Transient complexes of haem proteins

Volkov, O.M.

Chapter I

Introduction

Protein complexes: static or transient

Interactions between proteins mediate most of the cellular functions. The variety of such interactions is staggering: dissociation constants of protein-protein complexes span a range of twelve orders of magnitude[1]. At one extreme of this continuum are static complexes – characterized by tight binding, long life-times, and high specificity – like those formed between antigens and antibodies or enzymes and inhibitors. The complex of barnase, an extracellular RNAse from *Bacillus amyloliquefaciens*, and barstar, an intracellular inhibitor of barnase, is a typical example of the static complexes. The low dissociation constant (K_D) of 10^{-14} M and, consequently, a long life-time of this complex are necessary for shutting down a potentially devastating intracellular activity of barnase[2]. Furthermore, as the two proteins have co-evolved to selectively recognize each other in order to avoid potentially deleterious interactions with other cellular components, the barnase-barstar complex is highly specific[3].

At the other end of the protein interactions' scale are weak, short-lived complexes formed by the proteins that recognize multiple partners. As a rule, transient complexes orchestrate biochemical transmission processes, such as those taking place in electron transfer (ET) chains or signal-transduction cascades. Triggered by the interest in intermolecular ET, the studies of interactions between redox proteins provide most of the current knowledge on the transient complexes (for reviews see ref. 4 and 5).

Transient protein complex formation

In order to maintain a high turn-over in the ET or signal transduction processes, the interactions between the proteins involved occur transiently, resulting in formation of complexes with the lifetimes on a millisecond scale. Consequently, the dissociation rate constants (k_{off}) of transient complexes are high ($k_{off} \geq 10^3$ s^{-1}). Together with the high association rate constants (k_{on}) – varying from 10^7 to 10^9 $M^{-1}s^{-1}$ (see ref. 4 and references therein) – the large k_{off} rates convert into the K_D values in mM – μM range ($K_D = k_{off}/k_{on}$). This explains why transient interactions are weak. As k_{on} rates for transient complexes are

as high as those for the static ones – compare the above k_{on} values to those for the complexes of acetylcholinesterase and fasciculin-2 (10^7 $M^{-1}s^{-1}$)[6] or barnase and barstar (10^9 $M^{-1}s^{-1}$)[7] – it appears that most of the differences between the two types of protein interactions are determined solely by k_{off}.

In contrast to the closely-packed interfaces of static protein complexes[8], the interfaces of transient complexes are packed poorly in order to promote high k_{off} rates[9]. The low geometric complementarity of the proteins involved in transient interactions might further stem from the necessity to recognise several binding partners, which enables these proteins to act as messengers between biomolecules. Indeed, high optimisation of the binding site towards one partner would inevitably lead to the reduction of the affinity for another partner, thereby effectively inhibiting the turnover[9].

The promiscuity of transient protein-protein interactions, *i.e.* the ability to bind to more than one partner, poses an interesting dilemma. On one hand, the specificity of the interaction should be low enough to allow a protein to recognize several structurally distinct partners but, on the other hand, the specificity should be sufficiently high so that the protein interacts exclusively with its partners, avoiding all other cellular components. In fact, these two factors are finely balanced, so that only a narrow margin of specificities towards different partners is maintained, enabling an effective functioning of the transient complexes[4,10]. There is no simple relationship between affinity and specificity of transient protein-protein interactions[4], yet the two are clearly connected; in general, stronger complexes are often more specific and vice versa.

The reduced specificity of the transient interactions has an interesting ramification. If two proteins are separated *in vivo* – *e.g.* come from different organisms or are localized in different cellular compartments – but possess just enough specificity for each other – provided, for instance, by opposite overall charges or matching hydrophobic patches on protein surfaces – they can form a non-physiological protein complex *in vitro*. Such an unusual property, peculiar to the proteins involved in transient interactions, has been of great value to the investigation of interprotein ET that has often used non-physiological complexes as experimental and theoretical models (see below).

The reduced specificity is further manifested by the increased protein dynamics in transient complexes. As the partner proteins are only moderately optimised for each other, a number of nearly isoenergetic binding orientations might exist within one complex.

Therefore, the proteins may need to thoroughly explore each other's surface in search of the most favourable binding geometry. The extended conformational search of the optimal protein-protein orientation implies the enhanced protein mobility within transient, as compared to static, complexes. The extent of the search varies among different transient complexes: while some binding partners occupy a single orientation for most of the time (*e.g.* Cyt *c* – CcP, Chapters III-IV) , others explore a wide area around each other (*e.g.* Cyt *c* – Cyt b_5, Chapter II) or even sample a continuum of binding sites (*e.g.* Cyt b_5 – myoglobin[11,12]). Thorough understanding of dynamic processes within transient protein complexes is aided by the concept of the encounter complex, which is outlined in the following section.

Encounter complex

Diffusion paradox

Diffusion of the reactants is often the rate-limiting step of the protein-protein association. For two spherical proteins of 18 Å radius that form a complex upon every collision regardless of the orientation, the maximum diffusion-limited association rate constant, $k_{on} = 7 \times 10^9$ $M^{-1}s^{-1}$, is given by the Smoluchowski equation[13]. The specificity of the intermolecular interactions imposes an additional constraint: a successful association of two proteins requires the reactive patches on the surface of each partner to be properly aligned. As an average binding interface constitutes only 10 – 15 % of the total surface area available for the interaction[14], the probability that a random molecular collision will result in a productive protein-protein orientation is small, implying a 10^3 – 10^4-fold decrease of the Smoluchowski rate constant[4,7,15]. However, for a number of protein complexes – amongst which are barnase and barstar[7], Cyt *c* and CcP[16], and Cyt *c* and Cyt b_5[17] – the observed bimolecular association rates are as high as 10^9 $M^{-1}s^{-1}$. The paradoxical discrepancy between predicted and observed k_{on} values has stimulated the development of the "encounter complex" concept for accurate description of macromolecular binding in solution. Before discussing the theoretical framework of this concept, we need to define the terms used.

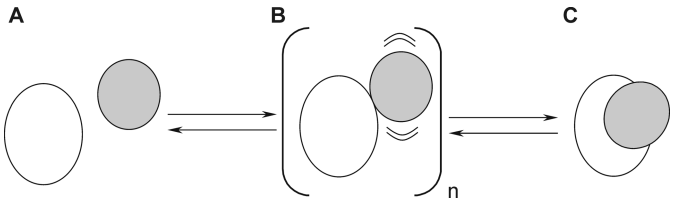

Figure 1.1. **Model for the formation of a protein complex.** Free proteins (**A**) enter a dynamic *encounter state* (**B**) composed of diverse *encounter complexes*, each comprising multiple protein *collisions*, some of which lead to the formation of a specific, *single-orientation complex* (**C**). The combination of (**B**) and (**C**) constitutes the *protein complex*. The terms in italic are defined in the text.

Definition of the encounter complex

There is much confusion in literature concerning what constitutes an encounter between proteins. According to Adam and Delbruck[18], two proteins form an encounter complex when they occupy the same solvation shell as a result of a diffusion-driven collision. A more precise definition is given by Northrup and Erickson[15], who discriminate between a *collision*, which is defined as taking place when a protein pair attains a specified small distance from one another (often taken to be equal to one hydration radius, *ca.* 2 Å), and an *encounter*, which is the complete set of interactions from the initial collision to the eventual separation of the protein molecules. We would like to refine the latter definition by discriminating between an *encounter complex*, which includes multiple protein collisions, and a specific, *single-orientation complex* formed as a result of some of these collisions (Figure 1.1). The combination of the two we will call a *protein complex*. Furthermore, we would like to define a dynamic *encounter state* as the combination of all encounter complexes, irrespectively of whether they lead to the formation of the single-orientation complex or not.

Theoretical models of protein-protein association

As mentioned above, the association rate constants observed for some protein interactions are three to four orders of magnitude higher than what is expected for a

diffusion-controlled process. For a long time it has been recognised that some additional forces must be responsible for accelerating the macromolecular association in solution. Back in 1968, Adam and Delbruck postulated "reduced-dimensionality" principle[18], according to which a shared solvation shell keeps interacting macromolecules in check and enables a 2D, rather than a 3D, diffusion along a protein surface, with the ensuing reduced-dimensionality search accelerating the association. This principle, extended to the case of 1D sliding, has been successfully used to explain the rate enhancements of protein-DNA interactions[19,20] and has stimulated a research into the mechanisms of protein-protein association.

The first mechanism to be proposed for proteins was the "lengthy collisions" mechanism of Sommer *et al.* purporting to account for the high homodimerization rates observed for a variety of proteins in solution[21]. It suggested that the interacting partners form a weakly bonded and rotationally non-specific encounter complex, termed "lengthy collision", that enables the proteins to rotate over each other's surface in search of the reactive regions. The duration of the collision was estimated to be 10 – 100 μs, while the dissociation constant for this weak interaction was expected to be close to 0.1 mM. The hypothesis relied on non-specific distance-dependent attractive potential to hold the molecules together during the lengthy collision; however, no physical meaning for this potential was provided.

These ideas were further developed by Berg, who, drawing on analogy of a protein sliding along a DNA strand, has suggested a surface sliding model in two dimensions for the protein-protein association[22]. This model proposed that the proteins sustain multiple collisions in repeated encounters, rather than one lengthy collision. While the individual encounters may be short-lived, their ensemble – *i.e.* the encounter state, see the definition above – has a lifetime of 10 – 100 μs, in agreement with the previous study[21]. Furthermore, the suggested mechanism did without an empirical attractive potential of the earlier model. Berg noticed that upon dissociation of an unproductive encounter complex, the dissociated proteins would remain spatially arranged. In other words, the protein would remember where it was on its partner's surface. Such molecular memory – peculiar to the surface diffusion and absent in the linear sliding – would substantially increase the chance of the second productive collision, thereby enhancing the overall association rate[22].

These earlier qualitative models have been finalized by Brownian dynamics simulation study that has provided a quantitative description of the rate enhancement mechanism[15]. It has shown that the rate enhancement of association of two neutral spherical proteins can be attributed to a diffusive entrapment effect, in which a protein pair surrounded and trapped by water undergoes multiple collisions with rotational re-orientation during each encounter[15]. An average encounter complex – with a lifetime of 6 ns, which is on the order of the rotational correlation time of an average-size protein – was shown to consist of 9 collisions. This simulation protocol, allowing for an accurate prediction of association rates for protein-protein interactions, was later refined by taking into account the partial desolvation of the intermolecular interface, which was shown to stabilize and accelerate the interaction[23,24].

When applied to the interaction between oppositely-charged proteins, the above model – developed for the association of neutral macromolecules – takes into account Coulombic attraction, which plays an important role in protein-protein interactions[25]. It has been shown that, while desolvation is often the main factor enhancing the association of weakly-charged molecules[23,24], electrostatic steering greatly enhances k_{on} and often determines the specificity of complex formation between highly-charged proteins[7,23,26,27]. As most proteins that form transient complexes are highly-charged[4,5], the electrostatic contribution is expected to dominate transient interactions in solution.

Transient ET complexes of haem proteins

Biological ET is mediated by the proteins that both insulate the redox centres and recognize the interaction partners[9,28]. Among the cofactors capable of a controlled redox behaviour is the haem, the complex of iron and protoporphyrin IX. In addition to its excellent catalytic qualities (employed, for example, in cytochromes P450 or cytochrome *c* peroxidase) and unique coordination behaviour (used for O_2 and CO_2 transport by myoglobin and haemoglobin), the haem possesses valuable physico-chemical and biological properties – among which are high stability, bio-availability of the metal, relative ease with which its redox potential can be modulated by various substituents, *etc.* – that make it a frequent cofactor in ET proteins.

The haem proteins are found in most biological processes that rely on ET, such as oxidative phosphorylation, photosynthesis, and manifold metabolic conversions[28]. As the diversity of the ET complexes of haem proteins is far beyond the scope of this work, the discussion is limited to Cyt *c* – Cyt b_5 and Cyt *c* – CcP, two haem protein complexes investigated in this study, and a more inquisitive reader is referred to the standard biochemistry texts and recent reviews[4,5,29-31].

The proteins

Cyt *c*

Cytochrome *c* (Cyt *c*) has been discovered[32] and isolated[33] in late 1920s by David Keilin, who recognized it to be a principal component of the eukaryotic respiratory chain. Since then the sequences of Cyt *c* from more than 100 different species have been reported[34]. The primary sequence of this protein is highly conserved among eukaryotes, where Cyt *c* functions as an electron carrier between the membrane-bound cytochrome *c* reductase and cytochrome *c* oxidase, two components of the mitochondrial oxidative phosphorylation ET chain[35]. In yeast, Cyt *c* has other physiological partners, such as cytochrome b_2 – also known as lactate dehydrogenase – and cytochrome *c* peroxidase (CcP)[34]. Thanks to its stability, availability, and attractive physico-chemical properties, yeast *iso*-1-Cyt *c* has been studied more that any other ET protein.

Cyt *c* is a positively-charged (the charge is +6 at pH 6.0 as estimated from protein titration curves[36]), low molecular weight (12.1 kDa) protein, consisting of 108 amino-acids[i] and the haem prosthetic group. The protein can be readily expressed and purified from *Saccharomyces cerevisiae* (baker's yeast)[38-40] or *Escherichia coli*[41]. In the latter expression system, *CYC1* and *CYC3* genes – encoding, respectively, for Cyt *c* and yeast Cyt *c* haem lyase, the enzyme that catalyses covalent insertion of haem into Cyt *c*[42] – are co-expressed, providing good yields of the recombinant protein[41].

[i] The amino-acid numbering adopted in this thesis is based on the sequence alignment with horse heart Cyt *c*[37] that generates a negative numbering for the first five residues (T-5, E-4, F-3, K-2, and A-1, followed by G1).

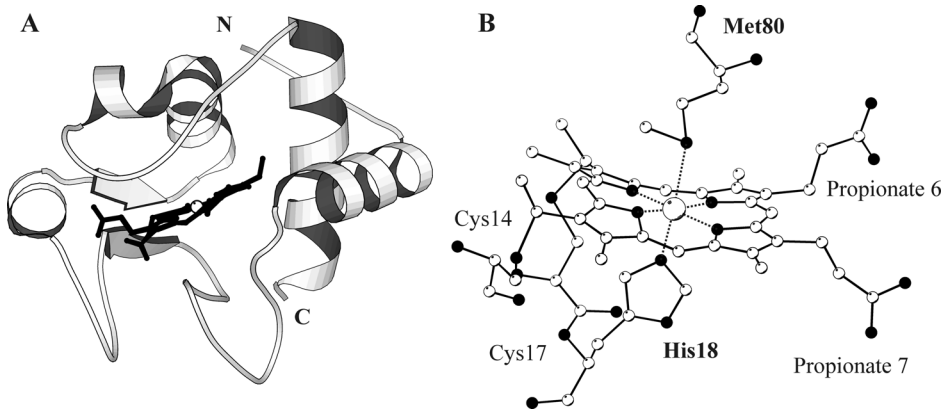

Figure 1.2. **Three-dimensional structure of Cyt *c*.** (**A**) Overall fold of the protein; the haem is shown in sticks, and the protein termini are indicated by the labels. (**B**) The view of the haem binding site showing the axial ligands, His 18 and Met 80 (bold labels), and two cysteine residues, Cys 14 and Cys 17, covalently bound to the haem group; haem propionates are indicated. The molecular coordinates were taken from the crystal structure of the oxidized Cyt *c* (PDB entry 2YCC[43]). This figure and Figures 1.3 – 1.6 were generated with MOLSCRIPT[44].

Cyt *c* has a nearly spherical shape and is formed by five α-helices and a short β-strand (Figure 1.2 A), the overall fold that is highly conserved across the protein family[45]. Cyt *c* contains a *c*-type haem group that is located near the N-terminus (Figure 1.2 A) and is attached to the polypeptide chain by covalent thioether bonds with two cysteine residues from the canonical C-X-X-C sequence (Figure 1.2 B). The haem contains a low-spin iron ion that has two physiologically relevant oxidation states, Fe(II) and Fe(III), and is diamagnetic in the reduced and paramagnetic in the oxidized form. In Cyt *c*, the haem iron is six-coordinated, with the coordination sphere consisting of four pyrrole nitrogens of the haem, $N_{\varepsilon 2}$ atom of His 18, and S_{δ} atom of Met 80 (Figure 1.2 B). To date, a number of high-resolution X-ray and solution NMR structures for the ferrous[46,47] and ferric[43,48,49] forms of Cyt *c* have been reported.

Cyt b_5

Cytochrome b_5 (Cyt b_5), discovered in 1952[50] and isolated four years later[51], exists in two forms: either as a membrane-bound or as a soluble protein. The first form is found in the mammalian liver, where it is involved in desaturation of stearyl Co-A[52] or in reduction

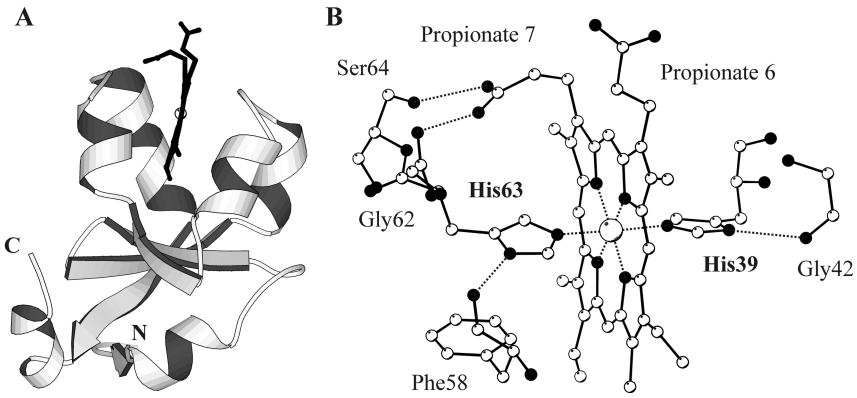

Figure 1.3. **Three-dimensional structure of Cyt b_5.** (**A**) Overall fold of the protein; the haem is shown in sticks, and the protein termini are indicated by the labels. (**B**) The view of the haem binding site showing the axial ligands, His 39 and His 63 (bold labels), and the residues that stabilize the haem environment by hydrogen bonding. The figure was drawn from the crystal structure of the oxidized Cyt b_5 (PDB entry 1CYO[61]).

of cytochrome P450[53], while the second, soluble, form is expressed in erythrocytes, where it functions in the reduction of adventitiously oxidised methaemoglobin[54,55].

The most-studied form of Cyt b_5 is traditionally prepared by lipase or trypsin digestion of the membrane-bound protein, yielding a fully active, soluble fragment of *ca*. 90 – 98 residues, with the length depending on the choice of the protease and the source of Cyt b_5 used[56]. More conveniently, the expression in *E. coli* of the synthetic genes coding for the rat liver Cyt b_5[57] or lipase-solubilized part of the bovine liver microsomal Cyt b_5[58] allows for a rapid isolation of the recombinant protein. The latter variant of Cyt b_5, studied in this work, is a small (10.1 kDa) protein that consists of 93 amino-acids – corresponding to residues 6-98 of the mature protein[56] – and a non-covalently attached (*b*-type) haem group. This protein bears an overall charge of –8 at pH 6.0 (Chapter II), and has a reduction potential of +20 mV at 26^o C and pH 7.0[59].

Cyt b_5 is an elongated molecule that contains six α-helices and five β-strands (Figure 1.3 A). The haem group, held in a hydrophobic pocket formed by four helices (Figure 1.3 A), is non-covalently bound to Cyt b_5 via coordination of the haem iron by $N_{\varepsilon 2}$ atoms of the two histidine side-chains, His 39 and His 63 (Figure 1.3 B). The orientations of the haem ligands are further stabilized by hydrogen bonds between their $N_{\delta 1}$ atoms and the peptide carbonyl oxygens of Gly 42 and Phe 58 (Figure 1.3 B). One of the haem propionates is buried, forming hydrogen bonds with the oxygen atoms of Gly 62 and Ser

64, while the other extends into solution (Figure 1.3 B). Since the original crystal structure of Cyt b_5 was solved in 1972 at 2.8 Å resolution[60], making it the first ET haem protein with a known 3D structure, a number of high-resolution X-ray and solution NMR structures for the ferric[61,62] and ferrous[63,64] forms of this protein have been reported.

CcP

Yeast cytochrome *c* peroxidase (CcP), discovered in baker's yeast in 1940[65], is localized in the intermembrane space of mitochondria, where it catalyses the reduction of hydrogen peroxide to water using the reducing equivalents from two molecules of ferrous Cyt *c*, its physiological partner[66]. It is believed that, together with catalase, CcP plays an important role in protecting the organism from high concentrations of hydrogen peroxide[67].

CcP is a moderately-sized (34.2 kDa), negatively-charged protein (at pH 6.0 the charge is -4 as estimated from protein titration curves[68]) consisting of 294 amino-acids and a non-covalently attached (*b*-type) haem. CcP can be isolated in a high yield from *S. cerevisiae*[69-71] or expressed in *E. coli*[72], from which the recombinant protein can be purified using a number of published procedures[73-75].

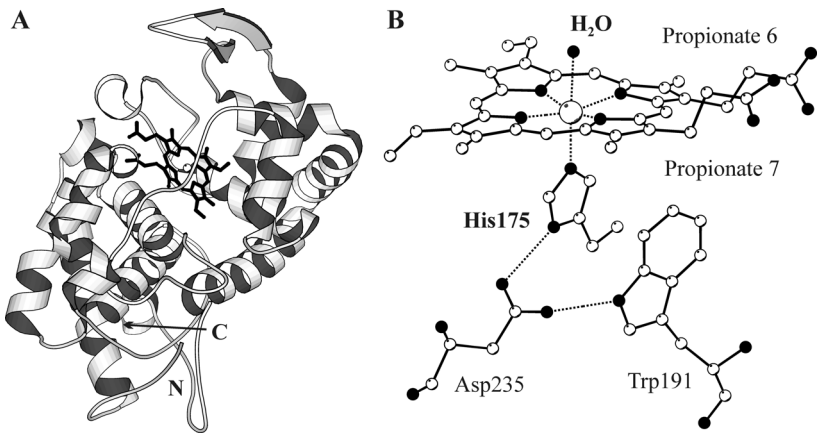

Figure 1.4. **Three-dimensional structure of CcP**. (**A**) Overall fold of the protein; the haem is shown in sticks, and the protein termini are indicated by the labels. (**B**) The view of the haem binding site showing the axial ligand, His 175 (bold label), the catalytically important Trp 191, and Asp 235 that forms stabilizing hydrogen bonds with the latter two; for clarity, only side-chain atoms of His 175 are shown. The figure has been generated from the crystal structure of the resting-state CcP (PDB entry 1ZBY[79]).

CcP, the first haem enzyme for which a crystal structure was solved[76,77] (and further refined to 1.7 Å[78] and, recently, to 1.2 Å[79]), is a highly helical protein of cylindrical shape (Figure 1.4 A). The haem group, buried in a hydrophobic pocket within the protein (Figure 1.4 A), is non-covalently bound to CcP via coordination of the haem iron by $N_{\varepsilon 2}$ atom of His 175, the only axial ligand (Figure 1.4 B). The sixth coordination position, occupied by a water molecule in the resting state of the enzyme (Figure 1.4 B), is vacant for the binding of the peroxide substrate. During the catalytic process, H_2O_2 oxidizes the resting-state CcP, with a high-spin (S= $^5/_2$) Fe(III) haem, to an intermediate called compound I[80,81], in which the haem iron is oxidized to Fe(IV) oxyferryl group[82,83] and the side chain of Trp 191 is oxidized to a cationic indole radical[84-86]. Two molecules of ferrous Cyt *c* transfer electrons to the Fe(IV) and Trp 191 radical sites, returning the enzyme to the native Fe(III) state. The aromatic ring of Trp 191 is parallel to, and in van der Waals contact with, the imidazole ring of His 175, and the orientations of both residues are stabilized by hydrogen bonds to Asp 235 (Figure 1.4 B). For a recent review of mechanistic and structural aspects of this enzyme see ref. 87.

Cyt c – Cyt b_5 complex

The non-physiological protein complex of Cyt *c* and Cyt b_5 was the first ET complex for which a detailed structural model was proposed. In 1976, Salemme matched oppositely-charged groups on the protein molecules taken from the available X-ray structures[60,88] and minimized the haem-to-haem distance, producing a 1:1 complex exhibiting four intermolecular electrostatic contacts and nearly coplanar haem groups (Figure 1.5)[89]. Since then, numerous experimental and theoretical studies have striven to prove or refute the hypothetical model of Salemme (for reviews see ref. 90-92).

Earlier experimental work – concerned with chemical modification or mutagenesis of Cyt *c* and Cyt b_5 residues involved in Salemme's complex – have supported the original model[93-95]. However, further investigation has provided an ample evidence for existence of at least two structurally similar 1:1 protein complexes[96-98], and alternative Cyt *c* – Cyt b_5 models were proposed[96,99]. Recent NMR studies, which use chemical shift perturbations upon complex formation to map the binding surfaces on both proteins, are at variance

regarding the residues that stabilize the complex; however, they all agree that the surface of Cyt b_5 involved in the interaction with Cyt *c* is more extensive than that predicted by any of the static models[100-102].

Another controversy around the Cyt *c* – Cyt b_5 complex concerns the stoichiometry of the interaction. As evidenced by earlier experimental studies[96,103], the stoichiometry of the complex formation was generally accepted to be 1 : 1 until an NMR titration analysis has proposed the presence of a ternary $(\text{Cyt } c)_2$ – Cyt b_5 complex at higher Cyt *c* concentrations[104]. The latter suggestion was strongly criticized, and the validity of the experimental data from which the 2 : 1 complex had been inferred was cast into doubt[91]. However, a recent NMR study has used several complementary approaches to show that the Cyt *c* – Cyt b_5 complex in solution consists of multiple equilibria with 1 : 1 and 2 : 1 stoichiometries[102].

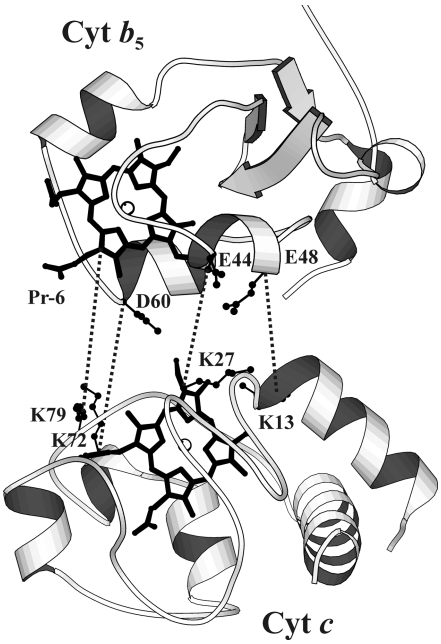

Figure 1.5. **Hypothetical model of Cyt *c* – Cyt b_5 complex.** The dashed lines indicate vectors between the haem ring atom or C_α atoms of the residues that were proposed to form intermolecular salt bridges in the model of Salemme[89]. The residues involved in the hypothetical complex formation are labelled and shown in ball-and-sticks. The haem groups for both proteins are in sticks. Molecular coordinates of Cyt *c* and Cyt b_5 were taken from the PDB entries 2YCC[43] and 1CYO[61], respectively.

The emerging view of the interaction between Cyt *c* and Cyt b_5 in solution, comprising dynamic interconversion among different forms of the Cyt *c* – Cyt b_5 complex and possible binding of two Cyt *c* molecules, highlights the inadequacy of the proposed static models[89,96,99] for an accurate description of this dynamic protein complex.

Cyt c – CcP complex

The widely-studied physiological complex of Cyt *c* and CcP has become a paradigm of the interprotein ET (for recent reviews see ref. 87,90,105,106). The popularity of this complex for the study of intermolecular ET was augmented by the reported Cyt *c* – CcP crystal structure[107], a rare achievement for a transient complex, that has allowed to assess the structure-function relationship by the combination of site-directed mutagenesis and biophysical techniques[87]. Contrary to an earlier electrostatic model[108] and Brownian dynamics simulations[26] that predicted numerous salt bridges across the interface, a single intermolecular hydrogen bond was observed in the crystal structure of the complex (Figure 1.6), stabilized mainly by van der Waals interactions[107].

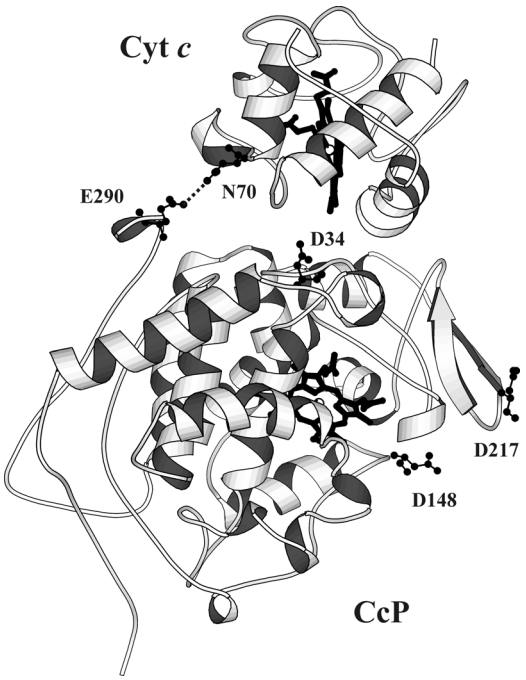

Figure 1.6. **Crystal structure of the Cyt *c* – CcP complex.** A single intermolecular hydrogen bond between $O_{\varepsilon 1}$ atom of E290 (CcP) and $N_{\delta 2}$ atom of N70 (Cyt *c*) observed in the crystal structure of the complex is indicated by a dashed line. The aspartates of CcP that define the Cyt *c* binding sites proposed by Northrup *et al.*[26] are indicated by labels and shown in ball-and-sticks. The haem groups for both proteins are in sticks. The figure has been generated from the crystal structure of the complex (PDB entry 2PCC[107]).

As suggested by the crystal structure and confirmed by a variety of experimental techniques, the interaction between Cyt *c* and CcP in the broad range of experimental conditions occurs with 1 : 1 stoichiometry[87]. However, several reports proposed the binding of the second Cyt *c* molecule at low-ionic strength[109-111], with the binding affinity 2 – 4 orders of magnitude weaker than that of the first Cyt *c*[87]. Since then, the relative ET activity of Cyt *c* bound at

different locations on the surface of CcP and the nature of the 1 : 1 and 2 : 1 complexes have been debated.

According to one view, the ET occurs only from the Cyt *c* bound to the high-affinity, crystallographically-defined site of CcP[112-114], while an alternative hypothesis suggests multiple forms of the ET active 1 : 1 complex and postulates that the ET from the second, weakly bound Cyt *c* is faster than that from the Cyt *c* bound to the high-affinity site in the 1 : 1 complex[109,115]. Recent studies of covalently cross-linked Cyt *c* – CcP complexes, pioneered by Poulos and co-workers[116,117], have confirmed the former hypothesis as they showed that only the Cyt *c* bound to the high-affinity site on CcP is ET active and that a possible binding of the second Cyt *c* does not contribute to the intermolecular ET[118,119].

Location of the second Cyt *c* binding site on the surface of CcP has been a matter of debate (see ref. 87,106 and references therein) that has given rise to two contrasting views on the Cyt *c* – CcP interaction. The first concept postulates a unique 1 : 1 protein complex with the high-affinity Cyt *c* binding site as seen in the crystal structure[107] and a second, much weaker binding site in the 2 : 1 complex, with the absence of interactions between the bound Cyt *c* molecules. Another hypothesis proposes multiple forms of the 1 : 1 complex, with Cyt *c* bound to different locations on the surface of CcP and suggests that the protein-protein orientation seen in the crystal structure represents only a sub-population of binding geometries[26,106]. In this view, formation of a 2 : 1 complex is inhibited by strong electrostatic repulsion between the bound Cyt *c* molecules.

The latter model, supported by an ample experimental evidence[120-125], suggests that the transient complex of Cyt *c* and CcP is dynamic. On the other hand, numerous studies indicate that in solution the proteins spend most of the time in a well-defined, single-orientation complex[126-129]. These two, seemingly contradictory, observations can be reconciled using the concept of the encounter complex (see pp. 10 – 11), according to which the dynamic part of the Cyt *c* – CcP interaction corresponds to the encounter state of the protein complex, while the dominant, well-defined binding geometry is analogous to the single-orientation complex.

Thesis outline

The aim of the research presented in this thesis was the structural characterization of two transient complexes of haem proteins, Cyt *c* – Cyt b_5 and Cyt *c* – CcP, in order to elucidate the principles governing transient interactions in solution, with a particular emphasis on the role of protein dynamics in the process of complex formation. To this end, we have used NMR spectroscopy, isothermal titration calorimetry (ITC), and site-directed mutagenesis.

Chapter II reports NMR chemical shift perturbation analysis and molecular docking of the Cyt *c* – Cyt b_5 complex that have resulted in a structural model for this protein complex.

Chapter III describes solution structure determination of Cyt *c* – CcP complex using site-specific spin-labelling in combination with paramagnetic relaxation enhancement (PRE) NMR spectroscopy and subsequent protein docking based on the experimental restraints.

Chapter IV describes the use of site-specific spin-labelling in tandem with a semi-quantitative PRE analysis to delineate the conformational space occupied by Cyt *c* and CcP in the dynamic encounter state of the complex.

Chapter V reports the binding of R13A and R13K variants of Cyt *c* to CcP, studied by ITC and NMR chemical shift perturbation and PRE analyses, and describes the solution structure of the R13A Cyt *c* – CcP complex solved by PRE NMR spectroscopy.

Chapter VI reports the binding of T12A variant of Cyt *c* to CcP, studied by the experimental techniques of Chapter V, and describes the solution structure of the T12A Cyt *c* – CcP complex solved by PRE NMR spectroscopy.

Finally, in the Concluding Remarks section, the results obtained in this thesis are brought into a broader perspective of the contemporary interdisciplinary research on transient ET protein complexes.