J|A|C|S ARTICLES

Published on Web 07/30/2010

Shifting the Equilibrium between the Encounter State and the Specific Form of a Protein Complex by Interfacial Point Mutations

Alexander N. Volkov,[†,§] Qamar Bashir,[†] Jonathan A. R. Worrall,[†,||] G. Matthias Ullmann,[‡] and Marcellus Ubbink*[,†]

Leiden Institute of Chemistry, Leiden University, Gorlaeus Laboratories, P.O. Box 9502, 2300 RA, Leiden, The Netherlands, and Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstr. 30, BGI, 95447 Bayreuth, Germany

Received February 1, 2010; E-mail: m.ubbink@chem.leidenuniv.nl

Abstract: Recent experimental studies have confirmed a long-held view that protein complex formation proceeds via a short-lived encounter state. The population of this transient intermediate, stabilized mainly by long-range electrostatic interactions, varies among different complexes. Here we show that the occupancy of the encounter state can be modulated across a broad range by single point mutations of interfacial residues. Using a combination of Monte Carlo simulations and paramagnetic relaxation enhancement NMR spectroscopy, we illustrate that it is possible to both enhance and diminish the binding specificity in an electron transfer complex of yeast cytochrome *c* (Cc) and cytochrome *c* peroxidase. The Cc T12A mutation decreases the population of the encounter to 10% as compared with 30% in the wild-type complex. More dramatically, the Cc R13A substitution reverses the relative occupancies of the stereospecific and the encounter forms, with the latter now being the dominant species with the population of 80%. This finding indicates that the encounter state can make a large contribution to the stability of a protein complex. Also, it appears that by adjusting the amount of the encounter through a judicious choice of point mutations, we can remodel the energy landscape of a protein complex and tune its binding specificity.

Introduction

Cellular function depends critically on the ability to transmit, receive, and integrate various signals arising within and outside the cell. In order to propagate a biochemical event such as phosphorylation or electron transfer (ET) to distant cellular sites rapidly and efficiently, protein complexes mediating relay processes are optimized for high turnover. As a consequence, such transient interactions are short-lived and typically have high association (k_{on}) and dissociation (k_{off}) rate constants, resulting in weak binding ($K_D = k_{off}/k_{on}$ in the μM−mM range).[1,2] Another constraint is imposed by the necessity to recognize multiple binding partners, which limits the specificity. Thus, transient complexes have evolved to function within narrow margins of affinity and specificity in order to maximize productive binding and minimize wasteful interactions with other components of the cellular milieu.

A good example of a biomolecular transient system is the ET protein complex of yeast cytochrome *c* (Cc) and cytochrome *c* peroxidase (CcP). The proteins are localized in the intermembrane space of yeast mitochondria, where CcP catalyzes the reduction of peroxides using the electrons donated by Cc, an important process mitigating the oxidative stress.[3] The catalytic cycle of CcP begins with the coordination of peroxide to the ferric heme. Then, bound peroxide undergoes rapid heterolytic cleavage, producing a molecule of water and an intermediate referred to as compound ES.[4] In this state, the heme iron is oxidized to an oxyferryl Fe(IV) group and an active-site tryptophan residue (W191) is oxidized to a cationic indole radical.[5−10] Then, in two successive one-electron transfer reactions, Cc reduces compound ES back to resting-state CcP.[11,12]

The stoichiometry of the Cc−CcP complex has been a subject of debate. The crystal[13] and solution[14] structures of the Cc−CcP complex show a single 1:1 binding site. However, several studies proposed that the interaction of Cc and CcP occurs with a 2:l stoichiometry, involving nonoverlapping binding domains.[15−21] Yet other studies led to the conclusion that only a single binding site exists,[22−24] and recent mutagenesis and cross-linking work by Erman and co-workers strongly supports this conclusion.[25−27]

Paramagnetic NMR experiments revealed a significant degree of dynamics within the Cc−CcP complex[14] and established that the encounter state, which precedes the formation of the specific

[†] Leiden University.
[‡] University of Bayreuth.
[§] Current address: Jean Jeener NMR Centre, Structural Biology Brussels, Vrije Universiteit Brussel and Department of Molecular and Cellular Interactions, VIB, Pleinlaan 2, 1050 Brussels, Belgium.
[||] Current address: Department of Biological Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, U.K.

(1) Crowley, P. B.; Ubbink, M. *Acc. Chem. Res.* **2003**, *36*, 723–730.
(2) Schreiber, G.; Haran, G.; Zhou, H.-X. *Chem. Rev.* **2009**, *109*, 839–860.
(3) Chance, B.; Devault, D.; Legallais, V.; Mela, L.; Yonetani, T. Kinetics of electron transfer reactions in biological systems. In *Fast reactions and primary processes in chemical reactions*; Claesson, S., Ed.; Interscience: New York, 1967; pp 437−464.
(4) Yonetani, T.; Schleyer, H.; Ehrenberg, A. *J. Biol. Chem.* **1966**, *241*, 3240–3243.

binding geometry,[28] represents 30% of the complex in solution.[29] The encounter state, believed to accelerate the search of the optimal binding geometry, is best represented by an ensemble of binary complexes, with the proteins assuming many different orientations. Recently, we have shown that this ensemble can be described well assuming only electrostatic interactions,[29] in agreement with earlier theoretical docking simulations.[30,31] Thus, our prior work has established the existence of minor species in the Cc−CcP complex[14] and contributed to the methodological foundations for PRE NMR analysis of transient encounter states in protein−protein interactions.[29]

The delicate balance between the encounter state and the specific complex may well be functionally relevant. An electrostatically driven encounter enhances the association rate but exhibits low ET activity due to the large distance between the redox centers. The specific complex brings these centers within the range of rapid ET,[29] yet it must not be very stable to avoid a slow turnover. Thus, we wondered whether the balance between an encounter and the specific complex can be modulated. A vast body of literature investigating the impact of interfacial mutations on protein−protein association reactions addresses the equilibrium between the free and bound forms, but not the transition between the encounter state and the specific binding geometry. Here we show that it is possible to both decrease and increase the encounter population by single point mutations of Cc residues in the binding interface, which suggests a subtle yet efficient way to remodel the energy landscape of transient protein−protein interactions. To the best of our knowledge, the equilibrium between the encounter and specific state, let alone its perturbation by surface mutations, has never been quantified experimentally, which is the focus of this study.

Methods

Sample Preparation. Both unlabeled and U-^{15}N labeled T-5A/C102T Cc, referred to as wt Cc, were expressed in *Escherichia coli* and purified as described before.[32,33] Cc mutations were introduced by site-directed mutagenesis using the Quik Change polymerase chain reaction protocol (Stratagene) with wt plasmid as a template. The constructs were verified by DNA sequencing, and the proteins produced and purified analogously to wt Cc. Wild-type CcP and its single-cysteine variants used in the PRE analysis were isolated from *E. coli* following the published procedures.[14,34] CcP mutants were conjugated, one at a time, with a paramagnetic spin-radical [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] as described.[14] The yield of labeling, determined by EPR or a spectrophotometric dithiodipyridine assay,[35] was close to 100%. Both Cc and CcP were in the ferric, oxidized form in all experiments.

Isothermal Titration Calorimetry. The experiments were performed on a Microcal VP-ITC calorimeter in 20 mM sodium phosphate, 0.1 M NaCl pH 6.0 at 303 K. A T12A Cc solution (0.3 mM) was placed in the syringe, rotated at 310 rpm during the experiment, and titrated into the sample cell containing 0.03 mM CcP, with an injection volume of 4 μL for the first and 10 μL for all subsequent titration points, with a 60 s initial equilibration delay and a 360 s delay between the injections. The binding isotherms were analyzed with Microcal Origin 5.0. In each case the first data point was discarded and the baseline was adjusted manually. The integrated data were corrected for the heat of Cc dilution, either determined by titrating Cc into the buffer or estimated from the last data points of a complete titration by averaging their integrated areas.

NMR Spectroscopy. PRE NMR samples contained 0.3−0.4 mM wt or mutant ^{15}N Cc and 1 molar equiv of a single-cysteine CcP variant labeled with a paramagnetic spin-radical (MTSL) or its diamagnetic analogue (MTS). For chemical shift perturbation analysis, samples containing 0.5 mM of ^{15}N Cc mutants or their 1:1 complexes with wt CcP were used. Measurements were performed in 20 mM sodium phosphate, 0.1 M NaCl pH 6.0, 6% D_2O for lock, and 0.1 mM each of TSP [3-(trimethylsilyl)-propionic acid-D4, sodium salt] and $CH_3CO^{15}NH_2$ as internal references at 303 K on a Bruker DMX-600 spectrometer equipped with a cryogenic probe. The pH of the sample was checked after each titration point and, if necessary, adjusted to 6.00 ± 0.05 with concentrated NaOH or HCl solutions. Two-dimensional [^{15}N, ^{1}H] HSQC spectra were acquired with 1024 and 128 complex points in the direct and indirect dimensions, respectively. All data were processed with Azara 2.7 (provided by W. Boucher and Department

(5) Mauro, J. M.; Fishel, L. A.; Hazzard, J. T.; Meyer, T. E.; Tollin, G.; Cusanovich, M. A.; Kraut, J. *Biochemistry* **1988**, *27*, 6243–6256.
(6) Scholes, C. P.; Liu, Y. J.; Fishel, L. A.; Farnum, M. F.; Mauro, J. M.; Kraut, J. *J. Isr. Chem.* **1989**, *29*, 85–92.
(7) Sivaraja, M.; Goodin, D. B.; Smith, M.; Hoffman, B. M. *Science* **1989**, *245*, 738–740.
(8) Miller, M. A.; Han, G. W.; Kraut, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11118–11122.
(9) Fitzgerald, M. M.; Churchill, M. J.; McRee, D. E.; Goodin, D. B. *Biochemistry* **1994**, *33*, 3807–3818.
(10) Huyett, J. E.; Doan, P. E.; Gurbiel, R.; Houseman, A. L. P.; Sivaraja, M.; Goodin, D. B.; Hoffman, B. M. *J. Am. Chem. Soc.* **1995**, *117*, 9033–9041.
(11) Hahm, S.; Miller, M. A.; Geren, L.; Kraut, J.; Durham, B.; Millett, F. *Biochemistry* **1994**, *33*, 1473–1480.
(12) Liu, R. Q.; Miller, M. A.; Han, G. W.; Hahm, S.; Geren, L.; Hibdon, S.; Kraut, J.; Durham, B.; Millett, F. *Biochemistry* **1994**, *33*, 8678–8685.
(13) Pelletier, H.; Kraut, J. *Science* **1992**, *258*, 1748–1755.
(14) Volkov, A. N.; Worrall, J. A. R.; Holtzmann, E.; Ubbink, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18945–18950.
(15) Kang, C. H.; Ferguson-Miller, S.; Margoliash, E. *J. Biol. Chem.* **1977**, *252*, 919–926.
(16) Kornblatt, J. A.; English, A. M. *Eur. J. Biochem.* **1986**, *155*, 505–511.
(17) Matthis, A. L.; Vitello, L. B.; Erman, J. E. *Biochemistry* **1995**, *34*, 9991–9999.
(18) Miller, M. A.; Geren, L.; Han, G. W.; Saunders, A.; Beasley, J.; Pielak, G. J.; Durham, B.; Millett, F.; Kraut, J. *Biochemistry* **1996**, *35*, 667–673.
(19) Stemp, E. D. A.; Hoffman, B. M. *Biochemistry* **1993**, *32*, 10848–10865.
(20) Zhou, J. S.; Hoffman, B. M. *J. Am. Chem. Soc.* **1993**, *115*, 11008–11009.
(21) Zhou, J. S.; Hoffman, B. M. *Science* **1994**, *265*, 1693–1696.
(22) Erman, J. E.; Vitello, L. B. *J. Biol. Chem.* **1980**, *255*, 6224–6227.
(23) Dowe, R. J.; Vitello, L. B.; Erman, J. E. *Arch. Biochem. Biophys.* **1984**, *232*, 566–573.
(24) Vitello, L. B.; Erman, J. E. *Arch. Biochem. Biophys.* **1987**, *258*, 621–629.
(25) Nakani, S.; Viriyakul, T.; Mitchell, R.; Vitello, L. B.; Erman, J. E. *Biochemistry* **2006**, *45*, 9887–9893.
(26) Pearl, N. M.; Jacobson, T.; Arisa, M.; Vitello, L. B.; Erman, J. E. *Biochemistry* **2007**, *46*, 8263–8272.
(27) Pearl, N. M.; Jacobson, T.; Meyen, C.; Clementz, A. G.; Ok, E. Y.; Choi, E.; Wilson, K.; Vitello, L. B.; Erman, J. E. *Biochemistry* **2008**, *47*, 2766–2775.
(28) Ubbink, M. *FEBS Lett.* **2009**, *583*, 1060–1066.
(29) Bashir, Q.; Volkov, A. N.; Ullmann, M.; Ubbink, M. *J. Am. Chem. Soc.* **2010**, *132*, 241–247.
(30) Northrup, S. H.; Boles, J. O.; Reynolds, J. C. L. *Science* **1988**, *241*, 67–70.
(31) Gabdoulline, R. R.; Wade, R. C. *J. Mol. Biol.* **2001**, *306*, 1139–1155.
(32) Pollock, W. B. R.; Rosell, F. I.; Twitchett, M. B.; Dumont, M. E.; Mauk, A. G. *Biochemistry* **1998**, *37*, 6124–6131.
(33) Morar, A. S.; Kakouras, D.; Young, G. B.; Boyd, J.; Pielak, G. J. *J. Biol. Inorg. Chem.* **1999**, *4*, 220–222.
(34) Goodin, D. B.; Davidson, M. G.; Roe, J. A.; Mauk, A. G.; Smith, M. *Biochemistry* **1991**, *30*, 4953–4962.
(35) Riener, C. K.; Kada, G.; Gruber, H. J. *Anal. Bioanal. Chem.* **2002**, *373*, 266–276.

Table 1. Binding Constants of Cc−CcP Complexes[a]

Cc	K_B, 10^5 M^{-1}
wt	1.9 ± 0.1[b]
T12A	17 ± 1
R13K	0.54 ± 0.03
R13A	0.06 ± 0.002[b]

[a] All values were determined in 20 mM sodium phosphate 0.1 M NaCl pH 6.0 at 303 K. [b] Values taken from ref 51.

of Biochemistry, University of Cambridge; available from http://www.bio.cam.ac.uk/azara/) and analyzed in Ansig for Windows.[36] Assignments of the ^{15}N and ^{1}H nuclei of wt Cc were taken from a previous work.[37] The assignments of backbone amide resonances of ^{15}N Cc mutants were verified by 3D NOESY-HSQC and 3D TOCSY-HSQC experiments.[38−40] The averaged amide chemical shift perturbations ($\Delta\delta_{avg}$) were derived from $\Delta\delta_{avg} = [(\Delta\delta_N)^2/50 + (\Delta\delta_H)^2/2]^{1/2}$, where $\Delta\delta_N$ and $\Delta\delta_H$ are the differences in the chemical shifts between the free and bound form for the amide nitrogen and proton, respectively. In order to obtain $\Delta\delta_{avg}$ extrapolated to the 100% bound form, the original $\Delta\delta_{avg}$ was divided by the molar fraction of Cc bound, which, at the protein concentrations used, was 0.9 (wt), 0.97 (T12A), 0.82 (R13K), or 0.57 (R13A). Binding constants from Table 1 were used to calculate the amount of Cc bound.

PRE Measurements. PREs were calculated from peak intensities in the HSQC spectra of paramagnetic (I_{para}) and diamagnetic control (I_{dia}) samples as

$$\frac{I_{para}}{I_{dia}} = \frac{R_2 e^{-t\Gamma_2}}{R_2 + \Gamma_2} \tag{1}$$

where R_2 is the transverse relaxation rate of Cc amide protons in the diamagnetic complex, Γ_2 is the PRE, and t is the total INEPT evolution time of the HSQC.[41] To account for minor differences in protein concentrations between the two samples, the obtained I_{para}/I_{dia} values were normalized as described before.[29] For the resonances that 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 $R_2 = \pi\Delta\upsilon_{1/2}$.

PRE Analysis. The observed PRE is the sum of the population-weighted contributions of the specific form and the encounter state:

$$\Gamma_{2,i}^{obs} = p\Gamma_{2,i}^{encounter} + (1 - p)\Gamma_{2,i}^{specific} \tag{2}$$

The $\Gamma_{2,i}^{encounter}$ and $\Gamma_{2,i}^{specific}$ were back-calculated from the simulated encounter state and the crystal structure of the complex, respectively, following a published protocol using a four-conformer ensemble representation of the attached SL to account for its flexibility.[14,29,42] To obtain the population of the encounter state (p) in a given complex, the agreement between the observed and the calculated PREs at varying p values was estimated using the Q-factor[29] (eq 3), and p values at the lowest Q were selected.

$$Q = \sqrt{\frac{\Sigma_i(\Gamma_{2,i}^{obs} - \Gamma_{2,i}^{calc})^2}{\Sigma_i(\Gamma_{2,i}^{obs} + \Gamma_{2,i}^{calc})^2}} \tag{3}$$

Monte Carlo Simulations. A Boltzmann ensemble of encounter complex geometries was generated by Metropolis Monte Carlo simulations.[29,43] Protein coordinates were taken from the crystal structure of the complex (PDB entry 2PCC[13]), and protein molecules were kept rigid during the simulations. Cc was allowed to move in the electrostatic potential of CcP, and moves were accepted or rejected according to the Metropolis criterion. The energy of each protein−protein orientation was calculated by multiplying the charges of Cc with the electrostatic potential of CcP. An exclusion grid was used to prevent steric overlap between the proteins. When the centers of mass of the two molecules were within 40 Å, the protein coordinates were saved every 1000 steps. In total, 1500 to 2000 configurations were retained for the analysis. In this approach, protein−protein orientations with large intermolecular distances contribute negligibly to the free energy change of protein complex formation and will also be invisible in the PRE analysis.

Mutations were introduced *in silico*, followed by the addition of hydrogen atoms and energy minimization in Xplor-NIH.[44] The simulation temperature and the ionic strength were set to 303 K and 0.12 M, respectively, to match the experimental conditions. The Poisson−Boltzmann electrostatic potential was calculated by the MEAD program suite[45] using atomic partial charges of the Charmm force field.[46] The dielectric constants used for the protein and water were 4 and 80, respectively.

Analysis of the Encounter State. For each CcP conjugation site, the orientation of the attached SL was randomized in Xplor-NIH and an ensemble of four representative conformers was used for subsequent data analysis. For every structure of the simulated encounter ensemble, distances between Cc backbone amide protons and the oxygen atom in each of the four SL conformers were measured and averaged as described before.[29] The effective distances, r_i, were used to calculate PREs for Cc atoms in the encounter state ($\Gamma_{2,i}^{encounter}$):

$$\Gamma_{2,i}^{encounter} = \frac{\gamma^2 g^2 \beta^2}{20 r_i^6}\left(4\tau_c + \frac{3\tau_c}{1 + \omega_h^2\tau_c^2}\right) \tag{4}$$

where τ_c is the correlation time of the electron−nucleus vector (16 ns);[14] ω_h and γ are the proton Larmor frequency and gyromagnetic ratio, respectively; g is the electronic g-factor; and β is the Bohr magneton. The total calculated PRE ($\Gamma_{2,i}^{calc}$) is computed as a sum of the population-weighted $\Gamma_{2,i}^{encounter}$ and $\Gamma_{2,i}^{specific}$ (eq 2), with the latter back-calculated from the crystal structure of the complex. Finally, an agreement between observed and calculated PREs at each p value is quantified (eq 3). The uncertainties in the Q factors were obtained by standard propagation of the error in $\Gamma_{2,i}^{obs}$, derived from the experimental uncertainties in I_{para} and I_{dia}.[29] For all Cc mutants, only the data from SLs that produced intermolecular paramagnetic effects (N38C, N200C, and T288C) were included in the analysis. As judged from the calculated Q factors, control MC runs employing proteins with randomized side-chain orientations or different sets of SL conformers produced results very similar to those presently described for all mutant complexes (data not shown). All structures were visualized with PyMOL.[47]

Structure of the Specific Form of T12A Cc−CcP Complex. The dominant T12A Cc−CcP structure in solution was determined using the PRE NMR analysis protocol for the wt complex.[14] Briefly,

(36) Helgstrand, M.; Kraulis, P. J.; Allard, P.; Hard, T. *J. Biol. NMR* **2000**, *18*, 329–336.
(37) Worrall, J. A. R.; Kolczak, U.; Canters, G. W.; Ubbink, M. *Biochemistry* **2001**, *40*, 7069–7076.
(38) Marion, D.; Kay, L. E.; Sparks, S. W.; Torchia, D. A.; Bax, A. *J. Am. Chem. Soc.* **1989**, *111*, 1515–1517.
(39) Zuiderweg, E. R. P.; Fesik, S. W. *Biochemistry* **1989**, *28*, 2387–2391.
(40) Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 6150–6156.
(41) Battiste, J. L.; Wagner, G. *Biochemistry* **2000**, *39*, 5355–5365.
(42) Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879–5896.
(43) Ullmann, G. M.; Knapp, E. W.; Kostic, N. M. *J. Am. Chem. Soc.* **1997**, *119*, 42–52.
(44) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 66–74.
(45) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environments*; Springer: Berlin, 1997; pp 233−240.
(46) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(47) http://www.pymol.org.

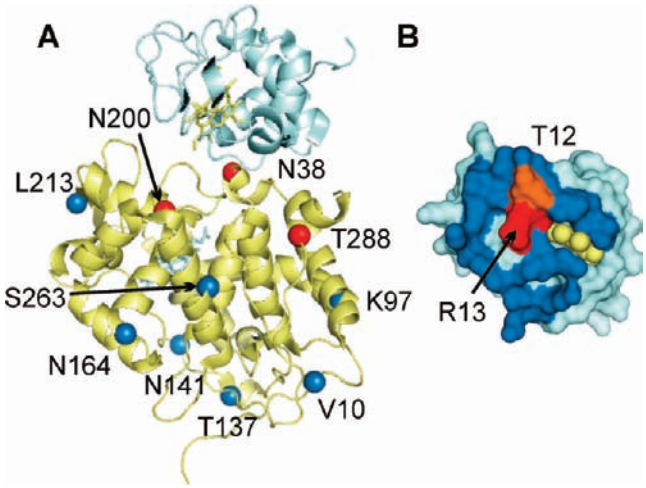

Figure 1. Specific form of Cc−CcP complex. (A) Crystallographic Cc−CcP orientation. Cc and CcP are in blue and yellow, respectively, with heme groups in sticks. C_α atoms of CcP residues used for spin-labeling are shown as spheres colored according to whether the attached SL exhibits intermolecular PREs (red) or not (blue). (B) Binding interface of Cc in the specific complex. The residues that lose solvent accessible surface area upon complex formation with CcP are dark-colored; heme group is space filled and shown in yellow; residues T12 and R13 are colored orange and red, respectively. The figure was drawn from the crystal structure of the complex (PDB entry 2PCC[13]).

for each observed Cc backbone amide proton, an intermolecular PRE was calculated from eq 1. The PREs were converted into distances using

$$r = \sqrt[6]{\frac{\gamma^2 g^2 \beta^2}{20\Gamma_2}\left(4\tau_c + \frac{3\tau_c}{1+\omega_h^2\tau_c^2}\right)} \qquad (5)$$

The distances obtained from eq 5 were used as restraints in the subsequent structure calculations, which were carried out following the procedure used for the wt complex.[14] Protein coordinates were taken from the Protein Data Bank (accession code 2PCC[13]). The T12A mutation was introduced *in silico*, followed by the addition of hydrogen atoms and energy minimization in Xplor-NIH. A set of distance restraints from four MTSL labels (placed at N38C, N200C, T288C, and S263C CcP) was used to drive the docking of the protein complex. For all spin-labeled CcP mutants, chemical shift perturbations in the complex with ^{15}N T12A Cc are equal to those in the ^{15}N T12A Cc−wt CcP complex at the same protein ratio (data not shown), indicating that an attached spin label does not alter the binding.

Results and Discussion

Characterization of Variant Cc−CcP Complexes. The structure of the specific Cc−CcP complex is shown in Figure 1A, with the binding surface of Cc detailed in Figure 1B. Cc residues T12 and R13, located in the center of the interface, mediate interactions with a variety of reaction partners,[48−50] and the latter is part of a binding hot spot in the complex with CcP.[51] Given the importance of these residues in macromolecular recognition, we sought to investigate their role in the specificity of Cc−CcP interaction. To this end, alanine-substituted (T12A and R13A) and charge-conserved (R13K) Cc mutants were prepared. To ascertain that the introduced mutations do not perturb the structure of the protein, we compared the NMR chemical shifts of Cc variants with those of the wt protein (Figure S1 in the Supporting Information). Except for several residues adjacent to the mutation sites, the chemical shifts of the variants are identical to those of wt Cc, indicating that the mutations do not perturb the global fold of the protein. R13K Cc exhibits the smallest $\Delta\delta$ (Figure S1C), which is expected given the similarity of lysine and arginine side chains.

The binding constants (K_B) of the wt and mutant complexes were determined by NMR or ITC titrations or taken from literature (Table 1 and Figure 2). In all cases, titration curves

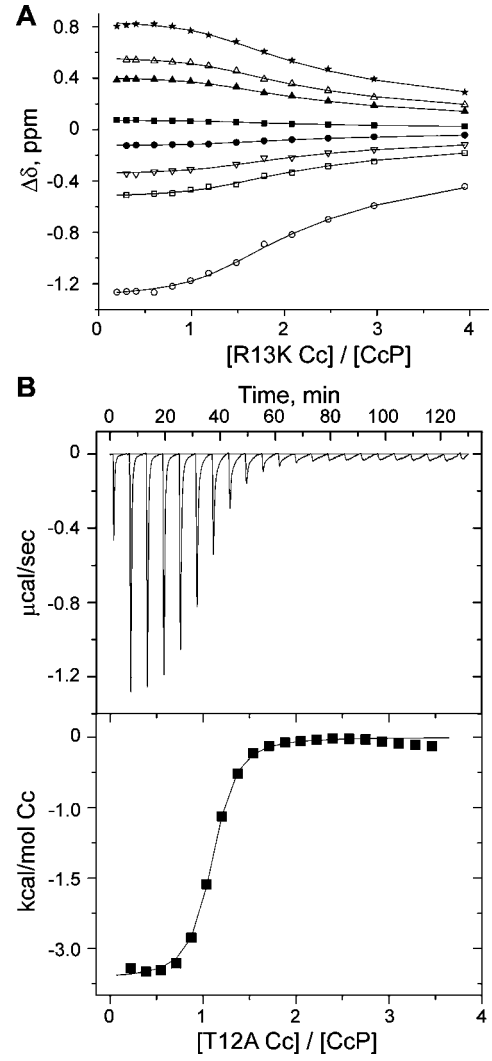

Figure 2. Binding constants. (A) Chemical shift perturbations ($\Delta\delta$) of ^{15}N R13K Cc backbone amide protons (open symbols) and nitrogen atoms (filled symbols) observed in a series of 2D [^{1}H, ^{15}N] HSQC spectra for residues L9 (squares), T12 (circles), C14 (inverted triangles), and Q16 (triangles) upon titration into wt CcP. Also plotted is $\Delta\delta$ for the Cc haem 3-methyl protons (stars), followed in 1D ^{1}H spectra. The curves represent the best fit to a 1:1 binding model[51] with the shared binding constant of $(5.4 \pm 0.3) \times 10^4$ M^{-1}. (B) Isothermal titration calorimetry of the T12A Cc−CcP complex. The top and bottom panels show, respectively, the raw data after the baseline correction and the integrated data corrected for the heat of dilution of Cc. The solid line in the bottom plot is the best fit to the 1:1 binding model.[52] The average values of the stoichiometry $n = 0.97 \pm 0.03$, equilibrium binding constant $K_B = (1.7 \pm 1) \times 10^6$ M^{-1}, and the binding enthalpy $\Delta H_B = -5.5 \pm 0.2$ kcal/mol were obtained from three separate titrations; the errors are standard deviations.

(48) Worrall, J. A. R.; Reinle, W.; Bernhardt, R.; Ubbink, M. *Biochemistry* **2003**, *42*, 7068–7076.
(49) Volkov, A. N.; Ferrari, D.; Worrall, J. A. R.; Bonvin, A. M. J. J.; Ubbink, M. *Protein Sci.* **2005**, *14*, 799–811.
(50) Solmaz, S. R. N.; Hunte, C. *J. Biol. Chem.* **2008**, *283*, 17542–17549.
(51) Volkov, A. N.; Bashir, Q.; Worrall, J. A. R.; Ubbink, M. *J. Mol. Biol.* **2009**, *385*, 1003–1013.

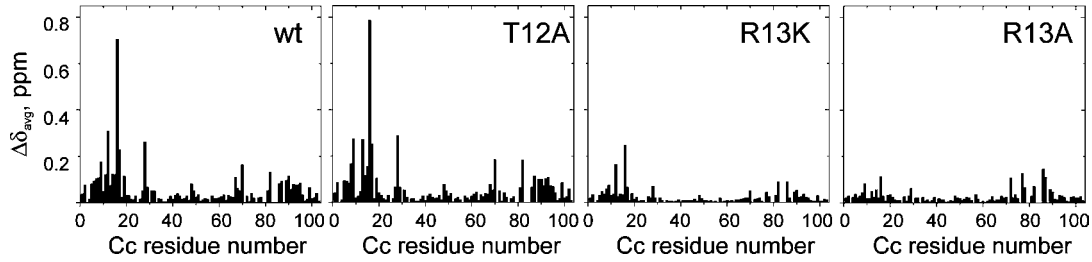

Figure 3. NMR chemical shift perturbations ($\Delta\delta_{avg}$) for the ^{15}N Cc−CcP complexes studied in this work. All $\Delta\delta_{avg}$ are extrapolated to the 100% bound form (see Methods) and thus do not reflect differences in the binding constants.

fit well to a 1:1 stoichiometry model, suggesting that no binding of a second Cc molecule takes place under our experimental conditions. As can be seen from Table 1, Cc R13K substitution leads to a 4-fold decrease in K_B, whereas R13A mutation reduces it 100-fold. Interestingly, a 10-fold increase in K_B is observed in the T12A Cc−CcP complex.

Chemical shift perturbations of Cc residues in variant and wt complexes with CcP exhibit similar patterns, indicating that mutant and wt Cc proteins employ the same interface to bind CcP (Figure 3). However, the magnitude of the binding shifts (extrapolated to 100% Cc bound; see Methods) differs, decreasing in the order T12A > wt > R13K > R13A Cc. The absolute size of chemical shift perturbations has been reported to correlate with the dynamics in transient protein complexes, with smaller shifts indicating increased mobility.[48,49,53] Thus, it appears that R13A and R13K Cc−CcP are more dynamic than the wt complex, while the protein mobility in T12A Cc−CcP is nearly unaffected.

Given the increased dynamics and the weaker binding in R13A(K) Cc−CcP complexes, solving their structures by either X-ray crystallography or solution NMR spectroscopy presents a difficult challenge. Attempts to determine the dominant protein orientations in these complexes using PRE NMR spectroscopy—an approach employed previously to elucidate the solution structure of the wt complex[14]—were unsuccessful, resulting in poor convergence. However, we were able to determine the backbone-resolution solution structure of the dominant form of the T12A Cc−CcP complex (Figure 4). Using a set of intermolecular distance restraints from the SLs placed at several positions on CcP (see Methods), multiple structure calculation runs consistently produced a single cluster of solutions. Overlay of the best (lowest energy) structures of T12A Cc−CcP (deposited in the Protein Data Bank under entry 2JTI) and wt Cc−CcP (PDB entry 2GB8[14]) shows that the relative protein orientations are very similar in both complexes (Figure 4). The positional backbone rmsd between the two cytochromes after superposition of CcP molecules is 1.6 Å.

Intermolecular PREs. In order to study the variant Cc−CcP complexes further, 10 single-cysteine CcP mutants were produced and conjugated with a paramagnetic spin label (SL) (Figure 1A). The unpaired electron of CcP-SL causes an intermolecular PRE of the nearby Cc nuclei, which depends on the distance between the unpaired electron and the observed nucleus and is readily quantified by NMR. For every Cc variant, multiple PRE profiles were measured in binary complexes with different CcP-SL (Figures 5 and S2−S4). The protein interactions studied here are in fast exchange on the NMR time scale;

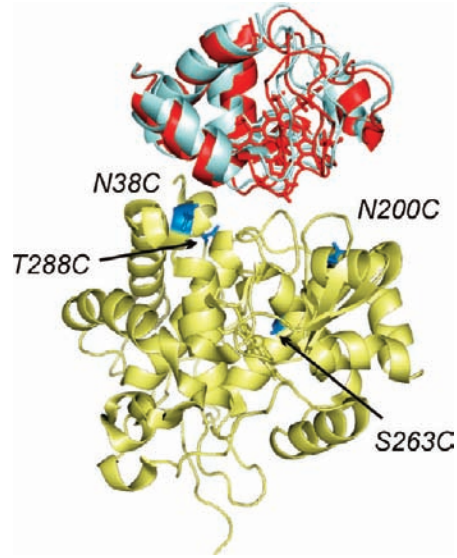

Figure 4. Specific form of the T12A Cc−CcP complex in solution. CcP residues used for conjugation with SLs are labeled and shown as blue sticks. The specific forms of wt and T12A Cc−CcP complexes are compared, with CcP molecules superimposed and Cc backbones in the variant and wt complexes colored red and light blue, respectively. Heme groups for both proteins are in sticks. The coordinates for the solution structure of the wt complex were taken from the PDB entry 2GB8.[14]

thus, the line width of Cc resonances is a weighted average of those in the free (12 kDa) and CcP-bound (46 kDa) forms. At the protein concentrations used, the fraction of bound Cc decreases in the order T12A (96 ± 1%) > R13K (79 ± 2%) > R13A (50 ± 3%), which explains higher signal-to-noise ratios for weaker R13-substituted complexes (Figures 5 and S2−S4). Varying noise levels for CcP-SL complexes with the same Cc mutant are due to different sample concentrations, total acquisition time, or both. Only five PRE profiles were obtained for T12A Cc. As T12A Cc−CcP data are very similar to those of the wt complex, the data sets with SL positions V10C, K97C, N141C, N164C, and L213C, which showed no effect in the wt complex,[29] were not collected.

For all Cc mutants, only three SLs attached close to the crystallographic binding site (positions N38C, N200C, T288C, colored red in Figure 1A) give rise to observable PREs, whereas others show no effects (Figures 5 and S2−S4). Overall, the distribution of the PREs is highly similar to that of wt Cc−CcP,[29] suggesting that the introduced Cc mutations do not significantly alter the conformational space sampled by the proteins within the complex.

Encounter Ensemble Simulations. The measured PRE (Γ_2^{obs}) is a population-weighted average of the contributions from all protein−protein orientations (eq 2) and, as such, contains the

(52) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L. N. *Anal. Biochem.* **1989**, *179*, 131–137.
(53) Xu, X.; Reinle, W.; Hannemann, F.; Konarev, P. V.; Svergun, D. I.; Bernhardt, R.; Ubbink, M. *J. Am. Chem. Soc.* **2008**, *130*, 6395–6403.

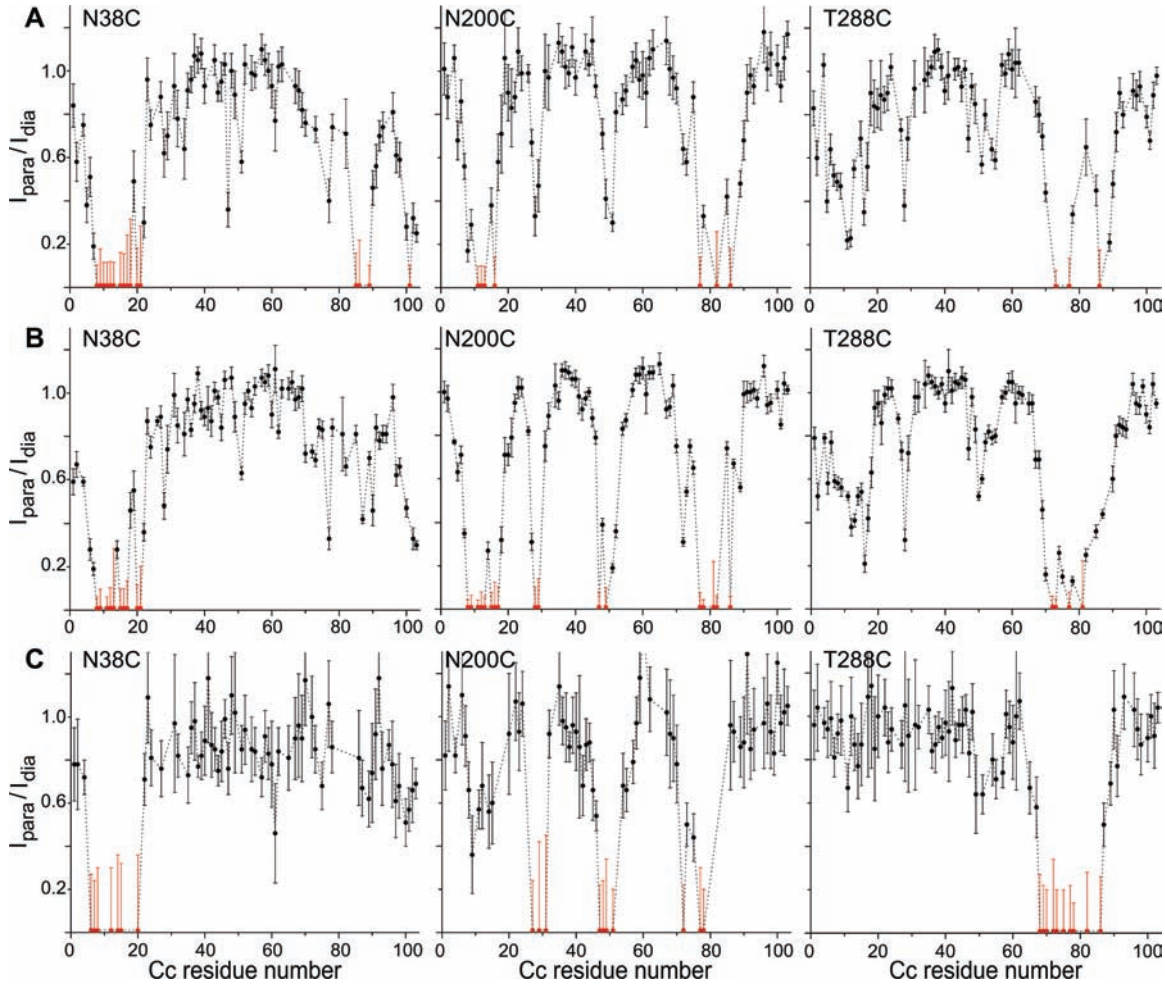

Figure 5. Paramagnetic effects in Cc−CcP-SL complexes. Relative [^{1}H, ^{15}N] HSQC intensities of (A) R13A, (B) R13K, and (C) T12A Cc backbone amides in the complex with CcP labeled with a paramagnetic SL (I_{para}) or a diamagnetic analogue (I_{dia}) at positions N38C, N200C, and T288C. For the residues whose resonances disappear in the paramagnetic spectrum (red), the upper limit of I_{para}/I_{dia} was estimated from the noise level. The error bars denote standard deviations, derived from spectral noise levels using standard error propagation procedures. Profiles for the other SL positions, showing no measurable effect, are shown in Figures S2−S4 in the Supporting Information.

information on both the specific form and the encounter state.[54] The PRE for the former ($\Gamma_2^{specific}$) can be back-predicted from the crystal structure of the complex.[13] To obtain the value for the population of the encounter state (p), its ensemble-averaged PRE ($\Gamma_2^{encounter}$, eq 2) needs to be determined. This PRE can be calculated from the spatial distribution of the constituent protein−protein orientations. Recently, we have shown that MC simulations provide an accurate description of the wt Cc−CcP encounter complex, which is in excellent agreement with experimental data and thus can be used to derive the required $\langle\Gamma_2^{encounter}\rangle$.[29] In the MC simulations, the proteins are treated as rigid bodies with partial charges assigned at the atomic level. Intermolecular electrostatic interactions and an exclusion grid, designed to prevent steric clashes, guide the subsequent protein−protein docking. Using this approach, here we simulated ensembles of encounter orientations for the three mutant complexes.

As can be seen from the energy versus the distance root-mean-square (drms)[55] plots, a single cluster of solutions is produced in all MC runs (Figure 6). The energy distributions of the MC solutions for R13K and T12A Cc−CcP are similar to that of the wt complex. For R13A Cc−CcP the distribution is shifted to higher energy values, indicating that removal of a positively charged side chain impairs protein association, as would be expected for an electrostatically driven process.

The spatial distributions of the simulated MC ensembles are shown in Figure 7 for R13A Cc−CcP and in Figures S5−S7 for the wt, R13K, and T12A Cc−CcP complexes. Similarly to what was seen for the wt complex,[29] most MC solutions are found surrounding the dominant Cc−CcP orientation observed by X-ray crystallography[13] and solution NMR spectroscopy.[14] In all complexes, Cc explores a narrow patch on the CcP surface, navigating between two local energy minima (panels A and C in Figures 7 and S5−S7), which were also observed in the classical Brownian dynamics study.[30] This patch contains the binding site of the specific complex, indicating that the electrostatic preorientation of CcP, occurring during a diffusional approach of Cc, steers Cc to the dominant binding site. As it probes the CcP surface, Cc experiences considerable rotational motion, as evidenced by broad distributions of CcP centers of mass around Cc (panels D−F in Figures 7 and S5−S7). Overall, the extent of the conformational space sampled by the proteins in the wt and variant encounter complexes appears to be very

(54) Clore, G. M. *Mol. Biosyst.* **2008**, *4*, 1058–1069.
(55) Kim, Y. C.; Tang, C.; Clore, G. M.; Hummer, G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 12855–12860.

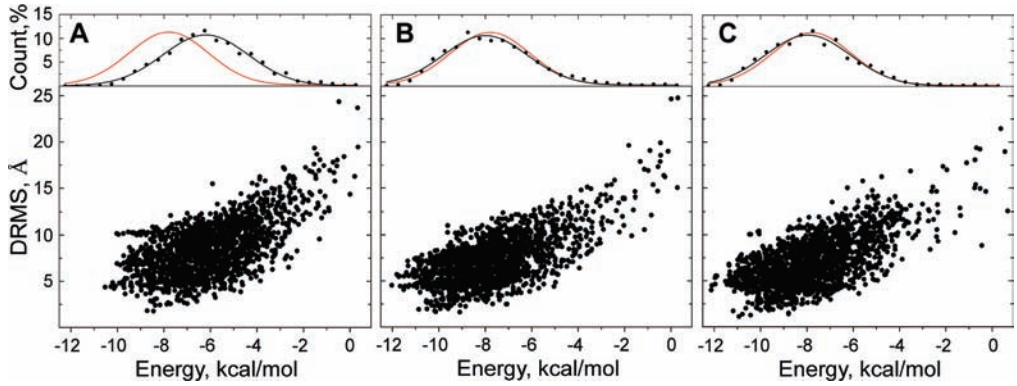

Figure 6. Drms versus energy plots for the Monte Carlo simulations of the encounter state. (A) R13A, (B) R13K, and (C) T12A Cc−CcP. In each panel the upper graph shows the energy distribution fitted with a Gaussian (black) with parameters (A) $r^2 = 0.98$ and $x_c = -6.2 \pm 0.1$ kcal/mol; (B) $r^2 = 0.98$, $x_c = -8.0 \pm 0.1$ kcal/mol); and (C) $r^2 = 0.97$, $x_c = -8.0 \pm 0.1$ kcal/mol. A Gaussian distribution for the wt complex (red; $r^2 = 0.98$, $x_c = -7.8 \pm 0.1$ kcal/mol) is given for comparison.

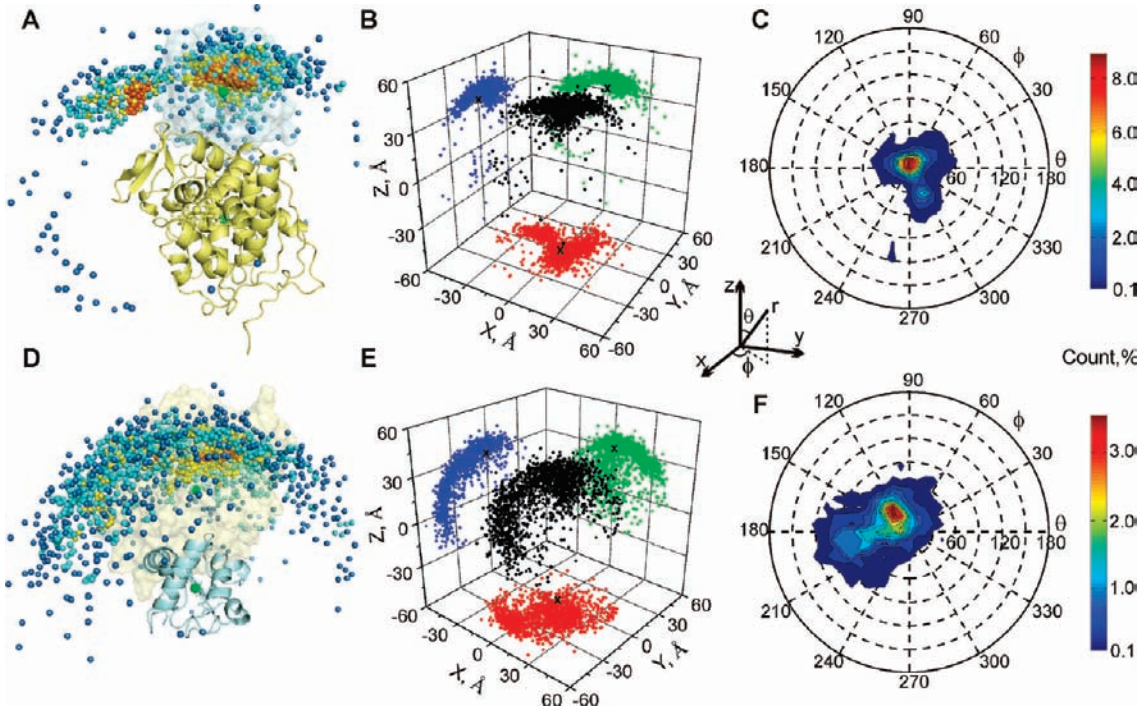

Figure 7. Monte Carlo simulation of R13A Cc−CcP encounter state. (A, D) Spatial distribution of the encounter complexes. Cc and CcP molecules in the specific complex are colored light yellow and blue, respectively, with one protein in cartoon and the other as a transparent surface. The centers of mass of (A) Cc around CcP and (D) CcP around Cc in the simulated encounter ensemble are shown as spheres, colored according to the density of the distribution increasing from blue to red. The highest densities denote the most favorable electrostatic orientations. Centers of mass of the proteins in the specific complex (green spheres) are aligned along the *z*-axis of the Cartesian coordinate system and correspond to spherical coordinates $(\theta, \phi) = (0,0)$ as defined in the inset. The view in (D) is obtained by rotating that in (A) by 160° around the *y* axis. (B, E) Distribution of (B) Cc centers of mass around CcP and (E) CcP centers of mass around Cc in the Cartesian coordinate space. Colored symbols denote projections on the coordinate planes, and crosses indicate the position of the center of mass of the partner protein in the specific complex. (C, F) Distribution of (C) Cc centers of mass around CcP and (F) CcP centers of mass around Cc in the spherical coordinate system. The contour levels refer to the counts (in percent) of the MC output structures binned in 10° steps along both θ and ϕ dimensions.

similar, implying that the introduced Cc mutations do not significantly alter the geometry of the encounter state.

Just as the case for wt Cc−CcP,[29] the intermolecular ET donor−acceptor distances in mutant encounters are large (Figure S8). At the observed donor−acceptor separations, the rate of intermolecular ET in the encounter state is expected to be very low.[56] This finding agrees well with the results of recent kinetics studies of Erman and co-workers, who concluded that most of the ET occurs in the dominant form of the Cc−CcP complex.[25−27]

Encounter State Populations. To obtain the encounter state population (*p*), the back-calculated PRE contributions of the specific and the encounter forms are fitted to the experimental Γ_2^{obs} (eq 2). The quality of the fit is estimated by calculating a *Q* factor (eq 3), and the *p* values at the lowest *Q* are selected.[29] A single, well-defined minimum in $Q = f(p)$ plots, corresponding to the optimal *p* values, is found for all studied complexes (Figure 8), and a combination of the specific form ($p = 0$) and the encounter state ($p = 1$) provides a better agreement with

(56) Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L. *Nature* **1992**, *355*, 796–802.

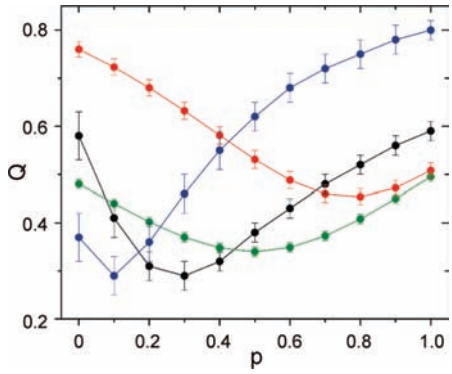

Figure 8. *Q*-factor analysis of the encounter complexes. *Q* factors for the PREs calculated at varying *p* for wt[29] (black), T12A (blue), R13K (green), and R13A (red) Cc−CcP. The errors are standard deviations.

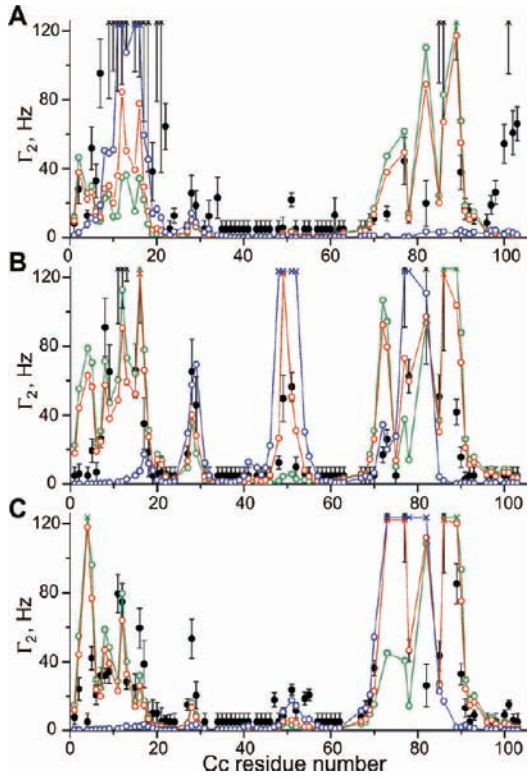

Figure 9. Observed and calculated PREs for R13A Cc−CcP-SL complexes. PRE profiles for SLs attached at position (A) N38C, (B) N200C, and (C) T288C. Experimental Γ_2^{obs} (black) and PREs back-calculated for the specific orientation (blue), simulated encounter state (green), and the combination of the two (red) at the optimal *p* value of 0.8. Crosses indicate the values of $\Gamma_2 \geq 125$ Hz for the calculated PREs or identify the residues whose resonances disappear in the paramagnetic spectrum. For the latter only the lower limit of Γ_2^{obs} could be estimated from the spectral noise level. For residues with $I_{ox}/I_{red} \geq 0.9$ (Figure 5A), Γ_2^{obs} is set to 5 ± 5 Hz; otherwise, the errors are standard deviations.

the observed PREs than either of these forms alone (Figures 9 and S9−S10).

Remarkably, there is a clear shift from $p = 0.3$ in the wt complex to $p = 0.1$ in T12A Cc−CcP and, in the other direction, to $p = 0.8$ in R13A Cc−CcP (Figure 8). In the latter complex, the encounter state is no longer a minor species but rather the dominant form. Thus, it appears that a single mutation can both decrease (T12A) and significantly increase (R13A) the population of the encounter state.

The crystallographic wt Cc−CcP orientation,[13] known to be the dominant species in solution,[14] was used to represent the

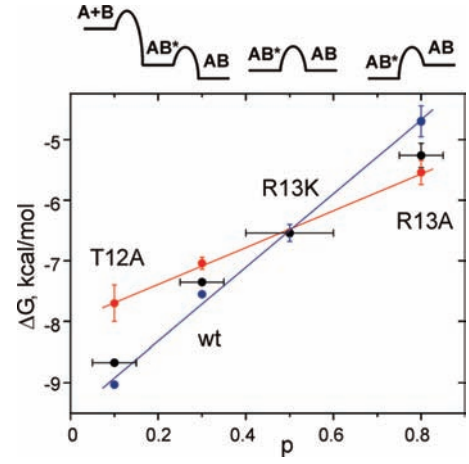

Figure 10. Energy contributions of the encounter states in variant Cc−CcP complexes. Experimental ΔG_B (black) decomposed into the contributions of the specific form (blue) and an encounter state (red). The former and the latter fit to linear trends with $r^2 = 0.98$, slope 6.1 kcal/mol and $r^2 = 0.96$, slope 3.1 kcal/mol, respectively. The errors are standard deviations derived from the uncertainties in experimental binding constants and determined *p* values (eq 10 in the Supporting Information). The diagrams above the plot schematically outline relative positions of the energy levels of the free proteins (A+B), the specific form (AB), and an encounter state (AB*) for the Cc−CcP complexes.

specific form in all complexes. This approach is justified as follows: (1) at the resolution of protein backbones, T12A mutation does not perturb the structure of the major form of the complex (Figure 4); (2) although no structures of the weak R13K and R13A Cc−CcP complexes could be obtained, given the overall similarity of their PRE profiles (Figure 5A,B) to those of the wt,[29] their specific binding geometries must not be greatly altered. Nevertheless, slight differences in the structures of the specific forms would give rise to higher *Q* factors observed for the R13-substituted complexes (Figure 8). Also, at higher encounter state populations, the quality of the agreement with the experimental data becomes sensitive to the level of detail for the simulated ensemble, which is limited by the approximations inherent in the simulation protocol.

Shifting the Equilibrium between the Encounter State and the Specific Form. There is a qualitative agreement between *p* and the size of the NMR chemical shift perturbations ($\Delta\delta_{avg}$) in Cc−CcP complexes (Figures 3 and 8). This finding strengthens the conclusions of earlier studies that correlated decreased $\Delta\delta_{avg}$ with an increased dynamics in transient ET protein complexes[48,49,53] and suggests that most of the binding shifts arise from the specific form of a complex. Interestingly, there is a linear correlation between *p* and the free binding energy (ΔG_B), with weaker complexes exhibiting higher *p* values (Figure 10). The T12A mutation increases the binding constant by an order of magnitude and decreases *p* to 0.1. R13K weakens the complex 4-fold and gives rise to equal populations of the specific and encounter forms. Finally, R13A decreases the binding 100-fold and increases *p* to 0.8 (Table 1 and Figure 10).

When ΔG_B is decomposed into individual contributions of the two forms (see Theory section in the Supporting Information), another trend can be discerned. The mutations exert a larger effect on the specific orientation than on the encounter state, and consequently, as complexes become progressively weaker, the contribution of the latter to ΔG_B steadily increases (Figure 10). In the extreme case of R13A, the encounter state becomes the dominant species and imparts most of the

thermodynamic stabilization to the overall protein binding. Thus, the introduced mutations reshape the energy landscape of Cc−CcP interaction and, in the case of R13A, can even interchange the energy levels of the specific and the encounter forms (Figure 10).

Earlier kinetic studies, mainly concerned with Cc F82 variants, already demonstrated that small changes in the interface can reduce the rate of intermolecular ET and cause Cc to assume different orientations in crystal structures.[57,58] These results show how easily the specific complex is perturbed, implying only a minor stabilization relative to the encounter state. A similar finding was reported in a study of the complex between cytochrome *f* and plastocyanin from *Prochlorothrix hollandica*: NMR data showed that the native complex exhibited dynamics, which was enhanced by flattening the plastocyanin interaction surface by mutagenesis.[59] Recently, Xiong et al.[60] converted a highly dynamic complex of myoglobin and cytochrome b_5 into a more static, specific one by introducing three charge-reversal mutations around the front face of myoglobin. In that case, reduction of the electrostatic repulsion between the proteins appears to be sufficient to form a specific complex.

Many experimental studies of protein−protein association (reviewed recently in ref 2) focused on the transition state (TS) of the free-energy barrier separating the encounter state and the specific form (AB* and AB in the schematic diagrams of Figure 10), which was believed to resemble the latter.[2] For example, it was shown that mutations that enhance the association rate in the TEM1−BLIP complex can make the TS less diffusive.[61] There is an essential difference between the TS and the encounter complex. By definition, the former is extremely short-lived and, due to its very low population, can only be probed by kinetic experiments. The encounter state, an intermediate state between free proteins and the specific form, can be populated significantly, which allows its characterization by PRE NMR spectroscopy. Here we used this tool to address the entire ensemble of protein−protein orientations constituting the encounter state, thus providing an insight into the equilibrium between this state and the specific form of the complex. Our present results suggest that, in addition to its accepted role in enhancing the binding kinetics,[2,28] the encounter can make a large contribution to the thermodynamic stabilization of weak protein interactions. This finding could rationalize the existence of highly dynamic macromolecular complexes consisting entirely of multiple encounter forms.[53,62]

On the Specificity of Protein−Protein Interactions. In the context of a single protein complex, specificity can be defined as an ability to adopt a unique, well-defined orientation as opposed to multiple, possibly overlapping, binding modes. We propose that p is a good metric for specificity in weak protein complexes, which can be used to classify a range of transient interactions varying from mostly a single orientation ($p < 0.5$),[14,63] to less specific ones ($p \approx 0.5$),[48,49] to those dominated by an encounter state ($p > 0.5$).[53,62] In fact, as we show here, p can be modulated within a single protein complex, thereby shifting the delicate balance of the specificity in either direction.

Acknowledgment. We thank Dr. I. Lasters and Dr. E. Blokhuis for assistance with the data analysis. A.V., J.A.R.W., and M.U. were supported by The Netherlands Organisation for Scientific Research, VIDI Grant 700.52.425 and Softlink Grant 98S1010. Q.B. was supported by a fellowship from the Higher Education Commission of Pakistan. G.M.U. was supported by the Deutsche Forschungsgemeinschaft Grant UL 174/7-1.

Supporting Information Available: Text detailing the theory behind the free energy decomposition; figures showing chemical shifts of Cc variants; intensity ratio plots for R13A, R13K, and T12A Cc−CcP for SL positions exhibiting no paramagnetic effect; coordinate analysis of the Monte Carlo simulations of the wt, T12A, and R13K Cc−Cc encounter state; histogram of ET distance distribution in the mutant encounters; PRE profiles for R13K and T12A Cc−CcP; and complete ref 46. This material is available free of charge via the Internet at http://pubs.acs.org.

JA100867C

(57) Everest, A. M.; Wallin, S. A.; Stemp, E. D. A.; Nocek, J. M.; Mauk, A. G.; Hoffman, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 4337–4338.
(58) Kang, S. A.; Crane, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15465–15470.
(59) Hulsker, R.; Baranova, M. V.; Bullerjahn, G. S.; Ubbink, M. *J. Am. Chem. Soc.* **2008**, *130*, 1985–1991.
(60) Xiong, P.; Nocek, J. M.; Griffin, A. K. K.; Wang, J.; Hoffman, B. M. *J. Am. Chem. Soc.* **2009**, *131*, 6938–6939.
(61) Harel, M.; Cohen, M.; Schreiber, G. *J. Mol. Biol.* **2007**, *371*, 180–196.
(62) Worrall, J. A. R.; Liu, Y.; Crowley, P. B.; Nocek, J. M.; Hoffman, B. M.; Ubbink, M. *Biochemistry* **2002**, *41*, 11721–11730.
(63) Tang, C.; Iwahara, J.; Clore, G. M. *Nature* **2006**, *444*, 383–386.