1664 *Biochemistry* **2011,** *50,* 1664–1671
DOI: 10.1021/bi101820d

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Experimentally Restrained Molecular Dynamics Simulations for Characterizing the Open States of Cytochrome $P450_{cam}$ [†,‡]

Eliana K. Asciutto,[§] Marina Dang,[∥] Susan Sondej Pochapsky,[∥] Jeffry D. Madura,[§] and Thomas C. Pochapsky*[,∥,⊥]

[§]*Department of Chemistry and Biochemistry, Center for Computational Sciences, Duquesne University, Pittsburgh, Pennsylvania 15282, United States,* [∥]*Department of Chemistry, MS 015, Brandeis University, 415 South Street, Waltham, Massachusetts 02454-9110, United States, and* [⊥]*Rosenstiel Basic Medical Sciences Research Institute, Brandeis University, Waltham, Massachusetts 02454-9110, United States*

Received November 12, 2010; Revised Manuscript Received January 19, 2011

ABSTRACT: Residual dipolar couplings (RDCs) 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 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 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)[1] 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 forms 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 sought 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

[†]This work was supported in part by a grant from the U.S. Public Health Service (R01-GM44191, T.C.P.). E.K.A. and J.D.M. acknowledge support from U.S. Public Health Service Grant R01DA027806, Department of Education Grant P116Z090309, and National Science Foundation support through Teragrid (MCB060061N).

[‡]The REP structure discussed in this paper has been deposited as Protein Data Bank entry 2L8M. NMR chemical shifts and RDC restraints are deposited as BMRB entry 17415.

*To whom correspondence should be addressed: Department of Chemistry, MS 015, Brandeis University, 415 South St., Waltham, MA 02454-9110. Telephone: (781) 736-2559. Fax: (781) 736-2516. E-mail: pochapsk@brandeis.edu.

[1]Abbreviations: CYP101A1, cytochrome $P450_{cam}$; CYP-S-CO, reduced camphor- and carbonmonoxy-bound CYP101A1; NMR, nuclear magnetic resonance; PDB, Protein Data Bank; Pdx, putidaredoxin; RDCs, residual dipolar couplings; rmsd, root-mean-square deviation; TROSY, transverse relaxation optimized spectroscopy.

 Published on Web 01/25/2011

from NOESY spectra obtained using otherwise perdeuterated CYP101A1 samples. Because of the 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 *Escherichia 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 medium 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% with an 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, and 2 mM camphor in a 90:10 H_2O/D_2O mixture. 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) was prepared in a gastight vial using degassed NMR buffer [50 mM Tris-HCl (pH 7.4), 100 mM KCl, and 2 mM camphor in a 90:10 H_2O/D_2O mixture]. 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, gastight vial. The atmosphere in the vial was flushed with CO for 3 min, alternating with vigorous vortexing for a few seconds. Using a metal syringe, 7 μL of *n*-hexanol (anhydrous, ≥99%, Sigma-Aldrich, Milwaukee, WI) was introduced in small aliquots, as little 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 and then turned completely clear with no bubbles. The final mixture appears opalescent, with a C12E5:hexanol molar ratio of 0.79. In an argon-filled glovebag, 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 and then transferred into a susceptibility-matched NMR tube (Shigemi, Inc.). 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 ^{2}H signal (~23 Hz after 0.5 h).

Alignment Using Filamentous Phage Pf1. Bacteriophage *Pf1* (Asla Biotech Ltd.) 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, and 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 1 M Tris buffer (pH 7.4) was added; 400 μL of 0.5 mM CYP-S-CO was added to the vial containing the phage and reducing agent. The mixture was stirred gently and then transferred to 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 ^{2}H signal (~9.5 Hz after 0.5 h).

NMR Experiments. A combination of TROSY and semi-TROSY data sets was acquired for both aligned and unaligned samples of reduced camphor and CO-bound CYP101A1 (CYP-S-CO). The pulse sequences were adapted from ref *13*. Spectra were recorded as data sets with 1024 (^{1}H) × 512 (^{15}N) complex points at 298 K on an 18.8 T (800 MHz ^{1}H) Bruker Avance NMR spectrometer equipped with a cryogenic probe. The interleaved data were separated prior to Fourier transformation, yielding two two-dimensional data sets with 1024 (^{1}H) × 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 optimal 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 (Ile88−Pro89) that precedes helix B′ 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

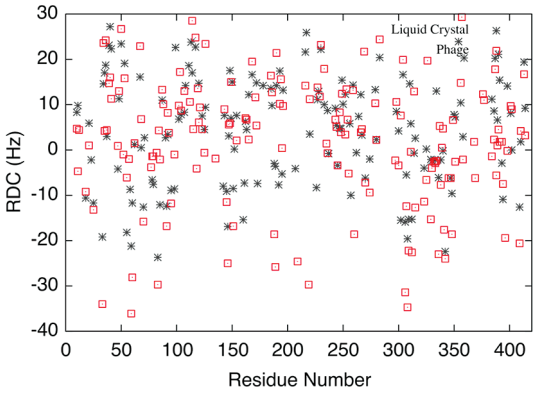

FIGURE 1: Measured H_N−N residual dipolar couplings (RDCs) resulting from an iterative procedure applied in both media: liquid crystal (C12E5) and bacteriophage *Pf1*. RDCs measured in C12E5 medium are shown (×), as are those measured using aligned bacteriophage (red squares).

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 16935 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 work presented here, we took snapshots from the 3 ns simulation as starting structures.

One hundred fifty-six 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 because of 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 indices u and l refer to the upper and lower limits, respectively. 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 three cycles of constant volume dynamics for 30 ps, gradually heating from 0 to 50 K, from 50 to 100 K, and from 100 to 300 K, with 5 kcal mol^{-1} $Å^{-2}$ harmonic restraints relative to the starting structure. A 1 ps time constant was used for coupling to the heat bath. Heating was followed by two cycles of constant pressure dynamics for 30 ps with a pressure relaxation time of 0.2 ps, which was then increased to 1.0 ps, and the harmonic constraints on the protein were released. Once the pressure was equilibrated, another 60 ps constant volume MD was performed. Last, 1 ns of productive constant volume dynamics at 300 K was performed. The described procedure was applied to nine 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*).

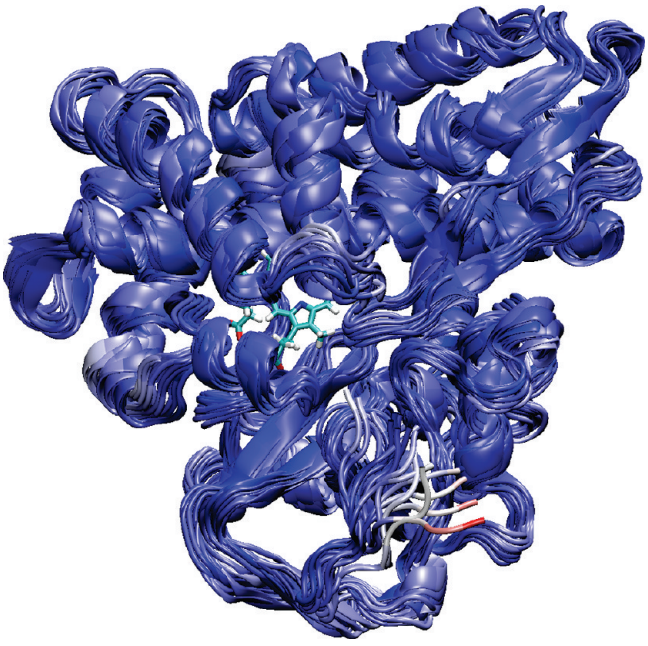

FIGURE 2: Eighteen lowest-restraint energy structures superimposed, colored by rmsd from blue (lower rmsd) to red (greater rmsd).

This simulation protocol was applied iteratively to obtain self-consistent results. Preliminary RDC data sets were applied as restraints, followed by analysis of the results in terms of the 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 reached 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 five 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 is shown in Figure 1. As one can see 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 nine initial structures were refined with the procedure described in Materials and Methods. For each simulation, we generated trajectories saving frames every 1 ps, giving a total of 1000 snapshots per simulation. From these 1000 snapshots, the two structures with the lowest restraint violation energies were selected, yielding an ensemble of 18 structures (Figure 2). The backbone rmsd between the structures is 1.298 Å. The 18 structures selected are representative of the ensemble, because 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

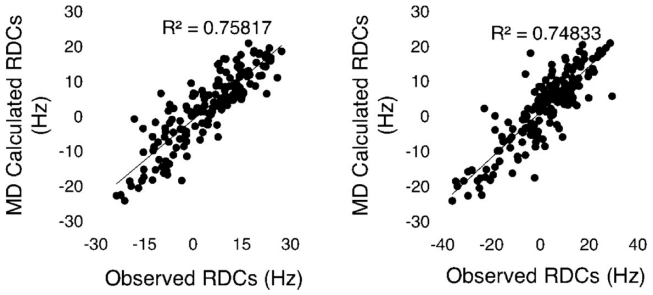

FIGURE 3: Experimental RDCs in C12E5 medium vs calculated (left) and experimental RDCs vs calculated in *pf1* medium (right) from a 1 ns simulation for representative structure REP.

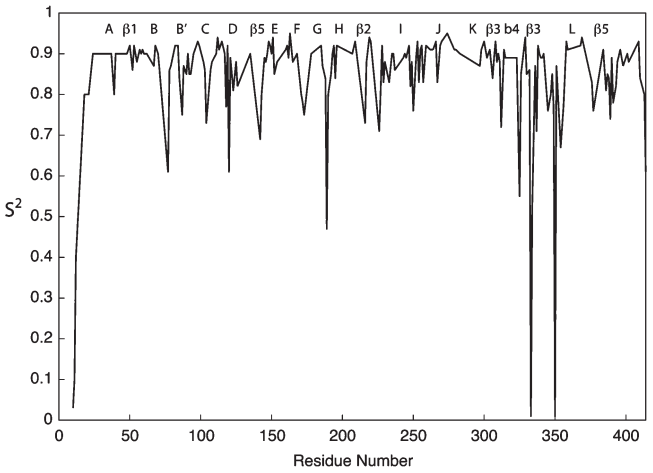

FIGURE 4: Order parameters (S^2) of H_N-N vectors calculated from the simulation plotted as a function of sequence. Secondary structural features are denoted atop the graph.

simulation (Table S1 of the Supporting Information). The first 10 residues display the greatest rmsd values. Other regions showing some discrepancies are the region between helices F and G, the end of helix B (Asp77), some residues of helix C, especially Met121, and 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 five 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 1 ns 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 exhibiting 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 are plotted in Figure 4. As expected, the first

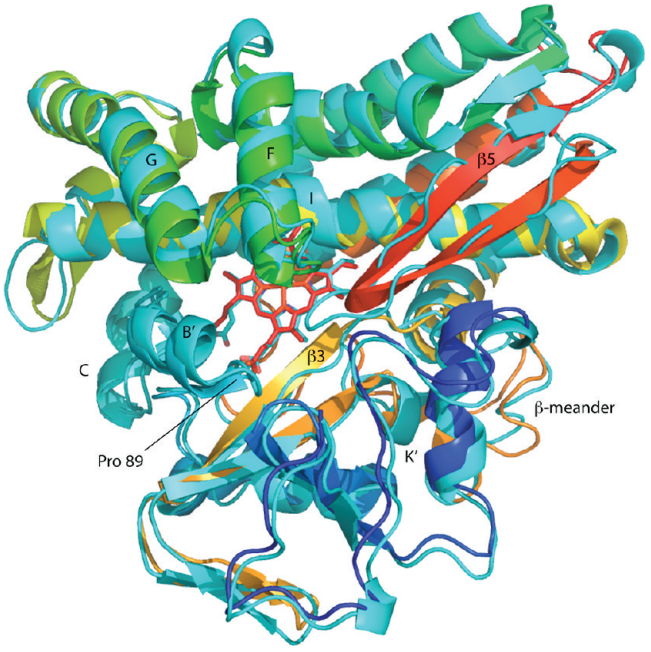

FIGURE 5: Alignment of the crystallographic structure of 3CPP of CYP-S-CO with representative structure (REP) from RDC-restrained dynamics. 3CPP is shown in rainbow from the N-terminus (blue) to the C-terminus (red). REP structure is colored cyan for all regions. Structural features and residues discussed in the text are labeled. Alignment yields an overall rmsd of 1.02 Å for 405 Cα atoms. This figure was generated using PyMOL (*20*).

10 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 helix K′ and precedes the loop containing the heme axial ligand Cys357. 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 the lowest restraint energy was identified, and the coordinates for this structure (REP) are deposited in the PDB and can be found in the Supporting Information.

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 Amber10, 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 of < 0.1 for both cases and in both (C12E5 and *pfl*) media. To ensure that the RDC restraints do not introduce local strain into bond angles, we compared the bond angle energy term for restrained simulations against that from unrestrained simulations from ref *3*. We found that the average bond angle energy does not change significantly (~1%), which is much smaller than average thermal fluctuations (~120 kcal/mol). Clearly, the presence of the RDC restraints changes the results of the simulations, at least

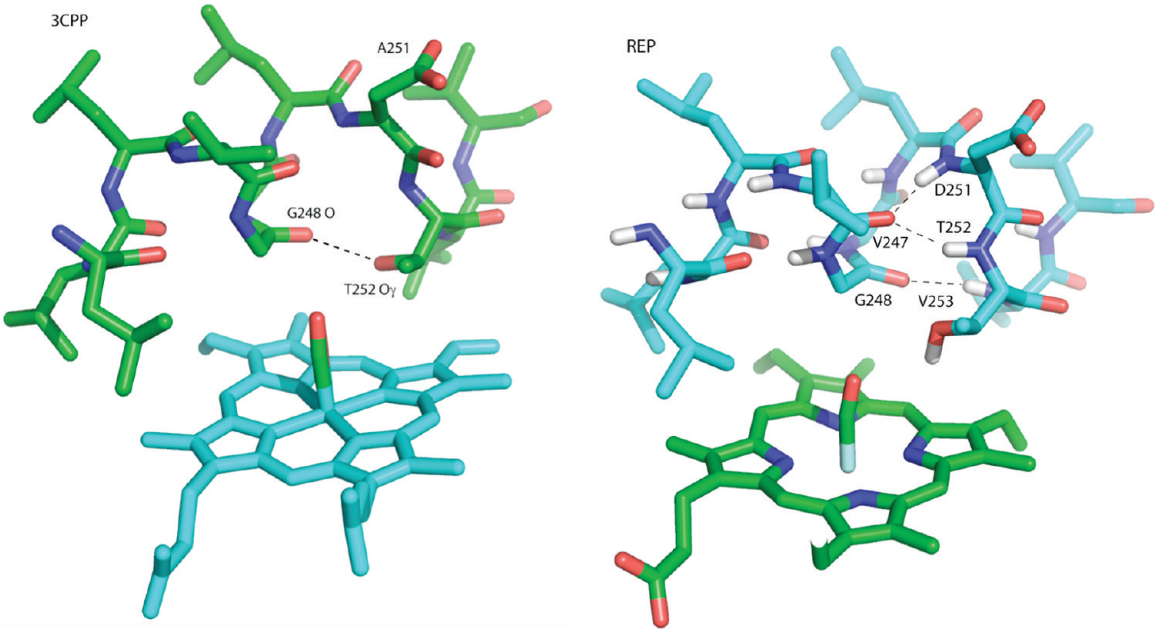

FIGURE 6: Comparison of the I helix kink in 3CPP (left) and the current structure REP (right). Hydrogen bonds stabilizing the gap in the helix in both cases are shown as dotted lines. So that hydrogen bonds could be visible, the structures are shown in slightly different orientations. This figure was generated using PyMOL (*20*).

with regard to N−H bond vectors. Structurally, the beginning of the B−B′ loop (residues 77−82) also shows considerable difference with respect to 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 because of the absence of the expected hydrogen bond between the NH group of Thr252 and the carbonyl oxygen of Gly248. Instead, the Gly248 carbonyl accepts a hydrogen bond from the γ-OH group of the Thr252 side chain (Figure 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 Gly248 C═O···H—N Thr252 hydrogen bond. In the calculations presented here, while the Thr252 side chain γ-OH group remains uninvolved in main chain hydrogen bonding, the Gly248 C═O group forms a well-developed hydrogen bond with the NH group of Val253, and the NH groups of Asp251 and Thr252 form a bifurcated hydrogen bond with the carbonyl oxygen of Val238. 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 the 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 the positioning of helix B′, the B′−C loop, and helix C (*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 Ile88−Pro89 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 access of the substrate 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 Ile88−Pro89 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.

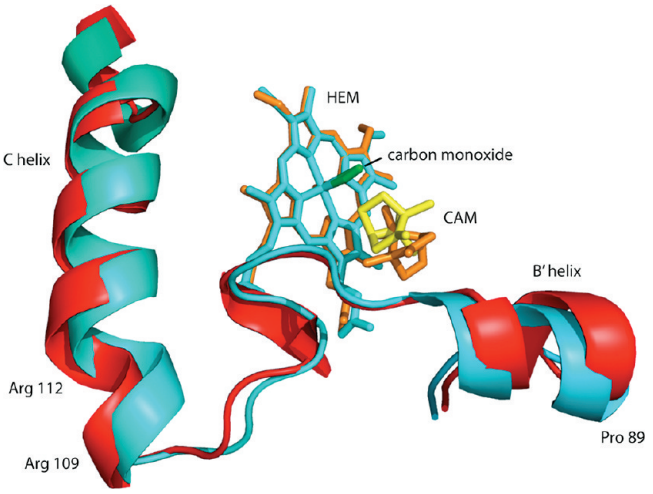

FIGURE 7: Comparison of helices B′ and C from aligned structures 3CPP and current structure REP. The 3CPP peptide is colored cyan, and the REP peptide is colored red. The Pdx binding site on helix C is defined by the positions of Arg109 and Arg112. The heme prosthetic groups (cyan for 3CPP and orange for REP) and camphor (yellow for 3CPP and orange for REP) are also shown. This figure was generated using PyMOL (*20*).

Differences in helix B′ and the B−C loop are shown in Figure 7. Helix B′ in REP is displaced by ~1 Å relative to its position in

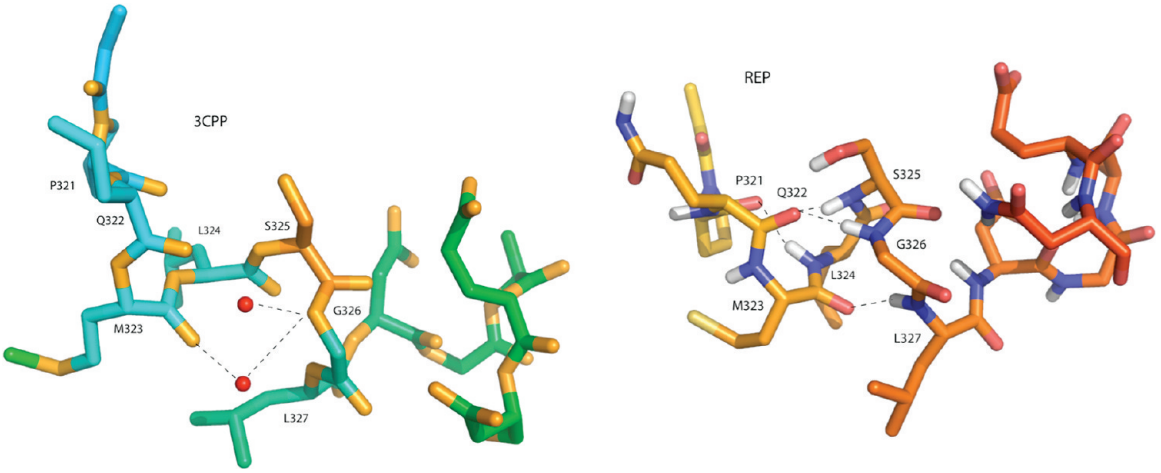

FIGURE 8: Comparison of the junction between sheet β3 and helix K′ in 3CPP and the current structure REP. Structures are rotated slightly with respect to each other for the sake of clarity. In the left panel (3CPP), a water molecule (red sphere) shares hydrogen bonds (dotted lines) with the carbonyl of Met323 in sheet β3 and Gly326 in helix K′. No hydrogen bonds to the NH group of Ser325 are observed. In REP (right), water is excluded from the junction, and multiple hydrogen bonds are observed at the sheet−helix junction. See the text for details. This figure was generated using PyMOL (*20*).

3CPP toward the proximal face of the enzyme, with heme positions almost superimposable. In turn, this displacement results in a similar outward movement of helix C into the Pdx binding site, setting up the "pressure switch" for driving *trans*−*cis* isomerization of the Ile88−Pro89 amide that has been discussed previously (*3*). As with the previous simulations, the Ile88−Pro89 amide bond is distorted slightly from planarity (average dihedral angle from the 18 structures of ~166°, compared to an expected value of 180° for a planar amide bond), thereby lowering the barrier to isomerization. The hydroxyl group of Tyr29 (which stabilizes the *cis* conformation of the Ile88−Pro89 bond in 3CPP by hydrogen bonding to the Ile88 C=O group) is 6.7 Å from the Ile88 C=O group in REP, and as in previous simulations, the Ile88−Pro89 amide is not solvated, destabilizing the $^{-}$O—C=N^{+} resonance structure that favors planarity of the peptide.

Other differences more apparent in the current simulations include the loss of definition in sheet β5 (Figure 5). This structure is well-defined in 3CPP, but in the earliest iterations of incorporation of RDCs 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 C_δ atom of Pro89 at the base of helix B′ is 6.2 Å from the C_γ atom of Pro187 in the F−G loop in 3CPP, but those atoms are 7.9 Å apart in REP. The δ1 methyl of Ile395 in sheet β5 is 7.6 Å from the Cα atom 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 sheet β3 and helix K′. Recent work in our laboratory has identified NH resonances assigned to Ser325 and Gly326 in helix K′ as being sensitive to the nature of the bound substrate, and mutations in and around helix K′ change substrate affinity and selectivity (M. Dang et al., manuscript submitted for publication). While remote from the active site, helix K′ fixes the position of sheet β3 adjacent to the substrate in the active site. In 3CPP, the interactions between Gly326 in helix K′ and Met323 in sheet β3 are mediated by a shared water molecule, and the Ser325 NH group is not involved in any hydrogen bonding interactions (Figure 8). Met323 is antiparallel in sheet β3 to Val295, which provides a primary substrate contact in 3CPP. In REP, the crystallographic water molecules are absent, and the carbonyl of Met323 forms a helix-initiating *i*, *i* + 4 hydrogen bond with the NH group of Leu327. The NH groups of Ser325 and Gly326 share a bifurcated hydrogen bond with the C=O group of Gln322 in sheet β3. This more direct and "drier" β3−K′ connection in the open form makes sense in light of both the sensitivity of Ser325 and Gly326 to substrate replacement and the fact that the amide NH group of Ser325 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 to 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 about 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 precatalysis form. As we have discussed previously, binding of effector and the concomitant conformational change (identified in previous work as isomerization of the Ile88−Pro89 bond at the N-terminus of helix B′) are 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 in which the *trans* conformation of the Ile88−Pro89 peptide bond was not observed have been characterized crystallographically (*26*). Instead, helix B′ is disordered, creating a larger opening into the active site. While we do not see evidence of the loss of helix B′ 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 helix B′ in other structures), although incorporation of camphor appears to be sufficient to stabilize helix B′. We have found that in the absence of potassium ion, the changes induced by Pdx binding (which drives the isomerization of the Ile88−Pro89 bond to the *cis* form) are more complex, with multiple resonances being observed for residues in the B−B′ loop and helix B′ (*15*). We proposed that potassium ion is required to maintain the integrity of the structure between helices B and C, 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 that of the *cis* form, a volume increase of ~1%. This suggests to us that the *cis* conformer may be favored in the crystal because of increased packing efficiency. We note that a number of P450 structures have been determined in which helix B′ or an 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, and *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 multidomain 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 this 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 under realistic solution conditions. We believe that the resulting ensemble of structures provides a reasonable sampling of the open conformations of CYP-S-CO under the conditions modeled. These 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.

ACKNOWLEDGMENT

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

SUPPORTING INFORMATION AVAILABLE

PDB format structure file for the REP structure, RDC restraints in Amber format, and two tables discussed in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Pauling, L. (1948) Nature of forces between large molecules of biological interest. *Nature 161*, 707–709.
2. Abeles, R. H., Frey, P. A., and Jencks, W. P. (1992) Biochemistry, Jones & Bartlett, Boston.
3. Asciutto, E. K., Madura, J. D., Pochapsky, S. S., OuYang, B., and Pochapsky, T. C. (2009) Structural and dynamic implications of an effector-induced backbone amide cis-trans isomerization in cytochrome P450$_{cam}$. *J. Mol. Biol. 388*, 801–814.
4. Pochapsky, S. S., Pochapsky, T. C., and Wei, J. W. (2003) A model for effector activity in a highly specific biological electron transfer complex: The cytochrome P450$_{cam}$-putidaredoxin couple. *Biochemistry 42*, 5649–5656.
5. OuYang, B., Pochapsky, S. S., Dang, M., and Pochapsky, T. C. (2008) A functional proline switch in cytochrome P450$_{cam}$. *Structure 16*, 916–923.
6. Wei, J. Y., Pochapsky, T. C., and Pochapsky, S. S. (2005) Detection of a high-barrier conformational change in the active site of cytochrome P450$_{cam}$ upon binding of putidaredoxin. *J. Am. Chem. Soc. 127*, 6974–6976.
7. Poulos, T. L., Finzel, B. C., and Howard, A. J. (1987) High-resolution crystal structure of cytochrome P450$_{cam}$. *J. Mol. Biol. 195*, 687–700.
8. Raag, R., and Poulos, T. L. (1989) Crystal structure of the carbon monoxide-substrate-cytochrome P450$_{cam}$ ternary complex. *Biochemistry 28*, 7586–7592.
9. Poulos, T. L., Finzel, B. C., and Howard, A. J. (1987) High-resolution crystal structure of cytochrome P450$_{cam}$. *J. Mol. Biol. 195*, 687–700.
10. Nickerson, D. P., and Wong, L. L. (1997) The dimerization of *Pseudomonas putida* cytochrome P450$_{cam}$: Practical consequences and engineering of a monomeric enzyme. *Protein Eng. 10*, 1357–1361.
11. Rui, L. Y., Pochapsky, S. S., and Pochapsky, T. C. (2006) Comparison of the complexes formed by cytochrome P450$_{cam}$ with cytochrome b5 and putidaredoxin, two effectors of camphor hydroxylase activity. *Biochemistry 45*, 3887–3897.
12. Ruckert, M., and Otting, G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc. 122*, 7793–7797.
13. Weigelt, J. (1998) Single scan, sensitivity- and gradient-enhanced TROSY for multidimensional NMR experiments. *J. Am. Chem. Soc. 120*, 10778–10779, 12706 (Erratum).
14. Raag, R., and Poulos, T. L. (1989) Crystal structure of the carbon monoxide-substrate-cytochrome P450cam ternary complex. *Biochemistry 28*, 7586–7592.
15. OuYang, B., Pochapsky, S. S., Pagani, G. M., and Pochapsky, T. C. (2006) Specific effects of potassium ion binding on wild-type and L358P cytochrome P450$_{cam}$. *Biochemistry 45*, 14379–14388.
16. Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F. V., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., and Kollman, P. A. (2008) Amber10, University of California, San Francisco.
17. Paulsen, R. B., Seth, P. P., Swayze, E. E., Griffey, R. H., Skalicky, J. J., Cheatham, T. E., and Davis, D. R. (2010) Inhibitor-induced structural change in the HCV IRES domain IIa RNA. *Proc. Natl. Acad. Sci. U.S.A. 107*, 7263–7268.
18. Xia, B., Tsui, V., Case, D. A., Dyson, H. J., and Wright, P. E. (2002) Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. *J. Biomol. NMR 22*, 317–331.

19. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF *chimera*: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.
20. DeLano, W. L. (2008) The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA.
21. Pochapsky, T. C., Kazanis, S., and Dang, M. (2010) Conformational plasticity and structure/function relationships in cytochromes P450. *Antioxid. Redox Signaling 13*, 1273–1296.
22. Fraser, J. S., Clarkson, M. W., Degnan, S. C., Erion, R., Kern, D., and Alber, T. (2009) Hidden alternative structures of proline isomerase essential for catalysis. *Nature 462*, 669–673.
23. Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., Karplus, M., Hubner, C. G., and Kern, D. (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature 450*, 838–844.
24. Pochapsky, S. S., Dang, M., OuYang, B., Simorellis, A. K., and Pochapsky, T. C. (2009) Redox-dependent dynamics in cytochrome $P450_{cam}$. *Biochemistry 48*, 4254–4261.
25. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) The catalytic pathway of cytochrome $P450_{cam}$ at atomic resolution. *Science 287*, 1615–1622.
26. Lee, Y. T., Wilson, R. F., Rupniewski, I., and Goodin, D. B. (2010) $P450_{cam}$ visits an open conformation in the absence of substrate. *Biochemistry 49*, 3412–3419.
27. Cupp-Vickery, J. R., and Poulos, T. L. (1995) Structure of cytochrome P450eryF involved in erythromycin biosynthesis. *Nat. Struct. Biol. 2*, 144–153.
28. Yano, J. K., Koo, L. S., Schuller, D. J., Li, H., Ortiz de Montellano, P. R., and Poulos, T. L. (2000) Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*. *J. Biol. Chem. 275*, 31086–31092.
29. Nagano, S., Li, H. Y., Shimizu, H., Nishida, C., Ogura, H., de Montellano, P. R. O., and Poulos, T. L. (2003) Crystal structures of epothilone D-bound, epothilone B-bound, and substrate-free forms of cytochrome P450epoK. *J. Biol. Chem. 278*, 44886–44893.
30. Belin, P., Le Du, M. H., Fielding, A., Lequin, O., Jacquet, M., Charbonnier, J. B., Lecoq, A., Thai, R., Courcon, M., Masson, C., Dugave, C., Genet, R., Pernodet, J. L., and Gondry, M. (2009) Identification and structural basis of the reaction catalyzed by CYP121, an essential cytochrome P450 in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A. 106*, 7426–7431.
31. Yang, W., Bell, S. G., Wang, H., Zhou, W., Hoskins, N., Dale, A., Bartlam, M., Wong, L.-L., and Rao, Z. (2010) Molecular Characterization of a Class I P450 Electron Transfer System from *Novosphingobium aromaticivorans* DSM12444. *J. Biol. Chem. 285*, 27372–27384.