Transient interactions between photosynthetic proteins

Hulsker, R.

Chapter I

Introduction

Static or transient

The vast number of biological processes that occur in living systems are mediated mostly by protein-protein interactions. In the cell, genetically encoded information is translated into proteins, which are crucial to the majority of cellular functions. Protein-protein complexes show a great variety in properties, depending on the function they perform. One aspect to consider when studying these complexes is their affinity, which is directly related to the lifetime of the complex. Historically, because protein complexes with a high affinity have long lifetimes, they were the first to be characterised. These so-called static complexes are exemplified by two well-studied RNase-inhibitor complexes. The low dissociation constant (K_d) of 10^{-14} M of the barnase-barstar complex[1] is essential to ensure tight binding and thereby inhibition of the potentially harmful intracellular activity of barnase (an extracellular RNase from *Bacillus*). The same is true for the human angiogenin – ribonuclease inhibitor complex[2]. Angiogenin induces blood vessel formation and is involved in many types of cancer. The angiogenin - inhibitor complex has an even lower[3] K_d of $<10^{-15}$ M, again indicative of the strict control organisms hold over RNases. A typical feature of static complexes are the so called 'hot spots', which are areas of interface residues that can be found by alanine scanning mutagenesis and that energetically contribute most to binding[4-7].

At the other extreme of the scale of affinities in protein interactions, apart from aspecific interactions, we find the transient protein complexes. Their affinities are in the μM^{-1} – mM^{-1} range and the associated lifetimes are on the millisecond time-scale. Transient protein complexes exhibit high turn-over and comprise proteins that interact with multiple partners: signal transduction cascades and electron transfer chains in photosynthesis and respiration for example. The field of metallo-proteins produced the initial questions and insights into the properties of proteins in redox chains and was at the basis of much current knowledge on transient protein complexes[8,9].

Transient protein complex formation and interface

In order to sustain the high turn-over rates required in electron transfer chains and signal transduction cascades, the chain components need to interact transiently. This is achieved by a high dissociation rate constant ($k_{off} \geq 10^3$ s^{-1}). Considering the association rate constants in the range of 10^7-10^9 M^{-1} s^{-1} found experimentally for several electron transfer complexes[10-13], we can explain the weak binding of transient complexes, with association constants ($K_a = k_{on}/k_{off}$) in the μM^{-1} – mM^{-1} range[8].

High k_{off} rates and the need to interact with multiple partners account for the poorly packed complex interfaces which are typical for these complexes[14]. Dissociation can be more rapidly achieved when water has increased access to the poorly packed interface. Promiscuity is also facilitated by the lack of geometric complementarity of the interface. In static as well as transient complex interfaces charged residues are found, but in transient interfaces they are more often located at the perimeter. There, they are usually not involved in salt-bridges and hydrogen bonds. These observations can be rationalised by two considerations. First, the presence of opposite charges on the partner proteins can facilitate complex formation by increasing the chance of the proteins finding each other, which is also true for static complexes[15-18]. Second, the need for high turn-over dictates that the proteins need to release easily, which is reflected in the high k_{off} rates. Strong electrostatic interactions in the complex interface would inhibit the release, so by not matching the charges in the interface exactly, these high k_{off} rates can be achieved.

While long-range electrostatic interactions are mainly involved in the initial stages of transient protein-protein association, hydrophobic interactions are more important for the specificity and stability of the transient complex. In electron transfer complexes a patch of hydrophobic residues is often found on the surface, close to the active site. This is thought to facilitate electron transfer by providing an entry point for the electrons. In order to prevent complete desolvation, as can be seen in static complexes, these patches are often kept small and are surrounded by a ring of polar residues, which enhances dissociation of the complex.

The encounter complex, kinetic models of binding

The theoretical diffusion-limited association rate for two spherical proteins of 18 Å proteins, given by the Smoluchowski equation[19] is 7 x 10^9 M^{-1} s^{-1}. In reality however, a successful binding event requires the interface residues on the proteins to align. As the interface area comprises on average 10-15% of the total surface area[20] and considering the high steric specificity needed to align the interface areas properly, the association rate is expected to drop by at least three or four orders of magnitude[8,21-23]. However, some protein-protein complexes, such as cytochrome *c* – cytochrome *c* peroxidase[24] and cytochrome *c* - cytochrome b_5 [25], exhibit association constants of ~10^9 M^{-1} s^{-1}. This discrepancy raises the question of how some of these complexes can achieve such high association rates. Several models have been proposed to explain the experimental data available. From laser flash kinetics studies on the electron transfer from Pc and cytc_6 to PSI three models were proposed that could account for the observed electron transfer reaction rates[12,26]. The type I model is a simple collisional model without the formation of a transient complex. The type II model involves complex formation and electron transfer, without a transitional rearrangement step. This suggests that the preformed complex is either sufficiently oriented for electron transfer or that rearrangement is too fast to be detected. The most elaborate (type III) model includes the formation of an ensemble of pre-oriented complexes guided by electrostatic interactions. Intermolecular rearrangement then forms the active electron transfer complex. This mechanism is reminiscent of the earlier Velcro-model[27], which rejected the classical 'lock and key' concept, that is more appropriate to describe static complexes and enzymatic processes. Instead, it suggested that the recognition and binding of electron transfer proteins originate from charged and hydrophobic complementary patches that stick on contact like Velcro.

From a cross-linking study on the plant Pc - cyt*f* complex, which causes the inhibition of electron transfer, it can be concluded that the initial orientation of the complex is not productive[28]. NMR structural studies of plant Pc - cyt*f* complexes[29,30] indicated that a large rearrangement from the well-defined state was unlikely. These results lead to the proposition of a two-step model[29,31] in which an ensemble of orientations is formed based

on electrostatic interactions, followed by a well-defined complex which is capable of electron transfer. In the case of plant Pc - cyt*f* the equilibrium between these steps lies towards the well-defined state, but in other cases the equilibrium is more toward the ensemble of orientations. The cytochrome b_5 – myoglobin complex can mainly be found in this state, existing as a number of isoenergetic binding orientations, described by the term dynamic docking[32-34]. More recently the investigation of the dynamics in the cytochrome *c*- cytochrome *c* peroxidase complex refined the two-step model and defined a dynamic encounter state as a combination of all encounter complexes, followed by a single-orientation or well-defined state[31,35] (Fig. 1.1).

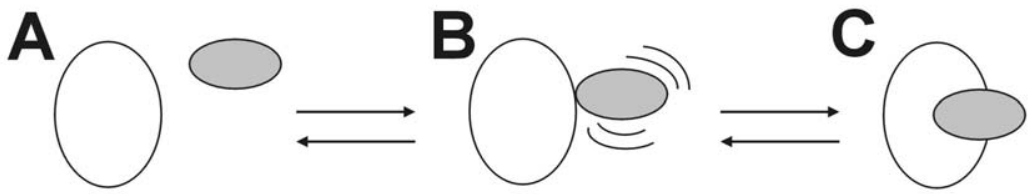

Figure 1.1. Model for the formation of a protein complex[29,31]. Free proteins (A) associate to form an encounter complex (B) consisting of an ensemble of protein orientations, which is in equilibrium with a single-orientation complex (C).

Dynamics in protein complexes studied by NMR spectroscopy

The term dynamics is used to describe movement in proteins and protein complexes on every timescale possible[36]. 'Dynamics' ranges from side chain and domain movements in proteins, to movement of proteins with respect to each other in protein complexes, but the term is also used for the turn-over of protein complexes on timescales that are involved in cellular processes. In this work, when the term dynamics is used it covers the movements of proteins in the (transient) protein complex. As mentioned above, these movements are thought to result in an ensemble of orientations, which together form the encounter state.
The field of NMR spectroscopy has had a great impact on the understanding of transient complexes as well as on the dynamics involved. Several types of NMR experiments have been deployed to study them[37]. The chemical shift perturbation experiment, where a ^{15}N-labelled protein is monitored while being titrated with its partner protein, is used to

determine which residues experience a change in their magnetic environment caused by the presence of the other protein. Plotting these changes on a surface representation of the protein then provides a map of the interface. Furthermore, titration curves can be fitted to produce a binding constant[16]. Interestingly, the overall size of the chemical shift changes was found to be a qualitative measure of the dynamics in the complex[33,38]. The more dynamic the complex, the more the chemical shift changes average out. In its extreme form the chemical shift changes are indistinguishable from aspecific interactions, as is the case for the myoglobin-cytochrome b_5 complex. Complex formation could be inferred from an increase in overall correlation time and the fact that the chemical shift perturbations titrate[33]. These NMR studies indicated that transient complexes cover a full range from single-orientation to fully dynamic.

Recently, paramagnetic spin labels have been utilised in attempts to quantify the dynamics in protein-protein complexes[31,39]. The paramagnetic properties of metals in electron transfer proteins were recognised earlier to be useful in structural studies of transient complexes[29,30,40-42]. This is being extended to proteins that do not naturally contain metals, by the attachment of metal-chelating tags[43-49] and nitroxide spin-labels[50-54] to the protein surface. Chemical shift perturbations give information on the proximity of the partner protein, but not to which specific nuclei of the partner protein is close to a particular nucleus in the ^{15}N-labelled protein. Paramagnetic effects on the other hand, give information on the distance of a nucleus to the paramagnetic centre, with the added advantage in the case of pseudocontact shifts that they also provide angular information. This is due to the anisotropic nature of the paramagnetic centre used to induce pseudocontact shifts. Both forms of information can be converted to distance and angular restraints and used in docking calculations to determine the structure of complexes. The use of paramagnetic relaxation enhancement has recently been discovered to be useful as a tool for investigating dynamics[31,39,55].

Photosynthesis

The proteins studied in this thesis transfer electrons as part of oxygenic photosynthesis. For recent reviews on the structural and mechanistical aspects of photosynthesis and photosynthesis as the inspiration for designing solar powered biotechnologies see references[56-59]. Oxygenic photosynthesis, performed by plants, green algae and cyanobacteria, converts sunlight into chemical energy (Fig. 1.2).

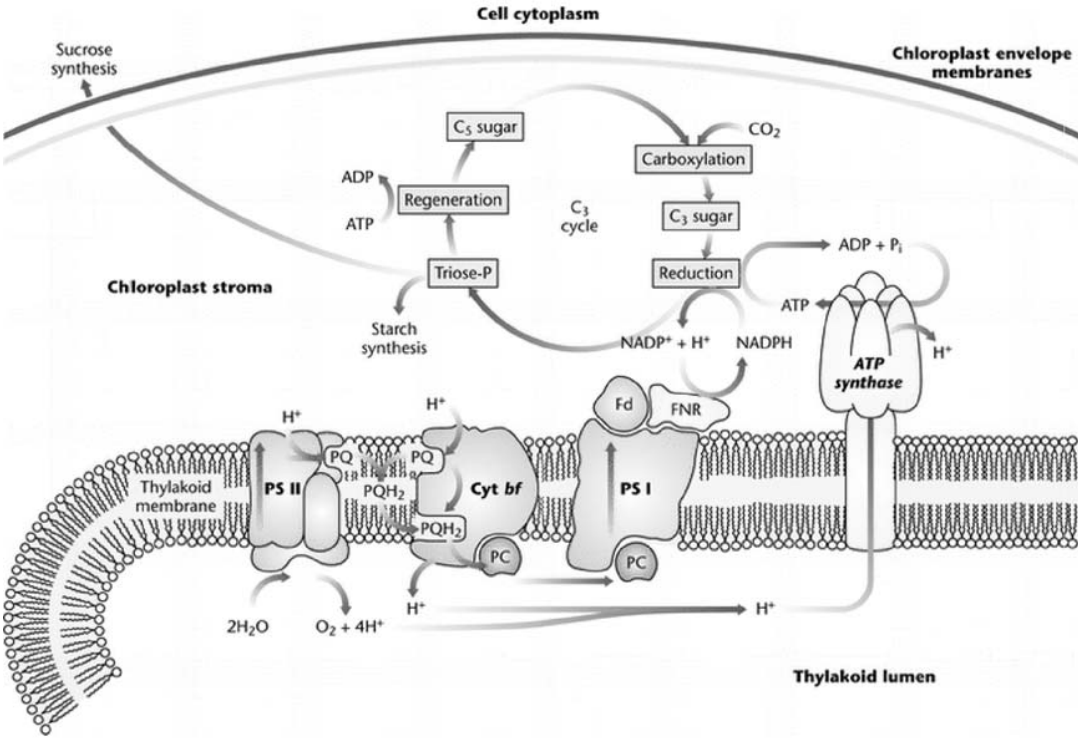

Figure 1.2. Schematic representation of oxygenic photosynthesis in plant chloroplasts[60].

The production of oxygen as a side product and the assimilation of carbon dioxide into organic matter are essential to all life forms. Photosynthesis in plants takes place in specialised organelles called chloroplasts, which are compartmentalised by stacked thylakoid membranes. These separate the lumen from the stroma, and facilitate formation of a proton gradient (also called proton motive force or pmf), which is the direct result of the light reactions. This starts with the splitting of water by PSII, providing the electrons

and producing O_2. The electrons are translocated across the membrane, initialised by the excitation of the reaction centres. This is achieved by light harvested by chlorophyll and other pigments that funnel the energy to specialised reaction centre chlorophylls in PSII. The electrons are then taken up by a quinone pool in the thylakoid membrane and shuttled to the cytochrome b_6f membrane complex where the redox energy is converted into the proton gradient. The proton gradient is used by an ATP synthase to form ATP. At the same time the electrons are shuttled by Pc to PSI, which produces reduced ferrodoxin at the stromal side of the membrane. Reduced ferrodoxin can be used for NADPH production, but also for other processes that need reductive power.

The proteins:

Plastocyanin

Plastocyanin (Pc) was first discovered in 1960 by Sakae Katoh[61], as a copper containing redox protein with an absorption maximum at 597 nm. It had been isolated from the green alga *Chlorella ellipsoidea* and was named plastocyanin[62] in 1961. The experiment that settled the discussion on whether Pc interacts with PSII or PSI was essential in determining the order of electron flow in the light reaction of photosynthesis[63]. It is now well established that Pc transports electrons from the cytochrome b_6f complex to PSI in green algae, cyanobacteria and plants. The divergent sequence of the *petE* gene coding for Pc is known for more than forty organisms. It codes for 97 to 105 amino acids and is usually longer in cyanobacteria. Crystal and solution structures are available for Pc from fifteen sources[65-81]. Despite its sequence divergence the structure of the ~11 kDa protein (Fig. 1.3) is conserved and forms a β-sandwich consisting of seven β-strands and an irregular strand that contains a small helix in the longer Pc's. An invariable feature is the type I 'blue' copper site, which is characterised by an intense blue colour at ~600 nm, an unusually narrow hyperfine coupling ($A_\parallel$ values of 0.0035-0.0063 cm^{-1}) in the EPR spectrum due to assymmetry at the metal and a high midpoint-potential (>350 mV at pH 7 for Pc's) as compared to the aqua Cu(II/I) couple[82]. Type II proteins also have a single copper atom, but are not characterised by an intense colour or unusual EPR spectrum.

Type 3 proteins are binuclear and therefore not EPR active. More types of copper sites have recently been discovered, but these will not be discussed here.

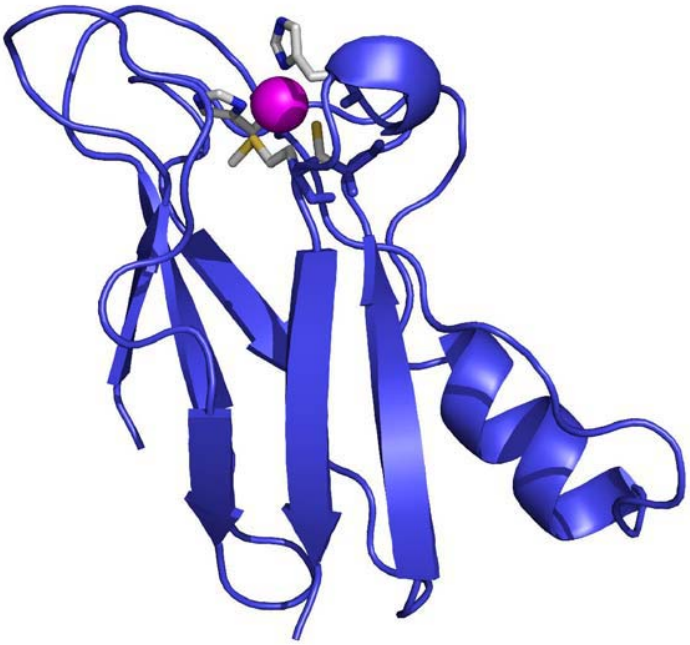

Figure 1.3. Three-dimensional representation of Pc. Pc is shown as a blue ribbon, with the Cu ion as a magenta sphere. The four copper ligands, a cysteine, methionine and two histidines, are shown as sticks. The molecular coordinates were taken from the crystal structure of reduced fern Pc (PDB entry 1KDI). This figure and Figures 1.4, 3.3, 3.5, 3.8, 4.1, 4.4, 4.6 and 5.3 were made in PyMOL v0.98 [64].

The copper ion in Pc is coordinated by a S_γ of a Cys, S_δ of a Met and two N_δ of two His side chains in a trigonal pyramid [83,84]. The colour of PCu(II) is caused by the Cys S_γ – Cu(II) charge transfer, which has a maximum at ~600 nm. The C-terminal histidine is the point of entry of the electron and becomes protonated in the reduced protein when the pH is lowered to <5. The only exception to that rule has been reported for Pc from Japanese fern[77,85] (but see Chapter II and reference [86]). This active site histidine is located at the so called 'north' end of the protein in a hydrophobic patch. Pc from the cyanobacterium *P. hollandica* contains two unique residues in this patch (Tyr12 and Pro14), that have not been found in other Pc's. When converted back to the otherwise conserved residues the complex with cyt*f* appears to show increased dynamics[87]. Another site of interest is the 'eastern' or acidic patch. It is usually located at the irregular strand 5 and includes residues 42-65. It was thought to be directly involved in electron transfer from cyt*f*, as deduced from studies with inorganic compounds[88-97], but this could be disproved once the first structure of the Pc - cyt*f* complex was determined[29]. The electrostatic interaction of this patch with oppositely charged residues that are located in and near the small domain of plant cyt*f* contribute to complex formation as discussed above. Curiously, the charges

on Pc and cyt*f* are reverted in the Pc - cyt*f* complex from cyanobacterium *Nostoc* sp. PCC 7119, with a basic patch on Pc interacting with acidic side chains of cyt*f* [41,98].

Cytochrome f

Figure 1.4. Three dimensional representation of cyt*f*. Cyt*f* is shown as a red ribbon, the haem in sticks and the Fe ion as a red sphere. The molecular coordinates were taken from a homology model[109] of *Nostoc* sp. PCC 7119 cyt*f* (see Chapter IV), based on the crystal structure of *Brassica rapa* (PDB entries 1CTM and 1HCZ).

Cytochrome *f* (cyt*f*) is a protein with a unique structure, which makes it the only member of a family and superfamily of protein structures[99]. It was discovered in the 1930s and first characterised in the 1950s by Robert Hill[100]. The discovery of its natural proteolysis in *Brassica* spp. in 1975, promoted by the presence of organic solvent, was an important step for the isolation of this membrane anchored protein[101]. The method was later perfected and extended to other plants by John Gray[102]. Cyt*f* is part of the membrane complex cytochrome b_6f, which was independently crystallised from two sources in 2003 [103,104]. This complex takes up two electrons from a quinone bound at the lumenal Q_0 site, which results in the release of two protons to the aqueous lumen. Cyt*f* is anchored in the thylakoid membrane by a stretch of thirty hydrophobic, C-terminal residues. The truncated soluble part (Fig. 1.4) is ~28 kDa and the crystal structure of cyt*f* from only three sources has been determined, one each from plant, green algae and cyanobacteria[105-108]. As for Pc, the sequence of cyt*f* shows diversity between plants, green algea and cyanobacteria, while the structure is much conserved. The differences in sequence reflect the differences at the surface and are linked with the surface properties of its partner Pc. The first crystal structure of cyt*f* [105]

revealed a few unique features of this protein that was classified as a *c*-type cytochrome, characterised by the haem co-factor that is covalently linked by two thioether bonds to two cysteine residues, which are part of a conserved CXXCH motif[110]. The histidine acts as the fifth ligand to the iron, but unlike the Met or His residue that is usually found as sixth ligand in *c*-type cytochromes, the N-terminal amino group of residue Tyr1 is the sixth ligand in cyt*f*. The structure of cyt*f* has a curious elongated shape and is mainly made up from β-sheets. It is composed of a small and large domain, the latter of which contains the active site haem.

Cytochrome c_6

Confused with cyt*f* in early literature[99] and also called cytochrome *c*-552 or *c*-553, this protein is a soluble *c*-type cytochrome, that performs the same functions as Pc in certain cyanobacteria and green algae. Other photosynthetic organisms can adapt to the copper availability in their environment by expressing either Pc or cytc_6[111]. Cytc_6 from green alga *Euglena gracilis,* was first characterised in 1959 as a *c*-type cytochrome with an α-band at 552 nm[112]. Recently, a homologous protein in plants was found, cytochrome c_{6A}, which has a lower midpoint potential[113,114] and is probably not involved in photosynthesis[115-117]. From the structure of *Ph. laminosum* cytochrome c_6 it was concluded that the replacement of a conserved Gln with Val in cytochrome c_{6A} caused its ~200 mV lower midpoint potential[114].

The structure of cytc_6 from eight different cyanobacteria or algae has been determined[114,118-126]. As mentioned before, in all class I cytochromes the haem is bound near the N-terminus by a CXXCH motif, of which the His is the fifth iron ligand. In cytc_6 a Met is the sixth ligand[127]. Contrary to plastocyanin the protein fold consist mainly of α-helices, typical of a cytochrome *c.*

When both proteins are available in one organism, they exhibit similar features. For instance, cytc_6 and plastocyanin share properties such as midpoint potential (~350 mV at pH 7) and isoelectric point. The isoelectric point is similar for both proteins within the same organism, but varies from one organism to another. It is believed the proteins underwent parallel surface mutations throughout evolution due to common

functionality[128,129]. Furthermore, cytc_6 is thought to have been replaced by plastocyanin when the character of the earth's atmosphere went from reducing to oxidative[129], because Cu^{II} is more readily available then Cu^{I} and the opposite is true for Fe^{III} and Fe^{II}.

The complexes of Pc and cytc_6 with cyt$f$

The transient nature of the Pc - cyt*f* and cytc_6 - cyt*f* complexes makes it difficult to co-crystallise them. For this reason NMR has been the method of choice for structural studies of these complexes. Before this was possible kinetics studies of the plant Pc - cyt*f* complex had already revealed the effect of ionic strength on the reaction rate[130] and electron transfer rate[16] showing the importance of electrostatics in the first step of association. For the determination of the first structure of the plant Pc - cyt*f* complex, pseudocontact shifts caused by the Fe^{III} in cyt*f* were used[29]. The copper in plastocyanin was replaced with cadmium to prevent reduction of cyt*f*. The structure revealed a so called 'side-on' orientation caused by the interaction of cyt*f* with the 'northern' hydrophobic patch and the eastern acidic patch on Pc as described above. When the same approach was used on the cyanobacterial, thermophilic *Ph. laminosum* Pc - cyt*f* [40], the structure showed an orientation of Pc which was 'head-on' instead, due to the lack of interactions with the eastern patch on Pc. The complex was found to be more dynamic, which is probably related to the fact that hydrophobic interactions seem to dominate in this complex, a view that is supported by results from kinetic studies[13,131]. The structure of the cyanobacterial *Nostoc* sp. PCC 7119 Pc - cyt*f* complex[109] revealed an orientation similar to the plant complex, but curiously the charges on Pc and cyt*f* are reversed. The differences between the Pc - cyt*f* complexes from different organisms can be attributed to differences of the surface properties of the proteins[41]. Theoretical studies, which often use Brownian dynamics, support the view from experimental work on the role of electrostatics and the ensemble of orientation formed in the initial association step[132-134]. The complex of cytc_6 - cyt*f* has received much less attention than the Pc - cyt*f* complex. The only reported structure is a docking model based on chemical shift perturbation data for the physiological cytc_6 - cyt*f* complex[135]. Despite the structural differences between Pc and cytc_6 the interaction with cyt*f* was found to be similar[136,137].

Outline of thesis

The aim of this thesis was to study several aspects that influence the transient interaction of Pc and cytc_6 with cyt*f*. In Chapter II the protonation of a copper histidine ligand of reduced fern Pc at low pH is investigated. This inactivation of Pc has been suggested as a way to down regulate photosynthesis. Protonation of the histidine ligand was reported for other Pc's but in fern Pc it seemed to be prevented by π–π stacking with nearby phenyl ring. This is investigated by NMR spectroscopy, X-ray crystallography and cyclic voltammetry for wt and mutant fern Pc. In Chapter III the complex of *P. hollandica* Pc - cyt*f* is structurally characterised by NMR spectroscopy. The structure of the physiological wt complex, determined by docking with NMR restraints, is presented. The mutation of two otherwise conserved residues in the hydrophobic patch of *P. hollandica* Pc is used to study the dynamics of this complex, which is balanced between well-defined and encounter state. The dynamics in the complex of *Nostoc* sp. PCC 7119 cytc_6 - cyt*f* are investigated in Chapter IV. Both pseudocontact shifts and PRE from five differently positioned spin labels on cyt*f* are used as distance restraints in docking calculations. In Chapter V the interaction of charged peptides with *S. pratensis* Pc is examined by NMR. These peptides have been shown to inhibit the electron transfer between Pc and cyt*f*. By addition of hydrophobic residues to the peptide, the effect of hydrophobics on the interaction is also investigated. In Chapter VI the effect of molecular crowding on the transient Pc - cyt*f* complex is investigated by NMR. The effects of addition of high molecular weight crowders on the binding constant and binding map of the complex are presented. Finally, in the concluding remarks the results from these chapters are put into a broader perspective and are discussed in relation with each other.