Published in final edited form as:
*Biochemistry*. 2011 March 15; 50(10): 1664–1671. doi:10.1021/bi101820d.

# Experimentally restrained molecular dynamics simulations for characterizing the open states of cytochrome $P450_{cam}$ [a,b]

**Eliana K. Asciutto**[1], **Marina Dang**[2], **Susan Sondej Pochapsky**[2], **Jeffry D. Madura**[1], and **Thomas C. Pochapsky**[*,2,3]

[1] Dept. of Chemistry and Biochemistry, Center for Computational Sciences, Duquesne University, Pittsburgh, PA

[2] Dept. of Chemistry, MS 015, Brandeis University, 415 South St., Waltham, MA 02454-9110

[3] Rosenstiel Basic Medical Sciences Research Institute, Brandeis University, Waltham, MA 02454-9110

## Abstract

Residual dipolar couplings were used as restraints in fully solvated molecular dynamics simulations of reduced substrate- and carbonmonoxy-bound cytochrome $P450_{cam}$ (CYP101A1), a 414-residue soluble monomeric heme-containing camphor monooxygenase from the soil bacterium *Pseudomonas putida*. The $^1D_{NH}$ residual dipolar couplings used as restraints were measured in two independent alignment media. A soft annealing protocol was used to heat the starting structures while incorporating the RDC restraints. After production dynamics, structures with the lowest total violation energies for RDC restraints were extracted to identify ensembles of conformers accessible to the enzyme in solution. The simulations result in substrate orientations different from that seen in crystallographic structures and a more open and accessible enzyme active site, and largely support previously reported differences between the open and closed states of CYP101A1.

---

The *sine qua non* of enzyme function is that the three-dimensional structure of an enzyme stabilizes the transition state of the reaction being catalyzed (1). This implies that the enzyme structure cannot stabilize the reactant (or product) ground states, at least not at the point of catalysis. However, enzymes have another job to do as well: They must concentrate reactants in the active site, that is, the thermodynamics of substrate binding must be sufficiently favorable so that substrate is bound selectively and with reasonable affinity. How then to square this particular circle? One way is to assume an enzyme is able to access different conformations in order to accomplish both tasks (2). If this is true, one would expect to find accessible “open” enzyme conformers that allow substrate access and bind substrate with a favorable free energy, and a different set of “closed” conformers that stabilize the appropriate transition states for catalysis. Furthermore, the activation barriers separating the open and closed conformers should be sufficiently high that each

---

[a]This work was supported in part by a grant from the USPHS (R01-GM44191, TCP). EKA and JDM acknowledge support from USPHS grant R01DA027806, a Department of Education grant P116Z090309 and NSF support through Teragrid (MCB060061N).

[b]The REP structure discussed in this paper has been deposited as PDB entry 2l8m. NMR chemical shifts and RDC restraints are deposited as BMRB entry 17415.

[*]Inquiries should be addressed to this author at Dept. of Chemistry, MS 015, Brandeis University, 415 South St., Waltham, MA 02454-9110. Tel.: 781-736-2559, fax: 781-736-2516, pochapsk@brandeis.edu, website: www.chem.brandeis.edu/pochapsky.

Supplementary Material Available: PDB format structure file for REP structure, RDC restraints in AMBER format and two tables discussed in the text are available in online at http://pubs.acs.org.

conformation is discretely accessible on the time scale of catalysis, but not so formidable as to preclude access to either.

We have used multidimensional nuclear magnetic resonance (NMR) spectroscopy to characterize the "open" conformations of the 46 kDa heme-containing monooxygenase enzyme cytochrome $P450_{cam}$ (CYP101A1) (3,4), and to identify a structural switch between the "open" and "closed" catalytically active form of this enzyme that reorients bound substrate into the appropriate position for the observed catalysis (3,5,6). In the course of this work, we also identified functionally important differences between crystallographically determined closed structures of CYP101A1 and the ensemble of conformers accessible to that enzyme in solution (5–8). We used solvated molecular dynamics simulations of CYP101A1 to characterize structural perturbations induced by the addition of the effector and redox partner putidaredoxin (Pdx) (3). We were gratified to find that differences between the closed Pdx-bound and free open forms of CYP101A1 observed in the simulations were for the most part faithfully matched by the NMR-detected perturbations observed in the TROSY-HSQC spectrum of CYP101A1 upon addition of Pdx. Still, those simulations of the open form did not include any experimental restraints, and we wished to generate ensembles of the enzyme that would accurately reflect available experimental data.

The differences we observed between the published structures of CYP101A1 (9) and the solution ensembles did not extend to major changes in secondary structure or tertiary structural rearrangements. Rather, they involve changes in the orientation of bound substrate (1*R*)-(+) camphor in the active site and in the relative orientation of secondary structural features that result in easier access and egress to and from the active site for substrates and products (3). Such changes do not result in significant differences in most NH-NH proton distances, and amide proton nearest-neighbor interactions provide the only nuclear Overhauser effects (NOEs) that can be identified in adequate numbers from NOESY spectra obtained using otherwise perdeuterated CYP101A1 samples. Due to efficient relaxation in protonated CYP101A1 samples, detectable NOEs due to carbon-attached protons are weak and difficult to compare meaningfully. On the other hand, our previous results suggested that a number of secondary structural features change their orientations sufficiently that residual dipolar couplings (RDCs) measured for $^{15}N$-$^{1}H$ pairs could be used to generate meaningful solution structural ensembles of CYP101A1.

We now show that a comprehensive set of RDC restraints measured in two alignment media and incorporated into a fully solvated restrained molecular dynamics simulation of the reduced camphor- and carbonmonoxy-bound form of CYP101A1 (CYP-S-CO) results in structural ensembles of "open" states of the enzyme that are consistent with experimental data and provide insight into the conformations accessible to this relatively large enzyme in solution. We expect that the methodology described here will be useful to others as a tool for investigating conformational equilibria in enzymes.

## Materials and Methods

### Expression and purification of CYP101A1

A plasmid construct encoding the C334A CYP101A1 mutant was transformed in *E. coli* NCM533 cells by electroporation. The C334A mutant has been found to be spectroscopically and enzymatically identical to the wild type, but it does not form dimers in solution (10). Uniformly $^{15}N$-labeled CYP101A1 was expressed in M9 minimal media containing $^{15}NH_4Cl$ along with trace metals, and a heme precursor (γ-aminolevulinic acid) was introduced prior to induction of CYP101A1 expression. Protein purification included a protamine sulfate cut to precipitate nucleic acids, and an ammonium sulfate cut to isolate CYP101A1. The desalted sample was passed through two gravity chromatography columns.

Expression and purification details have been published previously (11). All buffers contained 2 mM (1*R*)-(+)-camphor (98%, Sigma Aldrich). Protein purity was assessed spectroscopically and determined to be at least 90% pure with a $A_{391}/A_{280}$ ratio of at least 1.4.

### Sample preparation for NMR experiments

Buffer exchange was performed by passing the protein sample through a spin column containing P2 resin that was pre-equilibrated with 50 mM Tris-HCl (pH 7.4), 100 mM KCl, 2 mM camphor in 90:10 $H_2O$:$D_2O$. The sample was then reduced with microliter aliquots of freshly prepared sodium dithionite under an atmosphere of carbon monoxide in a septum-sealed vial. Typically, a total of ~8 μL of a 250 mM $Na_2S_2O_4$ solution is sufficient to completely reduce 250 μL of protein sample. CYP101A1 concentrations were typically ~250 μM in the final NMR sample. The reduced protein was anaerobically transferred to a susceptibility-matched NMR tube (Shigemi, Inc., Allison Park, PA).

### Alignment using an aqueous nematic liquid crystalline suspension

A 10% (v/v) solution of pentaethylene glycol monododecyl ether (C12E5, 98+%, Fluka, Switzerland) was prepared in a gas-tight vial using degassed NMR buffer (50 mM Tris-HCl, pH 7.4, 100 mM KCl, 2 mM camphor in 90:10 $H_2O$:$D_2O$). The vial was flushed with carbon monoxide for 2 min at a flow rate of approximately 6 mL/min. Camphor is volatile, so excessive CO flushing was avoided. This stock solution was vortexed and 180 μL of it was transferred into a clear, gas-tight vial. The atmosphere in the vial was flushed with CO for 3 min, alternating with a few seconds of vigorous vortexing. Using a metal syringe, 7 μL of *n*-hexanol (anhydrous, 99+%, Sigma-Aldrich, Milwaukee, WI) were introduced in small aliquots, as small as 0.5 μL as the final concentration was approached, with vortexing after each addition (12). The mixture, which was originally clear, became viscous with a milky appearance then turned completely clear with no bubbles. The final mixture appears opalescent, with a molar ratio of C12E5/hexanol = 0.79. In an argon-filled glove bag, a reduced 0.5 mM WT CYP101A1 sample in NMR buffer was added to the polyol/hexanol mixture (1:1, v/v). The solutions were mixed then transferred into a susceptibility-matched NMR tube (Shigemi, Inc., Allison Park, PA). All steps involving the C12E5/hexanol mixture were performed at room temperature. The sample alignment was monitored using the residual quadrupole splitting of the buffer $^2H$ signal (~23 Hz after ½ hour).

### Alignment using filamentous phage Pf1

Bateriophage *Pf1* (Asla Biotech Ltd., Latvia) was used as a second alignment medium. 3.2 mg of *Pf1* was transferred to a screw-top vial and flushed with CO for 10 min, then for an additional 5 min with shaking. The vial was transferred to an anaerobic chamber, where 3 μL of freshly prepared 0.25 M $Na_2S_2O_4$ in 1M Tris buffer (pH 7.4) were added. 400 μL of 0.5 mM CYP-1-CO was added to the vial containing the phage and reducing agent. The mixture was stirred gently then transferred into a susceptibility-matched NMR tube. The final phage concentration was ~8 mg/mL. The sample alignment was monitored using the residual quadrupole splitting of the buffer $^2H$ signal (~9.5 Hz after ½ hour).

### NMR experiments

A combination of TROSY and semi-TROSY datasets were acquired for both aligned and unaligned samples of reduced camphor and CO-bound CYP101A1 (CYP-S-CO). The pulse sequences were adapted from Weigelt (13). Spectra were acquired as 1024 ($^1H$) × 512 ($^{15}N$) complex point datasets at 298 K on an 18.8 T (800 MHz $^1H$) Bruker Avance NMR spectrometer equipped with a cryogenic probe. The interleaved data was separated prior to Fourier transformation, yielding two two-dimensional data sets with 1024 ($^1H$) × 256 ($^{15}N$)

points each. For each data set, one level of linear prediction was applied in the $^{15}N$ dimension (to 512 complex points), treated with a 90° shifted squared sine bell and zero-filled to 1024 points. Both data sets were treated with a Gaussian function in the $^{1}H$ dimension and zero-filled to 2048 points prior to Fourier transformation. This processing results in the optimum display for measuring the offset between the TROSY and semi-TROSY peaks in the $^{1}H$ dimensions of the two spectra. Data acquisition and processing were performed using the Topspin software package (Bruker Biospin, Inc.). One bond $H_N$-N RDCs were obtained from the differences between $^{1}J_{NH}$ coupling measured in the isotropic CYP-S-CO sample and in aligned medium.

## Molecular Dynamics Simulations

In previous work (3), we described a model for the solution structure of CYP-S-CO in which a single Ile-Pro bond (Ile 88-Pro 89) that precedes the B′ helix is isomerized from the crystallographically-observed *cis* conformation to distorted *trans*, based on spectroscopic evidence and the results of site-directed mutagenesis experiments (5). Summarizing that work, we started from a minimized coordinate set based on the crystallographic structure of CYP-S-CO (PDB entry 3CPP) (14). Because the first eight N-terminal residues are disordered in this structure, these are not included in the simulations. We note that NMR evidence indicates that these residues are also disordered in solution (3). We performed a series of restrained molecular dynamics simulations, varying the Ile88–Pro89 torsion angle ω by 5° at 350 K with a torsional restraint (force constant set to 50 kcal/mol), and relaxing the system at 300 K after each run until the *trans* conformer was obtained. For these runs, the structure was solvated with 16,935 TIP3P water molecules and sufficient $K^+$ and $Cl^-$ ions were added to neutralize the resulting structure and generate a KCl concentration of 0.1 M, the typical KCl concentration used to stabilize the active enzyme experimentally (15). After equilibration, a 3 ns production molecular dynamics simulation was performed. In the present work, we took snapshots from the 3 ns simulation as starting structures.

156 $H_N$-N RDCs from C12E5/hexanol medium and 167 $H_N$-N RDCs from filamentous phage *Pf1* were applied simultaneously to the initial structure by performing restrained molecular dynamics using Amber10 (16). RDC restraints were applied reflecting the expected uncertainty of the experimental values due to spectral line widths and digital resolution (+/− 3 Hz). This uncertainty was implemented by penalizing calculated couplings outside a +/− 3 Hz range, with a penalty function proportional to $(D_{calc} - D_{obs(u,l)})^2$, where the indices *u* and *l* refer to the upper and lower limits. The initial alignment tensor was calculated from 8000 minimization steps where only the elements of the alignment tensor are allowed to change. During the simulation the elements of the alignment tensor are optimized along with the structural parameters and updated manually after the end of each run (17).

Refinement was performed as follows. After minimization, the system was equilibrated through 3 cycles of 30 ps of constant volume dynamics, gradually heating from 0K to 50K, 50K to 100 K, and 100 K to 300 K, with 5 kcal/(mol $Å^2$) harmonic restraints relative to the starting structure. A 1 ps time constant was used for coupling to the heat bath. Heating was followed by 2 cycles of 30 ps constant pressure dynamics with a pressure relaxation time of 0.2 ps, which was then increased to 1.0 ps, and the harmonic constraints on the protein released. Once the pressure was equilibrated, another 60 ps constant volume MD was performed. Last, 1 ns of productive constant volume dynamics at 300K was performed. The described procedure was applied to 9 different structures taken from our previous simulation (3). The method described was adapted from that used for the refinement of the 215-residue glutaredoxin using dihedral and distance restraints (18).

This simulation protocol was applied iteratively in order to obtain self-consistent results. Preliminary RDC datasets were applied as restraints, followed by analysis of the results in terms of goodness of fit of the calculated structures to the experimental data. Outlying RDC values were checked for accuracy, and as more RDCs were identified, they were incorporated into the calculations. This process was continued until the deviations reach a plateau.

Molecular volumes were estimated as follows. From the previously published simulations of the *cis* conformer, a representative structure was calculated by clustering the trajectory into 5 clusters using the Amber10 module *ptraj*. The representative structure corresponds to a representative snapshot from the most populated cluster. Molecular volumes were calculated using Chimera (19). The volume of the *cis* representative structure is ~ 52410 $Å^3$. A similar calculation for the representative *trans* structure REP yields a volume of 52950 $Å^3$.

## Results

The iterative process described above resulted in a total of 323 $H_N$-N RDC restraints measured in two media for use in the final simulations. A scatter plot of the measured RDC values as a function of sequence are shown in Figure 1. As can be seen from the plot, the measured RDCs are fairly evenly distributed over the entire sequence, so no portion of the enzyme structure is devoid of restraints.

The initial nine structures were refined with the procedure described in the Methods Section. For each simulation we generated trajectories saving frames every 1 ps, giving a total of 1000 snapshots/simulation. From these 1000 snapshots, the two structures with the lowest restraint violation energies were selected, yielding an ensemble of 18 structures (Figure 2). Backbone rmsd between the structures is 1.298 Å. The 18 structures selected are representative of the ensemble, since the backbone rmsd values between each structure and the average structure in each simulation are similar to the backbone rmsd values between the average and the rest of the structures of the simulation (Table S1, Supporting Information). The first 10 residues display the greatest rmsd values. Other regions showing some discrepancies are: the region between the F and G helices, the end of the B helix: Asp 77, some residues at the C helix, especially Met 121, residues 304–309, and 343–344. It is interesting to note that the zone comprised by residues 304–309 has been reported previously to show resonances that shift upon addition of the effector (5), suggesting residues that would exhibit more mobility.

The 18 lowest restraint energy structures were further improved by clustering into 5 groups based on rmsd using the average linkage algorithm with the *ptraj* module of Amber10. From the most populated group, a representative structure was calculated. Ions and water molecules were added to the representative structure and the restraining procedure was repeated, including minimization, equilibration and a 1ns production run with all the RDCs applied. The lowest energy violations were seen in this final run, suggesting a more converged ensemble. Figure 3 shows experimental versus calculated $H_N$-N RDC values for a representative structure from this run.

The correlation between experimental RDCs and those calculated from the simulation is fairly good and similar in both media. Each member of the 18-member ensemble shows similar $R^2$ values between 0.7 and 0.8. Residues displaying greater deviations are the ones with greater rmsd values between the 18 structures mentioned before, suggesting more mobility in such regions. To investigate this point, $H_N$-N vector bond parameters were calculated and plotted in Figure 4. As expected, the first ten N-terminal residues, which exhibit random coil shifts in NMR spectra, exhibit the highest mobility. Residues 333 and

350 are near the beginning and end, respectively, of the β-meander. The β-meander is a region of irregular secondary structure that follows the K′ helix and precedes the loop containing the heme axial ligand Cys 357. Inspection of this region in Figure 5 shows the largest variations in structure relative to 3CPP of any region of the enzyme.

From the last simulation (starting from the representative structure), the snapshot with lowest restraint energy was identified, and the coordinates for this structure (REP) are deposited in the PDB database and can be found in Supplementary Material.

## Comparison of a representative structure with results of unrestrained dynamics simulations

To determine the effect of RDC restraints on the course of the simulations, we calculated best-fit alignment tensors for the current experimental RDC values using the open and closed structures reported in our previous work (3). Calculations were performed using AMBER 10, as described above. The positions of all the atoms involved in the RDC restraints were fixed, and the resulting alignment tensors were used to back-calculate expected RDC values. Comparison of expected RDCs obtained for both the open and closed structures with experimentally observed values shows that calculated and experimental values are almost uncorrelated, with $R^2$ values < 0.1 for both cases and in both (C12E5 and *pfl*) media. To ensure that the RDC restraints do not introduce local strain in bond angles, we compared the bond angle energy term for restrained simulations against unrestrained simulations from (3). We found that the average bond angle energy does not change significantly (~ 1%), which is much less than average thermal fluctuations (~ 120 kcal/mol). Clearly, the presence of the RDC restraints changes the results of the simulations, at least in regards to NH bond vectors. Structurally, the beginning of B-B′ loop (residues 77–82), also shows considerable difference with the previous results for the *trans* simulations, and in the current restrained simulations is much less perturbed from the original 3CPP structure.

The most significant difference between the unrestrained and restrained simulations is in the vicinity of the I helix “kink” in the active site. In crystallographic structures of CYP-S-CO, an interruption in regular α-helix hydrogen bonding is observed due to the absence of the expected hydrogen bond between the NH of Thr 252 to the carbonyl oxygen of Gly 248. Instead, the Gly 248 carbonyl accepts a hydrogen bond from the γ-OH of the Thr 252 side chain (Fig. 6). The gap introduced by this interruption in the helix provides the binding site for the Fe-bound carbon monoxide. In the original unrestrained calculations on the open form of the enzyme, we observed regularization of the helix hydrogen bonding pattern, with the formation of the Gly 248 C=O --- H-N Thr 252 hydrogen bond. In the current calculations, while the Thr 252 side chain γ-OH remains uninvolved in main chain hydrogen bonding, the Gly 248 C=O forms a well-developed hydrogen bond with the NH of Val 253, and the NH groups of Asp 251 and Thr 252 form a bifurcated hydrogen bond with the carbonyl oxygen of Val 238. The effect of these changes is to maintain the gap observed in the crystallographic structures, but the gap is now stabilized by two *i*, *i*+5 hydrogen bonds (Figure 6).

## Comparison of representative open structure with the crystallographic structure of CYP-S-CO

We have previously observed that the ensemble of conformations available to CYP-S-CO in solution differs from the crystallographically-determined structure of CYP-S-CO (PDB entry 3CPP (14)) primarily in positioning of the B′ helix, B′-C loop and C helix (3,6), as well as the I helix “kink” discussed in the previous section. We attributed the observed differences in large part to the isomerization of the Ile 88-Pro 89 bond discussed earlier, and proposed that the crystallographic structures represented the “closed” (and catalytically

active) form of the enzyme. In solution, CYP101A1 samples more open conformers, which allow easier substrate access to the active site and favor substrate binding but are not catalytically active. Upon binding of the effector protein Pdx on the C helix, the Ile 88-Pro 89 bond isomerizes to the *cis* conformation, generating the "closed" catalytically active conformer. The results of the RDC-restrained simulations described here confirm most of these earlier observations, and the incorporation of experimental restraints gives more confidence to our interpretation of the results of dynamics simulations.

Differences in the B′ helix and B-C loop are shown in Figure 7. The B′ helix in REP is displaced by ~1 Å relative to its position in 3CPP towards the proximal face of the enzyme with heme positions almost superimposable. In turn, this displacement results in a similar outward movement of the C helix into the Pdx binding site, setting up the "pressure switch" for driving *trans*-*cis* isomerization of the Ile 88-Pro 89 amide that has been discussed previously (3). As with the previous simulations, the Ile 88-Pro 89 amide bond is distorted slightly from planarity (average dihedral angle from the 18 structures is ~166°, as compared to an expected value of 180° for a planar amide bond), thereby lowering the barrier to isomerization. The hydroxyl group of Tyr 29 (which stabilizes the *cis* conformation of the Ile 88-Pro 89 bond in 3CPP by hydrogen bonding to the Ile 88 C=O) is located 6.7 Å from the Ile 88 C=O in REP, and as in previous simulations, the Ile 88-Pro 89 amide is not solvated, destabilizing the $^{-}$O-C=N$^{+}$ resonance structure which favors planarity of the peptide.

Other differences more apparent in the current simulations include the loss of definition in the β5 sheet (Fig. 5). This structure is well-defined in 3CPP, but in the earliest iterations of RDC incorporation into the simulations, residues in this region showed the largest deviations between observed and calculated RDC values. The REP structure, while not displacing the β5 residues significantly, indicates that the sheet is less well organized in solution.

As expected in the open form of the enzyme, the presumed entrance to the active site is considerably wider in REP than in 3CPP. The Cδ carbon of Pro 89 at the base of the B′ helix is 6.2 Å from the Cγ of Pro 187 in the F-G loop in 3CPP, but 7.9 Å apart in REP. The δ1 methyl of Ile 395 in the β5 sheet is 7.6 Å from the Cα of Pro 89 in 3CPP, while they are separated by 9 Å in REP.

Another interesting difference between the closed and open conformations of CYP-S-CO with potential functional implications is found at the interface between the β3 sheet and K′ helix. Recent work in our laboratory has identified NH resonances assigned to Ser 325 and Gly 326 in the K′ helix as being sensitive to the nature of the bound substrate, and mutations in and around the K′ helix change substrate affinity and selectivity (Dang et al., submitted for publication). While remote from the active site, the K′ helix fixes the position of the β3 sheet adjacent to substrate in the active site. In 3CPP, the interactions between Gly 326 in the K′ helix and Met 323 in the β3 sheet are mediated by a shared water molecule, and the Ser 325 NH is not involved in any hydrogen bonding interactions (Fig. 8). Met 323 is anti-parallel in the β3 sheet to Val 295, which provides a primary substrate contact in 3CPP. In REP, the crystallographic water molecules are absent, and the carbonyl of Met 323 forms a helix-initiating *i*, *i*+4 hydrogen bond with the NH of Leu 327. The NH groups of Ser 325 and Gly 326 share a bifurcated hydrogen bond with the C=O of Gln 322 in the β3 sheet. This more direct and "drier" β3-K′ connection in the open form makes sense in light of both the sensitivity of Ser 325 and Gly 326 to substrate replacement and the fact the amide NH of Ser 325 exhibits slow exchange with solvent water in CYP-S-CO.

## Discussion

It has become apparent that dynamics are critical to enzyme function. That different conformational ensembles of enzymes are selected depending upon the presence of substrate, cofactor/effector or other environmental variables is now beyond question (21). Nevertheless, the static presentation of enzymes in crystallographic structures tends to reinforce (unconsciously, perhaps) a monolithic view of enzyme structure. While NMR is ideally suited for characterizing the dynamics of proteins, the drawbacks of molecular and spectral complexity mean that relatively few enzymes have been analyzed dynamically in detail (22,23). Recently we described the redox-dependent dynamics of CYP101A1 as a function of heme oxidation and ligation state, and showed that it was possible to obtain residue-specific information on local dynamics in this enzyme (24). Still, such analysis does not provide us with direct structural information, and for the reasons described above, we chose to use RDC-based methods to provide insight into the structural ensembles available to this enzyme in solution. We focus here on the carbonmonoxy- and camphor-bound reduced enzyme CYP-S-CO, because it is diamagnetic and we have extensive sequential assignments for this form. Furthermore, CYP-S-CO is iso-electronic with the catalytically relevant oxygen-bound open pre-catalysis form. As we have discussed previously, binding of effector and the concomitant conformational change (identified in previous work as isomerization of the Ile 88-Pro 89 bond at the N-terminus of the B′ helix) is required for catalytic competence of the enzyme, and crystalline CYP101A1 is catalytically competent (25).

### Comparison of REP structure with other open forms of CYP101A1

Recently, other open forms of CYP101A1 have been characterized crystallographically (26) in which the *trans* conformation of the Ile 88-Pro 89 peptide bond is not observed. Instead, the B′ helix is disordered, creating a larger opening into the active site. While we do not see evidence for the loss of the B′ helix in solution by NMR, this region is sufficiently dynamic in the oxidized enzyme that we cannot preclude the possibility that a conformer with a disordered B′ helix is fractionally (<10%) populated. One of the open CYP101A1 crystal structures was determined in the absence of potassium ion (which stabilizes the B′ helix in other structures), although incorporation of camphor appears to be sufficient to stabilize the B′ helix. We have found that in the absence of potassium ion, the changes induced by Pdx binding (which drives the isomerization of the Ile 88-Pro 89 bond to the *cis* form) are more complex, with multiple resonances being observed for residues in the B-B′ loop and B′ helix (15). We proposed that potassium ion is required to maintain the integrity of the structure between the B and C helices, particularly in the oxidized enzyme.

Comparison of representative structures of the *cis* and *trans* conformers of CYP-S-CO indicates that the *trans* conformer has a volume ~540 $Å^3$ larger than the *cis*, a volume increase of ~1%. This suggests to us that the *cis* conformer may be favored in the crystal due to increased packing efficiency. We note that a number of P450 structures have been determined in which the B′ or analogous helix is present and is initiated by a φ-Pro peptide, where φ is a hydrophobic residue such as Val, Phe, Ile or Leu (21). In all cases that we have identified (27–30), except CYP101A1 and the closely related CYP101D1 (31), the φ-Pro amide is found to be in the *trans* conformation.

## Conclusions

Residual dipolar couplings have become a standard addition to the traditional restraint set used to refine NMR-based structures of biological macromolecules. Unlike other restraints (NOEs, chemical shifts, J-couplings), which are wholly local in character, RDCs can relate disparate parts of the macromolecule to a single frame of reference, making them

particularly valuable for determining large or multi-domain structures in solution. Typically, these restraint sets are used in simulated annealing protocols, whereby structures are heated to high temperatures and then cooled slowly, overcoming local barriers to find a global minimum. In the present case, we are not seeking to identify a single structure, but instead have used RDCs to limit the search space in a standard solvated molecular dynamics simulation of an enzyme at realistic temperatures and solution conditions. We believe that the resulting ensemble of structures provide a reasonable sampling of the open conformations of CYP-S-CO under the conditions modeled. The current results support most of the previous conclusions that we reached using unrestrained dynamics concerning the differences between closed and open conformations of CYP-S-CO. We have also demonstrated a general method for applying RDCs to the characterization of the solution behavior of enzymes via simulation, and expect incorporation of RDC data to become a valuable adjunct to future studies of this type.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank T. Cheatham III and D. Case for their help in implementing RDC restraints and alignment tensor refinement in Amber.

## Abbreviations used

| | |
|---|---|
| **CYP101A1** | cytochrome $P450_{cam}$ |
| **CYP-S-CO** | reduced camphor-and carbonmonoxy-bound CYP101A1 |
| **NMR** | nuclear magnetic resonance |
| **PDB** | Protein Data Base |
| **Pdx** | putidaredoxin |
| **RDC** | residual dipolar couplings |
| **rmsd** | root-mean square deviation |
| **TROSY** | transverse relaxation optimized spectroscopy |

## References

1. Pauling L. Nature of forces between large molecules of biological interest. Nature 1948;161:707–709. [PubMed: 18860270]
2. Abeles, RH.; Frey, PA.; Jencks, WP. Biochemistry. Jones & Bartlett; Boston: 1992.
3. Asciutto EK, Madura JD, Pochapsky SS, OuYang B, Pochapsky TC. Structural and dynamic implications of an effector-induced backbone amide cis-trans isomerization in cytochrome $P450_{cam}$. J Mol Biol 2009;388:801–814. [PubMed: 19327368]
4. Pochapsky SS, Pochapsky TC, Wei JW. A model for effector activity in a highly specific biological electron transfer complex: The cytochrome P450(cam)-putidaredoxin couple. Biochemistry 2003;42:5649–5656. [PubMed: 12741821]
5. OuYang B, Pochapsky SS, Dang M, Pochapsky TC. A functional proline switch in cytochrome $P450_{cam}$. Structure 2008;16:916–923. [PubMed: 18513977]
6. Wei JY, Pochapsky TC, Pochapsky SS. Detection of a high-barrier conformational change in the active site of cytochrome $P450_{cam}$ upon binding of putidaredoxin. J Am Chem Soc 2005;127:6974–6976. [PubMed: 15884940]

7. Poulos TL, Finzel BC, Howard AJ. High-resolution crystal structure of cytochrome $P450_{cam}$. J Mol Biol 1987;195:687–700. [PubMed: 3656428]
8. Raag R, Poulos TL. Crystal structure of the carbon monoxide-substrate-cytochrome $P450_{cam}$ ternary complex. Biochemistry 1989;28:7586–7592. [PubMed: 2611203]
9. Poulos TL, Finzel BC, Howard AJ. High-resolution crystal structure of cytochrome $P450_{cam}$. J Mol Biol 1987;195:687–700. [PubMed: 3656428]
10. Nickerson DP, Wong LL. The dimerization of *Pseudomonas putida* cytochrome $P450_{cam}$: Practical consequences and engineering of a monomeric enzyme. Prot Eng 1997;10:1357–1361.
11. Rui LY, Pochapsky SS, Pochapsky TC. Comparison of the complexes formed by cytochrome $P450_{cam}$ with cytochrome b5 and putidaredoxin, two effectors of camphor hydroxylase activity. Biochemistry 2006;45:3887–3897. [PubMed: 16548516]
12. Ruckert M, Otting G. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J Am Chem Soc 2000;122:7793–7797.
13. Weigelt J. Single scan, sensitivity- and gradient-enhanced TROSY for multidimensional NMR experiments (vol 120, pg 10778, 1998). J Am Chem Soc 1998;120:12706–12706.
14. Raag R, Poulos TL. Crystal structure of the carbon monoxide-substrate-cytochrome P450cam ternary complex. Biochemistry 1989;28:7586–7592. [PubMed: 2611203]
15. OuYang B, Pochapsky SS, Pagani GM, Pochapsky TC. Specific effects of potassium ion binding on wild-type and L358P cytochrome $P450_{cam}$. Biochemistry 2006;45:14379–14388. [PubMed: 17128977]
16. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Crowley, M.; Walker, RC.; Zhang, W.; Merz, KM.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, KF.; Paesani, FV.; Wu, X.; Brozell, SR.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, DH.; Seetin, MG.; Sagui, C.; Babin, V.; Kollman, PA. AMBER. Vol. 10. Univ. of California; San Francisco: 2008.
17. Paulsen RB, Seth PP, Swayze EE, Griffey RH, Skalicky JJ, Cheatham TE, Davis DR. Inhibitor-induced structural change in the HCV IRES domain IIa RNA. Proc Natl Acad Sci USA 2010;107:7263–7268. [PubMed: 20360559]
18. Xia B, Tsui V, Case DA, Dyson HJ, Wright PE. Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. J Biomol NMR 2002;22:317–331. [PubMed: 12018480]
19. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF *chimera* - A visualization system for exploratory research and analysis. J Comp Chem 2004;25:1605–1612. [PubMed: 15264254]
20. DeLano, WL. The PyMOL Molecular Graphics System. 2008.
21. Pochapsky TC, Kazanis S, Dang M. Conformational plasticity and structure/function relationships in cytochromes P450. Antioxidants & Redox Signalling 2010;13:1273–1296.
22. Fraser JS, Clarkson MW, Degnan SC, Erion R, Kern D, Alber T. Hidden alternative structures of proline isomerase essential for catalysis. Nature 2009;462:669–U149. [PubMed: 19956261]
23. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hubner CG, Kern D. Intrinsic motions along an enzymatic reaction trajectory. Nature 2007;450:838–U813. [PubMed: 18026086]
24. Pochapsky SS, Dang M, OuYang B, Simorellis AK, Pochapsky TC. Redox-dependent dynamics in cytochrome $P450_{cam}$. Biochemistry 2009;48:4254–4261. [PubMed: 19366254]
25. Schlichting I, Berendzen J, Chu K, Stock AM, Maves SA, Benson DE, Sweet RM, Ringe D, Petsko GA, Sligar SG. The catalytic pathway of cytochrome $P450_{cam}$ at atomic resolution. Science 2000;287:1615–1622. [PubMed: 10698731]
26. Lee YT, Wilson RF, Rupniewski I, Goodin DB. $P450_{cam}$ visits an open conformation in the absence of substrate. Biochemistry 2010;49:3412–3419. [PubMed: 20297780]
27. Cupp-Vickery JR, Poulos TL. Structure of cytochrome P450eryF involved in erythromycin biosynthesis. Nat Struct Biol 1995;2:144–153. [PubMed: 7749919]
28. Yano JK, Koo LS, Schuller DJ, Li H, Ortiz de Montellano PR, Poulos TL. Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*. J Biol Chem 2000;275:31086–31092. [PubMed: 10859321]

29. Nagano S, Li HY, Shimizu H, Nishida C, Ogura H, de Montellano PRO, Poulos TL. Crystal structures of epothilone D-bound, epothilone B-bound, and substrate-free forms of cytochrome P450epoK. J Biol Chem 2003;278:44886–44893. [PubMed: 12933799]

30. Belin P, Le Du MH, Fielding A, Lequin O, Jacquet M, Charbonnier JB, Lecoq A, Thai R, Courcon M, Masson C, Dugave C, Genet R, Pernodet JL, Gondry M. Identification and structural basis of the reaction catalyzed by CYP121, an essential cytochrome P450 in *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 2009;106:7426–7431. [PubMed: 19416919]

31. Yang W, Bell SG, Wang H, Zhou W, Hoskins N, Dale A, Bartlam M, Wong L-L, Rao Z. Molecular Characterization of a Class I P450 Electron Transfer System from *Novosphingobium aromaticivorans* DSM12444. J Biol Chem 2010;285:27372–27384. [PubMed: 20576606]

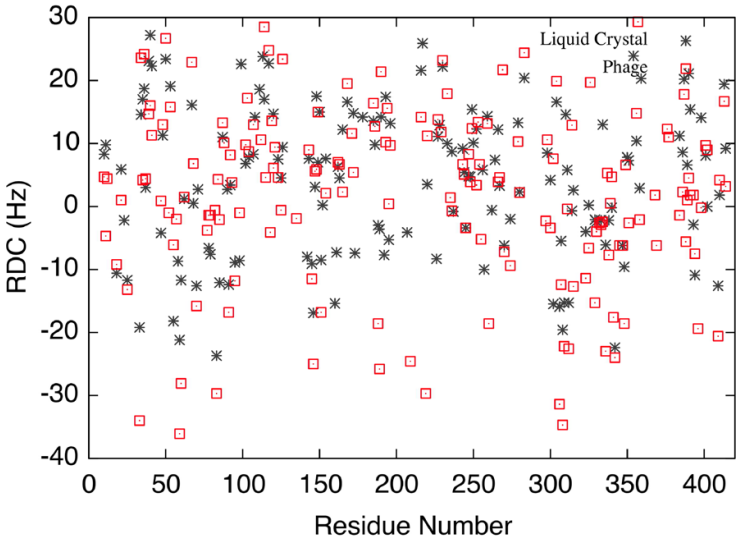


**Figure 1.**

Measured $H_N$-N residual dipolar couplings (RDCs) resulting from an iterative procedure applied in both media: liquid crystal (C12E5) and bacteriophage *Pf1*. RDCs measured in C12E5 medium are shown as (x), those measured using aligned bacteriophage are indicated by (□).

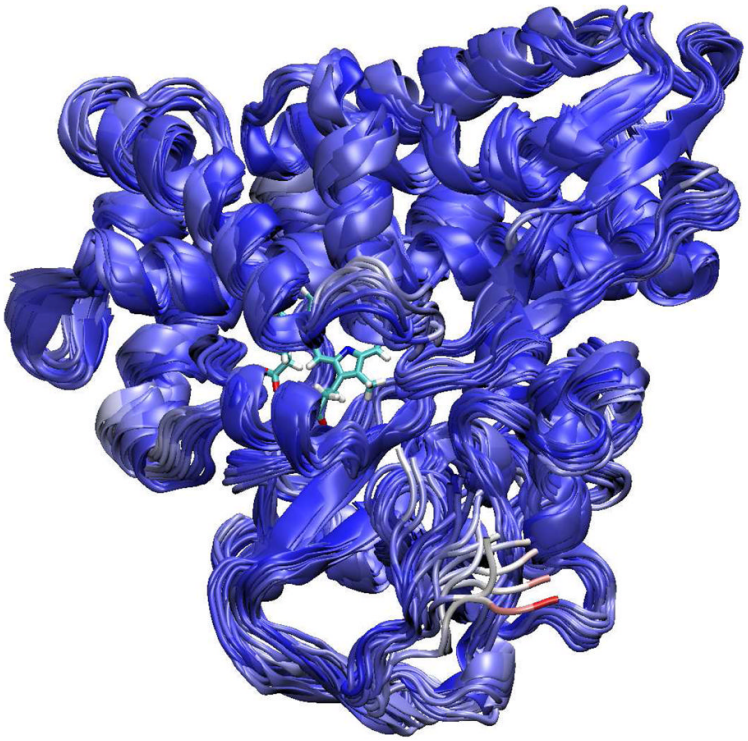

**Figure 2.**
The 18 lowest restraint energy structures superimposed, colored by rmsd from blue (lower rmsd) to red (greater rmsd).

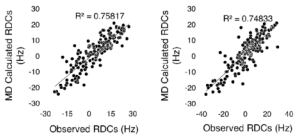


**Figure 3.**
Experimental RDCs in C12E5 medium versus calculated (left), and experimental RDCs versus calculated in *pf1* medium (right) from 1 ns simulation for representative structure REP.

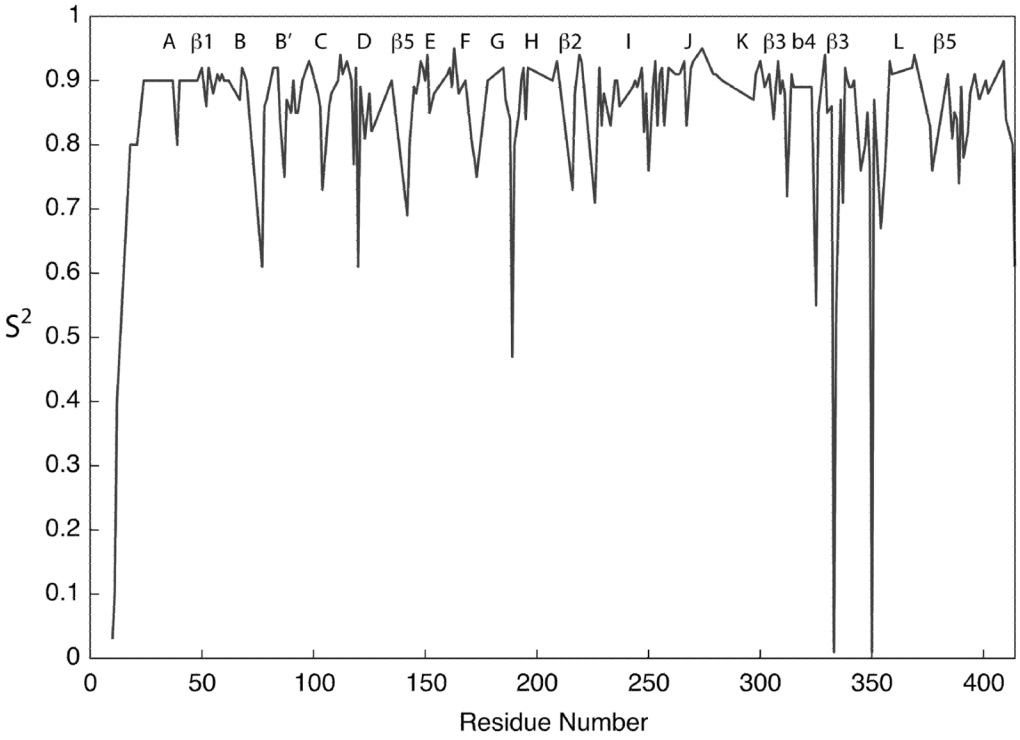


**Figure 4.**

Order parameters $S^2$ of $H_N$-N vectors calculated from the simulation plotted as a function of sequence. Secondary structural features are indicated at the top of the graph.

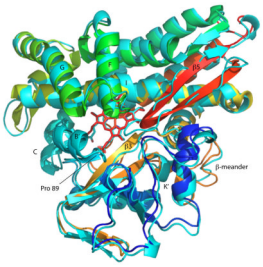


**Figure 5.**

Alignment of crystallographic structure 3CPP of CYP-S-CO with representative structure (REP) from RDC-restrained dynamics. 3CPP is shown in rainbow from N-terminus (blue) to C-terminus (red). REP structure is shown in cyan for all regions. Structural features and residues discussed in the text are labeled. Alignment yields an overall RMSD of 1.02 Å for 405 Cα carbons. Figure was generated using PyMOL (20).

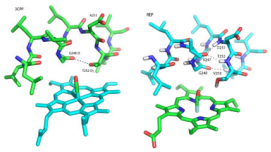

**Figure 6.**
Comparison of the I helix “kink” in 3CPP (left) and current structure REP (right). Hydrogen bonds stabilizing the gap in the helix in both cases are shown as dotted lines. For visibility of hydrogen bonds, the structures are shown in slightly different orientations. Figure was generated using PyMOL (20).

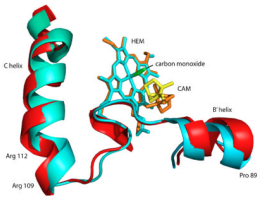


**Figure 7.**

Comparison of the B′ and C helices from aligned structures 3CPP and current structure REP. 3CPP peptide is shown in cyan, and REP peptide is shown in red. The Pdx binding site on the C helix is defined by the positions of Arg 109 and Arg 112. The heme prosthetic groups (cyan for 3CPP, orange for REP) and camphor (yellow for 3CPP, orange for REP) are also shown. Figure was generated using PyMOL (20).

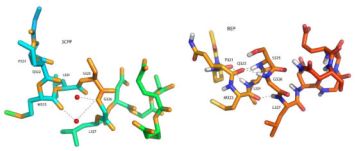

**Figure 8.**
Comparison of the junction between the β3 sheet and K′ helix in 3CPP and current structure REP. Structures are rotated slightly with respect to each other for clarity. Left, 3CPP, a water molecule (red sphere) shares hydrogen bonds (indicated by dotted lines) with the carbonyl of Met 323 in the β3 sheet and Gly 326 in the K′ helix. No hydrogen bonds are observed to the NH of Ser 325. In REP (right), water is excluded from the junction, and multiple hydrogen bonds are observed at the sheet-helix junction. See text for details. Figure was generated using PyMOL (20).