Transient complexes of haem proteins

Volkov, O.M.

Chapter V

Control of the specificity in the transient complex of yeast cytochrome c and cytochrome c peroxidase

Abstract

We have found that a single residue of yeast Cyt *c*, R13, controls the specificity of the interaction between Cyt *c* and CcP. This finding came from the binding studies of the wt Cyt *c* and two Cyt *c* variants, R13A and R13K, using isothermal titration calorimetry (ITC) and nuclear magnetic resonance (NMR) chemical shift perturbation mapping. While the interaction of the wt Cyt *c* with CcP is exothermic, the binding of both variants – resembling that of the horse Cyt c^{202} – is endothermic and entropy-driven, with the association constant four times lower than that of the wt complex. The size of the chemical shift perturbations in the NMR spectra of Cyt *c* variants upon binding to CcP is greatly reduced as compared to that of the wt complex, indicating that the variant complexes are more dynamic. For R13A Cyt *c* – CcP complex, this is further confirmed by site-specific spin-labelling in combination with paramagnetic relaxation enhancement (PRE) NMR spectroscopy. We show that, while the solution structure of the dominant form of this complex is the same as that of the wt complex (Chapter III), the proteins occupy the well-defined, single-orientation form for ~ 25 % of the lifetime of the complex, with the rest of the time spent in the dynamic encounter state. This contrasts sharply with the association of the wt Cyt *c* and CcP that spend > 70 % of the time in the single-orientation form (Chapter IV). Our results suggest that, while only moderately affecting the affinity, R13 controls the specificity of the Cyt *c* – CcP interaction. We propose that R13 stabilizes the complex by a cation-π interaction with Y39 residue of CcP, which would explain the drastic difference in the binding behaviour of the wt Cyt *c* and the charge-conserved R13K variant. Furthermore, we discuss a possible role of R13 in solvent organization at the Cyt *c* – CcP interface.

The results presented in this chapter will be published as:

Volkov, A. N., Worrall, J. A. R. & Ubbink, M. A specificity hot spot in the complex of yeast cytochrome *c* and cytochrome *c* peroxidase (manuscript in preparation).

Introduction

Protein-protein interactions are mediated by intermolecular interfaces of varying size and chemical diversity[8]. In some protein complexes, single interfacial residues can contribute a large fraction of the binding free energy[203,204]. Such "hot spots" of binding are characteristic of the static complexes, where the binding partners have co-evolved to be highly optimised for the interaction with each other[9,204]. As a result of this optimisation, the removal of a single residue from the interface of a static complex can shatter the intricate network of the steric and chemical complementarity between the two protein surfaces, thereby severely impairing the binding.

Unlike static protein complexes, transient complexes are formed between proteins that can interact with a variety of partners (Chapter I). Often being messengers between different biomolecules, these proteins are optimised for diversity, rather than specificity. Although the macromolecular recognition in the transient interactions remains poorly understood[4], it is believed that surface complementarity and specific intermolecular contacts are less important in stabilization of transient – as opposed to static – protein complexes[4,9]. However, it is unclear whether binding hot spots, ubiquitous in the static complexes, also orchestrate transient protein interactions. To clarify this point, we have investigated the role of a putative hot spot residue in the formation of a transient complex between Cyt *c* and CcP.

Identification of potential hot spot residues in protein interfaces has been made easier by an insightful analysis of Bogan and Thorn[204]. It was shown that most of the hot spots are located in the middle of the interface and are surrounded by a ring of energetically-unimportant residues that, like an O-ring in a pipe fitting, seal the hot spot from the bulk solvent. Furthermore, it has been shown that the hot spots are enriched in tryptophan, tyrosine, and arginine, the residues capable of forming multiple types of favourable interactions[204]. Interestingly, arginine is a typical feature of the interfaces in the transient protein complexes, for a number of which it is also the most buried residue in the interface[9]. Surrounded by a shell of atoms that form van der Waals contacts with the partner protein, R13 of Cyt *c* sits in the middle of the Cyt *c* – CcP interface (Figure 5.1) as

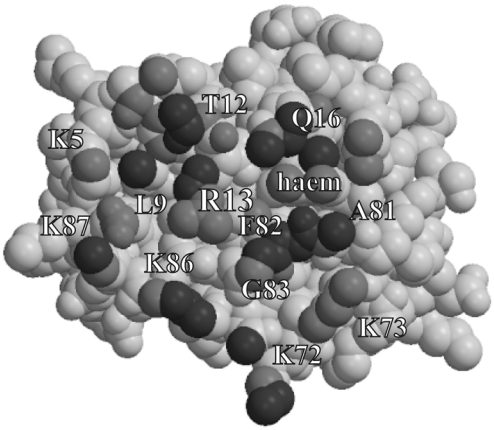

Figure 5.1. **The crystallographic binding site of Cyt *c* in the complex with CcP.** The Cyt *c* atoms that are buried in the interface upon complex formation are coloured light grey and those that make van der Waals contacts with CcP are coloured dark grey. The labels indicate Cyt *c* residues involved in binding to CcP. The molecular coordinates have been taken from PDB entry 2PCC[107]. The figure is based on Fig. 2 in reference 9.

observed in the crystal[107] and solution (Chapter III) structure of the complex and in the NMR chemical shift mapping study[129]. The central position of R13 and its possible occlusion from the solvent make it a suitable candidate for the binding hot spot in the Cyt *c* – CcP complex.

Previous mutational analyses of the yeast Cyt *c* – CcP[128] and horse Cyt *c* – CcP[205,206] complexes were mainly concerned with the charged residues – lysines on Cyt *c* and glutamates and aspartates on CcP – surrounding the crystallographic binding site[107] or the putative low-affinity binding region initially identified by Brownian dynamics study[26] (for a review see ref. 87). In this study we have replaced R13 of Cyt *c* by alanine (R13A), abolishing one positive charge on the surface of Cyt *c*, or lysine (R13K), introducing a charge-conserved mutation. The binding of the two Cyt *c* variants to CcP has been probed by ITC, NMR chemical shift mapping and, in the case of R13A, PRE NMR spectroscopy and compared to that of the wt protein.

Isothermal titration calorimetry (ITC) is the only technique that allows to determine the entire set of fundamental thermodynamic parameters – changes of the free Gibbs energy (ΔG), enthalpy (ΔH), and entropy (ΔS)[i] – for a chemical reaction in a single experiment (for a recent review see ref. 207). To study the binding, one protein is titrated into the other in small increments, producing a heat change associated with the binding[ii], which is registered, integrated, plotted against the protein molar ratio, and fitted to a binding model. From the fit of the experimental data, one obtains a stoichiometry (N), an

[i] And, with the experiments performed at different temperatures, constant-pressure heat capacity (ΔC_p).
[ii] Plus a number of non-specific processes, such as dilution, that are taken into account by performing a blank titration[207].

equilibrium association constant (K_a), and ΔH of the binding, which are used to calculate ΔG and ΔS from the equations $\Delta G = -RT\ln K_a$ (were R is the universal gas constant, 1.98 $cal \cdot K^{-1}mol^{-1}$) and $\Delta G = \Delta H - T\Delta S$. In this way all three thermodynamic parameters – two (ΔG and ΔH) obtained directly and another one (ΔS) derived from the two – are extracted from a single titration.

ITC has been successfully used before to study the interaction of both yeast[127,128,208] and horse[202,205,206] Cyt *c* with CcP. In this work we extend the previous ITC analysis and complement it by the data from solution NMR experiments. For the wt complex, our results agree with the findings of earlier ITC[127,128] and NMR[129] studies. Investigation of R13A and R13K Cyt *c* binding to CcP shows that substitution of R13 residue leads to only a moderate decrease in binding affinity (4-fold decrease in K_a). At the same time, it brings about a drastic decrease in the specificity of interaction, rendering the Cyt *c* – CcP complex more dynamic. We estimate that within the R13A Cyt *c* – CcP complex the proteins spend as much as 75 % of the time in the dynamic encounter state, with only quarter of the time spent in the well-defined, single-orientation complex. It transpires that, rather than being a binding hot spot, R13 is a specificity hot spot in the interaction of yeast Cyt *c* with CcP.

Results

Cyt c – CcP binding investigated by ITC

The ITC titration curve contains the information on both the heat of binding and the binding constant, which determine the amplitude and the shape of the curve, respectively. For accurate determination of both parameters, the product of protein concentration in the sample cell and the equilibrium association constant (K_a) should optimally be between 10 and 100 [209]. This was achieved by using CcP concentrations in the range of 30 – 200 μM (Materials and Methods).

Representative ITC curves for the binding of wt, R13K, and R13A Cyt *c* to CcP are shown in Figure 5.2, and the thermodynamic parameters for the three complexes are given in Table 5.1. All data have been corrected for the heat of dilution of Cyt *c* that was

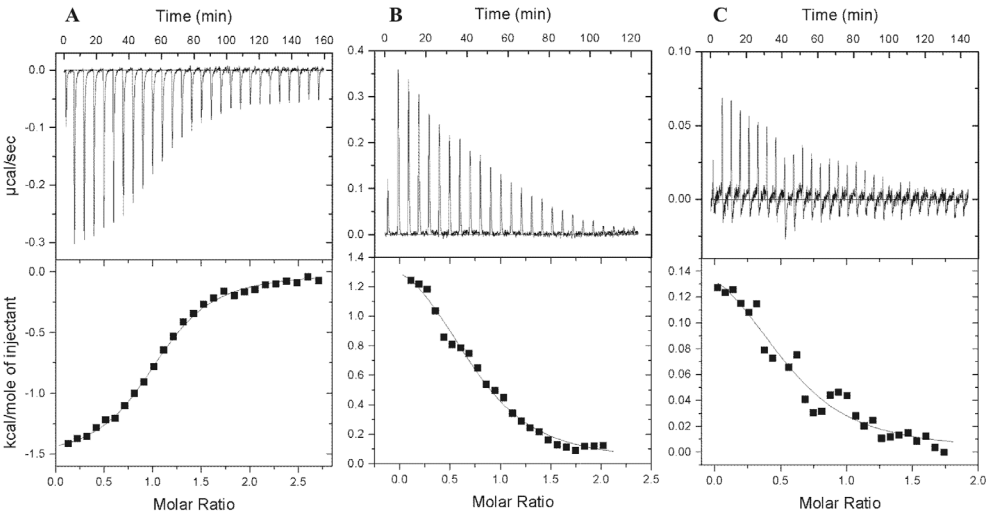

Figure 5.2. **Representative ITC binding curves.** Binding of wt (**A**), R13K (**B**), and R13A (**C**) Cyt *c* to CcP. The top and bottom panels show, respectively, the raw data after the baseline correction and the integrated data corrected for the heat of dilution of Cyt *c*. The solid line in the bottom panel is the best fit of the data to the 1:1 binding model. The thermodynamic binding parameters and experimental conditions are given in Table 5.1.

estimated from the titration of the corresponding Cyt *c* variant into the buffer solution. At 30^{o} C the binding of wt Cyt *c* is exothermic (Figure 5.2 A), while that of R13K is endothermic (Figure 5.2 B). The binding of R13A Cyt *c* is mildly endothermic (Figure 5.2 C); however, at this temperature the heat of binding is close to zero, making the analysis of the binding isotherm difficult. Therefore, the thermodynamic parameters for R13A Cyt *c* – CcP should be interpreted with care. For all Cyt *c* variants, the data – analysed with 1:1 and 2:1 binding models – fit best to the 1:1 model. The stoichiometry of binding of R13K and R13A Cyt *c* deviates from unity (Table 5.1). In the latter case, the deviation is most likely due to the poor quality of the binding isotherm, while in the former instance it may be due to an aspecific protein aggregation or precipitation in the sample cell.

Being a global thermodynamic property of the system, ΔH_B obtained from ITC does not only reflect the binding itself, but also contains contributions from all other processes taking place in solution, with one of the most significant contributions arising from buffer ionisation due to uptake or release of protons upon complex formation[207]. To correct for this, we have used Equation 5.1:

$$\Delta H_B = \Delta H_B' + n \times \Delta H_{ion} \qquad (5.1)$$

Table 5.1. **Thermodynamic parameters of Cyt *c* – CcP binding obtained from ITC experiments[a]**

Cyt *c* variant	N^b	$K_a{}^c$	ΔG_B	ΔH_B	$-T\Delta S_B$
wt (3)	1.06 ± 0.01	18 ± 2	-7.3 ± 0.1	-1.7 ± 0.1	-5.6 ± 0.2
R13K (2)	0.77 ± 0.02	5 ± 1	-6.5 ± 0.2	1.7 ± 0.1	-8.2 ± 0.1
R13A (1)	0.58 ± 0.05	3 ± 1	-6.2 ± 0.2	0.2 ± 0.02	-6.4 ± 0.1

[a] Experiments were performed at 30° C in 20 mM NaP_i 0.1 M NaCl pH 6.0. The units of ΔG_B, ΔH_B, and $T\Delta S_B$ are kcal·mol^{-1}. Number of experiments for each Cyt *c* variant is given in parentheses. For wt and R13K Cyt *c*, the errors are the standard deviations from the mean, while for R13A Cyt *c* the uncertainties are the errors of the fit of the integrated data. [b] Stoichiometry of binding. [c] Equilibrium association constant, 10^4 M^{-1}.

where ΔH_B' is the enthalpy change without the buffer ionisation contribution, n is the number of protons taken up (n > 0) or released (n < 0) upon complex formation, and ΔH_{ion} is the enthalpy of buffer ionisation[127]. Using the values of n = -0.35, determined potentiometrically for the yeast Cyt *c* – CcP complex at 25° C in KNO_3 pH 6.0 and 0.15 mM ionic strength[111], and ΔH_{ion} = 1.2 kcal·mol^{-1} for the phosphate buffer[210], we find that ΔH_B' is 0.42 kcal·mol^{-1} more positive than the ΔH_B values shown in Table 5.1. Thus, assuming that the introduced mutations do not influence the proton release upon complex formation, ΔH_B' for the binding of wt, R13K, and R13A Cyt *c* to CcP is -1.3, 2.1, and 0.6 kcal·mol^{-1}, respectively.

Cyt c – CcP binding investigated by NMR

The NMR spectra of the Cyt *c* variants are similar to those of the wt protein (Figure 5.3). The large differences in chemical shifts are found only for the resonances of the residues located around the mutation site, R13, indicating that the mutations do not perturb the overall structure of the protein. The relative chemical shift changes of R13K Cyt *c* (Figure 5.3 B) are smaller than those of R13A variant (Figure 5.3 A). In other words, R13K Cyt *c* resembles the wt protein more than does R13A variant, which is expected given the similarity of lysine and arginine side-chains.

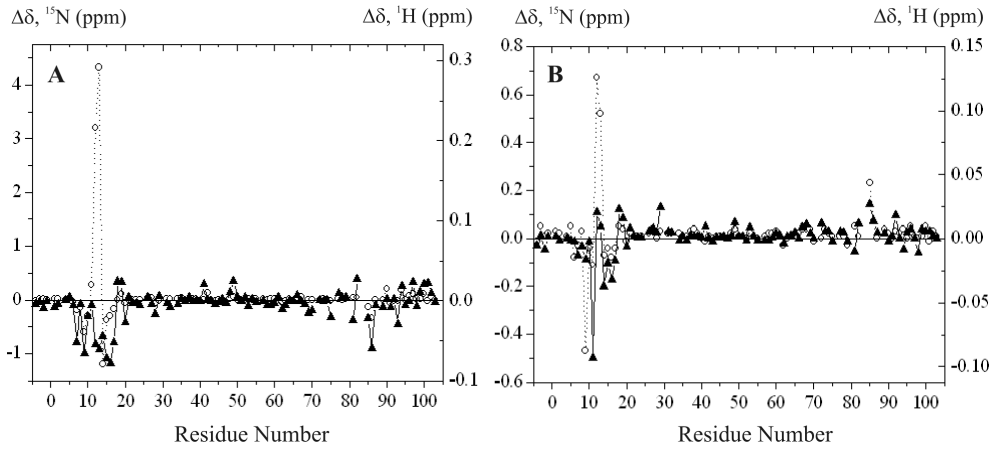

Figure 5.3. **Effects of mutations on the chemical shifts of Cyt *c* backbone amide atoms.** Chemical shift changes of the backbone amide nitrogen (open circles) and proton (filled triangles) nuclei for R13A (**A**) and R13K (**B**) ^{15}N Cyt *c*. Δδ is the difference between the chemical shifts of a given nucleus in the 2D [^{15}N, ^{1}H] HSQC spectra of a variant and the wt Cyt *c*.

Binding of Cyt *c* to CcP leads to the uniform line-broadening of all Cyt *c* resonances and chemical shift perturbations of certain Cyt *c* peaks, which can be monitored in the 1D and 2D NMR spectra (Figure 5.4). In each case a single set of resonances is observed, indicating that the system is in fast-exchange regime, which is consistent with the findings of an earlier study that reported a lower limit of the dissociation rate constant (k_{off}) of 1200 s^{-1} for the wt Cyt *c* – CcP complex[129]. In order to compare the size of the chemical shift perturbations in various Cyt *c* – CcP complexes, the binding shifts were extrapolated to the 100 % bound form of Cyt *c*. To do so, the chemical shift perturbations of wt, R13K, and R13A Cyt *c* were divided by 0.9, 0.82, and 0.77, respectively. Given the binding constants determined from ITC (Table 5.1), these values correspond to the fraction of Cyt *c* bound to CcP at the protein concentrations used in the NMR experiments (Materials and Methods). The size of the binding shifts in the Cyt *c* – CcP complexes diminishes in the order: wt >> R13K > R13A Cyt *c* (Figures 5.4 and 5.5). With the exception of the mutated R13, the residues perturbed by the interaction with CcP are the same for all Cyt *c* variants (Figure 5.5).

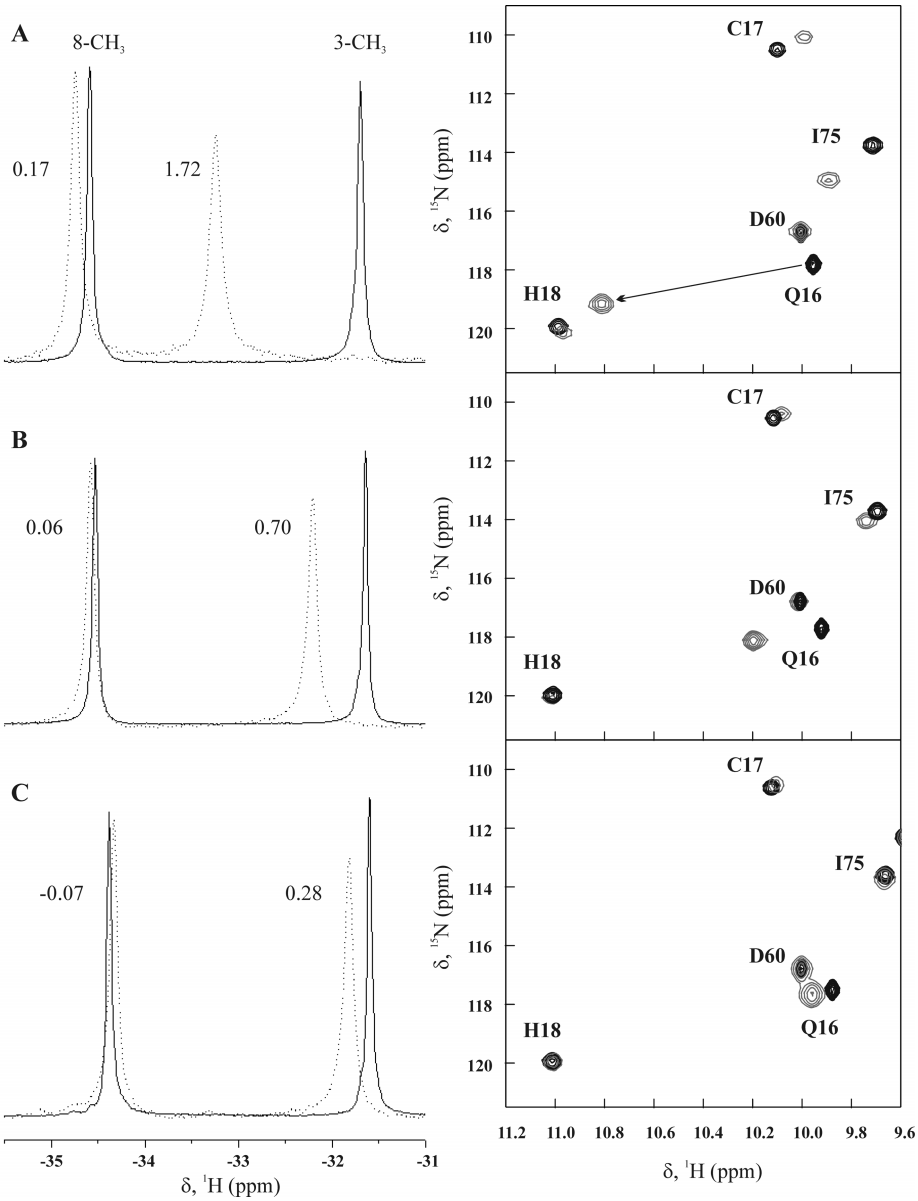

Figure 5.4. **Chemical shift perturbations of Cyt *c* resonances upon binding to CcP.** The binding shifts of wt (**A**), R13K (**B**), and R13A (**C**) Cyt *c* in the presence of 1 molar equivalent of CcP. (*left*) Downfield hyperfine-shifted resonances in the 1D ^{1}H NMR spectra of free (solid line) and bound (broken line) Cyt *c*. The assignments of the Cyt *c* haem peaks are indicated. The chemical shift perturbations, in ppm, extrapolated to the 100 % bound form (see text and Materials and Methods) are given next to each peak. (*right*) Part of the overlaid 2D [^{15}N, ^{1}H] HSQC spectra of the free (black) and bound (grey) ^{15}N Cyt *c*. The labels indicate the assignments of the Cyt *c* backbone amide resonances. Experiments were performed at 30^{o} C in 20 mM NaP_i 0.1 M NaCl pH 6.0.

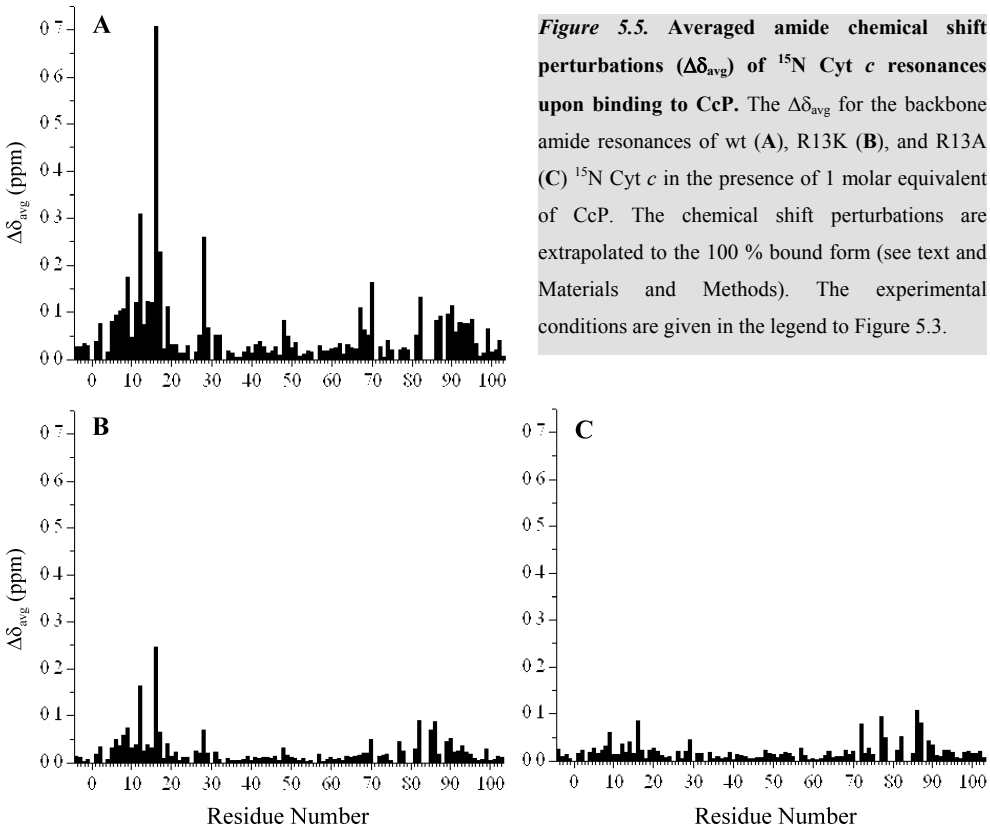

Figure 5.5. **Averaged amide chemical shift perturbations ($\Delta\delta_{avg}$) of ^{15}N Cyt *c* resonances upon binding to CcP.** The $\Delta\delta_{avg}$ for the backbone amide resonances of wt (**A**), R13K (**B**), and R13A (**C**) ^{15}N Cyt *c* in the presence of 1 molar equivalent of CcP. The chemical shift perturbations are extrapolated to the 100 % bound form (see text and Materials and Methods). The experimental conditions are given in the legend to Figure 5.3.

Solution structure and dynamics of the R13A Cyt c – CcP complex

We have used site-specific spin-labelling in combination with PRE NMR spectroscopy to solve the structure of R13A Cyt *c* – CcP complex in solution. Following the procedure described in detail in Chapter III, we have converted the observed intermolecular PREs – caused by an MTSL spin-label attached to the surface of CcP and detected on the Cyt *c* backbone amides – into distance restraints, which were used in the structure calculations. As explained in Chapter IV, the actual PREs (R_2^{para}) for Cyt *c* in the well-defined, single-orientation complex with CcP are obtained by dividing the observed PREs ($R_{2,obs}^{para}$) by f_{dom}, a fraction of the lifetime of the complex spent by the proteins in the single-orientation form. By varying the f_{dom}, multiple sets of distance restraints are generated and used in repeated structure calculations, which are verified for the meaningful

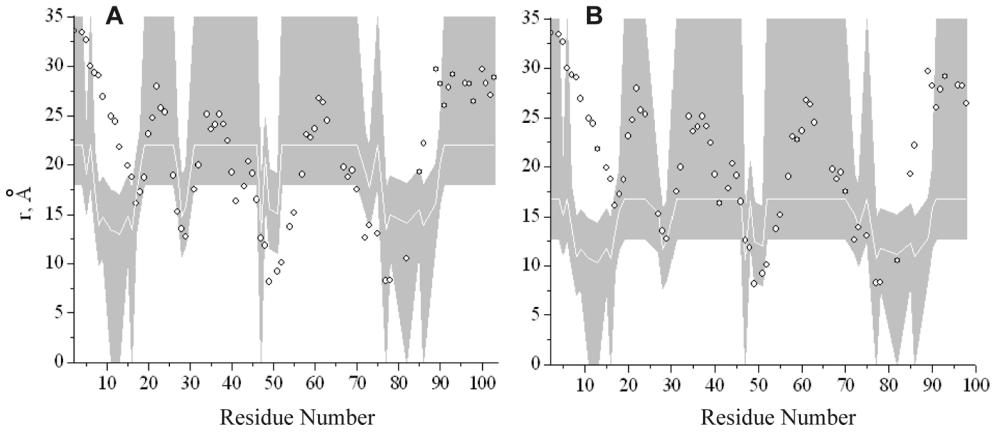

Figure 5.6. **Restraint violations of the lowest-energy R13A Cyt *c* – CcP structure.** The white lines and the shaded areas indicate the PRE-derived distances and error margins, respectively, generated at f_{dom} = 1.00 (**A**) and 0.25 (**B**). The distances from the Cyt *c* backbone amide protons in the lowest-energy structure to the averaged position of the oxygen atom of MTSL attached to N200C CcP are indicated by open circles.

solutions in terms of the restraint and van der Waals violations. In this way, an estimate for a range of f_{dom} occupied in the protein complex can be made (Chapter IV).

For R13A Cyt *c* – CcP complex, no meaningful solutions are produced in the structure calculations using the distance restraints based on $f_{dom} > 0.25$; the structure of the complex could only be computed for $f_{dom} \sim 0.25$. To clarify this point, we compare two sets of restraints generated with f_{dom} = 1 (Figure 5.6 A) and f_{dom} = 0.25 (Figure 5.6 B), corresponding to 100 % and 25 % of the time spent in a single-orientation complex, respectively. In both panels of Figure 5.6, the experimental distance restraints (white line), their bounds (shaded area), and the expected intermolecular distances in the lowest-energy structure calculated with the restraints based on f_{dom} = 0.25 (open circles) are shown. For a given residue, the restraint is satisfied if the corresponding circle is inside the shaded area and violated if the circle is outside the shadow (Figure 5.6).

For the f_{dom} = 0.25 restraints, there are only positive violations (Figure 5.6 B). In other words, the actual distances for some residues in the single-orientation complex are longer than those determined experimentally (*e.g.* residues 5 – 16 in Figure 5.6 B). As discussed in Chapter IV, such violations are indicative of additional protein-protein orientations sampled in the dynamic encounter state of the complex, in which the violated residues come close to the spin-label. The measured PRE is the sum of contributions from all protein-protein orientations – including the well-defined, single-orientation form –

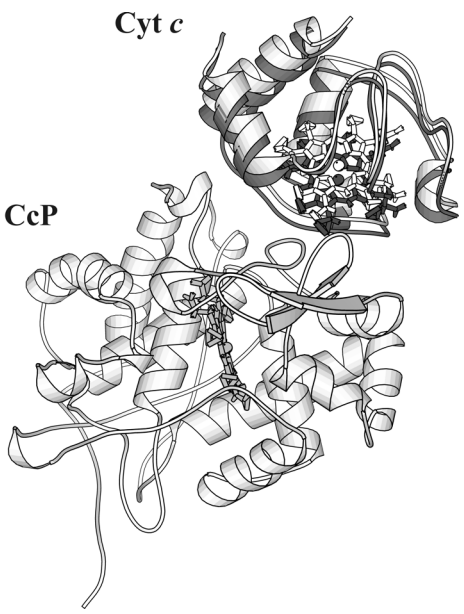

Figure 5.7. **Comparison of the solution structures of the wt Cyt *c* – CcP and R13A Cyt *c* – CcP complexes.** The CcP molecules are superimposed, and the Cyt *c* backbones in the variant and wt complexes are shown in dark and light grey, respectively. Haem groups for both proteins are in sticks. The positional backbone rmsd between the two Cyt *c* molecules is 1.54 Å. The coordinates for the solution structure of the wt complex were taken from the PDB entry 2PB8 (Chapter III).

corrected by the fraction of time spent in each of them (Equation 4.1, p. 80). Thus, the combination of PREs from all forms of the protein complex would satisfy the violated restraints. In other words, the restraints with positive violations are satisfied by additional protein-protein orientations sampled in the dynamic encounter state.

If we assume that the same structure of the single-orientation complex is valid for the restraints based on f_{dom} = 1, then some negative violations are observed (Figure 5.6 A), implying that the actual distances for some residues in the single-orientation complex are shorter than those determined experimentally (*e.g.* residues 49 – 55 in Figure 5.6 A). Unlike the positive violations, the negative violations cannot be resolved by additional protein-protein orientations. Due to the reciprocal of the sixth power distance-dependence of PREs, the shorter distances convert into very large PREs (Equation 3.3, p. 67); therefore, the contributions to the combined PRE from the structures with shorter distances are disproportionately higher than those from the orientations with the longer distances. As a result, the distances calculated from the combined PREs will still be short of the experimental values and, therefore, will not resolve the violations. Thus, if the structure of the complex computed for f_{dom} = 0.25 were a valid solution for the restraints based on f_{dom} = 100 %, the Cyt *c* residues with negative violations would experience large paramagnetic effects in the complex with N200C CcP-MTSL. However, the experiments show that those are only moderately affected (*e.g.* residues 49-51 in Figure 5.6 A) or not affected at all (*e.g.* residues 52 – 55 in Figure 5.6 A), implying the inadequacy of this structure for the case of f_{dom} = 1.

As mentioned above, the structure of the dominant form of the R13A Cyt *c* – CcP complex in solution was calculated using the experimental distance restraints based on the value of f_{dom} = 0.25 (Table C1 in Appendix C) following the protocol described in Chapter III. Multiple rigid-body structure calculations have consistently produced a single cluster of solutions (data not shown), analogously to the wt complex (Figure 3.9, p. 57). Overlay of the solution structures of the R13A Cyt *c* – CcP and wt Cyt *c* – CcP complexes shows that the two structures are very similar (Figure 5.7). The Cyt *c* in the variant complex is only slightly translated relative to that in the wt complex (Figure 5.7), and the positional backbone rmsd between the Cyt *c* molecules in the two structures is 1.5 Å.

Many distance restraints are violated in the lowest-energy structure of the R13A Cyt *c* – CcP complex (Figure 5.8 and Tables C1 and C2 in the Appendix C). As the violations are larger than the uncertainties associated with experimental and computational errors, it appears that the solution structure alone is not enough to explain the experimental paramagnetic effects, implying the presence of additional protein-protein orientations within this complex (Chapter IV). Distribution of the R13A Cyt *c* residues with violated restraints is similar to that observed in the wt Cyt *c* – CcP complex (Figure 5.8, *cf.* Figure 4.1, p. 74), suggesting that similar configurational space must be sampled in the dynamic encounter states of both complexes.

Discussion

Cyt c – CcP binding investigated by ITC

The complex formation between Cyt *c* and CcP is accompanied by the heat uptake or release, depending on the identity of the Cyt *c* residue at position 13. For wt and R13K Cyt *c* the binding heat is large (Figure 5.1 A and B) – allowing for an accurate analysis of the titrations curves – while that for R13A is close to zero (Figure 5.1 C), making the analysis difficult. One way to remedy this problem is to vary the experimental parameters, *e.g.* the temperature, in order to find the conditions under which $|\Delta H_B|$ is larger. Until such data are available, the thermodynamic parameters for R13A Cyt *c* – CcP complex determined in this work should be interpreted with care.

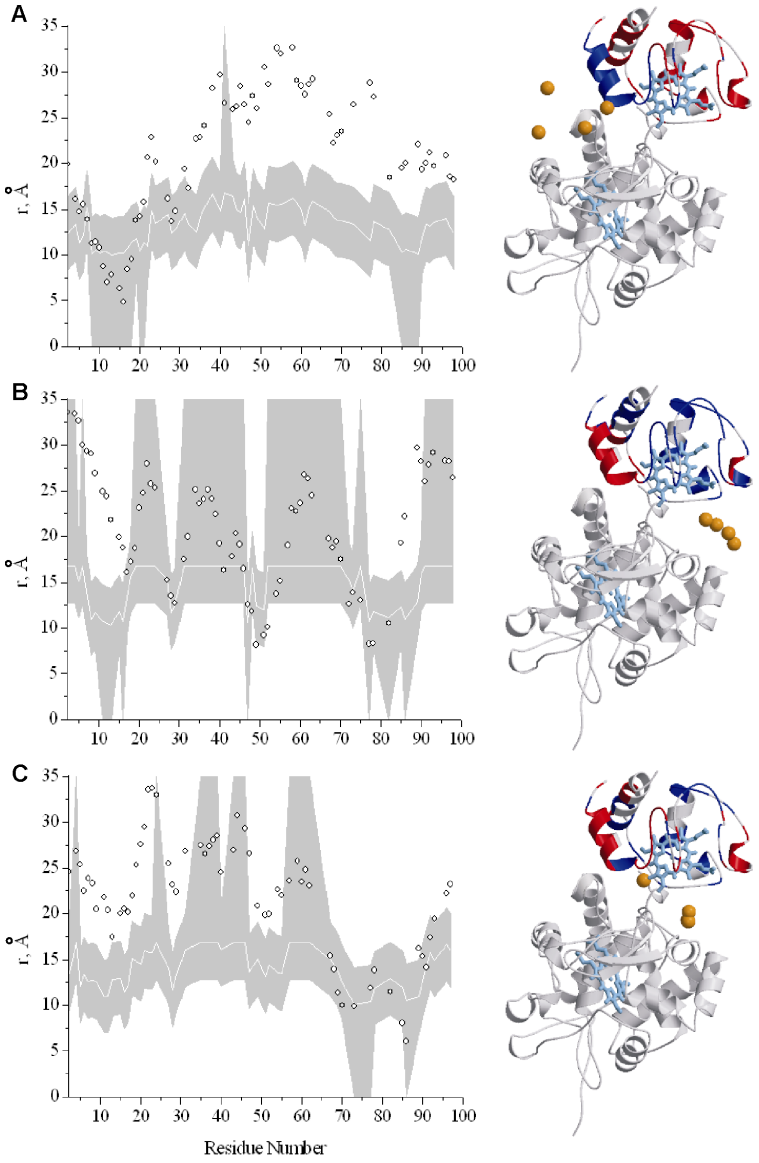

Figure 5.8. **Violation analysis of the best solution structure for the R13A Cyt *c* – CcP complex.** The graphs illustrate the distances from the Cyt *c* backbone amide protons in the lowest-energy structure (open circles) to the averaged position of the oxygen atom of MTSL attached to CcP at N38C (**A**), N200C (**B**), and T288C (**C**). The white line and the shaded area indicate the PRE-derived distances and error margins, respectively, used in the structure calculations. On the right are cartoon representations of the lowest-energy structure, indicating the residues with satisfied (blue) and violated (red) restraints. Haem groups for both proteins are in cyan. For each of the MTSL positions, four oxygen atoms that were used for ensemble averaging in the structure calculations are shown as orange spheres. For details see Materials and Methods of Chapter III.

For all protein complexes investigated in this work, the titration curves fit best to a 1:1 binding model, suggesting that no detectable binding of the second Cyt *c* takes place in solution. Under the conditions used, the dissociation constant for the second binding event is expected to be more than 1 mM[111]. If such weak binding does occur, it will not be observed in the ITC experiment[127].

The thermodynamic parameters for the interaction of wt Cyt *c* with CcP (Table 5.1) are similar to those determined in previous ITC studies[128,208]. In particular, the values of ΔH_B = -2.3 ± 0.4, $-T\Delta S_B$ = -5.3 ± 0.4, and ΔG_B = -7.7 ± 0.1 reported for this complex at 25^o C, pH 6.0 and I = 82 mM[208] are close to those obtained in this work (Table 5.1). The slight discrepancies in the values can be attributed to the differences in the experimental conditions. The complex formation between wt Cyt *c* and CcP is either ΔH- or ΔS-driven depending on the temperature[127]. Under the conditions used in this study, both ΔH and ΔS contribute favourably to the binding energy (ΔG_B), but the entropic term is larger (Table 5.1 and Figure 5.9), implying that the binding of wt Cyt *c* to CcP is mostly ΔS-driven.

The interaction of R13A and R13K Cyt *c* with CcP is clearly different from that of the wt protein: while the binding of the wt Cyt *c* is exothermic ($\Delta H_B < 0$), that of the variant proteins is endothermic ($\Delta H_B > 0$) and, therefore, is entirely ΔS-driven (Figure 5.9). Interestingly, the thermodynamic binding parameters for R13A and especially R13K yeast Cyt *c* look very similar to those reported for horse Cyt *c*[202]. It is worth noting that among the differences between the primary sequences of the yeast and horse Cyt *c* is the residue at position 13, which is arginine in the former and lysine in the latter. Our results suggest that, as far as the thermodynamic properties of the Cyt *c* – CcP complex formation are concerned, substitution of a single residue (R13K) can convert yeast Cyt *c* into horse Cyt *c*.

There has been a debate on theoretical grounds whether the changes in the thermodynamic properties of a system can be attributed specifically to the effect of a single mutation[211-213], with the general consensus that such analysis should be done with care[212,213]. In our case the effect of R13A and R13K mutations on the complex formation is net destabilizing (Figure 5.9). However, the ΔG_B increases by only 11 – 15 %, converting into a 4 – 6-fold reduction in K_a (Table 5.1), which is similar to the decrease in K_a observed for other mutants of Cyt *c* both from yeast[128] and horse[205]. The reason for only a mild effect of mutations on the binding affinity is enthalpy-entropy compensation, which until

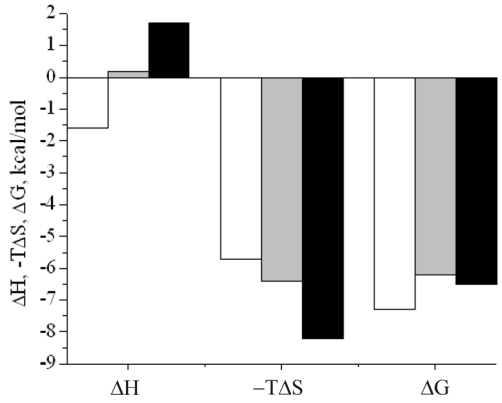

Figure 5.9. **Thermodynamic parameters for the Cyt *c* – CcP complex formation.** ΔH, -TΔS, and ΔG for the binding of wt (white), R13A (grey), and R13K (black) Cyt *c* to CcP. The histograms are based on the Table 5.1.

recently[214] was thought to be a general feature of the biomolecular interactions in solution[215]. The decrease in ΔG_B for the binding of wt Cyt *c* compared to that of R13A or R13K variant is caused by the favourable ΔH_B term, suggesting that some intermolecular interactions that are absent from the variant Cyt *c* – CcP complexes stabilize the wt complex.

The binding of R13A and R13K Cyt *c* to CcP is similar: for both variants it is endothermic and ΔS- driven (Figure 5.9), with the ΔG_B values being identical within the error of the experiment (Table 5.1). At the same time, the complex of R13K Cyt *c* exhibits an increase in ΔH_B and ΔS_B as compared to R13A variant. The observed unfavourable change in ΔH accompanied by a large, favourable change in ΔS is characteristic of the hydrophobic effect[216], one of the major factors promoting protein-protein assosiation[217]. It appears that the hydrophobic effect is enhanced in the binding of R13K Cyt *c* compared to that of R13A variant. As the contribution of a bulky lysine residue to the non-polar surface of the binding interface is larger than that of a short side-chain of an alanine, the entropic gain due to the removal of water from the nonpolar surface[218] is expected to be larger for R13K than for R13A Cyt *c*. As the release of a single water molecule from a protein surface can contribute up to 2 kcal·mol^{-1} to the binding energy[219], a few more water molecules could be displaced upon complex formation between R13K Cyt *c* and CcP as compared to R13A Cyt *c* – CcP.

Cyt c – CcP interaction investigated by NMR

The interaction of ^{15}N Cyt *c* with CcP, monitored in 1D and 2D NMR spectra (Figure 5.4), leads to chemical shift perturbations for some of the Cyt *c* resonances. While the residues perturbed by the binding to CcP are the same for all Cyt *c* variants (Figure 5.5),

the size of the binding shifts is greatly reduced for R13K and R13A Cyt *c* compared to the wt protein (Figures 5.4 and 5.5). Such differences in binding are due to different dynamics within Cyt *c* – CcP complexes, rather than an altered structure of the dominant form of the complex (see below).

The growing body of evidence suggests that small chemical shift perturbations – observed in the transient protein complexes of Cyt b_5 and myoglobin[12], Cyt *c* and adrenodoxin[146], and Cyt *c* and Cyt b_5[201] (Chapter II) – are indicative of dynamic complexes, while large binding shifts – like those in Cyt *f* and plastocyanin[166,220], Cyt *f* and *Anabaena* Cyt c_6[221], and wt Cyt *c* and CcP[129] (and this work) – are associated with predominantly single-orientation complexes[5]. Judging from the size of the chemical shift perturbations (Figure 5.5), the complexes formed by R13K and R13A Cyt *c* are more dynamic than that of the wt protein.

The increase in dynamics is further confirmed by PRE NMR analysis of the R13A Cyt *c* – CcP interaction. At the resolution of protein backbones, the R13A mutation does not alter the structure of the dominant form of the complex (Figure 5.7). At the same time, *ca.* 49 % of restraints are violated in the lowest-energy R13A Cyt *c* – CcP structure as compared to *ca.* 21 % violations in the wt complex (Table C2 in Appendix C). Also the size of the violations is larger in the case of R13A Cyt *c* (Figure 5.8, *cf.* Figure 4.1, p. 74). As discussed in Chapter IV, the extent of the violations is related to the protein dynamics within the complex, suggesting that the variant complex is more dynamic than that formed by the wt proteins. Indeed, an estimate for the fraction of time spent by R13A Cyt *c* and CcP in the well-defined, single-orientation form is ~ 25 %, which is greatly reduced compared to the value of > 70 % obtained for the wt complex (Chapter IV).

The interaction between R13K Cyt *c* and CcP has not been analysed by PRE NMR spectroscopy, so that no direct estimate for the fraction of time spent in the dynamic encounter state is possible. However, the overall similarity between R13A and R13K Cyt *c* variants (*e.g.* thermodynamics of binding to CcP, the size of chemical shift perturbations, the residues involved in binding) suggests that protein mobility is increased also in the R13K Cyt *c* – CcP complex. Our results indicate that, rather than changing the structure of the complex, the substitution of R13 shifts the equilibrium of the Cyt *c* – CcP complex formation towards the dynamic encounter state (Figure 1.1, p. 11).

The role of R13 residue of Cyt c in the Cyt c – CcP complex

The residue R13 of Cyt *c* clearly stabilizes the wt Cyt *c* – CcP complex and increases the specificity of the interaction between these proteins. R13K mutation of Cyt *c* provides an insight into how this dual function might be maintained. A large effect of the charge-conserved R13K mutation on the protein binding is surprising, considering the high similarity of arginine and lysine side-chains. The major difference between the two residues is the charged moiety, which is a primary amine in lysine and a guanidinium group in arginine. The latter renders arginine a potent partner in a cation-π interaction[222], which is increasingly recognized as an important type of non-covalent intermolecular interactions in protein-protein complexes[223]. Unlike arginine, lysine is a poorer performer in cation-π interactions as was suggested by earlier reports[222,223] and shown experimentally in a recent study[224].

In the crystal structure of Cyt *c* – CcP complex, R13 of Cyt *c* is in van der Waals contact with Y39 of CcP (Figure 5.10). While the two residues do not form a cation-π bond in the crystal, a small movement of the arginine side-chain would bring the two residues into the optimal orientation for this non-covalent interaction, which may indeed take place in solution. Such additional stabilization, which could contribute ~ 1 kcal·mol^{-1} to the binding energy[222,224], would explain the difference in binding of R13K and wt Cyt *c*. Moreover, the absence of a cation-π interaction could also account for the increased mobility and decreased specificity in the R13K Cyt *c* – CcP complex. Such conclusion is in agreement with an earlier suggestion that cation-π interactions are more important for specificity rather than affinity of transient protein associations[223].

Possible role of water ordering in the Cyt c – CcP complex

Binding-induced reorganisation of water molecules at the interface of transient protein complexes is expected to cause large NMR chemical shift perturbations[5,10,129]. Large differences in the size of the binding shifts for various Cyt *c* – CcP complexes

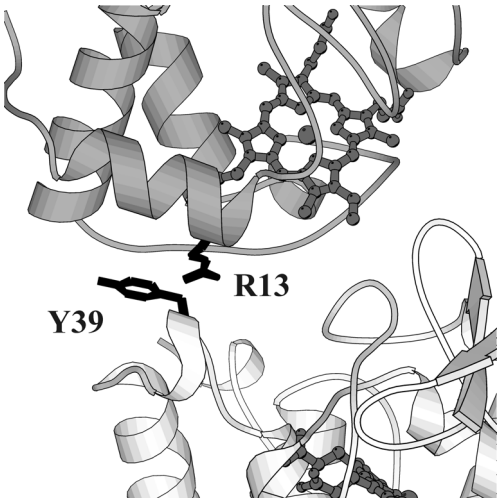

Figure 5.10. **The Cyt *c* – CcP interface as seen in the crystal structure of the complex.** With the haem groups in ball and sticks, Cyt *c* and CcP are coloured dark and light grey, respectively. The residues suggested to participate in the intermolecular cation-π interaction are labelled and shown in sticks. The molecular coordinates have been taken from PDB entry 2PCC[107].

observed in this work (Figure 5.5) could be attributed, in addition to dynamics (see above), to the difference in the hydration state of the binding interface.

A high-resolution crystallographic study of the Cyt *c* – CcP covalent cross-link has revealed several structured interfacial water molecules that mediate the interaction between the two proteins[118]. Interestingly, the water molecules are H-bonded to the Cyt *c* residues T12 and Q16 that experience the largest binding shifts in the wt Cyt *c* – CcP complex (Figure 5.5 A). If water ordering does occur at the interface of the non-covalent complex in solution, this would explain the decrease in ΔS_B for the wt complex relative to those formed by R13A or R13K Cyt *c*. In addition, the solvent reorganization, which can contribute significantly to the binding enthalpy[225], could partially account for the favourable ΔH_B observed for the wt complex. It is, therefore, possible that R13 residue of Cyt *c* plays a role in organisation of water structure at the binding interface.

One way to ascertain whether water is released or taken up upon complex formation is offered by analysis of protein binding in the presence of small molecular cosolutes, which are expected to promote the complex formation if water is released or inhibit it if water is taken up[208,210]. Previous study has shown that water uptake and release are nearly balanced for the wt Cyt *c* – CcP complex[208]. However, water release upon complex formation can counterbalance the binding-induced structuring of water molecules at the interface, with the resulting net water uptake or release close to zero as is indeed the case[208]. A comparative osmotic binding analysis of wt, R13A, and R13K Cyt *c* should clarify whether structuring of water takes place in the Cyt *c* – CcP complex.

Concluding remarks

R13 of Cyt *c* is a hot spot for specificity, rather than affinity, in the Cyt *c* – CcP interaction. Substitution of this residue impairs the protein recognition and shifts the equilibrium of the complex formation towards the dynamic encounter state. We propose that R13 maintains the high specificity of the Cyt *c* – CcP interaction by an intermolecular cation-π bond with Y39 of CcP. Further mutational analysis should clarify whether Y39 residue of CcP participates in the proposed interaction. We argue that R13 of Cyt *c* could also play a role in ordering of water at the Cyt *c* – CcP interface and suggest the experimental approach to test this hypothesis.

As suggested by other studies (ref. 5 and references therein) and confirmed by PRE analysis in this work, the size of the binding-induced NMR chemical shift perturbations (Δδ) can be used as a reliable diagnostic tool for the dynamics within protein-protein complexes. Qualitative interpretation of the magnitude of the binding shifts allows to discriminate between highly dynamic (small Δδ) and mostly single-orientation (large Δδ) protein complexes.

The large effect of the charge-conserved R13K mutation on the binding of Cyt *c* to CcP suggests that intermolecular electrostatic and hydrophobic forces – important determinants of the affinity in the transient protein complexes[4] – are not sufficient to maintain the high specificity of the Cyt *c* – CcP interaction and additional intermolecular contacts, such as the proposed cation-π bond, must facilitate molecular recognition in this protein complex.

Materials and Methods

Protein preparation

Cyt *c* R13A and R13K mutations have been introduced by site-directed mutagenesis using the Quik Change™ polymerase chain reaction protocol (Stratagene, La

Jolla, CA) with the plasmid pUCcc as a template (for details see Appendix D). All constructs have been verified by DNA sequencing. Both unlabelled and ^{15}N isotopically-enriched R13A, and R13K Cyt *c* have been expressed and purified analogously to the wt protein (Chapter II and Appendix B). Ferric Cyt *c* with a UV-vis peak ratio $A_{410}/A_{280} \geq 4.0$ was used throughout. Wild-type CcP has been expressed and purified following a published procedure[73] (for a detailed protocol see Appendix B). Single-cysteine CcP variants, used in the PRE analysis of the R13A Cyt *c* – CcP complex, have been prepared and spin-labelled following the procedure described in detail in Chapter III. 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 (ε = 106.1 $mM^{-1}cm^{-1}$)[152] and 408 nm (ε = 98 $mM^{-1}cm^{-1}$)[191], respectively.

ITC sample preparation and experiments

To prepare the protein samples, Cyt *c* and CcP were dialysed against 20 mM NaP_i 0.1 M NaCl pH 6.0 in the same beaker for ca. 48 hours with three buffer changes (the ratio of buffer to sample volume > 100), using Spectra/Por (Spectrum, Rancho Dominguez, CA) dialysis membranes with 3.5 and 8 kDa molecular weight cut-off for Cyt *c* and CcP, respectively. After dialysis the protein concentrations were determined by UV-vis spectrophotometry as described above.

All experiments were performed on a Microcal VP-ITC calorimeter at 30^o C in 20 mM NaP_i 0.1 M NaCl pH 6.0. Cyt *c* solution of 0.8 – 1.8 mM was placed into the syringe, which was rotated at 310 rpm during each experiment, and titrated into the sample cell containing 60 – 200 μM CcP. All titrations were performed with the injection volume of 4 μl for the first titration point and 10 μl for all subsequent injections, with the initial equilibration delay of 60 s and the delay between the injections of 360 s.

The binding isotherms were analysed with Microcal™ Origin® 5.0. In each case the first data point was discarded and the baseline adjusted manually. The integrated data were corrected for the heat of Cyt *c* dilution, determined by titrating Cyt *c* into the buffer, and fitted to 1 : 1 and 2 : 1 binding models included into the software package of the manufacturer[209].

NMR sample preparation and experiments

The protein concentrations in the NMR experiments were 0.8 – 1 mM for free ^{15}N Cyt *c*, 0.5 mM for 1 : 1 complexes of ^{15}N Cyt *c* – CcP, and 0.3 – 0.4 mM for 1 : 1 complexes of ^{15}N R13A Cyt *c* – CcP-MTSL. All NMR samples contained 20 mM NaP_i 0.1 M NaCl 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 pH of the samples was adjusted to 6.00 ± 0.05 with small aliquots of 0.1 M HCl or 0.1 M NaOH.

All NMR experiments were performed at 303 K on a Bruker DMX600 spectrometer equipped with a triple-resonance TCI-Z-GRAD CryoProbe (Bruker, Karlsruhe, Germany). The 1D ^{1}H spectra were acquired with a spectral width of 70 ppm and 4096 complex points. 2D [^{15}N, ^{1}H] HSQC spectra were obtained with 1024 and 256 complex points in the direct and indirect dimensions, respectively, and spectral widths of 42 ppm (^{15}N) and 16 ppm (^{1}H). All spectra were processed with AZARA 2.7[153], and chemical shift perturbations of ^{15}N and ^{1}H nuclei were analysed by overlaying the spectra of free and bound ^{15}N Cyt *c* in ANSIG for Windows[154,155]. The averaged amide chemical shift perturbations ($\Delta\delta_{avg}$) were derived from Equation 2.1 (p. 44).

Assignments of the ^{15}N and ^{1}H nuclei of the free and bound wt Cyt *c* were taken from the literature[129,226]. In this study Cyt *c* amides A3, H33, M80, G83, G84, and E88 were not observed. The HSQC spectra of R13A and R13K ^{15}N Cyt *c* – identical to that of ^{15}N wt Cyt *c*, except for the resonances of the several residues around the mutation site – have been unambiguously assigned using 3D TOCSY-HSQC and 3D NOESY-HSQC spectra.

PRE NMR analysis and structure calculations

PREs of the R13A Cyt *c* amide protons in the complexes with N38C, N200C, and T288C CcP-MTSL have been collected, converted into the intermolecular distance restraints (see Table C1 in Appendix C), and used in structure calculations following the protocol described in detail in Chapter III. Violation analysis of the lowest-energy R13A Cyt *c* – CcP structure has been performed following the procedure used in Chapter IV.