



Universiteit
Leiden
The Netherlands

Transient complexes of haem proteins

Volkov, O.M.

Citation

Volkov, O. M. (2007, February 28). *Transient complexes of haem proteins*. Leiden Institute of Chemistry/MetProt Group, Faculty of Mathematics and Natural Sciences, Leiden University. Retrieved from <https://hdl.handle.net/1887/11002>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/11002>

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

Chapter III

*Solution structure of the complex
between yeast cytochrome c and
cytochrome c peroxidase*

Abstract

The physiological complex of yeast *iso-1*-cytochrome *c* (Cyt *c*) and cytochrome *c* peroxidase (CcP) is a paradigm for the biological ET. In order to resolve a long-standing debate whether the crystal structure of this complex represents the dominant form in solution, we have solved the solution structure. To this end, site-specific spin-labelling in combination with paramagnetic relaxation enhancement (PRE) NMR spectroscopy were employed to generate intermolecular distance restraints, which then have been used as a sole input in structure calculations. We show that the solution structure is very similar to that observed in the crystal¹⁰⁷. In addition to that, further analysis of the intermolecular PREs provides the basis for assessing protein dynamics within this complex (Chapter IV).

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

Volkov, A. N., Worrall, J. A. R., Holtzmann, E. & Ubbink, M. Solution structure and dynamics of the complex between cytochrome *c* and cytochrome *c* peroxidase determined by paramagnetic NMR. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 18945-18950 (2006).

Introduction

Yeast Cyt *c* and CcP form a physiological complex, which plays a role in the organism's defence against high concentrations of intracellular peroxides⁶⁷ as CcP reduces H₂O₂ to H₂O using the electrons provided by ferrous Cyt *c*⁶⁶. This is perhaps the most studied ET protein complex (for the reviews see ref. 106 and 87), which has become a paradigm for the intermolecular ET in biological systems¹⁰⁵. Besides, it is one of the few transient complexes for which a crystal structure has been solved¹⁰⁷. A recent study has confirmed that the protein – protein orientation observed in the crystal is ET active¹¹⁸. However, it has remained a matter of debate whether this structure represents the dominant form in solution (ref. 87,106 and references therein). One of the reasons for the dispute is that, according to several studies, the complex is dynamic, and the crystal structure might therefore represent only a sub-population of protein orientations^{26,120-125}. In order to resolve this controversy, we report the structure of a well-defined Cyt *c* – CcP complex in solution.

The traditional NMR approach for structure determination of proteins and protein complexes in solution relies on intermolecular NOEs^{159,160}. However, this method has a limited use for weak protein complexes as it requires a relatively tight ($K_d \leq 10 \mu\text{M}$) interaction between the molecules¹⁶¹. For the last decade, structural characterisation of weak protein interactions has been greatly alleviated with the advent of the NMR techniques that use residual dipolar couplings and paramagnetic effects to provide long-range intermolecular restraints^{161,162}.

Paramagnetic NMR has been particularly useful for determining solution structures of metallo-proteins and their complexes as the paramagnetic metal cofactor, either native or substituted, exerts distance- and / or orientation-dependent effects, such as pseudocontact shifts and PREs^{163,164}. Usually, merely two 2D [¹H, ¹⁵N] HSQC spectra (one each for the paramagnetic form and a diamagnetic control) of the ¹⁵N labelled protein bound to its unlabelled partner suffice to determine intermolecular distance restraints for the amide protons. These can then be used in a rigid-body docking protocol to deduce the relative orientation of the two protein backbones in the complex¹⁶⁵. Despite undefined positions of the side-chains, this is often enough for an accurate description of a protein

complex. Using this simple approach, a number of metal-containing protein – protein complexes have been structurally characterised in solution¹⁶⁵⁻¹⁶⁹. Finally, with the use of various paramagnetic labels, such as genetically-engineered metal binding sites^{170,171}, metal-chelating tags¹⁷²⁻¹⁷⁶, and nitroxide spin-labels¹⁷⁷⁻¹⁸¹, the utility of paramagnetic NMR techniques has been further extended to structure determination of protein complexes that lack native paramagnetic centres.

The NMR resonances of Cyt *c* in the complex with CcP experience only small pseudocontact shifts, arising from the native, high-spin ($S = 5/2$) haem of CcP, that cannot be used for the structure determination (J. A. R. Worrall, personal communication). In order to introduce a stronger paramagnetic group, we have decided to attach a nitroxide spin-label (MTSL, Figure 3.1 A) to the surface of CcP and use the MTSL-induced PREs of the Cyt *c* backbone amides to generate intermolecular restraints for the structure calculation.

Site-specific spin-labelling is routinely employed for characterisation of protein interactions by electron paramagnetic resonance (EPR) spectroscopy^{182,183} and for the study of membrane proteins and protein-lipid associations by a tandem of EPR and solid-state NMR techniques^{184,185}. Although the past decade has seen the application of site-directed spin-labelling in combination with solution NMR to structural characterisation of individual proteins¹⁷⁷⁻¹⁸¹, to our knowledge, this approach has not yet been used for protein complexes. Here we show the utility of this method for the Cyt *c* – CcP complex and demonstrate that the intermolecular restraints derived solely from PREs suffice for an accurate structure determination of a protein complex in solution.

Results

Spin-labelling of CcP

Labelling of a protein with thiol-specific MTSL requires a surface-exposed cysteine (Figure 3.1 C). In order to attach MTSL at various positions on the surface of CcP, a number of single-cysteine variants have been produced (Figure 3.2). In each case the only

native cysteine residue of CcP (C128, Figure 3.2), which is partially accessible to MTSL, has been converted to alanine, so that each variant would be labelled at a single position.

For all variants, immobilization of the paramagnetic spin-label on the surface of CcP is evidenced by EPR spectroscopy (Figure 3.3). In each case, the room-temperature EPR signal of the attached MTSL (Figure 3.3 B-F) is broader than that of the free label (Figure 3.3 A). The line-widths of the EPR signals vary for different CcP-MTSL (Figure 3.3 B-F), reflecting the difference in the

rotational correlation time and, hence, the mobility of MTSL attached at different positions. For all CcP variants, the yield of spin-labelling was close to 100 % as estimated from double-integrated EPR spectra of mutant CcP-MTSL and a control sample containing a known amount of free MTSL (data not shown; for details see Materials and Methods). Another approach, based on comparison of the signal intensities and linewidths for the two samples (Equation 3.1 in Materials and Methods), was used to estimate the labelling efficiency from the EPR spectra acquired at 77 K. Both methods gave consistent results, differing by $\leq 5\%$.

In order to estimate rotational correlation time (τ_c) for MTSL attached to different CcP variants, we have attempted to simulate the RT EPR spectra using NLSL package¹⁸⁶. For N38C and N200C CcP-MTSL the spectra could be fitted with a minimum of two components, while for others no reliable fit has been obtained. An

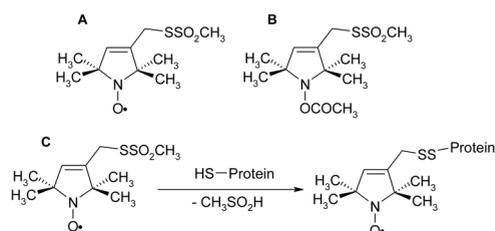


Figure 3.1. Labels used in the study. MTSL (A), MTS (B), and the reaction scheme of MTSL with a thiol group of a protein (C).

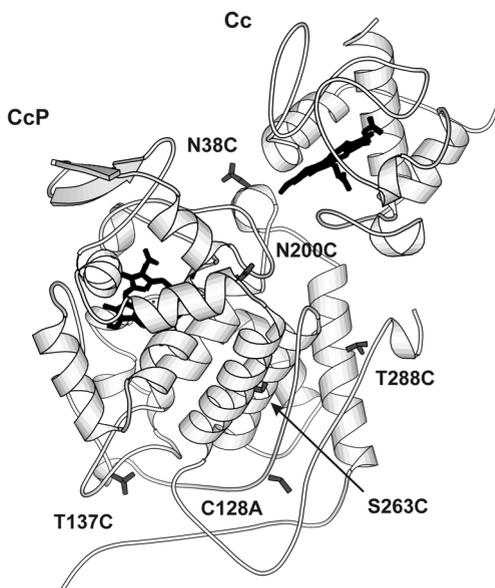


Figure 3.2. Positions of surface mutations on CcP. The residues of CcP used for mutagenesis are labelled and shown in grey sticks. Haem groups for both proteins are in black. Ribbon representation of the complex was drawn from the crystal structure (PDB entry 2PCC¹⁰⁷).

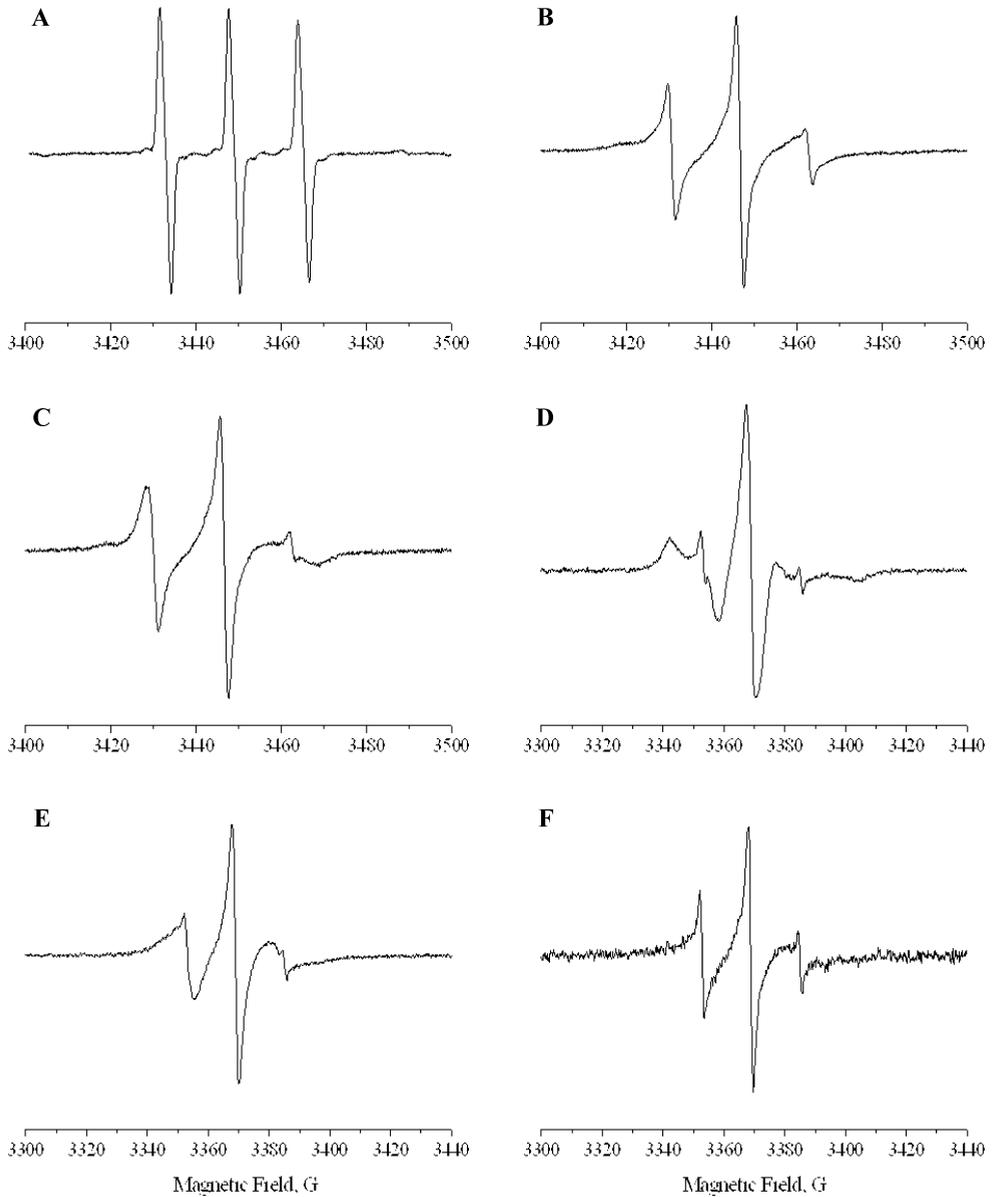


Figure 3.3. Room temperature X-band EPR spectra of free and bound MTSL. MTSL free (A) and attached to the CcP variants N38C (B), N200C (C), T288C (D), S263C (E), and T137C (F).

example of the simulated spectrum is shown in Figure 3.4.

To generate a diamagnetic control for the NMR experiments described below, all CcP variants have also been labelled with MTS, a diamagnetic analogue of MTSL (Figure 3.1 B). As evidenced by SDS-PAGE (data not shown), labelling of single-cysteine CcP with MTS abolishes the formation of the disulfide protein homo-dimers, indicating that all cysteine thiol groups are modified. Based on this observation and the similarity of the molecular structures of both labels, the yield of the labelling with MTS and MTSL was considered to be the same.

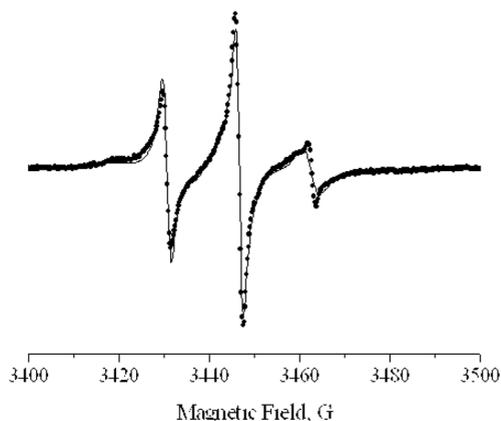


Figure 3.4. Simulated EPR spectrum of N38C CcP-MSTL. The solid line represents the global fit of two spectral components (with τ_c of 5 and 31 ns, and 41 and 59 % relative populations, respectively) to the data.

*PREs in the complex of ^{15}N Cyt *c* with CcP-MTSL*

In order to detect intermolecular paramagnetic effects of the spin-label attached to CcP on amide protons of ^{15}N Cyt *c*, 2D [^1H , ^{15}N] HSQC spectra of the paramagnetic complex (^{15}N Cyt *c* – CcP-MTSL) and the diamagnetic reference (^{15}N Cyt *c* – CcP-MTS) were recorded.

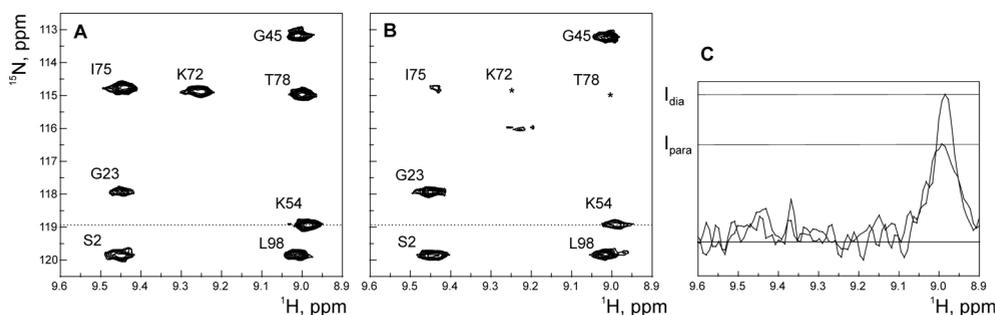


Figure 3.5. Intermolecular paramagnetic effects. (A, B) Parts of the ^{15}N – ^1H HSQC spectra of ^{15}N Cyt *c* bound to CcP N200C modified with a diamagnetic (A) or paramagnetic (B) label. The asterisks indicate resonances broadened beyond detection. (C) A slice through the amide resonance of K54 in the ^1H dimension, marked with a dotted line in (A, B). The lines marked I_{para} and I_{dia} indicate the intensities of the resonance in the paramagnetic and diamagnetic complex, respectively.

were acquired for all CcP variants. As illustrated for N200C CcP, comparison of the two spectra reveals that in the paramagnetic complex some Cyt *c* resonances experience decrease in intensity and concomitant increase in the line-width (Figure 3.5). These spectral changes are due to paramagnetic relaxation enhancement (PRE), the increase in the transverse relaxation rate ($R_{2,para}$), of the affected nuclei caused by the unpaired electron of MTSL. The paramagnetic effect is distance dependent. This explains why some of the Cyt *c* resonances – which must be close to the spin-label – are affected (*e.g.* K72, I75, or T78 in Figure 3.5 A and B), while others – which must be further away from MTSL – are not (*e.g.* S2 or L98 in Figure 3.5 A and B). The ratios of Cyt *c* peak intensities in the paramagnetic (I_{para}) and diamagnetic (I_{dia}) complexes with various CcP variants (Figure 3.6) are converted into PREs and then further into intermolecular distance restraints (for details see Materials and Methods), which are then used to calculate the structure of the protein complex.

To ascertain whether a spin-label attached to the surface of CcP interferes with Cyt *c* binding, we have monitored the chemical shift perturbations of ^{15}N Cyt *c* resonances in the complex with different CcP-MTS (Figure 3.7). In all cases the perturbations are identical to those in the wt complex at the same protein ratio, indicating that each of the attached MTSL molecules does not alter the binding.

Structure calculations of the Cyt *c* – CcP complex

A set of PRE-derived intermolecular distance restraints, defined from the averaged position of the oxygen atom of MTSL (see below) to the backbone amide protons of Cyt *c*, has been calculated for Cyt *c* in the complex with each of the CcP variants spin-labelled at positions N38C, N200C, S263C, and T288C (see Materials and Methods and Table C1 in Appendix C). The restraints from all four MTSL positions have been used simultaneously in the structure calculation of the protein complex.

Although MTSL attached to the protein is less mobile than that in solution (see above), it still possesses a degree of rotational freedom. In order to take into account the mobility of MTSL attached to the surface of CcP, we performed ensemble averaging of the intermolecular distance restraints using multiple starting positions of the MTSL oxygen atoms. As described in Materials and Methods, these starting positions have been generated

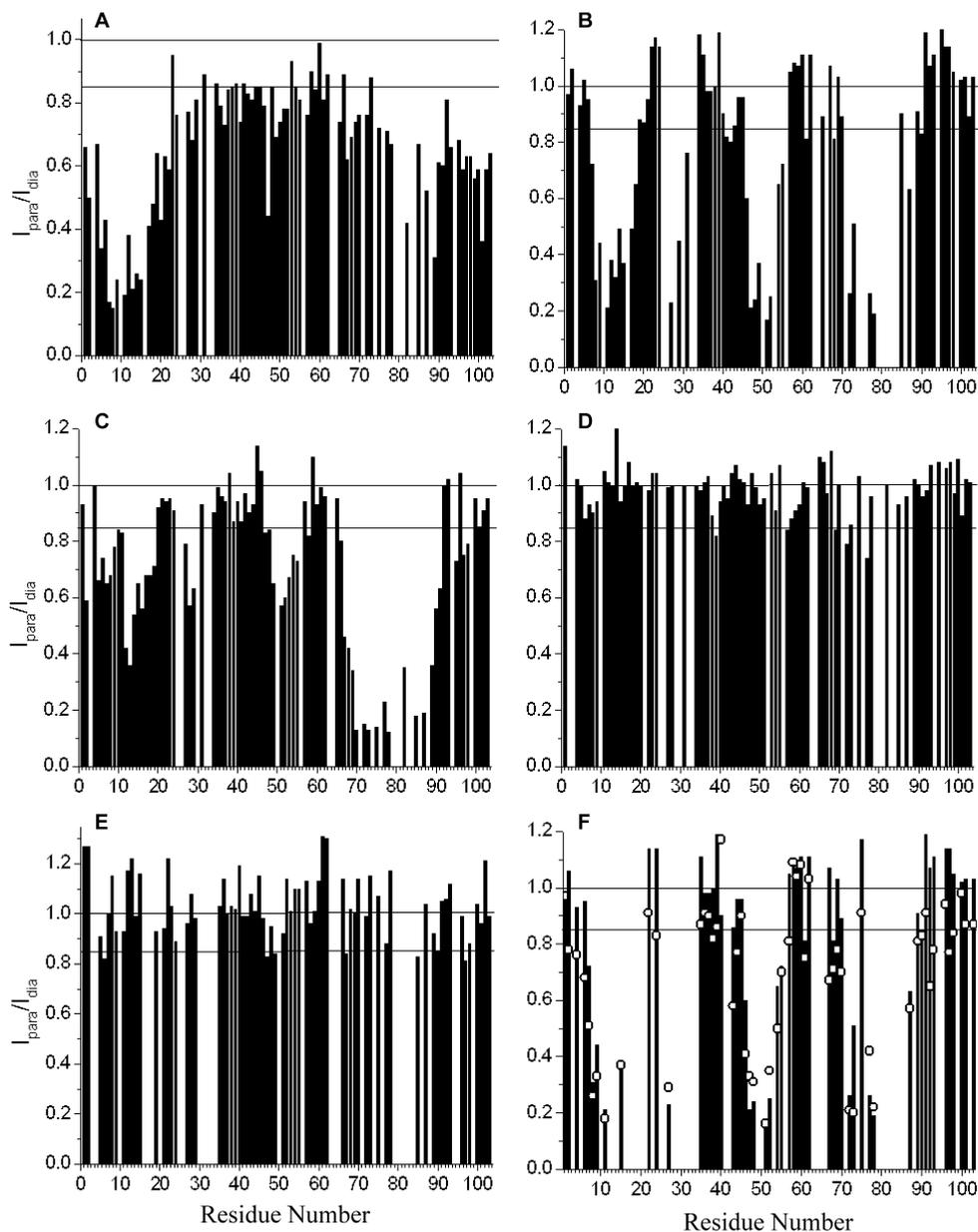


Figure 3.6. PREs of Cyt *c* amides in the complex with CcP-MTSL. Plots of intensity ratios for all Cyt *c* backbone amide protons in the complex with paramagnetically (I_{para}) and diamagnetically (I_{dia}) labelled CcP at position (A) N38C, (B) N200C, (C) T288C, (D) S263C, (E) T137C, and (F) N200C (the control sample; Chapter IV). Horizontal lines denote intensity ratios of 1.0 and 0.85. The residues with the intensity ratios below 0.85 are considered to be affected by MTSL (Materials and Methods and ref. 179). In (F) the intensity ratios of Cyt *c* in the 1:1 complex with CcP (open circles) are given for comparison.

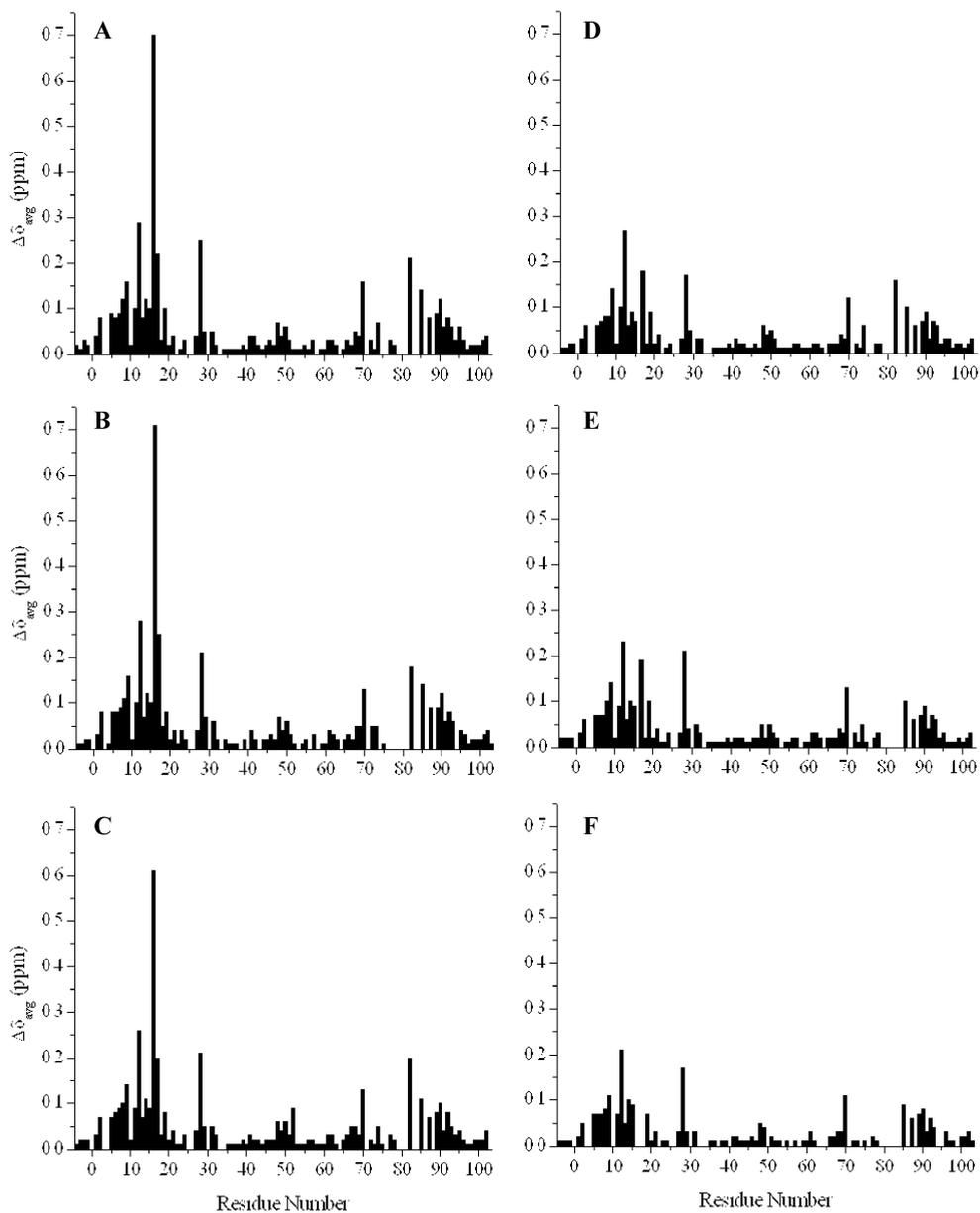


Figure 3.7 Averaged amide chemical shift perturbations ($\Delta\delta_{\text{avg}}$) of ^{15}N Cyt *c* resonances upon binding to CcP. The $\Delta\delta_{\text{avg}}$ for the backbone amide resonances of Cyt *c* in the 1:1 complex with wt (A) or T288C (B), S263C (C), N38C (D), N200C (E), and T137C (F) CcP-MTS. The chemical shift perturbations are extrapolated to the 100 % bound form (for details see Chapter V). Note that Q16 resonance, which exhibits the largest $\Delta\delta_{\text{avg}}$ in the wt complex, is not observed in Cyt *c* complexes with N38C (D), N200C (E), and T137C (F) CcP-MTS.

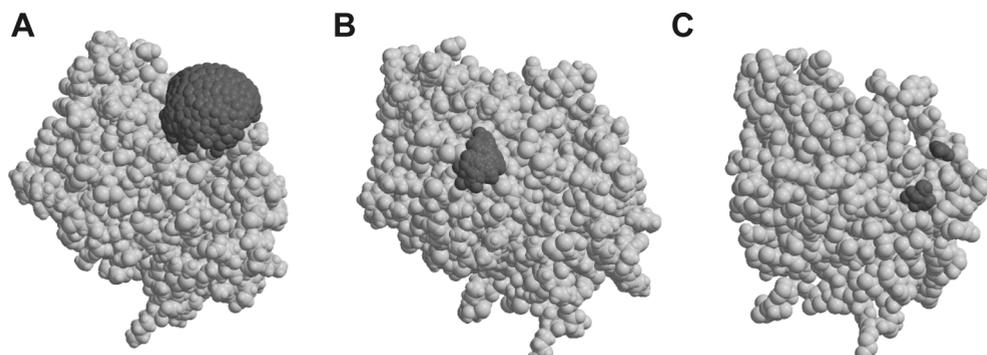


Figure 3.8. Possible positions of the MTSL oxygen atom. MTSL attached to CcP at N38C (A), N200C (B), and T288C (C). MTSL oxygen atom positions are shown in dark grey.

by systematic rotation of the attached MTSL. For three spin-labelled CcP variants, the ensembles of possible MTSL oxygen atom coordinates are shown in Figure 3.8. From each ensemble, four or ten representative oxygen positions have been selected and used for averaging (for details see Materials and Methods).

Rigid-body structure calculation of the Cyt *c* – CcP complex has produced a single cluster of solutions (68 structures, 0.19 ± 0.12 Å pair-wise backbone rmsd from the lowest

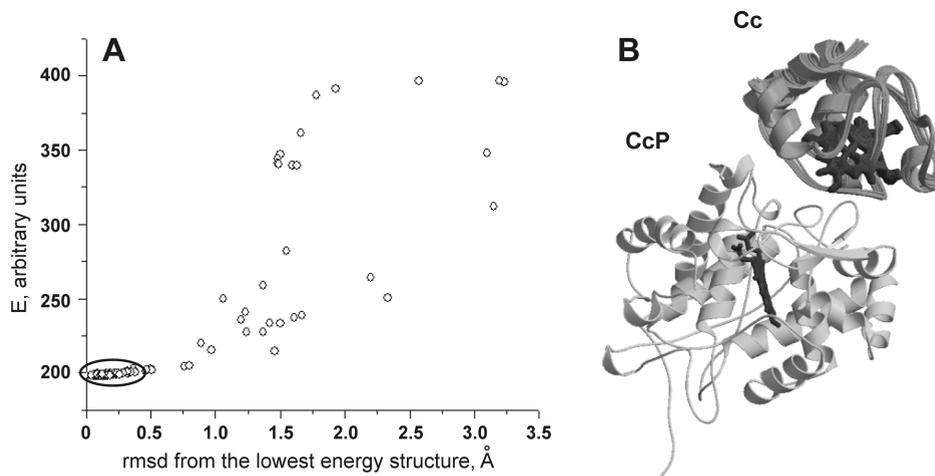


Figure 3.9. Restrained rigid-body structure calculations for Cyt *c* – CcP complex. (A) Total energy as a function of the backbone rmsd from the lowest energy structure. The lowest-energy cluster of solutions is circled. (B) Best 20 structures of the lowest-energy cluster, with CcP and Cyt *c* in light and dark grey, respectively. The average rmsd from the best structure for the backbone atoms of Cyt *c* is 0.7 ± 0.2 Å. Haem groups for both proteins are shown in sticks.

energy structure, and total energy of 199.2 ± 1.1 arbitrary units, where the error is a standard deviation from the mean; Figure 3.9 A). As no force-field has been implemented during the rigid-body calculations (only NMR restraints and van der Waals repulsion were active, see Materials and Methods), the energy has no physical meaning. It is merely used as a measure of how well a given structure satisfies the experimental restraints. The ensemble of twenty lowest-energy structures for the Cyt *c* – CcP complex, with an average rmsd of 0.7 ± 0.2 Å from the best structure for the backbone atoms of Cyt *c*, is shown in Figure 3.9 B. The molecular coordinates of the twenty best structures have been deposited to the Protein Data Bank (accession number 2GB8). The violations of restraints in the lowest energy structure are reported and discussed in Chapter IV.

Discussion

Spin-labelling of CcP

The reaction of an MTSL spin-label with a cysteine (Figure 3.1 C) involves deprotonation of the -SH group¹⁸⁷. As the pK_a for the cysteine thiol ranges from 7.4 to 9.1, the reaction is expected to be favoured by a high pH¹⁸⁷. For all single-cysteine variants of CcP, it was found that spin-labelling at pH 8.0 (see Materials and Methods) produced consistent results, with yields close to 100 % as estimated from the EPR spectra.

Due to the decrease in τ_c associated with immobilization of the spin-label on the protein surface, the X-band EPR signals of all spin-labelled CcP variants are broader than those of the free MTSL in solution (Figure 3.3). Differences in signal shapes for various CcP-MTSL reflect different mobility of the attached spin-label, with a broader signal corresponding to a less flexible MTSL. Visual inspection of the spectra B-F in Figure 3.3 suggests that for different CcP variants the mobility of the attached MTSL decreases in the order T137C > N38C > N200C > S263C > T288C. This agrees with the results of the molecular modelling of bound MTSL (Figure 3.8), which shows similar trend for the extent of the rotational freedom. Modelling of MTSL attached to T288C CcP reveals two sterically-separated populations (Figure 3.8 C), which are expected to have different mobility and, therefore, give rise to differing EPR signals. Interestingly, the EPR spectrum

of T288C CcP-MTSL exhibits superposition of two signals (Figure 3.3 D), which may indeed stem from two orientational sub-populations of the attached MTSL.

Only few RT EPR spectra of various CcP-MTSL could be simulated using NLSL algorithm¹⁸⁶ due to a cumbersome iteration procedure not optimised for surmounting local minima and an unknown hyperfine tensor, assumed to be equal to that determined for another spin-labelled protein (for details see Materials and Methods and ref. 188). As illustrated in Figure 3.4, simulations of the EPR spectra of CcP-MTSL suggest the presence of two components with markedly different τ_c values. In a recent high-resolution EPR study of the spin-labelled copper protein azurin, two spectral components have been directly observed in the J-band EPR spectra and attributed to two forms of MTSL differing in hydrogen bonding of the nitroxide group¹⁸⁸. According to another recent report, the two forms could correspond to two low-energy conformations of the disulfide bond between the cysteine and MTSL¹⁸⁹. As the rotation around a disulfide bond is expected to produce only a moderate change in mobility, the large difference in τ_c for the two components in our spectra is better explained by the hydrogen-bonding hypothesis.

Contrary to several reports¹⁷⁷⁻¹⁷⁹, the *in situ* reduction of the attached MTSL with up to 5-fold excess of ascorbate at RT has never been complete as evidenced by residual EPR signal (data not shown). The incomplete reduction has been also observed for another metallo-protein spin-labelled in our group (M. G. Finiguerra, personal communication). To circumvent this problem, we labelled the single-cysteine proteins *de novo* with MTS (an acetylated analogue of MSTL, Figure 3.1 B), and used those as diamagnetic controls. The same labelling strategy, necessitated by the solvent-inaccessibility of MTSL attached to lipid-covered sites of a membrane protein, has been employed in a recent NMR study¹⁸¹.

PRE-based distance restraints

The presence of a paramagnetic spin-label is clearly reflected in the NMR spectrum of ¹⁵N Cyt *c* bound to CcP-MTSL (Figure 3.5). Due to the relaxation enhancement brought about by the unpaired electron of MTSL, the resonances of the Cyt *c* residues that are close to the label decrease in intensity or vanish altogether (Figure 3.5 B). The change in intensity ($I_{\text{para}}/I_{\text{dia}}$) of the NMR peaks is used as a gauge of the distance-

dependent intermolecular paramagnetic effect and can be converted into the actual distances between the affected nuclei and MTSL (Equations 3.2 and 3.3 in Materials and Methods).

From the plots of $I_{\text{para}}/I_{\text{dia}}$ for Cyt *c* bound to different CcP-MTSL (Figure 3.6), it is clear that Cyt *c* is affected by the spin-label in the complexes with N38C, N200C, and T288C CcP-MTSL (Figure 3.6 A-C), while almost no effects are observed upon binding to S263C and T137C CcP-MTSL (Figure 3.6 D and E). The former spin-labelling positions are located close to the crystallographic binding site of Cyt *c*, while the latter two are further away towards the side and the back of CcP (Figure 3.2). Thus, the observed paramagnetic effects (plotted in Figure 3.6) suggest that the binding site of Cyt *c* in the complex with CcP in solution is similar to that observed in the crystal. In addition, the absence of paramagnetic effects on Cyt *c* in the complexes with S263C and T137C CcP-MTSL implies that no aspecific protein binding takes place in solution.

As explained in Materials and Methods, the $I_{\text{para}}/I_{\text{dia}}$ intensity ratios for the NMR resonances are converted into PREs and then further into distance restraints (Table C1 in Appendix C). As suggested by our data, MTSL allows to probe for distances in the range of 14 to 22 Å, in agreement with the values reported for other spin-labelled proteins^{177,179,181}. At the same time, our findings disagree with the results of yet another study that has suggested the broader range of 8 to 35 Å¹⁷⁸. The reported values seem unlikely as, due to the inverse sixth-power distance-dependence of the PRE (Equation 3.3 in Materials and Methods), the effects corresponding to the larger distances would be smaller than what could be measured in the NMR experiment. Similarly, the distances shorter than 12 – 13 Å could hardly be estimated as the large PREs would cause substantial intensity decrease of the affected resonances, with the resulting peak heights well below the noise level of the HSQC spectrum and, therefore, effectively zero.

For all restraints the calculated value of the distance, termed ‘target distance’, is assigned bounds, within which this value can vary. This is done in order to 1) take into account the experimental error and 2) increase the convergence of structure calculations. As discussed below, we have found that 4 Å bounds satisfy both requirements.

Among the contributions to the total error in the calculated distances are uncertainties in the measurements of the peak heights and widthsⁱ, error in the estimate of τ_c , and small discrepancies in concentrations of the compared paramagnetic and diamagnetic samples. Due to the sixth-power root dependence (Equation 3.3 in Materials and Methods), large uncertainties in $R_{2,\text{dia}}$ and τ_c are translated into only small distance errors¹⁷⁹. At the same time, the factors affecting the peak intensities, of which the concentration discrepancy is the largest, contribute greatly to the total error of this method.

In order to estimate the systematic and random errors arising from the use of two samples with matching protein concentrations, MTSL and MTS labelling and the NMR experiments have been repeated several times for some of the CcP variants. Each time consistent results have been produced (data not shown), suggesting that the observed paramagnetic effects are reproducible. The largest variation in $I_{\text{para}}/I_{\text{dia}}$, as seen for the repeated experiments, converts into 2 Å distance error, which is well within the 4 Å margin assigned to the distance restraints.

Conversion of a PRE (*i.e.* $R_{2,\text{para}}$) into the distance (Equation 3.3 in Materials and Methods) requires the rotational correlation time (τ_c) of the vector between the backbone amide proton and the unpaired electron of MTSL. Although τ_c is different for each amide proton, the use of a single value for all nuclei has been justified^{177,178}. As discussed above, an attempt to estimate τ_c of MTSL from the EPR spectra of CcP-MTSL has been compounded by the presence of two spectral forms with markedly different τ_c (Figure 3.4). This is why we have used the Einstein-Stokes equation (Equation 3.4 in Materials and Methods) to calculate the upper limit of τ_c for the Cyt *c* – CcP complex. The resulting value of 16 ns is in good agreement with that for the previously reported protein system of a similar size¹⁷⁹. Given a wide range of the reported τ_c values used for conversion of PREs into distances¹⁷⁷⁻¹⁷⁹, two control sets of distance restraints were generated with τ_c of 12 and 4 ns. In each case, structure calculations produced the same best solution as that for the restraints based on τ_c of 16 ns (data not shown).

ⁱ The latter, performed for the diamagnetic control, is required for the estimate of $R_{2,\text{dia}}$ used in PRE calculation; see Equation 3.2 in Materials and Methods.

Structure calculations

The present approach allows to define the distance restraints for the backbone amides of Cyt *c*. As our experiments did not provide any structural information on the protein side-chains, we have utilized a rigid-body protocol, which appears to be sufficient for the structure calculation of the Cyt *c* – CcP complex. Multiple rigid-body docking runs starting from random positions of the two proteins have consistently produced a single cluster of low-energy solutions (Figure 3.9), indicating that, in addition to accommodating the experimental error (see above), the restraint bounds of $\pm 4 \text{ \AA}$ provide a good convergence of the structure calculations. The use of tighter bounds has been avoided as this diminishes the convergence of solutions¹⁷⁹.

Use of the intermolecular distance restraints from several MTSL positions greatly increases both the accuracy and the precision of the structure calculations, as compared to the case of a single MTSL (data not shown). Similarly, averaging of the restraints over a number of possible MTSL orientations further enhances the quality of the final solutions. In the present work, four orientations for each MTSL position have been used in ensemble averaging (for details see Materials and Methods). Using ten MTSL orientations produced similar results (data not shown), consistent with the finding that averaging over a larger set of atoms does not further enhance the quality of the calculations¹⁹⁰.

Solution structure of the Cyt *c* – CcP complex

As the protein side-chains have not been analysed in this study, only the positions of the Cyt *c* and CcP backbones in the best solution structure and the crystal structure¹⁰⁷ are compared (Figure 3.10). The backbone rmsd of 2.2 \AA between the Cyt *c* molecules from the two structures implies that these are very similar. This finding provides the first direct evidence that the crystal structure represents the dominant form of the Cyt *c* – CcP complex in solution as was proposed by earlier studies^{127,128} and refutes the suggestion that, due to protein dynamics, the complex in solution is most likely unstructured¹²³. We believe that

our conclusions will close a long-standing debate whether the Cyt *c* – CcP crystal structure is a valid description of the complex in solution.

In the crystal structure of the Cyt *c* – CcP complex, Q16 residue of Cyt *c* forms a hydrogen bond with its own backbone¹⁰⁷, which, rather than being an artefact of crystallization, is induced by the binding¹²⁹. In order to verify the influence of the complex-induced conformational change of the Q16 side-chain on the results of the structure calculations, we have performed several runs with the molecular coordinates of the proteins taken from: 1) the crystal structure of the complex¹⁰⁷, as described above; 2) individual X-ray structures^{48,79}; and 3) CcP from an individual structure and Cyt *c* from that of the complex.

In the structure calculations with two individual proteins, the extended side-chain of Q16, located at the interface of the complex, protrudes from the surface of Cyt *c* thereby preventing a closer approach to CcP and forcing the proteins into the arrangement that is laterally translated relative to the crystal structure. This results in the 6.2 Å rmsd between the Cyt *c* molecules from the lowest-energy solution and the crystal structures of the complex. In contrast, docking of individual CcP and Cyt *c* taken from the structure of the complex (with Q16 “folded in”) produces the solution with the rmsd of 2.5 Å between the Cyt *c* molecules, which is very similar to that obtained in the run with both proteins coming from the crystal structure. It transpires that the use of Cyt *c* with the binding-induced, “folded” configuration of Q16 increases the accuracy of structure calculations and produces the solutions closely

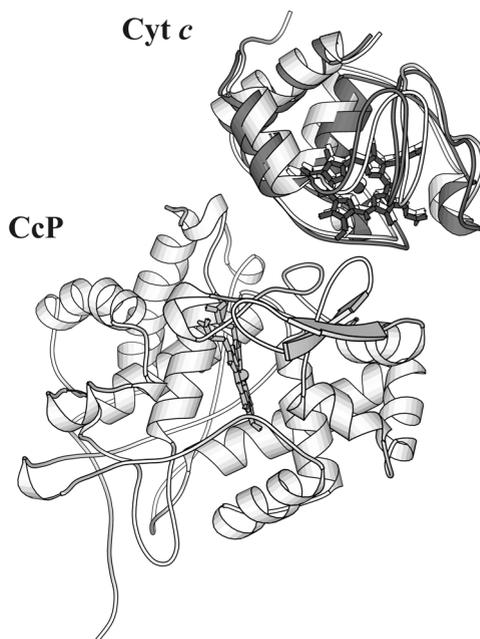


Figure 3.10. Comparison of the solution and crystal structures of the Cyt *c* – CcP complex. The CcP molecules are superimposed, and the Cyt *c* backbones in the solution and crystal structures are shown in dark and light grey, respectively. Haem groups for both proteins are in sticks. The positional backbone rmsd between the two Cyt *c* molecules is 2.2 Å. The coordinates for the crystal structure were taken from the PDB entry 2PCC¹⁰⁷.

resembling the crystal structure of the complex. Apparently, this rearrangement is crucial for the formation of the correct complex.

Concluding remarks

In this study we have illustrated the use of site-specific spin-labelling in combination with NMR spectroscopy for solving the structure of the Cyt *c* – CcP complex. The simplicity and the time-efficiency of this method, which requires introduction of a cysteine group and only two HSQC spectra per a single-cysteine variant, make it an attractive approach for rapid structure determination of other protein – protein, protein – DNA, and protein – RNA complexes in solution.

Materials and Methods

Protein preparation

In order to prepare single-cysteine CcP variants, site-directed mutagenesis was carried out using the Quik Change™ polymerase chain reaction protocol (Stratagene, La Jolla, CA) with the plasmid CCP(MKT) as a template⁷³ (for details see Appendix D). All constructs have been verified by DNA sequencing. CcP variants and isotopically-enriched ¹⁵N Cyt *c* (T-5A/C102T) were expressed in *E. coli* and purified following the published procedures^{41,73,150} (for the detailed protocols see Appendix B). Concentrations of ferric Cyt *c* and five-coordinated high-spin ferric CcP, used in this work, were determined according to the optical absorbance peaks at 410 nm ($\epsilon = 106.1 \text{ mM}^{-1} \text{ cm}^{-1}$)¹⁵² and 408 nm ($\epsilon = 98 \text{ mM}^{-1} \text{ cm}^{-1}$)¹⁹¹, respectively.

Purified single-cysteine CcP variants were incubated with 10 mM DTT in 0.1 M Tris-HCl pH 8.0, 0.1 M NaCl for 2 hours at RT, in order to reduce intermolecular disulfide bonds. DTT was removed by passing the CcP solution through a PD-10 column (Amersham Pharmacia, Uppsala, Sweden). The resulting monomeric protein was reacted with a 7 to 10-fold excess of MTSL or MTS (Figure 2.1) and incubated overnight at RT. MTSL [(1-Oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate] and MTS

[(1-Acetyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate] were purchased from Toronto Research Chemicals (North York, ON, Canada) and used without further purification. Upon completion of the reaction, protein solution was passed through a PD-10 column to remove any unreacted label, followed by exchange into 20 mM NaP_i 0.1 M NaCl pH 6.0, and concentrated by ultracentrifugation using Amicon Ultra concentrators (Millipore, Billerica, MA).

NMR experiments

NMR samples contained 0.3 to 0.5 mM of a 1:1 complex of ¹⁵N Cyt *c* with CcP-MTSL or CcP-MTS in 20 mM NaP_i 0.1 M NaCl pH 6.0, 6 % D₂O for lock, and 0.1 mM CH₃CO¹⁵NH₂ as an internal reference. The pH of the samples was adjusted to 6.00 ± 0.05 with small aliquots of 0.1 M HCl or 0.1 M NaOH. Measurements were performed at 301 K on a Bruker DMX600 spectrometer equipped with triple-resonance TXI-Z-GRAD probe (Bruker, Karlsruhe, Germany). 2D [¹⁵N, ¹H] HSQC spectra were obtained with 1024 and 256 complex points in the direct and indirect dimensions, respectively, and spectral widths of 32 ppm (¹⁵N) and 16 ppm (¹H). All data were processed with AZARA 2.7¹⁵³ and analysed in ANSIG for Windows^{154,155}. Assignments of the ¹⁵N and ¹H nuclei of Cyt *c* were taken from the previous work¹²⁹. In this study a number of Cyt *c* amides (A3, F10, H26, L32-H33, H39, N56, N63, M64, E66, Y74, K79-A81, G83-G84, K86, E88, and L94) were either not observed or not analysed due to spectral overlap.

EPR experiments

RT measurements have been performed on a Bruker Elexsys 600/608 spectrometer. The samples contained 10 – 20 µl of 0.2 – 0.5 mM CcP-MTSL or free MTSL in 20 mM NaP_i 0.1 M NaCl pH 6.0. X-Band EPR spectra were acquired with 3450 G or 3370 G central field, 100 G or 140 G sweep width, 1.5 G modulation amplitude, 82 s convolution time, and 5 to 25 scans. X-band EPR spectra of frozen solutions have been acquired on a JEOL JES-RE2X spectrometer operating at 77 K with 3200 G central field, 500 G sweep-width, 7.9 G frequency modulation, 4 minutes acquisition time, and 1 to 5

scans. The samples contained 100 – 150 μl of 0.2 – 0.5 mM CcP-MTSL or free MTSL in 20 mM NaP_i 0.1 M NaCl pH 6.0.

To estimate the yield of spin-labelling, the RT EPR spectra of CcP-MTSL and free MTSL were double-integrated in Microcal™ Origin® 6.0. Alternatively, the spin-labelling efficiency (η) was estimated from the EPR spectra of frozen solutions using Equation 3.1¹⁹²:

$$\eta = \frac{C_0 I W^2}{C I_0 W_0^2} \quad (3.1)$$

where C is the concentration of the CcP-MSTL, I is the intensity and W is the width of the M(0) hyperfine line of the EPR signal, while C_0 , I_0 , and W_0 are those for the free MTSL used as a reference. Simulations of the RT X-band EPR spectra have been performed with NLSL package¹⁸⁶ using the following g- and A-tensors for both spectral components: $g_{xx} = 2.0091$, $g_{yy} = 2.0068$, and $g_{zz} = 2.00297$; $A_{xx} = 4.0$ G, $A_{yy} = 4.5$ G, and $A_{zz} = 40.0$ G.

Determination of distance restraints

For each observed amide proton of Cyt c , an MTSL-induced PRE was calculated from Equation 3.2¹⁷⁹:

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2,dia} \exp(-tR_{2,para})}{R_{2,dia} + R_{2,para}} \quad (3.2)$$

where I_{para} and I_{dia} are measured intensities of HSQC peaks of Cyt c in the complex with CcP-MTSL and CcP-MTS, respectively (Figure 3.5 C); $R_{2,dia}$ is the transverse relaxation rate of Cyt c amide protons in the complex with CcP-MTS; $R_{2,para}$ is the paramagnetic contribution to the relaxation rate (PRE); and t is the total INEPT evolution time of the HSQC (9 ms). For the residues which resonances disappear in the paramagnetic spectrum, an upper limit for I_{para} was estimated from the noise level of the spectrum.

For all amide protons, the $R_{2,dia}$ has been estimated from the HSQC peaks of the Cyt c in the complex with CcP-MTS. The HSQC spectra have been zero-filled up to 2048 and 512 complex points in the direct and indirect dimensions, respectively, and processed with a 15 Hz line-broadening exponential multiplication window-function in the ¹H

dimension. For each peak the width at half-height ($\Delta\nu_{1/2}$) in the proton dimension has been extracted from the Lorentzian fit using Microcal™ Origin® 6.0. After correction for the artificial line-broadening, the $\Delta\nu_{1/2}$ has been converted into $R_{2,dia}$ using $R_{2,dia} = \pi\Delta\nu_{1/2}$.

Calculated PREs were converted into distances using Equation 3.3¹⁷⁹:

$$r = \sqrt[6]{\frac{\gamma^2 g^2 \beta^2 \tau_c}{20R_{2,para}} \left(4 + \frac{3}{1 + \omega_h^2 \tau_c^2} \right)} \quad (3.3)$$

where r is the distance between the unpaired electron of the MTSL and a given amide proton of Cyt c ; τ_c is the correlation time of the electron-nucleus vector; ω_h and γ are proton Larmor frequency and gyromagnetic ratio, respectively; g is the electronic g-factor; and β is the Bohr magneton.

Given an association constant $K_a = (1.8 \pm 0.2) \times 10^5 \text{ M}^{-1}$ (Chapter V), 88 to 91 % of Cyt c is bound to CcP in a 1:1 complex for the protein concentrations used. Thus, the observed PRE ($R_{2,obs}^{para}$) represents a fraction (F) of the actual PRE (R_2^{para}), so that $R_{2,obs}^{para} = FR_2^{para}$, with $F = 0.88 - 0.91$. To obtain the R_2^{para} used in Equation 3.3, all experimentally determined PREs were corrected by $1/F$.

The upper limit for the τ_c has been calculated from the Einstein-Stokes equation, Equation 3.4¹⁹³:

$$\tau_c (ns) = \frac{4\pi\eta_s r_h^3}{3k_B T} \quad (3.4)$$

where η_s is the viscosity of solution in centipoises, k_B is the Boltzmann constant, T is the temperature, and r_h is a hydrodynamic radius of the protein complex that was estimated from Equation 3.5¹⁹³:

$$r_h = \sqrt[3]{\frac{3\bar{V}M_r}{4\pi N_A}} + r_w \quad (3.5)$$

where \bar{V} is the specific weight of a protein, M_r is molecular weight of the complex, N_A is the Avogadro number and r_w is the hydration radius. Assuming $\bar{V} = 0.73 \text{ cm}^3/\text{g}$ and $r_w = 1.6 \text{ \AA}$ ¹⁹³ for the Cyt c – CcP complex ($M_r = 46.2 \text{ kDa}$), Equation 3.5 gives $r_h = 25 \text{ \AA}$. Feeding it into Equation 3.4 and assuming the normal conditions and $\eta_s = \eta_s (0.1 \text{ M NaCl}) = 1.013$ centipoises¹⁹⁴ gives a τ_c of 16 ns.

Three classes of intermolecular distance restraints¹⁷⁹ used in structure calculations were defined (Table C2 in Appendix C). Residues that are strongly affected by MTSL and which resonances disappear in the paramagnetic spectrum are restrained with only an upper bound. Those not affected by MTSL have only a lower limit. Finally, the residues affected by the spin-label and which resonances are observed in the paramagnetic spectra are restrained with both upper and lower bounds.

Structure calculations

Coordinates of both proteins were taken from the crystal structure of the complex (PDB accession code 2PCC¹⁰⁷). For the docking of individual proteins, the molecular coordinates from PDB entries 2YCC⁴⁸ for Cyt *c* and 1ZBY⁷⁹ for CcP were used. The surface cysteine mutations of CcP were introduced *in silico*, followed by addition of the MTSL atoms and energy minimization of the labelled protein in Xplor-NIH 2.13¹⁹⁵ using parallhdg.pro force-field of X-PLOR¹⁵⁷. For MTSL we have used the parameter and topology files that were available in our group.

In order to take into account mobility of the MTSL attached to the surface of CcP, we performed r^{-6} ensemble averaging of the intermolecular distance restraints, using four MTSL orientations. To generate these, the attached MTSL was systematically rotated around five single bonds that join its ring to the C $_{\alpha}$ atom of the cysteine and only the sterically-allowed conformers were retained for each mutant. From these, four representative orientations have been selected.

A set of distance restraints for MTSL attached to four positions on CcP (Table C1 in Appendix C) was used to calculate the structure of the protein complex in Xplor-NIH 2.13¹⁹⁵. First, a rigid-body docking of the protein molecules was carried out (for the input file see Appendix E). Only two energy terms, corresponding to restraints and van der Waals forces, were specified during this step. The latter were defined as repel forces for the protein atoms, while set to zero for MTSL. During the second step, 30 to 40 best structures were subjected to energy minimization and side-chain dynamics with fixed positions of backbone atoms for both proteins and active van der Waals parameters for MTSL. For the refined structures, the entire docking procedure was repeated until no further reduction in energy has been observed.