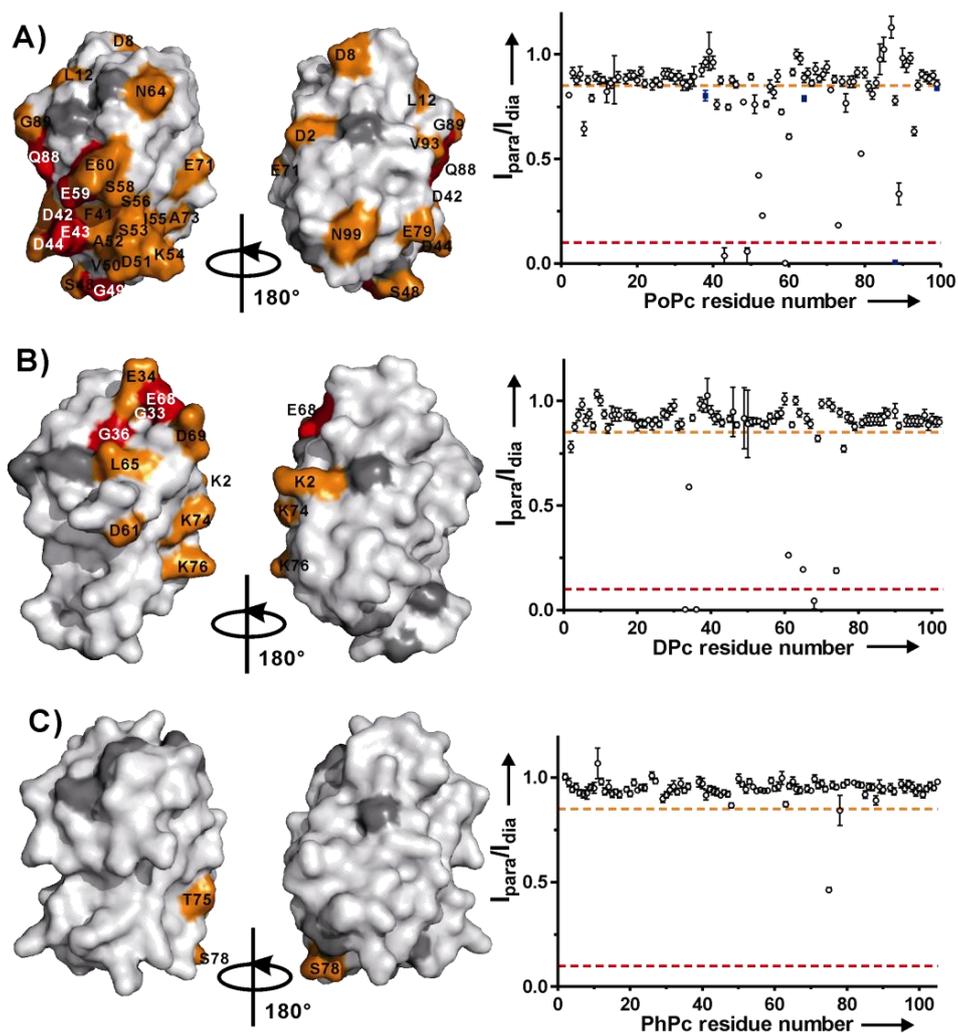


Figure S4: PRE effects in Pc- X-Lys₄ complexes. Left: PRE maps of PoPc (A, PDB entry 1TKW^[1]), DPc (B, PDB entry 1KDI^[2]) and PhPc (C, PDB entry 2Q5B) bound to X-Lys₄ peptide, color-coded on surface models of Pc: red, $I_{\text{para}}/I_{\text{dia}} < 0.1$; orange, $0.1 \leq I_{\text{para}}/I_{\text{dia}} < 0.85$; white, $I_{\text{para}}/I_{\text{dia}} \geq 0.85$; grey, prolines, unassigned, and overlapping resonances. Right: relative $[^1\text{H}, ^{15}\text{N}]$ -HSQC intensities of amides PoPc (A), DPc(B) and PhPc(C) in the complex with TOAC-containing peptides. For PoPc, the side chains are also included (blue squares). The dashed horizontal lines indicate $I_{\text{para}}/I_{\text{dia}} =$

0.85 (orange lines) and 0.1 (red lines). The error bars denote $2\times$ standard deviations, derived from spectral noise levels using standard error propagation procedures.

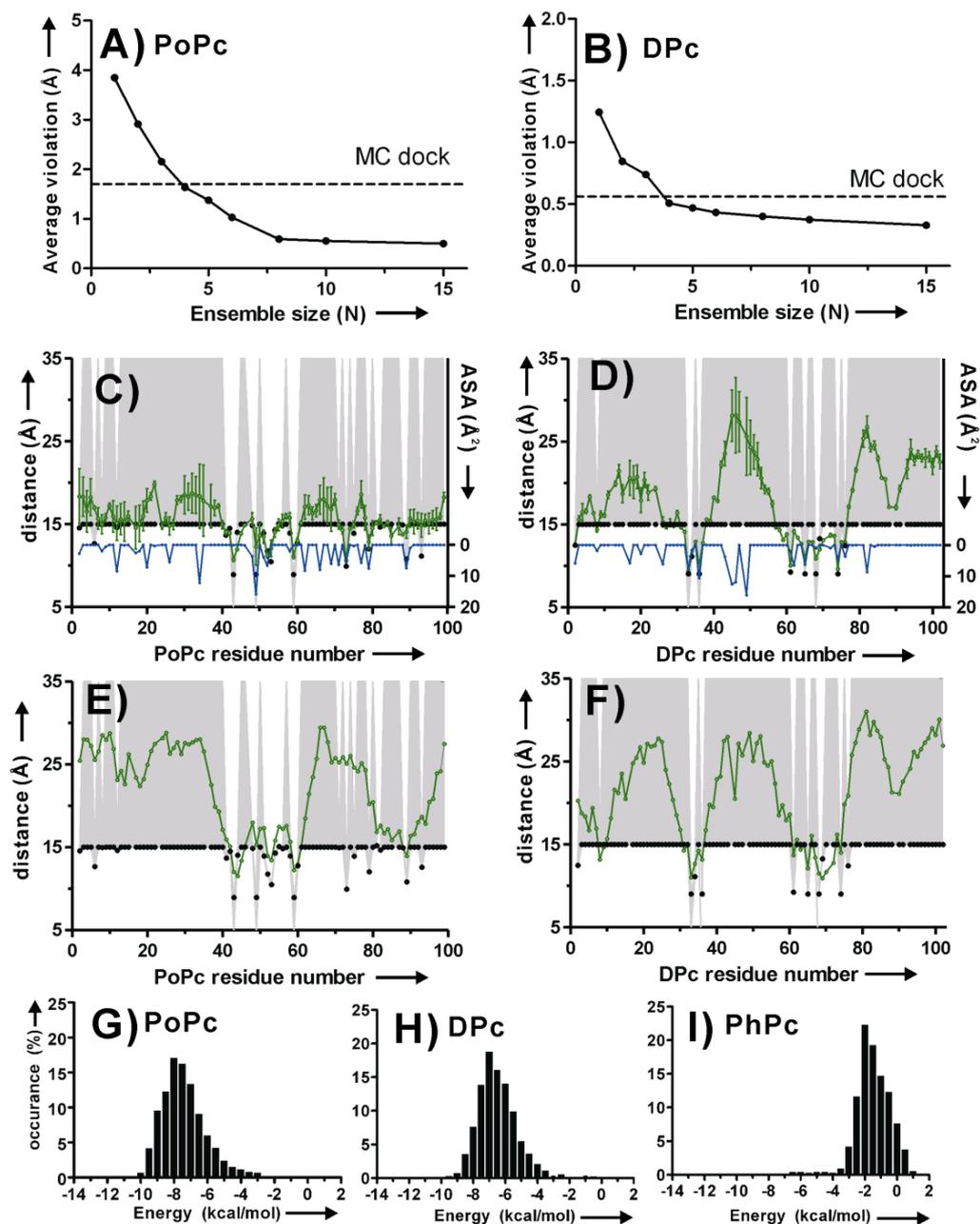


Figure S5: (A-B) Averaged distance violations against number of X-Lys₄ peptides (N=1-6,8,10,15) in the ensemble docking for PoPc (A) and DPc (B). (C-D) Correlation of experimental distances (black dots) and back-calculated average distances (green circles with connecting lines) from the ensemble docking (N=6) of

X-Lys₄ bound to PoPc (C) and DPc (D). The average distances from the 20 lowest-energy solutions of the PRE driven ensemble docking are shown as black circles connected by black lines with error bars representing the standard deviation. Right y axes show the accessible surface area (ASA) of each amide. Grey areas indicate the error margins of the experimental distances. (E-F) Comparison of experimental distances (black dots) and back-calculated average distances (green dots with connecting lines) between Pc amides and the 2000 ensembles of peptide paramagnetic oxygen atoms from MC simulations for PoPc (E) and DPc (F). Grey areas indicate the error margins of the experimental distances. (G-I) Histograms showing the energy distribution of 2000 ensembles from MC simulations: (G) PoPc-X-Lys₄, (H) DPc- X-Lys₄, (I) PhPc- X-Lys₄.

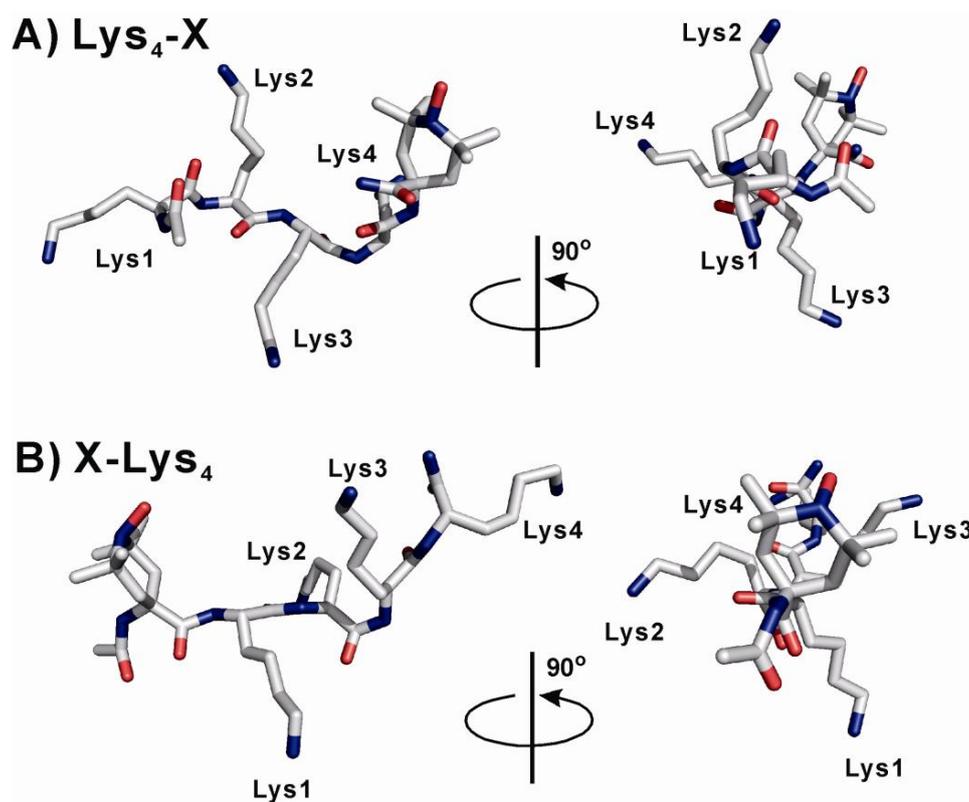


Figure S6: Modeled structures of tetralysine peptides with TOAC (X). (A) Lys₄-X; (B) X-Lys₄. The conformations were optimized in Swiss PDB-Viewer to separate the charges as far as possible. The position of each Lys residue is indicated as Lys1-Lys4.

References

- [1] C. Lange, T. Cornvik, I. Díaz-Moreno, M. Ubbink, *Biochim. Biophys. Acta* **2005**, *1707*, 179–188.
- [2] T. Kohzuma, T. Inoue, F. Yoshizaki, Y. Sasakawa, K. Onodera, S. Nagatomo, T. Kitagawa, S. Uzawa, Y. Isobe, Y. Sugimura, et al., *J. Biol. Chem.* **1999**, *274*, 11817–11823.

Chapter 8

Manuscript C

Role of hydrophobic interactions in the encounter complex formation of the plastocyanin and cytochrome *f* complex revealed by paramagnetic NMR spectroscopy

Sandra Scanu, Johannes M. Foerster, G. Matthias Ullmann
and Marcellus Ubbink

J. Am. Chem. Soc. 2013, 135 (20), 7681-7692

doi: 10.1021/ja4015452

Reprint with permission of

Journal of the American Chemical Society.

© 2013 Copyright by American Chemical Society.

Role of Hydrophobic Interactions in the Encounter Complex Formation of the Plastocyanin and Cytochrome *f* Complex Revealed by Paramagnetic NMR Spectroscopy

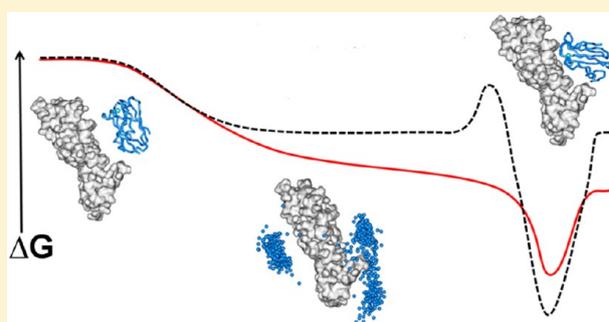
Sandra Scanu,[†] Johannes M. Foerster,[‡] G. Matthias Ullmann,[‡] and Marcellus Ubbink^{*†}

[†]Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

[‡]Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany

Supporting Information

ABSTRACT: Protein complex formation is thought to be at least a two-step process, in which the active complex is preceded by the formation of an encounter complex. The interactions in the encounter complex are usually dominated by electrostatic forces, whereas the active complex is also stabilized by noncovalent short-range forces. Here, the complex of cytochrome *f* and plastocyanin, electron-transfer proteins involved in photosynthesis, was studied using paramagnetic relaxation NMR spectroscopy. Spin labels were attached to cytochrome *f*, and the relaxation enhancements of plastocyanin nuclei were measured, demonstrating that a large part of the cytochrome *f* surface area is sampled by plastocyanin. In contrast, plastocyanin is always oriented with its hydrophobic patch toward cytochrome *f*. The complex was visualized using ensemble docking, showing that the encounter complex is stabilized by hydrophobic as well as electrostatic interactions. The results suggest a model of electrostatic preorientation before the proteins make contact, followed by the formation of an encounter complex that rapidly leads to electron-transfer active conformations by gradual increase of the overlap of nonpolar surface areas on cytochrome *f* and plastocyanin. In this model the distinction between the encounter and active complexes vanishes, at least in the case of electron-transfer complexes, which do not require a high degree of specificity.



INTRODUCTION

A general model for protein–protein interactions describes protein association as a stepwise process in which the formation of the final complex is preceded by that of a transient, lowly populated state encounter complex.¹ In the first step of association, when the freely diffusing proteins approach each other, they are steered toward certain encounter orientations by long-range electrostatic interactions. In the encounter state, proteins still show few specific interactions. They rather tend to assume multiple orientations to sample the surface of the partner and reduce the dimensionality of the search for the specific binding site.² The final complex is dominated by short-range, specific interactions, which stabilize it in a single orientation. An encounter complex will not always proceed toward the final complex. In some cases it is futile and will dissociate again.^{3,4} The applicability of this model has been theoretically and experimentally demonstrated for complexes with electrostatic-assisted association, both when the interaction partners showed high overall charge complementarity^{5,6} and when opposite charges are more localized in specific regions on the surface of the reactants.^{7,8} At the same time, given the wide variety in the electrostatic surface properties of proteins, this model cannot readily be generalized for all protein

complexes. For complexes in which the interaction partners do not present charge complementarity or apparent dipolar interactions, a desolvation-mediated association has been suggested on the basis of theoretical work,^{9,10} in which hydrophobic interactions guide both encounter complex formation and stabilization of the final complex. Experimental evidence of a predominantly hydrophobic-driven binding event is rare, although some has been reported.^{11,12}

To investigate the contribution of the different forces involved in the molecular recognition process for transient complexes involved in electron transfer (ET), the complex formed by plastocyanin (Pc) and cytochrome *f* (Cyt *f*) from the cyanobacterium *Nostoc* sp. PCC 7119 was studied. Pc and Cyt *f* are redox partners in oxygenic photosynthesis in plants, green algae, and cyanobacteria. Pc transfers electrons from Cyt *f* of the cytochrome *b₆f* complex to photosystem I (PSI).¹³ In both proteins the redox active sites are buried below extensive hydrophobic surface patches, which form the specific binding site. The overall electrostatic properties of the proteins vary significantly between different species and influence the final

Received: February 12, 2013

Published: April 29, 2013

orientation of the complex. In plants, the final complex has been shown to be electrostatically stabilized in a “side-on” orientation by complementary localized charges on the protein surfaces, negative in Pc and positive in Cyt *f*.^{14,15} Complementary charges tilt Pc toward the long side of Cyt *f* and align the hydrophobic binding sites, thus facilitating the ET reaction. In the cyanobacterium *Phormidium laminosum*, the final complex assumes a “head-on” orientation.¹² Pc is oriented perpendicular to the heme plane and comes into contact with Cyt *f* only with the hydrophobic patch. The “side-on” orientation was also observed in the complexes from the cyanobacteria *Nostoc*¹⁶ and *Prochlorothrix hollandica*,¹⁷ in which the charges are inverse compared to the plant counterparts, being positive in Pc and negative in Cyt *f*. Site-directed mutagenesis of key interface residues important for the overall electrostatic potential of the proteins from *Nostoc* demonstrated that electrostatic interactions heavily regulate the kinetics of complex formation.^{18,19} Interestingly, the loss of negatively charged residues in Cyt *f*, in which charges are spread over a large part of the surface instead of being localized in a specific region, caused only small changes of the association rate constant,¹⁹ whereas mutations of positively charged residues in a conserved region in Pc showed these charges to be fundamental for fast association.¹⁸ The surface charge properties of Pc and Cyt *f* appear to influence the degree of dynamics within the complexes.²⁰ The *Ph. laminosum* and *Pr. hollandica* complexes showed to be highly dynamic.^{12,17} The solution structures of the above-mentioned complexes have been obtained by taking advantage of the pseudocontact shift (PCS) caused by the paramagnetic heme iron of Cyt *f* on backbone amide protons of Pc. PCS from the heme are not very sensitive to dynamics, although the presence of many different Pc orientations will cause a decrease in the observed average PCS. In a recent paramagnetic relaxation enhancement (PRE) NMR study on the *Nostoc* system, we demonstrated that the Pc–Cyt *f* complex is more dynamic than was suggested by PCS, indicating the presence of a significantly populated encounter state.²¹

PRE NMR spectroscopy has proven to be a sensitive technique for the detection and visualization of lowly populated intermediates in protein–DNA²² and protein–protein complexes.^{8,23,24} PREs arise from magnetic dipolar interactions between the unpaired electron of a paramagnetic center and the observed nucleus, which causes an increase in the relaxation of the latter. Due to the large magnetic moment of the unpaired electron and the inverse sixth power distance dependence of the PRE, it is very large for nuclei that spend time in close proximity of the paramagnetic center. The sensitivity of PRE for lowly populated states is due to the fact that in the NMR fast exchange regime the observed PRE rate is a population weighted average of all species present in solution.²⁵ If in the lowly populated state the nucleus is close to the paramagnetic center, a PRE can be detected on the exchange averaged signal, even if the population is as low as 1%. Intermolecular PREs can thus be used to investigate transient intermediates in protein–protein complexes.²³ The observed PREs provide explicit qualitative evidence of the presence of the encounter state, but they do not provide a complete description of the encounter complex. Furthermore, the visualization of the encounter complex is an ‘inverse’ problem since many possible solutions can correspond to the observed PREs. To depict the encounter complex, experimental data need to be supported by theoretical models, generated by computational approaches. In the

ensemble docking approach, multiple conformers of a protein are simultaneously docked to the other protein on the basis of the experimental PRE.²³ The encounter complex is visualized as an ensemble of orientations that fit the experimental restraints. In purely theoretical methods, such as Brownian dynamics (BD) and Monte Carlo (MC) simulations, in which proteins are docked only on the basis of electrostatic interactions,^{7,26} the encounter complex is given as a distribution of the favorable electrostatic orientations.⁸

In our previous study on the *Nostoc* complex,²¹ three spin labels were attached to Cyt *f* at sites surrounding the binding site for plastocyanin. It was demonstrated that those spin labels did not affect the Pc chemical shift perturbation (CSP) map caused by binding to Cyt *f*. The PRE data were not in agreement with a single binding orientation of Pc, because the affected amide groups on Pc were very similar for each of the spin labels, despite their different locations on Cyt *f*. This finding showed that Pc was sampling multiple orientations within the complex. To map the range of the encounter orientations of Pc, in the present study the number of spin label positions was extended to nine. The results were used to visualize the encounter complex by ensemble docking, and this model was compared with MC simulations. The comparison indicates that long-range electrostatic interactions preorient Pc with the hydrophobic patch toward Cyt *f* and that Pc maintains the same orientation while sampling the surface of Cyt *f*. The ensemble docking visualization of the encounter complex showed that the encounter complex is stabilized by hydrophobic as well as electrostatic interactions. We propose a model for Pc–Cyt *f* complex formation in which long-range electrostatic interactions preorient the unbound proteins before they make contact. Pc diffusively binds to Cyt *f* forming an extended encounter complex stabilized by the overlap of the respective nonpolar surface areas, and the encounter complex rapidly evolves to ET active conformations. The ambiguous distinction between the encounter and the active complex in this system will be discussed in the context of the physiological cytochrome *b₆f* complex.

■ EXPERIMENTAL SECTION

Protein Production and Purification. ¹⁵N enriched, Zn-substituted Pc was produced and purified as described before.²¹ The concentration of the protein was determined by absorbance spectroscopy using $\epsilon_{280} = 5 \text{ mM}^{-1} \text{ cm}^{-1}$. The yield of pure protein was 10 mg/L of culture.

The pEAF-WT plasmid, containing the gene of the soluble domain (residue 1–254) of *Nostoc* sp. PCC7119 Cyt *f* was kindly provided by Prof. Dr. Miguel A. De la Rosa (University of Seville). Cyt *f* mutants were obtained using pEAF-WT plasmid as template for mutagenesis. The mutations to cysteine were introduced by using the QuikChange Site-Directed Mutagenesis kit (Stratagene). The primers used for the mutations at the positions N71, Q104, and S192 were described before.^{21,27} The primers employed for the introduction of a cysteine at the positions Q7, Q38, A63, Q125, S181, and Q242 are reported in Table S1.

Truncated Cyt *f* was produced in *E. coli* MV1190 (D(*lac-proAB*), *thi*, *supE*, D(*srl-recA*) 306::Tn10 (tet^r) [F':*traD36*, *proAB*+, *lacI^qZΔM15*]), transformed with pEAF-WT or mutant plasmids, and cotransformed with pEC86, containing a cassette for *c*-type cytochrome overexpression.²⁸ Production and purification of the protein and spin label attachment of (1-acetoxy-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) methanethiosulfonate (MTS) or (1-oxyl-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) methanethiosulfonate (MTSL) were performed as previously reported.^{18,21} The yield of protein production ranged from 1 to 2 mg/L of culture. The expression of

Q125C mutant was not reproducible, and only a small amount of protein was obtained and used for NMR experiments. The concentration of the protein was determined by absorbance spectroscopy using $\epsilon_{556} = 31.5 \text{ mM}^{-1} \text{ cm}^{-1}$ for ferrous Cyt *f*.

NMR Experiments. All NMR samples contained MES (20 mM, pH 6) and 6% D₂O for lock. The ferric state of Cyt *f* was preserved by addition of K₃[Fe(CN)₆] (50 μM). The pH of the sample was adjusted with small aliquots of HCl (0.5 M) and NaOH (0.5 M). For the chemical shift perturbation experiments Cyt *f* was titrated into Zn-substituted ¹⁵N Pc (50 μM). Spectra were recorded at multiple Cyt *f*:Pc molar ratios (0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 2.5, 5.0). CSP experiments were not performed for Q125C Cyt *f* because of lack of protein. Samples for PRE measurements contained 33 μM Cyt *f* for the Q125C mutant and 66 μM for the other mutants, labeled with either MTS or MTSL. Samples also contained Zn-substituted ¹⁵N Pc, 100 μM in the complex with Q125C Cyt *f* and 200 μM for the other Cyt *f* mutants. All NMR spectra were recorded at 298 K on a Bruker Avance III 600 MHz spectrometer equipped with a TCI-Z-GRAD Cryoprobe. The ¹H–¹⁵N HSQC spectra were acquired with 1024 and 80 complex points in the direct and indirect dimensions, respectively.

NMR Data Analysis. The NMR spectra were processed with NmrPipe²⁹ and analyzed with CcpNMR Analysis.³⁰ Chemical shift perturbation analysis was carried out as described before.²¹

The PREs were determined according to the procedure described by Battiste and Wagner.³¹ The intensity ratio I_p/I_d of the Pc resonances in the presence of MTSL–Cyt *f* (I_p) and MTS–Cyt *f* (I_d) was normalized by dividing them by the average value of the 10 largest I_p/I_d values (1.28 for Q7C, 1.63 for Q38C, 1.16 for A63C, 1.13 for N71C and Q104C, 1.37 for Q125C, 0.83 for S181C, and 1.06 for S192C and 0.92 for Q242C). The PRE (Γ_2) values were calculated according to the formula:

$$\frac{I_p}{I_d} = \frac{R_{2d} \exp(-\Gamma_2 t)}{R_{2d} + \Gamma_2} \quad (1)$$

R_{2d} represents the transverse relaxation rate in the diamagnetic sample, which was calculated from the line width at half height obtained from a Lorentzian peak fit in the direct dimension, by using FUDA (this software was kindly provided by Dr. D. Fleming Hansen, University College London). The symbol t indicates the time for transverse relaxation during the pulse sequence (9 ms). The Γ_2 values were extrapolated to the 100% bound state using the experimentally obtained K_D .

Monte Carlo Simulations of the Encounter Complex. The structure of the soluble part of Cyt *f* (residues 1–254) used for the calculation was taken from the crystal structure of the cytochrome b_6f complex from *Nostoc* sp. PCC 7120, PDB entry 2ZT9.³² The amino acid sequences of Cyt *f* from *Nostoc* sp. PCC 7120 and sp. PCC 7119 are identical. The structure file for Pc was taken from the PDB entry 2GIM.³³ The hydrogen atoms were added with the module HBUILD³⁴ of Charmm.³⁵ To preserve the original structure, only the hydrogen atoms were minimized with the Charmm force field,³⁶ while the other atoms were kept fixed in their original position. The iron of Cyt *f* and the copper of Pc were considered to be in the oxidized state, like in the experiments. The electrostatic potentials for the single proteins were calculated with APBS.³⁷ The dielectric constants for Cyt *f* and the water were set to 4 and 80, respectively. For all electrostatic potentials, a box with a diameter of 225 Å in x , y , and z directions, with Cyt *f* centered at the origin of the coordinates frame, was defined. The ionic strength was set to 0.02 M and the temperature to 298 K. The electrostatic potential was calculated with the linearized Poisson–Boltzmann equation.

The docking was performed with the program MC-Dock²⁶ and was carried out in a similar way as was done before.⁸ Cyt *f* was chosen as the receptor, and Pc was the ligand to dock. The simulation consisted of 250 runs with 1×10^6 steps each and was carried out at a temperature of 298 K. Only structures that respected the Metropolis MC criterion³⁸ were saved resulting in about 2.3×10^6 Cyt *f*:Pc orientations. The main difference to the previous simulation consisted in the use of an inclusion grid. The inclusion grid was created by

defining a grid with a distance to the surface of Cyt *f* of 3 Å and a grid point separation of 0.5 Å. If any atom of Pc is located within this inclusion grid, the structure was included in the final encounter ensemble, otherwise the orientation was not considered. An ensemble of 5000 Pc orientations, randomly selected, was considered for the calculations. The averaged distances were derived from the ensemble and compared to the experimental distances.

Ensemble Docking. Mutations and spin labels were modeled on the structure of Cyt *f* (PDB entry 2ZT9),³² and four conformations were used to represent the mobility of the spin label.³⁹ The structure of Pc was taken from PDB entry 2GIM.³³

The Γ_2 were converted into distances for structure calculations using eq 2:

$$r = \sqrt[6]{\frac{\gamma^2 g^2 \beta^2}{20 \Gamma_2} \left(4\tau_c + \frac{3\tau_c}{1 + \omega_h^2 \tau_c^2} \right)} \quad (2)$$

Where r is the distance between the oxygen atom of MTSL and the Pc amide proton, γ is the proton gyromagnetic ratio, g is the electronic g -factor, β is the Bohr magneton, ω_h is the Larmor frequency of the proton, and τ_c is the rotational correlation time of the MTSL oxygen-proton vector. τ_c was taken to be 30 ns on the basis of the HYDRONMR⁴⁰ prediction of the rotational correlation time for the Pc–Cyt *f* complex.

The restraints for the calculations were obtained according to eq 3:

$$\Gamma_2^{\text{obs}} = f_1 \Gamma_2^{\text{ens}} + f_2 \Gamma_2^{\text{final}} \\ f_1 + f_2 = 1 \quad (3)$$

The ensemble Γ_2 (Γ_2^{ens}) was calculated as the difference between observed Γ_2 (Γ_2^{obs}) and back-calculated Γ_2 from the model of the final complex (model 1, PDB entry 1TU2) (Γ_2^{final}). The calculations were carried out with f_2 values of = 0, 0.15, 0.25, 0.35, 0.5, 0.65, 0.75, 0.85, 0.95, and 1. The restraints were grouped into three classes as described before.²¹

A description of the encounter complex was obtained with restrained rigid-body docking in Xplor-NIH 2.9.9⁴¹ to minimize the difference between observed and back-calculated distances for all spin labels. Calculations were carried out using either a single Pc conformer or an ensemble of Pc molecules, with between 2 and 20 copies. The distances (r) between an amide proton and the oxygen atom of MTSL were r^{-6} averaged for all MTSL orientations and all Pc conformers. Cyt *f* and Pc were both considered as rigid bodies, the coordinates of Cyt *f* were fixed, and Pc ensemble members were allowed to move individually in a restrained rigid-body molecular dynamics calculation. Overlap of Pc copies was allowed, since the ensemble represents a distribution of states. Similarly, overlap of MTSL conformers with other MTSL or Pc copies was allowed. For the visualization of the final encounter complex ensemble 150 dockings were performed, yielding 144 ensembles of 7 Pc conformers, with a difference in the total restraint energy $\leq 20\%$.

The ensembles from separate dockings were evaluated by calculating the average violation over all experimental distances. Class 1 and 3 restraints are not easily expressed in a Q value. Violations provide a better representation of the fit of all three classes of restraints. Class 2 violations were defined as the absolute difference between experimental and calculated distances for a certain amide nucleus. Class 1 and 3 violations were defined as that difference only for back-predicted distances that were above 14 Å and below 23 Å, respectively. The ensemble violation is the average violation for all residues and all spin labels.

RESULTS

Introduction of Paramagnetic Probes on Cyt *f*. To determine the extent of surface area of Cyt *f* being sampled by Pc in the encounter complex, 9 cysteine mutants of Cyt *f* were made for the attachment of nitroxyl spin labels. The Cyt *f* mutants were created for the positions Q7, Q38, A63, N71,

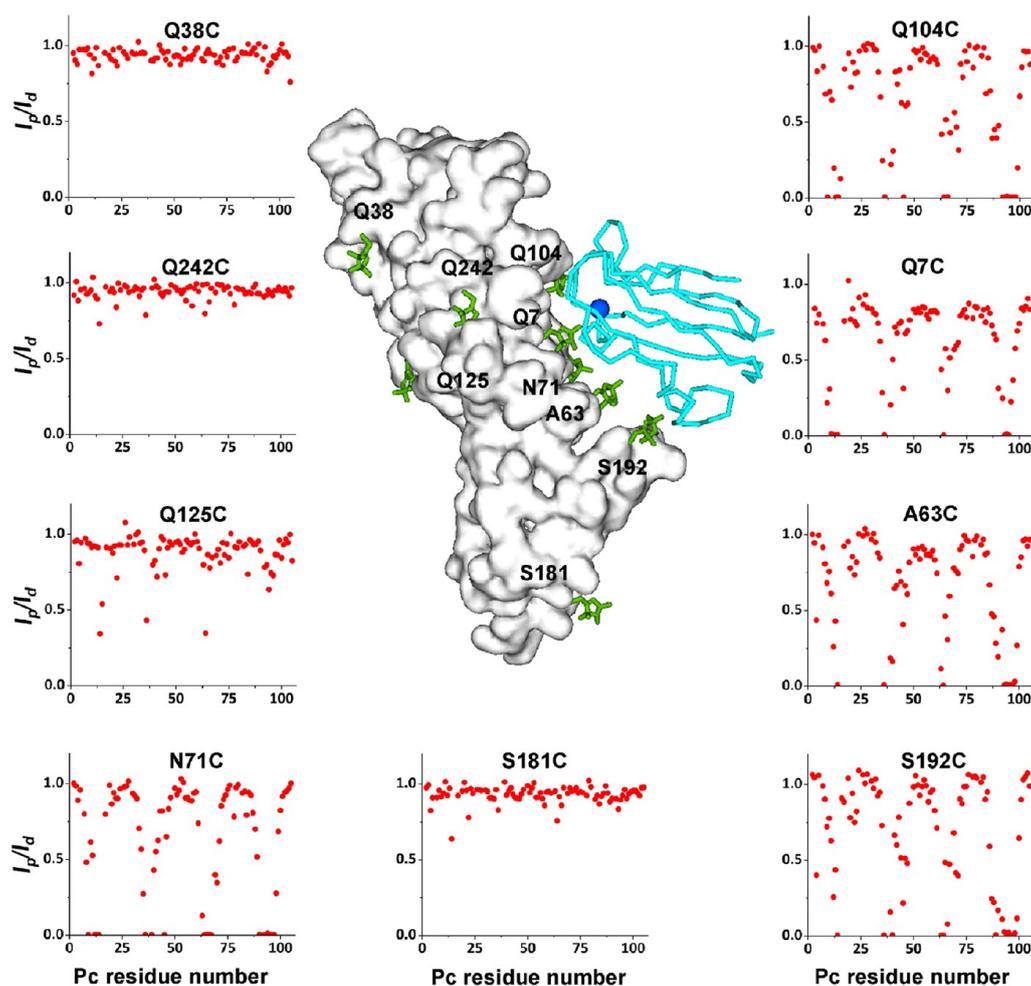


Figure 1. PRE in the Pc–Cyt *f* complex. Central panel. Location of the spin labels (green sticks) modeled on the *Nostoc* sp. PCC 7119 Pc–Cyt *f* complex (PDB entry 1TU2, model 1).¹⁶ Pc is shown in cyan C_{α} trace and the copper as a blue sphere. Cyt *f* is shown as white surface. This image and others of molecular structures were made with Discovery Studio Visualizer 2.5 (Accelrys). Side panels. The I_p/I_d ratios (red dots) are plotted against the Pc residue number for each of the spin label position on Cyt *f*.

Q104, Q125, S181, S192, and Q242. The mutation sites Q7, A63, N71, Q104, and S192 are near to the Pc binding site indicated by the solution model,¹⁶ whereas the remaining four mutations are located elsewhere (Figure 1, central panel). To preserve the original electrostatic potential of Cyt *f*, only polar, uncharged amino acids and one Ala were selected for mutation to cysteine. The copper in ¹⁵N Pc was substituted by Zn^{II} to eliminate the paramagnetic effect and possible interference of the ET reaction caused by the presence of Cu^{II}.⁴²

To test whether the presence of spin label interferes with the Pc–Cyt *f* binding, CSP analysis was performed for Pc bound to Cyt *f* wild-type and mutants conjugated to the diamagnetic control label MTS. Cyt *f* was thus titrated into a solution of ¹⁵N Zn-Pc and HSQC spectra were acquired at each titration point. The CSP curves for the most affected residues were fitted to obtain a dissociation constant for each complex (Figure S1). The K_D values are listed in Table 1.

The K_D value for the wt complex of $8 \pm (3) \times 10^{-5}$ M is similar to the reported values of 4×10^{-5} M for Cu-Pc⁴³ and 6×10^{-5} M for Cd-substituted Pc.⁴³

Most of Cyt *f* variants yielded K_D values within the experimental error of that of the wild-type. Moreover, the

Table 1. Dissociation Constants of the Complexes Formed by *Nostoc* Zn-Pc with Wild-Type and MTS-Conjugated Cyt *f*^a

Cyt <i>f</i> mutant	K_D (10^{-5} M)
wild-type	8 (3)
Q7C-MTS	5 (1)
Q38C-MTS	2 (1)
A63C-MTS	2 (1)
N71C-MTS	4 (1)
Q104C-MTS	3 (1)
S181C-MTS	6 (2)
S192C-MTS	4 (1)
Q242C-MTS	9 (2)

^aThe errors are indicated in parentheses and represent the precision of the fit.

binding maps, obtained by coloring the protein residues according to the size of CSP, present a pattern similar to the wild-type, indicating that the mutations and the attachment of MTS at these positions cause no significant effects on the affinity of Pc for Cyt *f* and orientation of Pc with the respect to Cyt *f* in the complex. In each case Pc binds predominantly via

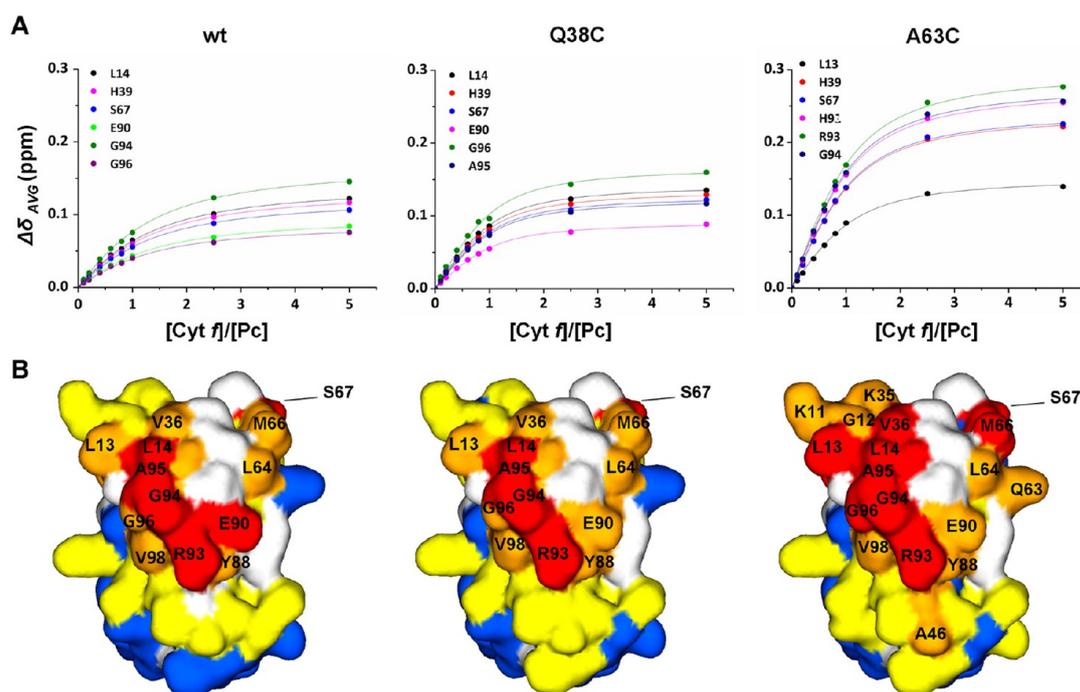


Figure 2. The interaction of *Nostoc* Zn-substituted Pc with wild-type Cyt *f* and Q38C and A63C MTS-conjugated variants. (A) Binding curves for selected residues were fitted globally to a 1:1 binding model.²¹ (B) Chemical shift perturbation maps color-coded on a surface model of Pc (PDB entry 2GIM),³³ with red, $\Delta\delta_{\text{avg}} \geq 0.10$ ppm; orange, $\Delta\delta_{\text{avg}} \geq 0.05$ ppm; yellow, $\Delta\delta_{\text{avg}} \geq 0.02$ ppm; and blue, $\Delta\delta_{\text{avg}} < 0.02$ ppm. Prolines and residues with overlapping resonances are in white.

the hydrophobic patch and the region around Arg 93 (Figure S2), similarly to previously reported data on Cd-substituted Pc in the presence of reduced Cyt *f*.¹⁶ The Q38C-MTS and A63C-MTS Cyt *f* variants exhibited K_D values of $2 \pm (1) \times 10^{-5}$ M, which represent a small increase of binding affinity. In the case of Q38C-MTS Cyt *f*, the average size of the CSP and binding map were similar to wild-type. The spin label position is located far from the final binding site (Figure 1), so the reason for the lower K_D values remains unclear. Larger perturbations of the resonance positions were observed for binding of Pc to Cyt *f* A63C-MTS than in the other studied cases. The largest CSPs were about twice as large as those in the presence of wild-type Cyt *f* (Figure 2A). Interestingly, the binding map is still similar to that of wt, although the effects of binding are stronger (Figure 2B). Under the assumption that CSPs predominantly represent the final state,^{44,45} this observation suggests that Pc binds Cyt *f* A63C-MTS in the same orientation as wt Cyt *f* but that the final state is more populated and the encounter state less.

Paramagnetic Relaxation Enhancements. To determine whether PREs could arise from unspecific interactions, free MTS (diamagnetic) and MTSL (paramagnetic), with a concentration corresponding to that of spin labeled Cyt *f* (66 μM , see below), were mixed with ¹⁵N-labeled Zn-Pc (200 μM). No line broadening of the resonances was observed in the presence of the paramagnetic spin label, indicating that unspecific interactions with the label are not significant under these conditions.

Then, MTSL was attached to each of the nine Cyt *f* mutants, and the tagged proteins were titrated to Pc to a molar ratio of 1:0.3 for Pc:Cyt *f*. At this ratio, the average fraction of Pc bound to Cyt *f* is 24%. The CSP studies indicated that association and dissociation are in the fast exchange regime, so the observed

PREs are a weighted average of free Pc, encounter complex and final complex. Thus, the PREs can be extrapolated to the 100% bound state (encounter complex + final complex) by dividing by the fraction of bound Pc.

Spin labels attached to Cyt *f* near the binding site for Pc in the final complex, namely at the positions Q7, A63, N71, Q104, and S192, caused an extensive broadening of Pc resonances, reflected in a decrease of the I_p/I_d ratio, the ratio of peak intensities in the spectra of the paramagnetic and diamagnetic samples (Figure 1). Surprisingly, mutants with the spin label located on the backside of Cyt *f* relative to the binding site also yielded moderate to strong PREs. For three of these mutants, Q38C, S181C, and Q242C, only moderate effects were observed, whereas the spin label at position 125 caused strong PRE on two Pc residues, L14 and L64, which are part of the hydrophobic patch. From the I_p/I_d ratios, the PRE (Γ_2) were determined and extrapolated to 100% bound Pc. The PREs were mapped on the surface of Pc, shown in Figure 3.

The PRE patterns observed in the presence of spin label attached near the main binding site are very similar. This result is surprising, because the labels are located on different sides of Cyt *f* in the structure of the final complex as shown in Figure 1, and thus it is expected that different regions of Pc would be affected. The results suggest that Pc samples an extensive area of Cyt *f* predominantly with one face oriented toward it, since no strong PREs were observed on the back-side of Pc (Figure S3). The maps also resemble the CSP maps in Figure 2, confirming that the hydrophobic patch and the region around Arg 93 are the interaction sites of Pc in the complex with Cyt *f*. The comparison of the PRE maps with the charge distribution map of Pc (Figure 3, top left) indicates that among the residues, which experience most PRE, only the minority is negatively (E90) or positively (K11, K35, and R93) charged,

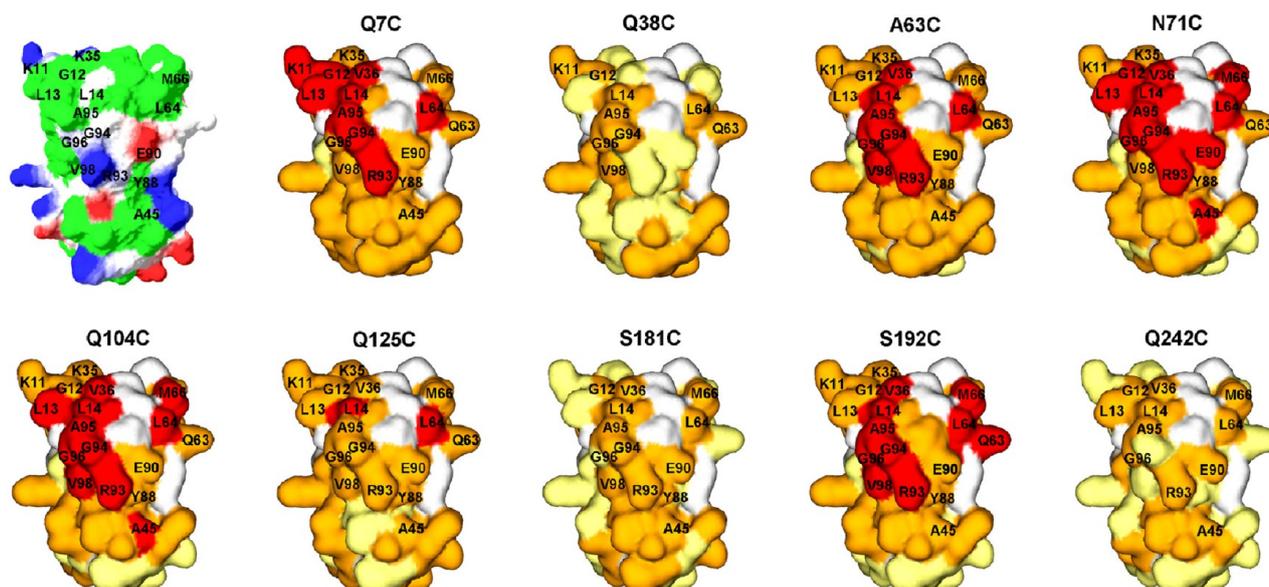


Figure 3. PRE maps of Zn-substituted Pc bound to MTS-conjugated Cyt *f*, color-coded on a surface model of Pc (PDB entry 2GIM). The sites of spin label attachment are indicated in Figure 1, central panel. Red, $\Gamma_2 \geq 200 \text{ s}^{-1}$; orange, $10 \text{ s}^{-1} < \Gamma_2 < 200 \text{ s}^{-1}$, and yellow $\Gamma_2 \leq 10 \text{ s}^{-1}$. Prolines and residues with overlapping resonances are white. Top left, the charge distribution of Pc with negatively and positively charged side chains shown in red and blue, respectively. Hydrophobic side chains are shown in green, and polar side chains are in white.

whereas the majority has a hydrophobic nature. Residues L13, L14, V36, L64, A95, and V98 are part of the hydrophobic patch, which also represents the main binding site and the likely site for ET.¹⁶

MC Simulations. Visualization of the encounter state on the basis of the PRE data is not straightforward, because the data represent a weighted average of all orientations of Pc within the complex, and thus, an infinite number of ensembles can produce the experimental data set. The encounter complex of Cyt *c* and Cyt *c* peroxidase was successfully visualized by combining PRE data and rigid-body MC simulations,⁸ showing that the formation of this encounter complex is solely driven by electrostatic interactions. In MC docking, a mobile protein is docked to a target molecule under the influence of an electrostatic field and MC sampling.²⁶ In this way, charge–charge interactions represent the only force that brings together the proteins. Following the same rationale, MC simulations for Pc–Cyt *f* complex were performed, and the Boltzmann distribution of orientations of Pc in complex with Cyt *f*, and vice versa, was obtained. The centers-of-mass of Pc (Figure S4A) and Cyt *f* (Figure S4B) are shown as blue and green spheres, respectively, around the interaction partner, shown as surface model. In the MC ensemble Pc is widely spread over the surface of Cyt *f* in correspondence with the negative charges distribution. Cyt *f* is overall negative with most charges in the region surrounding the heme on the large domain and lower charge density on the surface opposite of the heme. These results are inconsistent with the PREs observed in the presence of spin labels located at positions far from the heme, such as Q125 (backside) and S181C (small domain). The distribution of Cyt *f* around Pc is off-center from the CSP map obtained for binding to Cyt *f*. Thus, qualitatively the MC ensembles are not in complete agreement with the experimental data.

The observed PREs result from the contributions of both the encounter and the final complexes, whereas the MC ensemble is assumed to represent mostly the encounter state. To separate

the PRE contributions of the two states, the PREs from the solution model of the final complex were back-calculated and subtracted from experimental PREs assuming a population of the final state (f_2) varying from 0 to 1. The resulting PREs represent the encounter state at decreasing population, and these were converted into distances and compared with the average distances calculated from the MC ensemble (see Experimental Section for details). Ensembles composed of the 100, 1000, 2500, and 5000 randomly selected structures were considered for the analysis. Independent of the size of the MC ensemble and of the population of the two states, no good match with experimental data was found (Figure S5). These findings suggest that the formation of the Pc–Cyt *f* encounter complex is not exclusively driven by electrostatic forces. Other contributions must play a significant role, and therefore, MC simulations cannot provide a complete description of this encounter complex.

Ensemble Docking. The quantitative interpretation of PRE for the visualization of transient encounter complexes requires the use of PRE restraints in docking calculations.²⁵ An ensemble of orientations that represents the encounter state and agrees with the data is generated by docking several conformers of a protein (Pc in this work) simultaneously while minimizing the difference between the back-calculated PRE averaged over all conformers and the experimental data. This procedure is repeated many times, and because many different ensembles fit the experimental data, the result is a ‘cloud’ of orientations.^{23,46–51} An ensemble of non-interacting Pc structures was generated, and docking calculations were performed with PRE restraints arising from all nine spin labels at the same time. Though variants Cyt *f* Q38- and A63C-MTSL appeared to have some influence on the affinity or the equilibrium between encounter state and final complex (see above), they were included in the calculations, because we found that the description of the encounter complex was similar, whether or not these restraints were included in the

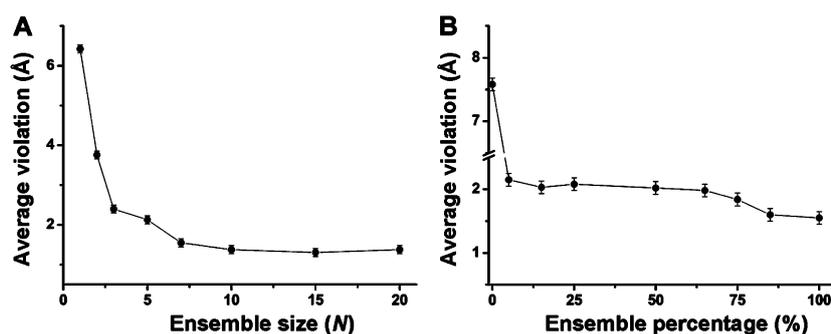


Figure 4. Plot of the average violation of all experimental distances versus the number (N) of Pc copies used in ensemble docking (A) and versus the ensemble percentage included in the restraints for the calculations (B). Error bars represent $2 \times$ SD of the average violations obtained from three independent calculations performed with $N = 1 + 7$ and an encounter percentage of 50% ($f_1 = 0.5$).

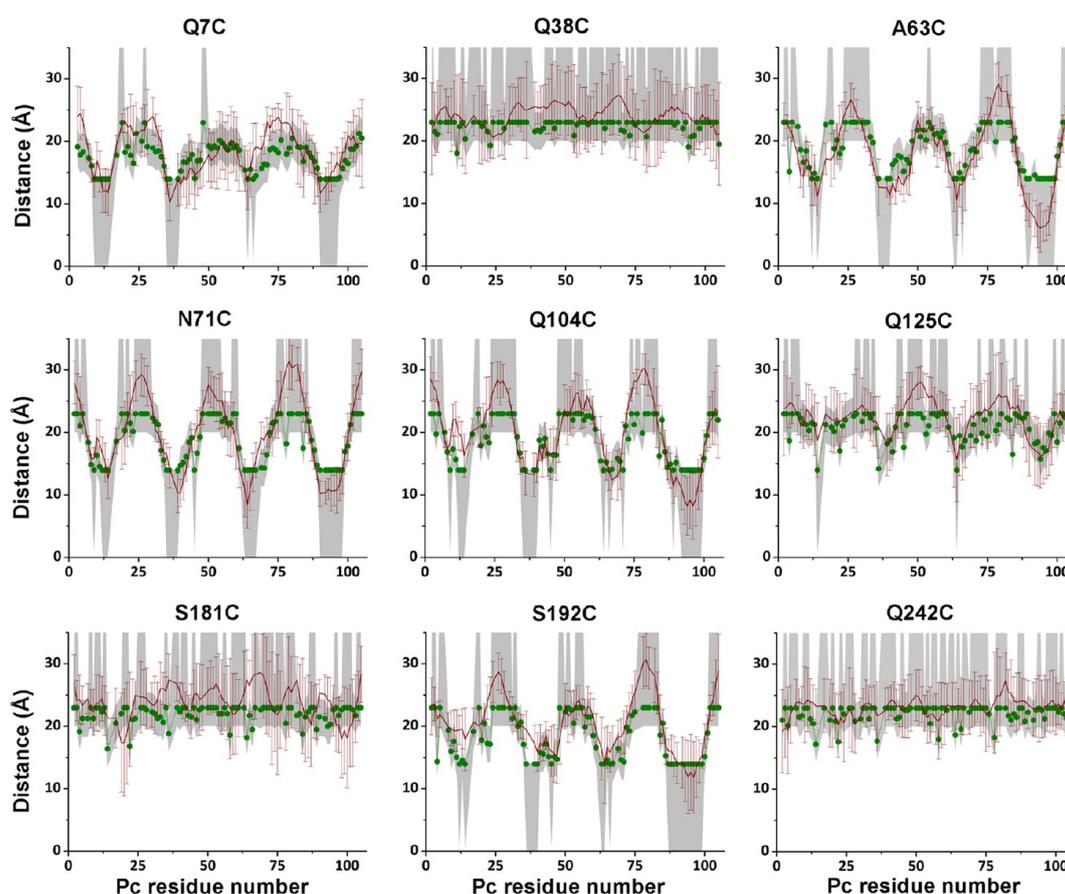


Figure 5. Ensemble docking. Experimental and back-calculated average distances between Pc amide protons and oxygen atoms of MTSL conjugated to Cyt *f* are plotted against the Pc residue number. The green circles and lines represent the experimental distances, and the gray areas indicate the error margins. The average distances back-calculated from 20 ensembles are shown as a red line with error bars representing the SD. Calculations were performed with $N = 7$ and $f_1 = 1.0$.

calculations. When the restraints derived either from variant Cyt *f* A63-MTSL or from both variants Cyt *f* Q38- and A63C-MTSL were excluded from the calculations, the distribution and the fit of the experimental data did not improve, yielding average ensemble violation values of 1.9 and 1.8, respectively, for $N = 7$ and $f_1 = 1$ (see below). In our hands, converting PREs to distances worked best, probably because Pc approaches some spin labels closely. Due to inverse sixth power distance dependence, very small movements at short distance result in

very large PRE changes that skew the outcome of the calculations. A repulsion function to avoid steric collision between Cyt *f* and the Pc molecules was the only other interaction included in the calculations. Calculations were performed by varying the size (N) of the docked ensemble, which ranged from 1 to 20 copies of Pc. The generated ensembles were evaluated by calculating the average violation over all experimental distances (see Experimental Section for details). As can be seen in Figure 4, a large decrease of the

average violation was observed, going from a single copy of Pc up to $N = 7$, while further increase of the number of Pc molecules ($N > 7$) did not improve the fitting. Thus, a combined docking of seven Pc copies simultaneously can produce a population distribution that can mostly satisfy the experimental PREs.

The solution structure of the final complex was previously reported,¹⁶ taking advantage of intermolecular PCS generated by the paramagnetic Fe^{III} of Cyt *f* on Pc nuclei. However, the results in Figure 4A show that a single orientation cannot account for the PRE data. In an earlier study we showed that the model of the final complex is already insufficient to explain the PRE data of only three nearby spin labels.²¹ Thus, the PRE data describe a combination of the final state and encounter complex. Following this rationale, calculations were carried out with $N = 1 + 7$, where 1 represents the final complex and 7 the number of copies in the ensemble.⁴⁸ The contribution of the final complex to the experimental PREs was subtracted to obtain the PREs of the encounter state only, in analogy to what was done for the MC calculations. Thus, the back-calculated PREs from the solution model¹⁶ were subtracted from the experimental values, assuming a population of the final complex (f_2) between 0 and 1. The resulting PREs were converted to distances and used for ensemble docking. Figure 4B presents the average violation as a function of the fraction of encounter state (f_1). The violations unequivocally indicate that the measured PREs do not derive from the final structure alone. Interestingly, a small fraction of encounter state is sufficient to decrease the average violation sharply. The average violation decreases slightly from 2.15 for $f_1 = 0.05$ encounter complex to 1.55 for $f_1 = 1$. In Figure 5 the results from the simulation carried out using seven conformers of Pc ($N = 7$) and assuming a pure encounter state ($f_1 = 1$) are shown.

The ensemble structures were used to back-calculate the distances between the oxygen atom of MTSL and Pc amide protons (red line), and these were compared with the experimental distances (green circles and line). The generated encounter complex fits the experimental data well, being within the error margins for most residues, although small deviations are observed for some residues at several spin label positions. The considerable standard deviations observed for the distances for the generated ensembles (error bars for red line) are noteworthy, because it is a clear illustration that rather different ensembles of seven Pc copies can fit the large experimental data set equally well, emphasizing the nature of the 'inverse problem' mentioned above.

Estimation of the Fraction of the Encounter Complex.

The structure of the final complex was based on experimental PCS, not on PRE, and consists of a single orientation of Pc relative to Cyt *f*, so by reducing the contribution of the final complex, it is expected that it is easier to create an ensemble that matches the experimental PREs. Therefore, the small decrease of the average violation with increasing fraction of the encounter complex (Figure 4B) may not be significant, indicating that the PRE data cannot distinguish between a fraction of the encounter complex of 5% and 100%. Since both PCS and PRE account for minor species present in solution, PCSs were back-calculated for the generated encounter complexes and compared with the experimental PCS data. To correlate experimental and back-calculated PCS, a Q factor (eq S2) was calculated for different fractions of final structure (f_2) (Figure S6). The size of the axial component of the magnetic susceptibility anisotropy ($\Delta\chi_{\text{ax}}$) of Cyt *f* Fe^{III} is not

known precisely, so a range of values was tested. For a final complex only ($f_1 = 0$), the best fit of the PCS is found for $\Delta\chi_{\text{ax}}$ values <50% of the one derived from EPR data, in line with earlier findings.¹⁶ The low-lying excited states for a low-spin ferric heme explain why the g tensor at 10 K cannot readily be used to calculate the $\Delta\chi_{\text{ax}}$ at 298 K. For all but very low values of $\Delta\chi_{\text{ax}}$ the combination of final complex and encounter ensemble (determined using the PRE ensemble docking) improves the fit between experimental and back-calculated PCS. Unfortunately, the lowest Q value achievable is always about 0.1 (see Figure S6), and this minimum is found at increasing values of f_1 for larger $\Delta\chi_{\text{ax}}$ values. Therefore, establishing the encounter state fraction on the basis of PCS is not possible as long as $\Delta\chi_{\text{ax}}$ cannot be established. For comparison, $\Delta\chi_{\text{ax}}$ for Cyt *c* is about $3.3 \times 10^{-32} \text{ m}^3$.⁵² If $\Delta\chi_{\text{ax}}$ of Cyt *f* would be the same, the fraction f_1 would be 0.25.

Visualization of the Encounter Complex. To represent the encounter complex, an ensemble of 144 solutions, for a total of 1008 Pc molecules was generated ($N = 7$; $f_2 = 0$). As shown in Figure 6, Pc visits a large area of Cyt *f*. The density

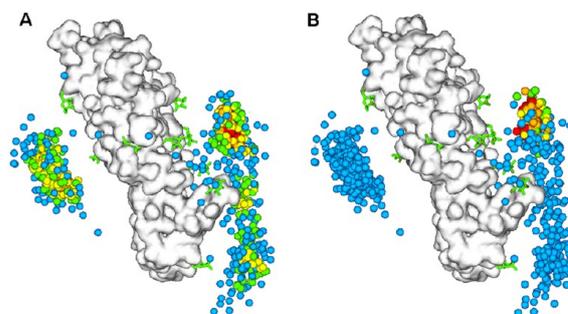


Figure 6. Encounter complex of the Pc–Cyt *f* complex. Cyt *f* is shown as a white surface and spin labels as green sticks. Pc centers-of-mass are represented by spheres. (A) Pc centers of mass are color-coded to indicate the density of the distributions, decreasing from red to blue. Densities were determined by counting the number of neighbors within 2.5 Å. (B) Pc centers-of-mass are color-coded to indicate the distance between Cu in Pc and Fe in Cyt *f*, increasing from red to blue (red ≤ 16 Å; orange ≤ 18 Å; yellow ≤ 20 Å; green ≤ 22 Å; blue > 22 Å).

plot (Figure 6A) shows the centers-of-mass of Pc colored according to the density, with red and blue representing the most and least populated positions, respectively.

Two defined encounter locations can be distinguished. The first one is close to the binding site in the final complex, and the second on the other side of Cyt *f*, opposite to the binding site in the final complex. These locations are discussed further below. It should be noted that even with nine spin labels, it was not possible to sample the encounter state at all locations on the Cyt *f* surface sufficiently. The cytochrome is a very elongated protein, and from the spin label positions in Figure 1 (central panel), the regions that were not sampled can be identified. It cannot be excluded that the area sampled in the encounter complex is still larger than the surface area covered in our experiments. Nevertheless, from the current analysis, it is obvious that Pc samples quite a significant fraction of the surface of its partner.

It is thought that in ET systems the formation of the encounter complex reduces the dimensionality of the search for the active site and increases the probability of ET.² To determine which of the encounter complex orientations were

compatible with rapid ET, the distances between Cu, in Pc, and Fe, in Cyt *f*, were calculated, and the centers-of-mass of Pc were color-coded accordingly (Figure 6B). Structures with Cu–Fe distance ≤ 16 Å (red dots), thus in principle suitable for fast ET,⁵³ are exclusively located in front of the heme, in the vicinity of the final structure.

DISCUSSION

The structure of Pc–Cyt *f* final complex from the cyanobacterium *Nostoc* sp. PCC7119 was solved by NMR on the basis of PCS data.¹⁶ The results were consistent with a conformation that accounts for the ET reaction between the two proteins. A recent PRE study²¹ on the same system demonstrated that the complex exists partly in a dynamic ensemble of orientations. That study was limited by the number and location of the spin labels, which were close to the site of the specific binding of Pc. It inspired the current study, in which we characterized the encounter complex of Pc and Cyt *f* by attaching the spin label MTSL to nine sites on Cyt *f*. MTSL is a small hydrophobic molecule, and its presence in particular regions of the protein could influence the complex formation with the interaction partner. Some interference by the spin label in the Pc–Cyt *f* complex formation was observed only at the position Q38 and A63. The exclusion of the data either from one (A63) or both (Q38 and A63) the spin labels had little influence on the results. PREs were mainly observed for spin labels situated near to the binding site indicated by the PCS-based model, although effects were also measured for the remaining spin label positions. The detection of widespread PREs clearly indicates that Pc samples a large surface area. The similarity of PRE patterns observed in the presence of spin labels close to the binding site (Figure 3) suggests that Pc is approaching Cyt *f* with the residues forming the hydrophobic patch and the region around R93. Independent of the location of the paramagnetic probes, the residues that experience the strongest PREs reside in these regions, implying that Pc is always oriented in the same way toward Cyt *f*.

The general model of protein association states that the formation of the encounter complex is an electrostatically driven process.² On the basis of this assumption, computational approaches, such as BD and MC docking, have been developed to describe the encounter complex as end-point of electrostatic steering toward the interacting partners.^{26,54} These methods successfully described the encounter complex in cases of protein association guided by charge–charge interactions.^{6–8,54–56} The MC simulations of the Pc–Cyt *f* encounter complex did not produce a result in accordance with the experimental data. It is important to note that despite the overall electrostatic potential of Pc and Cyt *f* being positive and negative, respectively, Pc shows a defined charge distribution, but Cyt *f* has a diffuse surface charge. For interaction partners with weak charge complementarity, it has been demonstrated that electrostatics do not play an exclusive role in protein association, but desolvation is the main driving force in binding.⁹ Furthermore, the interface of the encounter complex can bury a significant solvent-accessible surface area, suggesting a role for hydrophobic interactions in the stabilization of the encounter complex.⁵⁷ In the Pc–Cyt *f* complex from the cyanobacterium *Ph. laminosum*, the association is dominated by hydrophobic interactions, and only hydrophobic contacts stabilize the final complex.¹² In the *Nostoc* Pc–Cyt *f* system, the specific binding interface and the putative ET sites on both proteins comprise hydrophobic regions,¹⁶ similar in size and

composition to the *Ph. laminosum* counterparts, but electrostatic forces play a significant role in the association reaction^{18,19} and in the orientation of the final complex.⁴³ Since already in the encounter complex Pc is oriented toward Cyt *f* with its hydrophobic patch, we propose that during the initial stage of the encounter complex formation, long-range electrostatics preorient Pc toward Cyt *f*, and hydrophobic interactions keep Pc close to the surface of Cyt *f* and help to stabilize the encounter state.

The simulation of the encounter complex on the basis of the experimental PREs resulted in two distinct encounters: One is located at the side of the specific binding surface of Cyt *f* and the other one at the opposite side. In Figure 7 the encounter

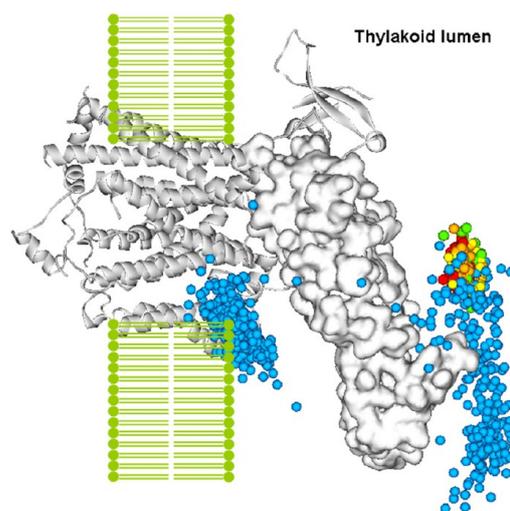


Figure 7. Encounter complex of the Pc–Cyt *f* complex superimposed on the structure of the cytochrome *b₆f* complex (PDB entry 2ZT9), embedded in the thylakoid membrane. Cytochrome *b₆f* complex is a dimer, but only a monomer is shown. Cyt *f* is shown as a white surface, and Pc centers-of-mass are represented by spheres, which are color-coded to indicate the distance between Cu in Pc and Fe in Cyt *f*, as in Figure 6B. The remaining components of the *b₆f* complex are shown as ribbons.

complex is superimposed with the cytochrome *b₆f* complex, as it is found in the thylakoid membrane, with the Pc centers-of-mass colored on the bases of the Cu–Fe distance.

The encounter complex located opposite to the ET site places Pc inside the thylakoid membrane. The Cyt *f* soluble part was shown to be only lightly flexible in the thylakoid lumen,⁵⁸ suggesting that Pc cannot interact with Cyt *f* at this site under the physiological conditions. The observation of encounters at this site is likely an artifact due to the use of the soluble part of Cyt *f* in isolation and not embedded in the thylakoid membrane. The encounters located in front of the binding site diffusely extends from the heme to the end of the small domain of Cyt *f*. In other studies, diffusive encounter complexes have also been described.⁵⁹ Either the partners sample large areas or form encounters at several more defined regions on the surface of the partner.^{3,4,60,61} Due to the diffuse distribution of negative charges on Cyt *f*, Pc probes a large part of the surface, instead of being attracted to specific sites by localized charges.

The formation of an encounter complex in many cases does not lead to a productive association, hence fruitful and futile encounter complexes can be distinguished.³ The encounter

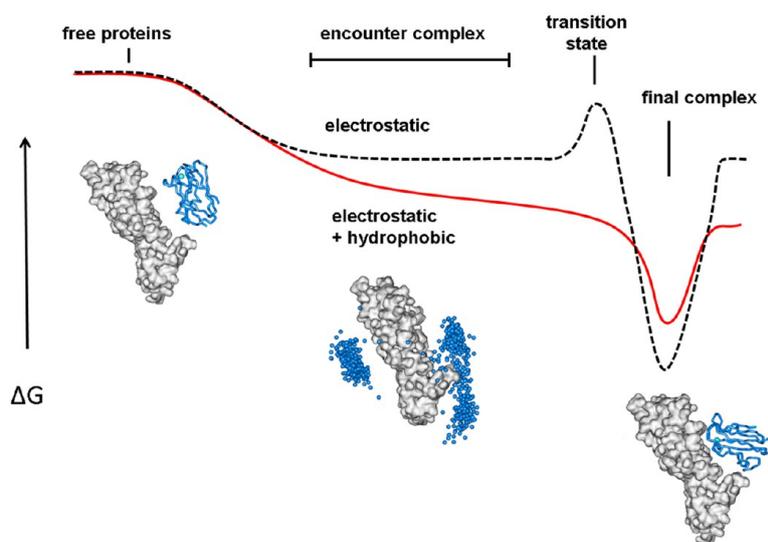


Figure 8. Free energy diagrams of proteins association pathways with high (dashed black line) and no (solid red line) energy barrier for the transition state from the encounter to the final complex.

ensemble shows a high density in the region directly in front of the heme, suitable for ET (red dots in Figures 6B and 7). These orientations can obviously be considered fruitful encounters. For many other encounters, the distances between the redox centers are unfavorable for ET, but not all of these are futile. The further away from the ET active site Pc binds, the less chance it has to diffuse to an ET active orientation before dissociation from Cyt *f*. Thus, the distinction between fruitful and futile encounter is a gradual one, and it is difficult to tell when these complexes can rearrange into a productive orientation.

The population of the encounter state, in relation with the final state population, significantly varies among different complexes. In some non-ET complexes the population of the encounter state has been determined to be 10%,^{23,57} and in the ET complex Cyt *c*–Cyt *c* peroxidase, it has been shown to be 30%.⁸ For Pc–Cyt *f* complex from *Nostoc*, it was not possible to establish the fraction accurately. Complexes that only exist in the encounter state have also been described for myoglobin–Cyt *b*₅⁴⁴ and adrenodoxin–Cyt *c*.⁶⁰ In *Nostoc*, the existence of the complex as merely an encounter state seems unrealistic. The size of CSP in the wild-type complex¹⁶ and in the presence of spin labeled Cyt *f* also in this study is indicative of the formation of a stereospecific complex. Moreover, PCS from the heme generated a converged structure stabilized in a defined orientation,¹⁶ in which hydrophobic contacts and electrostatic interactions are optimized within the structure. The sensitivity of this approach to lowly populated states is limited, but it clearly demonstrates the existence of a final state.

At the same time, the diffuse nature of the encounter complex on Cyt *f* surface suggests that in this system a final orientation may not be a fundamental requirement for the functionality of the complex. In fact, the efficient turnover required for rapid ET through the photosynthetic redox chain⁶² precludes the formation of a tight complex and favors the conditions for the existence of the ET active complex in multiple orientations⁶³ to enhance the probability of ET.² Against this background, the finding that hydrophobic contacts play a role not only in the final complex but also in the encounter complex is interesting. It blurs the distinction

between both states and would allow for a smooth transition from encounter to final complex via a gradual optimization of the hydrophobic contacts in the interface (Figure 8, solid line). This model of protein complex formation allows for more rapid formation of the final complex than in the case of a model with an activation energy barrier between both states (Figure 8, dashed line), used for other protein complexes.³ The interface area of Cyt *f* in encounter complex at the ‘front side’ of Cyt *f* comprises two hydrophobic patches: one located in the big domain (Y1, F3, W4, A63, V68, P118, L119, P120, Y161, and P162), and the other in the small domain (A184, L196, V197, V206, V207, P212, and A213). To test the proposed model, it could be of interest to mutate these residues to polar ones or smaller hydrophobic site chains. Substitution of residues that are only part of the encounter complex is expected to reduce the association rate, whereas mutations close to the heme can disrupt the formation of ET active complexes, reducing the ET rate and shifting the balance toward the non-ET active conformations.

CONCLUSIONS

The combination of paramagnetic NMR spectroscopy and theoretical calculations has enabled the characterization of the encounter complex of Cyt *f* and Pc. With nine spin label positions it was shown that the encounter complex is extended over a large area of Cyt *f* and even includes futile complexes far from the binding site, which are probably nonphysiological. It serves as a reminder that the natural context must be considered when studying the extent of surface areas sampled in the encounter complex. The MC docking calculations indicate that charge interactions play a role but are not dominant; instead hydrophobic contacts appear to guide Pc to the ET active positions on Cyt *f*. We speculate that the hydrophobic interactions in the encounter complex may ensure a relative flat energy landscape during all phases of association, without a clear distinction between the encounter and the active complexes. A flat energy landscape ensures rapid transitions between all states, which is relevant for systems that do not require a high degree of specificity, such as ET complexes.

■ ASSOCIATED CONTENT

● Supporting Information

Description of the primers used for Cyt *f* site-directed mutagenesis; binding curves and CSP maps of Zn-substituted Pc in the presence of Cyt *f*; representation of the MC-simulated Pc–Cyt *f* encounter complex; plots of the experimental and back-calculated distances from the MC-simulated Pc–Cyt *f* encounter; Q factors for experimental and back-calculated PCS. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

m.ubbink@chem.leidenuniv.nl

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

S.S. kindly acknowledges Maryam Hashemi Shabestari and Martin van Son for EPR measurements. S.S. and M.U. received financial support from The Netherlands Organisation for Scientific Research (NWO), grant 700.57.011 (S.S.) and 700.58.441 (M.U.). J.F. was supported by DAAD for the Erasmus Exchange Programme. G.M.U. and J.M.F. were supported by the German Science Foundation (DFG; GRK 1640).

■ REFERENCES

- Schreiber, G. *Curr. Opin. Struct. Biol.* **2002**, *12*, 41–47.
- Ubbink, M. *FEBS Lett.* **2009**, *583*, 1060–1066.
- Harel, M.; Spaar, A.; Schreiber, G. *Biophys. J.* **2009**, *96*, 4237–4248.
- Fawzi, N. L.; Doucleff, M.; Suh, J. Y.; Clore, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1379–1384.
- Schreiber, G.; Fersht, A. R. *Nat. Struct. Biol.* **1996**, *3*, 427–431.
- Gabdouline, R. R.; Wade, R. C. *Biophys. J.* **1997**, *72*, 1917–1929.
- Northrup, S. H.; Boles, J. O.; Reynolds, J. C. *Science* **1988**, *241*, 67–70.
- Bashir, Q.; Volkov, A. N.; Ullmann, G. M.; Ubbink, M. *J. Am. Chem. Soc.* **2010**, *132*, 241–247.
- Camacho, C. J.; Weng, Z.; Vajda, S.; DeLisi, C. *Biophys. J.* **1999**, *76*, 1166–1178.
- Camacho, C. J.; Vajda, S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10636–10641.
- Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021–1025.
- Crowley, P. B.; Otting, G.; Schlarb-Ridley, B. G.; Canters, G. W.; Ubbink, M. *J. Am. Chem. Soc.* **2001**, *123*, 10444–10453.
- Gorman, D. S.; Levine, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *54*, 1665–1669.
- Ubbink, M.; Ejdeback, M.; Karlsson, B. G.; Bendall, D. S. *Structure* **1998**, *6*, 323–335.
- Lange, C.; Cornvik, T.; Diaz-Moreno, I.; Ubbink, M. *Biochim. Biophys. Acta* **2005**, *1707*, 179–188.
- Diaz-Moreno, I.; Diaz-Quintana, A.; De la Rosa, M. A.; Ubbink, M. *J. Biol. Chem.* **2005**, *280*, 35784.
- Hulsker, R.; Baranova, M. V.; Bullerjahn, G. S.; Ubbink, M. *J. Am. Chem. Soc.* **2008**, *130*, 1985–1991.
- Albarran, C.; Navarro, J. A.; Molina-Heredia, F. P.; Murdoch, P. S.; De la Rosa, M. A.; Hervas, M. *Biochemistry* **2005**, *44*, 11601–11607.
- Albarran, C.; Navarro, J.; De la Rosa, M. A.; Hervas, M. *Biochemistry* **2007**, *46*, 997–1003.
- Gross, E. L.; Rosenberg, I. *Biophys. J.* **2006**, *90*, 366–380.
- Scanu, S.; Förster, J.; Finiguerra, M. G.; Shabestari, M. H.; Huber, M.; Ubbink, M. *ChemBioChem* **2012**, *13*, 1312–1318.
- Iwahara, J.; Clore, G. M. *Nature* **2006**, *440*, 1227–1230.
- Tang, C.; Iwahara, J.; Clore, G. M. *Nature* **2006**, *444*, 383–386.
- Suh, J. Y.; Tang, C.; Clore, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 12954–12955.
- Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108–4139.
- Ullmann, G. M.; Knapp, E. W.; Kostic, N. M. *J. Am. Chem. Soc.* **1997**, *119*, 42–52.
- Milikisyants, S.; Scarpelli, F.; Finiguerra, M. G.; Ubbink, M.; Huber, M. *J. Magn. Reson.* **2009**, *201*, 48–56.
- Arslan, E.; Schulz, H.; Zufferey, R.; Kunzler, P.; Thony-Meyer, L. *Biochem. Biophys. Res. Commun.* **1998**, *251*, 744–747.
- Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.
- Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins* **2005**, *59*, 687–696.
- Battiste, J. L.; Wagner, G. *Biochemistry* **2000**, *39*, 5355–5365.
- Baniulis, D.; Yamashita, E.; Whitelegge, J. P.; Zatsman, A. I.; Hendrich, M. P.; Hasan, S. S.; Ryan, C. M.; Cramer, W. A. *J. Biol. Chem.* **2009**, *284*, 9861–9869.
- Schmidt, L.; Christensen, H. E.; Harris, P. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 1022–1029.
- Brunger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148–156.
- Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- MacKerell, A. D.; Bashford, D.; Bellott, D.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kucera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037–10041.
- Metropolis, N.; Ulam, S. *J. Am. Stat. Assoc.* **1949**, *44*, 335–341.
- Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 12800–12808.
- García de la Torre, J.; Huertas, M. L.; Carrasco, B. *J. Magn. Reson.* **2000**, *147*, 138–146.
- Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65–73.
- Ubbink, M.; Lian, L. Y.; Modi, S.; Evans, P. A.; Bendall, D. S. *Eur. J. Biochem.* **1996**, *242*, 132–147.
- Diaz-Moreno, I.; Diaz-Quintana, A.; De la Rosa, M. A.; Crowley, P. B.; Ubbink, M. *Biochemistry* **2005**, *44*, 3176–3183.
- Worrall, J. A. R.; Liu, Y. J.; Crowley, P. B.; Nocek, J. M.; Hoffman, B. M.; Ubbink, M. *Biochemistry* **2002**, *41*, 11721–11730.
- Volkov, A. N.; Bashir, Q.; Worrall, J. A. R.; Ullmann, G. M.; Ubbink, M. *J. Am. Chem. Soc.* **2010**, *132*, 11487–11495.
- Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449*, 1078–1082.
- Tang, C.; Ghirlando, R.; Clore, G. M. *J. Am. Chem. Soc.* **2008**, *130*, 4048–4056.
- Tang, C.; Louis, J. M.; Aniana, A.; Suh, J. Y.; Clore, G. M. *Nature* **2008**, *455*, 693–696.
- Fawzi, N. L.; Doucleff, M.; Suh, J. Y.; Clore, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1379–1384.
- Anthis, N. J.; Doucleff, M.; Clore, G. M. *J. Am. Chem. Soc.* **2011**, *133*, 18966–18974.
- Volkov, A. N.; Ubbink, M.; van Nuland, N. A. J. *J. Biomol. NMR* **2010**, *48*, 225–236.
- Worrall, J. A. R.; Kolczak, U.; Canters, G. W.; Ubbink, M. *Biochemistry* **2001**, *40*, 7069–7076.
- Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L. *Nature* **1992**, *355*, 796–802.
- Gabdouline, R. R.; Wade, R. C. *J. Mol. Recognit.* **1999**, *12*, 226–234.
- Gabdouline, R. R.; Wade, R. C. *J. Mol. Biol.* **2001**, *306*, 1139–1155.

- (56) Spaar, A.; Dammer, C.; Gabdoulline, R. R.; Wade, R. C.; Helms, V. *Biophys. J.* **2006**, *90*, 1913–1924.
- (57) Kim, Y. C.; Tang, C.; Clore, G. M.; Hummer, G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 12855–12860.
- (58) Hasan, S. S.; Cramer, W. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13853–13860.
- (59) Harel, M.; Cohen, M.; Schreiber, G. *J. Mol. Biol.* **2007**, *371*, 180–196.
- (60) Xu, X. F.; Reinle, W. G.; Hannemann, F.; Konarev, P. V.; Svergun, D. I.; Bernhardt, R.; Ubbink, M. *J. Am. Chem. Soc.* **2008**, *130*, 6395–6403.
- (61) Liang, Z. X.; Nocek, J. M.; Huang, K.; Hayes, R. T.; Kurnikov, I. V.; Beratan, D. N.; Hoffman, B. M. *J. Am. Chem. Soc.* **2002**, *124*, 6849–6859.
- (62) Hope, A. B. *Biochim. Biophys. Acta* **2000**, *1456*, 5–26.
- (63) Crowley, P. B.; Ubbink, M. *Acc. Chem. Res.* **2003**, *36*, 723–730.

SUPPORTING INFORMATION TO:

Role of hydrophobic interactions in the encounter complex formation of plastocyanin and cytochrome *f* complex revealed by paramagnetic NMR spectroscopy

Sandra Scanu, Johannes Förster, G. Matthias Ullmann and Marcellus Ubbink

Details on mutagenesis

Table S1 gives the sequences of the primers used for mutagenesis. In each primer a silent mutation (bold) was designed to remove or to introduce an extra restriction site. In the cases of Q7C and Q38C mutations, the codon-changing mutations (bold, underlined) introduced at the same time a restriction site for the enzyme *Apa*LI and removed a restriction site for *Mn*II, respectively. For A63C and A125C mutations, restriction sites for the enzymes *Bst*XI and *Xma*I, respectively, were introduced at the 5' end of the forward primers. In the primers for the S181C mutation, the restriction site for the enzyme *Mn*II was inserted at the 3' end of the forward primer. In the case of the Q242C mutant, the restriction site for the enzyme *Taq*I was introduced next to the codon for the cysteine mutation. The presence of the mutations was verified by DNA sequencing.

Table S1. Nucleotide sequence of the primers used in site-directed mutagenesis of *Cyt.f.* Codon-changing mutations are shown in bold, italic and underlined; silent mutations are in bold.

Mutation	Primer sequence
Q7C	FWD: 5'-gcatatcctttctgggcgca <u>gtgc</u> acttacccag-3'
Q38C	FWD: 5'-gcccacagaagtggaagttc <u>cttgc</u> ctccgtactaccgacaccg-3'
A63C	FWD: 5'-ccagcgtccaacaagttg <u>gtgc</u> gatggctctaagg-3'
Q125C	FWD: 5'-cccggggaacagtat <u>tgca</u> aatcgtctccctgttcttctcccaacccc-3'
S181C	FWD: 5'-gcgctgctgctaccgggtacaatt <u>tgca</u> agattgctaacaagagggcg-3'
Q242C	FWD: 5'-ccctaacgttggtggttcg <u>gtgc</u> ctcgacgcagaaattgttctcc-3'

Calculation of PCS

The average intermolecular PCS from the ferric heme iron of Cyt *f* to the backbone amide atoms in all Pc conformers was calculated and compared with the experimental PCS previously measured in the wild type complex.^{S1} The equation used for the PCS calculation, assuming an axial magnetic susceptibility tensor oriented along the Fe-Y1 vector,^{S2} was:

$$\Delta\delta_{PCS} = \frac{\Delta\chi_{ax}}{12\pi r^3} (3 \cos^2 \theta - 1) \quad (\text{Equation S1})$$

In which $\Delta\delta_{PCS}$ is the size of the PCS, r is the distance between heme iron and observed Pc nucleus, and θ is the angle between Pc nucleus, heme iron and the nitrogen of the amine group of Y1 in Cyt *f*. $\Delta\chi_{ax}$ is the size of the axial magnetic component of the susceptibility tensor, derived from the g-tensor values measured by EPR spectroscopy on plant Cyt *f* and taken to be $7 \times 10^{-32} \text{ m}^3$, as previously reported for *Nostoc* Cyt *f*.^{S1} To correct for the possible difference in tensor size for the temperatures of EPR and NMR measurements, 10 K and 298 K, respectively, the $\Delta\chi_{ax}$ was varied from 0.7 to $8.4 \times 10^{-32} \text{ m}^3$.

The agreement between observed (PCS^{obs}) and calculated (PCS^{calc}) PCS was expressed by the PCS Q factor, defined as:

$$Q_{PCS} = \frac{\sqrt{\sum (PCS^{obs} - PCS^{calc})^2}}{\sqrt{\sum (|PCS^{obs}| + |PCS^{calc}|)^2}} \quad (\text{Equation S2})$$

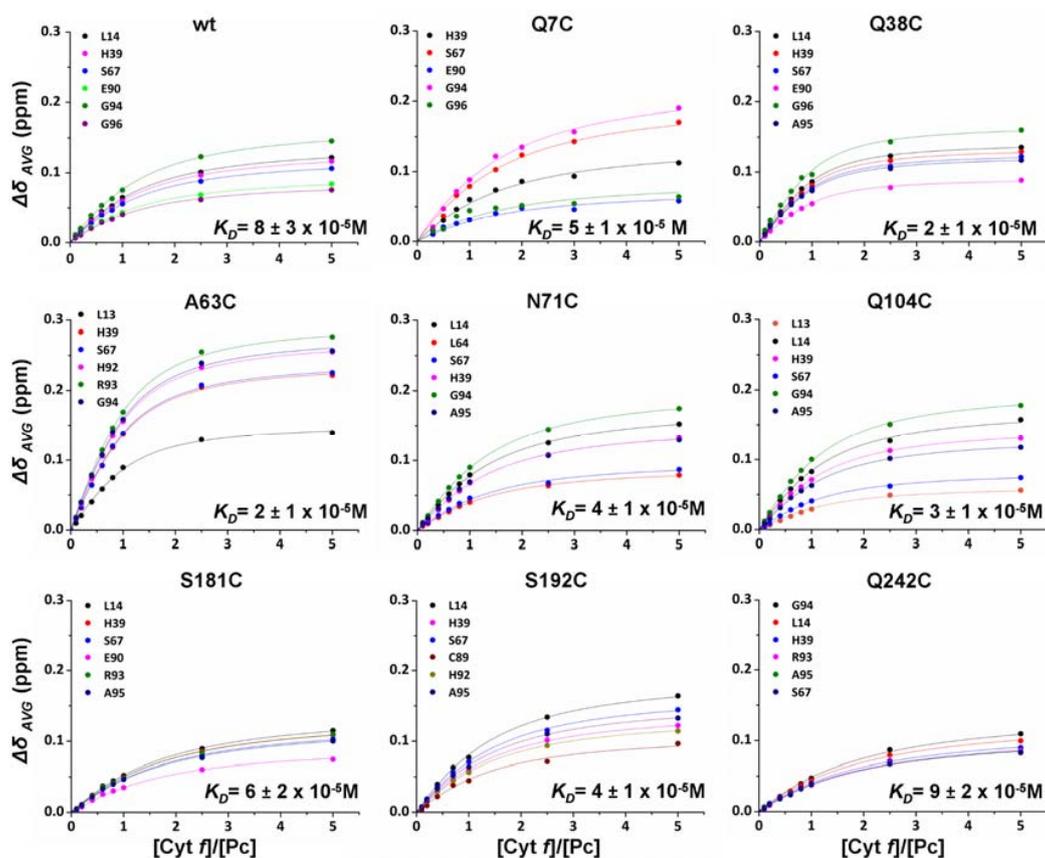


Figure S1. The interaction of *Nostoc* Zn-substituted Pc with wild-type Cyt *f* and MTS-conjugated variants. The binding curves for selected residues were fitted globally to a 1:1 binding model.

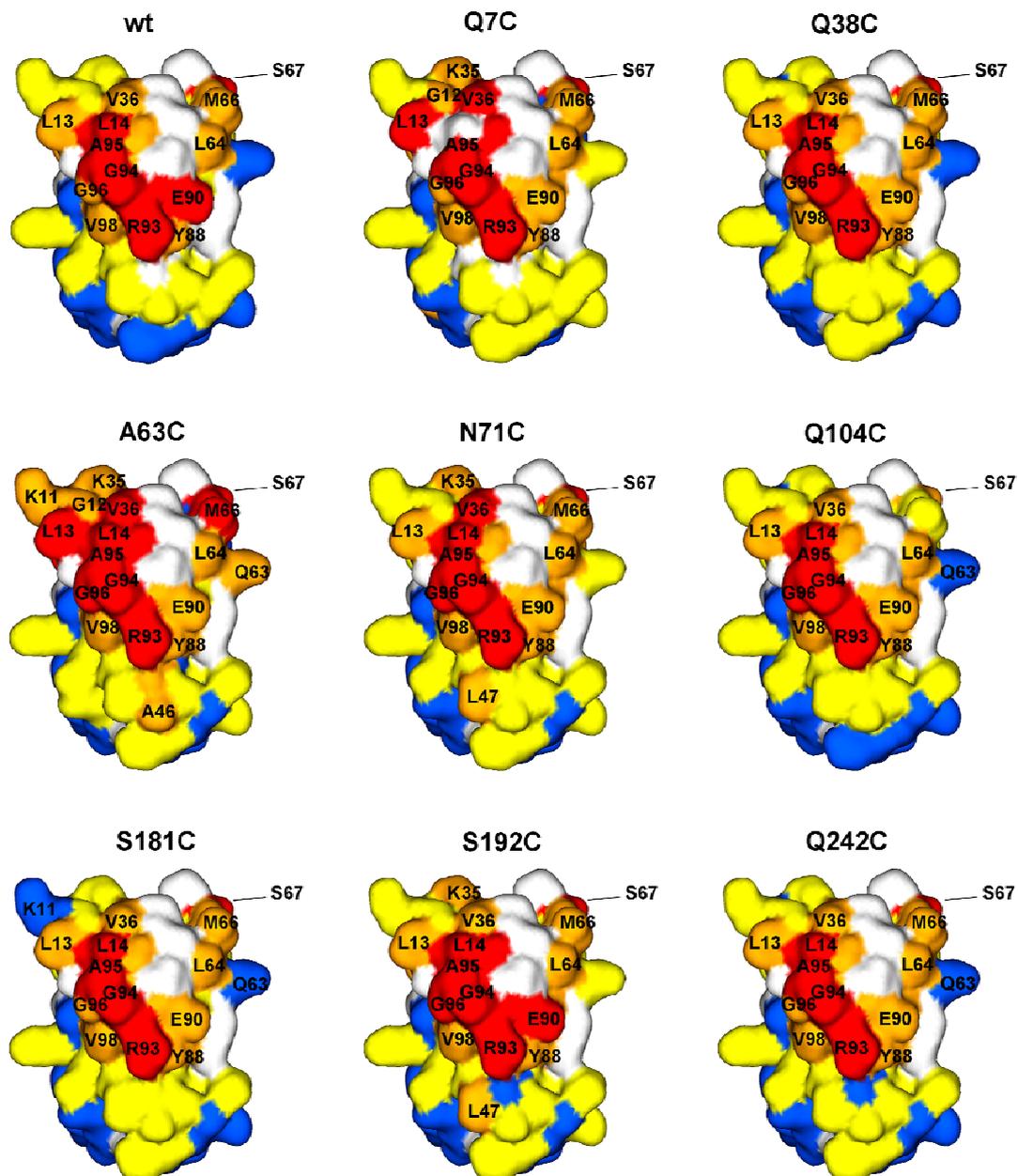


Figure S2. Chemical shift perturbation maps of *Nostoc* Zn-substituted Pc in the presence of wild-type and MTS-conjugated Cyt *f*, colour-coded on a surface model of Pc (PDB entry 2GIM), with red, $\Delta\delta_{avg} \geq 0.10$ ppm; orange, $\Delta\delta_{avg} \geq 0.05$ ppm; yellow, $\Delta\delta_{avg} \geq 0.02$ ppm; blue, $\Delta\delta_{avg} < 0.02$ ppm. Prolines and residues with overlapping resonances are in white.

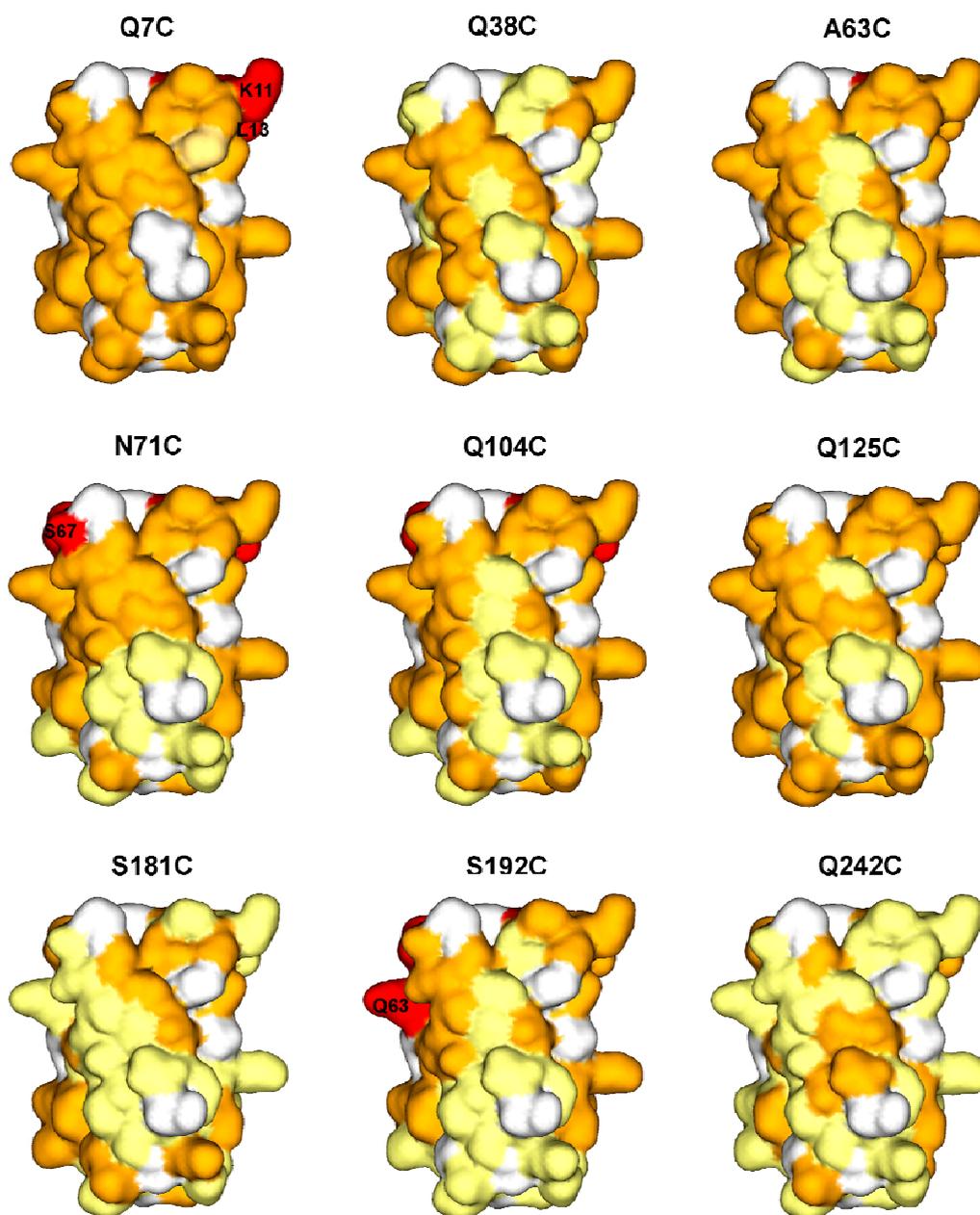


Figure S3. PRE maps of Zn-substituted Pc bound to MTSL-conjugated Cyt *f*, color-coded on a surface model of Pc (PDB entry 2GIM), the sites of spin label attachment are indicated in Figure 1, central panel. Red, $\Gamma_2 \geq 200 \text{ s}^{-1}$; orange, $10 \text{ s}^{-1} < \Gamma_2 < 200 \text{ s}^{-1}$ and yellow $\Gamma_2 \leq 10 \text{ s}^{-1}$. Prolines and residues with overlapping resonances are white.

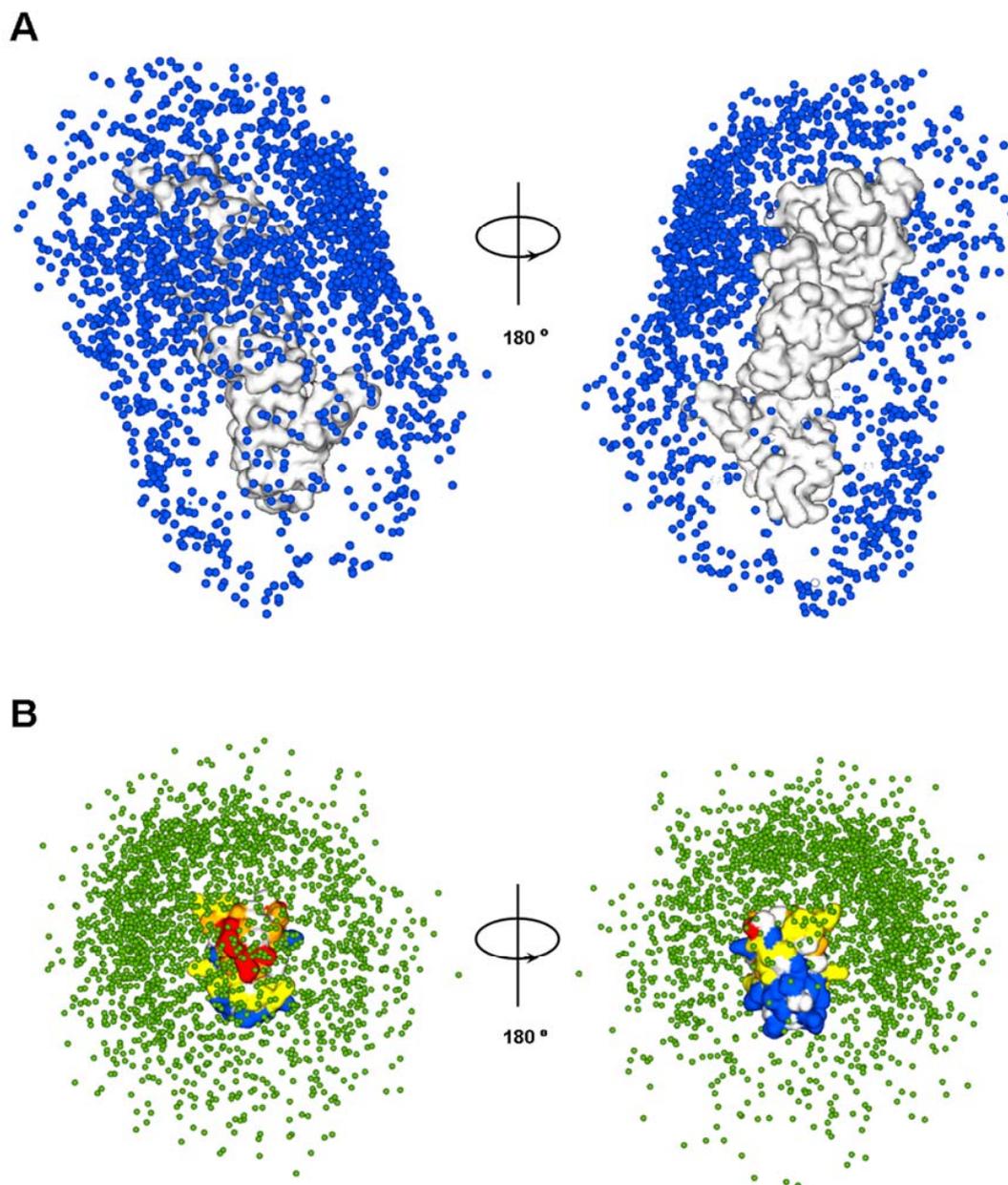


Figure S4. Encounter complex of the *Nostoc* Pc-Cyt *f* complex obtained by random selection of 2000 structures from the MC simulations. A) Cyt *f* is shown as a white surface and Pc centers-of-mass are represented by blue spheres. B) Pc is shown as a surface color-coded according to the CSP in the presence of wild-type Cyt *f* and Cyt *f* centers-of-mass are represented by green spheres.

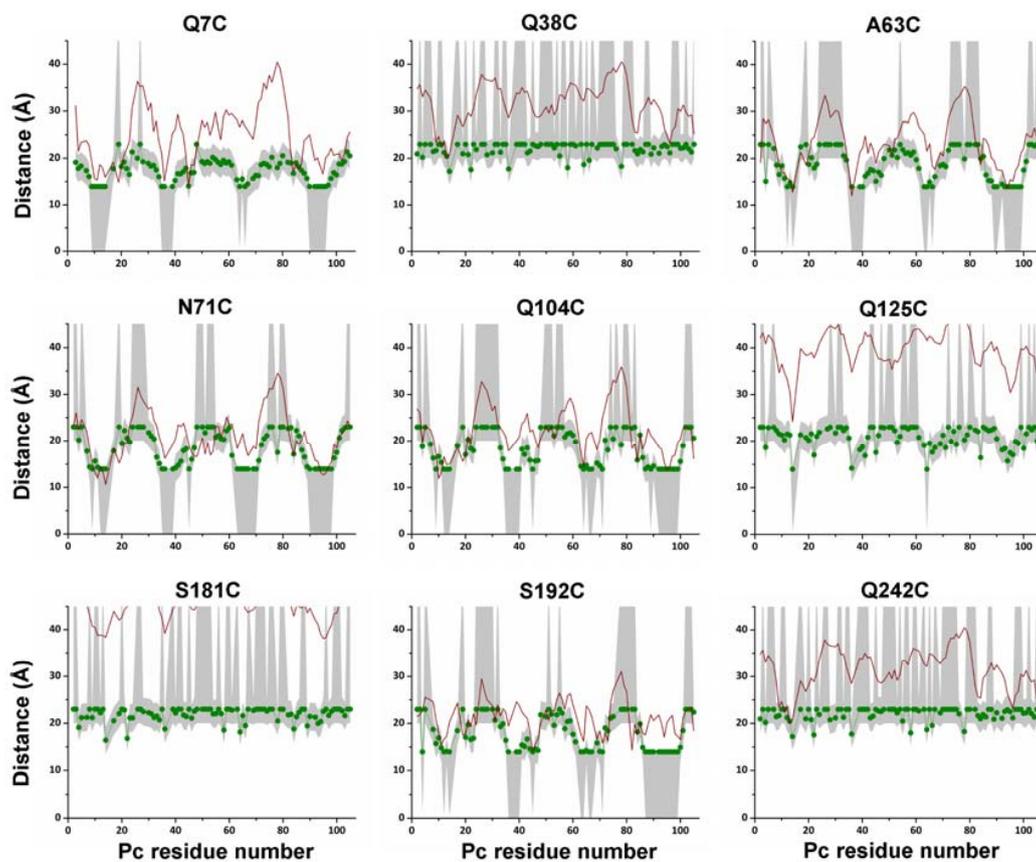


Figure S5. Comparison between back-calculated averaged distances from 2000 randomly selected structures of the MC simulation (red line) assuming $f_i = 1$ and the experimental distances (green circles and lines). The grey areas indicate the error margins of experimental data.

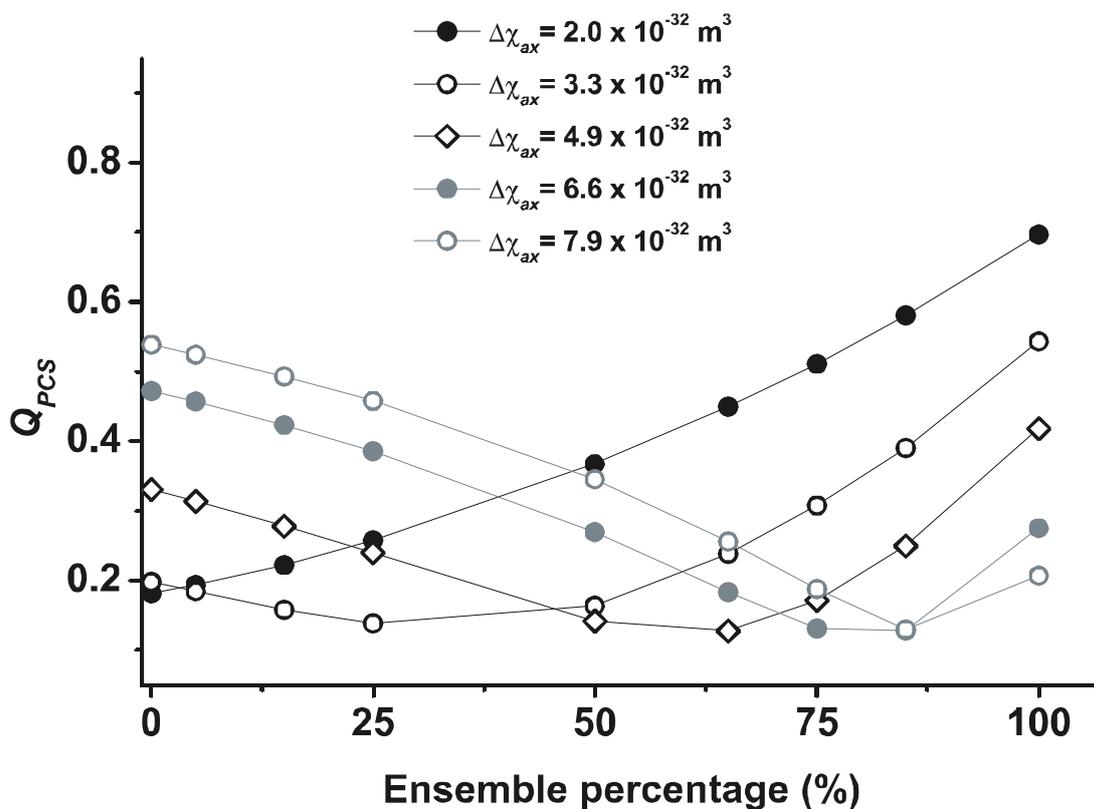


Figure S6. Q factors calculated for a combination of experimental PCS measured for the specific complex and back-calculated PCS from the encounter complex obtained at different fraction population of the encounter complex (f_i). The Q factors were calculated for different values of a scaling factor for the size of the axial component of the magnetic susceptibility tensor.

Reference List

- S1. Diaz-Moreno, I.; Diaz-Quintana, A.; De la Rosa, M. A.; Ubbink, M. *J. Biol. Chem.* **2005**, *280*, 35784.
- S2. Ubbink, M.; Ejdeback, M.; Karlsson, B. G.; Bendall, D. S. *Structure* **1998**, *6*, 323-335.

Chapter 9

Manuscript D

Loss of electrostatic interactions causes increase of dynamics within the plastocyanin-cytochrome *f* complex

Sandra Scanu, Johannes M. Foerster, Monika Timmer,
G. Matthias Ullmann and Marcellus Ubbink
Biochem. 2013, 52 (38), 6615-6626
doi: 10.1021/bi400450q

Reprint with permission of *Biochemistry*.

© 2013 Copyright by American Chemical Society.

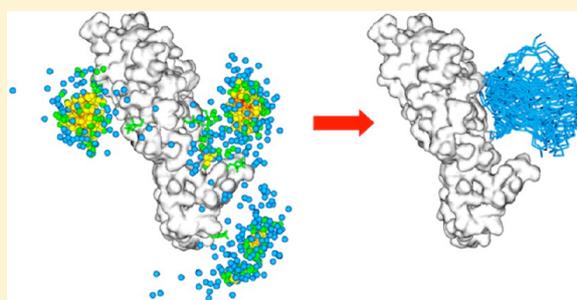
Loss of Electrostatic Interactions Causes Increase of Dynamics within the Plastocyanin–Cytochrome *f* Complex

 Sandra Scanu,[†] Johannes M. Foerster,^{†,‡} Monika Timmer,[†] G. Matthias Ullmann,[‡] and Marcellus Ubbink^{*,†}
[†]Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

[‡]Computational Biochemistry, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany

S Supporting Information

ABSTRACT: Recent studies on the electron transfer complex formed by cytochrome *f* and plastocyanin from *Nostoc* revealed that both hydrophobic and electrostatic interactions play a role in the process of complex formation. To study the balance between these two types of interactions in the encounter and the final state, the complex between plastocyanin from *Phormidium laminosum* and cytochrome *f* from *Nostoc* sp. PCC 7119 was investigated using NMR spectroscopy and Monte Carlo docking. Cytochrome *f* has a highly negative charge. *Phormidium* plastocyanin is similar to that from *Nostoc*, but the net charge of the protein is negative rather than positive. NMR titrations of Zn-substituted *Phormidium* plastocyanin and *Nostoc* cytochrome *f* indicated that a complex with an affinity intermediate between those of the *Nostoc* and *Phormidium* complexes is formed. Plastocyanin was found in a head-on orientation, as determined using pseudocontact shifts, similar to that in the *Phormidium* complex, in which the hydrophobic patch represents the main site of interaction on plastocyanin. However, the interaction in the cross-complex is dependent on electrostatics, similar to that in the *Nostoc* complex. The negative charge of plastocyanin decreases, but not abolishes, the attraction to cytochrome *f*, resulting in the formation of a more diffuse encounter complex than in the *Nostoc* case, as could be determined using paramagnetic relaxation spectroscopy. This work illustrates the subtle interplay of electrostatic and hydrophobic interactions in the formation of transient protein complexes. The results are discussed in the context of a model for association on the basis of hydrophobic contacts in the encounter state.



Protein association involves the formation of a dynamic encounter complex that is in equilibrium with the final, single-orientation complex.¹ In the encounter state, the proteins sample the surface of the partner, thus reducing the dimensionality of the search for the specific binding site.² Protein complex formation has been commonly described with a general model, in which the formation of the encounter complex is dominated by long-range electrostatic interactions, whereas the final state is determined by short-range interactions.³ However, theoretical studies demonstrated that desolvation can be a dominant interaction in the process of complex formation for systems with weak charge complementarity.^{4–6} Furthermore, partial desolvation of the binding interface was reported for some encounter complexes.^{7,8} The recent characterization of the encounter state of cytochrome *f* (^NCyt *f*) and plastocyanin (^NPc) complex from the cyanobacteria *Nostoc* sp. PCC 7119 (^{N–N} complex) demonstrated experimentally that electrostatic interactions alone cannot describe the encounter complex, suggesting that hydrophobic interactions also contribute to its formation.⁹ In the proposed model, long-range electrostatics result in the preorientation of ^NPc relative to ^NCyt *f*, and hydrophobic interactions stabilize

the encounter complex by promoting the overlap of the extended nonpolar surfaces of both proteins. ^NPc can diffuse in the hydrophobic interface and smoothly reach orientations capable of electron transfer (ET). The identification of hydrophobic interactions in the encounter state contrasts the view in which short-range interactions occur only in the final complex.²

Pc and *Cyt f* are photosynthetic redox partners in oxygenic organisms, such as plants, green algae, and cyanobacteria. *Pc* is a soluble electron carrier, which shuttles electrons from *Cyt f* of the cytochrome *b₆f* complex to photosystem I.^{10–12} The association of *Pc* and *Cyt f* is on the border between electrostatic-assisted¹³ and desolvation-mediated association,⁴ therefore representing a good model to elucidate the balance between electrostatic and hydrophobic interactions in protein complex formation. *In vitro*, electrostatic interactions enable fast association,^{14–19} and nonpolar interactions favor the stabilization of the complex in an ET active conformation.^{20–22} The

Received: April 9, 2013

Revised: July 8, 2013

Published: August 28, 2013

characterization of *Pc*-Cyt *f* complexes from several organisms revealed that small differences in the electrostatic surface properties of the proteins in the individual proteins strongly influence both the binding equilibrium and the final orientations of the complexes. Both in plants^{20,23} and in the cyanobacteria *Nostoc*²² and *Prochlorothrix hollandica*,²⁴ electrostatic interactions influence the final orientation of *Pc* with respect to Cyt *f* within the complex and tilt the long side of *Pc* toward the small domain of Cyt *f* in the so-called side-on orientation. The complex from the cyanobacterium *Phormidium laminosum* (^{Ph}*Pc* complex) was found instead in the head-on orientation, in which solely the hydrophobic ET site represents the binding site.²¹ Neutralization of charged residues on the surface of ^{Ph}*Pc*¹⁶ and ^N*Pc*¹⁸ has shown to have greater effect on the kinetics of the reaction than similar modifications on ^{Ph}Cyt *f*¹⁷ and ^NCyt *f*,¹⁹ respectively. ^{Ph}Cyt *f* and ^NCyt *f* are electrostatically similar, with an overall charge of -13 and -15,²⁵ respectively, and a rather even distribution of the negative charges over the surfaces. The two *Pc*'s show 63% amino acid sequence identity and very similar three-dimensional structures, but they vary considerably in their electrostatic properties. ^N*Pc* is overall positively charged with six lysines (K6, K11, K20, K24, K35, and K100) forming an extended charged patch, which juxtaposes the long side of ^NCyt *f* in the side-on orientation. In ^{Ph}*Pc*, K11 and K20 are substituted by serine and asparagine, respectively, and the positively charged patch is composed of four lysines (K6, K30, K35, and K100), yielding a protein with a net negative charge (-1 at pH 6). To evaluate the effects that these electrostatic differences between the two *Pc*'s can cause along the association pathway of Cyt *f* and *Pc*, the complex of ^NCyt *f* and ^{Ph}*Pc* (^{N-Ph} complex) was studied using NMR spectroscopy and computational approaches. The consequences for binding affinity, final complex orientation, and encounter complex are discussed in light of the recent findings for the ^{N-N} complex.⁹

EXPERIMENTAL SECTION

Protein Production and Purification. The plasmid pET11PC,²⁶ which contains the gene for wild-type ^{Ph}*Pc*, was transformed in *E. coli* BL21 pLysS. ¹⁵N enriched-Zn substituted *Pc* was produced as described before for ^N*Pc*,²⁷ with the difference that ampicilline (100 mg/L) and chloramphenicol (20 mg/L) were added to the growth media instead of kanamycin. The purification procedure was reported before.²¹ The concentration of the protein was determined by absorbance spectroscopy using $\epsilon_{280} = 5 \text{ mM}^{-1}\text{cm}^{-1}$. The yield of pure protein was 4 mg/L of culture.

The pEAF-WT plasmid, containing the gene of the soluble domain (residue 1–254) of *Nostoc* sp. PCC 7119 Cyt *f* was kindly provided by Professor Dr. Miguel A. De la Rosa (University of Seville). Cyt *f* mutants were obtained using pEAF-WT plasmid as template for mutagenesis as described before.^{9,27,28} Production and purification of the protein and spin label attachment were performed as previously reported.^{9,18,27}

NMR Experiments. All NMR samples contained 2-(*N*-morpholino) ethanesulfonic acid (MES, 20 mM, pH 6) and 6% D₂O for lock. The pH of the sample was adjusted with small aliquots of HCl (0.5 M) and NaOH (0.5 M). For the chemical shift perturbation (CSP) experiments, Cyt *f* was titrated into Zn-substituted ¹⁵N *Pc* (40 μ M). Spectra were recorded at multiple Cyt *f*/*Pc* molar ratios (0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 2.5, 5.0, 7.5, and 10). For measurements of the pseudocontact shifts

(PCSs), HSQC spectra of the free *Pc* and in the presence of ferric and ferrous Cyt *f* were acquired on the same sample. Ferric Cyt *f* was oxidized with K₃[Fe(CN)₆] and loaded on a PD10 column to remove the oxidant, concentrated, and then added to *Pc* (final concentration of 135 μ M) to a Cyt *f*/*Pc* molar ratio of 3:1. Ferric Cyt *f* was then reduced by adding 10 mol equiv of ascorbic acid directly into the sample. For the paramagnetic relaxation enhancement (PRE) experiments, the ferric state of Cyt *f* was preserved by the addition of K₃[Fe(CN)₆] (50 μ M). These samples contained 135 μ M Cyt *f* for the Q125C mutant and 300 μ M for the other mutants, labeled with either (1-acetoxy-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) methanethiosulfonate (MTS) or (1-oxyl-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) methanethiosulfonate (MTSL). Samples also contained Zn-substituted ¹⁵N *Pc*, 45 μ M in the sample with Q125C Cyt *f*, and 100 μ M for the other Cyt *f* mutants. All NMR spectra were recorded at 298 K on a Bruker Avance III 600 MHz spectrometer equipped with a TCI-Z-GRAD CryoProbe. The ¹H-¹⁵N HSQC spectra were acquired with 1024 and 80 complex points in the direct and indirect dimensions, respectively.

NMR Data Analysis. The NMR spectra were processed with NmrPipe²⁹ and analyzed with CcpNMR Analysis.³⁰ CSP analysis was carried out as described before.²⁷ PCS was defined as the chemical shift difference for a resonance in the presence of paramagnetic and diamagnetic Cyt *f*, according to previously reported procedures.^{20,22}

The PREs were determined according to the procedure described by Battiste and Wagner.³¹ The intensity ratio I_p/I_d of the *Pc* resonances in the presence of MTSL-Cyt *f* (I_p) and MTS-Cyt *f* (I_d) were normalized by dividing them by the average value of the 10 largest I_p/I_d values (1.09 for Q7C, 1.05 for Q38C, 2.21 for N71C, 1.41 for Q125C, 1.16 for S181C, and 1.25 for S192). The PRE (Γ_2) values were calculated according to eq 1:

$$\frac{I_p}{I_d} = \frac{R_{2d} \exp(-\Gamma_2 t)}{R_{2d} + \Gamma_2} \quad (1)$$

R_{2d} represents the transverse relaxation rate in the diamagnetic sample, which was calculated from the line width at half height obtained from a Lorentzian peak fit in the direct dimension by using FuDA (this software was kindly provided by Dr. D. Fleming Hansen, University College London). The symbol t indicates the time for transverse relaxation during the pulse sequence (9 ms). The Γ_2 values were extrapolated to the 100% bound state using the experimentally obtained K_p . The uncertainty for I_p/I_d ratios ($\Delta\sigma_{p/d}$) was determined by error propagation according to eq 2 in which σ_p and σ_d represent the noise level of paramagnetic and diamagnetic spectra, respectively.:

$$\Delta\sigma_{p/d} = \frac{I_p}{I_d} \sqrt{\left(\frac{\sigma_p}{I_p}\right)^2 + \left(\frac{\sigma_d}{I_d}\right)^2} \quad (2)$$

The noise level of each spectrum is represented by the standard deviation of the intensities measured at 10 randomly chosen positions on the baseline.

Docking Calculations. The structure of the soluble part of Cyt *f* (residues 1–254) used for the calculation was taken from PDB entry 2ZT9³² as described before.²⁷ The structure of ^{Ph}*Pc* was taken from PDB entry 2Q5B. The orientation of ^{Ph}*Pc* in complex with ^NCyt *f* was determined by rigid body docking

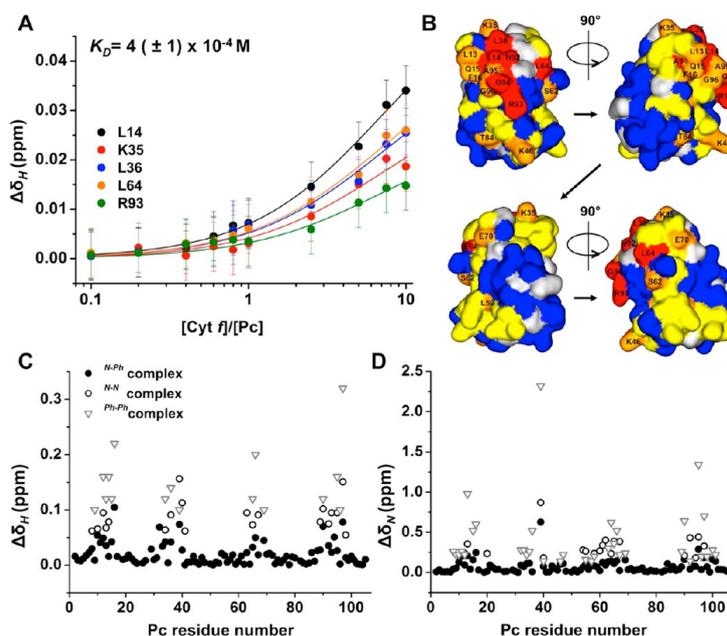


Figure 1. Interaction of Zn-substituted ^{15}N $^{\text{Ph}}\text{Pc}$ with $^{\text{N}}\text{Cyt } f$. (A) CSP curves for Zn $^{\text{Ph}}\text{Pc}$ binding to $^{\text{N}}\text{Cyt } f$ for selected residues fitted to a 1:1 interaction model. (B) Binding map of $^{\text{Ph}}\text{Pc}$ in the presence of wild-type $^{\text{N}}\text{Cyt } f$ (Fe III), color-coded on a surface model of Pc (PDB entry 2Q5B). The red color corresponds to $\Delta\delta_{\text{AVG}} \geq 0.030$ ppm, orange to $\Delta\delta_{\text{AVG}} \geq 0.015$ ppm, yellow to $\Delta\delta_{\text{AVG}} \geq 0.0075$ ppm, and blue to $\Delta\delta_{\text{AVG}} \leq 0.0075$ ppm. Prolines and overlapping residues are colored in light gray. This image and others of molecular structures were made with Discovery Studio Visualizer 2.5 (Accelrys). (C,D) Chemical shift perturbations of $^{\text{Ph}}\text{Pc}$ resonances upon the binding of $^{\text{N}}\text{Cyt } f$ (Fe II). The CSPs in ^1H dimension (C) and in ^{15}N dimension (D) observed for the $^{\text{N-Ph}}$ complex are shown as black dots, for the $^{\text{N-N}}$ complex as black circles, and for $^{\text{Ph-Ph}}$ complex as gray triangles.

using solely PCS restraints with the option PARAAstraints³³ in Xplor-NIH 2.9.9.³⁴ For this reason, the observed $^1\text{H}\Delta\delta_{\text{PCS}}$ were extrapolated to 100% bound Pc by dividing them by the fraction bound (0.47). The size of the axial magnetic component of the magnetic susceptibility anisotropy tensor ($\Delta\chi_{\text{ax}}$) was derived from the g-tensor values measured by EPR spectroscopy on plant $\text{Cyt } f$ ($7 \times 10^{-32} \text{ m}^3$).²² However, the value required to obtain convergence of the structure calculations is much smaller. One reason for this is the temperature difference between the EPR measurements (10 K) and NMR spectra (taken at 298 K). The second reason is an averaging effect occurring in the encounter state, which reduces the PCS considerably. The $\Delta\chi_{\text{ax}}$ was varied in the range $(0.61-3.3) \times 10^{-32} \text{ m}^3$. The best convergence was found for $\Delta\chi_{\text{ax}} = 0.87 \times 10^{-32} \text{ m}^3$, whereas it was taken to be $3.8 \times 10^{-32} \text{ m}^3$ in the $^{\text{N-N}}$ complex.²² The intermolecular PCSs from the ferric heme iron of $\text{Cyt } f$ to the backbone amide atoms in Pc were back-calculated from the best 20 structures and compared with the experimental PCSs. Equation 3 was used for the PCS calculation, assuming an axial magnetic susceptibility tensor oriented along the vector defined by the iron and the N-atom of Y1 of $\text{Cyt } f$.²⁰

$$\Delta\delta_{\text{PCS}} = \frac{\Delta\chi_{\text{ax}}}{12\pi r^3} (3\cos^2\theta - 1) \quad (3)$$

In eq 3, $\Delta\delta_{\text{PCS}}$ is the PCS, r is the distance between the heme iron and the observed Pc nucleus, and θ is the angle among the Pc nucleus, heme iron, and the nitrogen of the amine group of Y1 in $\text{Cyt } f$. The degree of agreement between observed (PCS^{obs}) and back-calculated (PCS^{calc}) PCSs was determined by the PCS Q factor, defined as follows:

$$Q_{\text{PCS}} = \sqrt{\frac{\sum (\text{PCS}^{\text{obs}} - \text{PCS}^{\text{calc}})^2}{\sum (|\text{PCS}^{\text{obs}}| + |\text{PCS}^{\text{calc}}|)^2}} \quad (4)$$

Ensemble docking was performed as described for the $^{\text{N-N}}$ complex with seven Pc copies per docking.⁹ The restraints for the calculations were obtained according to equation 5:

$$\Gamma_2^{\text{obs}} = f_1 \Gamma_2^{\text{ens}} + f_2 \Gamma_2^{\text{final}} \quad (\text{Sa})$$

$$f_1 + f_2 = 1 \quad (\text{Sb})$$

The ensemble Γ_2 (Γ_2^{ens}) was calculated as the difference between observed Γ_2 (Γ_2^{obs}) and average back-calculated Γ_2 from the models of the PCS-based final complex (Γ_2^{final}). The calculations were carried out with f_2 values = 0, 0.15, 0.25, 0.35, 0.5, 0.65, 0.75, 0.85, 0.95, and 1. The restraints were grouped into three classes as described before.²⁷ For the visualization of the encounter complex ensemble, 150 docking runs were performed, yielding 148 ensembles of 7 Pc conformers, with a difference in the total restraint energy $\leq 20\%$.

The ensembles from separated dockings were evaluated by means of the averaged violation for all experimental restraints as described before.⁹

Monte Carlo Simulations of the Encounter Complex.

The structure files for $\text{Cyt } f$ and Pc were taken from the PDB entries 2ZT9³² and 2Q5B, respectively. Monte Carlo (MC) simulations generate a Boltzmann distribution of encounter complexes according to their electrostatic interaction energy.³⁵ The simulations were performed using a previously described approach.^{9,36}

RESULTS AND DISCUSSION

Affinity and Binding Site. For the characterization of the nonphysiological cyanobacterial N^{Ph} complex formed by N Cyt *f* and ^{Ph}Pc , ^{15}N -enriched Zn- ^{Ph}Pc was titrated to either oxidized N Cyt *f* (Fe^{III}) or reduced N Cyt *f* (Fe^{II}) to molar ratios $Pc/Cyt f$ 1:10 and 1:3, respectively. ^{Ph}Pc was produced with Zn rather than the Cu in the binding site to avoid the effects of ET and the disappearance of important resonances due to the line-broadening caused by the paramagnetic Cu.³⁷ Each titration point was monitored through the acquisition of ^{15}N - 1H HSQC spectra. Upon addition of Cyt *f*, a number of resonances shifted in the spectrum, indicating complex formation. The appearance of shifting resonances indicates that free and bound Pc are in fast exchange on the NMR time scale. The binding curves for the most affected residues were obtained by plotting the chemical shift perturbation (CSP, $\Delta\delta_{H^1}$) versus Cyt *f*(Fe^{III})/ Pc molar ratio, as shown in Figure 1A.

The CSP curves did not reach saturation, indicating a low affinity. The global fit of the binding curves to a 1:1 binding model yielded a dissociation constant of $4 (\pm 1) \times 10^{-4}$ M. This value is in-between the reported values for the $^{Ph-Ph}$ complex and $N-N$ complex, being $\approx 10 \times 10^{-4}$ M²¹ and 0.8×10^{-4} M,²⁷ respectively. Whereas the cross-complex formed by $^{Ph}Cyt f$ and $N^{Ph}Pc$ ($^{Ph-N}$ complex) was reported to have similar affinity to the $N-N$ complex ($K_D = 0.8 \times 10^{-4}$ M),²⁵ the N^{Ph} complex shows an affinity intermediate to that of the two physiological complexes but closer to that of the $^{Ph-Ph}$ complex. The experimental K_D was used to determine whether the fraction of ^{Ph}Pc bound to N Cyt *f* at the last point of the titration was 0.52, and the average amide CSPs ($\Delta\delta_{AVG}$) were extrapolated to the 100% bound form. The CSP map of Zn- ^{Ph}Pc was obtained by color coding each residue according to the size of $\Delta\delta_{AVG}$ (Figure 1B). The largest effects were observed for residues surrounding the metal binding site, namely, A9, L14, L36, H39, L64, H92, R93, and G94, colored in red. Most of these residues are hydrophobic and make up the hydrophobic patch of Pc , which was also identified as the main binding site in the structural models of the $N-N$ complex and $^{Ph-Ph}$ complex.^{21,22} Clearly, the hydrophobic patch plays a fundamental role in the formation of the Cyt *f*- Pc complexes. The CSP map is qualitatively similar to that of the $N-N$ complex^{22,27} with a prominent perturbation for R93, known to be involved in the binding in both the $N-N$ complex^{18,19,22} and $^{Ph-Ph}$ complex.¹⁶ Interestingly, a significant CSP was also observed for K46 in the N^{Ph} complex. K46 is located far from the hydrophobic patch, well below R93, and kinetic studies suggested its involvement in the electrostatic modulation of the binding of $^{Ph-Ph}$ complex.¹⁶

Structure of the Final Complex. The orientation of the ^{Ph}Pc in complex with N Cyt *f* was determined by taking advantage of the intermolecular PCSs caused by the paramagnetic oxidized iron of Cyt *f* on Pc backbone amide protons, in a way similar to that done previously for other Pc -Cyt *f* complexes.^{20,21,23,24} PCSs arise from the through-space interaction between the spin of the unpaired electron and that of the observed nucleus. PCS is distance and orientation dependent and provides restraints for structural calculations. The calculations converged to an ensemble of structures. The best 20 structures exhibit a difference in the restraint energy of less than 6% and are shown in Figure 2A.

The resulting model shows a high degree of variability, but in all structures, the hydrophobic patch of Pc is making contact with the hydrophobic patch surrounding the heme of Cyt *f* and

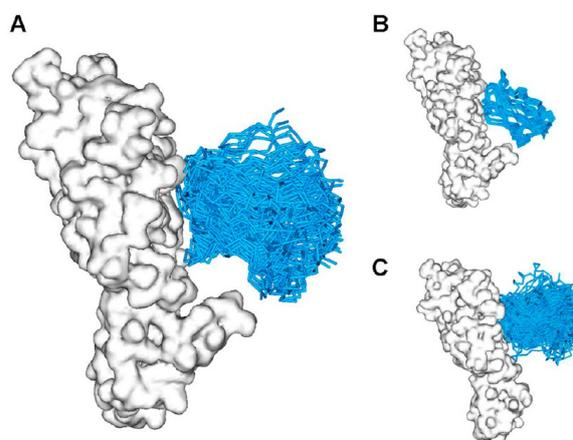


Figure 2. Comparison of the structures of Pc -Cyt *f* complexes, showing the structure obtained for the N^{Ph} complex (A), the physiological $N-N$ complex (PDB entry 1TU2²²) (B), and $^{Ph-Ph}$ complex (C).²¹ N Cyt *f* and $^{Ph}Cyt f$ are shown as white surface models of PDB entries 2ZT9 and 1CI3,³⁸ respectively, and Pc is represented by the ensemble of the 20 (A), 10 (B), and 25 (C) lowest energy conformations, shown as cyan C_α traces.

represents the entire complex interface. All structures showed an interaction between H92 of Pc and F3 of Cyt *f*, also found in the $^{Ph-Ph}$ complex²¹ and $N-N$ complex.²² The binding interface is composed of polar and hydrophobic residues, located in the regions 11–14, 36–39, 64–68, and 90–95 on ^{Ph}Pc . R93 represents the only charged interfacial residue. The averaged Cu–Fe distance in the ensemble was 15.3 (± 0.5) Å. In Figure 3A, the observed (black dots) and the back-calculated PCSs for the best 20 structures (gray lines) are plotted versus Pc residue numbers.

For most residues, experimental and back-calculated PCSs agree within the error margins. Small deviations are observed for F16, V29, W31, V32, and A90, which form the edge of the hydrophobic binding site, and M97, which coordinates the metal. Considering the relative vicinity of these residues to the heme, it is possible that the approximations made for the size, axiality, and orientation of the magnetic susceptibility tensor cause these deviations. The overall quality of the structures was evaluated by calculating a quality (Q) factor for the back-calculated PCSs for each structure of the final model and the experimental PCSs (see Experimental Section, eq 4). The average Q value was calculated to be 0.23 (± 0.01).

The orientation of Pc in the complex is more similar to the head-on orientation found in the $^{Ph-Ph}$ complex (Figure 2C) than to the side-on orientation of the $N-N$ complex (Figure 2B). ^{Ph}Pc is oriented perpendicular to the heme with a slight tilt toward the small domain of N Cyt *f*. In the $N-N$ complex, the specific electrostatic contacts between K57 and K62 of $N^{Ph}Pc$ and E189 and D64 of N Cyt *f* appear to be responsible for the long side of Pc to be tilted toward Cyt *f*.^{22,25} In ^{Ph}Pc , these lysines are substituted with D57 and S62, respectively, and the loss of these important electrostatic contacts may lead to the head-on orientation in the N^{Ph} complex. In the N^{Ph} complex, only the bottom part of ^{Ph}Pc (relative to the hydrophobic patch) is turned toward the small domain of N Cyt *f*, probably as a consequence of the charge–charge interaction between K46 of ^{Ph}Pc and E189 and D190 of N Cyt *f*. The soluble part of $^{Ph}Cyt f$ is shorter than N Cyt *f*, comprising 249 instead of 254 residues.

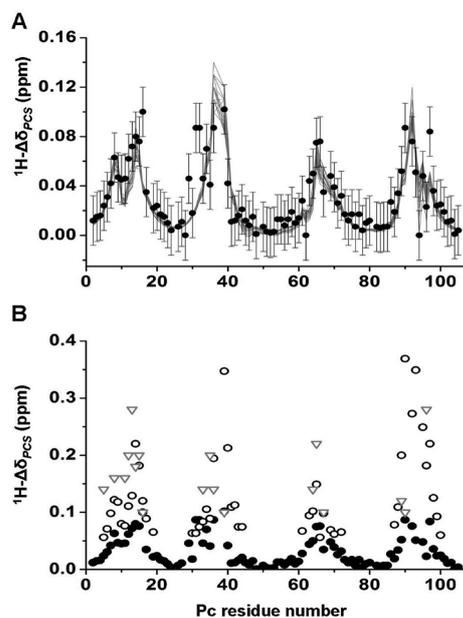


Figure 3. Evaluation of $N\text{-}^{Ph}$ complex. In panel A, the observed ${}^1\text{H}\Delta\delta_{\text{PCS}}$, which were extrapolated to 100% bound Pc , are shown as black dots, and the back-calculated ${}^1\text{H}\Delta\delta_{\text{PCS}}$ for the 20 lowest PCS energy structures are shown as gray lines. The error bars represent the estimated experimental errors in the resonance positions. In panel B, the observed ${}^1\text{H}\Delta\delta_{\text{PCS}}$ for the $N\text{-}^{Ph}$ complex are shown as black dots, for the $N\text{-}N$ complex²² as black circles, and for the $Ph\text{-}Ph$ complex²¹ as gray triangles. All PCSs were extrapolated to the 100% bound form and plotted versus Pc residue numbers.

This causes the small domain to be less extended and not in direct contact with Pc in the $Ph\text{-}Ph$ complex.²¹ In the $N\text{-}^{Ph}$ complex, K46 is in a favorable position to have electrostatic interactions with E189 and D190 in the prominent small domain of *Cyt f*.

Since PCSs depend on the orientation of the observed nucleus with respect to the paramagnetic iron, the presence of multiple orientations is expected to influence the size of PCSs. In the $Pc\text{-}Cyt f$ complex from *Prochlorothrix hollandica*, the mutation of Y12 and P14 in Pc to Gly and Leu, respectively, caused an increase of dynamics, as judged by the decrease of PCSs for nuclei in certain regions of Pc .²⁴ In Figure 3B, the observed ${}^1\text{H}\Delta\delta_{\text{PCS}}$ of the $N\text{-}^{Ph}$ complex were compared with the reported values for the $N\text{-}N$ complex²² and $Ph\text{-}Ph$ complex,²¹ each extrapolated to the 100% bound state. The pattern of the ${}^1\text{H}\Delta\delta_{\text{PCS}}$ is similar for all complexes, but the sizes of ${}^1\text{H}\Delta\delta_{\text{PCS}}$ are comparable only for the two physiological complexes, whereas they are considerably lower for the $N\text{-}^{Ph}$ complex. This indicates that in the $N\text{-}^{Ph}$ complex the dynamics of Pc is larger than that in both the $N\text{-}N$ complex and the $Ph\text{-}Ph$ complex.

Encounter Complex. To map the distribution of the encounter intermediates on the $N\text{-}^{Ph}$ *Cyt f* surface in the $N\text{-}^{Ph}$ complex, six spin labels were attached on $N\text{-}^{Ph}$ *Cyt f*, one at a time, and PREs were measured on the amide backbone protons of ^{15}N -labeled Pc . *Cyt f* was added to Pc in a molar ratio $Pc/Cyt f$ of 1:3. PRE causes line broadening of Pc resonances resulting in a low ratio of peak intensities in the spectra of the paramagnetic and diamagnetic samples (I_p/I_d). In Figure 4 (central panel), the positions of spin labels are shown on a surface model of *Cyt f* with respect to Pc oriented as found in the lowest energy structure of the PCS-based final complex (cyan C_α trace). Spin

labels attached to *Cyt f* on the same side as the binding site for Pc , at positions Q7, N71, and S192, caused a large decrease of I_p/I_d ratios of Pc resonances.

It is noteworthy also that spin labels attached on the backside of *Cyt f*, at the positions Q38 and Q125, or located far away, S181, showed a moderate to large decrease for some resonances. The large error bars calculated for the ratios in the presence of the Q125C mutant are due to the lower concentration of Pc in this sample (45 μM) as compared to that in the other mutants (100 μM) resulting in a low signal-to-noise ratio. The I_p/I_d ratios were used to determine the PRE (Γ_2). In the fast exchange regime (see above), the observed PREs are weighted averages of free Pc , the encounter complex, and the final complex. The PREs were extrapolated to the 100% bound state (encounter complex + final complex) by dividing by the fraction of bound Pc . The PREs caused by each spin label were mapped on the surface of Pc (Figure 5).

Even though the three spin labels located at the same side of *Cyt f* as the binding site (Q7C, N71C, and S192C) are relatively far from each other, the PRE patterns are very similar and resemble the CSP map in the presence of wild-type *Cyt f* (Figure 1B). The qualitative similarity of the PRE patterns suggests that Pc samples a large area of the *Cyt f* surface, while maintaining the same relative orientation to *Cyt f*. The highest PREs were observed for residues located in the hydrophobic patch of Pc , indicated as main binding site in the PCS-based final complex. Most of these residues are hydrophobic or polar, with the exception of R93 that was strongly affected by PRE in the presence of spin labels in N71 and S192. The same residue exhibited a high CSP in the presence of wild-type *Cyt f* (Figure 1B). Interestingly, for most of these residues moderate PREs were also observed in the presence of spin labels attached to the backside of *Cyt f* with respect to the PCS-based binding site of Pc , indicating that Pc also visits this part of *Cyt f*.

The encounter complex was visualized by ensemble docking. This approach is based on the fact that PREs result from the weighted average contribution of all species in solution, both the encounter and the final complexes.³⁹ To represent all species that contribute to the observed PREs, multiple conformers of a protein are simultaneously docked on the other protein to obtain a population distribution that fits the experimental data. Each docking yields a unique ensemble of orientations that account for the experimental PREs. To separate the PRE contribution of the complex in the final state, the averaged back-calculated PREs from the PCS-based models of the final complex were subtracted from experimental PREs, and the resulting PREs were used. A series of ensemble docking calculations were then carried out by varying the population of the final state (f_2) from 1 to 0. The resulting ensembles were evaluated by calculating the average distance violation over all experimental distances. The average distance violations were plotted versus the percentage of the encounter complex (Figure S1, Supporting Information). The violations show that the observed PREs are not explained by the PCS-based structure alone. A significant decrease in the average violation is already observed when the encounter complex is taken to be 5%. Further increase of the encounter complex fraction in the restraints did not improve the fit of the data. For all generated ensembles, an average violation of about 2 Å was observed.

The calculations for the representation of the encounter complex were performed assuming a pure encounter state ($f_1 = 1$). The comparison of the back-calculated distances between the oxygen atom of the spin labels and the amide protons of all

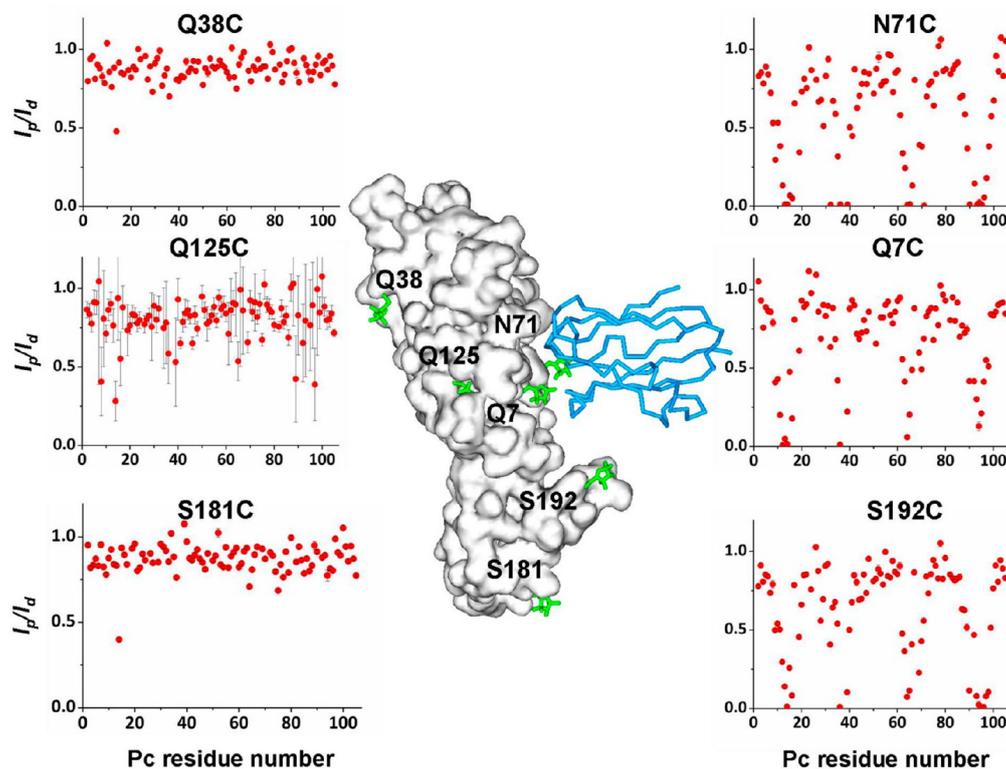


Figure 4. PRE in the N^{Ph} complex. Central panel: location of the spin labels (green sticks) modeled on the $N^{Cyt f}$ (PDB entry 2ZT9). *Cyt f* is shown as a white surface, and *Pc* is represented as a cyan C_{α} trace, oriented as the PCS-based final complex. Side panels: the I_p/I_d ratios are plotted versus the *Pc* residue number for each of the spin label positions on *Cyt f*. The error bars represent the uncertainty for I_p/I_d ratios based on the noise levels of the spectra. For most points, the error bar is within the symbol.

Pc conformers of the generated encounter complex (red line in Figure 6) and the back-calculated distances in the PCS-based models of the final complex (blue line) shows that only the generated encounter complex fits the experimental distances (green dots and line).

The main deviation is represented by S192C, indicating that ^{Ph}Pc spends more time close to this spin label than expected from the PCS-based models. This suggests that PREs from S192C mainly arise from the encounter complex. Most of calculated distances from the generated encounter complex lie within the error margins of the experimental values. Deviations were observed for the spin label Q125C, likely due to the poor data quality (see above).

To represent the encounter complex, an ensemble from 145 docking solutions, with a total of 1015 *Pc* conformers was generated (Figure 7A).

The centers of mass (CoMs) of *Pc* were colored according to the density of distribution, with red and blue representing the largest and smallest densities, respectively. It should be noted that the incomplete coverage of spin labels on the *Cyt f* surface implies that also other *Cyt f* surface areas could be involved in the encounter complex. The current analysis shows that the encounter complex is at least distributed over three extensive areas of *Cyt f* surface. The most extended area is located in the vicinity of the binding site found in the final complex models, the second is in front of the small domain of *Cyt f*, and the third on the backside relative to the final complex. The third area is an artifact due to the use of the soluble part of $N^{Cyt f}$. *In vivo*,

Cyt f is embedded on the thylakoid membrane that will prevent *Pc* from binding on this side.⁹ In all three areas, the interface comprises large patches of polar and hydrophobic residues. Despite the fact that in this study a less extensive portion of the *Cyt f* surface was monitored, the encounter complex resembles the one found for the $N-N$ complex (Figure 7B). The encounter ensemble of N^{Ph} complex is more extensive and covers a larger area of the hydrophobic regions of *Cyt f*. In the $N-N$ complex, stronger charge interactions may lead to more defined encounter regions. In the $N-N$ complex, one continuous diffusive encounter region is present on the side of the binding site, while in the N^{Ph} complex, two distinct diffusive encounter areas can be seen. To evaluate the distribution of the ET active complexes, the CoMs of ^{Ph}Pc are colored according to the calculated distance between Cu in *Pc* and Fe in *Cyt f*, with red and blue representing the smallest and largest distances, respectively (Figure S2A, Supporting Information). As for the $N-N$ complex (Figure S2B, Supporting Information), the encounter complex orientations compatible with rapid ET (red dots, Cu–Fe distance ≤ 16 Å) are located only in front of the heme, in close vicinity of the binding site found in the final complex.

Role of Electrostatic Interactions in Complex Formation. The effect of ionic strength (I) on the binding shifts of ^{Ph}Pc in the presence of reduced $N^{Cyt f}$ at a *Cyt f*/*Pc* molar ratio of 3:1 was investigated at NaCl concentrations of 100 mM ($I = 110$ mM) and 200 mM ($I = 210$ mM). The CSPs ($\Delta\delta_H$) were defined relative to the control measurements recorded on free

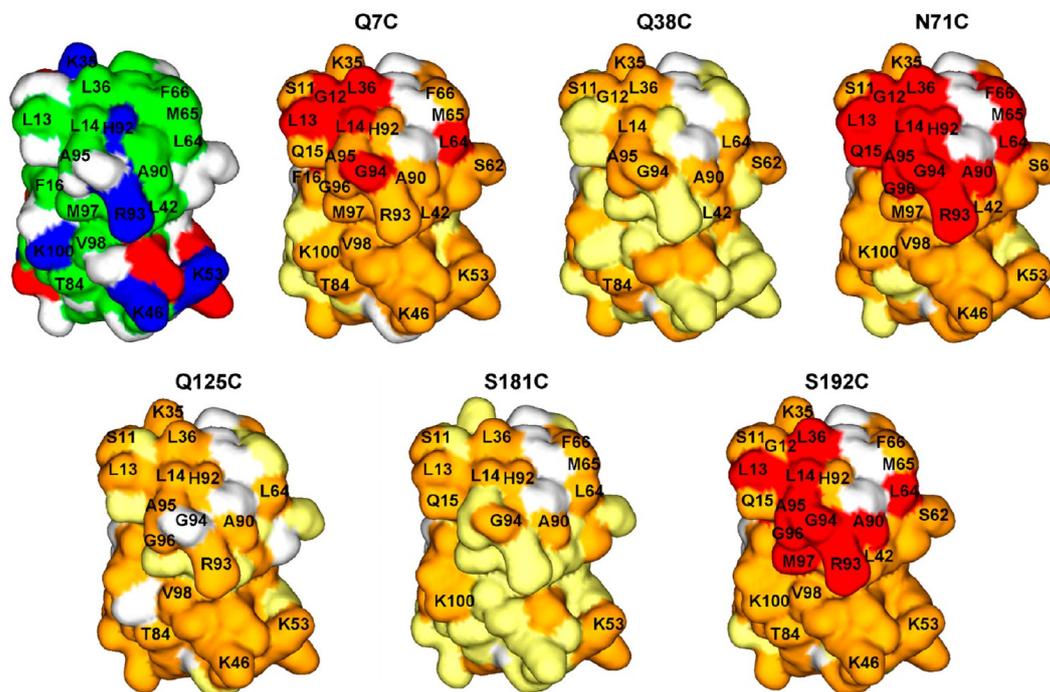


Figure 5. PRE maps of ^{15}N enriched Zn-substituted ^{ph}Pc in the presence of MTSL-conjugated $^N\text{Cyt } f$, color-coded on a surface model of Pc (PDB-entry 2Q5B). Experimental PREs were extrapolated to 100% bound Pc . Residues with $\Gamma_2 \geq 200 \text{ s}^{-1}$ are colored in red, with $10 \text{ s}^{-1} < \Gamma_2 < 200 \text{ s}^{-1}$ in orange and with $\Gamma_2 \leq 10 \text{ s}^{-1}$ in light yellow. Prolines and residues with overlapping resonances are colored in white. On the top left, Pc is colored according to its charge distribution. Negatively and positively charged residues are shown in red and blue, respectively. Hydrophobic residues are in green and polar residues in white.

^{ph}Pc at the same NaCl concentration. The $\Delta\delta_{\text{H}}$ values at the different salt concentrations were plotted versus ^{ph}Pc residue numbers (Figure 8A).

The addition of 200 mM NaCl did not affect the perturbation pattern observed in the absence of salt. The residues showing the strongest ionic strength dependence, namely, V15, F16, V32, L36, H39, A90, and M97, are located in the loops that make up the hydrophobic patch. Similarly to what was observed for the $^{N-N}$ complex,²⁵ all major shifts ($\Delta\delta_{\text{H}} \geq 0.03 \text{ ppm}$) decreased by about 60% upon the addition of 200 mM NaCl. These findings suggest that at low ionic strength favorable electrostatic interactions play a role also in $^{N-ph}$ complex formation but do not influence the relative orientation of the proteins within the final complex. Apparently, attractive interactions exist despite the fact that $^N\text{Cyt } f$ and ^{ph}Pc both have an overall negative charge at pH 6. This observation suggests that charge distribution plays a critical role in the association process at low ionic strength. For this reason, MC simulations were performed.

In rigid-body MC simulations, the association of two proteins is simulated on the basis of their electrostatic potentials.³⁵ On the assumption that the formation of the encounter complex is purely driven by long-range electrostatic forces,² PRE and MC simulations were successfully combined for the visualization of the encounter complex of cytochrome *c* and cytochrome *c* peroxidase, demonstrating that the formation of this complex could be explained by electrostatic interactions alone.³⁶ The same approach on the $^{N-N}$ complex was revealed to be inadequate to describe the encounter complex, which appears to be stabilized by electrostatic as well as hydrophobic

interactions.⁹ At the same time, MC simulations provided evidence of the electrostatic preorientation of Pc toward $\text{Cyt } f$, as was found on the basis of CSP and PRE data. MC simulations were performed for the $^{N-ph}$ complex to establish whether electrostatic preorientation of ^{ph}Pc can occur despite the negative charge of both proteins. The calculations produced an ensemble consisting of the Boltzmann distribution of orientations of $\text{Cyt } f$ around Pc . An ensemble of 5000 structures was randomly selected from the entire set of two million solutions, and the positions of $\text{Cyt } f$ CoMs were plotted in Figure 8C. The position in the plot is determined by two angles. The first is the cone angle (α) formed by the $\text{Cyt } f$ CoM, the Pc CoM, and the Ne atom of copper ligand H92, taken as the center of the hydrophobic patch. The larger this angle is, the further the $\text{Cyt } f$ CoM is rotated away from the hydrophobic patch. The α angle is represented by the circles in Figure 8C and D. The second angle, β , indicates the position on the cone and represents the side of Pc to which the $\text{Cyt } f$ CoM is rotated. The hydrophobic patch is delineated by a red line marked with residue numbers. Figure 8C shows that $\text{Cyt } f$ binds in a diffusive manner but more toward the hydrophobic patch side of ^{ph}Pc than toward the other end. Clearly, preorientation occurs due to electrostatic interactions. This finding is also illustrated in Figure 8B. The cumulative fraction of $\text{Cyt } f$ CoMs for the α angle is plotted (black bars). The red line represents the cumulative fraction for a completely random distribution around a sphere. The fraction of CoMs with α angles of less than 90° is larger than 50%; therefore, more than half of the CoMs is present around the half of Pc that comprises the hydrophobic patch due to electrostatic preorientation. This

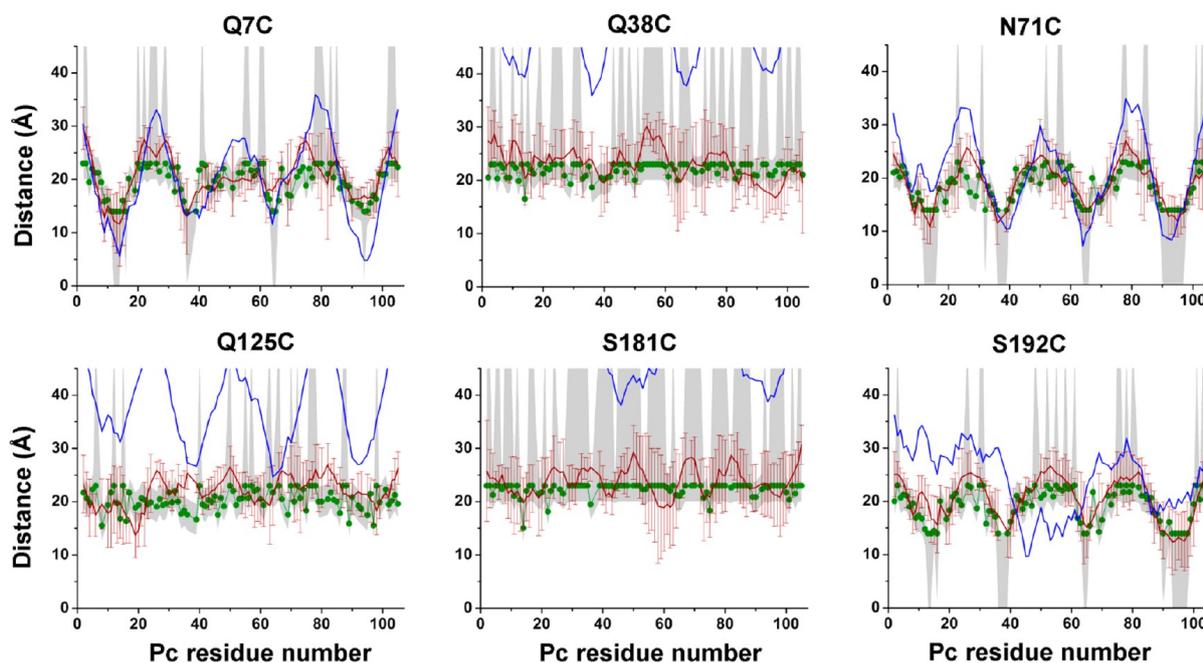


Figure 6. Ensemble docking. Experimental and back-calculated average distances between *Pc* amide protons and oxygen atoms of MTSL conjugated to *Cyt f* are plotted against the *Pc* residue number. The green circles and lines represent the experimental distances, and the gray areas indicate the error margins. The average distances back-calculated from the 20 lowest-energy solutions of the PRE driven ensemble docking are shown as a red line with error bars representing the SD. The average back-calculated distances from the PCS-based final complex models are shown as a blue line. Calculations were performed with $f_1 = 1$.

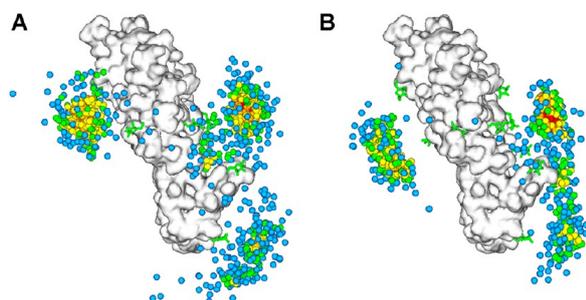


Figure 7. Comparison of the encounter complexes of N^{Ph} complex (A) and N^N complex (B). $N^{Cyt f}$ is shown as a white surface and spin labels as green sticks. *Pc* CoMs are represented by spheres, color-coded to indicate the density of the distributions, decreasing from red to blue. Densities were determined by counting the number of neighbors within 2.5 Å.

suggests that despite the net negative charge of ^{Ph}Pc , the localization of positive charges promotes the formation of an oriented complex. For comparison, the same calculations, on the basis of an earlier study,⁹ are shown for the N^N complex in Figures 8B (gray bars) and 8D. For this complex, the preorientation is stronger and shows a more defined binding spot for $\alpha = 60^\circ$ – 80° and $\beta = 30^\circ$ – 120° . The primary reason for this difference between the complexes of $N^{Cyt f}$ with NPc and ^{Ph}Pc is the presence of two Lys residues (K11 and K20) in this region of NPc , which are substituted by serine and asparagine, respectively, in ^{Ph}Pc .

To compare the importance of ionic strength on the formation of the different *Cyt f*-*Pc* complexes, MC simulations

were performed for the N^N complex, the N^{Ph} complex, and the $^{Ph-Ph}$ complex at ionic strength values of 10 mM, 110 mM, and 210 mM (Figure 8E–G). In the cases of the N^{Ph} complex (Figure 8F) and $^{Ph-Ph}$ complex (Figure 8G), very sparsely distributed encounter complexes were observed at higher values of I . For the N^N complex (Figure 8E), though the increase in ionic strength resulted in the production of more diffusive encounters, in which the *Cyt f* distribution covers a wider area of the *Pc* surface than observed at low ionic strength, a preferable docking area could be still recognized and related to the electrostatic properties of NPc . The histograms of the electrostatic interaction energies show that at an ionic strength of 210 mM (green bars), the N^{Ph} complex (Figure 8F) and $^{Ph-Ph}$ complex (Figure 8G) have lost all electrostatic attraction. For the N^N complex (Figure 8E), it is strongly reduced but not completely zero.

Comparison among *Cyt f*-*Pc* Complexes. Recently, we proposed a model for the formation of the N^N complex on the basis of the available kinetic and NMR data. Upon approach of the proteins, NPc is rotated by electrostatic interactions to face $N^{Cyt f}$ with its hydrophobic patch leading to the formation of the encounter complex. This state is stabilized not only by charge interactions but also hydrophobic interactions, allowing a smooth transition from encounter to ET-capable orientations by gradual increase of the hydrophobic overlap and sliding over the hydrophobic interface. It is interesting to interpret the data for the N^{Ph} complex in light of this model.

The most important difference between NPc and ^{Ph}Pc is the net positive and negative charges, respectively. Given the highly negative charge on $N^{Cyt f}$, a poor interaction with ^{Ph}Pc is expected if charge interactions are dominant. It was found that the affinity is 5-fold lower for ^{Ph}Pc , suggesting that charges

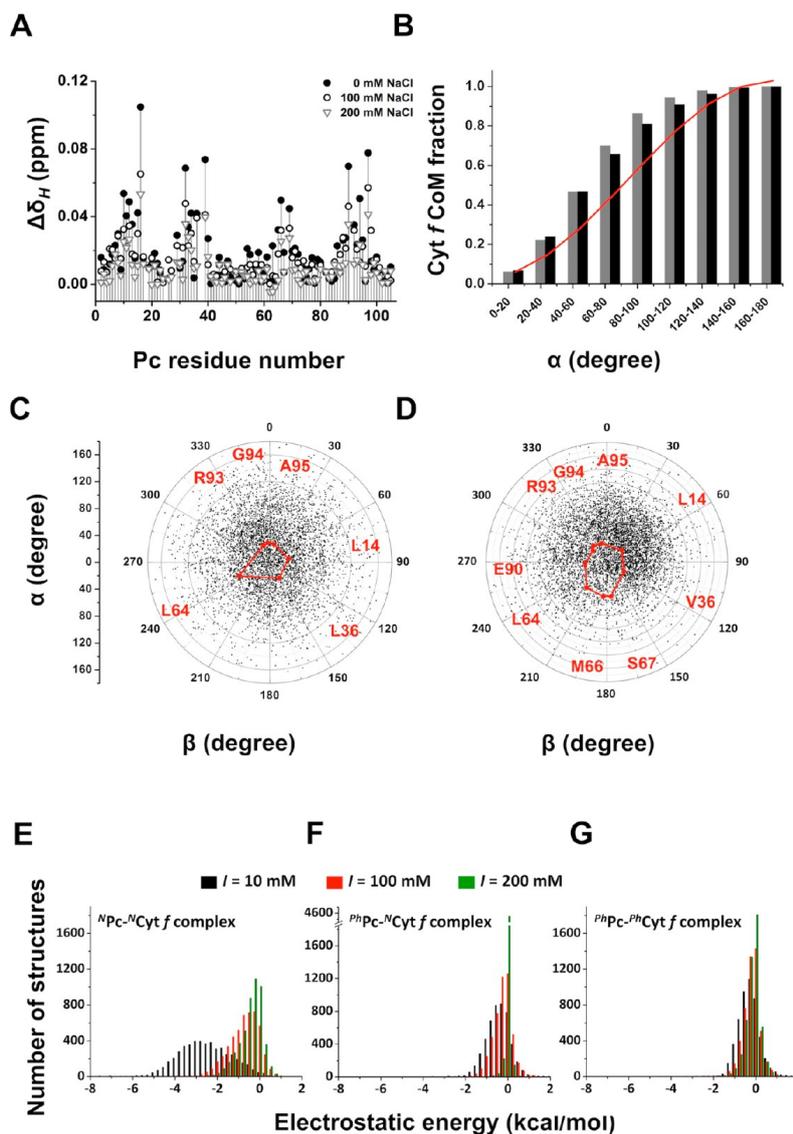


Figure 8. Role of electrostatic interactions in Cyt *f*-Pc complexes. (A) Ionic strength dependence of $\Delta\delta_H$ for ^{Ph}Pc backbone amide protons in the presence of reduced $^{N}Cyt f$ at 0 mM NaCl (black dots), 100 mM NaCl (black circles), and 200 mM NaCl (gray triangles). (B,C,D) Analysis of the encounter complex generated by MC simulations. (B) The cumulative fraction of Cyt *f* CoMs for the α angle is plotted for the $^{N-N}$ complex (black bars) and $^{N-Ph}$ complex (gray bars). The red line represents the cumulative fraction for a completely random distribution around a sphere. Plots of the position of the $^{N}Cyt f$ CoMs with respect to ^{Ph}Pc (C) and ^{N}Pc (D) in the MC ensembles. The red line connects the positions of hydrophobic patch residues. The $N\epsilon$ of H92 is at the center of the plots in panels C and D. (E,F,G) Electrostatic interaction energy histograms for MC simulations performed at $I = 10$ mM (black bars), $I = 100$ mM (red bars), and $I = 200$ mM (green bars) for $^{N-N}$ complexes (E), $^{N-Ph}$ complexes (F), and $^{Ph-Ph}$ complexes (G).

indeed play a role. This is also supported by the MC calculations that show less preorientation for ^{Ph}Pc than for ^{N}Pc . Nevertheless, some preorientation is still observed, indicating that the dipolar nature of the charge distribution is important in complex formation at low ionic strength. The MC results are supported by the CSP and PCS data, which clearly demonstrate that the hydrophobic patch is the side of ^{Ph}Pc that is in contact with $^{N}Cyt f$. However, the MC results do not agree quantitatively with the PRE data, indicating that electrostatic interactions alone are not sufficient to describe the encounter ensemble and the final complex.

The PCS-based final complex shows predominantly hydrophobic contacts, and the ^{Ph}Pc orientation is different from that in the $^{N-N}$ complex, which can be explained by the substitution of several Lys residues on ^{Ph}Pc , resulting in the absence of several charge–charge interactions with negative residues on $^{N}Cyt f$. The encounter complex produced using PRE-driven ensemble docking is similar to that of the $^{N-N}$ complex, though even more diffusive. In both encounter complexes, *Pc* is found in contact with the nonpolar surfaces of Cyt *f*, strongly suggesting that hydrophobic interactions indeed contribute to the encounter complex.

The PCS are much smaller in the $N\text{-}P_h$ complex than in the $N\text{-}N$ complex, suggesting that the encounter complex is more populated. The size of PCS strongly depends on the distance between the heme iron and the *Pc* nucleus that experiences the PCS. Thus, it is expected that in the encounter complex, which is spread over a large surface area of *Cyt f*, the PCS will be smaller than that in the final complex. Orientation averaging may reduce the PCS further. The size of the PCS is about 3-fold less for P_hP_c than for NP_c (Figure 3B).

In encounter complexes that are of an electrostatic nature, CSPs are very small, compared to those in the final complex,^{40,41} and increasing the fraction of the encounter complex strongly reduces the average size of the CSP in those complexes.⁴² The CSPs for P_hP_c in complex with $N\text{Cyt } f$ are also reduced compared to those of NP_c (Figure 1C,D) but not by very much, much less than 3-fold. This is an important observation because significant CSPs may be expected also in the encounter complex if it is stabilized by hydrophobic contacts. The chemical shift of amide groups is particularly sensitive to polarity and hydrogen bond formation; therefore, the desolvation of the protein surface that accompanies the formation of hydrophobic contacts is expected to cause significant CSPs.

It is interesting to compare the effects of ionic strength in the $N\text{-}N$ complex, $N\text{-}P_h$ complex, and $P_h\text{-}P_h$ complex. The addition of 200 mM NaCl to the $P_h\text{-}P_h$ complex ($I = 210$ mM) had essentially no effect on the fraction of bound *Pc*,²¹ suggesting that hydrophobic contacts strongly dominate the interaction. Our simulations are in line with that observation (Figure 8G). The K_D was difficult to determine accurately and was reported to be about 1 mM. Here, we use a range of 1–3 mM. If it is assumed that the hydrophobic contribution to the binding is similar in the three complexes, the contribution of the electrostatic interactions can be estimated for the $N\text{-}N$ complex and the $N\text{-}P_h$ complex. An affinity of 1–3 mM equals a change in free energy of binding of 4.1–3.4 kcal/mol. The K_D values for the $N\text{-}N$ complex and the $N\text{-}P_h$ complex are 80²⁷ and 400 μM in the absence of salt ($I = 10$ mM), suggesting an additional contribution from the charge interactions of 1.5–2.2 kcal/mol and 0.55–1.2 kcal/mol, respectively. Thus, the electrostatic interaction represents 27%–38% and 12%–26% of the total binding energy in the $N\text{-}N$ complex and the $N\text{-}P_h$ complex.

The addition of 160 mM NaCl ($I = 170$ mM) to the $N\text{-}N$ complex reduced the fraction bound by about 50%,²⁵ and it can be calculated on the basis of the protein concentrations used in that experiment that the binding energy decreased with 1.6 kcal/mol, nearly abolishing the charge–charge contribution. The same is observed for the $N\text{-}P_h$ complex, where the addition of 200 mM NaCl ($I = 210$ mM) reduces the fraction bound by 60%, which translates to a loss of –0.8 kcal/mol of binding energy under the given experimental conditions. Thus, under the assumption that the hydrophobic contribution is conserved among these complexes, it can be concluded that the electrostatic contribution represents one-third of the binding energy for the $N\text{-}N$ complex at low ionic strength and much less at more physiological values. For the $N\text{-}P_h$ complex, this fraction is even smaller. The trend is qualitatively supported by the electrostatic interaction histograms from the MC calculations (Figure 8E–G). This is an important finding in relation to earlier *in vivo* studies, in which no significant effects of mutation of charged residues in the interface of an algal *Cyt f*–*Pc* complex could be detected in the activity assay.^{43,44} These results

suggest that charge interactions are not relevant for the complex. However, the results on the cross-complex show that even weak electrostatic interactions are effective in preorienting *Pc* to face *Cyt f* with its hydrophobic patch. Furthermore, many charged residues on both proteins are conserved, especially among plants, suggesting that at least under some circumstances the charge interactions contribute significantly to the electron transfer process in photosynthesis.

In conclusion, the current study fully supports the model complex formation described for the $N\text{-}N$ complex. In the $N\text{-}P_h$ complex, the role of charges has not been abolished at low ionic strength, but it is reduced in favor of hydrophobic contacts, creating a complex with biophysical properties that is a mixture of the $N\text{-}N$ complex and the $P_h\text{-}P_h$ complex. The variation that is observed between mechanisms of complex formation observed for the same complex from different species shows that several ways exist to achieve both fast ET and rapid turnover in protein complexes. The common denominator may be low affinity and low energy barriers between the subsequent states in the reaction.

■ ASSOCIATED CONTENT

● Supporting Information

Plot of the average violation of all experimental distances versus the ensemble percentage included in the restraints for the calculations and a comparison of the encounter complexes of the $N\text{-}P_h$ complex and $N\text{-}N$ complex with *Pc* CoMs color-coded according to Cu–Fe distances. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +31 (0) 71 527 4628. E-mail: m.ubbink@chem.leidenuniv.nl.

Funding

S.S., M.T., and M.U. received financial support from The Netherlands Organisation for Scientific Research (NWO) and grants 700.57.011 (S.S.) and 700.58.441 (M.T. and M.U.). G.M.U. and J.M.F. were supported by the German Science Foundation (DFG; GRK 1640).

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

NMR, nuclear magnetic spectroscopy; *N*, *Nostoc* sp. PCC 7119; *Ph*, *Phormidium laminosum*; *Cyt f*, cytochrome *f*; *Pc*, plastocyanin; ET, electron transfer; MES, 2-(*N*-morpholino) ethanesulfonic acid; HSQC, heteronuclear single quantum coherence; CSP, chemical shift perturbation; PRE, paramagnetic relaxation enhancement; PCS, pseudocontact shift; MTS, (1-acetoxy-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) methanethiosulfonate; MTSL, (1-oxyl-2,2,5,5-tetramethyl- δ -3-pyrroline-3-methyl) methanethiosulfonate; MC, Monte Carlo; CoM, center of mass

■ REFERENCES

- (1) Northrup, S. H., and Erickson, H. P. (1992) Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U.S.A.* 89, 3338–3342.
- (2) Ubbink, M. (2009) The courtship of proteins: understanding the encounter complex. *FEBS Lett.* 583, 1060–1066.
- (3) Schreiber, G. (2002) Kinetic studies of protein-protein interactions. *Curr. Opin. Struct. Biol.* 12, 41–47.

- (4) Camacho, C. J., Weng, Z., Vajda, S., and DeLisi, C. (1999) Free energy landscapes of encounter complexes in protein-protein association. *Biophys. J.* 76, 1166–1178.
- (5) Camacho, C. J., Kimura, S. R., DeLisi, C., and Vajda, S. (2000) Kinetics of desolvation-mediated protein-protein binding. *Biophys. J.* 78, 1094–1105.
- (6) Camacho, C. J., and Vajda, S. (2001) Protein docking along smooth association pathways. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10636–10641.
- (7) Kim, Y. C., Tang, C., Clore, G. M., and Hummer, G. (2008) Replica exchange simulations of transient encounter complexes in protein-protein association. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12855–12860.
- (8) Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021–1025.
- (9) Scanu, S., Foerster, J. M., Ullmann, G. M., and Ubbink, M. (2013) Role of hydrophobic interactions in the encounter complex formation of the plastocyanin and cytochrome complex revealed by paramagnetic NMR spectroscopy. *J. Am. Chem. Soc.* 135, 7681–7692.
- (10) Hope, A. B. (2000) Electron transfers amongst cytochrome *f*, plastocyanin and photosystem I: kinetics and mechanisms. *Biochim. Biophys. Acta* 1456, 5–26.
- (11) Díaz-Quintana, A., Hervás, M., Navarro, J. A., and De la Rosa, M. A. (2008) Plastocyanin and Cytochrome *c₆*: The Soluble Electron Carriers between the Cytochrome *b₆f* Complex and Photosystem I, in *Photosynthetic Protein Complexes: A Structural Approach* (Fromme, P., Ed.), pp 181–200, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
- (12) Crowley, P. B., and Ubbink, M. (2003) Close encounter of the transient kind: protein interactions in the photosynthetic redox chain investigated by NMR spectroscopy. *Acc. Chem. Res.* 36, 723–730.
- (13) Schreiber, G., and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* 3, 427–431.
- (14) Gong, X. S., Wen, J. Q., Fisher, N. E., Young, S., Howe, C. J., Bendall, D. S., and Gray, J. C. (2000) The role of individual lysine residues in the basic patch on turnip cytochrome *f* for electrostatic interactions with plastocyanin in vitro. *Eur. J. Biochem.* 267, 3461–3468.
- (15) Kann, A., Young, S., and Bendall, D. S. (1996) The role of acidic residues of plastocyanin in its interaction with cytochrome *f*. *Biochim. Biophys. Acta* 1277, 115–126.
- (16) Schlarb-Ridley, B. G., Bendall, D. S., and Howe, C. J. (2002) Role of electrostatics in the interaction between cytochrome *f* and plastocyanin of the cyanobacterium *Phormidium laminosum*. *Biochemistry* 41, 3279–3285.
- (17) Hart, S. E., Schlarb-Ridley, B. G., Delon, C., Bendall, D. S., and Howe, C. J. (2003) Role of charges on cytochrome *f* from the cyanobacterium *Phormidium laminosum* in its interaction with plastocyanin. *Biochemistry* 42, 4829–4836.
- (18) Albarran, C., Navarro, J. A., Molina-Heredia, F. P., Murdoch, P. S., De la Rosa, M. A., and Hervas, M. (2005) Laser flash-induced kinetic analysis of cytochrome *f* oxidation by wild-type and mutant plastocyanin from the cyanobacterium *Nostoc* sp. PCC 7119. *Biochemistry* 44, 11601–11607.
- (19) Albarran, C., Navarro, J., De la Rosa, M. A., and Hervas, M. (2007) The specificity in the interaction between cytochrome *f* and plastocyanin from the cyanobacterium *Nostoc* sp. PCC 7119 is mainly determined by the copper protein. *Biochemistry* 46, 997–1003.
- (20) Ubbink, M., Ejdeback, M., Karlsson, B. G., and Bendall, D. S. (1998) The structure of the complex of plastocyanin and cytochrome *f*, determined by paramagnetic NMR and restrained rigid-body molecular dynamics. *Structure* 6, 323–335.
- (21) Crowley, P. B., Otting, G., Schlarb-Ridley, B. G., Canters, G. W., and Ubbink, M. (2001) Hydrophobic interactions in a cyanobacterial plastocyanin-cytochrome *f* complex. *J. Am. Chem. Soc.* 123, 10444–10453.
- (22) Diaz-Moreno, I., Diaz-Quintana, A., De la Rosa, M. A., and Ubbink, M. (2005) Structure of the complex between plastocyanin and cytochrome *f* from the cyanobacterium *Nostoc* sp. PCC 7119 as determined by paramagnetic NMR. The balance between electrostatic and hydrophobic interactions within the transient complex determines the relative orientation of the two proteins. *J. Biol. Chem.* 280, 35784.
- (23) Lange, C., Cornvik, T., Diaz-Moreno, I., and Ubbink, M. (2005) The transient complex of poplar plastocyanin with cytochrome *f*: effects of ionic strength and pH. *Biochim. Biophys. Acta* 1707, 179–188.
- (24) Hulsker, R., Baranova, M. V., Bullerjahn, G. S., and Ubbink, M. (2008) Dynamics in the transient complex of plastocyanin-cytochrome *f* from *Prochlorothrix hollandica*. *J. Am. Chem. Soc.* 130, 1985–1991.
- (25) Diaz-Moreno, I., Diaz-Quintana, A., De la Rosa, M. A., Crowley, P. B., and Ubbink, M. (2005) Different modes of interaction in cyanobacterial complexes of plastocyanin and cytochrome *f*. *Biochemistry* 44, 3176–3183.
- (26) Schlarb, B. G., Wagner, M. J., Vijgenboom, E., Ubbink, M., Bendall, D. S., and Howe, C. J. (1999) Expression of plastocyanin and cytochrome *f* of the cyanobacterium *Phormidium laminosum* in *Escherichia coli* and *Paracoccus denitrificans* and the role of leader peptides. *Gene* 234, 275–283.
- (27) Scanu, S., Foerster, J., Finiguerra, M. G., Shabestari, M. H., Huber, M., and Ubbink, M. (2012) The complex of cytochrome *f* and plastocyanin from *Nostoc* sp. PCC 7119 is highly dynamic. *ChemBioChem* 13, 1312–1318.
- (28) Milikisyants, S., Scarpelli, F., Finiguerra, M. G., Ubbink, M., and Huber, M. (2009) A pulsed EPR method to determine distances between paramagnetic centers with strong spectral anisotropy and radicals: the dead-time free RIDME sequence. *J. Magn. Reson.* 201, 48–56.
- (29) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- (30) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687–696.
- (31) Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labelling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* 39, 5355–5365.
- (32) Baniulis, D., Yamashita, E., Whitelegge, J. P., Zatsman, A. I., Hendrich, M. P., Hasan, S. S., Ryan, C. M., and Cramer, W. A. (2009) Structure-function, stability, and chemical modification of the cyanobacterial cytochrome *b₆f* complex from *Nostoc* sp. PCC 7120. *J. Biol. Chem.* 284, 9861–9869.
- (33) Banci, L., Bertini, I., Cavallaro, G., Giachetti, A., Luchinat, C., and Parigi, G. (2004) Paramagnetism-based restraints for Xplor-NIH. *J. Biomol. NMR* 28, 249–261.
- (34) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* 160, 65–73.
- (35) Ullmann, G. M., Knapp, E. W., and Kostic, N. M. (1997) Computational simulation and analysis of dynamic association between plastocyanin and cytochrome *f*. Consequences for the electron-transfer reaction. *J. Am. Chem. Soc.* 119, 42–52.
- (36) Bashir, Q., Volkov, A. N., Ullmann, G. M., and Ubbink, M. (2010) Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome *c* and cytochrome *c* peroxidase. *J. Am. Chem. Soc.* 132, 241–247.
- (37) Ubbink, M., Lian, L. Y., Modi, S., Evans, P. A., and Bendall, D. S. (1996) Analysis of the 1H-NMR chemical shifts of Cu(I)-, Cu(II)- and Cd-substituted pea plastocyanin. Metal-dependent differences in the hydrogen-bond network around the copper site. *Eur. J. Biochem.* 242, 132–147.
- (38) Carrell, C. J., Schlarb, B. G., Bendall, D. S., Howe, C. J., Cramer, W. A., and Smith, J. L. (1999) Structure of the soluble domain of cytochrome *f* from the cyanobacterium *Phormidium laminosum*. *Biochemistry* 38, 9590–9599.

(39) Clore, G. M., and Iwahara, J. (2009) Theory, practice and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* 109, 4108–4139.

(40) Worrall, J. A. R., Liu, Y. J., Crowley, P. B., Nocek, J. M., Hoffman, B. M., and Ubbink, M. (2002) Myoglobin and cytochrome b5: a nuclear magnetic resonance study of a highly dynamic protein complex. *Biochemistry* 41, 11721–11730.

(41) Xu, X. F., Reinle, W. G., Hannemann, F., Konarev, P. V., Svergun, D. I., Bernhardt, R., and Ubbink, M. (2008) Dynamics in a pure encounter complex of two proteins studied by solution scattering and paramagnetic NMR spectroscopy. *J. Am. Chem. Soc.* 130, 6395–6403.

(42) Volkov, A. N., Bashir, Q., Worrall, J. A. R., Ullmann, G. M., and Ubbink, M. (2010) Shifting the equilibrium between the encounter state and the specific form of a protein complex by interfacial point mutations. *J. Am. Chem. Soc.* 132, 11487–11495.

(43) Soriano, G. M., Ponamarev, M. V., Tae, G. -S., and Cramer, W. A. (1996) Effect of the interdomain basic region of cytochrome f on its redox reactions in vivo. *Biochemistry* 35, 14590–14598.

(44) Soriano, G. M., Ponamarev, M. V., Piskorowski, R. A., and Cramer, W. A. (1998) Identification of the basic residues of cytochrome f responsible for electrostatic docking interactions with plastocyanin in vitro: relevance to the electron transfer reaction in vivo. *Biochemistry* 37, 15120–15128.

SUPPORTING INFORMATION TO:

Loss of electrostatic interactions causes increase of dynamics within the plastocyanin-cytochrome *f* complex

Sandra Scanu, Johannes M. Foerster, Monika Timmer, G. Matthias Ullmann and Marcellus Ubbink

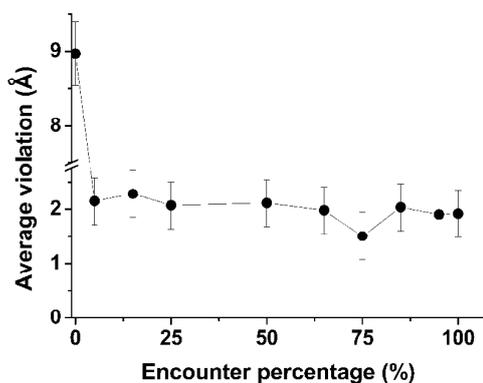


Figure S1. Plot of the average violation of all experimental distances versus the ensemble percentage included in the restraints for the calculations. Error bars represent $2 \times \text{SD}$ of the average violations obtained from three independent calculations performed with $N=7$ and $f_2=0$.

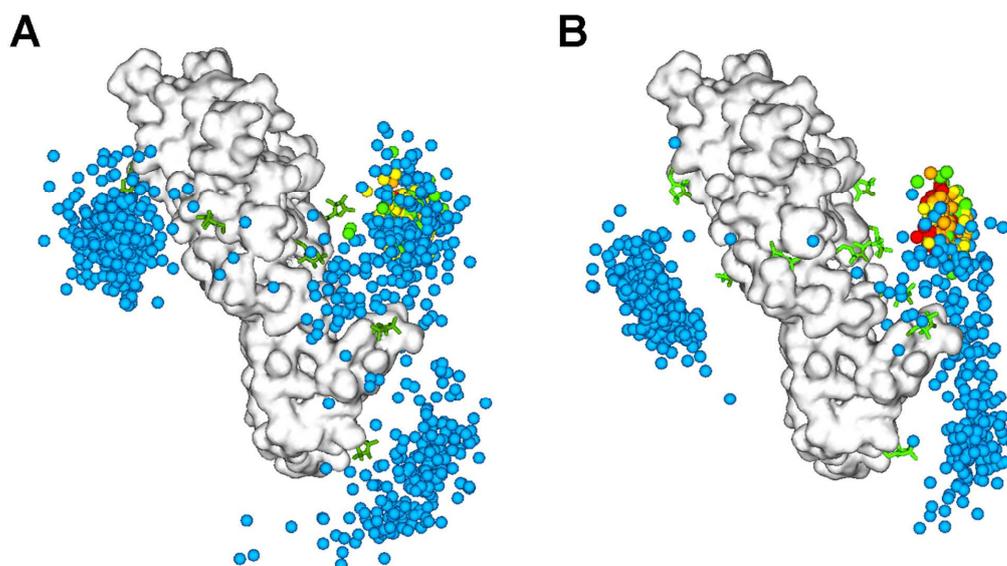


Figure S2. Comparison of the encounter complexes of $N\text{-}Ph$ complex (A) and $N\text{-}N$ complex (B). Cyt f is shown as a white surface and spin labels as green sticks. Pc CoMs are represented by spheres. Pc CoMs are color-coded to indicate the distance between Cu in Pc and Fe in Cyt f , increasing from red to blue (red ≤ 16 Å; orange ≤ 18 Å; yellow ≤ 20 Å; green ≤ 22 Å; blue > 22 Å).

Chapter 10

Manuscript E

The dynamic complex of cytochrome c_6 and cytochrome f studied with paramagnetic NMR spectroscopy

Irene Díaz-Moreno, Rinske Hulsker, Pavol Skubak, Johannes M. Foerster, Davide Cavazzini, Michelina G. Finiguerra, Antonio Díaz-Quintana, Blas Moreno-Beltrán, Gian-Luigi Rossi, G. Matthias Ullmann, Navraj S.Pannu, Miguel A. De la Rosa and Marcellus Ubbink

Biochim. Biophys. Acta. 2014, 1837 (8), 1305-1315
doi: 10.1016/j.bbabi.2014.03.009

Reprint with permission of
Biochimica et Biophysica Acta (BBA) - Bioenergetics.
© 2014 Copyright by Elsevier B.V.



Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

The dynamic complex of cytochrome c_6 and cytochrome f studied with paramagnetic NMR spectroscopy

Irene Díaz-Moreno ^a, Rinske Hulsker ^b, Pavol Skubak ^b, Johannes M. Foerster ^c, Davide Cavazzini ^d, Michelina G. Finiguerra ^b, Antonio Díaz-Quintana ^a, Blas Moreno-Beltrán ^a, Gian-Luigi Rossi ^d, G. Matthias Ullmann ^c, Navraj S. Pannu ^b, Miguel A. De la Rosa ^a, Marcellus Ubbink ^{b,*}

^a Instituto de Bioquímica Vegetal y Fotosíntesis, Universidad de Sevilla-CSIC, Avda. Américo Vesputio 49, Sevilla 41092, Spain

^b Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

^c Computational Biochemistry, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany

^d Department of Life Sciences, Laboratory of Biochemistry, Molecular Biology and Bioinformatics, Parco Area delle Scienze 23/a, University of Parma, 43124 Parma, Italy

ARTICLE INFO

Article history:

Received 27 January 2014

Received in revised form 13 March 2014

Accepted 16 March 2014

Available online 28 March 2014

Keywords:

Electron transfer

Photosynthesis

Crystallography

Paramagnetic relaxation enhancement

Monte-Carlo modelling

Protein interaction

ABSTRACT

The rapid transfer of electrons in the photosynthetic redox chain is achieved by the formation of short-lived complexes of cytochrome b_6f with the electron transfer proteins plastocyanin and cytochrome c_6 . A balance must exist between fast intermolecular electron transfer and rapid dissociation, which requires the formation of a complex that has limited specificity. The interaction of the soluble fragment of cytochrome f and cytochrome c_6 from the cyanobacterium *Nostoc* sp. PCC 7119 was studied using NMR spectroscopy and X-ray diffraction. The crystal structures of wild type, M58H and M58C cytochrome c_6 were determined. The M58C variant is an excellent low potential mimic of the wild type protein and was used in chemical shift perturbation and paramagnetic relaxation NMR experiments to characterize the complex with cytochrome f . The interaction is highly dynamic and can be described as a pure encounter complex, with no dominant stereospecific complex. Ensemble docking calculations and Monte-Carlo simulations suggest a model in which charge–charge interactions pre-orient cytochrome c_6 with its haem edge toward cytochrome f to form an ensemble of orientations with extensive contacts between the hydrophobic patches on both cytochromes, bringing the two haem groups sufficiently close to allow for rapid electron transfer. This model of complex formation allows for a gradual increase and decrease of the hydrophobic interactions during association and dissociation, thus avoiding a high transition state barrier that would slow down the dissociation process.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Protein complex formation is at least a two-step process [3] in which the formation of a final, well-defined complex – dominated by short-range interactions – entails the initial formation of a dynamic encounter complex. The lifetime of the protein complex is determined by the dissociation rate. Highly transient complexes, with lifetimes on the order of milliseconds, exhibit moderate or low binding affinities, with

dissociation constants in the μM – mM range. Electron transfer (ET^{\dagger}) reactions mediated by soluble redox proteins exchanging electrons between large membrane complexes in photosynthesis and respiration are excellent examples of transient interactions. The purpose of the protein complex formation in these cases is two-fold. A complex must be formed that is sufficiently specific to allow rapid electron transfer and at the same time the complex needs a high dissociation rate to enable rapid turn-over in order not to limit the flow of electrons through the redox chain. The electron transfer rate is exponentially dependent on the distance between the redox centres. Thus, bringing the centres in close approximation ($< 16 \text{ \AA}$) [4] is essential, but the formation of a well-defined complex is not required if multiple orientations exist in which ET can occur. In fact, such a specific complex is not desirable from the point of view of fast dissociation, because a well-defined state has a lower free energy than all similar states and thus a higher transition state energy to be overcome to dissociate. In other words, high specificity opposes rapid turnover. The study of transient complexes enables the understanding of the biophysical mechanisms that exist to reach the right compromise between these two properties of a complex.

Abbreviations: Cb6f, Cytochrome b_6f ; Cc6, Cytochrome c_6 ; Cf, Cytochrome f ; ET, Electron transfer; PRE, Paramagnetic relaxation enhancement; MTS, (1-Acetyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate; MTSL, (1-Oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate

* Corresponding author at: Tel.: +31 715274628.

E-mail addresses: idiazmoreno@us.es (I. Díaz-Moreno), hulsker_rinske@yahoo.co.uk (R. Hulsker), p.skubak@chem.leidenuniv.nl (P. Skubak), johannes.foerster@uni-bayreuth.de (J.M. Foerster), davide.cavazzini@unipr.it (D. Cavazzini), mfinig@hotmail.com (M.G. Finiguerra), qzaida@us.es (A. Díaz-Quintana), joseblas.moreno@ibvf.csic.es (B. Moreno-Beltrán), gianluigi.rossi@unipr.it (G.-L. Rossi), Matthias.Ullmann@uni-bayreuth.de (G.M. Ullmann), raj@chem.leidenuniv.nl (N.S. Pannu), marosa@us.es (M.A. De la Rosa), m.ubbink@chem.leidenuniv.nl (M. Ubbink).

<http://dx.doi.org/10.1016/j.bbabio.2014.03.009>

0005-2728/© 2014 Elsevier B.V. All rights reserved.

In oxygenic photosynthesis, the ET from the cytochrome *b₆f* (Cb6f) complex to Photosystem I (PSI) – both membrane-embedded complexes – is carried out by two soluble metalloproteins, plastocyanin (Pc) and cytochrome *c*₆ (Cc6) [5–8]. Most cyanobacteria and green algae synthesize either Pc or Cc6, depending on the availability of copper and iron, their respective cofactor metals [9,10]. Higher plants only contain Pc, although a Cc6-like protein has been identified in *Arabidopsis* [11], but it is unable to transfer electrons to PSI [12].

The hetero-oligomeric Cb6f complex contains eight tightly bound polypeptide subunits that couple the ET to proton translocation, generating a proton electrochemical potential gradient necessary for ATP synthesis. The three-dimensional crystal structure has been determined for the Cb6f complex from the green alga *Chlamydomonas reinhardtii* [13], the cyanobacteria *Mastigocladus laminosus* [14] and *Nostoc* sp. PCC 7120 [15]. The main difference between the cyanobacterial Cb6f crystallographic structures is the acetylation of the *Nostoc* Rieske Fe–S protein at the N terminus, a post-translational modification unprecedented in cyanobacterial membrane and ET proteins [15]. Cytochrome *f* (Cf) is a subunit of the Cb6f complex, anchored to the thylakoid membrane by a C-terminal transmembrane helix leaving a 28-kDa soluble portion exposed to the lumen with a clear two-domain structure. The large domain harbours the haem and the small domain possesses a patch of charged residues. Cf is considered an unusual c-type cytochrome because of its β -sheet structure, elongated form and particular haem axial coordination with the amino group of the N-terminus, residue Tyr1 [13–17].

Cc6 is a more typical 10-kDa single haem c-type cytochrome with the cofactor covalently bound to the cysteine residues in a CXXCH motif. The Fe atom is hexacoordinated with His and Met residues acting as axial ligands, as revealed by the available cyanobacterial and green algal Cc6 structures [18–22]. One of the most important functional characteristic of Cc6 is its midpoint redox potential (E_m) around +335 mV at physiological pH value, with the exception of that present in plants whose E_m is substantially lower (ca. +100 mV) despite having the same axial ligands [23]. This finding can be partly explained by the replacement of a highly conserved Gln in cyanobacterial Cc6 by a Val residue in the plant Cc6-like protein, which regulates the Fe–S(Met) bond stability and causes a 100 mV-drop in the E_m [20,24]. A more drastic E_m change occurs when the sixth axial ligand Met is substituted by His, leading to inhibition of both the spontaneous self-reduction of Cc6 mutant and its reduction by the Cb6f complex [25].

Cc6-involving physiological interactions have been extensively studied in recent years as a model to understand the nature of protein–protein interactions in ET chains. The Cc6–PSI interaction from *Nostoc* has been well-characterized from the structural and functional point of view [26–30]. Fast-kinetics studies combined with Brownian dynamics using a *Chlamydomonas* Zn–Cc6 derivative and Cf have been reported [31,32], concluding that the nature of this complex is dynamic and that hydrophobic contacts are important. Two NMR-based structural approaches using haem proteins from different cyanobacterial sources also suggest that the binding site on *Nostoc* Cc6 involves the predominantly hydrophobic patch surrounding the Cf haem [33,34]. *In silico* data on *Chlamydomonas* Cc6–Cf complex show not only the relevance of hydrophobic and electrostatic interactions in bringing both haem proteins sufficiently close to allow efficient ET [35, 36], but also the key role of the Cf small domain in binding to Cc6, suggesting that Cc6 explores different positions on Cf [37].

Here, experimental approaches using NMR spectroscopy are combined with charge-driven docking simulations to study the molecular recognition processes in ET complexes, using the physiological *Nostoc* Cc6–Cf interaction as a model system. Our paramagnetic relaxation enhancement (PRE) NMR data are not compatible with a well-defined Cc6–Cf complex. The complex is best described by a highly dynamic ensemble, first formed by electrostatic pre-orientation and stabilized mainly by hydrophobic contacts.

2. Materials and methods

2.1. Mutagenesis

The expression vector pEAC-WT for wt Cc6 from *Nostoc* sp. PCC 7119 [38] was used as the template for site-directed mutagenesis to obtain the M58H and M58C variants using the QuikChange PCR protocol (Stratagene, La Jolla, CA). The following primer pairs were used: 5' CCGTAAGAACGCCACCTGCTTTCAAAGG and its complement for M58H and 5' CCGTAAGAACGCCTGCCCTGCTTTCAAAGG and its complement for M58C. For the introduction of Cys residues in Cf, the pEAF-wt [39] expression plasmid encoding the soluble domain of Cf from *Nostoc* sp. PCC7119 was used as template. The single-cysteine variants Q7C, A63C, N71C, Q104C and S192C have been described before [40,41]. All constructs were verified by DNA sequencing.

2.2. Protein production and purification

Uniformly ¹⁵N-labelled *Nostoc* sp. PCC 7119 Cc6 wt and its mutants were produced as described before [26] in *Escherichia coli* JM109 cells co-transformed with pEAC-WT [38] and pEC86 [42]. Culture conditions and protein purification methods were as reported previously [26,33]. Protein concentrations were determined by absorption spectrophotometry using a ϵ_{553} of 26.2 mM⁻¹ cm⁻¹ for the ferrous form of Cc6 wt [38], a ϵ_{554} of 20.8 mM⁻¹ cm⁻¹ for the ferrous form of M58H and a ϵ_{540} of 7.2 mM⁻¹ cm⁻¹ for the ferric form of M58C mutant. The Cc6 wt ϵ_{278} was estimated using protein concentration values from Bradford assays. A A_{278}/A_{553} ratio of 1.05 of the wt ferrous Cc6 indicated sufficient purity for characterization by NMR.

To obtain a high yield of holo-Cf and promote the correct insertion of the haem group, *E. coli* strain MV1190 (Bio-Rad) was co-transformed with plasmids pEC86 and (mutated) pEAF plasmid. The cells were plated on Lysogeny Broth (LB) medium plates and incubated at 37 °C for 24 h. All media were supplemented with 20 mg/L ampicillin and chloramphenicol. Several pre-cultures were prepared in 100 mL flasks with 20 mL of LB medium and incubated at 37 °C and 250 rpm for 5–6 h. The pre-cultures with the highest OD₆₀₀ were used to inoculate 1.7 L (in 2 L Erlenmeyer flasks) of LB, ratio 1:100. The cultures were incubated at 25 °C and 150 rpm under semi-anaerobic conditions and high antibiotic pressure by adding further ampicillin and chloramphenicol after 20 h and 40 h. Expression was induced 20 h after the inoculation of the large culture using 1 mM IPTG (isopropyl- β -thiogalactopyranoside). More than 80 h after the induction the cultures appeared brown because of the presence of Cf. The cells were harvested by centrifugation and the periplasmic fraction was extracted by osmotic shock. The pink water fraction (about 200 mL per 1.7 L of culture), was dialyzed against 2 L of 5 mM Tris–HCl buffer, pH 8 and 3 mM dithiothreitol (DDT). The yield in the periplasmic fraction was 10 mg/L of culture of protein for N71C and Q7C and 5 mg/L for Q104C, S192C and A63C. The resulting dialysate was cleared by centrifugation and loaded on a DEAE column equilibrated in the same buffer. Elution was performed with a gradient of 20–500 mM NaCl and 3 mM DTT. The fraction containing the Cf was concentrated and loaded on a gel-filtration (G75 Superdex) column and eluted in the same buffer containing 150 mM NaCl. The protein fractions were pooled, concentrated, dialyzed against 5 mM MES, pH 6 and 3 mM DTT and loaded on a DEAE column equilibrated in the same buffer. The Cf was eluted with a gradient of 0–500 mM NaCl. Pure fractions showed a A_{280}/A_{556} of 1.3 under reducing conditions. Protein concentrations were determined by optical spectroscopy using ϵ_{419} of 85.5 mM⁻¹ cm⁻¹ for Cc6 M58C and $\epsilon_{556} = 31.5$ mM⁻¹ cm⁻¹ for reduced Cf [39].

2.3. Labelling of Cf with spin labels

For attachment of spin label to Cf, DTT was first removed by ultrafiltration (Amicon, MW cut-off 10 kDa). The protein was subsequently

exchanged to 10 mM sodium phosphate, pH 6.0 and concentrated to ~40 μ M. The protein was oxidized by a 100-fold excess of $K_3[Fe(CN)_6]$ and a 10-fold excess was added of either MTSL [(1-oxy-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate] or MTS [(1-acetyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate] (TRC, North York, Ontario, Canada). Stock solutions of 0.1 M MTS(L) in DMSO were used. The protein solution was kept for 2 h at RT and O/N at 4° after which the excess $K_3[Fe(CN)_6]$ and MTS(L) were removed by ultrafiltration.

2.4. Electrochemistry

The redox potential value for the haem group in each Cc6 wt and mutants was determined as reported previously [38], for which the differential absorbance changes at 553 minus 570 nm were followed. Menadione, diaminodurool and *p*-benzoquinone, at 20 μ M final concentration, were used as redox mediators. Errors in the experimental determinations were less than 20 mV.

2.5. Crystallization and data collection

Crystals were obtained with the sitting drop method. The final protein concentration was 10 mg/mL in the following solutions: Cc6 wt: 0.1 M Tris/HCl pH 7.0, 2.3 M ammonium sulphate, 0.1 M lithium sulphate; Cc6 M58H: 0.1 M Tris/HCl pH 8.0, 2.5 M ammonium sulphate, 0.1 M lithium sulphate; and Cc6 M58C: 0.1 M citrate pH 5.0, 2.5 M ammonium sulphate. The crystals were frozen and diffraction data for the wild type and M58H crystals were collected at the BM16 beamline of the ESRF synchrotron on a MAR 165 CCD detector at the peak wavelength of the iron (1.5418 Å), whereas M58C crystals were collected in-house on an Enraf-Nonius FR591 generator and MAD 345 image plate detector. All crystals were collected with a 1.0° oscillation at 100 K. For the wild type crystals 360 images were collected, for the M58H crystals 230 and for the M58C crystals 200 images. All data sets were processed by MOSFLM [43] and SCALA [44] from CCP4 [45].

Molecular replacement for the wild type Cc6 structure was unsuccessful due to the presence of translational non-crystallographic symmetry, as noted by a large off-origin peak in the Patterson map. However, the anomalous signal from the intrinsic iron atoms was sufficient for structure determination by SAD phasing. The CRANK [46] software pipeline was used to solve the structure and CRUNCH2 was used [47] for substructure detection, BP3 [46] for heavy atom refinement, and density modification by DM [48] estimates. Automated model building with ARP/wARP [49] using the iterative refinement with the SAD target [50] in REFMAC [51] provided a good quality model of the structure consisting of 483 backbone residues, 478 of which were (correctly) docked in the 6 wild type molecules present in the asymmetric unit. Some of the chains traced missed several residues from either the C- or N-terminus, however, one chain contained all the residues fitting well in the electron density. The haem group was fitted manually into the chain that was built completely at this stage. The resulting completely built Cc6 molecule was superimposed on the five other cytochrome molecules present in the asymmetric unit to fit in any missed residues present in the density as well as the other haem groups. The model obtained in this way was refined by REFMAC5 with tight NCS restraints. Manual corrections to the model were performed with XtalView [52], followed by refinement with REFMAC with loose NCS restraints.

The point mutation M58C crystal was isomorphous to the wild type crystal, thus the final wild type model was used as a starting model in the refinement of the M58C mutant. The structure of M58H Cc6 was solved by molecular replacement using the wild type structure and contained two molecules in the asymmetric unit. Manual fitting and refinement of both mutant structures was done with COOT [1] and REFMAC. Data collection and refinement statistics are reported in Table 1. The coordinates have been deposited in the protein data bank

with PDB IDs: 4GYD for wildtype, 4H0J for M58C, and 4H0K for M58H Cc6.

2.6. NMR spectroscopy and data analysis

Cc6 wt and M58C mutant protein solutions were concentrated to the required volume by ultrafiltration methods (Amicon, YM3 membrane) and exchanged into 10 mM sodium phosphate, pH 6.0, H₂O/D₂O 95:5 (v/v) solutions. The soluble domain of Cf was concentrated using an Amicon YM10 membrane and exchanged into 10 mM sodium phosphate, pH 6.0, 3 mM sodium ascorbate, H₂O/D₂O 95:5 solutions. A 3.7 mM ferrous Cf stock solution with a A₂₇₈/A₅₅₆ ratio of 0.9 was used. Cf was kept in a reduced form with a few equivalents of sodium ascorbate and was stable in this form for days. The ferric form was prepared by the addition of a 5-fold excess of potassium ferricyanide ($K_3[Fe(CN)_6]$) followed by gel filtration (Amersham Biosciences Superdex G75) to remove ferrocyanide. Complete oxidation was verified by the disappearance of the absorption band at 556 nm. Then, a 2.0 mM ferric Cf stock solution was prepared.

All NMR experiments were performed on a Bruker DMX 600 NMR spectrometer with a TXI or TCI-cryo triple resonance probehead operating at 298 K. The ¹H and ¹⁵N assignments of the backbone amide resonances from ferric *Nostoc* M58C Cc6 mutant (Table S1) were elucidated by recording 2D ¹H, ¹⁵N HSQC-NOESY with 150 ms mixing time and 2D ¹H, ¹⁵N HSQC-TOCSY with 80-ms mixing time spectra. The effects of complex formation on M58C were followed by acquiring 2D ¹H, ¹⁵N HSQC spectra during titrations of aliquots of Cf stock solutions into a solution of 0.2 mM ¹⁵N-labelled M58C. For the measurements of PRE, samples contained 0.3 mM ¹⁵N M58C Cc6 and 0.1 mM Cf-MTS(L).

All data processing was performed with AZARA 2.7 (www2.ccpn.ac.uk/azara), and spectral analysis was performed with Ansig [53,54]. The spectra were calibrated against the internal standard [¹⁵N]acetamide (0.5 mM).

Titration curves were obtained by plotting chemical-shift perturbations ($\Delta\delta_{\text{bind}}$) against the molar ratio of Cf and Cc6 M58C for the most strongly affected signals. Non-linear least squares fits to a 1:1 binding model [55] were performed in Origin 8.0 (Microcal Inc.). The chemical-shift perturbations (CSP) observed in the complex M58C-Cf with 3 eq. of Cf were extrapolated to 100% bound for all residues using the K_a obtained from the fits. The average chemical-shift perturbation ($\Delta\delta_{\text{avg}}$) of each amide was calculated using the following equation: $\Delta\delta_{\text{avg}} = (((\Delta\delta_N / 5)^2 + (\Delta\delta_H)^2) / 2)^{1/2}$ in which $\Delta\delta_N$ is the change in the ¹⁵N chemical shift, and $\Delta\delta_H$ is the change in the ¹H chemical shift when the protein is 100% bound to Cf.

PREs were derived from the ratio of intensities in the spectra from paramagnetic and diamagnetic samples and converted into distances, as described [56,57]. The correlation time assumed for the Cf and Cc6 complex was 20 ns. PREs are only observed for the fraction of Cc6 that is bound to Cf. The binding is in fast exchange, so the PREs are weighted by the fraction bound. Therefore, the PREs were extrapolated to the 100% bound state of Cc6 for docking calculations.

2.7. Ensemble docking

Cf from *Nostoc* sp. PCC 7120 is identical to that of PCC 7119. The crystal structure of Cf of the former species from PDB ID 2ZT9 [15], residues 1–254, was modified to introduce Cys residues and MTSL spin labels for Gln 7, Ala 63, Asn 71, Gln 104 and Ser 192. Each spin label was built in four orientations, to represent its mobility [58]. For Cc6 the structure from the wt protein from *Nostoc* sp. PCC 7119 (this study) was used. Protons were added to both structures.

In the combined single structure/ensemble docking, first the experimental distances were assigned as restraints between the oxygens of the four MTSL conformers of a spin label and an amide proton of a single copy of Cc6. During the docking, the spin labels were free to rotate. After docking, the distances were measured in the lowest energy structure

Table 1

Data collection and refinement statistics for Cc6 crystals.

	Wild type	M58C	M58H
Space group	P 21 21 2	P 21 21 2	P 32 2 1
Cell dimensions a, b, c (Å)	77.72, 79.80, 80.15	78.82, 80.16, 80.15	60.37, 60.37, 95.37
Resolution (Å)	18.00–1.80 (1.90–1.80) ^a	28.33–2.00 (2.10–2.00) ^a	35.23–1.95 (2.06–1.95) ^a
R _{pim}	0.026 (0.113)	0.021 (0.129)	0.022 (0.221)
I/σI	27.0 (10.1)	23.8 (6.3)	18.8 (3.4)
Completeness (%)	97.3 (94.0)	98.6 (93.8)	99.7 (99.1)
Redundancy	3.4 (2.4)	7.5 (5.2)	12.4 (11.4)
N ^o unique reflections	44311	32761	14317
N ^o molecules in ASU	6	6	2
R _{work} /R _{free}	0.182/0.216	0.207/0.246	0.213/0.268
R.m.s. deviations			
Bond lengths (Å)	0.021	0.017	0.018
Bond angles (°)	1.92	1.72	1.96
Ramachandran favored ^b	90.7%	93.7%	91.6%
Ramachandran outliers ^b	0.40%	0.20%	2.99%

^a Values from the highest resolution shell are given in brackets.^b As defined by COOT.

and converted back to PREs. These back-calculated values were subtracted from the experimental PREs and the difference served as the input for the ensemble docking. Five copies of Cc6 were used and the input PREs were assigned as distance restraints between the nitroxy oxygens and the five copies of the amide protons simultaneously. All averaging was done using the sixth power of the distance. The Cf structure resulting from the single-structure docking was used as the input for the ensemble docking and the spin label orientations were fixed during the ensemble docking. Because of this, the compatibility between the two stages of the docking is ensured. The top ten ensembles were used to calculate the average distance violations. To ensure that the ensemble docking is not strongly influenced by the starting structure, the ensemble docking was repeated several times using the next best structure from the single structure docking as input and repeating the entire procedure. These results yielded the average distance violation and error margins (SD) shown in Fig. 3. For the 0% and 100% ensemble fractions, only a single-structure or ensemble docking was used, respectively, and the experimental PREs served as the restraints. All docking calculations treated the proteins as rigid bodies, using the rigid body dynamic routine in XPLOR-NIH [59]. PREs were divided into three classes, as described before [57]. Resonances that were not significantly affected yielded a distance that served as lower bound only. Resonances that were affected but not completely broadened yielded a distance (with an upper and lower bound of 3 Å in the docking calculations) and signals that were strongly affected (PRE of 100% bound state > 200 s⁻¹) or completely broadened provided only an upper bound distance. The distance violations were defined as the difference between the target distance or range and the back-calculated value. For the first and the third class, this criterion means a violation for values below or above the distance bound, respectively, and for the second class the violations are the absolute difference between target and back-calculated values (so the error margins were not considered for the violation).

2.8. Monte Carlo docking

In the Monte-Carlo simulations the PDB IDs 2ZT9 [15] and 4GYD (this work) were used for Cf and Cc6, respectively. The structure preparation and the Monte Carlo simulation [60] was similarly performed as was done before [40,61]. The iron of Cf and Cc6 were considered to be in the oxidized state. In order to match the experimental conditions, the electrostatic potential was calculated for an ionic strength of 0.01 M and a temperature of 298 K with APBS [62]. 10,000 randomly chosen encounters of the simulation were used for the analysis.

3. Results

3.1. Self-reduction of Cc6 and ligand mutagenesis

For the characterization of the complex of Cf and Cc6, paramagnetic relaxation enhancement (PRE) was used. With this method intermolecular PREs of Cc6 nuclei are measured that are caused by spin labels attached to the surface of Cf at various positions. The PREs are then converted into distance restraints for docking calculations. The spin label MTSL is linked to site-specific Cys residues engineered on Cf via a disulphide bridge. To maintain the spin label in the paramagnetic state and the disulphide bridge intact, it is essential that both cytochromes remain in the oxidized, ferric state. In the past we experienced rapid self-reduction with c-type cytochromes, a phenomenon that has been described also by others for yeast iso-1-cytochrome c [63]. To avoid the problem, it was decided to use a mutant Cc6 with a much lower redox potential to prevent self-reduction. Two mutants were produced in which the purposed axial Met ligand (M58) to the haem iron was replaced with either His or Cys.

The midpoint potential of wt Cc6 is 335 mV at pH 7.0 [38]. Both mutations result in a very large decrease of the midpoint potential, with E_m = -140 mV (pH 7.0) and -235 mV (pH 6.5) [25] for M58H and M58C Cc6, respectively. Thus, the replacement of the Met with a Cys ligand decreases the potential by 570 mV. To determine the effect of the mutation on the structure, the crystal structures were determined. Table 1 reports the refinement statistics and Fig. 1A illustrates the quality of the data. The structure of the wt Cc6 of *Nostoc* sp. PCC 7119 consists of the classic Cc fold, with 4 α-helices and 3 coils. The haem group is attached covalently to Cys14 and Cys17, and His18 and Met58 coordinate the iron. The structure is very similar to Cc6 from other sources, including green algae [18–22]. The closest resemblance is to Cc6 from *Phormidium laminosum* with an RMSD for the backbone heavy atoms of 0.58 Å (PDB ID: 2V08) [20]. The structures of the mutants are similar to that of the wt Cc6, with backbone RMSDs of 0.95 and 0.13 Å for M58H and M58C, respectively. Clearly, the M58C Cc6 structure is essentially identical to that of the wild type, although the thiolate–iron distance in M58C is longer for all six of the Cc6 molecules in the asymmetric unit (the average distance is 3.27 ± 0.04 Å) than the distance between the iron and the S^o of the Met ligand (2.38 ± 0.06 Å). A picture of the electron density in this region for one of the molecules is shown in Fig. 1A.

Significant differences are observed for M58H Cc6 (Fig. 1, panels B and C). The His 58 N^c is coordinated to the iron (2.02 Å), resulting in a backbone change around Lys 55, rotating it to a more solvent-exposed

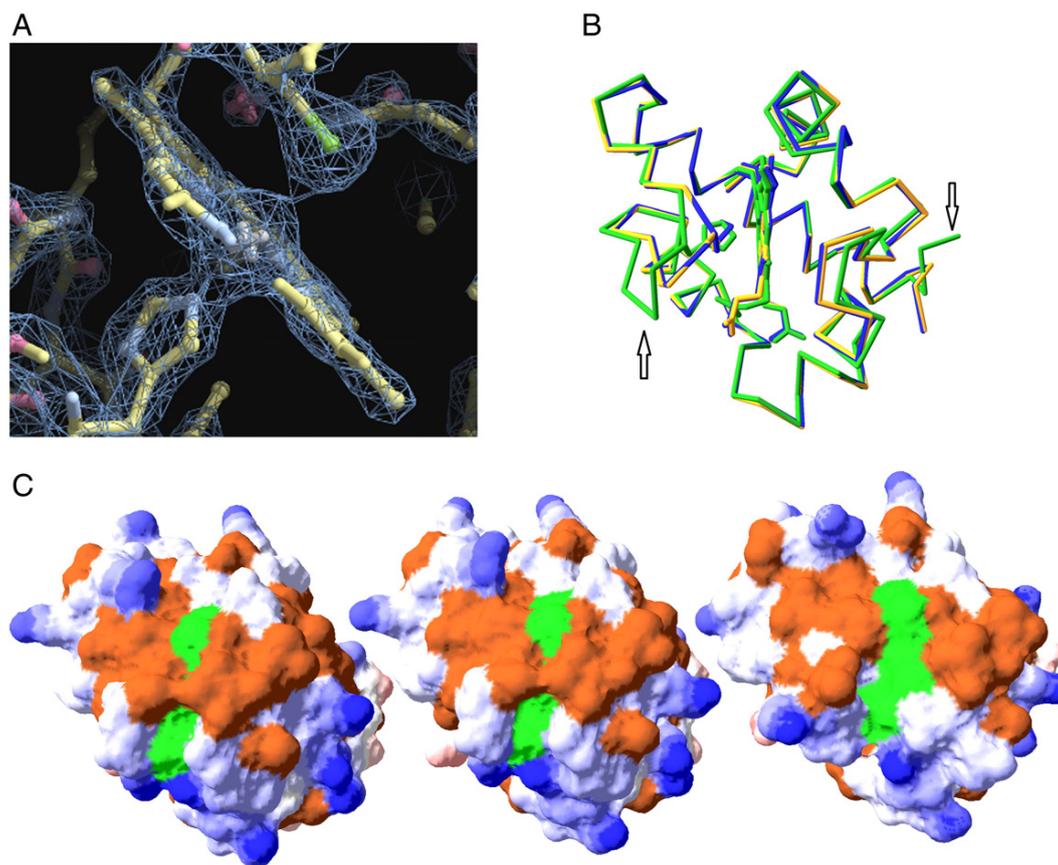


Fig. 1. Comparison of Cc6 structures. A) Electron density contoured at 2.2 sigma with the final M58C Cc6 model in the region around the haem iron. This figure was created with COOT [1]. B) Overlay of C α traces of wt (blue), M58C (yellow) and M58H (green). Residue 58 and the haem are shown in sticks. The arrows indicate the large differences between M58H and wt Cc6. C) Surface representations of M58C (left), wt (middle) and M58H (right). The surface is coloured according to the surface potential from red to blue (negative to positive, calculated with DeepView (<http://www.genebee.msu.su/spdbv>)). Non-polar residues are in brown and the haem is in green.

orientation. Surprisingly, also Trp85 and the C-terminal residue Lys86, on the other side of the protein, far from the His58, show a large displacement compared to the wt structure.

It was decided to use M58C Cc6 as a redox inactive substitute for the wt protein, because of its low midpoint potential and structural similarity. M58C Cc6 was enriched in ^{15}N and the amide nuclei were assigned on the basis of NOESY- ^{15}N HSQC (150 ms mixing time) and TOCSY- ^{15}N HSQC spectra with 80 ms mixing. For the M58C Cc6 variant, no self-reduction was observed, as expected from its low midpoint potential. The backbone amide assignments are reported in Table S1 in the supplementary information. Those corresponding to the two loops surrounding the haem edge are partly missing or tentative. Sequential connections were difficult to make in these regions, most likely due to some dynamics on the micro-millisecond timescale.

3.2. Dissociation constant and chemical shift map

To determine the affinity between M58C Cc6 and Cf a titration was performed. Ferrous Cf was titrated into M58C Cc6 and HSQC spectra were recorded at every point. Fig. 2, panel A shows the chemical-shift perturbations of several Cc6 amide nuclei plotted as a function of the Cf–Cc6 ratio. The curves can be fitted with a 1:1 binding model, yielding a binding constant of $7(2) \text{ mM}^{-1}$, identical within error to that for the wt complex, $8(2) \text{ mM}^{-1}$ [34]. Panel B in Fig. 2 shows the binding map, in which the surface of the protein has been coloured according to the size of the average amide chemical-shift perturbation for each residue, extrapolated to the 100% bound state of Cc6 ($\Delta\delta_{\text{Avg}}$). The map is similar to the one

reported for wt Cc6 binding to *Nostoc* [34] as well as *P. laminosum* Cf [33] and shows that the complex uses mainly the haem edge region for binding, although some residues at the other sides of the protein are affected, most notably Glu 68. The overall size of the shifts and this distribution of residues on the surface suggest that Cc6 mostly binds with one side toward Cf, in a relatively well-defined orientation [64]. Whether it binds on a single site of Cf cannot be established on the basis of these data.

3.3. Paramagnetic relaxation enhancements

To obtain intermolecular PREs for docking and structure determination of the complex, five mutants of Cf were produced in which cysteine residues were introduced on the surface of Cf in the region surrounding the haem. To avoid changes in the pI, which could affect the protein–protein interactions, only neutral amino acid residues were selected, Gln7, Ala63, Asn71, Gln104 and Ser192. Either the paramagnetic spin label MTSL or the non-paramagnetic analogue MTS was linked to the Cys residues. By measuring the intensity ratio of Cc6 M58C amide resonances in spectra recorded on samples with MTSL–Cf and MTS–Cf, the PRE was determined [56]. All the spin labels have large effects on Cc6 signals (Fig. S1). Interestingly, the effects are all found in the same region of Cc6, the loops centred around residues 20 and 55, which is also the area exhibiting the largest chemical-shift perturbations (Fig. 2). This observation suggests that Cc6 is always oriented toward Cf with this surface patch, comprising the region where the haem penetrates the surface, the haem edge. It also implies that Cc6 samples a significant surface area of Cf with this patch, because it is affected by the

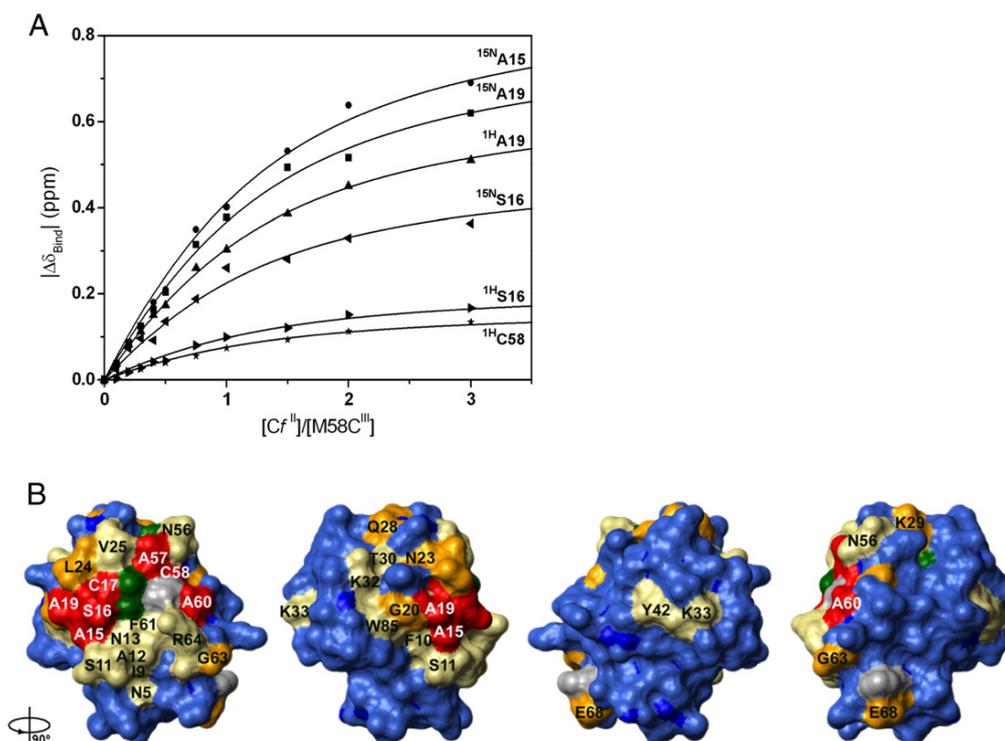


Fig. 2. Binding of M58C Cc6 to Cf. A) Binding curves for the interaction of M58C mutant with ferrous Cf. The data were fitted globally to a 1:1 model (non-linear, least-squares), yielding an association constant of $7 (\pm 2) 10^3 \text{ M}^{-1}$. B) Chemical shift perturbation map. Residues for which a $\Delta\delta_{\text{avg}}$ (ppm) was calculated are colour-coded on the structure of M58C mutant according to the following categories: blue for <0.025 ppm, yellow for ≥ 0.025 ppm, orange for ≥ 0.050 ppm, red for ≥ 0.125 ppm. Prolines are shown in grey and the haem group in dark green. Residues are identified with the single-letter amino acid code, and the surfaces have been rotated in anti-clockwise 90° steps around the vertical axis, with respect to the one on the left. Surface representations were generated using MOLMOL [2].

spin labels at every position. If it would assume a single orientation on Cf, located between the spin label positions, PREs from each spin label would affect a different side of Cc6.

3.4. Docking Cc6 in a single orientation

The PREs were converted into distances and docking calculations were performed. The spin label dynamics were represented by using four orientations per spin label [58] and the spin labels were allowed to rotate during the docking to avoid steric clashes with Cc6. The proteins were treated as rigid bodies. It turned out that Cc6 docks so closely to the spin labels (see below) that tiny distance variations led to very large changes in PRE, a consequence of dependence of the PRE on the sixth power of the distance between nucleus and spin label. For this reason it was preferable to use the PRE-based distances and not the PREs themselves as restraints in the docking calculations and evaluation of the quality of the fit with the observed data.

First, a simple model of Cc6 and Cf interaction was tested, by assuming that the PREs represent only a single, well-defined orientation of the proteins within the complex. The resultant structure that best fits the experimental data is physically unrealistic, because Cc6 does not make contact with Cf, but rather remains at a distance of several Ångström from the surface. The degree to which the structure represents the experimental data was expressed as the average distance violation, which represents the difference between the experimental and the back-calculated distances for all residues and all spin labels (the definition is given in the Materials and methods section). Thus, the larger the average violation, the poorer the quality of the fit. Fig. 3A shows that the average distance violation for the simple model, with a fraction of a well-defined structure of 100% ($p = 0$), is 2.0 Å. Figs. 4A and S2 plot the experimental and back-calculated distances for each residue. Several regions of Cc6 experience more paramagnetic effects than predicted by this model and

these residues are expected to be closer to the spin label than is found in the model. It is clear that the PRE data cannot be described by Cc6 in a single orientation within the complex with Cf.

3.5. Ensemble docking

Ensemble docking was used to obtain a better fit to the experimental data. This approach assumes that the complex exists in more than a

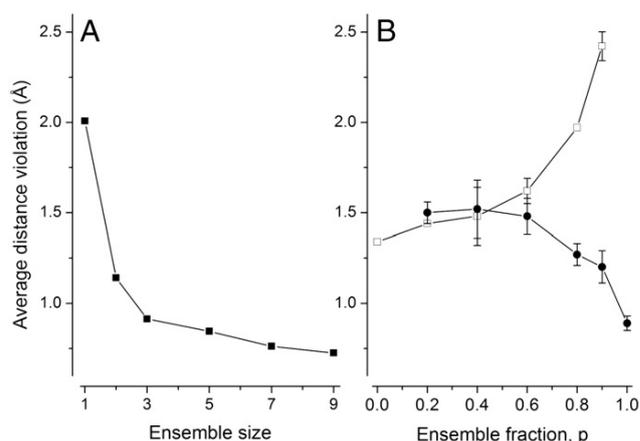


Fig. 3. Average violation plots. A) The average distance violation plotted for increasing ensemble size. B) The average distance violation is plotted for different ensemble fractions (p). The results for the average violation for the single structure only (open symbols) and for the combination of single structure and ensemble (solid symbols) are shown. For $p = 1$, the complex consists entirely of an ensemble. The ensemble size used in these calculations was five. The definition of the error bars is given in the Materials and methods section.

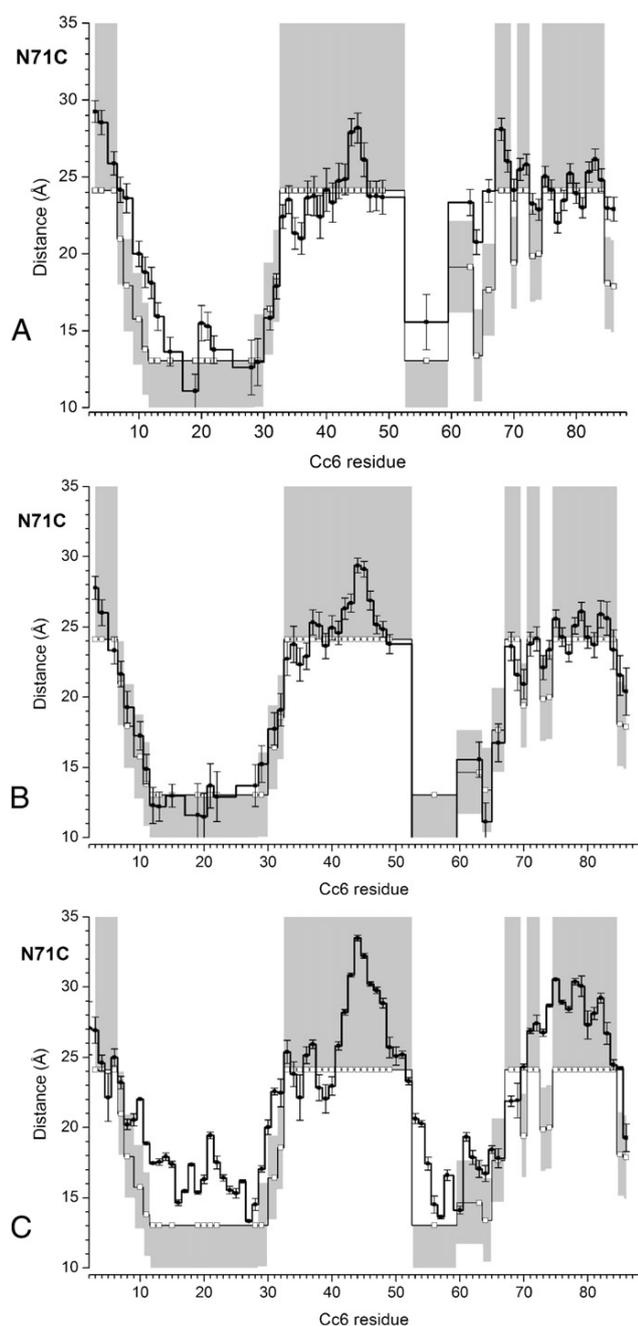


Fig. 4. Fitting of the experimental distances. The average distances between the oxygen atom of spin label N71C and Cc6 amide protons were back-calculated from various ensembles (thick line and solid symbols), plotted against the residue number and compared with the experimental distances (open symbols). The grey areas indicate the experimental error margins. In (A, B) the solid line represents the averages for the ten best structures obtained from docking a single copy (A), or ten ensembles of five copies (B) of Cc6. In (C), it represents the averages from three sets of 10,000 randomly selected structures from the MC ensemble. The error bars indicate the SD of the ensembles.

single orientation and can be represented by an ensemble of orientations [65]. Several copies of Cc6 were docked simultaneously, driven by the PRE-derived distances. The average distance between the spin label and the set of identical nuclei in the ensemble was compared with the experimental distance and minimized, resulting in an ensemble of orientations of Cc6 around Cf. This procedure was repeated many times, creating a cloud of orientations. The quality of the fit is

expressed as the average violation of the experimental distances. The size of the ensemble, which is the number of copies of Cc6 used in the docking, was varied, demonstrating that the fit improvement by adding more copies levels off quickly (Fig. 3A). In the subsequent calculations an ensemble size of five Cc6 copies was used. Adding more copies does not significantly improve the fits and results in superfluous copies of Cc6 that are placed far from the surface of Cf.

The ensemble docking can be executed assuming that all the PREs derive from this ensemble (100% ensemble, $p = 1$) or that a two-state model is applicable, with a dynamic encounter state and a well-defined state ($0 < p < 1$). In the latter case, a single Cc6 molecule is docked first, assuming a certain population, for example 40%, and then the back-calculated PRE effects for this structure are subtracted from the observed PREs and the remaining PREs are input for the subsequent ensemble docking to obtain the encounter complex that represents the remaining 60% of the complex [66]. The back-calculated distances from the combined single structure + ensemble are compared with the experimental distances to evaluate the quality of the fit. Fig. 3B shows the results for such calculations. The average distance violation is plotted as a function of the fraction of encounter complex (p), for the single structures only (open symbols) and for the single structure + ensemble, representing the entire complex (solid symbols). As expected, ensemble docking yields lower average violations, because more degrees of freedom are added in such calculations. Interestingly, the fit does not improve for p -values up to 0.6. That means that a large fraction of the PRE is attributable to the encounter complex. The best results are found for $p = 1.0$, so in the absence of a single, well-defined orientation of Cc6. These results suggest that a single dominant orientation is not present and that the complex of Cf and Cc6 is best described by an ensemble of orientations.

The average distances back-predicted from the ensemble ($p = 1.0$) fit the experimental distance much better than the single structure (Figs. 4B and S3), indicating that the ensemble is an acceptable representation of the PRE data.

3.6. Monte Carlo docking calculations

We then wondered whether the ensemble is of purely electrostatic nature. Before, it had been shown that the encounter complex of cytochrome *c* peroxidase and cytochrome *c* could be described by a theoretical ensemble obtained via Monte Carlo (MC) calculations that only considered electrostatic interactions between the proteins [61]. Similar calculations were performed for the Cf–Cc6 complex and a large ensemble was created. From this ensemble the average distances from the Cc6 amide protons and the spin labels were calculated and compared with the experimental distances (Figs. 4C and S4). The distance patterns roughly follow the experimental ones, but it is clear that the MC ensemble cannot describe the experimental data very well. In most of the MC orientations Cc6 is oriented with its haem edge face oriented toward the Cf surface, in line with the conclusion from the experimental PRE patterns. Thus, it can be concluded that Cc6 pre-orientates this face toward Cf upon its approach, due to electrostatic interactions. However, the poor quantitative match with the experimental distances indicates that the places where Cc6 is located on the Cf surface in the MC ensemble do not agree well with the real complex, suggesting that within the encounter complex, other interactions than electrostatics contribute significantly.

3.7. Analysis of the ensemble

The centres-of-mass of the 610 Cc6 molecules obtained from 122 runs of ensemble docking show that Cc6 visits an area including and surrounding the hydrophobic surface patch near the haem of Cf (Fig. 5, green). Clearly, a single structure cannot describe this ensemble. It is possible that the Cc6 samples an even larger area, because the five spin labels did not cover the whole surface of Cf, so in the calculations

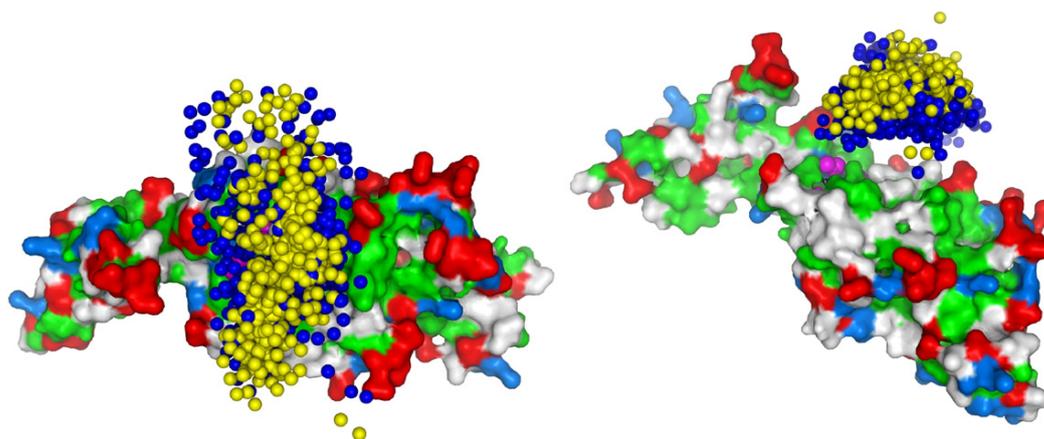


Fig. 5. The ensemble of Cc6 around Cf shown in two orientations. Cf is shown in surface representation with the haem in magenta spheres. Negative, positive and hydrophobic residues (including Tyr and Trp) of Cf are displayed in red, marine and green, respectively. The yellow and dark blue spheres represent the centres-of-mass and iron atoms of Cc6 in the ensemble, respectively.

there are no restraints to guide Cc6 to those regions. The charge distribution on Cf shows that it has an overall negative charge, with strong negative potential around the hydrophobic patch. Cc6 has a ring of positive charges around its hydrophobic patch, close to the haem edge (Fig. 1C). Cc6 is always oriented with this region toward Cf, as can be seen in Fig. 5. The blue spheres represent the iron atoms of Cc6 in the ensemble. These atoms are always closer to Cf than the centres-of-mass (yellow spheres), indicating that Cc6 has a preferred binding orientation. Most of the Cc6 molecules are found to interact with the hydrophobic patch of Cf, not with the charged regions, in accord with the finding that the MC calculations, which are based purely on charge-charge interactions, do not produce an ensemble that fits the experimental data well.

4. Discussion

Characterization of the complex of Cf and Cc6 was hindered by self-reduction of Cc6 during NMR experiments. A solution was found by replacement of the methionine ligand of the haem iron. Substitution with a His or a Cys residue resulted in a large decrease of the midpoint potential. The crystal structures of the *Nostoc* wt Cc6 and both variants were determined. In particular the Cys mutant is remarkably similar in structure to the wt protein, despite the shorter side chain length of Cys. Mutant M58C was used in the interactions studies with Cf, because its low midpoint potential abolished the problem of self-reduction. Substitution of haem ligands by Cys has been reported for haem enzymes to study the effects on enzyme activity ([67] and references therein) and the spectroscopic characterization of semi-synthetic Cc with a Cys replacing the ligand Met has been described [68,69]. Very recently, an extensive study of yeast Cc with a co-ordinating Cys was published. Replacement of the Thr at position 78 by Cys yielded a protein that was more stable than Met80Cys Cc [70]. It is difficult to produce redox-inactive cytochrome *c* analogues by metal substitution, because removal of the haem iron can only be achieved under harsh conditions, requiring denaturation and refolding of the protein. The substitution of Met co-ordination by that with a Cys thiolate can be used as a convenient alternative. In each of the described cases the midpoint potentials dropped by many hundreds of mV.

The PRE data indicate that the complex is not dominated by a single well-defined orientation, but instead it is best described by an ensemble of orientations. The chemical-shift perturbation results as well as the PRE data indicate that Cc6 pre-oriens upon approaching Cf. It is expected that the long-range electrostatic attraction causes Cc6 to orient with its positive charges toward the overall negative Cf. Given the dipolar nature of Cc6, this movement results in Cc6 facing Cf with its hydrophobic

patch located around the haem edge. It has been known for a long time that the haem edge provides strong coupling for electron transfer, so the pre-orientation primes Cc6 for rapid electron transfer from the Cf haem. At a short distance, hydrophobic interactions appear to be important, because the ensemble that fits the PRE data is located mostly above the hydrophobic region around the Cf haem group. Contrary to the encounter complex of the electron transfer complex of cytochrome *c* and cytochrome *c* peroxidase, which can be described with electrostatic interactions only [57,61], the Cf–Cc6 complex also involves hydrophobic interactions. These findings are in line with recent studies on the Cf–Pc complex from the same *Nostoc* species [40,41]. Also in that case the complex is at least partly in an encounter state and electrostatic MC calculations cannot fully describe the PRE data. In this complex hydrophobic interactions appear to play a similarly important role, which became clear from a recent comparative study on Cf–Pc complexes from *Nostoc* and *P. laminosum* [71]. Other studies on Cc6 from *Nostoc* and *C. reinhardtii*, using NMR, kinetic measurements and docking calculations also indicated the importance of hydrophobic interactions with its partners, Cf [31–33], photosystem I [26,29] and cytochrome *c* oxidase [72]. The data further suggest that the Cf–Cc6 complex consists predominantly or entirely of an ensemble of orientations, whereas in the Cf–Pc complex a stereospecific complex is also present for a significant amount of the time.

These studies raise the question whether the two state model of an electrostatic encounter complex and a well-defined complex, with specific hydrophobic interactions, hydrogen bonding and van der Waals forces as well as electrostatic interactions, is generally applicable to electron transfer complexes. At least in the *Nostoc* case, of Cf reacting with Cc6 and Pc, the complex is better described by electrostatic pre-orientation when the proteins are still approaching and an ensemble of orientations in which the proteins exhibit a form of hydrophobic sliding, with increasing desolvation of the hydrophobic patches and thus a gradual transition to the most stably bound orientation (Cf–Pc) or orientations (Cf–Cc6) [40]. The model is illustrated in Fig. 6. Such a hydrophobic search mechanism has been suggested on the basis of theoretical studies by Camacho and co-workers [73,74]. Although the encounter complex is normally considered to be dominated by electrostatic interactions [3,75,76], the involvement of hydrophobic contacts in the encounter state has been reported before [40,66,71,77].

In the ensemble some orientations will exhibit optimal coupling between the redox centres, resulting in rapid electron transfer. As long as the distance between the haems is short and the space jump between the proteins is small, electron transfer will be fast. A single, well-defined orientation is not required in that case. In fact, it is not desirable, because a well-defined complex needs to be stabilized by multiple

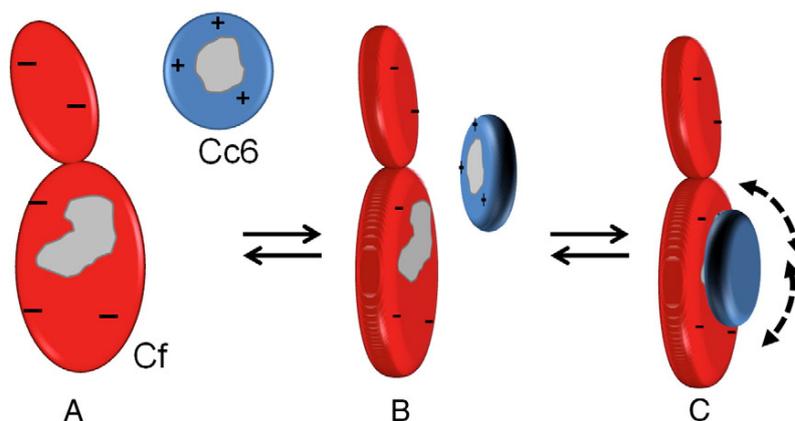


Fig. 6. Model for complex formation of Cf and Cc6. A) Free proteins approach via diffusion. B) Electrostatic pre-orientation during approach. C) Cc6 slides over the surface of Cf optimizing hydrophobic contacts (grey areas). Some of the positions in this state are compatible with fast electron transfer.

interactions, making it too stable for rapid dissociation. The view that in some complexes multiple orientations, part of a larger ensemble, are suited for ET, is similar to the dynamic docking model described for myoglobin and cytochrome b_5 , although that complex appears to be dominated by charge–charge interactions [78,79]. It is supported by a recently published kinetic study demonstrating that multiple conformations of Cc6 contribute to electron transfer within the Cf–Cc6 complex [31]. In the described model of gradual desolvation, the reverse process happens upon dissociation. Gradual resolution avoids a high transition state barrier between the electron transfer active state and the free proteins, thus ensuring rapid dissociation. For photosynthetic electron transfer proteins, a high off-rate is as important as fast electron transfer to avoid product inhibition and reduction of the electron flow rate in the redox chain [80]. The relatively low affinity between Cf and its partners is in line with the idea that rapid turn-over is important. In the thylakoids, Cf is present in a tilted orientation, with the side shown in Fig. 5 facing the lumen. The other side is close to the membrane and not accessible to Cc6 and Pc. Given the considerable confinement within the lumen, it is expected that the local concentrations are high and affinity is not a stringent requirement for complex formation.

5. Conclusions

The findings on the complexes of Cf with Pc and Cc6 show that encounter complexes represent an important part of the photosynthetic ET protein complexes. Experimental evidence for a role of hydrophobic interactions in the encounter complex is accumulating, blurring the distinction between encounter complex and stereospecific complex. This is true in particular for ET complexes, because a distance between the metal ions that is sufficiently short for rapid ET is all that is required for activity, so there is no reason for the presence of a single active orientation within the complex.

Acknowledgements

We thank Daniel de Geus, Ellen Thomassen and Irakli Sikharulidze for assistance in the X-ray diffraction data collection. Financial support was provided by the Netherlands Organisation for Scientific Research (NWO), through VIDI Grant Nos. 700.52.425 (RH, MU) and 700.55.425 (NSP), VICI Grant No. 700.58.441 (MU), and Open Competition Grant No. 700.50.026 (MF), by the Spanish Ministry of Economy and Competitiveness (Grant Nos. BFU2003-00458/BMC, BFU2006-01361/BMC, BFU2009-07190/BMC and BFU2012-31670/BMC) and by the Andalusian Government (Grant PAI, BIO198). IDM was supported by the Human Potential Programme and the Mobility of Researchers Programme of the European Commission (Contract No. HPRN-CT-1999-

00095) and by the Ministry of Education, Culture and Sports, Spain (Grant No. AP2000-2937). BMB was awarded with a PhD fellowship (AP2009-4092) from the Ministry of Education, Culture and Sports, Spain, co-funded by the European Social Fund, ERDF (2007–2013). DC and GLR were supported by the Italian Ministry of University and Research, MIUR (Grant PRIN 20074TJ3ZB_004). JMF and GMU were supported by Grant DFG RTG 1640 of the Deutsche Forschungsgemeinschaft.

Appendix A. Supplementary data

Supplementary data include a table with assignments of Cc6 amides, intensity ratio plots (I_{para}/I_{dia}), distance plots for single structure and ensemble docking and Monte-Carlo simulations, an example of an ensemble docking script for XPLOR and an input file with PRE restraints.

Deposits

PDB IDs: Cc6 wt (4GYD), M58C (4H0J) and M58H (4H0K). Supplementary data to this article can be found online at DOI: <http://dx.doi.org/10.1016/j.bbabi.2014.03.009>.

References

- [1] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, *Acta Crystallogr. D* 66 (2010) 486–501.
- [2] R. Koradi, M. Billeter, K. Wuthrich, MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graph.* 14 (1996) 51–55.
- [3] M. Ubbink, The courtship of proteins: understanding the encounter complex, *FEBS Lett.* 583 (2009) 1060–1066.
- [4] C.C. Moser, J.M. Keske, K. Warncke, R.S. Farid, P.L. Dutton, Nature of biological electron transfer, *Nature* 355 (1992) 796–802.
- [5] A. Diaz-Quintana, J.A. Navarro, M. Hervás, F.P. Molina-Heredia, B. De la Cerda, M.A. De la Rosa, A comparative structural and functional analysis of cyanobacterial plastocyanin and cytochrome c(6) as alternative electron donors to Photosystem I – Photosystem I reduction in cyanobacteria, *Photosynth. Res.* 75 (2003) 97–110.
- [6] B. De la Cerda, J.A. Navarro, M. Hervás, M.A. De la Rosa, Changes in the reaction mechanism of electron transfer from plastocyanin to photosystem I in the cyanobacterium *Synechocystis* sp. PCC 6803 as induced by site-directed mutagenesis of the copper protein, *Biochemistry* 36 (1997) 10125–10130.
- [7] J. Sun, W. Xu, M. Hervás, J.K. Navarro, M.A. De la Rosa, P.R. Chitnis, Oxidizing side of the cyanobacterial photosystem I – evidence for interaction between the electron donor proteins and a luminal surface helix of the Psab subunit, *J. Biol. Chem.* 274 (1999) 19048–19054.
- [8] G.M. Ullmann, M. Hauswald, A. Jensen, N.M. Kostic, E.W. Knapp, Comparison of the physiologically equivalent proteins cytochrome c6 and plastocyanin on the basis of their electrostatic potentials. Tryptophan 63 in cytochrome c6 may be isofunctional with tyrosine 83 in plastocyanin, *Biochemistry* 36 (1997) 16187–16196.
- [9] R.V. Duran, M. Hervás, M.A. De la Rosa, J.A. Navarro, The efficient functioning of photosynthesis and respiration in *Synechocystis* sp PCC 6803 strictly requires the presence of either cytochrome c6 or plastocyanin, *J. Biol. Chem.* 279 (2004) 7229–7233.
- [10] M.A. De la Rosa, J.A. Navarro, A. Diaz-Quintana, B. De la Cerda, F.P. Molina-Heredia, A. Balme, P.D. Murdoch, I. Diaz-Moreno, R.V. Duran, M. Hervás, An evolutionary analysis of the reaction mechanisms of photosystem I reduction by cytochrome c(6) and plastocyanin, *Bioelectrochemistry* 55 (2002) 41–45.
- [11] R. Gupta, Z.Y. He, S. Luan, Functional relationship of cytochrome c(6) and plastocyanin in *Arabidopsis*, *Nature* 417 (2002) 567–571.

- [12] F.P. Molina-Heredia, J. Wastl, J.A. Navarro, D.S. Bendall, M. Hervás, C.J. Howe, M.A. De la Rosa, A new function for an old cytochrome? *Nature* 424 (2003) 33–34.
- [13] D. Stroebel, Y. Choquet, J.L. Popot, D. Picot, An atypical haem in the cytochrome b(6)f complex, *Nature* 426 (2003) 413–418.
- [14] G. Kurisu, H.M. Zhang, J.L. Smith, W.A. Cramer, Structure of the cytochrome *b₆f* complex of oxygenic photosynthesis: tuning the cavity, *Science* 302 (2003) 1009–1014.
- [15] D. Baniulis, E. Yamashita, J.P. Whitelegge, A.I. Zatsman, M.P. Hendrich, S.S. Hasan, C. M. Ryan, W.A. Cramer, Structure–function, stability, and chemical modification of the cyanobacterial cytochrome b(6)f complex from *Nostoc* sp. PCC 7120, *J. Biol. Chem.* 284 (2009) 9861–9869.
- [16] S.E. Martinez, D. Huang, A. Szczepaniak, W.A. Cramer, J.L. Smith, Crystal-structure of chloroplast cytochrome *f* reveals a novel cytochrome fold and unexpected heme ligation, *Structure* 2 (1994) 95–105.
- [17] I. Díaz-Moreno, S. Diaz-Moreno, G. Subias, M.A. De la Rosa, A. Diaz-Quintana, The atypical iron-coordination geometry of cytochrome *f* remains unchanged upon binding to plastocyanin, as inferred by XAS, *Photosynth. Res.* 90 (2006) 23–28.
- [18] M. Beissinger, H. Sticht, M. Sutter, A. Ejchart, W. Haehnel, P. Rosch, Solution structure of cytochrome c(6) from the thermophilic cyanobacterium *Synechococcus elongatus*, *EMBO J.* 17 (1998) 27–36.
- [19] M.R. Sawaya, D.W. Krogmann, A. Serag, K.K. Ho, T.O. Yeates, C.A. Kerfeld, Structures of cytochrome *c-549* and cytochrome c(6) from the cyanobacterium *Arthrospira maxima*, *Biochemistry* 40 (2001) 9215–9225.
- [20] J.A.R. Worrall, B.G. Schlarb-Ridley, T. Reda, M.J. Marcaida, R.J. Moerland, J. Wastl, J. Hirst, D.S. Bendall, B.F. Luisi, C.J. Howe, Modulation of heme redox potential in the cytochrome c(6) family, *J. Am. Chem. Soc.* 129 (2007) 9468–9475.
- [21] W. Bialek, S. Krzywdka, M. Jaskolski, A. Szczepaniak, Atomic-resolution structure of reduced cyanobacterial cytochrome c(6) with an unusual sequence insertion, *FEBS J.* 276 (2009) 4426–4436.
- [22] L. Banci, I. Bertini, M.A. De la Rosa, D. Koulougliotis, J.A. Navarro, O. Walter, Solution structure of oxidized cytochrome c(6) from the green alga *Monoraphidium braunii*, *Biochemistry* 37 (1998) 4831–4843.
- [23] M.J. Marcaida, B.G. Schlarb-Ridley, J.A.R. Worrall, J. Wastl, T.J. Evans, D.S. Bendall, B.F. Luisi, C.J. Howe, Structure of cytochrome c(6A), a novel dithio-cytochrome of *Arabidopsis thaliana*, and its reactivity with plastocyanin: implications for function, *J. Mol. Biol.* 360 (2006) 968–977.
- [24] B.S. Rajagopal, M.T. Wilson, D.S. Bendall, C.J. Howe, J.A.R. Worrall, Structural and kinetic studies of imidazole binding to two members of the cytochrome c(6) family reveal an important role for a conserved heme pocket residue, *J. Biol. Inorg. Chem.* 16 (2011) 577–588.
- [25] A. Kranich, H. Naumann, F.P. Molina-Heredia, H.J. Moore, T.R. Lee, S. Lecomte, M.A. De la Rosa, P. Hildebrandt, D.H. Murgida, Gated electron transfer of cytochrome c(6) at biomimetic interfaces: a time-resolved SERR study, *Phys. Chem. Chem. Phys.* 11 (2009) 7390–7397.
- [26] I. Díaz-Moreno, A. Diaz-Quintana, F.P. Molina-Heredia, P.M. Nieto, O. Hansson, M.A. De la Rosa, B.G. Karlsson, NMR analysis of the transient complex between membrane photosystem I and soluble cytochrome c(6), *J. Biol. Chem.* 280 (2005) 7925–7931.
- [27] I. Díaz-Moreno, A. Diaz-Quintana, G. Subias, T. Mairs, M.A. De la Rosa, S. Diaz-Moreno, Detecting transient protein–protein interactions by X-ray absorption spectroscopy: the cytochrome c(6)–photosystem I complex, *FEBS Lett.* 580 (2006) 3023–3028.
- [28] M. Hervás, J.A. Navarro, M.A. De la Rosa, Electron transfer between membrane complexes and soluble proteins in photosynthesis, *Acc. Chem. Res.* 36 (2003) 798–805.
- [29] F.P. Molina-Heredia, A. Diaz-Quintana, M. Hervás, J.A. Navarro, M.A. De la Rosa, Site-directed mutagenesis of cytochrome c(6) from *Anabaena* species PCC 7119 – identification of surface residues of the heme protein involved in photosystem I reduction, *J. Biol. Chem.* 274 (1999) 33565–33570.
- [30] F.P. Molina-Heredia, M. Hervás, J.A. Navarro, M.A. De la Rosa, A single arginyl residue in plastocyanin and in cytochrome c(6) from the cyanobacterium *Anabaena* sp. PCC 7119 is required for efficient reduction of photosystem I, *J. Biol. Chem.* 276 (2001) 601–605.
- [31] T.Z. Grove, G.M. Ullmann, N.M. Kostic, Simultaneous true, gated, and coupled electron-transfer reactions and energetics of protein rearrangement, *J. Inorg. Biochem.* 106 (2012) 143–150.
- [32] T.Z. Grove, N.M. Kostic, Metalloprotein association, self-association, and dynamics governed by hydrophobic interactions: simultaneous occurrence of gated and true electron-transfer reactions between cytochrome *f* and cytochrome *c₆* from *Chlamydomonas reinhardtii*, *J. Am. Chem. Soc.* 125 (2003) 10598–10607.
- [33] P.B. Crowley, A. Diaz-Quintana, F.P. Molina-Heredia, P. Nieto, M. Sutter, W. Haehnel, M.A. De la Rosa, M. Ubbink, The interactions of cyanobacterial cytochrome *c₆* and cytochrome *f*, characterized by NMR, *J. Biol. Chem.* 277 (2002) 48685–48689.
- [34] I. Díaz-Moreno, A. Diaz-Quintana, M. Ubbink, M.A. De la Rosa, An NMR-based docking model for the physiological transient complex between cytochrome *f* and cytochrome c(6), *FEBS Lett.* 579 (2005) 2891–2896.
- [35] E.L. Gross, D.C. Pearson, Brownian dynamics. Simulations of the interaction of *Chlamydomonas* cytochrome *f* with plastocyanin and cytochrome *c₆*, *Biophys. J.* 85 (2003) 2055–2068.
- [36] E.J. Haddadian, E.L. Gross, Brownian dynamics study of cytochrome *f* interactions with cytochrome c(6) and plastocyanin in *Chlamydomonas reinhardtii* plastocyanin, and cytochrome c(6) mutants, *Biophys. J.* 88 (2005) 2323–2339.
- [37] E.J. Haddadian, E.L. Gross, A Brownian dynamics study of the effects of cytochrome *f* structure and deletion of its small domain in interactions with cytochrome c(6) and plastocyanin in *Chlamydomonas reinhardtii*, *Biophys. J.* 90 (2006) 566–577.
- [38] F.P. Molina-Heredia, M. Hervás, J.A. Navarro, M.A. De la Rosa, Cloning and correct expression in *Escherichia coli* of the *petE* and *petJ* genes respectively encoding plastocyanin and cytochrome *c₆* from the cyanobacterium *Anabaena* sp. PCC 7119, *Biochem. Biophys. Res. Commun.* 243 (1998) 302–306.
- [39] C. Albarran, J.A. Navarro, F.P. Molina-Heredia, P.S. Murdoch, M.A. De la Rosa, M. Hervás, Laser flash-induced kinetic analysis of cytochrome *f* oxidation by wild-type and mutant plastocyanin from the cyanobacterium *Nostoc* sp. PCC 7119, *Biochemistry* 44 (2005) 11601–11607.
- [40] S. Scanu, J.M. Foerster, G.M. Ullmann, M. Ubbink, Role of hydrophobic interactions in the encounter complex formation of the plastocyanin and cytochrome *f* complex revealed by paramagnetic NMR spectroscopy, *J. Am. Chem. Soc.* 135 (2013) 7681–7692.
- [41] S. Scanu, J. Förster, M.G. Finiguerra, M.H. Shabestari, M. Huber, M. Ubbink, The complex of cytochrome *f* and plastocyanin from *Nostoc* sp. PCC 7119 is highly dynamic, *ChemBioChem* 13 (2012) 1312–1318.
- [42] E. Arslan, H. Schulz, R. Zufferey, P. Kunzler, L. Thony-Meyer, Overproduction of the *Bradyrhizobium japonicum* c-type cytochrome subunits of the *cbb(3)* oxidase in *Escherichia coli*, *Biochem. Biophys. Res. Commun.* 251 (1998) 744–747.
- [43] A.G.W. Leslie, Integration of macromolecular diffraction data, *Acta Crystallogr. D* 55 (1999) 1696–1702.
- [44] P. Evans, Scaling and assessment of data quality, *Acta Crystallogr. D* 62 (2006) 72–82.
- [45] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, *Acta Crystallogr. D* 67 (2011) 235–242.
- [46] N.S. Pannu, W.J. Waterreus, P. Skubak, I. Sikkharulidze, J.P. Abrahams, R.A.G. de Graaff, Recent advances in the CRANK software suite for experimental phasing, *Acta Crystallogr. D* 67 (2011) 331–337.
- [47] R.A.G. de Graaff, M. Hilge, J.L. van der Plas, J.P. Abrahams, Matrix methods for solving protein substructures of chlorine and sulfur from anomalous data, *Acta Crystallogr. D* 57 (2001) 1857–1862.
- [48] K. Cowtan, Error estimation and bias correction in phase-improvement calculations, *Acta Crystallogr. D* 55 (1999) 1555–1567.
- [49] A. Perrakis, R. Morris, V.S. Lamzin, Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol.* 6 (1999) 458–463.
- [50] P. Skubak, G.N. Murshudov, N.S. Pannu, Direct incorporation of experimental phase information in model refinement, *Acta Crystallogr. D* 60 (2004) 2196–2201.
- [51] G.N. Murshudov, P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr. D* 67 (2011) 355–367.
- [52] D.E. McRee, XtalView Xfit – a versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol.* 125 (1999) 156–165.
- [53] P.J. Kraulis, Ansig – a program for the assignment of protein H-1 2D-NMR spectra by interactive computer-graphics, *J. Magn. Reson.* 84 (1989) 627–633.
- [54] M. Helgstrand, P. Kraulis, P. Allard, T. Hard, Ansig for Windows: an interactive computer program for semiautomatic assignment of protein NMR spectra, *J. Biomol. NMR* 18 (2000) 329–336.
- [55] A. Kannt, S. Young, D.S. Bendall, The role of acidic residues of plastocyanin in its interaction with cytochrome *f*, *Biochim. Biophys. Acta Bioenerg.* 1277 (1996) 115–126.
- [56] J.L. Battiste, G. Wagner, Utilization of site-directed spin labeling and high-resolution heteronuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data, *Biochemistry* 39 (2000) 5355–5365.
- [57] A.N. Volkov, J.A.R. Worrall, E. Holtzmann, M. Ubbink, Solution structure and dynamics of the complex between cytochrome *c* and cytochrome *c* peroxidase determined by paramagnetic NMR, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 18945–18950.
- [58] J. Iwahara, C.D. Schwieters, G.M. Clore, Ensemble approach for NMR structure refinement against H-1 paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule, *J. Am. Chem. Soc.* 126 (2004) 5879–5896.
- [59] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, G.M. Clore, The Xplor-NMR molecular structure determination package, *J. Magn. Reson.* 160 (2003) 65–73.
- [60] G.M. Ullmann, E.W. Knapp, N.M. Kostic, Computational simulation and analysis of dynamic association between plastocyanin and cytochrome *f*. Consequences for the electron-transfer reaction, *J. Am. Chem. Soc.* 119 (1997) 42–52.
- [61] Q. Bashir, A.N. Volkov, G.M. Ullmann, M. Ubbink, Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome *c* and cytochrome *c* peroxidase, *J. Am. Chem. Soc.* 132 (2010) 241–247.
- [62] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 10037–10041.
- [63] S.J. Moench, J.D. Satterlee, A comparison of spectral and physicochemical properties of yeast iso-1 cytochrome *c* and Cys 102-modified derivatives of the protein, *J. Protein Chem.* 14 (1995) 567–582.
- [64] Q. Bashir, S. Scanu, M. Ubbink, Dynamics in electron transfer protein complexes, *FEBS J.* 278 (2011) 1391–1400.
- [65] C. Tang, J. Iwahara, G.M. Clore, Visualization of transient encounter complexes in protein–protein association, *Nature* 444 (2006) 383–386.
- [66] Y.C. Kim, C. Tang, G.M. Clore, G. Hummer, Replica exchange simulations of transient encounter complexes in protein–protein association, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 12855–12860.
- [67] S.W. Vetter, A.C. Terentis, R.L. Osborne, J.H. Dawson, D.B. Goodin, Replacement of the axial histidine heme ligand with cysteine in nitrophorin 1: spectroscopic and crystallographic characterization, *J. Biol. Inorg. Chem.* 14 (2009) 179–191.
- [68] A.L. Raphael, H.B. Gray, Semisynthesis of axial-ligand (position 80) mutants of cytochrome *c*, *J. Am. Chem. Soc.* 113 (1991) 1038–1040.

- [69] C.J. Wallace, I. Clark-Lewis, Functional role of heme ligation in cytochrome *c*. Effects of replacement of methionine 80 with natural and non-natural residues by semisynthesis, *J. Biol. Chem.* 267 (1992) 3852–3861.
- [70] F.F. Zhong, G.P. Lisi, D.P. Collins, J.H. Dawson, E.V. Pletneva, Redox-dependent stability, protonation, and reactivity of cysteine-bound heme proteins, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) E306–E315.
- [71] S. Scanu, J.M. Foerster, M. Timmer, G.M. Ullmann, M. Ubbink, Loss of electrostatic interactions causes increase of dynamics within the plastocyanin–cytochrome *f* complex, *Biochemistry* 52 (2013) 6615–6626.
- [72] S.E. Hart, C.J. Howe, K. Mizuguchi, J. Fernandez-Recio, Docking of cytochrome *c*(6) and plastocyanin to the aa(3)-type cytochrome *c* oxidase in the cyanobacterium *Phormidium laminosum*, *Protein Eng. Des. Sel.* 21 (2008) 689–698.
- [73] C.J. Camacho, S.R. Kimura, C. DeLisi, S. Vajda, Kinetics of desolvation-mediated protein–protein binding, *Biophys. J.* 78 (2000) 1094–1105.
- [74] C.J. Camacho, Z.P. Weng, S. Vajda, C. DeLisi, Free energy landscapes of encounter complexes in protein–protein association, *Biophys. J.* 76 (1999) 1166–1178.
- [75] G. Schreiber, G. Haran, H.X. Zhou, Fundamental aspects of protein–protein association kinetics, *Chem. Rev.* 109 (2009) 839–860.
- [76] J.Y. Suh, C. Tang, G.M. Clore, Role of electrostatic interactions in transient encounter complexes in protein–protein association investigated by paramagnetic relaxation enhancement, *J. Am. Chem. Soc.* 129 (2007) 12954–12955.
- [77] K. Sugase, H.J. Dyson, P.E. Wright, Mechanism of coupled folding and binding of an intrinsically disordered protein, *Nature* 447 (2007) 1021–1025.
- [78] Z.X. Liang, J.M. Nocek, K. Huang, R.T. Hayes, I.V. Kurnikov, D.N. Beratan, B.M. Hoffman, Dynamic docking and electron transfer between Zn-myoglobin and cytochrome *b₅*, *J. Am. Chem. Soc.* 124 (2002) 6849–6859.
- [79] J.A.R. Worrall, Y.J. Liu, P.B. Crowley, J.M. Nocek, B.M. Hoffman, M. Ubbink, Myoglobin and cytochrome *b₅*: a nuclear magnetic resonance study of a highly dynamic protein complex, *Biochemistry* 41 (2002) 11721–11730.
- [80] S. Kuhlert, F. Drepper, C. Fufezan, F. Sommer, M. Hippler, Residues psaB Asp612 and psaB G1u613 of Photosystem I confer pH-dependent binding of plastocyanin and cytochrome *c*(6), *Biochemistry* 51 (2012) 7297–7303.

List of Abbreviations

Cytc ₆	cytochrome <i>c</i> ₆
K _A	association constant
K _D	dissociation constant
Cc	cytochrome <i>c</i>
CcP	cytochrome <i>c</i> peroxidase
CSP	chemical shift perturbation
Cytf	cytochrome <i>f</i>
DPc	<i>Dryopteris crassirhizoma</i> plastocyanin
EPR	electron paramagnetic resonance
HSQC	heteronuclear single quantum coherence
MC	Monte Carlo
MTS	(1-acetyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)- methanethiosulfonate
MTSL	(1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)- methanethiosulfonate
NMR	nuclear magnetic resonance
Pc	plastocyanin
PCS	pseudocontact shift
PhPc	<i>Phormidium laminosum</i> plastocyanin
PoPc	<i>Populus nigra</i> plastocyanin
PQH ₂	plastoquinone
PRE	paramagnetic relaxation enhancement
PSI	Photosystem I
PSII	Photosystem II
RDC	residual dipolar coupling
TOAC	2,2,6,6-tetramethyl-N-oxyl-4-amino-4-carboxylic acid

Bibliography

- [1] W. W. Fischer, J. Hemp, and J. E. Johnson. Evolution of oxygenic photosynthesis. *Annu. Rev. Earth Planet. Sci.*, 44(1):647–683, 2016.
- [2] S. B. Hedges, J. E. Blair, M. L. Venturi, and J. L. Shoe. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol. Biol.*, 4(1):2, 2004.
- [3] R. G. Saer and R. E. Blankenship. Light harvesting in phototrophic bacteria: structure and function. *Biochem. J.*, 474(13):2107–2131, 2017.
- [4] J. Whitmarsh and Govindjee. *The Photosynthetic Process*, pages 11–51. Springer Netherlands, Dordrecht, 1999.
- [5] E. Yamashita, H. Zhang, and W. A. Cramer. Structure of the cytochrome b_6f complex: Quinone analogue inhibitors as ligands of heme c_n . *J. Mol. Biol.*, 370(1):39–52, 2007.
- [6] R. Morales, G. Kachalova, F. Vellieux, M.-H. Charon, and M. Frey. Crystallographic studies of the interaction between the ferredoxin-nadp⁺ reductase and ferredoxin from the cyanobacterium *anabaena*: looking for the elusive ferredoxin molecule. *Acta Cryst. D*, 56(11):1408–1412, 2000.
- [7] Y. Umena, K. Kawakami, J.-R. Shen, and N. Kamiya. Crystal structure of oxygen-evolving photosystem ii at a resolution of 1.9 Å. *Nature*, 473:55–60, 2011.
- [8] P. Jordan, P. Fromme, H. T. Witt, O. Klukas, W. Saenger, and N. Krauß. Three-dimensional structure of cyanobacterial photosystem i at 2.5 Å resolution. *Nature*, 411:909–917, 2001.
- [9] S. S. Hasan, E. Yamashita, and W. A. Cramer. Transmembrane signaling and assembly of the cytochrome b_6f -lipidic charge transfer complex. *Biochim. Biophys. Acta, Bioener.*, 1827(11):1295–1308, 2013.
- [10] L. Dall’Osto, M. Bressan, and R. Bassi. Biogenesis of light harvesting proteins. *Biochim. Biophys. Acta, Bioener.*, 1847(9):861–871, 2015.
- [11] M. Weigel, C. Varotto, P. Pesaresi, G. Finazzi, F. Rappaport, F. Salamini, and D. Leister. Plastocyanin Is Indispensable for Photosynthetic Electron Flow in *Arabidopsis thaliana*. *J. Biol. Chem.*, 278(33):31286–31289, 2003.
- [12] M. A. Schöttler, H. Kirchhoff, and E. Weis. The Role of Plastocyanin in the Ad-

- justment of the Photosynthetic Electron Transport to the Carbon Metabolism in Tobacco. *Plant Physiol.*, 136(4):4265–4274, 2004.
- [13] I. M. A. Nooren and J. M. Thornton. Diversity of protein–protein interactions. *EMBO J.*, 22(14):3486–3492, 2003.
- [14] E. Fischer. Einfluss der configuration auf die wirkung der enzyme. *Ber. Dtsch. Chem. Ges.*, 27(3):2985–2993, 1894.
- [15] J. R. Perkins, I. Diboun, B. H. Dessailly, J. G. Lees, and C. Orengo. Transient protein-protein interactions: Structural, functional, and network properties. *Structure*, 18(10):1233–1243, 2010.
- [16] P. B. Crowley and M. Ubbink. Close encounters of the transient kind: Protein interactions in the photosynthetic redox chain investigated by nmr spectroscopy. *Acc. Chem. Res.*, 36(10):723–730, 2003.
- [17] M. Ubbink. The courtship of proteins: understanding the encounter complex. *FEBS Lett.*, 583(7):1060–1066, 2009.
- [18] S. E. Hart, B. G. Schlarb-Ridley, C. Delon, D. S. Bendall, and C. J. Howe. Role of charges on cytochrome f from the cyanobacterium *phormidium laminosum* in its interaction with plastocyanin. *Biochemistry*, 42(17):4829–4836, 2003.
- [19] G. McLendon. *Control of biological electron transport via molecular recognition and binding: The “velcro” model*, pages 159–174. Springer Berlin Heidelberg, Berlin, Heidelberg, 1991.
- [20] T. Selzer and G. Schreiber. New insights into the mechanism of protein-protein association. *Proteins: Struct., Func., Bioinf.*, 45(3):190–198, 2001.
- [21] G. Schreiber and A. R. Fersht. Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.*, 3:427–431, 1996.
- [22] Q. Bashir, A. Volkov, G. M. Ullmann, and M. Ubbink. Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome c and cytochrome c peroxidase. *J. Am. Chem. Soc.*, 132:241–247, 2010.
- [23] J. A. R. Worrall, W. Reinle, R. Bernhardt, and M. Ubbink. Transient protein interactions studied by nmr spectroscopy: The case of cytochrome c and adrenodoxin. *Biochemistry*, 42(23):7068–7076, 2003.
- [24] X. Xu, W. Reinle, F. Hannemann, P. V. Konarev, D. I. Svergun, R. Bernhardt, and M. Ubbink. Dynamics in a pure encounter complex of two proteins studied by solution scattering and paramagnetic nmr spectroscopy. *J. Am. Chem. Soc.*, 130(20):6395–6403, 2008.
- [25] J. A. R. Worrall, Y. Liu, P. B. Crowley, J. M. Nocek, B. M. Hoffman, and M. Ubbink. Myoglobin and cytochrome b5: A nuclear magnetic resonance study of a highly dynamic protein complex. *Biochemistry*, 41(39):11721–11730, 2002.
- [26] A. D. MacKerel Jr., C. L. Brooks III, L. Nilsson, B. Roux, Y. Won, and M. Karplus. *CHARMM: The Energy Function and Its Parameterization with an Overview of*

- the Program*, volume 1 of *The Encyclopedia of Computational Chemistry*, pages 271–277. John Wiley & Sons: Chichester, 1998.
- [27] J. Ponder and D. A. Case. Force fields for protein simulations. protein simulations. *Adv. Prot. Chem*, 66:27–85, 2003.
- [28] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren. The gromos biomolecular simulation program package. *J. Phys. Chem. A*, 103(19):3596–3607, 1999.
- [29] K. A. Sharp and B. Honig. Electrostatic interactions in macromolecules: Theory and applications. *Annu. Rev. Biophys. Biophys. Chem.*, 19(1):301–332, 1990.
- [30] G. M. Ullmann and E.-W. Knapp. Electrostatic models for computing protonation and redox equilibria in proteins. *Eur. Biophys. J.*, 28(7):533–551, 1999.
- [31] M. R. Gunner, J. Mao, Y. Song, and J. Kim. Factors influencing the energetics of electron and proton transfers in proteins. what can be learned from calculations. *Biochim. Biophys. Acta, Bioener.*, 1757(8):942–968, 2006.
- [32] G. M. Ullmann and E. Bombarda. *Continuum Electrostatic Analysis of Proteins*, pages 135–163. Springer International Publishing, Cham, 2014.
- [33] J. G. Kirkwood. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *J. Chem. Phys.*, 2(7):351–361, 1934.
- [34] F. B. Sheinerman, R. Norel, and B. Honig. Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.*, 10(2):153–159, 2000.
- [35] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.
- [36] A. N. Volkov, J. A. R. Worrall, E. Holtzmann, and M. Ubbink. Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic nmr. *Proc. Natl. Acad. Sci. USA*, 103:18945–18950, 2006.
- [37] J. L. Battiste and G. Wagner. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry*, 39:5355–5365, 2000.
- [38] I. Díaz-Moreno, A. Díaz-Quintana, M. A. De la Rosa, and M. Ubbink. Structure of the complex between plastocyanin and cytochrome f from the cyanobacterium *nostoc* sp. pcc 7119 as determined by paramagnetic nmr. *J. Biol. Chem.*, 280(19):18908–18915, 2005.
- [39] C. Lange, T. Cornvik, I. Díaz-Moreno, and M. Ubbink. The transient complex of poplar plastocyanin with cytochrome f: effects of ionic strength and pH. *Biochim. Biophys. Acta, Bioener.*, 1707(2-3):179–188, 2005.

- [40] M. Ubbink, M. Ejdebäck, B. G. Karlsson, and D. S. Bendall. The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic nmr and restrained rigid-body molecular dynamics. *Structure*, 6(3):323–335, 1998.
- [41] P. B. Crowley, G. Otting, B. G. Schlarb-Ridley, G. W. Canters, and M. Ubbink. Hydrophobic interactions in a cyanobacterial plastocyanin-cytochrome f complex. *J. Am. Chem. Soc.*, 123(43):10444–10453, 2001.
- [42] J.-Y. Suh, C. Tang, and G. M. Clore. Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement. *J. Am. Chem. Soc.*, 129(43):12954–12955, 2007.
- [43] G. M. Clore and J. Iwahara. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.*, 109(9):4108–4139, 2009.
- [44] N. L. Fawzi, M. Doucleff, J.-Y. Suh, and G. M. Clore. Mechanistic details of a protein-protein association pathway revealed by paramagnetic relaxation enhancement titration measurements. *Proc. Natl. Acad. Sci. U.S.A.*, 107(4):1379–1384, 2010.
- [45] J. Iwahara, D. E. Anderson, E. C. Murphy, and G. M. Clore. Edta-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to dna by intermolecular paramagnetic relaxation enhancement. *J. Am. Chem. Soc.*, 125(22):6634–6635, 2003.
- [46] P. H. J. Keizers, J. F. Desreux, M. Overhand, and M. Ubbink. Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. *J. Am. Chem. Soc.*, 129(30):9292–9293, 2007.
- [47] P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand, and M. Ubbink. Design, synthesis, and evaluation of a lanthanide chelating protein probe: Clanp-5 yields predictable paramagnetic effects independent of environment. *J. Am. Chem. Soc.*, 130(44):14802–14812, 2008.
- [48] P. A. Kosen. [5] spin labeling of proteins. In *Nuclear Magnetic Resonance Part B Structure and Mechanism*, volume 177 of *Methods in Enzymology*, pages 86–121. Academic Press, 1989.
- [49] C. Tang, J. M. Louis, A. Aniana, J.-Y. Suh, and G. M. Clore. Visualizing transient events in amino-terminal autoprocessing of hiv-1 protease. *Nature*, 455:693–696, 2008.
- [50] Y. Takayama and G. M. Clore. Intra- and intermolecular translocation of the bi-domain transcription factor oct1 characterized by liquid crystal and paramagnetic nmr. *Proc. Natl. Acad. Sci. U.S.A.*, 108(22):E169–E176, 2011.
- [51] A. N. Volkov, Q. Bashir, J. A. R. Worrall, G. M. Ullmann, and M. Ubbink. Shifting the equilibrium between the encounter state and the specific form of a protein complex by interfacial point mutations. *J. Am. Chem. Soc.*, 132(33):11487–11495,

2010.

- [52] D. S. Gorman and R. P. Levine. Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA*, 54(6):1665–1669, 1965.
- [53] I. Bertini, D. A. Bryant, S. Ciurli, A. Dikiy, C. O. Fernández, C. Luchinat, N. Safarov, A. J. Vila, and J. Zhao. Backbone dynamics of plastocyanin in both oxidation states: Solution structure of the reduced form and comparison with the oxidized state. *J. Biol. Chem.*, 276(50):47217–47226, 2001.
- [54] F. De Rienzo, R. R. Gabdoulhine, M. C. Menziani, and R.C. Wade. Blue copper proteins: A comparative analysis of their molecular interaction properties. *Protein Sci.*, 9(8):1439–1454, 2000.
- [55] M. Ejdebäck, A. Bergkvist, B. G. Karlsson, and M. Ubbink. Side-chain interactions in the plastocyanin-cytochrome f complex. *Biochemistry*, 39(17):5022–5027, 2000.
- [56] I. Díaz-Moreno, A. Díaz-Quintana, M. A. De la Rosa, P. B. Crowley, and M. Ubbink. Different modes of interaction in cyanobacterial complexes of plastocyanin and cytochrome f. *Biochemistry*, 44(9):3176–3183, 2005.
- [57] C. Albarrán, J. A. Navarro, M. A. De la Rosa, and M. Hervás. The specificity in the interaction between cytochrome f and plastocyanin from the cyanobacterium *Nostoc sp. pcc 7119* is mainly determined by the copper protein. *Biochemistry*, 46(4):997–1003, 2007.
- [58] M. A. De la Rosa, J. A. Navarro, A. Díaz-Quintana, B. De la Cerda, F. P. Molina-Heredia, A. Balme, P. del S. Murdoch, I. Díaz-Moreno, R. V. Durán, and M. Hervás. An evolutionary analysis of the reaction mechanisms of photosystem i reduction by cytochrome c6 and plastocyanin. *Bioelectrochemistry*, 55(1):41–45, 2002.
- [59] M. A. De la Rosa, J. A. Navarro, and M. Hervás. *The Convergent Evolution of Cytochrome c6 and Plastocyanin Has Been Driven by Geochemical Changes*, pages 607–630. Springer Netherlands, Dordrecht, 2011.
- [60] C. A. Kerfeld and D. W. Krogmann. Photosynthetic cytochromes c in cyanobacteria, algae, and plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 49(1):397–425, 1998.
- [61] F. P. Molina-Heredia, M. Hervás, J. A. Navarro, and M. A. De la Rosa. Cloning and correct expression in *Escherichia coli* of the *petE* and *petJ* genes respectively encoding plastocyanin and cytochrome c6 from the cyanobacterium *Anabaena sp. pcc 7119*. *Biochem. Biophys. Res. Commun.*, 243(1):302–306, 1998.
- [62] M. Hervás, J. A. Navarro, A. Díaz, and M. A. De la Rosa. A comparative thermodynamic analysis by laser-flash absorption spectroscopy of photosystem i reduction by plastocyanin and cytochrome c6 in *Anabaena pcc 7119*, *Synechocystis pcc 6803*, and spinach. *Biochemistry*, 35(8):2693–2698, 1996.

- [63] J. A. R. Worrall, B. G. Schlarb-Ridley, T. Reda, M. J. Marcaida, R. J. Moorlen, J. Wastl, J. Hirst, D. S. Bendall, B. F. Luisi, and C. J. Howe. Modulation of heme redox potential in the cytochrome c6 family. *J. Am. Chem. Soc.*, 129(30):9468–9475, 2007. PMID: 17625855.
- [64] F. P. Molina-Heredia, J. Wastl, J. A. Navarro, D. S. Bendall, M. Hervás, C. J. Howe, and M. A. De la Rosa. A new function for an old cytochrome? *Nature*, 424:33–34, 2003.
- [65] M. R. Sawaya, D. W. Krogmann, A. Serag, K. K. Ho, T. O. Yeates, and C. A. Kerfeld. Structures of cytochrome c-549 and cytochrome c6 from the cyanobacterium *arthrospira maxima*. *Biochemistry*, 40(31):9215–9225, 2001.
- [66] C. A. Kerfeld, H. P. Anwar, R. Interrante, S. Merchant, and T. O. Yeates. The structure of chloroplast cytochrome c6 at 1.9 Å resolution: Evidence for functional oligomerization. *J. Mol. Biol.*, 250(5):627–647, 1995.
- [67] J. Schnackenberg, M. E. Than, K. Mann, G. Wiegand, R. Huber, and W. Reuter. Amino acid sequence, crystallization and structure determination of reduced and oxidized cytochrome c6 from the green alga *scenedesmus obliquus* edited by a. r. fersht. *J. Mol. Biol.*, 290(5):1019–1030, 1999.
- [68] J. A. Navarro, M. Hervás, and M. A. De la Rosa. Co-evolution of cytochrome c6 and plastocyanin, mobile proteins transferring electrons from cytochrome b6f to photosystem i. *J. Biol. Inorg. Chem.*, 2(1):11–22, 1997.
- [69] W. A. Cramer, G. M. Soriano, M. Ponomarev, D. Huang, H. Zhang, S. E. Martinez, and J. L. Smith. Some new structural aspects and old controversies concerning the cytochrome b6f complex of oxygenic photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47(1):477–508, 1996. PMID: 15012298.
- [70] G. Kurisu, H. Zhang, J. L. Smith, and W. A. Cramer. Structure of the cytochrome b6f complex of oxygenic photosynthesis: Tuning the cavity. *Science*, 302(5647):1009–1014, 2003.
- [71] C. J. Carrell, B. G. Schlarb, D. S. Bendall, C. J. Howe, W. A. Cramer, and J. L. Smith. Structure of the soluble domain of cytochrome f from the cyanobacterium *phormidium laminosum*. *Biochemistry*, 38(30):9590–9599, 1999. PMID: 10423236.
- [72] D. Stroebel, Y. Choquet, J.-L. Popot, and D. Picot. An atypical haem in the cytochrome b6f complex. *Nature*, 426:413–418, 2003. Article.
- [73] G. R. Moore and G. W. Pettigrew. *Cytochromes c*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1990.
- [74] S. Scanu, J. Förster, M. G. Finiguerra, M. H. Shabestari, M. Huber, and M. Ubink. The complex of cytochrome f and plastocyanin from *nostoc* sp. pcc 7119 is highly dynamic. *ChemBioChem*, 13(9):1312–1318, 2012.
- [75] G. M. Ullmann, E.-W. Knapp, and N. M. Kostić. Computational Simulation and

- Analysis of Dynamic Association between Plastocyanin and Cytochrome f. Consequences for the Electron-Transfer Reaction. *J. Am. Chem. Soc.*, 119(1):42–52, 1997.
- [76] D. C. Pearson and E. L. Gross. Brownian dynamics study of the interaction between plastocyanin and cytochrome f. *Biophys. J.*, 75(6):2698–2711, 1998.
- [77] I. Díaz-Moreno, A. Díaz-Quintana, M. Ubbink, and M. A. De la Rosa. An nmr-based docking model for the physiological transient complex between cytochrome f and cytochrome c6. *FEBS Lett.*, 579(13):2891–2896, 2005.
- [78] X.-S. Gong, J. Q. Wen, N. E. Fisher, S. Young, C. J. Howe, D. S. Bendall, and J. C. Gray. The role of individual lysine residues in the basic patch on turnip cytochrome f for electrostatic interactions with plastocyanin in vitro. *Eur. J. Biochem.*, 267(12):3461–3468, 2000.
- [79] B. G. Schlarb-Ridley, D. S. Bendall, and C. J. Howe. Role of Electrostatic in the Interaction between Cytochrome f and Plastocyanin of the Cyanobacterium *Phormidium laminosum*. *Biochemistry*, 41(10):3279–3285, 2002.
- [80] I. Cruz-Gallardo, I. Díaz-Moreno, A. Díaz-Quintana, and M. A. De la Rosa. The cytochrome f-plastocyanin complex as a model to study transient interactions between redox proteins. *FEBS Lett.*, 586(5):646–652, 2012.
- [81] P. B. Crowley, A. Díaz-Quintana, F. P. Molina-Heredia, P. Nieto, M. Sutter, W. Haehnel, M. A. De la Rosa, and M. Ubbink. The interactions of cyanobacterial cytochrome c6 and cytochrome f, characterized by NMR. *J. Biol. Chem.*, 277(50):48685–48689, 2002.
- [82] R. A. Marcus. On the theory of oxidation-reduction reactions involving electron transfer. v. comparison and properties of electrochemical and chemical rate constants. *J. Phys. Chem.*, 67(4):853–857, 1963.
- [83] R. A. Marcus and N. Sutin. Electron transfers in chemistry and biology. *Biochim. Biophys. Acta, Rev. Bioener.*, 811(3):265–322, 1985.
- [84] S. Nakani, L. B. Vitello, and J. E. Erman. Characterization of Four Covalently-Linked Yeast Cytochrome c /Cytochrome c Peroxidase Complexes: Evidence for Electrostatic Interaction between Bound Cytochrome c Molecules. *Biochemistry*, 45(48):14371–14378, 2006.
- [85] H. Mei, K. Wang, S. McKee, X. Wang, J. L. Waldner, G. J. Pielak, B. Durham, and F. Millett. Control of formation and dissociation of the high-affinity complex between cytochrome c and cytochrome c peroxidase by ionic strength and the low-affinity binding site. *Biochemistry*, 35(49):15800–15806, 1996.
- [86] T. Ž. Grove, G. M. Ullmann, and N. M. Kostić. Simultaneous true, gated, and coupled electron-transfer reactions and energetics of protein rearrangement. *J. Inorg. Biochem.*, 106(1):143–150, 2012.
- [87] S. J. Moench and J. D. Satterlee. A comparison of spectral and physicochemical properties of yeast iso-1 cytochrome c and cys 102-modified derivatives of the

protein. *J. Prot. Chem.*, 14(7):567–582, 1995.

- [88] T. Ž. Grove and N. M. Kostić. Metalloprotein association, self-association, and dynamics governed by hydrophobic interactions: Simultaneous occurrence of gated and true electron-transfer reactions between cytochrome f and cytochrome c6 from *Chlamydomonas reinhardtii*. *J. Am. Chem. Soc.*, 125(35):10598–10607, 2003.

Danksagungen/ Acknowledgements

Mein herzlicher Dank gilt

Prof. Dr. Matthias Ullmann für die ausgezeichnete Betreuung und wissenschaftliche Weiterbildung während meiner Zeit in deiner Arbeitsgruppe. Du hast immer ein offenes Ohr und hast maßgeblich zur wissenschaftlichen Ausrichtung während meiner Zeit in Bayreuth beigetragen. Danke dafür.

Ina Pöhner für die vielen Ideen zur Weiterentwicklung von Monty-Dock.

Prof. Dr. Marcellus Ubbink for the good and interesting collaboration about the encounter complex and the introduction to the world of paramagnetic NMR.

Dr. Sandra Scanu for the great collaboration and the informative discussions. I have learned a lot from you. Thanks also for the proofreading of the thesis.

Dr. Jia-Ying Guan for the continuative collaboration and the work on the encounter complex analysis.

Lars Müller für die zahlreichen Diskussionen und hilfreichen Tipps, die eine Arbeit wie Diese erst ermöglicht haben. Weiterhin Danke für

das Korrekturlesen der Arbeit.

allen aktuellen und früheren Kollegen in der B14, insbesondere Elisa, Florian, Jan, Martin, Rajeev, Simon, Steffen, Thomas und Timm für die zahlreichen Ideen und Diskussionen.

Prof. Dr. Stephan Kümmel und Prof. Dr. Jürgen Köhler für die begleitende Betreuung als Mentoren während meiner Zeit im Graduiertenkolleg 1640.

Ingo Schelter für die gute Zusammenarbeit im Rahmen des Graduiertenkollegs.

Claudia Geier und Violaine Zigan für die Unterstützung bei administrativen Aufgaben.

dem Grauiertenkolleg 1640, der Deutschen Forschungsgemeinschaft und der Bayreuther Graduiertenschule für die finanzielle Unterstützung und der Förderung wissenschaftlicher Fähigkeiten.

Yvonne Nerbl, meinem Fels in der Brandung. Du unterstützt mich bei all meinen Handlungen. Danke für das Korrekturlesen der Arbeit.

meiner Familie, die immer für mich da ist und mich ermutigt. Danke für das Vertrauen und die Geduld.

all meinen Freunden, die mich in meinem Leben begleiten und immer für mich da sind. Danke für die gemeinsame Zeit.

nochmals allen, die zum Gelingen dieser Arbeit beigetragen haben.
Thanks for everything!

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Weiterhin erkläre ich, dass ich die Hilfe von gewerblichen Promotionsberatern bzw. Promotionsvermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe, noch künftig in Anspruch nehmen werde.

Ferner erkläre ich, dass ich nicht bereits anderweitig mit oder ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

Bayreuth, den 7. Februar 2018,

Johannes M. Förster