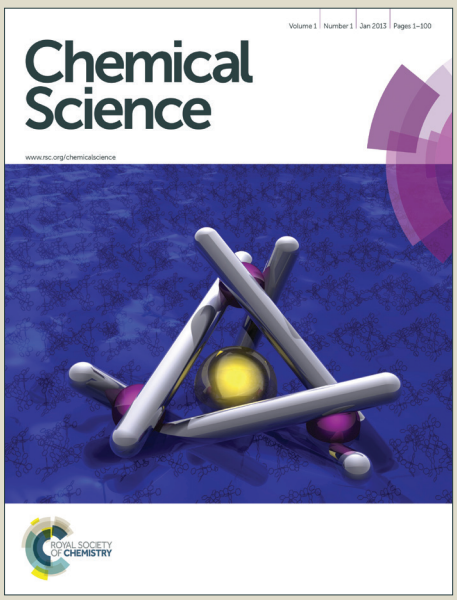
Chemical Science

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemicalscience

Chemical Science

RSCPublishing

ARTICLE

Transient Protein Encounters Characterized by Paramagnetic NMR

Cite this: DOI: 10.1039/xoxxoooooox

K. Van de Water,[a,b] N. A. J. van Nuland[a,b] and A. N. Volkov[a,b,*]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

Invisible to most biophysical techniques, transient intermediates formed on the path of biomolecular association orchestrate protein recognition and binding. Here we study such minor species mediating the interaction between physiological partners cytochrome *c* and cytochrome *c* peroxidase by paramagnetic relaxation enhancement NMR spectroscopy. Visualization of multiple protein-protein orientations constituting the transient encounter state reveals a broad spatial distribution, which is in striking agreement with that obtained in earlier theoretical simulations. Being inactive in the intermolecular electron transfer, the encounter complex pre-orients the interacting molecules, enabling the reduced dimensionality search of the dominant, functionally active bound form. The encounter complex is insensitive to the redox and spin states of the interacting molecules, suggesting that its properties are determined by the protein polypeptides rather than their heme cofactors.

Introduction

For a long time it has been recognized that many biomolecular interactions proceed via lowly-populated, transient intermediates.[1] Believed to govern the macromolecular recognition and binding, these short-lived species are inaccessible to conventional biophysical and structural techniques and, until recently, could only be studied by theoretical simulations.[2,3] Recent development of the experimental approaches sensitive to the presence of minor species – in particular relaxation dispersion and paramagnetic relaxation enhancement (PRE) NMR spectroscopy – has enabled direct detection and detailed structural characterization of such transient intermediates.[2,3] Thanks to the methodological advances made in the past decade, the hitherto inaccessible short-lived biomolecular encounters were visualized by PRE NMR in a number of protein-DNA and protein-protein complexes.[4-7]

The PRE is caused by a dipolar interaction between a nucleus and the unpaired electron of the paramagnetic center, either present in the native protein or introduced by bioconjugation techniques as a paramagnetic label. Because of the large magnetic moment of an unpaired electron and the r^{-6} distance dependence, the PRE is a long-range effect, which can extend up to 35 Å and is exquisitely sensitive to the presence of minor species.[2,3] For a number of practical reasons, the PRE is generally measured as the difference in the transverse relaxation rates of a proton in a paramagnetic sample and a diamagnetic reference, and a set of such PREs (^{1}H Γ_2) is used for the subsequent structural analysis.[2,3] If the major and minor forms of the protein complex are in fast exchange, the observed PRE (Γ_2^{obs}) is given by the sum of the population-weighted contributions from the major (Γ_2) and minor (Γ_2^*) species (Eq. 1):

$$\Gamma_2^{obs} = (1 - p)\Gamma_2 + p\Gamma_2^* \qquad (1),$$

where p is the fractional population of the latter. Provided that the distance between the paramagnetic center and the observed nucleus is shorter in the minor form than in the major one, the Γ_2^{obs} will contain a significant contribution from the minor species.

Here we use PRE NMR spectroscopy to study transient intermediates in the complex formation between the yeast cytochrome *c* peroxidase (CcP) and cytochrome *c* (Cc). Located in the mitochondrial intermembrane space, CcP is a 34.2 kDa heme enzyme, which catalyzes reduction of hydroperoxides using the electrons provided by its 12 kDa physiological partner Cc. The catalytic mechanism of H_2O_2 reduction involves formation of CcP compound I (CpdI), a low-spin species oxidized two equivalents above the CcP(Fe^{3+}) resting state (RS) and containing the Fe(IV)=O heme oxyferryl and the W191 cation radical.[8,9] Subsequent CpdI reduction in two one-electron steps involves complex formation with ferrous Cc, intermolecular electron transfer (ET), and product dissociation. The Cc-CcP system has been widely investigated and become a paradigm for understanding biological ET.[8,9]

Earlier studies revealed that the dominant Cc-CcP binding form in solution – essentially the same as that observed by X-ray crystallography[10] – is in equilibrium with an ensemble of lowly-populated protein-protein orientations, which constitute an encounter state populated for 30 % of the total lifetime of the complex.[6,11] However, due to the technical and conceptual limitations imposed by the weak, redox-sensitive paramagnetic label and the experimental design wherein effects from the larger, paramagnetically tagged protein were observed on the smaller one – the protein complex could not be studied in its functionally relevant redox and spin states, and the collected PRE data were insufficient for the full description of the conformational space sampled in the protein encounter.[12]

In this work, the above limitations were overcome by labeling the smaller Cc with EDTA(Mn) tag – which provides stronger PREs and has a higher reduction potential than the nitroxide spin label used in previous studies – and observing the effects on the larger CcP, whose HSQC spectra have been assigned recently.[13,14] This setup enables experiments with ferrous Cc (Cc^{red}), not feasible before due to the spin label

Chemical Science Accepted Manuscript

reduction, and affords the complete coverage of the encounter's conformational space. Here, we studied the complex of the cyanide-bound CcP [CcP(CN)] and Cc^{red}, which is in the same redox, spin, and ligation states as the active Cc^{red}-CpdI species and was shown to be a suitable structural mimic.[15] In addition, to investigate the role of the protein redox and spin states on the encounter complex formation, the CcP(CN)-Cc^{ox} and the original, resting-state CcP(RS)-Cc^{ox} systems were studied.

Results and discussion

The EDTA(Mn^{2+}) attachment does not perturb the native Cc-CcP interaction

Employed for protein modification in several PRE NMR studies,[5,16] the metal-chelating probe used here has a good leaving group (thiopyridine), enabling the specific and high-yield reaction with protein thiols, and harbors the strongly chelating EDTA moiety, allowing tight binding of the metal of choice. As established by a spectrophotometric assay of the free thiols remaining after the conjugation, the labeling yields were consistently close to 100 %, and a single, well-defined product was observed throughout by mass spectrometry (Fig. S1). With K_Ds of $2.8 \cdot 10^{-14}$ M and $3.2 \cdot 10^{-17}$ M for EDTA(Mn^{2+}) and EDTA(Zn^{2+}), respectively,[17] the Mn^{2+} and Zn^{2+} metal ions were used here to generate highly stable, isosteric protein conjugates for the paramagnetic and diamagnetic samples.

To ensure that the introduced EDTA(metal) group does not perturb the protein-protein interaction, the probe attachment sites on Cc were designed outside the crystallographic Cc-CcP binding interface. Furthermore, to preserve the overall charge of the native protein upon conjugation to the metal-containing chelate with a net charge of -1, an acidic protein residue (Asp or Glu) was mutated to cysteine, required for the probe attachment, in each of the Cc variants. All CcP – Cc-EDTA(Mn^{2+}) complexes studied in this work have virtually the same binding constants and chemical shift perturbations as those of the wt complex (Fig. 1), confirming that the introduced probe does not affect the native protein binding.

Ambiguity in the NMR resonance assignments resolved by the PRE analysis

At earlier stages of this work, we observed an anomaly in the Γ_2 profiles of CcP complexes with E66C and E88C Cc-EDTA(Mn^{2+}): residues 11 and 12 seemed to feel very strong PREs (their resonances disappeared in the paramagnetic spectra), while their neighbors did not and, similarly, residues 292 and 293 appeared to be unaffected, while the adjacent stretch experienced strong paramagnetic effects (Fig. 2A, top). The PRE profiles were reproducible, which ruled out experimental errors, and ensemble refinement against this dataset failed to produce solutions that satisfied restraints for residues 11, 12, 292, and 293. These outliers raised a suspicion, which prompted us to re-examine the resonance assignments of the CcP backbone amides. Close scrutiny of the 3D NMR spectra acquired and analyzed in the earlier work[13] revealed that – due to the very similar Larmor frequencies of the carbon resonances of G293 and G13 – the backbone atom assignments for two triplets of residues, G13-K12-E11 and G293-Q292-E291, appeared to be swapped (Fig. 2B). Once this mistake was corrected, the resonances in question showed sequential connections to neighboring residues and revealed hitherto unobserved $C\alpha_{i-1}/C\alpha_i$ and CO_{i-1}/CO_i connectivities between E291 and E290 resonances, confirming the assignment. With this correction in place, the PRE profiles showed a consistent

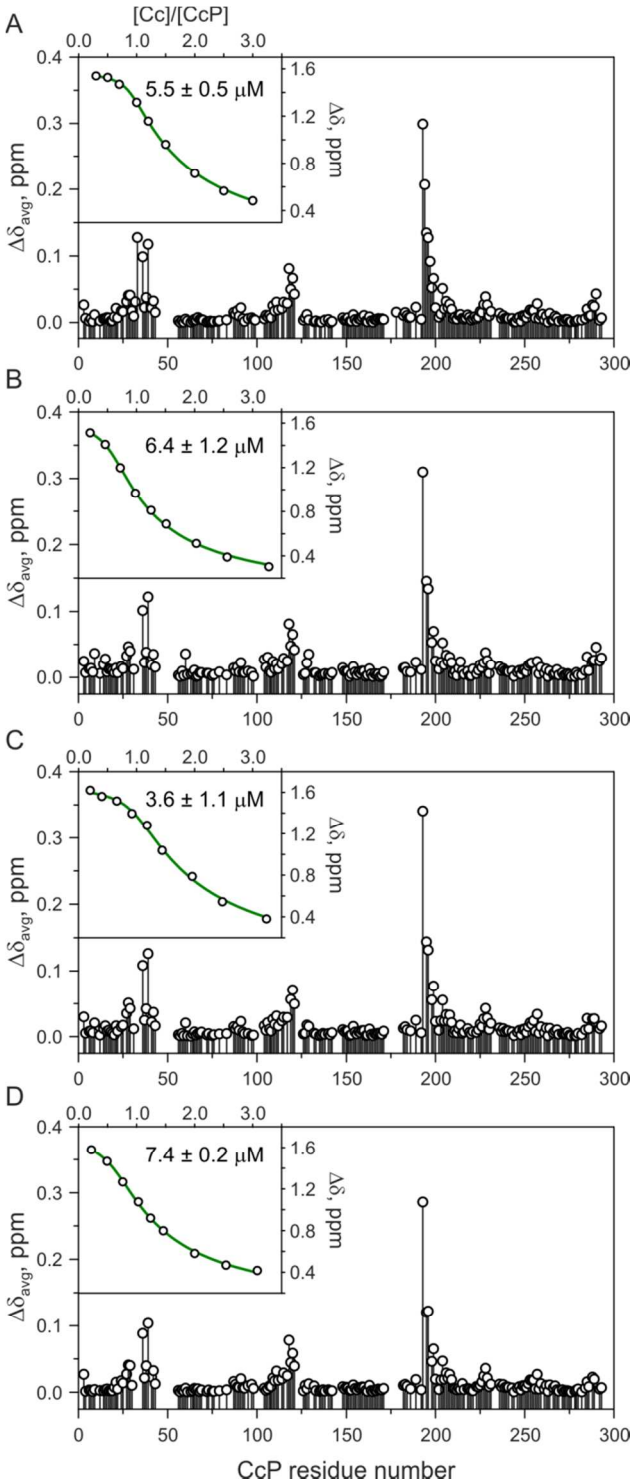


Fig. 1. CcP binding properties of the Cc-EDTA(Mn^{2+}) constructs studied in this work. Average chemical shift perturbations for the [D,^{15}N] CcP(RS) complexes with (A) wt Cc^{ox} and (B) D50C, (C) E66C, and (D) E88C Cc^{ox}-EDTA(Mn^{2+}) conjugates. The insets show chemical shift changes ($\Delta\delta$) of the 3-CH_3 heme resonance of Cc upon binding to the CcP and the best fits of the data to the binding model (Eq. 5). The values of the equilibrium dissociation constants derived from the titration curves are indicated. All experiments were conducted in 20 mM sodium phosphate, 100 mM NaCl, pH 6.0 at 303 K.

Chemical Science Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2014

pattern without strong outliers (Fig. 2A, bottom), which was well reproduced by the ensemble refinement solutions (see below). We have updated BMRB entries 19004, 19005, 19075, and 19076, containing the backbone chemical shift assignments of CcP(CN), CcP(RS), and the iron-free protoporphyrin IX CcP.

Intermolecular PREs in the Cc-CcP complex

The intermolecular PREs were obtained from a series of experiments on the CcP bound to paramagnetically tagged, single-cysteine Cc variants D50C, E66C, and E88C (one at a time). As discussed above, the label attachment at any of these sites did not perturb the native Cc-CcP binding, and distinctive PRE profiles were obtained for the CcP complexes with each of the Cc-EDTA(Mn) conjugates (Fig. 3). Most of the PREs originate from the dominant, crystallographic orientation of the protein complex; yet several regions (highlighted in Fig. 3) feature Γ_2^{obs} that cannot be accounted for by a single Cc-CcP structure. Arising from the PRE contributions from the encounter complex, such discrepancies are the footprint of the minor species.[2,3] Being very similar for different complexes studied in this work, the Γ_2^{obs} profiles are reproducible and largely insensitive to the choice of the PRE measurement method (Supplementary Text and Figs. S2-S5), confirming that the additional effects are real and not experimental artifacts. Furthermore, the control experiments with an unrelated paramagnetically labeled protein ubiquitin show absence of any PREs, ruling out aspecific interactions of the attached EDTA(Mn) label with the CcP (Fig. S6).

Estimating the population of the encounter state

Extracting the Γ_2^* values – necessary for the structural analysis on the encounter state – from the measured PREs requires the knowledge of the Γ_2 and p (Eq. 1). While the former can be back-calculated from the structure of the dominant form of the complex (solved by X-ray crystallography),[10] the latter is not easily accessible experimentally. In practice, the p is varied during the ensemble refinement against the measured PREs, and the value yielding the best match with Γ_2^{obs} data is selected. The agreement between the calculated, ensemble-based PREs and the experimental Γ_2^{obs} is quantified by calculating the Q factor (Eq. 7), with smaller Q values indicating a better match.

Intermolecular pseudocontact shifts. Here, in essentially the same experimental setup as that used for the PRE measurements, we infer the encounter state population from the intermolecular pseudocontact shifts (PCSs) caused by the heme iron of the Cc and detected on the backbone amide protons of CcP. Arising from a paramagnetic center with an anisotropic magnetic susceptibility ($\Delta\chi$) tensor, the PCSs are readily detectable as the differences in chemical shifts of the protein resonances in the oxidized and reduced forms (Fig. 4A), providing that no other factors (redox-dependent structural changes, discrepancies in the experimental conditions of the two samples, etc.) contribute to the measured chemical shifts. Due to the r^{-3} distance dependence, the PCSs are much less sensitive than PREs to the presence of minor species and, thus, mainly report on the dominant form of the protein complex, enabling estimates of its population.

In a general approach, similar to that used here for the Mn^{2+} labeling, the PCS-causing metals (e.g. lanthanides) bound to appropriate chelating groups can be introduced into the

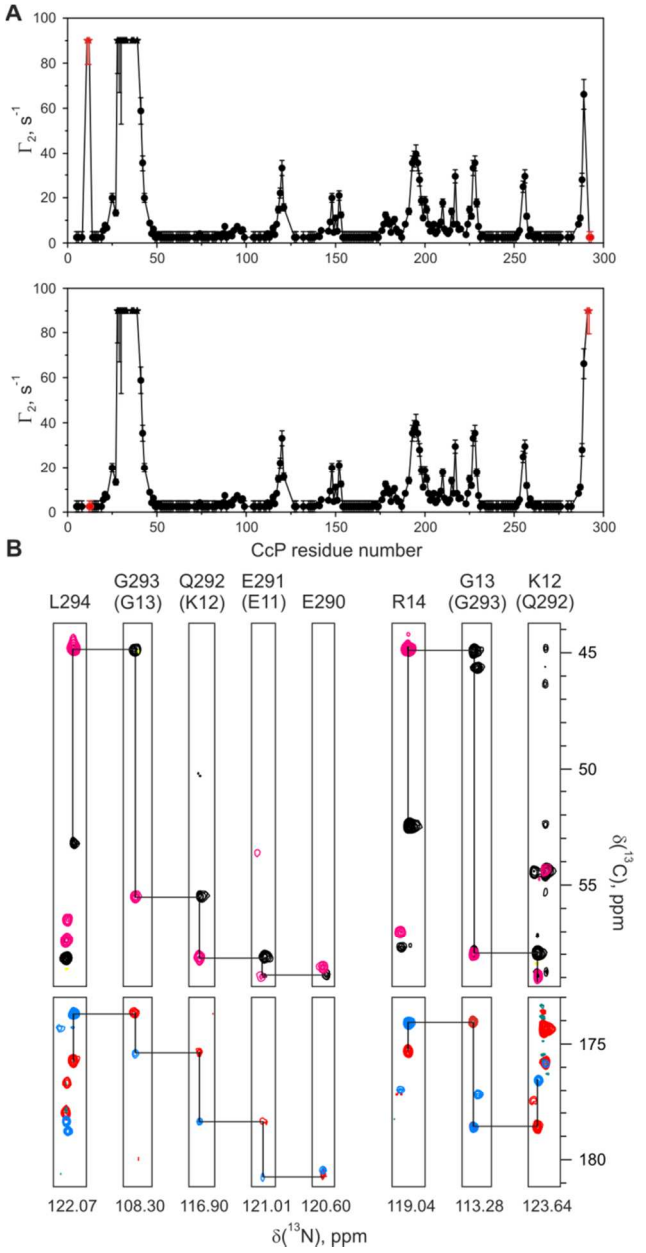


Fig. 2. Corrected assignments of the NMR resonances of CcP(CN). (A) PRE profiles of [D, ^{15}N] CcP(CN) in complex with E88C-EDTA(Mn^{2+}) Cc^{red} before (top) and after (bottom) the assignment correction. The Γ_2 data points for the CcP residues E11, K12, E291, and Q292, whose assignments were rectified, are shown in red. Stars indicate the residues whose resonances disappear in the paramagnetic spectrum. The errors are standard deviations. The HSQC spectra for the PRE analysis were recorded with d_1 = 10 s. (B) Selected ^{1}H-^{13}C planes of the overlaid 3D HNCA (black), HN(CO)CA (pink), HNCO (blue), and HN(CA)CO (red) spectra of [D, ^{13}C, ^{15}N] CcP(His)$_6$-CN taken at the nitrogen frequencies of the indicated backbone amide resonances. The spectra were acquired and analyzed in the previous work.[13] Vertical lines connect pairs of peaks in complementary Cα (above) and CO (below) correlation experiments, while the horizontal lines indicate the corresponding sequential connectivities. The former, incorrect, assignments are given in parentheses.

Chemical Science Accepted Manuscript

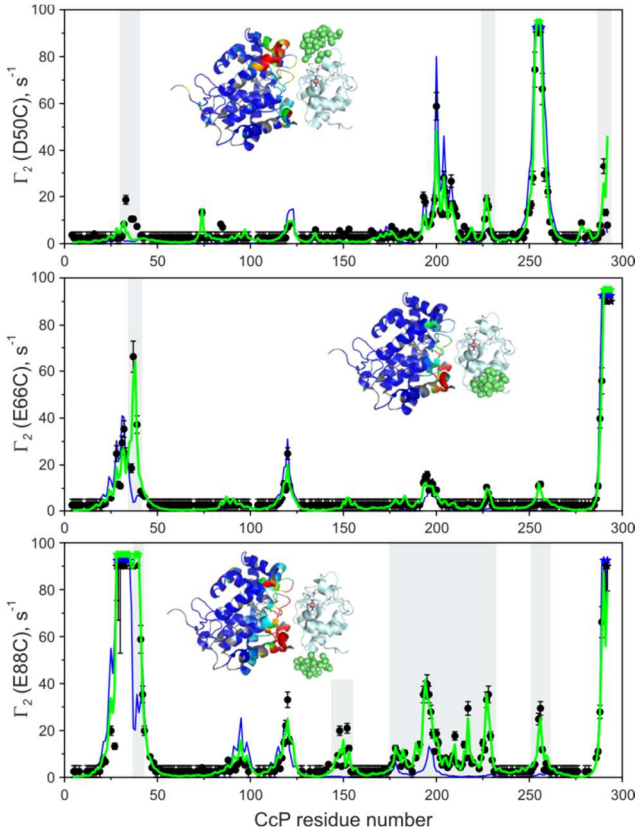


Fig. 3. Intermolecular PREs for the CcP(CN) in the complex with Cc paramagnetically labeled at D50C (top), E66C (middle), and E88C (bottom). The plots show measured PREs (Γ_2^{obs}, black symbols), Γ_2 values back-calculated from the crystallographic orientation (blue line), and the PREs calculated for the combination of the specific and the encounter complexes (Γ_2^{calc}, green line), obtained in a typical ensemble refinement run (p = 0.4, Q = 0.3). Stars indicate the values of $\Gamma_2 \geq 90$ s^{-1} for the calculated PREs or identify the residues whose resonances disappear in the paramagnetic spectrum. The errors are standard deviations. The regions exhibiting the PRE contributions from the encounter complex are highlighted. The insets show the X-ray structure of the complex, with Cc in light blue and CcP coloured by the plotted Γ_2^{obs} (from ≤ 2 s^{-1} in blue to ≥ 30 s^{-1} in red). The CcP residues with no measured PREs are in gray. The green spheres correspond to the Mn^{2+} atoms of multiple conformers of the attached paramagnetic label used for the ensemble averaging.

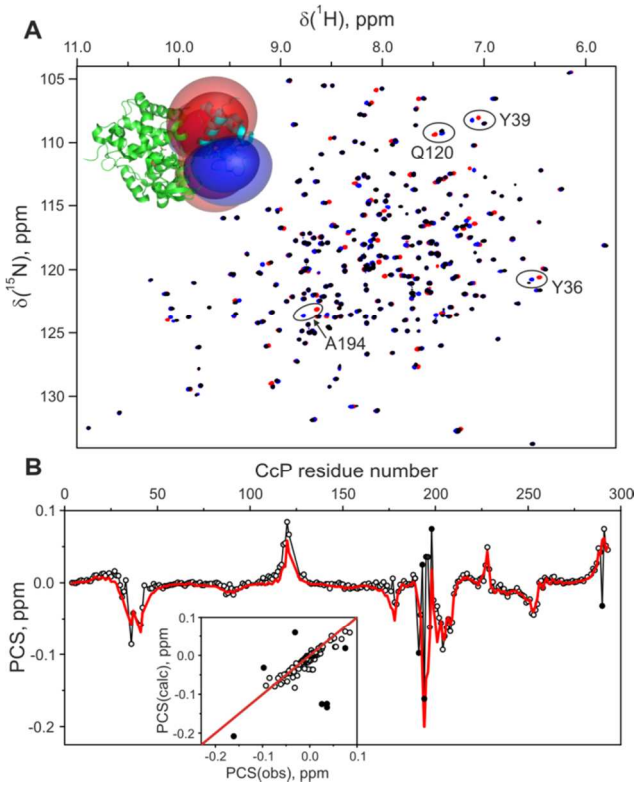


Fig. 4. The PCS analysis of the Cc-CcP complex. (A) An overlay of the HSQC spectra of [D, ^{15}N] CcP(CN) in the free form (black) or bound to Cc^{ox} (red) or Cc^{red} (blue). Several resonances experiencing dissimilar chemical shifts in the two complexes are indicated by the labels. The inset shows the crystallographic Cc-CcP orientation and the PCS isosurfaces of ± 0.1 ppm (dark shade) and ± 0.05 ppm (light shade) emanating from the heme iron of Cc. Calculated from the magnetic susceptibility ($\Delta\chi$) tensor of the CcP-bound Cc,[15] the blue and red isosurfaces indicate positive and negative PCSs, respectively. (B) Intermolecular PCSs of the CcP backbone amide protons. Black circles show the experimental PCSs, PCS(obs), while the red line indicates the PCS(calc) values back-calculated from the X-ray structure of the complex using the encounter state population p = 0.4. The inset shows the correlation plot of the observed and calculated PCSs. The strongest outliers – corresponding to the residues W191, A193, A194, N195, N196, V198, and E290, most of which are located in the crystallographic binding interface – are indicated by filled circles both in the inset and the main graph.

molecular frame by bioconjugation techniques. However, the mobility of the attached label can obscure the measured effects,[18] and the large PCSs – exhibited by most lanthanide ions[19] – might contain non-negligible contributions from the minor species, both of which decrease the accuracy and precision of the p estimates. Here, these obstacles were overcome by observing the PCSs arising from the native Cc heme co-factor, firmly embedded in the protein matrix and harboring a low-spin (S = 1/2) Fe^{3+} atom causing small paramagnetic effects.

Comparison of the measured PCSs with those calculated from the known $\Delta\chi$ tensor of Cc[15] for the dominant, crystallographic Cc-CcP orientation[10] reveals that the best match is obtained when p = 0.4 (Fig. 4B). The largest discrepancies are found for the residues located in the binding interface, most likely reflecting subtle differences in the binding of the ferric and ferrous Cc.[15] Independently verified by an exhaustive, cross-validated ensemble refinement against the intermolecular PRE data (see below), the derived p value of 0.4 ± 0.1 is close to p = 0.3 estimated before for the CcP(RS)-Cc^{ox} complex.[6,12]

Cross-validated ensemble refinement. As illustrated in Fig. 3 and evidenced by the large Q factor of 0.75, the crystallographic Cc-CcP structure does not fully account for the experimental PRE data. Furthermore, the PRE-based rigid-body docking of the individual protein molecules yields best solutions with Q of only 0.55, which is not improved in control runs that allow a partial overlap of protein sidechains. Thus, it appears that a single Cc-CcP orientation is not sufficient to explain all Γ_2^{obs} values. To find the optimal solution satisfying the PRE restraints, we used an ensemble refinement

Chemical Science Accepted Manuscript

This journal is © The Royal Society of Chemistry 2014

protocol where multiple Cc copies are docked simultaneously to a single CcP molecule.

The CcP(CN)-Cc^{red} refinement runs at each ensemble size, N, were repeated with p values incremented in 0.1 steps from 0.1 to 0.9. At every (p, N) combination, the Q factor was computed as an average over 50 solutions with the lowest Q values. The resulting Q = f(p, N) plot (Fig. 5A) shows a well-defined minimum at N = 1-3 (p = 0.4) and a number of less distinct minima at higher N values. To verify whether the decrease in the Q factors at increasing N represents a genuine improvement in the fit of the calculated and observed PREs or is simply due to an over-fitting, we performed the complete cross-validation.[20] As explained in the Experimental section, this was achieved by randomly omitting 10 % of the Γ_2^{obs} data and verifying how well these 'free' PREs are predicted by the remaining, 'working' data set (i.e. 90 % included in the refinement), with Q_{free} as a measure of the fit. As can be seen in Q = f(N) plots at p = 0.3 and 0.4 (Fig. 5B), after the initial, steep decrease from N = 0 to N = 1, the Q_{free} values level off at N = 3-5, and then rise sharply at N = 10 and 15. Such increase in Q_{free} with the rise in N indicates over-fitting.[20] Thus, the cross-validation analysis suggests that N = 5, corresponding to the smallest Q value at the Q_{free} plateau (i.e. the smallest Q + Q_{free} sum), is the optimal size of the Cc ensemble required to satisfy the experimental restraints. At N = 5, the minimum value of Q = 0.35 ± 0.03 is found at p = 0.4 (the intersection of the white isolevels in Fig. 5A), confirming the value of the encounter state population obtained by the PCS analysis.

Finally, to assess the performance of the HSQC-based single-time-point PRE measurement scheme employed here for the quantitative structural analysis of the protein encounters, we investigated the impact of the errors in the measured peak intensities on the outcome of the Cc-CcP ensemble calculations Presented in Fig. S7 and the Supplementary Text, the results of the error analysis suggest that the ensemble refinement procedure is relatively insensitive to the specific Γ_2 values used in the calculations and, instead, is driven by the set of intermolecular PREs as a whole.

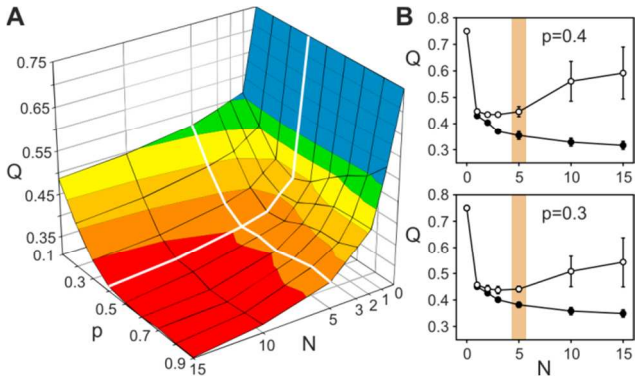


Fig. 5. Analysis of the encounter state population in the Cc-CcP(CN) complex by cross-validated ensemble refinement. (A) The Q = f(p, N) surface plot for the PRE-based ensemble refinement of the encounter complex, colored by the Q factors (the hotter the color, the lower the Q factor). The orthogonal isolevels connect the Q values at the same p (vertical axis) or N (horizontal axis). The intersection of the white isolevels indicates the best solution (p, N) = (0.4, 5) with Q = 0.35 ± 0.03. (B) The Q = f(N) plots for the ensemble calculation runs with p = 0.4 (top panel) and p = 0.3 (bottom panel). The Q and Q_{free} are given by filled and open symbols, respectively. The tan bars indicate the optimal value of N = 5, corresponding to the smallest Q + Q_{free} sum.

Visualizing the transient Cc-CcP(CN) encounters

Ensemble refinement of the encounter state with p = 0.4 against the Γ_2^{obs} yields solutions that agree well with the experiment and feature the PRE profiles accounting for the Γ_2 contributions from the minor species (highlighted in Fig. 3). Presented as an atomic probability density map, the spatial distribution of the protein-protein orientations sampled in the encounter complex is strikingly similar to that obtained in early Brownian dynamics (BD) simulations (Fig. 6 A-B).[22] This agreement appears even more remarkable considering that no intermolecular electrostatic forces – which guided the BD simulations – were used in the ensemble refinement protocol, which relied solely on the experimental PREs and steric properties of the interacting molecules.

However, contrary to the conclusions of the BD work, which imposed the ET reaction criteria to select the successful docking geometries,[22] the encounter complex described here is essentially ET inactive, with only 0.4 % of the ensemble members displaying distances between the redox centers of ≤ 15 Å required for an efficient ET (Fig. 6 C-D).[23] This conclusion is further supported by the control ensemble refinement run that allows partial overlap of protein sidechains (Fig. S8), which suggests that the large ET donor-

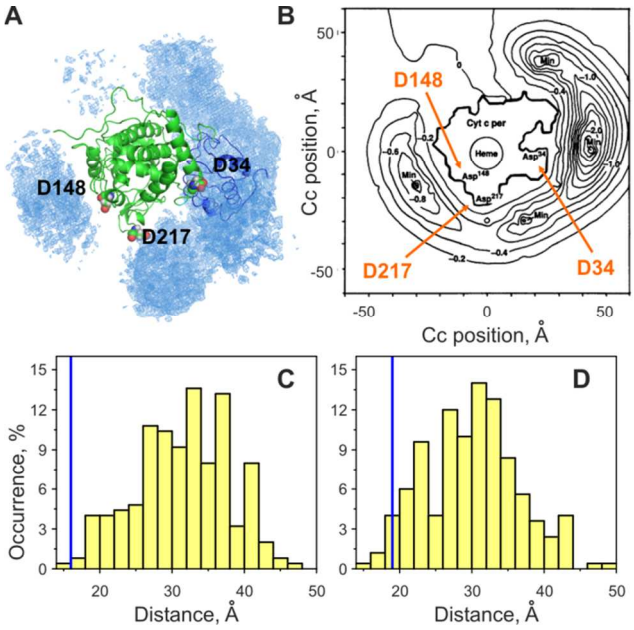


Fig. 6. The CcP(CN)-Cc^{red} encounter complex. (A) Reweighted atomic probability density map, plotted at a threshold of 30 % maximum,[21] for the overall distribution of the Cc molecules obtained from the PRE-based ensemble refinement with p = 0.4. The CcP and Cc in the crystallographic orientation are shown in green and blue cartoons, respectively. The CcP residues D34, D148, and D217 are spacefilled and indicated by the labels. (B) Boltzmann-averaged total electrostatic potential energy of the interaction between CcP and Cc obtained by Brownian dynamics simulations.[22] The contours correspond to different energy levels in units of k_BT, and the axes indicate the postition of the Cc centre of mass. Panel (B) is reproduced with modification from ref.[22] with permission from Science. (C)-(D) Intermolecular ET donor-acceptor distances in the encounter state. The histograms show distributions of the edge-to-edge (C) W191 (CcP) – heme (Cc) and (D) heme-heme distances in the encounter ensemble. Solid vertical lines mark the corresponding ET distances in the crystallographic orientation.

Chemical Science Accepted Manuscript

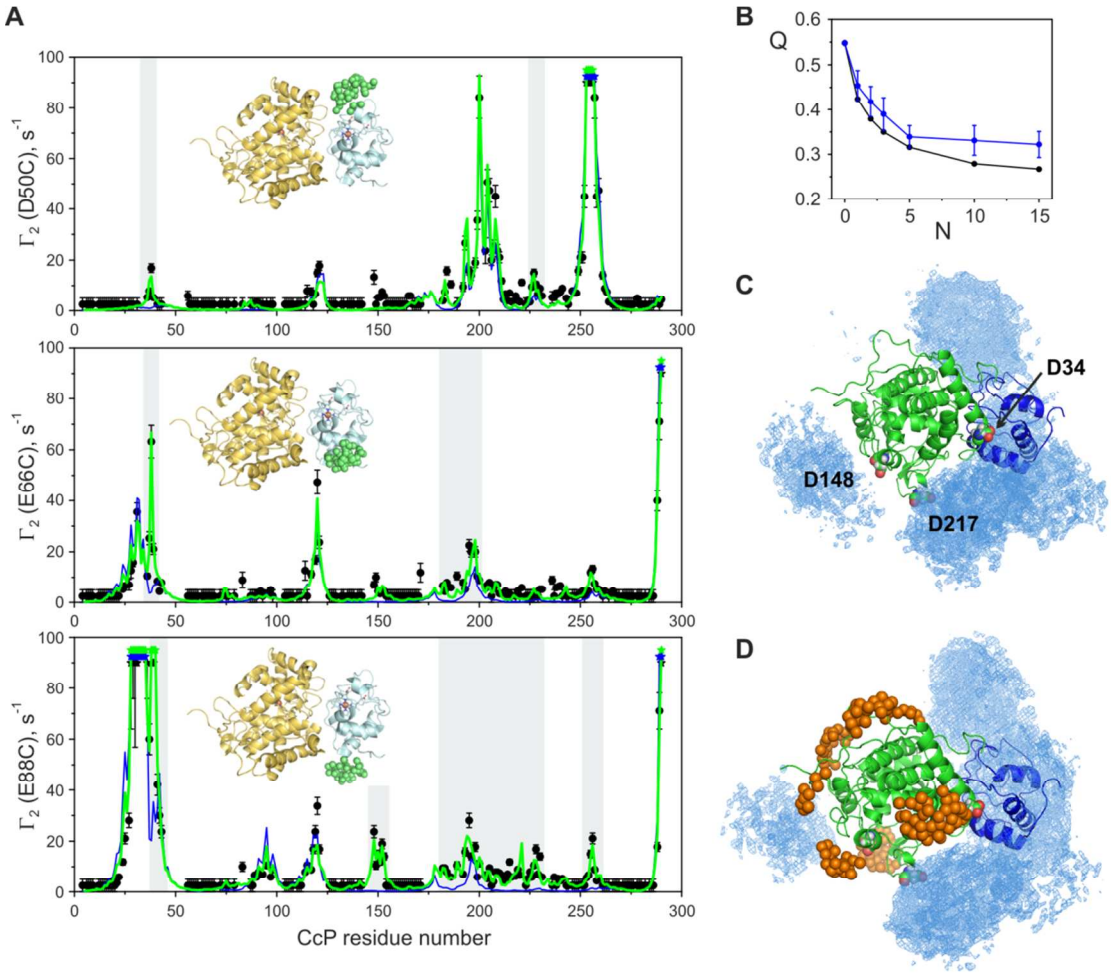


Fig. 7. The CcP(RS)-Cc^{ox} encounter complex. (A) Intermolecular PREs for the CcP(RS) in the complex with Cc paramagnetically labeled at D50C (top), E66C (middle), and E88C (bottom). The plots show measured PREs (Γ_2^{obs}, black symbols), Γ_2 values back-calculated from the crystallographic orientation (blue line), and the PREs calculated for the combination of the specific and the encounter complexes (Γ_2^{calc}, green line), obtained in a typical ensemble refinement run (p = 0.3, Q = 0.29). Stars indicate the values of $\Gamma_2 \geq 90$ s^{-1} for the calculated PREs or identify the residues whose resonances disappear in the paramagnetic spectrum. The errors are standard deviations. The regions exhibiting the PRE contributions from the encounter complex are highlighted. The insets show the X-ray structure of the complex, with CcP in yellow and Cc in blue; the green spheres correspond to the Mn^{2+} atoms of multiple conformers of the attached paramagnetic label used for the ensemble averaging. (B) The Q = f(N) plots for the ensemble calculation runs with p = 0.3. The Q and Q_{free} are given by the black and blue symbols, respectively. (C)-(D) Reweighted atomic probability density maps, plotted at a threshold of 30 % maximum,[21] for the overall distribution of the Cc molecules obtained from the ensemble refinement (p = 0.3, N = 5) against the PRE dataset obtained in this work and (C) used alone or (D) in combination with the PRE data from earlier studies.[11,12] The CcP and Cc in the crystallographic orientation are shown in green and blue cartoons, respectively. The CcP residues D34, D148, and D217 are spacefilled and indicated by the labels in (C). The orange spheres in (D) correspond to the oxygen atoms of multiple conformers of the nitroxide spin labels attached to the CcP at positions V10, K97, T137, N141, N164, L213, and S263 used for the ensemble averaging.

acceptor separations are not an artifact of the rigid-body docking protocol. It appears that, while Cc explores the same CcP surface region as that determined in the BD study, the relative orientations of the interacting molecules differ, which explains the disparity in the ET properties of the protein encounter as shown by molecular simulations and described here. Our findings suggest that, rather than directly contributing to the functional ET activity, the encounter complex ensures electrostatically favorable pre-orientation of the interacting molecules and enables the ensuing reduced dimensionality search of the dominant, ET active protein-protein orientation.

Encounter complexes of Cc and CcP in different oxidation and ligation states

To explore the influence of the protein redox and spin states on the encounter complex formation, we performed PRE analysis of the CcP(CN)-Cc^{ox} and the CcP(RS)-Cc^{ox} systems. Although the latter was studied before,[6,11] due to the paucity of the intermolecular PREs collected in the previous experimental setup, the encounter state could not be fully characterized.[12] Ensemble refinement against the Γ_2^{obs} dataset obtained in this work provided complete description of the conformational space sampled by the interacting molecules and revealed that the distribution of the encounter Cc-CcP orientations closely resembles that of the CcP(CN)-Cc^{red} described above.

Chemical Science Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2014

Similarly to those of the CcP(CN)-Cc^{red} system (Fig. 3), the PRE profiles of CcP(RS)-Cc^{ox} complex cannot be accounted for by a single structure and feature a number of regions with contributions from the encounter state (Fig. 7A). Ensemble refinement against the Γ_2^{obs} data with p = 0.3 determined in the earlier work[6,12] yields solutions that agree well with the experiment and exhibit the PRE profiles that capture the Γ_2 contributions from the minor species (highlighted in Fig. 7A). Complete cross validation of the refinement runs shows that Q_{free} levels off at N = 5 – 15 (Fig. 7B), suggesting that the Cc ensemble size in this range is optimal for the description of the experimental data. The atomic probability density map (drawn for N = 5 ensembles at a threshold of 30 % maximum to enable direct comparison with the map in Fig. 6A) shows that the spatial distribution of the protein-protein orientations sampled in the CcP(RS)-Cc^{ox} encounter complex (Fig. 7C) is very similar to that of the CcP(CN)-Cc^{red} system (Fig. 6A). Furthermore, as evidenced by the results of the ensemble refinement against the combination of the PRE data obtained in this and previous studies,[11,12] the encounter state distribution derived here is consistent with the earlier data. In particular, inclusion of intermolecular restraints from seven spin-labeled CcP variants that showed no PREs on the Cc[11] yields solutions that both satisfy the EDTA(Mn^{2+}) Γ_2^{obs} restraints (Q = 0.31 ± 0.01) and do not violate the nitroxide spin-label data (Fig. 7D).

Thus, it appears that essentially the same regions of the conformational space are sampled by the interacting molecules in the CcP(RS)-Cc^{ox} and CcP(CN)-Cc^{red} encounter complexes. Together with the observation that the PRE profiles of the CcP(CN) complexes with the Cc^{ox} and Cc^{red} are highly similar (Fig. S9), this finding suggests that the properties of the encounter state are dictated by the polypeptide components of the partner proteins rather than redox and spin states of their heme cofactors.

Experimental

Sample preparation

Single-cysteine D50C, E66C, and E88C mutants of the yeast *iso*-1-cytochrome *c* (Cc) were prepared by site-directed mutagenesis of the pUCcc expression vector[24] using whole plasmid synthesis polymerase chain reaction protocol.[25] The uniformly-labeled [D,^{15}N] wild-type (wt) CcP, the wt Cc, and the single-cysteine Cc variants were produced in *Escherichia coli* and purified as described elsewhere.[13,24] For the purification of the latter, 1 mM of dithiothrietol (DTT) was included in all working solutions.

To attach the metal-containing probe, the purified single-cysteine Cc variants were incubated with 10 mM of DTT at room temperature (RT) for 15-30 minutes to reduce the possible intermolecular disulfide bonds; passed through a desalting column (HiTrap, GE Healthcare) equilibrated in 20 mM Tris-HCl pH 8.0 to remove the reducing agent; and incubated overnight at RT with a 5-fold molar excess of the thiol-reactive label N-[S-(2-pyridylthio)cysteaminyl] ethylenediamine-N,N,N',N'-tetraacetate monoamide (Toronto Research Chemicals, catalogue number P996250) and a 3-fold molar excess (relative to the label) of $MnCl_2$ or $ZnCl_2$ (both from Sigma) for the paramagnetic and diamagnetic samples, respectively. The modified proteins were purified by ion-exchange chromatography on a 5 ml pre-packed SP column (GE Healthcare) equilibrated in 20 mM sodium phosphate pH 6.0 and eluted with a linear gradient from 0 to 1 M NaCl in the same buffer. As a rule, the chromatogram recorded at 280 nm showed a single major peak, containing the Cc-EDTA(metal) conjugate. All buffers for the protein labeling and subsequent purification were pre-treated with chelating beads (Chelex, Sigma) to remove traces of heavy metals. As established by the quantitative assay of the free protein thiols with 4,4'-dithiodipyridine[26] and confirmed by mass spectrometry (Fig. S1), the labeling yields of all preparations were close to 100 %. Freshly-prepared Cc-EDTA(metal) conjugates were used throughout. The same labeling protocol was followed to obtain the D32C-EDTA(Mn^{2+}) ubiquitin conjugate.

Mass Spectrometry

The samples were prepared by desalting the purified proteins on a C18 SPE column (Thermo Scientific) and then diluted with the 50 : 50 % (v/v) acetonitrile/water mixture containing 0.1 % formic acid to an approximate concentration of 5 μM. The protein solutions were introduced by an off-line infusion using a capillary electrospray at 1.5 μl/min. An LTQ XL mass spectrometer (Thermo Fisher Scientific) was used to acquire mass spectra from *m/z* 400 to 2000 in centroid mode. Electrospray source conditions such as "source fragmentation" voltage and the tube lens voltage were optimized to help desolvation, but without fragmenting the intact protein. Default values were used for most of the other data acquisition parameters. The resulting spectra were averaged up to 200 scans and deconvoluted with ProMass software (Thermo Fisher Scientific).

NMR Experiments and Data Analysis.

For the PRE experiments, 0.4 mM [D,^{15}N] CcP and 1 molar equivalent of the Cc-EDTA(Mn^{2+}) (for the paramagnetic samples) and Cc-EDTA(Zn^{2+}) or wt Cc (for diamagnetic references) were used. Control experiments were conducted on samples containing 0.4 mM [D,^{15}N] CcP and 1 molar equivalent of the D32C-EDTA(Mn^{2+}) ubiquitin (paramagnetic sample) or wt ubiquitin (the diamagnetic reference). For the PCS measurements, 0.4 mM [D,^{15}N] CcP(CN) and 1 molar equivalent of the wt Cc^{ox} or Cc^{red} were used. The CcP(CN) and Cc^{red} were generated from the CcP(RS) and Cc^{ox} by addition of a 2-fold excess of buffered NaCN and sodium ascorbate solutions, respectively. All NMR samples contained 20 mM sodium phosphate, 100 mM NaCl, pH 6.0 and 6 % D_2O for the lock.

The NMR experiments were conducted at 303 K on a Varian NMR Direct-Drive System 600 MHz spectrometer [control PRE measurements with ubiquitin and 1D ^{1}H titrations of CcP with Cc-EDTA(Mn^{2+})] or an 800 MHz spectrometer equipped with a salt-tolerant PFG-Z cold probe (all other experiments). The NMR data were processed in NMRPipe[27] and analysed in CCPN.[28] The assignments of the backbone amide resonances of [D,^{15}N] CcP(RS) and CcP(CN) were taken from our earlier work.[13,14]

The residue-based average chemical shift perturbations presented in Fig. 1 were calculated as $\Delta\delta_{avg} = (\Delta\delta_N^2/50 + \Delta\delta_H^2/2)^{0.5}$, where $\Delta\delta_N$ and $\Delta\delta_H$ are the chemical shift perturbations of the backbone amide nitrogen and proton, respectively, for a given residues of [D, ^{15}N] CcP(RS) upon binding to the Cc constructs.

Γ_2 PRE Measurements. The ^{1}H Γ_2 PREs were obtained from two identical [^{1}H,^{15}N] transverse relaxation optimized spectroscopy (TROSY)-selected heteronuclear single-quantum correlation (HSQC) experiments run on the paramagnetic and diamagnetic samples. For the reasons given in Supplementary Text and Fig. S2, the HSQC spectra were acquired with the repetition delays of 3 or 10 s, with the former used for qualitative comparison among the different systems studied here and the quantitative analysis of the

Chemical Science Accepted Manuscript

Cc-CcP(RS) complex, while the latter was used for ensemble refinement of the Cc-CcP(CN) encounter state. The ratios of the signal intensities were converted into the Γ_2 values with an in-house script according to Eq. 2:[29]

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2,dia}\exp(-\Gamma_2 t)}{R_{2,dia}+\Gamma_2} \qquad (2),$$

where I_{para} and I_{dia} are the measured resonance intensities (heights) of CcP in complex with paramagnetic and diamagnetic Cc, respectively; the $R_{2,dia}$ is the transverse relaxation rate of the CcP amide protons in the diamagnetic sample (estimated from the width at half-height, $\Delta\nu_{1/2}$, of its Lorentzian fit in the proton dimension as $R_{2,dia}= \pi\Delta\nu_{1/2}$); and t (9 ms) is the total insensitive nuclei enhanced by polarization transfer (INEPT) evolution time of the HSQC.[29]

The resonances showing strong spectral overlap were excluded from the analysis. When applicable, the intensity ratios were rescaled prior to the Γ_2 conversion (see Supplementary Text). To obtain full PREs for the protein complex, the Γ_2 values were divided by the fraction of the Cc bound to CcP at the present experimental conditions, estimated from protein concentrations and the reported binding constants.[15] The uncertainties on the Γ_2 values, $\delta\Gamma_2$, were propagated from the signal intensities and the spectral noise levels as described elsewhere[11] and set to $\delta\Gamma_2 = \max(\delta\Gamma_2, 0.1\Gamma_2)$ to accommodate the experimental errors. For the residues experiencing no PREs (i.e. $I_{para}/I_{dia} > 0.85$), the Γ_2 values were set to 2.5 ± 2.5 s^{-1}. For the backbone amides whose resonances disappear in the paramagnetic spectrum, the upper limit of the I_{para} was estimated from the spectral noise level, and the corresponding I_{para}/I_{dia} intensity ratios were converted into open-ended, lower-limit Γ_2 restraints for the subsequent ensemble refinement.

Alternatively, the PREs were obtained from TROSY-selected four-time-point Γ_2 measurement experiments (run with interleaved relaxation delays of 0, 6, 9, and 12 ms), adapted from the published pulse sequences.[30,31] For each backbone amide, the Γ_2 value was obtained either as a difference in transverse relaxation rates – determined from the fit of the four signal intensities to the exponential decay function – in the paramagnetic ($R_{2,dia} + \Gamma_2$) and the diamagnetic ($R_{2,dia}$) samples or as an average of the Γ_2 values derived from a series of two-time-point data analyses.[31]

Intermolecular Pseudocontact Shifts. The pseudocontact shifts (PCSs) are anisotropic paramagnetic effects given by the Eq. 3:

$$PCS = 1/(12\pi)r^{-3}[\Delta\chi_{ax}(3\cos^2\theta - 1) + 1.5\Delta\chi_{rh}\sin^2\theta\cos(2\varphi)] \qquad (3),$$

where r, θ, and φ are the polar coordinates of the nuclear spin with respect to the principal axes of the magnetic susceptibility tensor ($\Delta\chi$), and $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are, respectively, the axial and rhombic $\Delta\chi$ components defined as $\Delta\chi_{ax}= \chi_{zz}-0.5(\chi_{xx} + \chi_{yy})$ and $\Delta\chi_{rh}= \chi_{xx} - \chi_{yy}$, where χ_{xx}, χ_{yy}, and χ_{zz} are the principal components of the χ tensor.[32] Arising from the Cc heme iron and observed on the CcP backbone amide protons, intermolecular PCSs were calculated as $PCS^{obs} = \Delta\delta_{ox} - \Delta\delta_{red}$, where $\Delta\delta_{ox}$ and $\Delta\delta_{red}$ are the residue-based chemical shift perturbations in the [^{1}H, ^{15}N] HSQC spectra experienced by CcP(CN) upon binding to the Cc^{ox} and Cc^{red}, respectively, extrapolated to the 100 % bound form. The PCS^{calc} values for the dominant, crystallographic Cc-CcP orientation were calculated with Numbat[33] from the $\Delta\chi$ tensor of the CcP-bound Cc[15] and the X-ray structure of the complex.[10] The agreement between the PCS^{obs} and PCS^{calc} was assessed by calculating a Q factor, Eq. 4:

$$Q_{PCS} = \sqrt{\sum_i \left[PCS_i^{obs} - (1-p)PCS_i^{calc}\right]^2 \Big/ \sum_i \left(PCS_i^{obs}\right)^2} \qquad (4),$$

where p is the encounter state population, and the sums run over all backbone amides with measured PCS^{obs}. In practice, the p was varied, and the value yielding the lowest Q_{PCS} (p = 0.4) – indicating the best match between the PCS^{obs} and PCS^{calc} – was selected. The resulting, population-weighted PCS^{calc} profile is shown in Fig. 4B.

1D ^{1}H NMR Titrations. The Cc-observed, reverse titrations (Fig. 1) were performed in 20 mM sodium phosphate, 100 mM NaCl, pH 6.0 at 303 K. A concentrated stock solution (0.5 - 0.8 mM) of the different Cc^{ox}–EDTA(Mn^{2+}) conjugates or the wt Cc (as a control) was incrementally added to the CcP(RS) samples at the initial concentration of 0.1 mM. At each increment, changes in the position of the downfield-shifted resonance of the Cc heme 3-CH_3 group (31.6 ppm in free Cc^{ox}) were monitored in 1D ^{1}H NMR spectra. The titration curves were analyzed with a two-parameter nonlinear least squares fit using a one-site binding model corrected for the dilution effect[34] as given in Eq. 5:

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

$$A = 1 + 1/R + K_D\frac{[Cc]_0 + R[CcP]_0}{R[Cc]_0[CcP]_0} \qquad (5),$$

where $\Delta\delta_{binding}$ is the chemical shift perturbation at a given protein ratio; $\Delta\delta_0$ is the chemical shift perturbation at 100 % Cc bound; R is the [Cc]/[CcP] ratio at a given point; $[CcP]_0$ and $[Cc]_0$ are the concentrations of the starting sample and the titrant stock solution, respectively; and K_D is the equilibrium dissociation constant. Thus, $\Delta\delta_{binding}$ and R are the dependent and independent variables, respectively, and $\Delta\delta_0$ and K_D are the fitted parameters.

Ensemble Refinement against Intermolecular PREs. The transverse paramagnetic relaxation enhancement, Γ_2, is given by the Solomon-Bloembergen equation, Eq. 6:[35,36]

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2\gamma_I^2 g^2\mu_B^2 S(S+1)r^{-6}\left(4\tau_c + \frac{3\tau_c}{1+\omega_h^2\tau_c^2}\right) \qquad (6),$$

where r is the distance between the paramagnetic center and the observed proton, μ_0 is the permeability of vacuum, γ_I is the proton gyromagnetic ratio, g is the electron g-factor, μ_B is the electron Bohr magneton, S is the electron spin number, τ_c is the rotational correlation time, and ω_h is the proton Larmor frequency. The rotational correlation time (τ_c = 6 ns) is defined as $\tau_c = (\tau_r^{-1} + \tau_s^{-1})^{-1}$, where τ_r is the rotational correlation time of the Cc-CcP complex (equal to 16 ns)[6] and τ_s is the effective electron relaxation time, which is 9.6 ns for the EDTA(Mn^{2+}) group.[5]

For the exchange between the major, crystallographic protein-protein orientation and the minor, encounter form, the lower limit of the exchange constant can be estimated as $k_{ex} >> 1/\tau$ = 1,200 s^{-1}, where τ is the lifetime of the Cc-CcP complex.[9] As $k_{ex} >> |\Gamma_2 - \Gamma_2^*|_{max} \approx 100$ s^{-1}, where $|\Gamma_2 - \Gamma_2^*|_{max}$ is the largest difference in the PREs for the major and minor species observed in this work, the major and minor forms of the protein complex are in fast exchange. Thus, for each CcP backbone amide, the observed PRE (Γ_2^{obs}) is the sum of the population-weighted contributions from the specific form (Γ_2) and the encounter state (Γ_2^*), Eq. 1.

The coordinates of Cc-CcP complex were taken from the X-ray structure (PDB 2PCC).[10] Using the Γ_2^{obs} dataset obtained from the three EDTA(Mn^{2+}) conjugation sites (D50C, E66C, and E88C Cc), the rigid-body simulated annealing refinement of the Cc-CcP encounter state was carried out in Xplor-NIH[37,38] following the published procedure.[5] Briefly, the position of the crystallographic

Chemical Science Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2014

Cc-CcP complex was fixed, and multiple copies of Cc molecules, representing ensembles with N = 1, 2, 3, 5, 10, 15, were docked to minimize the energy function consisting of the PRE target term, van der Waals repulsion term to prevent atomic overlap between Cc and CcP, and a weak radius-of-gyration restraint used to encourage intermolecular Cc-CcP contacts.[5] Note that this procedure allows for the atomic overlap among Cc molecules constituting an ensemble. To allow a partial overlap of protein sidechains in control runs, the van der Waals potential for all side-chain atoms extending beyond Cb was set to zero. To account for the mobility of the attached label, the calculated effects were averaged over an ensemble of 50 EDTA(Mn^{2+}) conformers generated by simulated annealing in torsion angle space.[39] As a rule, 100 independent refinement runs were performed, and 50 solutions with lowest Q factors (see below) were selected for further analysis.

To assess the agreement between the observed PREs and the PREs back-calculated from Cc ensembles generated in each run, we calculated the Q factor, Eq. 7:

$$Q = \sqrt{\sum_j \sum_i \left(\Gamma_{2,ij}^{obs} - \Gamma_{2,ij}^{calc}\right)^2 \Big/ \sum_j \sum_i \left(\Gamma_{2,ij}^{obs}\right)^2} \qquad (7),$$

where j = 1 - 3 runs over the three EDTA(Mn^{2+}) attachment sites on the Cc (D50C, E66C, and E88C) and $\Gamma_{2,ij}^{calc}$ is given by Eq. 8:

$$\Gamma_{2,ij}^{calc} = \frac{p}{N} \sum_{k=1}^{N} \Gamma_{2,ijk}^{*} + (1-p)\Gamma_{2,ij} \qquad (8),$$

where p is the total population of the encounter state, N is the size of the encounter ensemble, $\Gamma_{2,ijk}^{*}$ is the PRE from EDTA(Mn^{2+}) (j) back-calculated for the residue (i) of the Cc ensemble member (k), and $\Gamma_{2,ij}$ is the PRE back-calculated from EDTA(Mn^{2+}) (j) for the residue (i) of Cc in the dominant form of the complex.

The complete cross-validation[20] was performed by randomly omitting 10 % of Γ_2^{obs} data and verifying how well these 'free' PREs are predicted by the remaining, 'working' data set (i.e. 90 % included in the refinement), with Q_{free} as a measure of the fit. Typically, three cross-validation runs at each N were performed, each time excluding a different subset of the starting data.

Acknowledgements

K.V.d.W. acknowledges the VUB funding (BAP mandate). A.N.V. is an FWO Post-Doctoral Researcher. We thank Dr. Didier Vertommen (De Duve institute) for the mass spec analysis, Dr. Evgeny Tishchenko (Agilent) for setting up the 4-point PRE measurement experiments, and Dr. Maria-Angeles Ceregido-Perez (VUB) for providing the ubiquitin samples. This work was supported by the VIB and the Flemish Hercules Foundation.

Notes and references

[a] Jean Jeener NMR Centre, Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.
[b] Molecular Recognition Unit, Structural Biology Research Centre, VIB, Pleinlaan 2, 1050 Brussels, Belgium.

Electronic Supplementary Information (ESI) available: deconvoluted mass spectra of the EDTA(Mn^{2+})-conjugated proteins, PRE profiles for the control experiments, the distribution of the intermolecular ET distances in the control ensemble refinement run, discussion of the practical aspects of PRE measurements, and the results of the error analysis of ensemble calculations from the measured PREs. See DOI: 10.1039/b000000x/

1 G. Adam and M. Delbruck, M. in *Structural Chemistry and Molecular Biology*, ed. A. Rich and N. Davidson, Freeman, San Francisco, 1968, pp. 198-215.
2 G. M. Clore, *Mol. Biosyst.*, 2008, **4**, 1058-1069.
3 G. M. Clore, *Prot. Sci.*, 2011, **20**, 229-246.
4 J. Iwahara and G. M. Clore, *Nature*, 2006, **440**, 1227-1230.
5 C. Tang, J. Iwahara and G. M. Clore, *Nature*, 2006, **444**, 383-386.
6 A. N. Volkov, J. A. R. Worrall, E. Holtzmann and M. Ubbink, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 18945-18950.
7 Q. Bashir, S. Scanu and M. Ubbink, *FEBS J.*, 2011, **278**, 1391-1400.
8 J. E. Erman and L. B. Vitello, *Biochim. Biophys. Acta*, 2002, **1597**, 193-220.
9 A. N. Volkov, P. Nicholls and J. A. R. Worrall, *Biochim. Biophys. Acta*, 2011, **1807**, 1482-1503.
10 H. Pelletier and J. Kraut, *Science*, 1992, **258**, 1748-1755.
11 Q. Bashir, A. N. Volkov, G. M. Ullmann and M. Ubbink, *J. Am. Chem. Soc.*, 2010, **132**, 241-247.
12 A. N. Volkov, M. Ubbink and N. A. J. van Nuland, *J. Biomol. NMR*, 2010, **48**, 225-236.
13 A. N. Volkov, A. Wohlkonig, S. H. Soror and N. A. J. van Nuland, *Biochemistry*, 2013, **52**, 2165-2175.
14 S. Vanwetswinkel, N. A. J. van Nuland and A. N. Volkov, *J. Biomol. NMR*, 2013, **57**, 21-26.
15 A. N. Volkov and N. A. J. van Nuland, *J. Biomol. NMR*, 2013, **56**, 255-263.
16 D. Yu, A. N. Volkov and C. Tang, *J. Am. Chem. Soc.*, 2009, **131**, 17291-17297.
17 T. Furia, in *CRC Handbook of Food Additives* , ed. T. Furia, CRC Press, Boca Raton, 1973, pp. 271-294.
18 M. A. S. Hass, P. H. J. Keizers, A. Blok, Y. Hiruma and M. Ubbink, *J. Am. Chem. Soc.*, 2010, **132**, 9952-9953.
19 G. Otting, *J. Biomol. NMR*, 2008, **42**, 1-9.
20 A. T. Brünger, G. M. Clore, A. M. Gronenborn, R. Saffrich and M. Nilges, *Science*, 1993, **261**, 328-331.
21 C. D. Schwieters and G. M. Clore, *J. Biomol. NMR*, 2002, **23**, 221-225.
22 S. H. Northrup, J. O. Boles and J. C. L. Reynolds, *Science*, 1988, **241**, 67-70.
23 C. C. Page, C. C. Moser, X. Chen and P. L. Dutton, *Nature*, 1999, **402**, 47-52.
24 A. N. Volkov, S. Vanwetswinkel, K. Van de Water and N. A. J. van Nuland, *J. Biomol. NMR*, 2012, **52**, 245-256.
25 M. P. Weiner, G. L. Costa, W. Schoettlin, J. Cline, E. Mathur and J. C. Bauer, *Gene*, 1994, **151**, 119-123.
26 C. K. Riener, G. Kada and H. J. Gruber, *Anal. Bioanal. Chem.*, 2002, **373**, 266-276.
27 F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, *J. Biomol. NMR*, 1995, **6**, 277-293.

Chemical Science Accepted Manuscript

28 W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, *Proteins*, 2005, **59**, 687-696.
29 J. L. Battiste and G. Wagner, *Biochemistry*, 2000, **39**, 5355-5365.
30 L. W. Donaldson, N. R. Skrynnikov, W. Y. Choy, D. R. Muhandiram, B. Sarkar, J. D. Forman-Kay and L. E. Kay, *J. Am. Chem. Soc.*, 2001, **123**, 9843-9847.
31 J. Iwahara, C. Tang and G. M. Clore, *J. Magn. Reson.*, 2007, **184**, 185-195.
32 M. Ubbink, J. A. R. Worrall, G. W. Canters, E. J. J. Groenen and M. Huber, *Annu. Rev. Biophys. Biomol. Struct.*, 2002, **31**, 393-422.
33 C. Schmitz, M. J. Stanton-Cook, X.-C. Su, G. Otting and T. Huber, *J. Biomol. NMR*, 2008, **41**, 179-189.
34 A. Kannt, S. Young and D. S. Bendall, *Biochim. Biophys. Acta*, 1996, **1277**, 115-126.
35 I. Solomon, *Phys. Rev.*, 1955, **99**, 559-565.
36 I. Solomon and N. Bloembergen, *J. Chem. Phys.*, 1956, **25**, 261-266.
37 C. D. Schwieters, J. J. Kuszewski, N. Tjandra and G. M. Clore, *J. Magn. Reson.*, 2003, **160**, 66-74.
38 C. D. Schwieters, J. J. Kuszewski and G. M. Clore, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2006, **48**, 47-62.
39 J. Iwahara, C. D. Schwieters and G. M. Clore, *J. Am. Chem. Soc.*, 2004, **126**, 5879-5896.

Chemical Science Accepted Manuscript

 This journal is © The Royal Society of Chemistry 2014