Dynamics in photosynthetic transient complexes studied by paramagnetic NMR spectroscopy

Scanu, S.

Citation

Scanu, S. (2013, October 10). *Dynamics in photosynthetic transient complexes studied by paramagnetic NMR spectroscopy*. Retrieved from https://hdl.handle.net/1887/21915

Version:	Not Applicable (or Unknown)
License:	Leiden University Non-exclusive license
Downloaded from:	https://hdl.handle.net/1887/21915

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/21915 holds various files of this Leiden University dissertation.

Author: Scanu, Sandra
Title: Dynamics in photosynthetic transient complexes studied by paramagnetic NMR spectroscopy
Issue Date: 2013-10-10

Chapter I

Introduction

Based on the review article:
Bashir Q., Scanu S., and Ubbink M. Dynamics in electron transfer protein complexes. *FEBS J*. **278**, 1391-1400 (2011).

Transient complexes

Protein-protein complexes can be classified as static or transient, on the basis on their binding characteristics. The equilibrium dissociation constant (K_D) is given by the ratio between the dissociation rate constant (k_{off}) and the association rate constant (k_{on}).[1] The k_{off} values range over many orders of magnitude and, consequently, so does the life-time of the complex (1/ k_{off}). Static complexes can have K_D values as low as 10^{-16} M, indicating life-times that can be as long as days, in the extreme case of toxins bound to their targets.[2] An example of a static complex formed by an enzyme and its inhibitor, is the complex of TEM1-β-lactamase (TEM1) and β-lactamase inhibitor protein (BLIP).[3] Association is not very fast (k_{on}= 2.6 x 10^5 M^{-1} s^{-1}), but the low k_{off} (2 x 10^{-4} s^{-1}) ensures that the proteins dissociate slowly. Even though the thermodynamic affinity is not very high (K_D = 10^{-9}M), the enzyme is kinetically shut down. On the other hand, K_D values for transient complexes vary in the range of μM to mM,[4] resulting from a high k_{off} (≥ 10 s^{-1})[5-8] and k_{on} in the range of 10^7-10^9 M^{-1} s^{-1}.[4,9] The lifetimes are thus on the μs-ms time scale. Proteins involved in signal transduction and electron transfer pathways are examples of such complexes. In all cases, the lifetime of the complex is closely related to its function. The transmission of a signal and the transfer of an electron are fast events and a high turnover of the involved partners is required, therefore, rapid dissociation is cardinal.

High geometrical surface complementarity between the interaction partners is a prerequisite in static complex formation. A closely packed interface is formed and stabilized by a dense network of short range forces, such as hydrogen bonds and van der Waals, hydrophobic and short-range electrostatic interactions. At the same time the multitude of short-range interactions leads to high specificity. There are specific residues, which contribute most to the binding,[10] so-called "hot spots" of binding and their presence is characteristic in static complexes.[11,12] Due to the lack of crystal structures, the ubiquity of "hot spots" in transient complexes has not yet been definitively established,[13] and they have been visualised only in a couple of cases.[14,15] Transient electron transfer (ET) protein complexes do not present a high level of geometric complementarity on the binding surface and the specificity is low. Instead, the protein surfaces are often optimized for the recognition of multiple interaction partners.[16] Such promiscuity is of paramount importance for the physiological function of the proteins.[14] For example, an electron carrier in photosynthesis or mitochondrial respiration needs to transfer an electron from a donor to an acceptor protein(s), so that it has to recognize specifically at least two partners, carry out ET and quickly dissociate. Proteins involved in ET usually hold the redox centre in the proximity of a surface hydrophobic patch, which represents the interaction patch for the reaction partner(s).[17] In this way, the ET pathway from the redox centre to the protein surface is optimised to be as short as possible. In the cupredoxin family,[18,19] as in many other cases, this hydrophobic region is surrounded by polar amino acids, which assist in dissociation since they facilitate the entry of water molecules in the binding interphase.[4] Finally, in order to be biologically functional a protein-protein complex requires a fine balance of association and dissociation rates, of affinity and specificity.[20]

Due to the interest in the ET-phenomenon and associated events, and the availability of biochemical and biophysical techniques applicable to these systems, the redox partners in photosynthesis and mitochondrial respiration are, at the moment, the best understood transient complexes.

Protein-protein interactions

Protein-protein interactions represent a fundamental biological phenomenon, being the basis of essentially all cellular processes, such as signal transduction regulation, electron transfer, chemical reactions, molecular recognition in the immune system, cytoskeletal movement, and more. A thorough understanding of the ways in which proteins recognise each other and the physical forces involved is paramount in addressing many questions pertaining to the events behind the interaction processes. The formation of a protein complex is a multistep reaction.[21] Initially, proteins come in contact with each other by random collisions and form an encounter complex, which is the end point of diffusional association.[22] In the first step, proteins diffuse by thermal motions and the presence of opposite charges leads the pre-orientation of the proteins by long-range electrostatics.[23] The second step consists in the transition from the encounter (AB*) to the final complex (AB), in which short-range interactions, such as van der Waals interactions, hydrogen bonds, hydrophobic interactions and short-range electrostatic interactions, dominate (Figure 1.1). An encounter complex does not always evolve to a final complex, in many cases it actually represents a futile complex (AB**).[22]

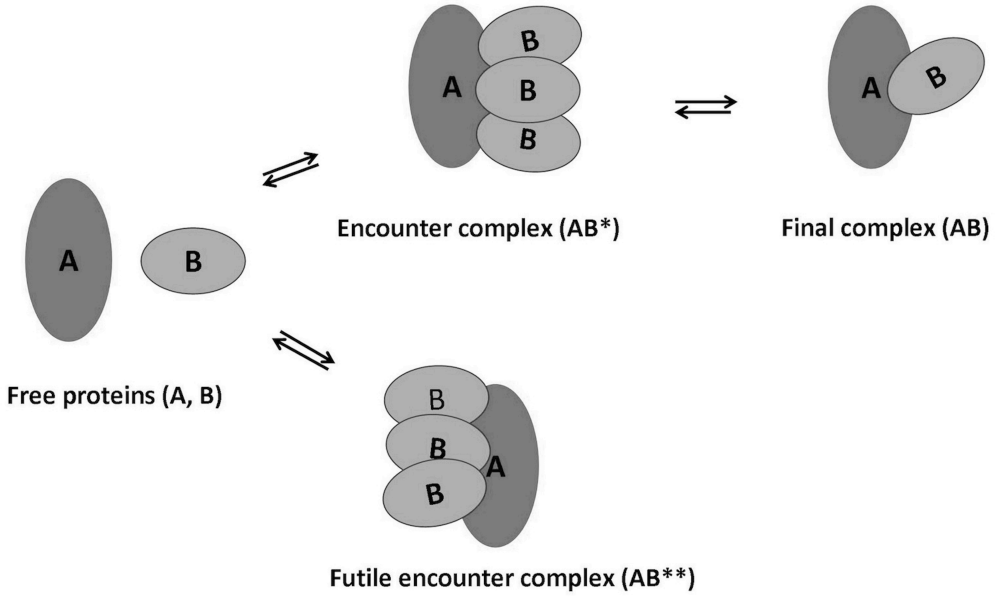

Figure 1.1. Two-step model of protein complex formation.[21] Free proteins (A and B) diffuse, and form either an encounter complex (AB*), which is in equilibrium with the final complex (AB) or a futile complex (AB**), which is in equilibrium with the free proteins.

As transient complex formation involves fast association and dissociation, proteins need to finely balance specificity and rapid turnover. When proteins meet by diffusion, they stay associated for some time, even in the absence of interactions. This is called the macrocollision. During the macrocollision, proteins have to find the target site on the counterpart surface to form the final complex. The target site of a protein represents only a small percentage of the overall surface (often <1%), so the search requires the sampling of an extended area, which takes time. The lifetime of macrocollisions is often insufficient to guarantee a productive search (futile encounter).[24] The formation of the encounter complex, consisting of an ensemble of multiple protein-protein orientations, optimises the number of productive collisions by increasing the chance of final complex formation via two mechanisms, both depending upon electrostatic forces. In the first mechanism, the encounter complex formation can decrease the dimensionality of the search. Electrostatic forces extend the lifetime of the macrocollisions, therefore the three-dimensional diffusion and search of the binding site is reduced to two-dimensional surface translations and reorientations (microcollisions). In the second mechanism, opposite charges pre-orientate the surfaces of the proteins in a way that the active sites face each other. The area to sample is thus drastically reduced and proteins can use the available time optimally to find the active site, thus increasing k_{on} up to four orders of magnitude. In this way, the formation of the encounter complex enables the proteins to orientate their interaction sites into a potentially fruitful position for interaction, via translational and rotational movements on the reciprocal surfaces. The formation of this intermediate leads to a stabilised active complex via the aforementioned short-range interactions. Consequently, the encounter complex is in equilibrium between the dissociated components and the specific state[25] and the equilibrium depends on the balance between non-specific long-range electrostatic forces and specific short-range forces. The population of the encounter state varies significantly among different complexes, dependent on the nature of the complex, so that it may correspond to a large fraction of the complex.[26-29] In highly dynamic systems, the encounter complex can form the larger population of the complexes, or the entire population is encounter complex.[30] Mutation of surface residues involved in the association of the complex, for example, those important for electrostatic interactions, can shift the equilibrium towards the specific or the encounter complex.[31,32] Moreover, such mutations can produce a fruitful or futile encounter complex.[24]

Therefore, the encounter complex must be explicitly considered in the study of protein complexes, in order to achieve a complete description of the interaction process.

Photosynthetic electron transfer proteins

The subject of this thesis is the transient complex formed by plastocyanin (Pc) and cytochrome *f* (Cyt *f*), two redox partners in oxygenic photosynthesis. Photosynthesis takes place at the thylakoid membrane of chloroplasts in plants, algae and cyanobacteria (Figure 1.2). It is a fundamental biochemical process,

through which light energy is converted into chemical energy, leading to the production of ATP and NADPH molecules. These compounds provide the energy and reducing power for biosynthetic pathways, including carbon assimilation.[33] Dioxygen is produced as a waste product. The photosynthetic machinery consists of a chain of membrane-bound protein complexes and mobile partners, which accomplish the long-range electron transfer process. When light hits photosystem II (directly or indirectly via the light harvesting proteins), the water splitting reaction is triggered and two water molecules are consecutively oxidized to generate four electrons, one oxygen molecule and four protons. Protons contribute to the membrane proton gradient, which is the driving force for the ATP production. The electrons are transferred through the membrane from photosystem II (PSII) to the cytochrome b_6f complex (b_6f) with the help of the lipid-soluble compound plastoquinone (PQ). Plastocyanin (Pc) is a soluble mobile electron carrier, which shuttles electrons from b_6f complex to photosystem I (PSI). Finally, electrons are transferred to ferredoxin-NADP reductase (FNR) via ferrredoxin (Fd) for the production of NADPH. The rapid ET in the photosynthetic chain requires a fine balance between the specificity of the interactions among different partners and the rapid turnover of the complexes, which is ensured by the transient interactions.

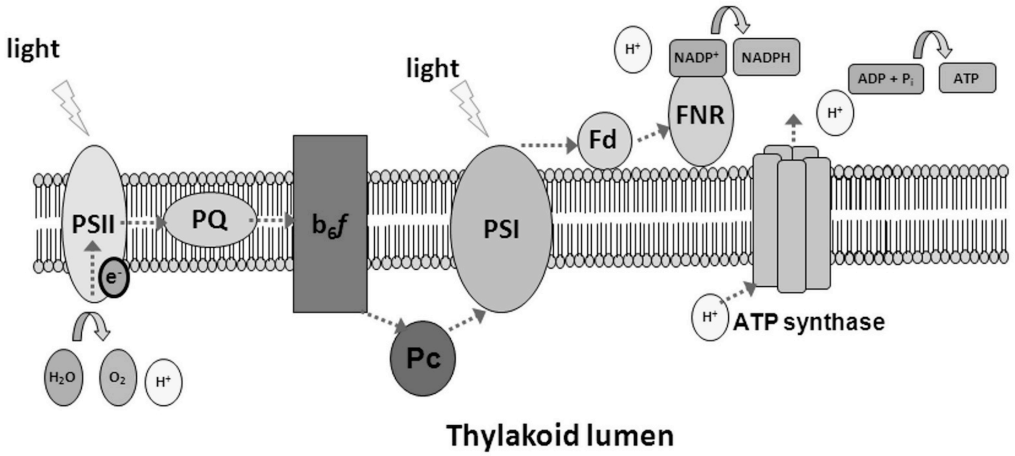

Figure 1.2. Schematic representation of electron transfer in the photosynthetic process.

In some cyanobacteria and green algae, the copper protein Pc can be substituted by the haem protein cytochrome c_6 (cyt c_6), which has the same function. Although the two proteins have no structural similarities, their interaction surfaces seem to be evolutionary correlated.[34] It has been proposed that the alternative expression of the genes coding for the proteins depends on the relative environmental abundance of copper and iron. In this way cyanobacteria and green algae may adapt to chemically different environments. In recent years, various new c-type cytochromes have been identified in plants (cyt c_{6A}), cyanobacteria (cyt c_{6B}) and algae (cyt c_{6C}).[35] They have structural features similar to cyt c_6 but their function has not been yet elucidated, though their low haem mid-point redox potential (E_m)[36] excludes a possible functional analogy with cyt c_6 and Pc.[37] The possibility of inter-exchange between Pc and cyt c_6 within the

photosynthetic chain underlines the importance of the transient and dynamic nature of the complexes formed by Pc and cyt c_6 with both b_6f complex and PSI.[34]

Plastocyanin

Pc was first isolated from the green alga *Chlorella ellipsoidea*[38] and then identified as electron carrier from Cyt *f* of the b_6f complex to the chlorophyll pair P_{700}^+ of Photo-system I (PSI).[39] It is a small, type I blue copper protein,[40] with a maximum UV absorption at 595 nm. The metal is coordinated in a trigonal pyramidal manner by the N_δ atoms of two histidines, one S_γ of a cysteine and one S_δ of a methionine and this motif is highly conserved among the different species (Figure 1.3).[41]

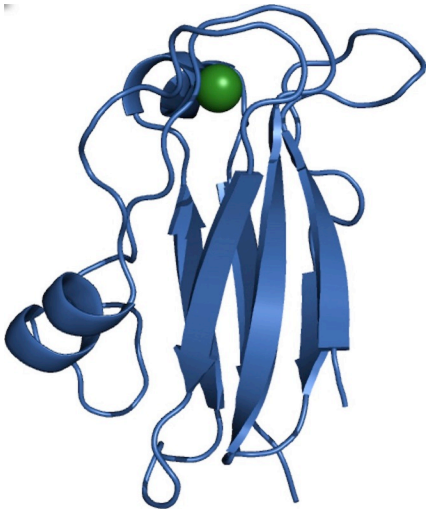

Figure 1.3. Plastocyanin from the cyanobacterium *Nostoc* (PDB entry 2GIM[42]). The backbone is represented as a blue ribbon. The copper ion is shown as a cyan sphere. This image and others of molecular structures were made with Discovery Studio Visualizer 2.5 (Accelrys®).

More than 100 genes in GenBank(www.ncbi.nlm.nih.gov/genbank) are clearly labeled as *pet*E-plastocyanin, from as many sources. Among these organisms the amino acids chain length of the mature protein varies from 97 to 105, from which the longer chains belong to cyanobacteria. Pcs have a molecular weight of 11 kDa and a highly conserved rigid β-sandwich structure[43] composed of seven β-strands and one irregular strand containing a small helix, mainly found in cyanobacteria. The metal ion and the proximal histidine ligand are surrounded by 25 amino acids that form an important region for electron transfer and protein-protein interactions, named site 1 or the hydrophobic patch,[44] which contains 16 hydrophobic residues and for which some small dynamics were observed.[45] The hydrophobic patch is present in both eukaryotic and prokaryotic Pcs, being a more extensive feature in the latter.[18] In fact, the hydrophobic patches in cyanobacterial Pcs comprise residues with long aliphatic chains, such as leucines and methionines, which make the overall region more expanded than in plant Pc. Another important feature of this protein is the presence of a ridge of charged amino acids at the side of the reaction centre, known as site 2. The nature of this site varies between eukarya and cyanobacteria. It is involved in the interaction with redox partners and influences the final orientation of the protein within the complexes it forms.[32,46-50] In plants, the acidic character of the site 2 surface exposed residues creates a negatively charged area, which interacts with a corresponding positively charged region on the small domain of Cyt *f*. In cyanobacteria, the charges are reversed with a predominance of basic exposed residues at site 2 in Pc and negatively charged amino acids in Cyt *f*.

Cytochrome *f*

Cyt *f* is the largest component of the b_6f complex.[51,52] It is embedded in the thylakoid membrane by a single C-terminal transmembrane α-helix, while the soluble part is exposed in the intraluminal space where the interaction with Pc occurs. Soluble Cyt *f*, being 254 amino acids, was first crystallized from turnip leaves (*Brassica rapa*)[53] and the structure revealed that it consists of a 28 kDa β-sheet protein with an elongated shape organized in two domains, named small and large domains. It belongs to the *c*-type cytochromes family because the haem is covalently attached to two cysteines in the conserved motif CXXCH of the protein through thioether bridges. The fifth ligand of the iron is an axial histidine and the sixth ligand of the iron is the N-terminus of the protein (Y1).[53] The presence of the haem provides the characteristic absorption profile with a Soret band at 420 nm and a α-band with a maximum at about 556 nm, in the reduced form of the protein. The haem iron in the oxidized state represents a resident paramagnetic source, that can be used for NMR studies as it gives rise to intra- and inter-molecular pseudocontact shifts, from which distance information from the iron to the observed protons can be obtained.[46] The *c*-type haem is located in the large domain, close to the hinge region with the small domain (Figure 1.4).

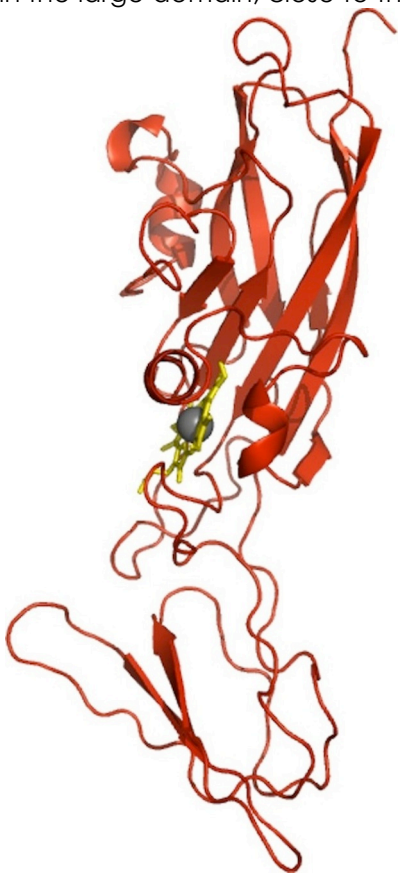

Figure 1.4. The structure of the soluble part of Cyt *f*, derived from the b_6f complex crystal structure from the cyanobacterium *Nostoc* (PDB entry 2ZT9). The backbone is shown as red ribbons, the haem in yellow stick representation and the iron as a grey sphere.

The porphyrin ring is located underneath a hydrophobic patch, which participates in the reaction with Pc. The contact between the hydrophobic surface on Cyt *f* and the corresponding one in Pc creates a hydrophobic interface for the ET. Cyt *f* is overall very negative, but in the small domain a charged region is present, with a composition that varies between different species, but generally is of a charge that is opposite to Pc. Therefore, it is positively charged in plants and negatively charged in cyanobacteria. The surface charge complementarity between the small domain area on Cyt *f* and site 2 on Pc plays an important role in the final orientation of the proteins with respect to each other. Hence, the architecture of the complex can be distinguished as "head on", as in the cyanobacterium *Phormidium laminosum*,[54] where Pc participates in the interaction with only the hydrophobic patch due to the lack of basic amino acids at site 2, and "side on", in plants[46-48] and in the cyanobacteria *Nostoc*[49,50] and *Prochlorothrix hollandica*,[32] where electrostatic attractions also

occur between the aforementioned sites, orientating Pc toward the small domain of Cyt *f*.

Pc-Cyt *f* complex

The Pc-Cyt *f* complex represents an excellent and challenging system for studying the influence and the contribution of different forces in ET, binding kinetics and in the structure of the complex.[55] In fact, even though the three dimensional structure of the interaction partners is highly conserved, the surface properties vary significantly in the different species.
In plants and green algae, Cyt *f* presents a ridge of positively charged amino acids, mainly composed of lysines. The electrostatic features within the complex have been the focus of extensive studies, due to the role they play in the association and in the function of the complex itself. The ET pathway from Cyt *f* into Pc was initially debated. Two main routes were proposed, via H87 (numbering of spinach Pc), which is a direct ligand for the Cu, and via Y83, in which case the electron tunneling is enhanced by a passage through C84 (which also coordinates the Cu). The ET pathway from Cyt *f* into Pc was initially debated. Two main routes were proposed, via H87 (numbering of spinach Pc), which is a direct ligand for the Cu, and via Y83, in which case the electron tunneling is enhanced by a passage through C84 (which also coordinates the Cu). The latter pathway indicated a cation-π interaction between K65 on Cyt *f* and Y83 on Pc as possible mechanism for ET.[56] A Monte Carlo approach, based on the electrostatically most favorable complexes,[57] suggested the initial formation of a high binding affinity complex in a configuration, which favored the interaction between these residues followed by a rearrangement that promoted the intermolecular ET through H87 instead of Y83. The role of electrostatics was thought to enhance the formation of the encounter complex, which would give rise to the ET-active complex. Subsequent Brownian Dynamics-based docking studies supported the ET pathways between Y1 on Cyt *f* and H87.[58] The first solution structure of the complex, published in the same year, indicated that ET must occur predominantly via H87.[46]
The transient nature of the Pc-Cyt *f* complex appears to impair the co-crystallization of the two proteins because until today no crystal structures of the complex are available. However, a plant complex was investigated via NMR using pseudocontact shifts[46] and diamagnetic chemical shift perturbation analysis.[47,48] The large size of the chemical shifts observed for the residues belonging to the hydrophobic patch on Pc, defined it, in terms of specificity, as the central binding site, suggesting a specific interaction interface with at least partial exclusion of solvent molecules. In contrast, the modest size of the shifts of the amino acids in the charged areas of the interaction partners suggested a charge-charge interaction in which the residues do not undergo desolvation. The complex was shown to exist predominantly in a state with "side on" orientation (Figure 1.5A) promoted by Coulombic attractions from the negatively charged Pc and the positively charged Cyt *f*. The role of electrostatics in the orientation of the proteins matched with kinetic measurements on Pc[59] and Cyt *f* mutants.[60] Single or bulk mutations of charged residues important for the overall electrostatic

potential of the two proteins underlined the contribution of long range electrostatic forces in the formation of the encounter complex and consequently in the overall ET rate.[61]

Further NMR investigations on the side-chain interactions validated these results and enabled the deduction of the roles of electrostatic interactions in the molecular recognition step and the contribution to the binding affinity. The complex showed the same behavior even at different ionic strength and pH conditions.[62] On the other hand, *in vivo* studies on *Chlamydomonas reinhardtii*[63,64] had opposing outcomes. The ET rate was measured for several Cyt *f* mutants, which concerned specific residues of the positively charged site, and no influence of electrostatics was observed, while *in vitro* kinetics experiments showed a monotonical decrease in the oxidation of Cyt *f* with dependence on ionic strength.[65] *In vivo* the intralumen ionic strength is higher than the one used for the *in vitro* experiments, corresponding to ~350 mM,[63] a value at which electrostatic contribution is limited. Furthermore, the volume of mature thylakoids is small and long-distance Pc diffusion is not required, contrary to situation of initial chloroplasts formation, before the membranes are enclosed, when the thylakoidal space is wider. The contradiction between *in vitro* and *in vivo* studies resides in the different composition of the media where Pc and Cyt *f* were studied, and in the latter case, the concomitant presence of thylakoidal elements of diverse nature, could somehow mask the effect of electrostatics.[62]

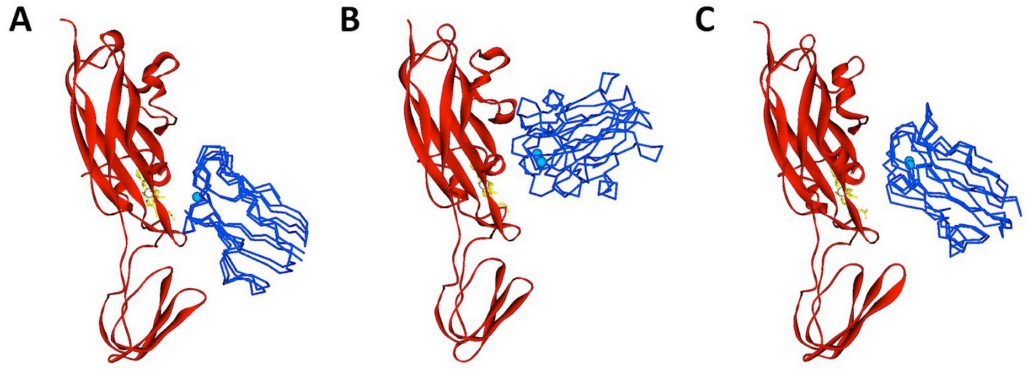

Figure 1.5. Orientations of Pc (shown as blue Cα traces) in the complex with Cyt *f* (shown as red ribbon) from different organisms: "side on" in plants (A) (PDB entry 2PCF[46]) and in the cyanobacterium *Nostoc* (C) (PDB entry 1TU2[50]), and "head on" in the cyanobacteria *Phormidium laminosum* (B).[54]

Also the complex from *Nostoc sp*. PCC7119 revealed to be electrostatically stabilised by complementary localised charges. Within this structure, Pc was found in "side-on" orientation (Figure 1.5C),[50] despite the surface charges being reversed in this organism, with a highly basic character of the copper protein (pI=8.8) compared with its eukaryotic analogues (pI=3.8 for spinach Pc). In this case, complex formation is heavily dependent on hydrophobic and electrostatic interactions in both the final state and the encounter complex.[49] The electrostatic nature of the interaction was shown by kinetic experiments on Pc[8] and Cyt *f* mutants.[66] Mutations offsetting or inverting the electrostatic charges in the charged areas and in the hydrophobic patch of both proteins highlighted the

importance of electrostatics in the complex formation and hydrophobics in ET. The ionic strength dependence of the rate constant is mainly determined by the pronounced positive charges of Pc, which are also responsible for the specificity of the interaction.[66]

Pc from *Ph. laminosum* presents a more symmetric charge distribution compared to plant and *Nostoc* orthologs. Positively and negatively charged amino acids are not found in patches but are balanced. Thus, Pc has more neutral character, with a pI= 5.0. As a consequence, the complex showed a weaker affinity (in the millimolar range, in contrast with the micromolar range in *Nostoc*) and a moderate dependence on electrostatic interactions in complex formation. Thus, electrostatic interactions only confer a minimal contribution to the final organisation of the complex as indicated by the NMR structure, [54] with Pc in a "head on" orientation (Figure 1.5B), and solely the hydrophobic patch goes to form the interaction surface with Cyt *f*. The fit of PCS data did not converge into a single structure, indicating a highly dynamic complex and suggesting that the encounter complex must have a decisive role in leading the reduction of Pc by Cyt *f*. Although the electrostatics do not appear to be important for the orientation of the proteins within the complex, kinetics[67] and theoretical[68] studies pointed to an electrostatic influence on the association process (effects on k_{on}). Also in *Ph. laminosum* Pc shows the highest dependence on electrostatics as well as a bigger influence on specificity.[69] Mutagenesis studies involving the N-terminal hydrophobic residues of Cyt *f*,[70] responsible for the shielding of the haem from the solvent, indicated that hydrophobic interactions have a role in the encounter complex formation and showed the prominent role of Cyt *f* in the fast ET rate observed at high ionic strength. The explicit treatment of the non-polar desolvation (hydrophobic) force together with the electrostatic force on the *Nostoc* and *Phormidium*[71] complexes highlighted a stronger contribution of electrostatic interactions to the overall ET rate in *Nostoc*, due to their contribution in the encounter complex formation. The role of hydrophobic interactions contributed most to the formation and stabilization in the *Phormidium* complex instead.

The complex from the cyanobacterium *Prochlorothrix hollandica* emerged to be highly dynamic as well, with a "side on" orientation.[32] The role of electrostatic interactions in the complex formation was established by observing a decrease in the affinity with increasing ionic strength, although the influence is less prominent than in plants and *Nostoc*.

The different orientations of Pc and Cyt *f* within the various complexes seem to depend on both the long-range forces involved in the intermolecular recognition and the short-range forces important for the formation of the final complex. A surprising variation in the structures and contributing interaction forces has been observed among different species. Also the degree of dynamics in the various complexes appears to differ.

NMR methods for studying transient complexes

^{15}N labeling of proteins is a well-established technique that allows to perform two-

dimensional ^{15}N-^{1}H heteronuclear single-quantum correlation experiments (HSQC). The HSQC spectrum could be considered as a two-dimensional picture of the protein. Each HSQC cross-peak correlates the chemical shift of the amide proton with the chemical shift of the attached ^{15}N amide nitrogen of a certain ^{15}N labeled amino acid residue of the protein, in which is possible to observe and quantify changes in the position and in the intensities of the HSQC resonances due to interaction with one or more partners.
Solution NMR spectroscopy is an excellent technique for the characterization and structural determination of transient complexes. In the fast exchange regime on the NMR time scale, a signal is an averaged representation of all species in solution, but it is still a powerful tool to study dynamics within a transient complex. The NMR techniques used in this thesis are diamagnetic chemical shift perturbations (CSP) and paramagnetic NMR, comprising pseudocontact shift (PCS) and paramagnetic relaxation enhancement (PRE).

Diamagnetic chemical shift perturbations

CSP occurs as a result of changes in the chemical environment of the observed nucleus. When a protein interacts with its partner, the residues involved in the interaction will usually experience a change in their chemical environment, which will cause a shift of the respective NMR resonances. CSP analyses are commonly performed on ^{15}N-^{1}H HSQC spectra of the ^{15}N labeled protein recorded in the absence and presence of increasing amounts of the unlabelled partner, or vice versa, in which the labeled protein is titrated into the unlabelled partner. In this way, it is possible to obtain a map of the amino acids involved in the interaction and visualise the binding sites. Furthermore, the analysis of chemical shifts as a function of protein concentration provides information both on stoichiometry of the complex and the binding affinity. The size of the shifts is also related to the dynamics of the complex; large chemical shift perturbations indicate well-defined, short-range interactions, whereas small ones denote high dynamics and weaker interactions.[32,72,73] Short range interactions are involved in the formation of the final complex, which is stabilized in a single orientation through a network of H-bounds, salt bridges and van der Waals interactions. Also changes in the solvation of the protein occur. Together, these interactions cause large CSPs. According to current model, the encounter complex forms by long range electrostatics interactions and the proteins are still largely solvated, so that the chemical environment at the binding sites undergoes minimal alterations and the size of perturbations is small. As a consequence, average chemical shifts values also reflect the population of the encounter state[9,30,72] and therefore can be also used as a reliable diagnostic tool for the dynamics within transient protein-protein complexes.[72-74]
The information gained from chemical shift perturbations analysis mainly has a qualitative value, therefore other NMR methods, such as paramagnetic NMR, are required for detailed structure determination.

Paramagnetic NMR

The presence of a paramagnetic centre in a molecule or in the solvent causes distance-dependent effects on the NMR signals of observed nuclei, which can provide very precise structural information. The paramagnetic source can be intrinsic to the protein, for example, the Fe^{III}-haem of Cyt *f*, or can be inserted into the protein via site-directed spin labeling (SDSL). In the latter case, a paramagnetic probe is linked to a cysteine residue that either is present on the protein or introduced by site-directed mutagenesis. A paramagnetic probe is a small molecule containing an unpaired electron on a metal chelating tag[75-79] or a stable organic radical.[80-83] The chemical-physical properties of the paramagnetic centre determine the nature of the effects on the NMR spectrum. Anisotropic electron g-tensors give rise to direction dependent effects such as residual dipolar coupling (RDC), associated to paramagnetic alignment, and PCS, together with PRE. Isotropic electron g-tensors only cause PRE.[84] PCSs arise from through-space dipolar interactions between the (time-averaged) magnetic fields of the paramagnet and the observed nucleus, which cause additive paramagnetic shifts of the observed NMR signals. The PCS effect is proportional to the inverse third power of the distance (r^{-3}) between the unpaired electron and the observed nucleus.[85] Therefore, PCSs can be measured for residues that are far from the paramagnetic centre.[86] up to 60 Å.[87] In the case of PRE, the longitudinal electron spin relaxation as well as the dipolar coupling cause nuclear relaxation,[88] which gives rise to line broadening of the nuclei in its proximity.[84] PRE shows r^{-6} distance dependency, hence the peaks corresponding to amino acids close to the paramagnetic centre will exhibit a decrease in peak height or will disappear entirely from the spectrum. PREs can be very strong and can affect nuclei up to 35 Å from the paramagnetic centre.[84]

When a complex is studied, one of the interaction partners contains the paramagnetic centre and PCSs, or PREs, are observed on the isotopically labelled counterpart monitored in the NMR spectrum.[81] This approach has been extensively used for intermolecular structure determination, taking advantage of both PCSs[46,50,54,87,89-91] and PREs.[26,77,92-97] PCS has been the method used for the determination of the structures of Pc-Cyt *f* complexes from several plants,[46,48,62] *Ph. laminosum*,[54] *Nostoc*[98] and *Pr. hollandica*.[32] It showed to be appropriate for the visualization of the complex in the highest populated state (AB in Figure 1.1). However, considering the high dynamics within this complex, as suggested by NMR CSP studies and kinetic data, PRE measurements on this complex will certainly provide a more complete picture of the overall complex. The PRE method is an exquisitely sensitive technique to detect lowly populated states of proteins orientations, provided the paramagnetic effect is much larger for the lowly populated state than the ground state.[92] This is not the case of the haem derived PCS, so minor states are not easily detected using only these shifts. The introduction of a paramagnetic source in different locations on the Cyt *f* surface allows to sample the whole surface and to visualise all the possible locations of Pc.

Computational methods for studying protein complexes

Computational methods are essential for determining protein structures and studying the association of protein complexes, either in combination with experimental methods or in a purely theoretical manner. Furthermore, due to the limitations of experimental techniques, they represent a primary tool for the visualization of the encounter complex.[99,100]

In many structure determination studies, experimental restraints, such as those gained from X-ray crystallography, SAXS or NMR, are used as inputs for docking software packages, some of which are based on molecular dynamics procedures. The structures of plant,[46,62] *Ph. laminosum*[54] *Nostoc*[50] and *Pr. hollandica*[32] Pc-Cyt *f* complexes were obtained by rigid-body docking of the proteins, on the basis of experimentally obtained paramagnetic restraints and chemical shift perturbations. The experimental restraints were the driving force for the association of proteins, rigid-body dynamics was applied at each docking step and the structures were saved according to the energy values calculated for each complex. The same concept is applied in the ensemble docking procedure. In this case an ensemble of a variable number of elements is concurrently docked to the interaction partner, which is considered as single and fixed object. This method has been successfully employed to study intermolecular motions[101,102] and for the visualization of various kinds of encounter complex, such as DNA-protein[103,104] and protein-protein complexes.[27,92,105] This method is discussed further in Chapter III.

Purely theoretical approaches have supported both structural and functional experimental studies and in many cases they have enriched the understanding of ET pathways[56-58] and ET rates[61,68,71,106-108] in the Pc-Cyt *f* complex. Brownian dynamics (BD) and Monte Carlo (MC) simulations represent the most widespread methods used for studying protein-protein interactions. A mobile diffusing protein is docked to a target molecule under the influence of an electrostatic field and Brownian motions[109] or Monte Carlo sampling,[57] respectively. Both techniques allow the screening of a large number of conditions in a relatively short time. In this way, different mutations or ionic strength effects can be included in the simulations providing an extensive data set, which is hard to achieve experimentally. Assuming that the formation of the encounter complex is mainly driven by electrostatics, those methods allow to study the different configurations of the encounter intermediates. The combination of experimental NMR data and Monte Carlo simulations was successfully employed in the visualization of the encounter complex formed by another pair of ET proteins, cytochrome *c* and cytochrome *c* peroxidase, for which also the crystal structure had been obtained.[110] This encounter complex has proven to be absolutely dependent on electrostatic forces, in agreement with the theoretical predictions.[28] Furthermore, mutagenesis of the interface residues involved in the interaction, caused a change in the equilibrium between encounter and final state populations.[31]

Thesis outline

This thesis focuses on the characterization of the dynamic aspects of the transient complex of Cyt *f* and Pc in order to elucidate the nature of the mechanisms behind protein complexes formation.
In chapter II the Cyt *f*-Pc complex from *Nostoc* was characterized by PRE NMR spectroscopy, revealing a high degree of dynamics within the complex. The influence of spin label in the binding was also determined by chemical shift perturbations analysis.
Chapter III describes the visualization of the encounter state of the *Nostoc* Cyt *f*-Pc complex achieved through the combination of PRE NMR spectroscopy, Ensemble docking and Monte Carlo simulations. The role of hydrophobic interactions in the encounter complex formation has been experimentally demonstrated and a new model for protein complexes formation has been proposed.
The effect of electrostatic interactions on dynamics of the complex has been studied in the cross complex formed by *Nostoc* Cyt *f* and *Phormidium laminosum* Pc and it is presented in Chapter IV.
Finally, the results of this thesis have been contextualized in terms of encounter complex and its role in the association of ET systems (Chapter V).