Dynamics in electron transfer protein complexes

Bashir, Q.

Chapter 6

Engineering specificity in the non-physiological complex of horse cytochrome *c* and yeast cytochrome *c* peroxidase by a single conserved mutation

Abstract

It has been demonstrated that the complex of yeast cytochrome c (Cc) and cytochrome c peroxidase exists as a delicate equilibrium between a specific state and dynamic encounter state. Replacement of the interface residue Cc Arg 13 to Lys or Ala decreases the specificity and shifts the complex towards the encounter state. Horse Cc has a Lys residue at the equivalent position (Lys 13) and its complex with yeast CcP is less specific than the physiological one. The aim of this study was to determine whether the horse Cc-CcP complex could be made more specific. It is demonstrated that the conservative mutation K13R enhances the specificity of the complex of horse Cc with CcP and yields a specific complex that resembles the physiological yeast Cc-CcP complex. In contrast, the K13A mutation reduces the specificity of the complex. These results show that changes in the specificity of weak protein complexes can be achieved by rational engineering of interface residues.

Introduction

Many biological processes such as electron transfer and signal transduction occur through transient protein-protein interactions. The proteins involved recognize multiple partners and have to compromise between forming a complex with high specificity and high turn-over rate. Specificity is achieved via multiple non-covalent interactions and increases with the affinity between the partners, expressed as the binding constant K_A, the ratio of the association and dissociation rate constants, k_{on}/k_{off}. The turn-over rate is directly proportional to k_{off}. Consequently, to achieve as much specificity as possible k_{on} is high in such complexes. Association is enhanced by the formation of an encounter complex prior to the formation of specific complex[17]. In recent studies, the encounter complex has been visualized for a number of protein-protein[19,20,23,184] and protein-DNA[21,22] interactions. Recently, we have characterized the encounter complex of yeast cytochrome *c* (yCc) and yeast cytochrome *c* peroxidase (yCcP) using paramagnetic relaxation enhancement (PRE) NMR spectroscopy and Monte Carlo simulations[19,184] (chapter 4). Combining the PRE NMR data with the theoretical simulation provided a more detailed description of the encounter state and enabled us to calculate the relative populations of the encounter and specific complexes.

Pelletier and Kraut have determined the crystal structures of yCc and horse cytochrome *c* (hCc) in complexes with yCcP. A distinction between these structures is a rotation and translation of the Cc molecules relative to CcP. A single intermolecular hydrogen bond between Glu290 of CcP and Asn70 of Cc was identified in the yCc-yCcP complex. The structure of the hCc-yCcP complex shows three intermolecular hydrogen bonds, between (i) Glu290 of CcP and Lys72 of Cc, (ii) Asn38 of CcP and Lys8 of Cc, and (iii) Glu35 of CcP and Lys87 of Cc. With a single weak intermolecular hydrogen bond, the yCc-yCcP complex appears to be dominated by the hydrophobic and van der Waals interactions, whereas the hCc-yCcP complex shows stronger charge-charge interactions between CcP and Cc. Another remarkable difference is found for Cc residue 13, which is buried in the complex interface. In hCc, Lys13 forms an intramolecular hydrogen bond with Glu90. In yCc, the residues number 13 and 90 are an arginine and aspartate, respectively, and there is no intramolecular hydrogen bond between them.

Previously we characterized the dynamic complex of yCc and yCcP[184] (chapter 4). In solution the transient yCc-yCcP complex was found to exist as equilibrium of the specific and the encounter complex, with the specific complex as the dominant form. In the wild type complex the relative fraction of the specific complex and the encounter complex was found to be 70% and 30%, respectively. Another study revealed that this relative population could be altered by single residue mutations in the interface (chapter 5). The mutation of yCc Arg13 to either Lys or Ala resulted in a reduced specificity and enhanced population of the encounter complex. The hCc has a Lys residue at the equivalent position. The aim of the present study was to investigate the role of this particular residue in the specificity of the hCc-yCcP complex. It is revealed that the specificity of the hCc-yCcP complex can be enhanced or reduced by mutation of this residue. The mutation of Lys13 to Arg increases the specificity of the complex, resembling that of the physiological yCc-yCcP complex. On the other hand, the K13A mutation decreases the specificity of the complex, making the complex more dynamic. These results demonstrate that the specificity of a transient complex can be modified by interfacial mutations.

Materials and methods

Protein preparation

Cc and CcP mutants were created by site-directed mutagenesis using Quik Change PCR protocol (Stratagene, La Jolla, CA) with wt Cc (H26N, H33N) and wt CcP (C128A) plasmids as template[19,188]. The constructs were verified by DNA sequencing. The proteins were expressed in *Escherichia coli* and purified as described earlier[144,164]. During the gel filteration step in the purification of single cysteine CcP mutants, 1 mM DTT was always present in the buffer and the protein samples. Concentrations of ferric Cc and high spin-ferric CcP were determined according to their optical absorbance peaks at 410 nm[114,165]. The final yield for unlabelled and ^{15}N labelled Cc was 15-20 mg/L and 8-10 mg/L respectively. The final yield for unlabelled CcP mutants was 20-25 mg/L.

Purified single cysteine CcP mutants were labelled with a paramagnetic spin label, MTSL [(1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanethiosulfonate], or a diamagnetic analogue, MTS [(1-acetyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanethiosulfonate], both purchased from Toronto Research Chemicals (Toronto, Canada), using the previously described protocol[19]. The yield of labelling was more than 90% as determined by EPR or dithio-dipyridine assay[166].

X-ray crystallization and structure determination

Crystals of the K13R hCc:yCcP complex were grown at room temperature in sitting drops prepared by mixing 1 µl of reservoir solution and 1 µl of 0.5 mM 1:1 complex. The reservoir solution contained 15% PEG 3350 and 150 mM NaCl. Crystallization drops were equilibrated against 80 µl of reservoir solution. Cc and CcP were in Fe(III) state. Both proteins were thoroughly exchanged in water prior to the crystallization trials. The crystals were flash-frozen and kept at 100 K during data collection. Crystals of the complex diffracted to 2.1 Å and belong to space group $P2_12_12_1$ with unit cell parameters a = 44.9 Å, b = 87.7Å and c = 104.7Å. Reflections were indexed and integrated with iMosflm[189]. Scaling and merging was performed with SCALA[190] from the CCP4 suite[191]. The data statistics are given in Table 1. Structure solution was performed by molecular replacement with the structure of yCc-yCcP (Protein Data Bank ID codes 2PCC)[81] as search model using MOLREP[192] from the CCP4 suite[191]. Manual adjustments to the model were done with COOT[193]. Refinement was done using REFMAC[194] and water molecules were added using ARP/wARP[195] and COOT[193]. The structure solution and refinements were performed by Elisabeth Meulenbroek and Dr. Navraj Pannu. The X-ray structure of the K13R hCc-yCcP complex has been submitted to Protein Data Bank and has received the PDB ID code as 3N6P.

NMR spectroscopy

For PRE experiments, NMR samples contained 0.3 mM of a 1:1 ^{15}N Cc-CcP-MTS(L) complex. For chemical shift perturbation analysis, NMR samples contained 0.2 mM CcP and varying concentration of ^{15}N Cc from 0.04 mM to 1 mM. The NMR experiments were performed in 20 mM sodium phosphate, 100 mM NaCl pH 6.0, 6%

D_2O for lock and 0.1 mM $CH_3CO^{15}NH_2$ as internal reference. ^1H-^{15}N HSQC spectra were recorded at 303 K on a Bruker DMX 600 MHz spectrometer equipped with a TCI-Z-GRAD cryoprobe. The spectra were acquired with 1024 and 128 complex points in direct and indirect dimensions, respectively. The pH of the sample was checked before each titration point and adjusted to 6.00 ± 0.05 if necessary, with small aliquots of 0.1 M NaOH or 0.1 M HCl. All data were processed with AZARA 2.7 (http://www.bio.cam.ac.uk/azara/) and analyzed in Ansig for Windows[115]. Assignments of the free HSQC spectrum of wt Cc were taken from a previous report[196]. The HSQC spectra of K13R and K13A were similar to that of the wild type, and the distinct peaks for mutated residues could be assigned unambiguously. The average chemical shift perturbations were derived from eq 6.1:

$$\Delta\delta_{avg} = \sqrt{\frac{(\Delta\delta_N / 5)^2 + \Delta\delta_H^2}{2}} \qquad (6.1)$$

where $\Delta\delta_N$ and $\Delta\delta_H$ are the chemical shift perturbations for the amide nitrogen and proton, respectively. Chemical shift titration curves were analysed with a two-parameter non-linear least square fit using a one site binding model:

$$\Delta\delta_{binding} = 0.5\Delta\delta_0\left(A - \sqrt{A^2 - 4/R}\right) \qquad (6.2a)$$

$$A = 1 + 1/R + \frac{[Cc]_0 + R[CcP]_0}{R[Cc]_0[CcP]_0 K_B} \qquad (6.2b)$$

where R is the [Cc]/[CcP] ratio, $\Delta\delta_{binding}$ is the chemical shift perturbation at a given R, $\Delta\delta_0$ is the chemical shift perturbation at 100% bound Cc, $[CcP]_0$ is the starting concentration of CcP, $[Cc]_0$ is the stock concentration of Cc and K_B is the binding constant.

Calculation of PRE

From the intensities of peaks in paramagnetic and diamagnetic HSQC spectra, PREs were calculated using eq. 6.3[168]

$$\frac{I_{para}}{I_{dia}} = \frac{R_2 \exp(-t\Gamma_2)}{R_2 + \Gamma_2} \qquad (6.3)$$

where I_{para} and I_{dia} are the intensities of the ^{15}N Cc amide peaks in the presence of CcP-MTSL and CcP-MTS, respectively, R_2 is the transverse relaxation rate of Cc amide protons in the complex with CcP-MTS, Γ_2 is the PRE, and t is the total INEPT evolution time of the HSQC. For the amides of which the resonances disappear in the paramagnetic spectrum, an upper limit for I_{para} was estimated from the standard deviation of the noise level of the spectrum. For each peak, R_2 was estimated from the width at half-height ($\Delta\nu_{1/2}$) of a Lorentzian fit in the proton dimension by using $R_2 = \pi\Delta\nu_{1/2}$. To account for the minor differences in protein concentrations of the two samples, PRE values were calculated after normalization of the I_{para}/I_{dia} values. The residues with $I_{para}/I_{dia} > 0.85$ were used for this correction. The upper and lower 10% of I_{para}/I_{dia} values were removed, and the average of the remaining values was obtained. All I_{para}/I_{dia} values were divided by this average to get the normalized PREs.

The Γ_2 can be converted into distance restraints by the following equation

$$\Gamma_2 = \frac{\gamma^2 g^2 \beta^2}{20 r^6}\left(4\tau_c + \frac{3\tau_c}{1+\omega_h^2\tau_c^2}\right) \qquad (6.4)$$

where τ_c is the correlation time of the electron-nucleus vector (16 ns)[19], ω_h and γ are proton Larmor frequency and gyromagnetic ratio, respectively, g is the electronic g factor, r is the distance between the unpaired electron of MTSL and a given amide proton of Cc, and β is the Bohr magneton.

Structure determination on the basis of PRE restraints

The coordinates for Cc and CcP were taken from the Protein Data Bank [PDB ID 2PCB][81]. Four cysteine mutations (N38C, N200C, T288C and S263C) in CcP were introduced *in silico*, followed by the addition of MTSL atoms and energy minimization of the labelled protein. Docking of Cc onto CcP was done using restrained rigid body molecular dynamics in XPLOR-NIH[150]. The intermolecular distance restraints were divided into three groups. Residues that were strongly affected by the MTSL and whose resonances disappear in the paramagnetic spectrum were given only upper limit derived from an I_{para}/I_{dia} based on the noise level in the paramagnetic spectrum plus a 4 Å error margin. Those not affected by MTSL had only lower limit (18 Å). Residues that were affected by MTSL and whose resonances are observed in the paramagnetic spectra were given a PRE derived distance (equation 6.5) with error margins of ±4 Å. During docking, the coordinates of CcP were fixed, while Cc was placed at a random position and allowed to move under the forces of restraints and a van der Waals repel function. The repel function was set to zero for MTSL. To account for the mobility of the spin label, four orientations representative of the sterically allowed orientations of the oxygen atom of spin label were used. The restraint for a particular residue was defined as the averaged distance $(r^{-6})^{-1/6}$ between the O atoms of four representative spin label orientations and the amide protons. Structures below an energy threshold were saved.
In an alternative approach, short stretches of rigid docking were alternated with short intervals of molecular dynamics of the Cys residues with spin labels, while keeping the remainder of CcP as well as Cc fixed. In these runs all-atom descriptions of the SL were used. This procedure accounts for the spin label mobility during the docking. Four SL were present simultaneously for each position. The atoms of the four labels were mutually excluded from the van der Waals interaction energy term, allowing them to move through each other.

Monte Carlo simulations

A Boltzmann distribution of orientations of one protein around the other based on electrostatic interactions was generated by Monte Carlo simulations as described

earlier[169]. Proteins were treated as rigid bodies and the coordinates (PDB entry 2PCB[81] for wt hCc and K13A hCc, and 3N6P for K13R hCc) were taken from Protein Data Bank and this study. The atomic partial charges of the CHARMM force field[171] were assigned to all atoms including those in the haem and hydrogen atoms. For histidine residues that are not coordinated to the haem iron, average Charmm charges were used, which were calculated as 0.25 × (His_N^{ε} + His_N^{δ}) + 0.5 × protonated His. The dielectric constant used for the protein and water were 4 and 80, respectively. In order to match the simulation conditions to those of the NMR experiments, the simulation temperature and ionic strength were set to 303 K and 0.12 M, respectively. CcP coordinates were fixed and Cc was allowed to move in the electrostatic potential of CcP, calculated with program package MEAD[170]. The moves were accepted or rejected according to the Metropolis algorithm[197]. An exclusion grid was used to avoid steric overlap between the proteins. About 2500 configurations were saved for analysis.

Analysis of the ensemble

The structures in the ensemble were compared using the distance root mean square (DRMS) metric defined by equation 6.5

$$DRMS = \frac{1}{N}\sum_{i,j}\left|d_{ij}^{ens} - d_{ij}^{xrd}\right| \qquad (6.5)$$

where d_{ij} is the distance between the C^{α} atoms of two residues i and j from different proteins, N is the total number of pairs (i, j), and d_{ij}^{xray} and d_{ij}^{ens} are the distance matrix elements from the specific and ensemble structures, respectively.

For each structure in the ensemble, the distance between oxygen atom of the spin label on CcP and the Cc backbone amide hydrogen of all the residues was measured in X-PLOR. To account for the mobility of the spin label, four representative positions of the oxgen atom were used and the average distance was selected. From these distances, Γ_2 was back-calculated using equation 6.5. In order to compare with the observed values, the calculated Γ_2 was multiplied by the fraction of Cc bound to CcP, 62%, 79% and 49% for the wt, K13R and K13A hCc, respectively.

The quality of the ensemble was assessed by calculating a Q factor defined by equation 6.6

$$Q = \sqrt{\sum_i \left(\Gamma_{2,i}^{obs} - \Gamma_{2,i}^{calc}\right)^2 / \sum_i \left(\Gamma_{2,i}^{obs} + \Gamma_{2,i}^{calc}\right)^2} \qquad (6.6)$$

where $\Gamma_{2,i}^{obs}$ and $\Gamma_{2,i}^{calc}$ are the observed and calculated PREs.

The fraction of the encounter complex was found by calculating the Q factor for each value of *p* value between 0 and 1, with increments of 0.1, according to equation 6.7

$$\Gamma_{2,i}^{calc} = p\Gamma_{2,i}^{calc,ens} + (1-p)\Gamma_{2,i}^{calc,specific} \qquad (6.7)$$

where *p* and (1 - *p*) are the population fractions of the encounter and specific complexes, respectively.

Results

X-ray structure of K13R hCc-yCcP complex

The structure of the K13R hCc-yCcP complex was determined by X-ray diffraction of crystals. Data collection and refinement statistics of the K13R hCc-yCcP complex are listed in table 6.1. The structure was refined to an R factor of 0.18 (R_{free} of 0.24) at 2.1 Å. A representative electron density map of complex interface is shown in figure 6.1. The electron density showed that K13R hCc binds to CcP in a rotated orientation as compared to the wt hCc. Figure 6.1B shows the crystal structure of the K13R hCc-yCcP complex aligned with the structures of the wt hCc-yCcP and yCc-yCcP complexes. The figure was generated by superimposing the CcP molecules in order to show how the orientations of Cc molecules relative to CcP differ between the structures. In the K13R hCc-yCcP complex, the Cc haem group is shifted toward CcP. Interestingly, the orientation of K13R hCc and the position of its haem group resemble that of the yCc in yCc-yCcP complex, which has an Arg residue at the equivalent position. The K13R hCc-yCcP complex differs from the wt hCc-yCcP complex with an rms deviation of 2.4

Å, and from the yCc-yCcP complex with an rms deviation of 0.8 Å for the C^{α} atoms. The hCc-yCcP complex differs from the yCc-yCcP complex with rms deviation of 2.7 Å.

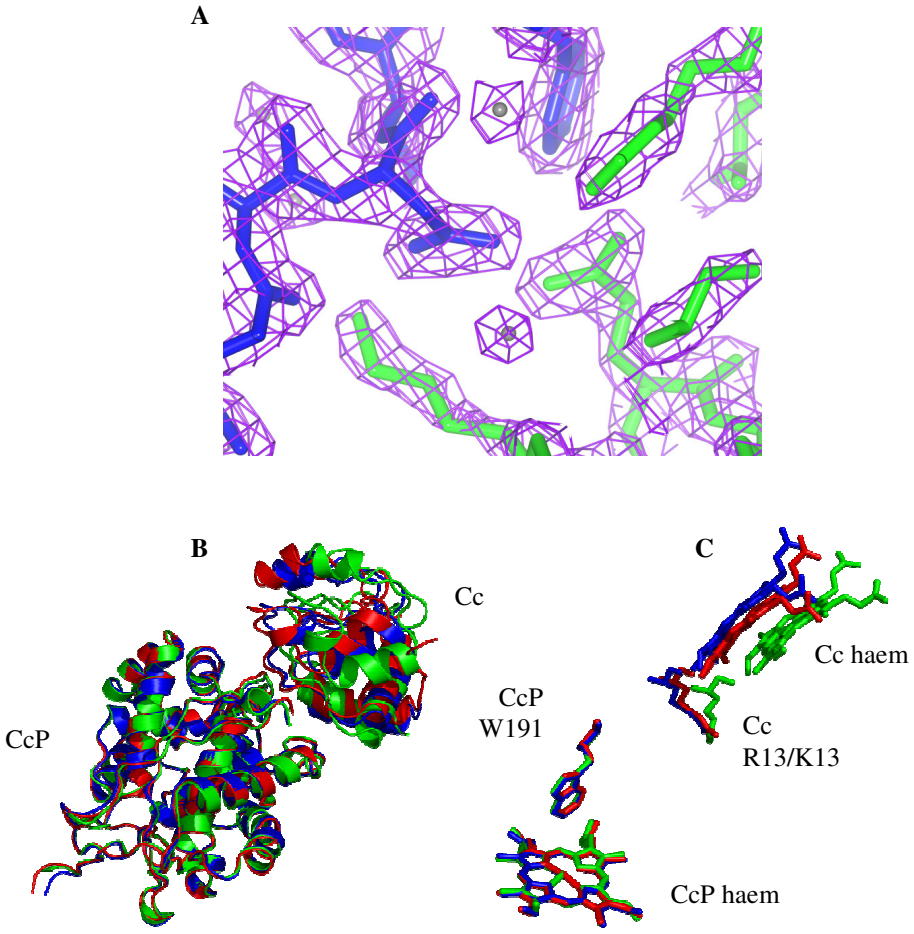

Figure 6.1. **(A)** Representative electron density at the interface of the K13R hCc and yCcP complex. Green and blue sticks represent Cc and CcP, respectively. The magenta lines represent the electron density. **(B)**, **(C)** Comparison of the crystal structures of yCc-yCcP (red), hCc-yCcP (green) and K13R hCc-yCcP (blue) complexes with the CcP molecules aligned. **(B)** Protein backbones shown as ribbons. **(C)** Showing only the haem groups, Cc residues R13/K13 and CcP residue W191.

Table 6.1. Statistics for the structure of K13R hCc-yCcP complex.

a. Data collection	
Source	Rotating anode ENRAF NONIUS FR 591
Detector	Image plate MAR345
Resolution range (Å)	67.25-2.10 (2.22-2.10)[a]
Multiplicity	11.3 (5.8)
Completeness (%)	98.9 (92.3)
R_{merge} [b] (%)	0.072 (0.198)
R_{pim} [c](%)	0.021 (0.086)
I/sigma	28.2 (9.5)
Space group	$P2_12_12_1$
Unit cell parameters a/b/c (Å)	44.9/87.7/104.7
Crystallization condition	15% PEG 3350, 150 mM NaCl
b. Molecular Replacement	
R factor of correct solution/second peak	0.464/0.592
c. Refinement	
No. Of reflections used in refinement	23186
Resolution range (Å)	52.37-2.11
R factor[d]	0.1835
R_{free} [e]	0.2372
R.m.s. deviations (bonds, Å/ angle, °)[f]	0.0093/1.133

[a]Values in parentheses are for the highest resolution bin, where applicable.

[b] $R_{merge} = \sum |\mathrm{I} - \langle \mathrm{I} \rangle| / \sum |\mathrm{I}|$

[c]Refs[198,199].

[d] $R = \sum \| F_{obs}(hkl) | - | F_{calc}(hkl) \| / \sum | F_{obs}(hkl) |$

[e]About 5% of the reflections were used for the cross validation set. These reflections were randomly chosen.

[f]Estimates provided by the program REFMAC[200].

The potential intermolecular hydrogen bonds noted in the K13R hCc-yCcP complex are different from those in the complex of wt hCc or yCc. The potential intermolecular hydrogen bonds are listed in the table 6.2. In the wt hCc-yCcP complex (at 2.8 Å resolution), Pelletier and Kraut[81] identified three intermolecular hydrogen bonds in the interface, between Asn38 of CcP and Lys8 of Cc, Glu35 of CcP and Lys87 of Cc and between Glu290 of CcP and Lys72 of Cc. The closest intermolecular contact observed

in the yCc-yCcP complex is a potential hydrogen bond between Asp290 of CcP and Asn70 of Cc. Inspection of the K13R hCc-yCcP complex interface reveals three potential hydrogen bonds. It retains the hydrogen bond between Asn38 of CcP and Lys8 of Cc, as in the wt hCc-yCcP complex. However, the other two hydrogen bonds noted in the wt complex are not observed in the K13R hCc-yCcP complex. The other potential hydrogen bonds observed in the K13R hCc-yCcP complex are between Glu34 of CcP and Arg13 of Cc, and between Glu34 of CcP and Lys87 of Cc, which are not observed in the wt hCc-yCcP or yCc-yCcP complexes. Although the charged groups on side chains of K13 in hCc and R13 in yCc and K13R hCc all locate in the same position relative to CcP (Figure 6.1C), only R13 of K13R hCc has been noted to be involved in the intermolecular hydrogen bond with CcP. However, in an earlier study of the yCc-yCcP complex by NMR spectroscopy (chapter 3) and isothermal titration calorimetry, the Cc R13 has been identified as a hot-spot residue which interacts with CcP Y39 [141].

Table 6.2. Comparison of the dissociation constants and intermolecular distances of interest in Cc-CcP complexes.

Protein Complex	K_d (μM)	Fe-Fe distance (Å)	Trp191-Haem distance (Å)	Potential hydrogen bond CcP-Cc
yCc-yCcP (2PCC)	5[a]	26.4	13.7	E290-N70 (3.3 Å)
hCc-yCcP (2PCB)	86	29.9	18.8	E290-K72 (3.3 Å) N38-K8 (2.9 Å) E35-K87 (3.2 Å)
K13R hCc-yCcP (3N6P)	44	26.7	14.1	N38-K8 (2.7 Å) D34-R13 (3.3 Å) D34-K87 (2.6 Å)
K13AR hCc-yCcP	240			

[a] Data taken from previous work (chapter 3).

Chemical shift perturbation analysis

Complex formation leads to changes in the chemical environment of residues in the interface due to solvation changes and molecular interactions between partner proteins. The resulting chemical shift perturbations can be used to identify residues involved in binding, and to estimate the binding constant. The size of NMR chemical shift perturbations also provides information about the dynamics in the complexes. Highly dynamic protein complexes result in very small chemical shift perturbations[29]. We have studied the complexes of yCcP with ^{15}N enriched wt hCc, K13R hCc and K13A, in reverse titrations using a fixed concentration of yCcP and varying the concentration of hCc. A single set of amide peaks in the [^{1}H, ^{15}N] HSQC spectra was observed indicating that hCc-yCcP complex is in fast exchange on NMR time scale. Chemical shift perturbations for the amide resonances were monitored in a series of 2D [^{15}N, ^{1}H] HSQC spectra. The chemical shift perturbations for some of the residues were plotted against the Cc/CcP ratio and the titration curves were fitted to a 1:1 binding model (Figure. 6.2). From these titration curves, the values of dissociation constants K_d were calculated. The estimated dissociation constants for the wt, K13R and K13A complex are 86 ± 7 μM, 44 ± 5 μM and 240 ± 14 μM, respectively. The binding curves of the complexes are shown in figure 6.2.

Chemical shift perturbations of Cc residues in the K13R, K13A and wt complexes exhibit a similar pattern, indicating that the proteins employ a similar interface to bind CcP. However, the size of average chemical shift changes extrapolated to fully bound Cc is much higher for K13R complex and much lower for the K13A complex as compared to the wt complex (Figure 6.3). These results suggest that the K13R mutation makes the complex more specific as compared to the wt complex, whereas the K13A mutation makes the complex more dynamic.

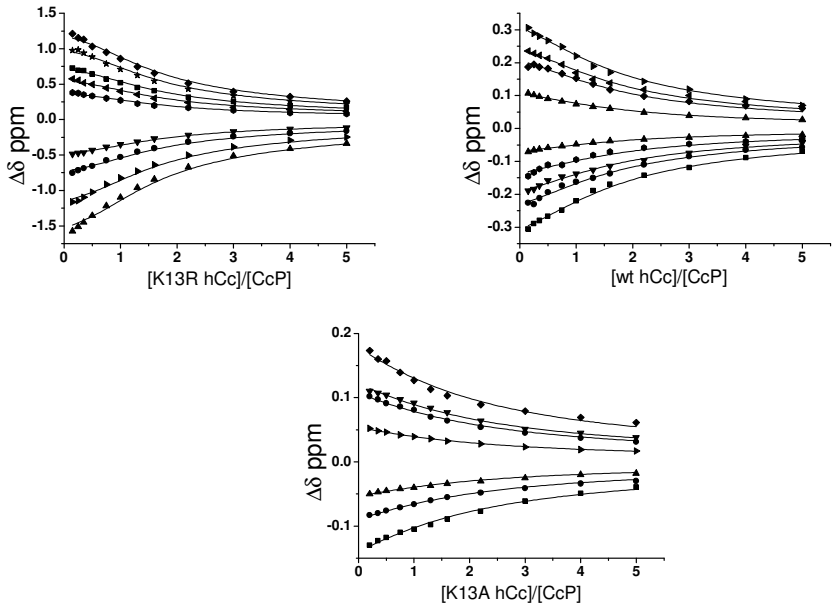

Figure 6.2. NMR titrations of K13R hCc, wt hCc and K13A hCc with yCcP. The solid lines represent the best simultaneous fit of the curves to 1:1 binding model with a shared K_d value.

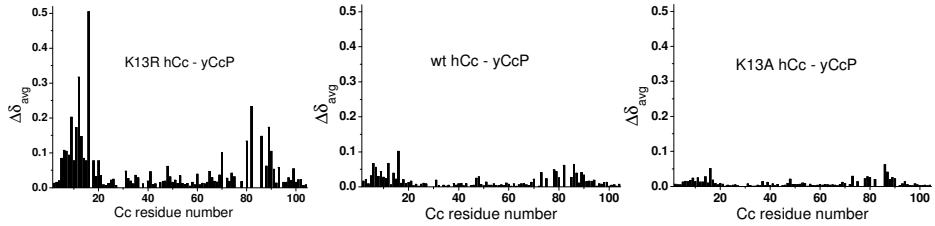

Figure 6.3. Chemical shift perturbations of K13R hCc, wt hCc and K13A hCc in complex with yCcP. The average chemical shifts (eq 6.1) are plotted versus the Cc residue number.

Paramagnetic relaxation analysis

The hCc-yCcP complex in solution was further studied using paramagnetic relaxation enhancement (PRE) NMR spectroscopy. Nitroxide spin labels were attached, one at a time, at ten positions spread over the entire surface of CcP (Figure 6.4) and their effects on the Cc backbone amide protons were monitored in [^{1}H, ^{15}N] HSQC experiments. I_{para}/I_{dia} ratios measured for the Cc backbone amide protons for K13R hCc wt hCc, and

K13A hCc are shown in figures 6.5, 6.6 and 6.7, respectively. Out of all spin label positions, only three located near the crystallographic binding site resulted in large paramagnetic effects. These were located at position 38C, 200C and 288C of CcP. Small effects were also observed for CcP with MTSL at position 213C. These findings indicate that the conformational search in the encounter complex is limited to an area surrounding the interface in the specific complex. In figure 6.8 the observed PREs (Γ_2) are compared with the Γ_2 back-calculated from the crystal structure of the complex. While many of the observed paramagnetic effects are satisfied by the specific complex, there are certain regions showing paramagnetic effects which do not arise from the formation of the specific complex. For the spin label attached to 288C, the residues 11-18 show higher paramagnetic effects than those expected for the specific complex. Similarly, for 200C the paramagnetic effects in the region 8-18 do not correlate with those for the specific complex. Similar regions can be found in wt, K13R and K13A complexes. These effects are attributed to contributions from the encounter complex, which exists in equilibrium with the specific complex, in analogy to the situation observed for the yCc-yCcP complex (chapter 4)[19,184].

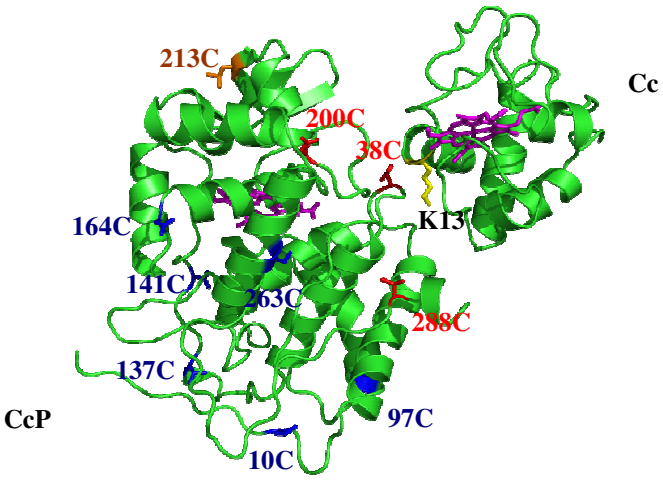

Figure 6.4. Spin label attachment sites on surface of CcP. Cc and CcP are shown as green ribbons. The spin label attachment sites are shown as red or blue sticks depending on whether they resulted in observable PREs or not. The 213C site which is shown as orange sticks, for which only few paramagnetic effects were

observed. The interface mutation site (K13) on Cc is shown as yellow sticks. The haem groups are shown as magenta sticks.

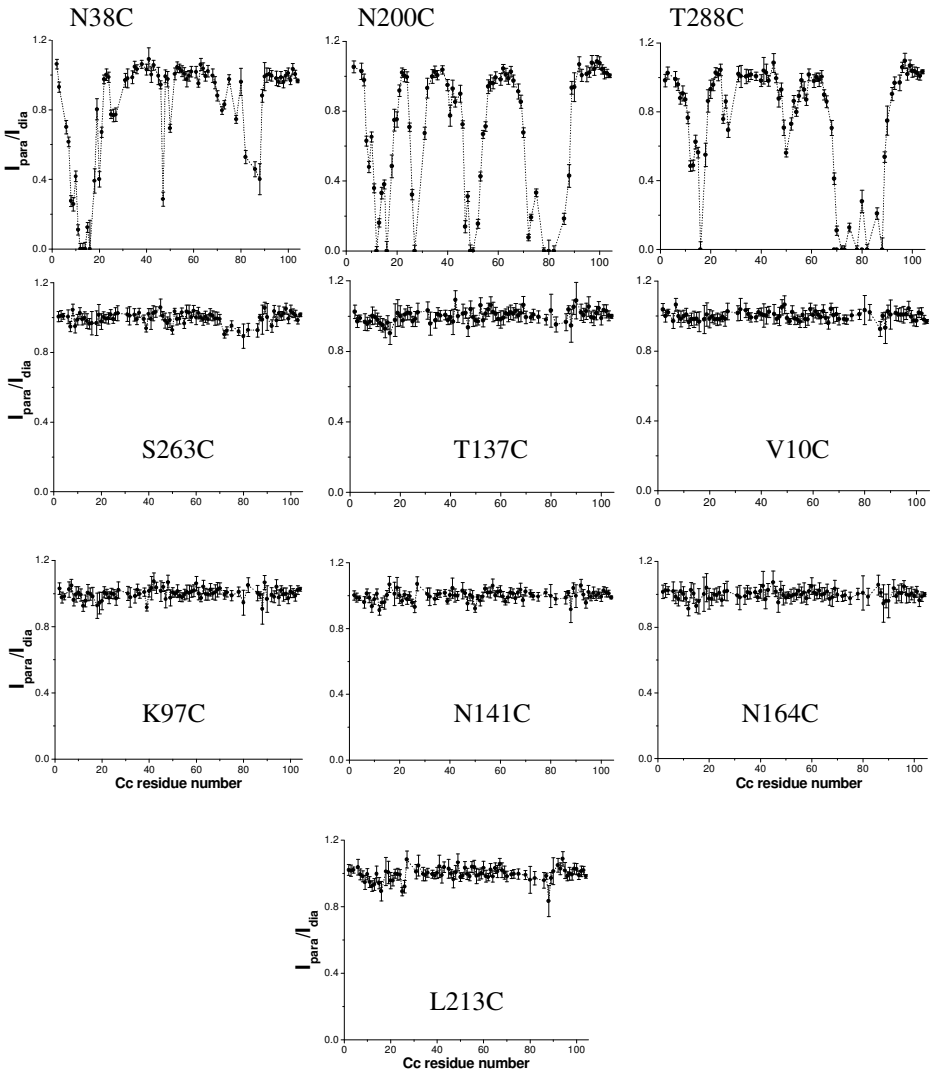

Figure 6.5. Paramagnetic effects from CcP-SL on K13R hCc. [^{1}H,^{15}N] HSQC intensities ratios are plotted for backbone amides of Cc in complex with CcP labeled with a paramagnetic spin label (I_{para}) or a diamagnetic analogue (I_{dia}).

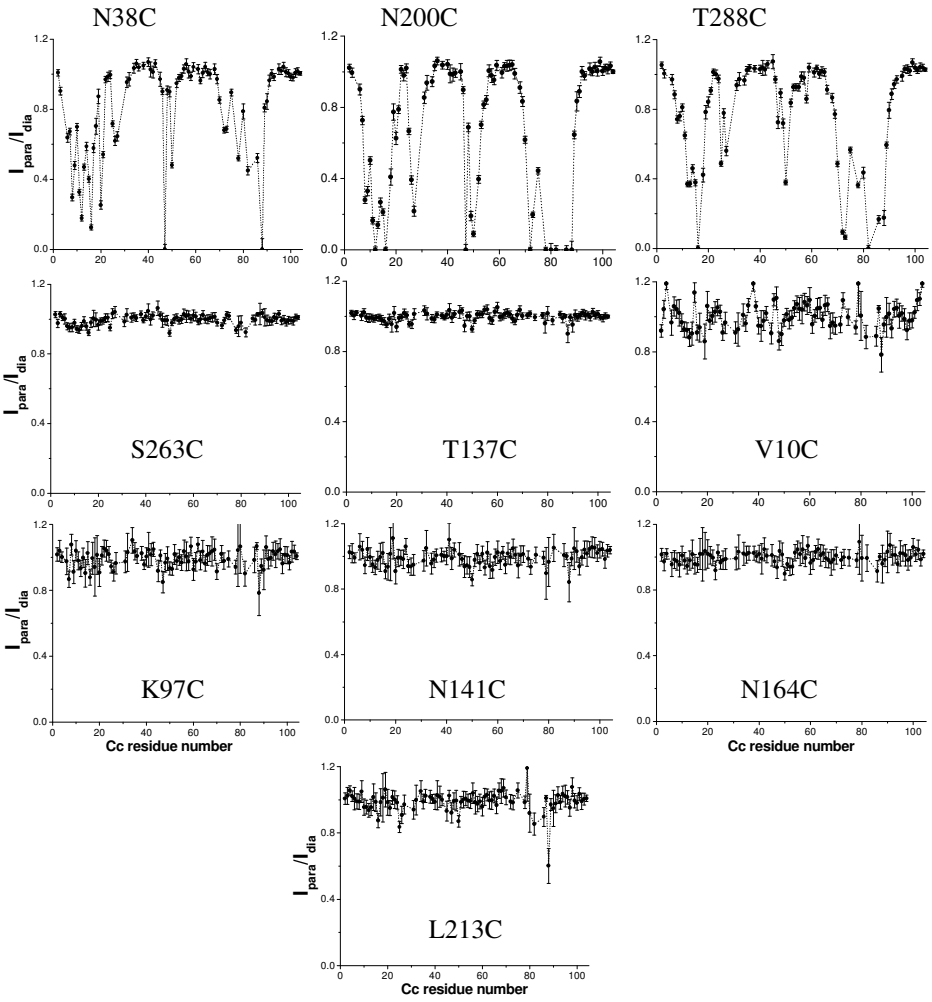

Figure 6.6. Paramagnetic effects from CcP-SL on wt hCc. [^{1}H,^{15}N] HSQC intensities ratios are plotted for backbone amides of Cc in complex with CcP labeled with a paramagnetic spin label (I_{para}) or a diamagnetic analogue (I_{dia}).

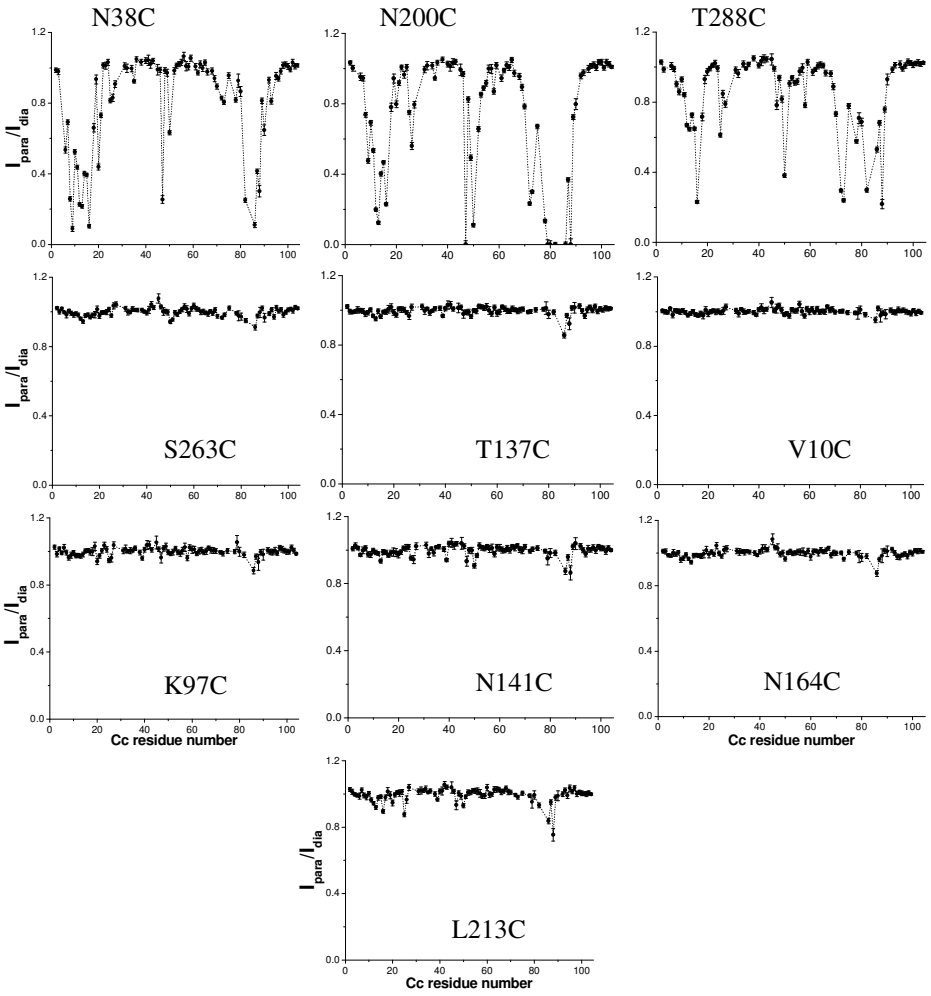

Figure 6.7. Paramagnetic effects from CcP-SL on K13A hCc. [^{1}H,^{15}N] HSQC intensities ratios are plotted for backbone amides of Cc in complex with CcP labeled with a paramagnetic spin label (I_{para}) or a diamagnetic analogue (I_{dia}).

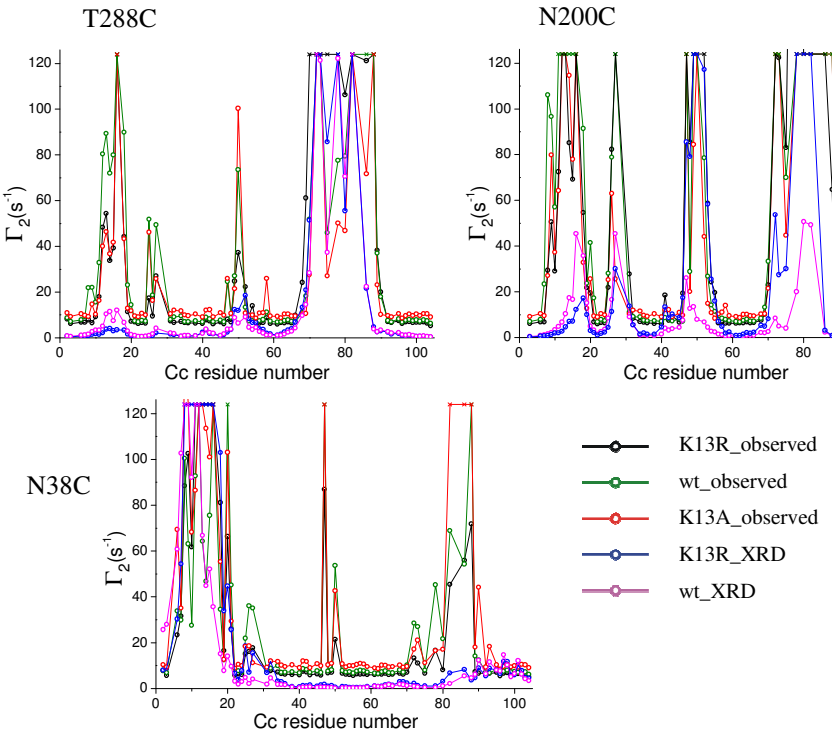

Figure 6.8. Paramagnetic relaxation analysis. The observed PREs (corrected for 100% bound) for the three spin label positions are compared with the calculated PREs from the crystal structures of wt and the K13R hCc-CcP complex.

Solution structure of hCc-yCcP complex

We have attempted to determine the solution structure of the hCc-yCcP (wt and the mutants) complexes by molecular docking using intermolecular distance restraints obtained from PRE NMR spectroscopy. We have used distance restraints from four spin label positions including N38C, N200C, T288C and S263C. As the spin label is attached to the protein through a single covalent bond, it can show a high degree of mobility. To account for the mobility of the spin label, we have used four orientations for each spin label position. In order to check whether the spin label orientations affect the final structure, we determined the structure of the complex several times with different sets of four spin label orientations. The lowest energy structure was selected as

the best solution. Each set of the spin label orientations resulted in a different solution (Figure 6.9). The structures using two different sets of spin label orientations differ from each other with rms deviation of 2.8 Å and from the crystal structure with rms deviation of 3.5 Å and 3.7 Å, respectively. We also performed docking calculations using all the spin label atoms and allowing them to rearrange during docking. Every run resulted in a different structure, yielding an ensemble of lowest energy structures, shown in figure 6.9. The five lowest energy structures differ from each other with rms deviation ranging from 2.3 Å to 4.5 Å, with an average rms deviation of 3.0 Å. These structures differ from the crystal structure with rms deviation of 2.0 Å to 3.8 Å. We have done these calculations with wt hCc and K13R hCc with similar results. This analysis demonstrates the resolution limits of protein complex structure calculation using freely rotatable paramagnetic centres. Clearly, insufficient detail can be obtained for comparison of the interfaces. More rigid paramagnetic probes are likely to be better suited for structure determination.

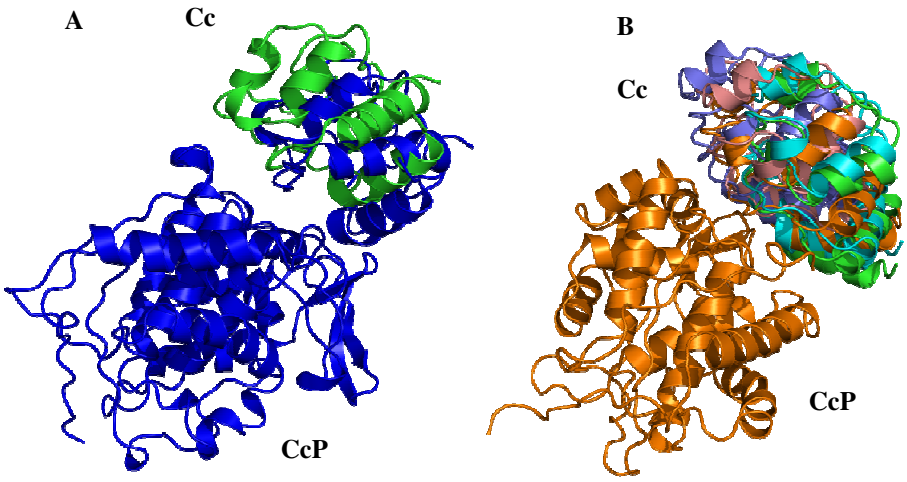

Figure 6.9. Solution structure of the hCc-yCcP complex using restrained rigid body docking. (A) The lowest energy structures using different sets of fixed orientations of the spin labels in two separate runs. (B) The five lowest energy structures from a single run using sets of flexible spin labels.

Simulation of the encounter ensemble

The PREs represent an average of the paramagnetic effects in the specific complex and all orientations of the encounter complex. PRE NMR spectroscopy does not provide information about the individual orientations in the encounter complex, and the relative population of the encounter and the specific complex. Previous studies[17] have shown that the encounter complex is dominated by the long range electrostatic forces, contrary to the specific complex which is dominated by the short range interactions. In our previous study[184] (chapter 4), we have successfully simulated the encounter ensemble using Monte Carlo simulations and considering electrostatic and steric interactions only. The Monte Carlo simulations provide the individual orientations of the two proteins in the encounter complex, from which the averaged contribution of the encounter complex to the PRE can be calculated, and the fractions of the encounter and the specific states estimated.

Figure 6.10 shows the simulated encounter ensemble of hCc and yCcP with centers of mass of one protein around the partner protein. The simulated ensemble shows the Cc distribution restricted near the crystallographic binding site on CcP, consistent with the experimentally observed data. In the encounter ensemble of hCc-yCcP, the distribution of Cc shows three electrostatically attractive regions around CcP. Two regions are similar to those observed in the yCc-yCcP ensemble (chapter 4, figure 4.3). However, the third region near Asp148 of CcP was not observed in the yCc-yCcP ensemble[184] (chapter 4). The small paramagnetic effects observed for spin label attached to 213C support this observation.

The encounter ensembles of the K13R hCc-yCcP wt hCc-yCcP, and K13A hCc-yCcP were compared using the distance root mean square (DRMS)[23] of the structures relative to the crystal structures. The DRMS values for the K13A hCc-yCcP ensemble structures were calculated relative to the crystal structure of the wt complex. The plots of DRMS versus energy of the structures are shown in figure 6.11. The structures in the K13R hCc-yCcP ensemble have lower energy compared to those in wt and K13A ensemble.

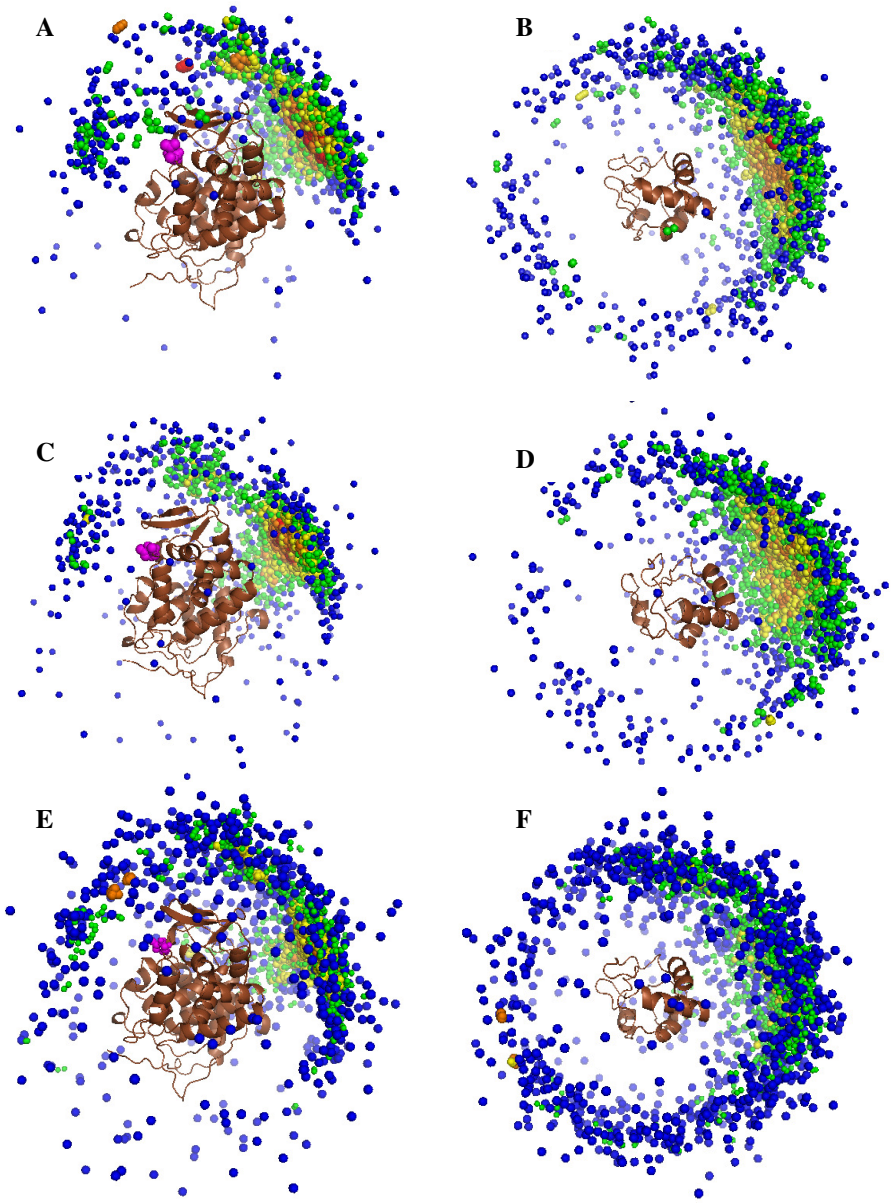

Figure 6.10. Monte Carlo simulations of the K13R hCc-yCcP **(A,B)**, wt hCc-yCcP **(C,D)** and K13A hCc-yCcP **(E,F)**, showing the centers of mass of Cc **(A,C,E)** or CcP **(B,D,F)** around the partner protein for the K13R, wt and K13A complex. The centers of mass are colored to show the density of distribution decreasing from red to blue. The highest densities denote the most favorable electrostatic orientations. Asp148 of CcP is shown as magenta spheres.

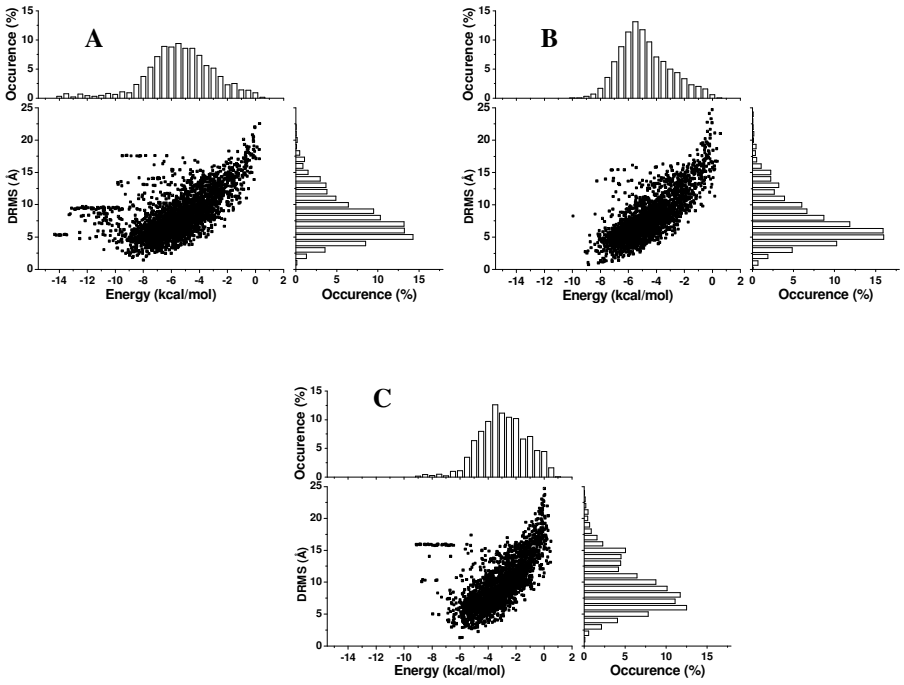

Figure 6.11. DRMS versus energy plots for the structures in the encounter ensembles of K13R **(A)**, wt **(B)** and K13A **(C)** complex.

Population of the encounter complex

The relative populations of the encounter and the specific complexes in solution were obtained by fitting the back-calculated Γ_2^{calc} of the specific complex and the encounter ensemble, to the experimentally observed Γ_2^{obs}. Γ_2^{calc} for the encounter complex and the specific complex can be back-calculated from the simulated ensemble and the crystal structure of the complex, respectively using equation 6.5. The back-calculated Γ_2^{calc} for the encounter ensemble (p=1) were linearly combined with the Γ_2^{calc} from the specific complex (p=0) by varying the population of p from 0 to 1 and compared to the observed Γ_2^{obs}. The quality of the fit was determined by calculating a quality factor (Q factor)[184] using equation 6.7. The relative population that provided the best Q fator was selected as the population of the encounter complex (Figure 6.12). As shown in figure 6.13, the back-calculated data for the K13R mutant provided a good fit with the experimental

data. The population of the encounter complex obtained for the K13R hCc-yCcP is 40%. Thus the K13R hCc-CcP complex exists as equilibrium of 40% encounter complex and 60% specific complex. The same relative population was obtained using Γ_2^{obs} from the three spin label positions (38C, 200C and 288C). However, we could not obtain a good fit of the data for wt and the K13A mutant. The individual SL show very different behaviour and for some no good agreement between the observed and predicted PREs can be attained.

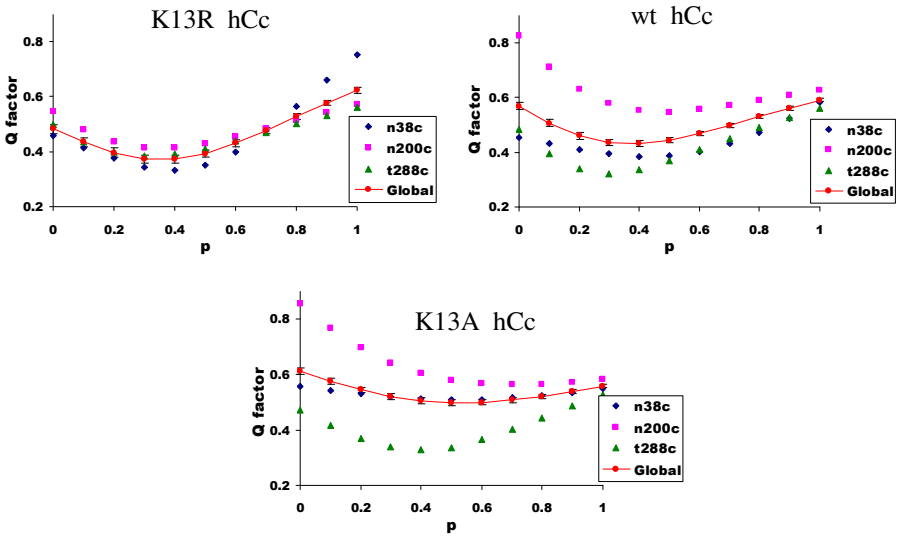

Figure 6.12. Population of the encounter complex. The Q factors (eq 6.7) for three spin label positions are plotted against p for the complexes of K13R, wt and K13A hCc with yCcP.

Discussion

A single residue interface mutation has been shown to alter the specificity of the complex. The results demonstrate that the K13 mutation can turn the highly dynamic non-physiological hCc-yCcP complex into a less dynamic and more specific complex more like the physiological yCc-yCcP complex. Thus, the R13 is crucial for the specificity of the complex. In contrast, the complex can be made more dynamic by the K13A mutation. Thus a careful selection of the interface mutation can enhance or reduce the specificity of the complex. These results suggest that transient protein-

protein interactions are highly tuned and sensitive to small changes at the interface. Introduction of a larger or a smaller side chain at the interface can affect the structure, affinity and specificity of the complex.

The encounter ensemble of the hCc-yCcP complex shows a diffusive distribution of Cc structures near Asp148 of CcP. In the Monte Carlo simulations of yCc and yCcP by Bashir et al (chapter 4)[184] and the the Brownian dynamics performed by Gabdouline and Wade[103] this has not be observed. However, Northrup et al[102] have observed Cc structures in this region in their Brownian dynamic simulations of tuna Cc and yCcP. This may be related to the the difference in the amino acid sequence of the proteins. The closest spin label to this region is the one attached to 213C. This spin label results in a few small paramagnetic effects. These effects can be improved by application of a stronger spin label in this region.

The K13R mutation has been shown to alter the orientation of the Cc molecule in complex with CcP. The K13R hCc acquires an orientation that resembles closely the one found for yCc. The crystal structure of the K13R hCc-yCcP complex differs from the yCc-yCcP complex and the wt hCc-yCcP complex with rms deviation of 0.8 Å and 2.3 Å, respectively. However, the structure determination by molecular docking using PRE restraints resulted in different solutions having rms deviation of 2.0 Å to 3.8 Å from the crystal structure of the wt hCc-yCcP complex, making it impossible to study the effect of the mutation on the structure of the complex.

We could not obtain a reliable solution structure of the complex because of the mobility of the spin label. The nitroxide spin label is attached to the protein surface by a single disulphide bond and have higher mobility in the three dimensional space due to freedom of rotation around the single bond. We have used four representative positions of the spin label to account for the mobility. However, this does not seem to be sufficient to get a reliable structure as the different set of spin label orientations results in a different structure. The structure determination by this method requires a more rigid spin label. Recently Caged Lanthanide NMR probes[201] are being used for structural studies in protein-protein and protein-ligand interactions. These probes are attached to the protein surface via two cysteine residues and are more rigid. Application of such probes may offer a more reliable method to determine the solution structure of protein complexes.