Transient complexes of haem proteins

Volkov, O.M.

Chapter II

The highly dynamic complex of cytochrome c with cytochrome b_5

Abstract

The interaction of bovine ferric cytochrome b_5 (Cyt b_5) with ferric and ferrous yeast *iso*-1-cytochrome *c* (Cyt *c*) has been investigated by heteronuclear NMR spectroscopy. Chemical shift perturbations for ^{1}H and ^{15}N nuclei of both cytochromes, arising from the interactions with the respective unlabelled partner protein, have been used to map the interacting surfaces on both proteins. The observed binding shifts for the oxidized and reduced Cyt *c* are similar, indicating that the complex formation is not influenced by the oxidation state of Cyt *c*. Protein – protein docking simulations have been performed for the binary Cyt b_5 – Cyt *c* and ternary Cyt b_5 – (Cyt *c*)$_2$ complexes using HADDOCK[130]. The docking algorithm, which is driven by the experimental data, has identified a range of orientations assumed by the proteins in the complex. We demonstrate that Cyt *c* uses a confined surface patch for interaction with a much more extensive surface area of Cyt b_5. The experimental data are best interpreted by the presence of a dynamic ensemble of protein – proteins orientations within the complex rather than a single, well-defined structure.

The results presented in this chapter have been published in part as:

Volkov, A. N., Ferrari, D., Worrall, J. A. R., Bonvin, A. M. J. J. & Ubbink, M. The orientations of cytochrome *c* in the highly dynamic complex with cytochrome b_5 visualized by NMR and docking using HADDOCK. *Prot. Sci.* **14**, 799-811 (2005), corrected in *Prot. Sci.* **15**, 1563 (2006).

The majority of the NMR titration experiments and chemical shift perturbation analyses presented in this chapter have been performed by Dr. D. Ferrari (current address: Department of Biochemistry and Molecular Biology, University of Parma, Parco Area delle Scienze 23/A, 43100, Parma, Italy).

Introduction

The Cyt *c* – Cyt b_5 complex is the first ET complex for which a detailed structural model was proposed. In 1976, starting from the available X-ray structures of the individual proteins[60,88], Salemme optimised the electrostatic contacts between the protein molecules by visual inspection and least-squares minimization of the haem-to-haem distance[89]. The resulting model of a well-defined 1:1 complex exhibited four intermolecular complementary charge interactions (E48 – K13, E44 – K27, D60 – K72 and the 6-Pr – K79; Cyt b_5 residues listed first), nearly coplanar haem groups separated by 8 Å, and water exclusion from the binding interface (Figure 1.5, p. 19). This early model stimulated a vigorous development of experimental approaches for structural and mechanistic characterisation of the ET complex between Cyt *c* and Cyt b_5.

The first direct evidence for formation of the non-physiological complex between Cyt b_5 and Cyt *c* came from the gel permeation and ultracentrifugation studies[131] and was later supported by electronic[132] and NMR[103] spectroscopy. Early attempts to identify the residues involved in the complex formation were concerned with chemical modification of specific lysine residues on the surface of Cyt *c*[93,133], esterification of Cyt b_5 haem propionates[94,96], and mutagenesis of the Cyt b_5 carboxylate groups[95,134] combined with steady-state kinetics, spectrophotometry, and hyperbaric spectroscopy, respectively. Although these studies supported the general validity of the Salemme model both in terms of stoichiometry and the residues involved in the interaction, it was suggested that several other interprotein electrostatic contacts might contribute to the complex formation[96,131,133]. The latter suggestion went in line with the earlier findings that five to seven intermolecular charge interactions would account for the observed ionic strength dependence of the Cyt *c* reduction by Cyt b_5[131].

During further investigation of the complex, it became clear that the experimental data from 1D NMR spectroscopy[97,103,104,135], spectrophotometry[96], and kinetics of the ET from Cyt *c* to Cyt b_5[98] could not be accurately interpreted by the simple model of Salemme. The existence of at least two structurally similar 1:1 complexes or of a dynamic ensemble of nearly isoenergetic protein – protein orientations was suggested[96-98,103], and the structure

of an alternative complex proposed[96]. Brownian dynamics simulations of this ET complex led to the conclusion that there is "a single binding domain contributing to electron-transfer dynamics, but not a single conformation as would be observed in a crystal"[99].

From the analysis of NMR titration curves and the line broadening at higher Cyt *c* concentrations, it has been suggested that, in addition to a binary complex, a ternary Cyt b_5 – (Cyt *c*)$_2$ product is also formed[104,135]. Despite the strong criticism of the proposed 1 : 2 stoichiometry[91], a recent NMR study supported the concept of the ternary complex[102]. In particular, analysis of the molecular tumbling correlation rates of the complex in the millimolar concentration range has indicated the presence of multiple equilibria with 1 : 1 and 1 : 2 stoichiometries[102].

Recently, 2D NMR has been used to characterise the ferrous/ferrous[100-102] and ferric/ferric[100] Cyt *c* – Cyt b_5 complexes, and the chemical shift perturbation maps for both proteins have been determined[102]. Although these studies are at variance concerning what residues participate in the complex formation, they all agree in 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[89,95,96].

In the present work, the complexes of ferric Cyt b_5 with ferric and ferrous yeast Cyt *c* are investigated, and the binding-induced chemical shift perturbations are mapped for both proteins. Analysis of the binding maps allows to define the surface areas involved in the complex formation and to perform protein – protein docking simulations using HADDOCK[130], a novel approach that, unlike traditional docking algorithms, enables direct incorporation of the experimental restraints to drive the docking. For all studied complexes, the docking results in a number of nearly isoenergetic protein – protein orientations rather than a single, well-defined structure. The observed ensemble of structures reflects the dynamic nature of the Cyt *c* – Cyt b_5 complex and represents the relative orientations sampled by the protein molecules within the complex in solution.

Results

In order to generate docked structures of the complex between Cyt b_5 and Cyt *c* using HADDOCK, chemical shift mapping of the binding interfaces of both proteins is a pre-requisite. For this purpose, four titration experiments, differing in the ^{15}N labelled

protein and oxidation state of Cyt *c*, have been performed. In each case, complex formation is evidenced by an increase in line-width (7 – 9 Hz) for all peaks in the NMR spectrum, as well as chemical shift perturbations of certain resonances. For all titrations the magnitude of the line broadening in the bound form is the same for both shifted and unaffected peaks, and a single set of amide peaks in the 2D [^{15}N, ^{1}H] HSQC spectrum is observed. This indicates that the Cyt b_5 – Cyt *c* complex is in fast exchange on the NMR time scale. The estimate for the lower limit of the dissociation rate [$k_{off} >> \sqrt{2}\pi\Delta\delta_{max}/2$, where $\Delta\delta_{max}$ is the maximal observed chemical shift perturbation ($\Delta\delta_{binding}$) in Hz] of 200 s^{-1} for both ferric Cyt b_5 – ferric Cyt *c* and ferric Cyt b_5 – ferrous Cyt *c* complexes is in good agreement with the recently reported value of 180 s^{-1} for the ferrous – ferrous complex[101].

Interaction of ferric Cyt b_5 with ferric Cyt c

In the downfield region of the 1D ^{1}H NMR spectrum of ferric Cyt b_5, a number of well-resolved hyperfine-shifted haem resonances and several less intense signals are observed (Figure 2.1). The latter are due to the minor form of Cyt b_5, in which the haem is rotated by 180° about the α-γ-meso axis in respect to the major form[136-138]. The peaks of the major form assigned to the 1-CH_3 (11.7 ppm), 6-α-Pr CH_2 (14.2 ppm), and 5-CH_3 (21.5 ppm)[138] shift during the titration with ferric Cyt *c*, while the position of 7-α-Pr CH_2 signal (18.8 ppm) does not change (Figure 2.1). While chemical shift perturbations of the signal assigned to the 6-α-Pr CH_2 do not follow a 1:1 binding model (Figure 2.2 B), those of the resonances assigned to the 1-CH_3 and 5-CH_3 in the major form fit well with the shared equilibrium association constant (K_a) of (5 ±

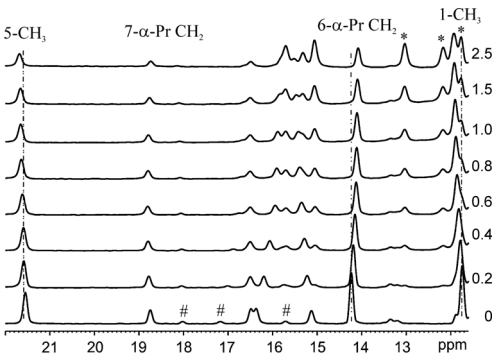

Figure 2.1. **Downfield hyperfine-shifted ^{1}H NMR signals of ferric Cyt b_5 upon titration with ferric Cyt *c*.** The Cyt *c* / Cyt b_5 ratio is given next to each spectrum. The assignments of several peaks of the major form of Cyt b_5 are indicated. The resonances marked with an asterisk and a hash arise from the Cyt *c* protons and the minor form of Cyt b_5, respectively.

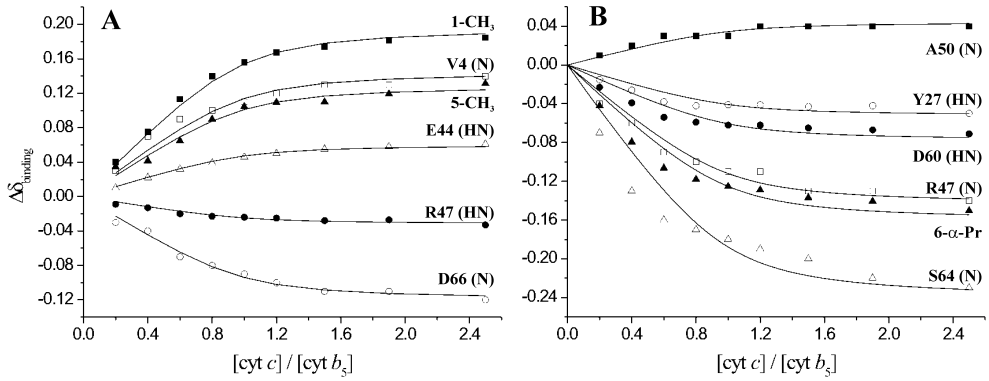

Figure 2.2. **Chemical shift perturbations of ferric Cyt b_5 resonances upon titration with ferric Cyt *c*.** Titration profiles for several Cyt b_5 signals that fit well (**A**) or fit poorly (**B**) to the 1:1 binding model. The curves represent the best fit to a 1:1 binding model with a shared K_a of $(5 \pm 3) \times 10^4$ M^{-1}.

3) × 10^4 M^{-1} (Figure 2.2 A), which is in a good agreement with the value of $(6 \pm 3) \times 10^4$ M^{-1} reported recently for the ferrous – ferrous complex[101].

Chemical shift perturbations for the amide resonances were monitored in a series of 2D [^{15}N, ^{1}H] HSQC spectra. Small $\Delta\delta_{binding}$ for 41% of ferric Cyt b_5 amides were observed, with the largest averaged chemical shift perturbation ($\Delta\delta_{avg}$) of 0.054 ppm for Y27. The $\Delta\delta_{avg}$ values for Cyt b_5 amides in the presence of 2.5 molar equivalents of ferric

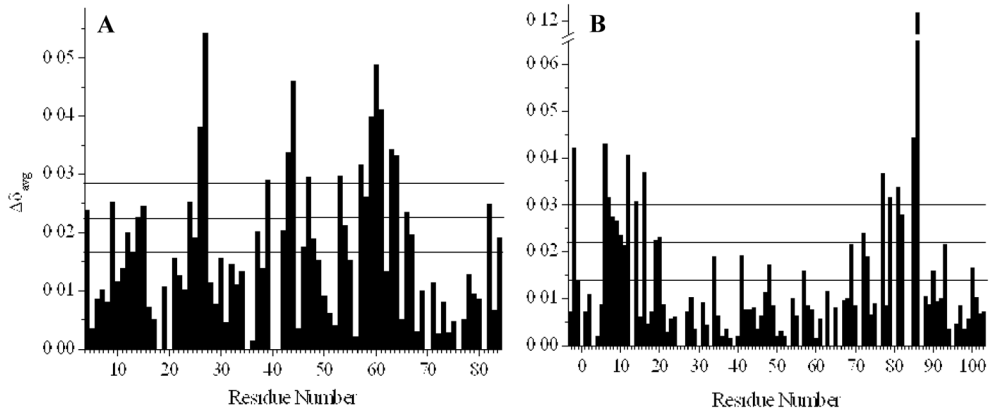

Figure 2.3. **Averaged amide chemical shift perturbations ($\Delta\delta_{avg}$) for ferric Cyt b_5 and ferric Cyt *c*.** (**A**) Ferric ^{15}N Cyt b_5 in the presence of 2.5 molar equivalents of ferric Cyt *c*. (**B**) ^{15}N ferric Cyt *c* in the presence of 2.5 molar equivalents of ferric Cyt b_5. The horizontal lines denote the significance levels of the mean $\Delta\delta_{avg}$ value, the mean plus one half of the standard deviation, and the mean plus one standard deviation.

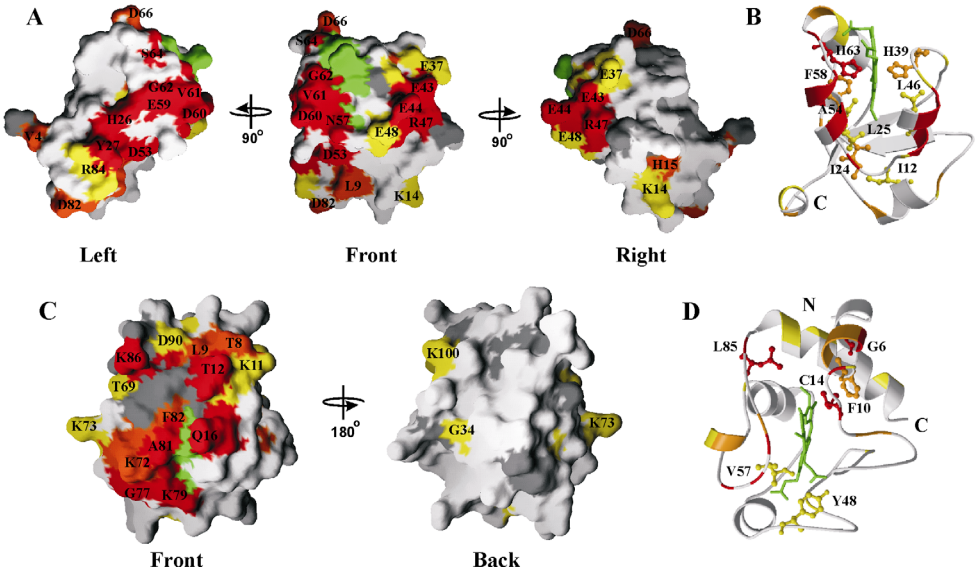

Figure 2.4. **Chemical shift mapping of ferric ^{15}N Cyt b_5 in the presence of ferric Cyt *c*, and ^{15}N ferric Cyt *c* in the presence of ferric Cyt b_5.** (**A**) Surface representations of Cyt b_5 (PDB entry 1CYO[61]); the residues are coloured according to $\Delta\delta_{avg}$ experienced upon binding of ferric Cyt *c* ($\Delta\delta_{avg} \geq 0.029$ ppm in red, $\Delta\delta_{avg} \geq 0.023$ ppm in orange, $\Delta\delta_{avg} \geq 0.017$ ppm in yellow). Unassigned and proline residues are shown in dark grey; the haem group is in green. (**B**) The same as in (**A**), but in ribbon representation. Buried residues affected upon the interaction are labelled and shown in ball-and-stick. The positions of the N- and C-termini are indicated. (**C**) Surface representations of Cyt *c* (PDB entry 2YCC[48]), with residues coloured according to $\Delta\delta_{avg}$ experienced upon binding to Cyt b_5 ($\Delta\delta_{avg} \geq 0.030$ ppm red, $\Delta\delta_{avg} \geq 0.022$ ppm orange, $\Delta\delta_{avg} \geq 0.014$ ppm in yellow). (**D**) The same as in (**B**), but for ferric Cyt *c*. Surface representations in this figure were generated with GRASP 1.3[139], ribbon representations in this figure and Figures 2.8 – 2.12 were created using MOLSCRIPT[44] and rendered in Raster3D[140].

Cyt *c* are plotted in Figure 2.3 A. Several representative titration curves were fitted globally to a 1:1 binding model. A number of residues fit well with the resulting shared binding constant of $(5 \pm 3) \times 10^4$ M^{-1} (Figure 2.2 A), indicating that, at the protein concentrations used, *ca.* 80% of Cyt b_5 is bound to Cyt *c* at the molar ratio of 1.0. For most residues, however, the titration curves fit poorly to the 1:1 binding model (Figure 2.2 B).

The chemical shift changes for Cyt b_5 were mapped onto surface representation of the protein and coloured according to the increasing $\Delta\delta_{avg}$ values (Figure 2.4 A). Two adjacent surface areas on Cyt b_5 that are influenced by binding of Cyt *c* are found on the front and left sides of the protein. The patch on the front side of Cyt b_5 contains the affected

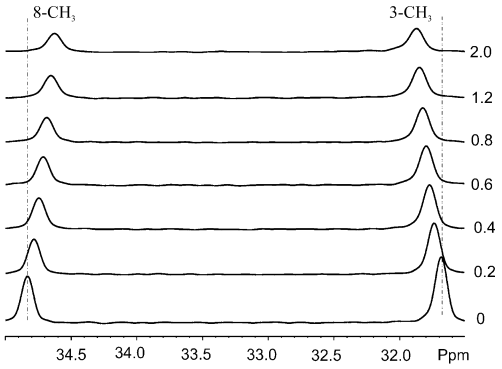

Figure 2.5. **Downfield hyperfine-shifted signals of ferric Cyt *c* upon titration with ferric Cyt b_5.** The Cyt b_5 / Cyt *c* ratio is given next to each spectrum. The assignments of the Cyt *c* peaks are indicated.

residues E43 – E44, R47 – E48, N57, D60 – V61, S64, and D66, which surround the haem group, and L9 and D53, located somewhat farther from the haem (Figure 2.4 A). On the left side, the affected surface residues H26 – Y27 and E59 – D60 form a single patch, while V4 and D82 are spatially distant from other interacting groups (Figure 2.4 A). In addition to the solvent-exposed Cyt b_5 amino acids described above, several buried residues show significant $\Delta\delta_{binding}$ upon interaction with Cyt *c*. These are the haem ligands (H39 and H63), the groups that form part of the hydrophobic haem-binding pocket (L25, L46, A54, and F58), and two amino acids (I12 and I24) that are located towards the bottom of the protein (Figure 2.4 B).

In order to ascertain whether intermolecular pseudocontact shifts arising from the paramagnetic ferric Cyt *c* contribute to the observed chemical shift perturbations of Cyt b_5 resonances, Cyt *c* was reduced by a stoichiometric amount of sodium ascorbate at the end of the titration. Comparison of the HSQC spectra of the ferric – ferric and ferric – ferrous complexes did not show any significant differences, implying the absence of the pseudocontact shifts.

Complex formation between ferric Cyt b_5 and ferric Cyt *c* has also been investigated from the Cyt *c* side. The same set of experiments as for the ^{15}N Cyt b_5 – Cyt *c* system was performed, but with ^{15}N Cyt *c* as the observed protein. In the 1D NMR spectra of ferric Cyt *c*, two signals at 35.0 and 31.5 ppm, assigned to the haem 8-CH_3 and 3-CH_3 groups, respectively[141], shift during titration with Cyt b_5 (Figure 2.5). The $\Delta\delta_{binding}$ for the two peaks can be fitted to a 1:1 binding model with a K_a of $(3 \pm 1) \times 10^4$ M^{-1} (Figure 2.6), in agreement with the value obtained for the corresponding ^{15}N Cyt b_5 – Cyt *c* titration.

Upon addition of ferric Cyt b_5, chemical shift perturbations of ferric Cyt *c* amides were monitored in the HSQC spectra. Small $\Delta\delta_{binding}$ for 32% of the Cyt *c* amides were observed, with the largest $\Delta\delta_{avg}$ of 0.123 ppm for K86. The $\Delta\delta_{avg}$ values for Cyt *c* residues in the presence of 2.5 equivalents of Cyt b_5 are plotted in Figure 2.3 B. The chemical shift

perturbations in both ^{15}N and ^{1}H dimensions fit well to the 1:1 binding model with a shared K_a of $(6 \pm 3) \times 10^4$ M^{-1} (data not shown), which is again consistent with the corresponding Cyt b_5 titration.

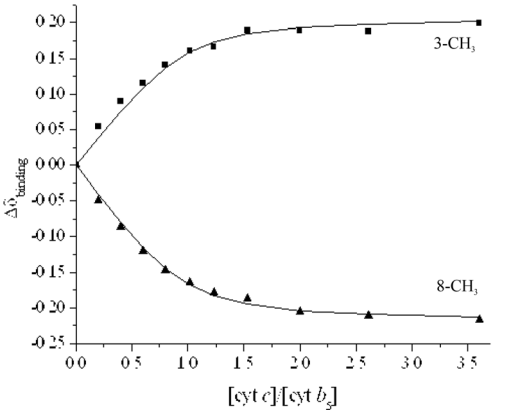

Figure 2.6. **Chemical shift perturbations of ferric Cyt *c* resonances upon titration with ferric Cyt b_5.** The curves represent the best fit to a 1:1 binding model with a shared K_a of $(3 \pm 1) \times 10^4$ M^{-1}.

Using the same procedure as for ^{15}N Cyt b_5, Cyt *c* residues were colour-coded and mapped onto the surface representation of the protein (Figure 2.4 C). The groups exhibiting the largest $\Delta\delta_{avg}$ are T8 – L9, T12, Q16, K72, G77, K79, A81 – F82, and K86, all of which are located in a patch surrounding the haem edge. A few weakly affected amides are found on the opposite side of the protein (Figure 2.4 C). Several buried residues also exhibit significant chemical shift perturbations during the titration with Cyt b_5 (Figure 2.4 D). These are G6, F10, C14, Y48, V57 and L85, all of which sit in the core of the protein.

Interaction of ferric Cyt b_5 with ferrous Cyt c

To investigate the influence of the oxidation state of Cyt *c* on complex formation, ^{15}N ferric Cyt b_5 was titrated with ferrous Cyt *c*. Reduced Cyt *c* does not have signals in the downfield region of the 1D NMR spectrum; therefore, chemical shift perturbations of several Cyt b_5 peaks that overlapped with Cyt *c* resonances in the ferric – ferric experiment (Figure 2.1) could now be monitored. Again, these spectral changes were attributed to complex formation and fitted to a 1:1 binding model. The signals arising from 1-CH_3, 5-CH_3, and the 7-α-Pr CH_2 of the minor form (15 ppm) fit well to this model (data not shown).

Analogously to the ferric – ferric titration, small $\Delta\delta_{binding}$ for various ^{15}N and ^{1}H nuclei were observed in the HSQC spectra. The values of $\Delta\delta_{avg}$ are very similar to those for the ferric – ferric titration for all Cyt b_5 residues, with the exception of I24, D53, and N57,

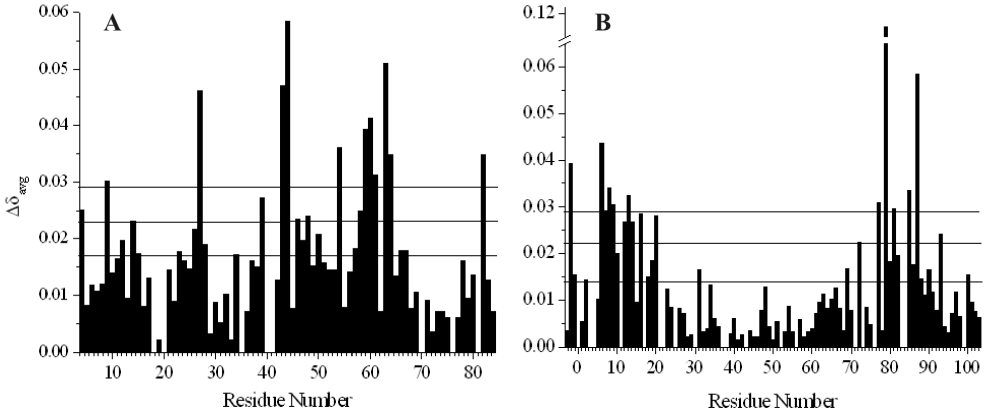

Figure 2.7. **Averaged amide chemical shift perturbations ($\Delta\delta_{avg}$) for ferric Cyt b_5 and ferrous Cyt *c*.** (**A**) Ferric ^{15}N Cyt b_5 in the presence of 2.5 molar equivalents of ferrous Cyt *c*. (**B**) ^{15}N ferrous Cyt *c* in the presence of 2.5 molar equivalents of ferric Cyt b_5. The horizontal lines denote the significance levels of the mean $\Delta\delta_{avg}$ value, the mean plus one half of the standard deviation, and the mean plus one standard deviation.

which are not affected in the titration with ferrous Cyt *c* (Figure 2.7 A). The values of $\Delta\delta_{binding}$ for several amide resonances and three 1D peaks were fitted globally to a 1:1 binding model with a resulting K_a of $(4 \pm 1) \times 10^4$ M^{-1} (data not shown), implying that, in the range of the protein concentrations used, *ca.* 80% of Cyt b_5 is bound to ferrous Cyt *c* at the molar ratio of 1.0.

Analysis of the HSQC spectra for the ^{15}N ferrous Cyt *c* – Cyt b_5 titration shows that the oxidation state of Cyt *c* has little effect on the binding properties of the two cytochromes. Plotting $\Delta\delta_{avg}$ for ferrous Cyt *c* in the presence of 2.5 molar equivalents of Cyt b_5 gives results very similar to those for the ferric Cyt *c* (Figure 2.7 B). The main difference is found for the amides of K79 and K86, with the former more affected by the binding of the reduced Cyt *c* and the latter of the oxidised Cyt *c*. Several residues of ferrous Cyt *c*, which exhibit significant $\Delta\delta_{binding}$, fit well to the 1:1 binding model with a shared K_a of $(3 \pm 1) \times 10^4$ M^{-1} (data not shown), which is consistent with that of the corresponding Cyt b_5 – Cyt *c* titration.

Finally, to check for intermolecular pseudocontact shifts arising from ferric Cyt b_5 and affecting the amides of ferrous Cyt *c*, sodium dithionite was added at the end of the titration in order to generate the completely reduced complex. Similarly to the ^{15}N ferric Cyt b_5 – Cyt *c* titration, no intermolecular paramagnetic effects were observed.

Effects of ionic strength on complex formation

In order to assess the influence of the ionic strength on the complex formation, two salt titrations, differing in the ^{15}N labelled protein, were performed. In each experiment, the sample containing the ^{15}N labelled protein and 2.1 molar equivalents of the unlabelled partner was titrated by adding small aliquots of NaCl solution to a total salt concentration of 20 – 220 mM. To correct for the ionic strength effect on the free protein, the $\Delta\delta_{binding}$ values for each titration point were determined by using the isotonic sample containing only ^{15}N labelled protein as a reference. For both ionic strength titrations, increasing salt concentration leads to the uniform decrease of the $\Delta\delta_{binding}$ for all affected residues of both proteins (data not shown).

Protein – Protein Docking

NMR experiments performed in this study have been used to map the interacting surfaces on both proteins. In order to generate the possible protein – protein orientations sampled by Cyt *c* and Cyt b_5 within the complex, *in silico* docking simulations have been carried out using the binding maps as an input. Two docking runs for the binary complexes of Cyt b_5 with ferric and ferrous Cyt *c* and two runs for the ternary complexes of Cyt b_5 – (ferric Cyt $c)_2$ and Cyt b_5 – (ferrous Cyt $c)_2$ have been performed.

For both binary and ternary complexes of Cyt b_5 with ferric and ferrous Cyt *c*, the ambiguous intermolecular restraints (AIRs) were generated using the active and passive residues specified for both proteins as described in Materials and Methods (for the list of AIRs see Table A1 in Appendix A). The docking procedure – governed by the combination of electrostatic, van der Waals, and AIR restraints – generated solutions that were clustered using a backbone rmsd cut-off criterion (for details see Materials and Methods and ref. 130).

For the binary Cyt b_5 – ferric Cyt *c* complex, ten clusters of solutions were obtained with a 2.0 Å rmsd cut-off. As can be appreciated from the plots of the non-bonded energy terms versus the backbone rmsd from the lowest-energy structure (Figure 2.8 A), the

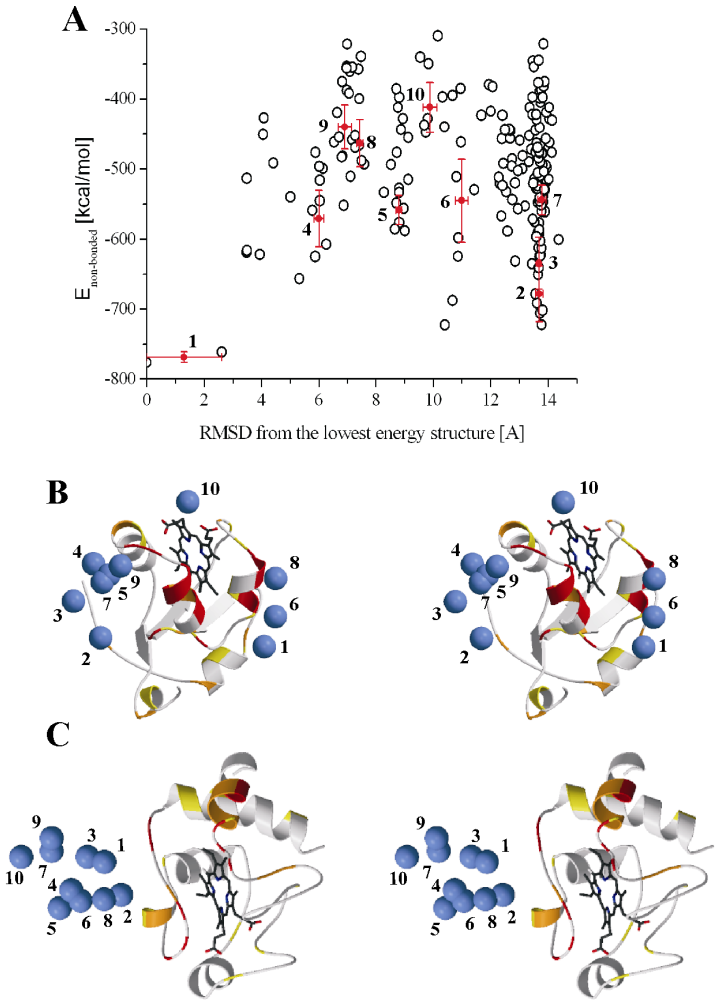

Figure 2.8. **HADDOCK docking solutions for the ferric Cyt b_5 – ferric Cyt *c* complex.** (**A**) Intermolecular non-bonded energy as a function of the backbone rmsd from the lowest energy structure. The clusters' averages are indicated by red circles with error bars representing the standard deviation from the mean for the five lowest-energy structures of each cluster. The clusters are numbered according to the increasing energy. Stereo views of the centres of mass of (**B**) Cyt *c* from each cluster (blue spheres) superimposed onto the ribbon representation of Cyt b_5, and (**C**) Cyt b_5 from each cluster (blue spheres) superimposed onto the ribbon representation of Cyt *c*. The ribbon representations of the proteins in (**B**) and (**C**) are coloured according to the increasing $\Delta\delta_{avg}$ as in Figure 2.4, with the haem group shown in sticks. The centres of mass were calculated for the lowest energy structure of each cluster. The numbering of the clusters corresponds to that in (**A**).

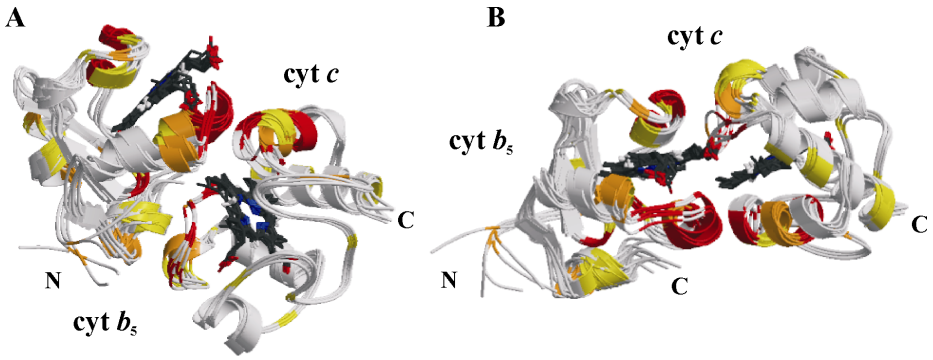

Figure 2.9. **Docking models of the ferric Cyt b_5 – ferric Cyt *c* complex.** "Side-on" (**A**) and "head-on" (**B**) orientations. Ensembles of the five lowest-energy structures for the cluster 3 (**A**) and cluster 8 (**B**) are coloured according to the increasing $\Delta\delta_{avg}$ as in Figure 2.4, with the haem group shown in sticks. Cluster numbers correspond to those in Figure 2.8. The labels indicate positions of the protein termini.

docking solutions consist of several groups of structures that have nearly equal energies (for the structural statistics see Table A2 in Appendix A). The spatial distribution of the docking solutions is shown in Figures 2.8 B and C, with the centres of mass of Cyt *c* molecules from the lowest-energy structure of each cluster superimposed on the Cyt b_5 structure, and vice versa. Overall, the docking solutions consist of two sub-populations of clusters: one in which Cyt *c* molecule docks to the front face of Cyt b_5 ("head-on" orientation as seen is cluster 8, Figure 2.9 B) and the other with Cyt *c* binding to the left side of Cyt b_5 ("side-on" orientation as that in cluster 3, Figure 2.9 A).

In order to verify the possibility of a ternary complex formation, the protein – protein docking of the Cyt b_5 – (Cyt c)$_2$ heterotrimer was performed with a modified version of HADDOCK allowing direct docking of ternary complexes (see Material and Methods). The set of AIRs used for the docking of the trimer was the same as that for the binary complex (Table A1 in Appendix A) with the modification such that each active residue of Cyt b_5 had a restraint to all active and passive residues of both Cyt *c* molecules.

The results for the simultaneous docking of two ferric Cyt *c* molecules to one ferric Cyt b_5 are presented in Figure 2.10. The increased number of degrees of freedom in a ternary system makes clustering of the docking solutions more difficult, as can be appreciated from Figure 2.10 A showing a broader structure distribution compared to that of the binary complex (Figure 2.8 A). Clustering of the docked structures with a backbone

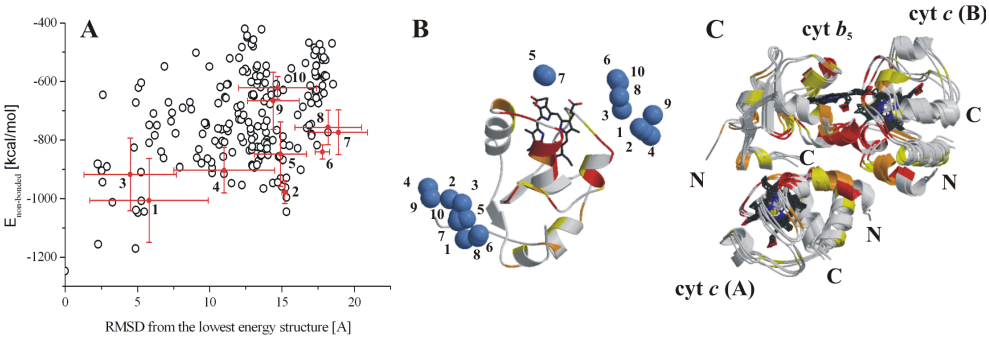

Figure 2.10. **HADDOCK docking solutions for the ternary ferric Cyt b_5 – (ferric Cyt c)$_2$ complex**. (**A**) Intermolecular non-bonded energy as a function of the backbone rmsd from the lowest energy structure. (**B**) Centres of mass of ferric Cyt *c* from each cluster (blue spheres) superimposed onto the ribbon representation of Cyt b_5. (**C**) Ensemble of the three lowest-energy structures of the lowest-energy cluster [number 1 in (**A**) and (**B**)]. The positions of the protein termini are indicated. The ribbon representations are coloured as in Figure 2.4. For more details see the legend to Figure 2.8.

rmsd cut-off of 3.75 Å resulted in ten partially overlapping clusters (for the structural statistics see Table A2 in Appendix A). The distribution of Cyt *c* molecules around Cyt b_5 is shown in Figure 2.10 B. Compared with the binary complex, the same surface areas on Cyt b_5 sustain the interactions with Cyt *c* molecules. An ensemble of structures for the ternary complex (Figure 2.10 C) looks like a combination of binary complexes with the "side-on" and "head-on" orientations (Figure 2.9), implying that similar interactions

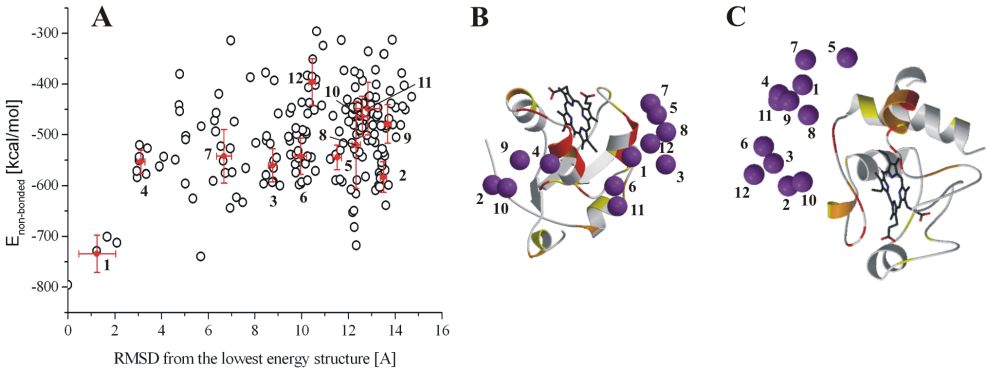

Figure 2.11. **HADDOCK docking solutions for the ferric Cyt b_5 – ferrous Cyt *c* complex.** (**A**) Intermolecular non-bonded energy as a function of the backbone rmsd from the lowest energy structure. Centres of mass of (**B**) Cyt *c* from each cluster (purple spheres) superimposed onto the ribbon representation of Cyt b_5, and (**C**) Cyt b_5 from each cluster (purple spheres) superimposed onto the ribbon representation of Cyt *c*. For more details see the legend to Figure 2.8.

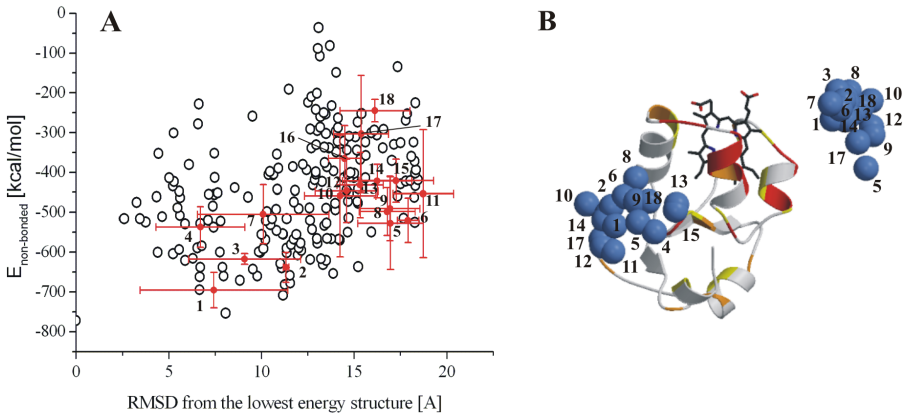

Figure 2.12. **HADDOCK docking solutions for the ternary ferric Cyt b_5 – (ferrous Cyt c)$_2$ complex.** (**A**) Intermolecular non-bonded energy as a function of the backbone rmsd from the lowest energy structure. (**B**) Centres of mass of ferrous Cyt *c* from each cluster (blue spheres) superimposed onto the ribbon representation of Cyt b_5. For more details see the legend to Figure 2.8.

between the partner molecules could take place in the binary and ternary complexes.

As the binding maps for the complexes of Cyt b_5 with ferric and ferrous Cyt *c* are almost the same (see above), it could be expected that docking of the complexes differing only in the oxidation state of the Cyt *c* would also give similar results. Indeed, docking runs for the binary (Figure 2.11) and ternary complexes (Figure 2.12) of Cyt b_5 with ferrous Cyt *c* produced solutions similar to those for the runs with ferric Cyt *c* described above (for the structural statistics see Table A2 in Appendix A). The only difference between the runs with ferric and ferrous Cyt *c* is a broader cluster distribution for the latter (Figure 2.11 A and Figure 2.12 A). Because of the similarity of the docking results, only the ferric – ferric complex is discussed below.

Discussion

NMR study of the Cyt b_5 – Cyt c complex

Previous attempts to study by heteronuclear NMR the complex between Cyt b_5 and Cyt *c*, in which one or both proteins are in the oxidized state, have been severely hampered by the auto-reduction of Cyt *c*[102]. Under the experimental conditions used in this

work [20 mM sodium phosphate pH 6.0, pre-treated with Chelex (Sigma) chelating beads], both proteins are stable, and no self-reduction of ferric Cyt *c* occurs during the experiments. The binding constants, as estimated from fitting the titration curves to a 1:1 binding model, are $(5 \pm 3) \times 10^4$ M^{-1} and $(3 \pm 1) \times 10^4$ M^{-1} for ferric Cyt b_5 – ferric Cyt *c* and ferric Cyt b_5 – ferrous Cyt *c* complexes, respectively. These are in good agreement with the previously reported values in the range of $(4 – 8) \times 10^4$ M^{-1} at low ionic strength, pH 7.0, and 298 - 303 K[101,132,142]. The variations in these values can be attributed to different protein sources and experimental conditions, especially the ionic strength, which is notoriously difficult to control at high protein concentrations[101].

The NMR titrations show that the chemical shift perturbations of some Cyt b_5 resonances follow a 1:1 binding model, while others clearly deviate (Figure 2.2). At the same time, the binding curves for most Cyt *c* resonances exhibit a clear 1:1 stoichiometry (*e.g.* Figure 2.6). A plausible explanation for this phenomenon is the formation of higher-order protein complexes. Existence of a ternary complex, in which Cyt b_5 interacts with two molecules of Cyt *c*, has been suggested before[104,135] and later criticised[91]. Although the experimental data presented in this work are not sufficient for reliable fitting to a 1:2 binding model, combined with the protein – protein docking they support the formation of a Cyt b_5 – (Cyt $c)_2$ complex, in agreement with the results of a recent NMR study[102].

Assessing the stoichiometry of the Cyt b_5 – Cyt *c* complex from the NMR titration curves should be done with care, because the complex is sensitive to variations in the solution ionic strength arising from addition of the partner protein in the course of the titration[91]. In this study the approximate ionic strength of solution, as estimated using the approach of Eley & Moore[103], increased from 45 to 55 mM with the addition of protein aliquots up to two molar equivalents. A further ionic strength increase of *ca.* 1 mM at the end of the titration occurred due to the pH correction after each titration step. From the ionic strength dependence experiments, it could be concluded that a total ionic strength increase of 11 mM causes negligible effects on the binding curves. Thus, changes of the ionic strength during the titrations cannot account for the deviation from 1:1 binding observed for some of the curves (Figure 2.2 B).

Comparison of the amide binding shifts ($\Delta\delta_{avg}$) for the Cyt b_5 residues upon interaction with either ferric or ferrous Cyt *c* shows that the only differences are found for three amino acids (I24, D53, and N57), which are perturbed by the binding of oxidised but

not reduced Cyt *c*. Analogously, $\Delta\delta_{avg}$ values for ferric and ferrous Cyt *c* in the complex with Cyt b_5 are similar for all but two residues, K79 and K86, whose NMR resonances are very sensitive to small changes in the ionic strength (data not shown, also ref. 143). Therefore, the variations in $\Delta\delta_{avg}$ between ferric and ferrous forms is likely due to slight discrepancies in ionic strengths during the two titrations, rather than to the redox-dependent differences in the binding of Cyt *c* to Cyt b_5.

The binding surface of Cyt b_5 in the complex with either reduced or oxidised Cyt *c* is more extensive than that reported recently for the ferrous – ferrous complex[101] or predicted by the static models of Salemme[89] and Rodgers *et al.*[95]. The perturbed surface residues of Cyt b_5 are located at two adjacent regions (Figure 2.4 A), which are similar to those described in a recent report[102]. In contrast to the results of yet another NMR study[100], no binding at the suggested "cleft" site is observed in our experiments.

Many affected residues are found on the front face of Cyt b_5, in a patch surrounding the exposed haem edge (Figure 2.4 A). Among those are the charged groups implied by the models of Salemme[89] and Northrup *et al.*[99] – E44, E48, D60, and the most exposed 6-α-Pr CH_2. Involvement of these residues in the complex formation was shown in an earlier study by a combination of site-directed mutagenesis and NMR techniques[144], and a recent report has confirmed the pivotal role played by these groups in stabilization of the complex[145].

The other affected area is located on the left side of Cyt b_5, where most of the perturbed residues cluster in a single patch (Figure 2.4 A). Although the majority of the affected Cyt b_5 resonances are charged, signifying the importance of the intermolecular electrostatic interactions for this complex, several surface-exposed polar and hydrophobic groups are also perturbed by Cyt *c*. Interestingly, the binding surface map of Cyt b_5 in the complex with Cyt *c* looks very similar to that for the highly dynamic complex with myoglobin[12]; residues E43, E44, D60, and V61 are strongly affected in both, implying the similarity of the binding mode employed by Cyt b_5 in both complexes.

The Cyt *c* residues affected by the interaction with Cyt b_5 are confined to a single surface patch on the front side of the molecule (Figure 2.4 C). This region is similar to that observed in a recent NMR study[102], with the main difference being the α_4 helix (residues 90-101), which is unaffected in our experiments. Among the most affected Cyt *c* residues are K72, K79, and K86 – three lysines stabilizing the complex in the models of Salemme[89]

and Northrup *et al.*[99]. Interestingly, binding maps of Cyt *c* in the complexes with Cyt b_5 (this work) and the non-physiological partner adrenodoxin[146] are strikingly similar. Moreover, chemical shift mapping studies of Cyt *c* in the complexes with yeast cytochrome *c* peroxidase[129], cyanobacterial cytochrome *f*[143], pea plastocyanin[147], and Cyt b_5 (this work) indicate that T12 (Q12 in horse heart Cyt *c* used in ref. 147), Q16, and K79 give the biggest binding shifts. This finding suggests that Cyt *c* employs a conserved set of surface-exposed residues for the interactions with a variety of proteins.

The importance of intermolecular electrostatic interactions for the complex formation between negatively-charged Cyt b_5 and positively-charged Cyt *c* is further confirmed by the series of NMR experiments performed at varying ionic strengths. As would be expected for a complex dominated by electrostatic interactions, $\Delta\delta_{binding}$ of the affected residues uniformly decreases with the increasing salt concentration, reaching the values of the chemical shifts of the free protein at the end of the titration.

For both Cyt b_5 and Cyt *c*, there are several buried groups whose chemical shift perturbations cannot be explained by the direct interaction with a partner protein. Most likely, these are caused by transmittance of the binding effects from the surface of the protein to its core via covalent and hydrogen bonds[147]. In the case of Cyt b_5 and Cyt *c*, most of the affected buried residues are located close to the haem group (Figures 2.4 B and D). It is conceivable that the observed effects are caused by small changes in the haem environment induced by the complex formation.

The binding surface of Cyt b_5 is more extensive than that of Cyt *c* (Figure 2.4). As it is believed that there are no structural rearrangements of Cyt b_5 and Cyt *c* upon complex formation[91], the binding of these proteins in a single, well-defined orientation with a 1:1 stoichiometry does not account for the observed chemical shift perturbation maps. Moreover, as there are two simultaneously affected areas on the surface of Cyt b_5, the interaction implies the co-existence of at least two different complexes with Cyt *c*, or the presence of an ensemble of structures with similar energies[91,96,97,102]. It seems likely that Cyt *c* uses a single surface patch to explore a much wider area on Cyt b_5. The reason for such behaviour is probably the broader charge distribution on the latter.

The presence of a dynamic ensemble of structures rather than a single, well defined Cyt b_5 – Cyt *c* complex is further confirmed by the absence of intermolecular pseudocontact shifts. Comparison of the HSQC spectrum of the ferric Cyt b_5 – ferric Cyt *c*

complex with those of the complexes in which either only Cyt *c* or both proteins are reduced shows no significant differences. Contrary to the previously published data[148], this means that no pseudocontact shifts caused by the ferric haem of Cyt *c* is experienced by the amide protons of Cyt b_5 and *vice versa*. The absence of the intermolecular pseudocontact shifts can be explained by multiple fast-exchanging protein – protein orientations within the complex. For such a system the paramagnetic, orientation-dependent effects would average out to zero[147].

A dynamic ensemble of Cyt b_5 – Cyt *c* structures could also explain small $\Delta\delta_{avg}$ observed for both proteins. For multiple, fast-exchanging, isoenergetic protein – protein configurations, the observed $\Delta\delta_{avg}$ would be averaged over all orientations. Together with the scarce close intermolecular contacts and limited desolvation of the interface, this would explain the observed small $\Delta\delta_{avg}$ in the complex[146]. Comparison of $\Delta\delta_{avg}$ with those of other dynamic complexes places the Cyt b_5 – Cyt *c* interaction between that of Cyt *c* – CcP[129], in which the proteins occupy a single orientation for a large fraction of the lifetime of the complex, and the highly dynamic complex of Cyt b_5 with myoglobin[12]. The magnitude of the $\Delta\delta_{avg}$ for the Cyt b_5 – Cyt *c* system is similar to that for the dynamic complexes of Cyt *c* with adrenodoxin[146] and horse heart Cyt *c* with pea plastocyanin[147], suggesting that Cyt b_5 and Cyt *c* adopt different relative orientations within the complex, rather than form a single, well defined structure.

Protein – protein docking

HADDOCK is a new *in silico* docking approach that allows direct incorporation of biochemical and / or biophysical information to drive the docking process[130]. Unlike traditional docking algorithms that use *ab initio* docking simulations followed by *a posteriori* scoring based on the experimental information, HADDOCK *a priori* includes the available experimental data to drive the subsequent docking simulations.

The distribution of the centres of mass of docked proteins (Figures 2.8 B and C) shows that Cyt *c* uses a single surface patch for interaction with two areas on the surface of Cyt b_5. Only a combination of both sub-populations of docking solutions – “head-on” and

"side-on" orientations with Cyt *c* binding to the front and left side of Cyt b_5, respectively (Figure 2.9) – satisfies all restraints derived from the binding maps.

Docking of the ternary Cyt b_5 – (Cyt c)$_2$ complex results in a broader distribution of clusters (Figure 2.10 A) as compared to the binary complex (Figure 2.8 A), which is not unexpected for a system with more degrees of freedom. The resulting structures of the ternary complex combine the features of the "head-on" and "side-on" orientations for the binary complex (Figures 2.10 B and C), indicating that one Cyt b_5 can sustain the interaction with two Cyt *c* molecules. As mentioned above, the formation of the ternary complexes may indeed take place at higher Cyt *c* concentrations, in which case the geometry of the docked Cyt b_5 – (Cyt c)$_2$ complex would account for the observed chemical shift perturbation maps.

In a very recent study, NMR cross-saturation and chemical shift perturbation data have been used as an input for HADDOCK docking of ferrous Cyt b_5 – ferrous horse Cyt *c* complex[149]. While the reported binding map of Cyt *c* is similar to that determined in our work, the interacting surface of Cyt b_5 is much narrower than in our case[101,149]. Protein docking produces two clusters of solutions, one of which is rejected. The dismissal is based on the predicted chemical shifts changes of Cyt *c* amides induced by the haem ring current of Cyt b_5. However, for both clusters the predicted values do not accurately reproduce the observed binding shifts (Figure 8 in ref. 149); therefore, the decision to retain only one solution that better matches the sign of the predicted shifts is, in our view, not well substantiated. It appears that there is insufficient data to discriminate between the two clusters and, therefore, both of them should be retained. In this case the reported results suggest the presence of more than one structure for the Cyt b_5 – Cyt *c* complex and are best interpreted using the concept of multiple protein – protein orientations illustrated in the present report. Moreover, the sheer fact that the magnitude of the observed chemical shift perturbations is several times lower than predicted could be explained by motional averaging as shown above.

Binary and ternary HADDOCK solutions are optimised for the interaction energy, yielding 4 to 5 salt bridges and 10 to 12 hydrogen bonds between the proteins. However, it is unrealistic to assume that such extensive intermolecular contacts are formed. The apparent binding constant of 5×10^4 M^{-1} is equivalent to a binding energy of mere 6 kcal/mol, suggesting that on average only a few weak interactions are formed. Furthermore,

small $\Delta\delta_{avg}$ can be interpreted as an evidence for the limited desolvation of the interface[4]. Extensive H-bonding and ample desolvation would result in much larger chemical shift perturbations as observed in many cases for specific complexes of biomolecules. Thus, the HADDOCK structures should not be seen as detailed snapshots of the complex; rather, they provide a broader picture of the range of orientations sampled by both proteins within this dynamic complex.

Materials and Methods

Preparation of proteins and NMR samples

Both unlabelled and isotopically-enriched ^{15}N Cyt b_5 and ^{15}N T-5A/C102T variant of Cyt *c* were produced in *E. coli* and purified as reported previously[41,58,150] (for the detailed protocols see Appendix B). Cyt b_5 with a UV-vis peak ratio $A_{412.5} / A_{280} \geq 4.0$ was used for the NMR experiments. Concentrations of the Cyt b_5 were determined according to the absorbance peak at 412.5 nm ($\varepsilon = 117$ mM^{-1}cm^{-1})[151]. Ferric Cyt *c* with a UV-vis peak ratio $A_{410} / A_{280} \geq 4.0$ was used throughout. Cyt *c* was oxidised by a 2 – 3-fold excess of $K_3[Fe(CN)_6]$, followed by ultrafiltration under nitrogen against 20 mM NaP_i pH 6.0 using a YM3 membrane (Millipore, Billerica, MA). Ferrous Cyt *c* was prepared similarly, using sodium ascorbate as a reductant. Protein concentrations were determined from the absorbance peaks at 410 nm ($\varepsilon = 106.1$ mM^{-1}cm^{-1}) and 550 nm ($\varepsilon = 27.5$ mM^{-1}cm^{-1}) for ferric and ferrous Cyt *c*, respectively[152].

All NMR samples contained 0.49 – 0.53 mM of the ^{15}N labelled protein and varying amounts of the unlabelled partner protein in 20 mM NaP_i pH 6.0, 6% D_2O for lock, and 0.1 mM $CH_3CO^{15}NH_2$ and 0.1 mM TSP as internal references. The buffer for the samples and protein stock solutions was pre-treated with Chelex 100 (Sigma, Zwijndrecht, Netherlands) chelating agent in order to remove traces of metals that could catalyse self-reduction of Cyt *c*. The pH of the samples was checked before and after each titration step and adjusted, if necessary, with small aliquots of 0.1 M NaOH or 0.1 M HCl solutions.

For the NMR titrations, microlitre aliquots of the stock solution of the unlabelled protein (1.76 mM and 1.86 mM of Cyt b_5 in titrations with ^{15}N ferric and ^{15}N ferrous Cyt *c*,

respectively, and 2.15 mM of ferric Cyt *c* and 2.2 mM of ferrous Cyt *c* in titrations with ^{15}N Cyt b_5) were added to the sample containing 0.5 ml of the ^{15}N labelled protein with an initial concentration of 0.5 mM. Each titration consisted of ten experimental points with 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 1.9, and 2.5 molar equivalents of the unlabelled protein. For the ionic strength titrations, aliquots of 2 M NaCl were added to 0.5 ml sample containing 0.5 mM of the ^{15}N labelled protein and 2.1 molar equivalents of the unlabelled partner protein in 20mM NaP_i pH 6.0 to the total salt concentrations of 20, 70, 120, 170, and 220 mM. For each salt titration, a sample containing only ^{15}N labelled protein was titrated in the same manner and used as a reference.

NMR experiments

All NMR experiments were performed at 301 K on a Bruker DMX600 spectrometer equipped with triple-resonance TXI-Z-GRAD probe (Bruker, Karlsruhe, Germany). For each titration, 1D ^{1}H spectra were acquired with a spectral width of 48.08 kHz and 2048 complex points. The spectral widths (in Hz) for 2D [^{15}N, ^{1}H] HSQC spectra were 1,900 (^{15}N) and 9,620 (^{1}H) for ^{15}N Cyt b_5 titration and 2,680 (^{15}N) and 8,390 (^{1}H) for ^{15}N Cyt *c* titration. All 2D spectra were acquired with 1024 and 216 complex points in the ^{1}H and ^{15}N dimensions, respectively.

Data processing of the 1D ^{1}H and 2D [^{15}N, ^{1}H] HSQC spectra was performed in XWINNMR (Bruker) and AZARA 2.7[153], respectively. Assignments of the ^{15}N and ^{1}H nuclei of the free ferric Cyt b_5 and ferric and ferrous Cyt *c* were taken from the previous work[12,129]. The amides not observed in the present work were A3, S18, S20, F35, G41, L70, and I76 for Cyt b_5; F-4, A3, H33, H39, N56, E66, M80, G83-G84, K87, and I95 for ferric Cyt *c* and A3, K11, K22, R38, A43, N52, N56, K73, M80, G83, G84, and K99 for ferrous Cyt *c*.

Chemical shift perturbations of ^{15}N and ^{1}H nuclei were analysed by overlaying the spectra of bound ^{15}N labelled protein with that of the free protein in ANSIG[154,155]. The averaged amide chemical shift perturbations ($\Delta\delta_{avg}$) were derived from Equation 2.1:

$$\Delta\delta_{avg} = \sqrt{\frac{(\Delta\delta^{N}_{binding}/5)^2 + (\Delta\delta^{H}_{binding})^2}{2}} \qquad (2.1)$$

where $\Delta\delta^{N}_{binding}$ is the chemical shift perturbation of the amide nitrogen and $\Delta\delta^{H}_{binding}$ is that of the amide proton.

Chemical shift titration curves were analysed with a two-parameter non-linear least-squares fit using a one site binding model corrected for the dilution effect[156], Equation 2.2:

$$\Delta\delta_{binding} = \tfrac{1}{2}\Delta\delta_{\infty}(A - \sqrt{A^2 - 4R}) \qquad (2.2)$$

$$A = 1 + R + \frac{LR + U}{LUK_a}$$

where R is the ratio between the unlabelled and the ^{15}N labelled proteins, $\Delta\delta_{binding}$ is the chemical shift perturbation at a given protein ratio, $\Delta\delta_{\infty}$ is the chemical shift perturbation at $R \rightarrow \infty$, L is the initial concentration of the ^{15}N labelled protein, U is the concentration of the unlabeled protein stock solution used in the titration, and K_a is the association constant of the complex.

Protein-protein docking

The coordinates for ferric Cyt b_5, ferric and ferrous Cyt *c* were taken from the RCSB Protein Data Bank (PDB) entries 1CYO[61], 2YCC[48], and 1YCC[46], respectively. The charges on the individual haem atoms and all force constants were taken from top19x.heme and parallhdg.pro files of X-PLOR[157]. The net charges of Cyt *c* and Cyt b_5 used in the calculations were +6 and –8, respectively. The average relative solvent-accessible surface area for each residue was calculated using NACCESS[158]. The docking, performed with HADDOCK 1.2[130], is driven by ambiguous intermolecular restraints (AIRs), which are defined as follows. For both proteins, the residues with $\Delta\delta_{avg}$ above the mean plus half of the standard deviation and with a relative solvent accessibility above 50% are defined as 'active'. Then all neighbour residues of active residues exhibiting a relative surface accessibility above 50% are defined as 'passive'. The AIRs are defined as ambiguous distance restraints between the active residues of one protein and both the active and the passive residues of the other (see ref. 130 for details). The flexible segments for docking are defined from the active and passive residues used in the definition of AIRs ± two sequential residues.

The docking of Cyt b_5 with ferric or ferrous Cyt *c* was performed following the standard HADDOCK 1.2 protocols[130] with the modification that, due to the expected high dynamical nature of these complexes, the docking was performed during the semi-flexible simulated annealing part of the protocol rather than during the rigid body energy minimization. For this, the separation distance between the respective molecules was decreased to 50 Å and only rotational minimization was allowed in the initial rigid-body energy minimization to allow the molecules to face each other properly prior to docking. The 200 solutions from the rigid-body energy minimization were then subjected to the standard semi-flexible simulated annealing in the torsion angle space with an increased number of integration steps to allow for the docking during the simulated annealing. The entire docking run consisted of a rigid-body molecular dynamics search (12000 steps at 2000 K); first rigid-body simulated annealing (8000 steps from 2000 K to 500 K); second semi-flexible simulated annealing during which side chains at the interface are free to move (2000 steps from 2000 K to 50 K); and third semi-flexible simulated annealing during which both side chains and backbone at the interface are free to move (1000 steps from 500 K to 50 K).

The resulting structures were subjected to a final refinement in explicit water, clustered using a 2.0 Å backbone rmsd cut-off criterion, and sorted according to the intermolecular energy (sum of the van der Waals, electrostatic, and AIRs energy). The five lowest-energy structures of each cluster were selected for further analysis.

The docking of the tertiary Cyt b_5 complex with two ferric or ferrous Cyt *c* was performed with the standard HADDOCK 1.2 protocol[130] extended to deal with three molecules. 2000 docking solutions were generated during the rigid body energy minimization step and the best 200 were further subjected to the semi-flexible simulated annealing and final refinement in explicit solvent. For docking of the tertiary complex the AIR restraints were modified such as each active residue of Cyt b_5 had a restraint to all active and passive residues of both Cyt *c* molecules.