A Molecular Dynamics Study of Fe_2S_2 Putidaredoxin: Multiple Conformations of the C-Terminal Region

Adrian E. Roitberg
National Institute for Standards and Technology, Bioprocess Engineering Group, Gaithersburg, Maryland 20899 USA

ABSTRACT Putidaredoxin (Pdx) plays an essential role as an electron donor and effector in the biochemical cycle involving cytochrome P450cam. Only recently has an NMR-derived structure for this protein been published, but because of the presence of a paramagnetic Fe_2S_2 center, the NMR assignment could not be completed for residues within a region of 8 Å around the active site. That region was modeled by homology with a related protein. The structural refinement for those experiments was done in vacuum, without the use of electrostatic terms in the force field. The present manuscript will describe and discuss a series of long-time, unrestrained, solution molecular dynamic runs for this system. Results will be presented that construct a molecular-level picture that rationalizes experimental results concerning the conformation and mobility of the C-terminal residue Trp^{106}. At least two different conformers are found for this residue during the simulations. The time scale for interconversion between them is found to be in the subnanosecond regime. The results presented here open the possibility for studying binding and electron transfer between Pdx and P450cam, in a framework that allows for dynamical information to be used during the computational process, instead of the single structures deposited on the protein data base.

INTRODUCTION

The cytochrome P450 system presents an example of exquisite control of the production of metabolic compounds. Pathways involving this enzyme provide a very interesting biochemical system for investigating problems related to electron transfer, substrate binding, redox, and protein-protein interactions.

The *in vivo* function of the P450cam enzymatic complex is to catalyze the stereoselective oxidation of camphor to 5-exo-hydroxycamphor. It has recently been shown that P450cam can also metabolize a number of compounds (not all related to camphor). The reactions that this enzyme is capable of catalyzing include dehalogenation, carbon hydroxylation, epoxidation, and sulfoxidation (Gelb et al., 1982; Collins and Loew, 1988; Jones et al., 1992; Fruetel et al., 1992; Koe and Vilker, 1993; Li and Wackett, 1993; Fruetel et al., 1994; Loida et al., 1995; Grayson et al., 1996). The driving force behind our interest in these protein complexes is to use their catalytic power for practical purposes, such as biocatalysis.

The full catalytic cycle of P450cam starts with reducing power provided by NADH, which in turn reduces the FAD-containing protein putidaredoxin reductase (PdR). PdR then transfers an electron to putidaredoxin (Pdx). The final step in this cascade involves a complex formation between Pdx and P450cam and a subsequent electron transfer. Two electrons are required per turnover, transferred in two discrete steps.

Received for publication 13 March 1997 and in final form 25 June 1997.
Address reprint requests to Dr. Adrian E. Roitberg, Bioprocess Engineering Group, National Institute for Standards and Technology, Building 222-A353, Gaithersburg, MD 20899. Tel.: 301-975-4469; Fax: 301-975-5449; E-mail: adrian@nist.gov.

0006-3495/97/10/2138/11 $2.00

Of the three major proteins in this cycle, only the structure of P450cam has been elucidated by X-ray methodology (Poulos et al., 1987). There has been no reported X-ray structure for Pdx or PdR. Recently a model has been proposed for the structure of Pdx by NMR methods (Pochapsky et al., 1994). A discussion of this structure and the methods used for its elucidation will be presented in the rest of this paper.

Pdx is a 106-amino acid protein (MW 11,594) containing a Fe_2S_2. The Fe_2S_2 of Pdx has unique electronic structure properties. The most important one is the fact that the oxidized protein has a silent EPR signal at low temperatures, indicative of antiferromagnetic coupling between the two iron atoms, each one with its own spin, $S = 5/2$. However, as the temperature is raised, an EPR signal appears, and the cluster becomes paramagnetic (Poe et al., 1971). The main effect derived from the presence of a paramagnetic center is to shift and broaden the NMR signals. All nuclei within 8 Å of the paramagnetic center will be affected, and thus dependable structural data for those regions cannot be extracted from NMR experiments (Chae and Markley, 1995).

Recently the structure of Pdx has been elucidated by multidimensional ^1H NMR (Pochapsky and Mei, 1991; Mei et al., 1992; Pochapsky et al., 1994; Lyons et al., 1996). Despite the problems associated with the paramagnetic center, Pochapsky's group successfully found 878 NOE-derived distance restraints, 66 ϕ angular constrains (from NH—C_αH coupling constants), and a number of paramagnetic broadening constraints. There are data to support the idea that the regions close to the metallic cluster are homologous among a series of ferredoxins. The overall sequence homology between Pdx and the ferredoxins of *Spirulinas platensis* and *Anabaena* is low (around 10%), but the metal cluster regions have large homology. This enabled Po-

chapsky to model the active site region of Pdx based on homology with *Anabaena* 7120 ferredoxin (whose X-ray structure, 1FRD is known), by using ϕ—ψ restraints for the homologous (nonglycinil) residues to find the optimal structure for Pdx. The protocol proceeded through a simulated annealing refinement, with ramped van der Waals repulsion terms. No solvent or electrostatic terms were included. Twelve possible structures that satisfy these constraints were deposited in the PDB data base under the name 1PUT. These structures will be referred to as the "PDB structures" for the rest of this article. The average pairwise RMSD (root mean square deviation) for these structures was 1.14 Å for the backbone atoms (C_α, N, and C) and 1.80 Å for all nonhydrogen atoms (including side-chain heavy atoms). This indicates that the proposed structures are very similar, and are certainly within normal thermal fluctuations for most of the protein.

In cases where restrained molecular dynamics has been used to optimize particular structures, it becomes necessary to compute trajectories with the restraints turned off, and with a full force field, without the usual neglect of electrostatic interactions. In this way, all available conformers will be sampled, with no bias toward those that satisfy NOE restraints. It is important to determine the stability and topological integrity when the experimental restraints are removed. In the case of NMR, it becomes important to note that the strong NOE dependence on the interatomic distances tends to bias the conformations toward short distances between related nuclei. This means that in cases where there are large fluctuations of a particular region (as determined by MD), the relevant distances can change drastically during the time scale of the experiment, and only those conformations that actually give rise to NOEs will be seen. A case of such large fluctuations will be presented later in this paper.

One of the most striking properties of Pdx is the existence of a tryptophan as the C-terminal residue. Experiments have shown that this residue is found at the binding interface between Pdx and P450cam. Enzymatic removal of this residue greatly reduces the activity in the reconstituted hydroxylase system (Sligar et al., 1974). Recently, site-directed mutagenesis experiments have been performed, and modifications have been introduced in the C-terminal site (Davies and Sligar, 1992; Holden et al., 1997). These data suggest that the existence of aromatic residues in this position is important to the binding process, but that the role of W106 in electron transfer is less essential. The main effect of the presence of W as a C-terminal residue is to provide for differential affinities between P450cam and reduced or oxidized Pdx.

Another set of important data regarding W106 is found in the fluorescence experiments of Sligar's group (Stayton and Sligar, 1991). They have performed steady-state fluorescence experiments in both redox states for Pdx. The emission spectrum has a maximum at 358 nm, indicative of a solvent-exposed indole ring. They have also found a steady-state anisotropy of 0.04, suggesting a tryptophan residue rotating free of protein constraints. Three different lifetimes for the fluorescence decay were observed (0.3, 2, and 5 ns). The most striking observation was the fact that the components with the two longest lifetimes present differential quenching rates with iodine ion, indicating the presence of more than one conformation for W106. Direct molecular dynamics results supporting this observation will be presented, and a discussion of this structural microheterogeneity will be included in the Results.

The work presented here consists of the description and analysis of a molecular dynamics simulation for Pdx. Parameters were developed for the treatment of the metallic center and its interaction with its four cysteine anchors. The protein was immersed in a cubic box of water molecules, and counterions were added for overall charge neutrality. A particle mesh Ewald (PME) method was used to account for the long-range electrostatic forces. Three different simulations were run, each lasting ~1 ns. A detailed analysis of the structure and dynamics of these MD simulations was performed by standard techniques. In particular, given the importance of the C-terminal region of the protein, a locally enhanced sampling (LES) analysis was performed involved Q105 and W106.

The importance of this work resides in showing the need for a study of protein dynamics when dealing with systems that have been shown experimentally to be highly mobile. The information obtained from the present simulations will allow us to study the more important issues of binding and electron transfer properties between Pdx and P450cam, in a framework that goes beyond the use of single structures, thus incorporating important dynamical effects.

METHODS

All molecular dynamics simulations have been run using the program AMBER 4.1 (Pearlman et al., 1995) from UCSF. The parameters for the molecular dynamics simulations were used as described in the all-atom set contained in AMBER 4.1 (Cornell et al., 1995), except as noted.

An important issue when dealing with metal center proteins is the availability of potential energy parameters. In the case of iron-containing proteins, this problem is compounded because high-level ab initio calculations on sizable systems become extremely difficult. A number of experiments have been performed in proteins related to the present one, or in model systems designed to mimic the active center (Han et al., 1989; Fu et al., 1992). Some theoretical studies have computed effective charges in the metallic clusters (Mouesca et al., 1994).

It has been experimentally determined that the Fe_2S_2 cluster is bonded to four cysteine residues numbered Cys^{39}, Cys^{45}, Cys^{48}, and Cys^{86} (Gerber et al., 1990). A set of parameters for these residues as well as the Fe_2S_2 cluster is available from the author on request. All other parameters have been extracted from the regular AMBER data base. The charges of the ionizable residues were assigned as their default values at pH 7. The overall charge for Pdx is −9.

The simulation was set up as follows. The structure named 1PUT was extracted from the PDB data base (Bernstein et al., 1977). The protein was immersed in a box of TIP3 water molecules. The wetting was set up such that at least 8 Å was available between the surface of the protein and the box sides. Furthermore, at least 2 Å was left between the surface of the protein and the closest O atom of any water molecule, so as to prevent unfavorable overlaps. Counterions (in the form of Na^+ and Cl^-) were introduced close to the point charges of opposite sign, whenever possible.

After this procedure is finished, the system is neutral in charge. This procedure results in a box containing 1569 protein atoms, 19 counterion atoms, and 3155 water molecules, totaling 11,054 atoms in the system, including hydrogens.

The system was heated from 0 K to 300 K over 100 ps, using periodic boundary conditions and a constant pressure algorithm (with a pressure relaxation time of 0.2 ps), and a particle mesh Ewald (PME) implementation of the Ewald sum for long-range electrostatics (Darden et al., 1993; Kollman et al., 1995). The use of PME is crucial for long molecular dynamics runs (longer than 200 ps), especially in systems that have a large concentration of charges, such as this one.

A dielectric constant of 1 was used throughout. A spherical cutoff of 8 Å was used for the Lennard-Jones nonbonded potentials, which were updated every 20 fs. The time step was 2 fs. The bonds that involved a H atom were constrained to their equilibrium lengths by the SHAKE method (Ryckaert et al., 1977). After the heating period was completed, another 50 ps of equilibration was carried out under the same set of parameters. During the equilibration period, the pressure was stabilized around 1 atm and the temperature around 300 K. All components of the energy behaved properly, with the total energy fluctuating less than 1%. The temperatures of the solute and solvent were independently coupled to a heath bath, with both couplings using a 0.2 ps time constant. After the system was equilibrated, the box size stabilized around 52.34(0.06) × 48.97(0.06) × 43.79(0.05) (with the numbers in parentheses representing single standard deviations during the MD runs). The overall volume fluctuation was 0.3%.

All MD runs were performed on either a 4 Proc SGI Power Challenge or on eight processors of an IBM SP2. The system was optimized for a throughput of 1 ns/week.

RESULTS AND DISCUSSION

Overall behavior of the PDB structures

As an indicator of the overall stability of the PDB structures, the RMSD, computed using only the C_α for all residues, was calculated during the first nanosecond of dynamics. The deviation of the dynamical structures against the PDB structure is shown as a thick line in Fig. 1. The value of this RMSD after heating and equilibrating is already 2 Å, and reaches values close to 3 Å during the dynamics. The fact that the RMSD seems to converge to a value that it is not extremely large lends the first hard evidence that the structure, as deposited in the PDB data base, is essentially correct. The thin line in Fig. 1 represents the C_α RMSD during the dynamics, computed against a structure that represents the MD average. This RMSD stabilizes around 1 Å, meaning that the MD runs fluctuate very little around their own average.

Once the overall stability of the structure has been determined, it becomes of interest to investigate how different regions in the protein fluctuate. Fig. 2 shows the average RMSD fluctuations for each residue, computed against the average structure. As shown in Fig. 1, the overall RMSD is on the order of 1 Å. This means that residues that fluctuate more than that are more mobile than the average, whereas residues with a RMSD lower than 1 Å are more rigid. The region around residues 64 and 73 and the last two residues in the C-terminal region are the more mobile segments in this protein. Fig. 3 presents a different view of the same effects. Equispaced time frames, spanning a total of 1-ns dynamics, have been superimposed. Only the backbone is shown for all residues, except W106, for which the side chain is also included. Labels for selected residues are included to aid visualization. The most important feature of

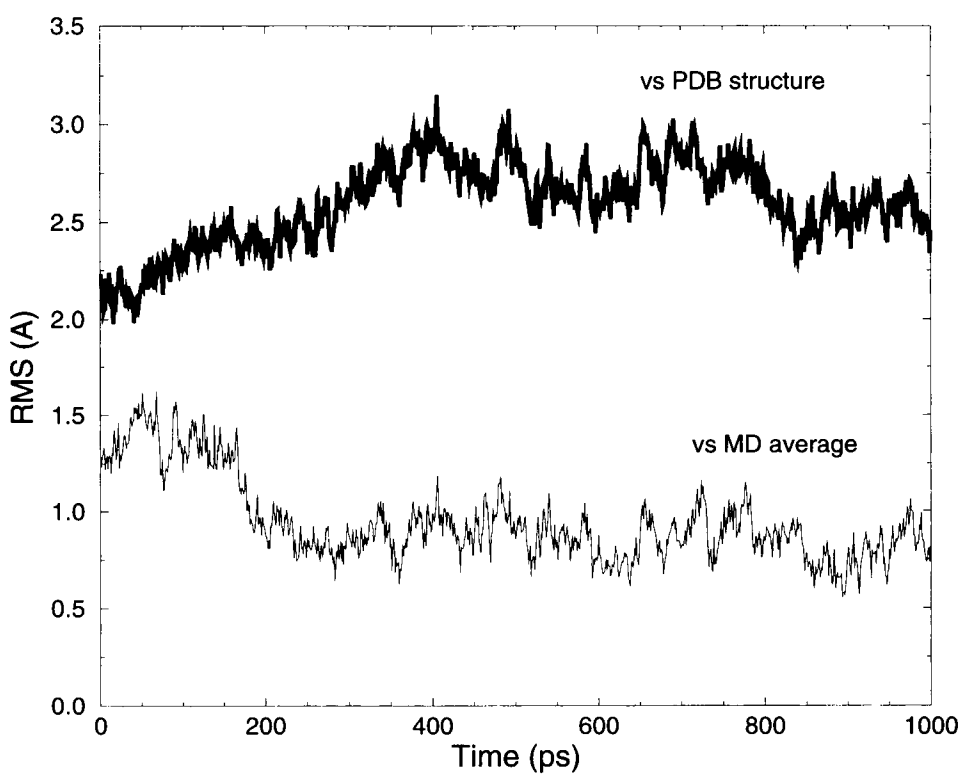

FIGURE 1 C_α RMSD versus time. The thick line is computed against the PDB structure and the thin line is computed against the MD average structure.

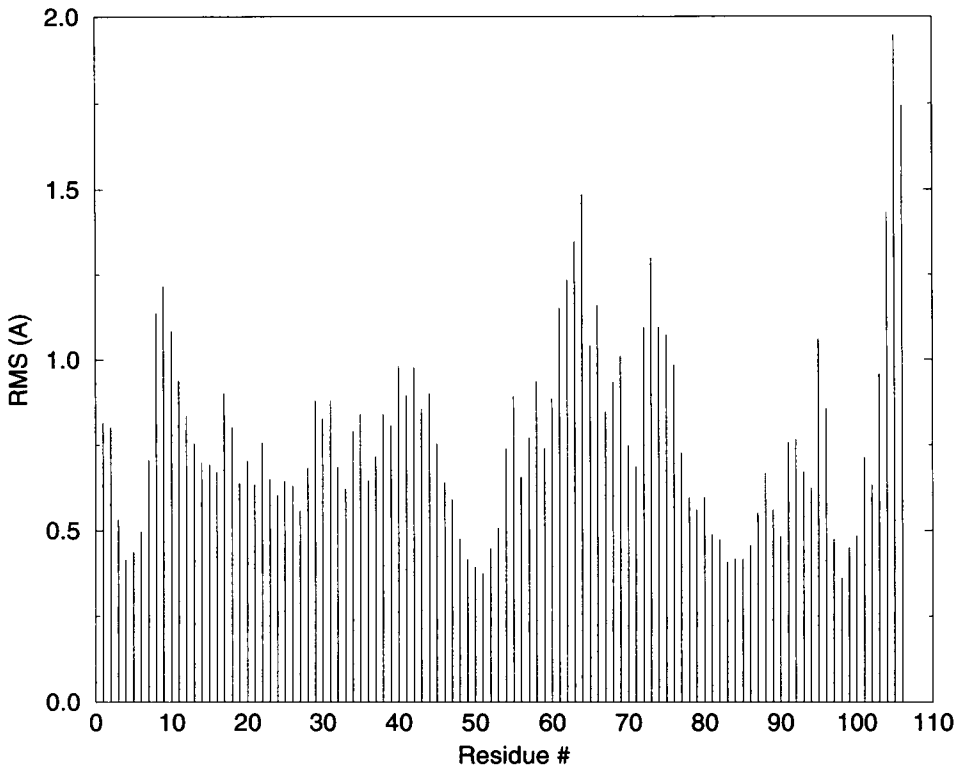

FIGURE 2 RMSD fluctuations as a function of residue number during the dynamics.

Fig. 3 is the fact that the overall structural integrity seems to be preserved very well. It can be seen from this picture that some areas are highly mobile (the Asp^{58} region and the side chain for W106), and others seem to be quite rigid (the region around the Fe_2S_2 cluster). This latter fact adds support to Pochapsky's assumption regarding the low flexibility of the cluster region.

The results from the dynamics and the apparent high disorder at the C-terminal region shown in Fig. 3 provide a wealth of information about the flexibility of the Pdx system. It shows that for this protein (as well as for most others), the concept of a single structure is probably highly misleading, and that information concerning the dynamics of the system cannot be transmitted by a single structure.

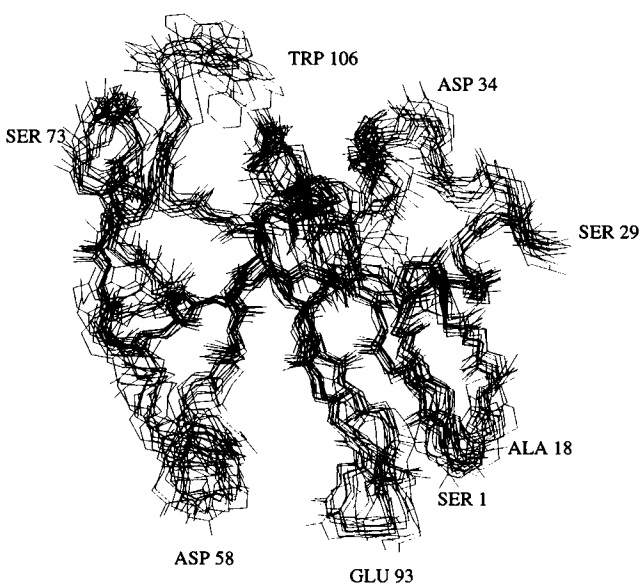

FIGURE 3 Superimposed frames during the dynamics. The frames are equispaced in time. Only the backbone is shown for all residues except W106, for which the side chain is also included.

Microconformers

Of the regions of high mobility, the most relevant for the function of Pdx is the C-terminal region. Detailed analysis of the dynamical trajectories is warranted in this region, and some results will be presented here. Fig. 4 is a zoomed-in version of Fig. 3, showing only residues Q105, W106, and the Fe_2S_2 cluster. The W106 side chain appears to be completely disordered, with no particular preferred conformation. A clustering analysis shows a very different picture. The idea is to identify clusters of structures that can minimize the intracluster RMSD deviation, while differentiating themselves from other clusters. Two clusters are said to be different if the average RMSD deviation between members of different clusters is larger than a given cutoff. If the cutoff is set to zero (or a very small value), then the number of clusters will equal the number of structures, with each structure belonging to its own cluster. On the other hand, if the cutoff is chosen to be very large, only one cluster, containing all structures, is obtained. Neither of these extreme cases provide much useful information. A cutoff of 3.5 Å was chosen as giving the most stable distribution of

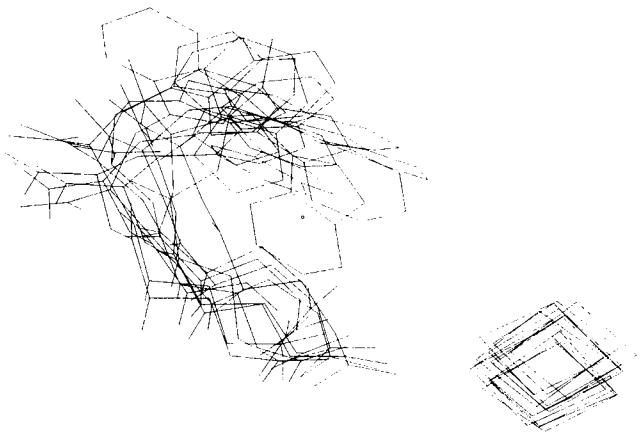

FIGURE 4 Expanded version of Fig. 3, showing only residues Q105 and W106 and the Fe_2S_2 cluster.

clusters. The clustering was performed using routines included in the program Moil-View (Simmerling et al., 1995). The results of this clustering technique are useful in understanding the dynamical processes taking place on a nanosecond time scale. When we focus on the last two residues in Pdx, six clusters are obtained. Two of them are highly populated, whereas the other four are sparsely so. The two main clusters were studied. Cluster A and cluster B appear 49% and 40% of the time, respectively. Table 1 presents the values of the dihedral angles for structures representative of these two clusters. The clustering technique can then serve to unscramble the seemingly random form of Fig. 3, by showing that there are mostly two different conformations, each with enough flexibility to look like one single disordered system. The two clusters differ from one another mostly in the values for χ_1 and χ_2 for residue W106. It is then useful to reanalyze the dynamics runs to find signatures for this behavior. Fig. 5 shows the values of χ_1 and χ_2 for residue W106 during the dynamics. We see that there is a "flip" in the value of χ_1 from positive to negative values during the dynamics, and that the equilibrium values for χ_2 in the two cases are slightly different. If we count the number of occurrences in each of the two clusters, we find the same ratio as in the clustering analysis, that is 49/40. This particular type of flip occurs in the nanosecond time scale.

Of particular interest is the comparison of these data with experimental results from Sligar's group. According to Stayton and Sligar (1991), the time-dependent fluorescence of the only tryptophan residue in Pdx (W106) can be deconvoluted into at least three different components, with two of them showing different quenching rates with I^-. This difference in quenching rates can be ascribed to different accessible surface areas (ASA) (described below in detail). In fact, the two clusters described above have ASAs of 109 and 120 $Å^2$, respectively. It is important to note that neither the time scale nor the sampling provided in this calculation warrants the complete identification of such clusters with Sligar's microconformers.

TABLE 1 Dihedral angles for the two main C-terminal clusters encountered during the dynamics

Dihedral angle	Cluster A	Cluster B
ϕ_{Q105}	−129.17	−95.12
ψ_{Q105}	−16.20	−37.50
ϕ_{W106}	−121.77	−151.27
$\chi_{1,W106}$	−66.94	41.05
$\chi_{2,W106}$	114.60	82.86

Fluorescence depolarization

The use of fluorescence depolarization as a tool to study the dynamics of proteins has a long and fruitful history (Lakowicz, 1983). It uses the preferential absorption of photons whose electric vectors are aligned parallel to the transition moment of the fluorophore. In a homogeneous solution, this will also result in a partial polarization of the fluorescence emission. Measurement of the fluorescence depolarization parameter provides information about the relative angular displacement of the fluorophore during the time between absorption and emission.

According to Stayton and Sligar (1991), not only does the steady-state emission fluorescence spectrum of the only tryptophan residue in Pdx (W106) peak at a wavelength of 358 nm, indicative of a completely solvent exposed residue, but its anisotropy of 0.04 is characteristic of a rotation that is partially free of overall protein tumbling. The data provided by the molecular dynamics simulations described above can be used to compute the fluorescence anisotropy. The time-dependent anisotropy can be written as

$$\alpha(t) = \frac{I_{\parallel}(t) - I_{\perp}(t)}{I_{\parallel}(t) + 2I_{\perp}(t)} \tag{1}$$

where $I_{\parallel}$ and $I_{\perp}$ are the fluorescence intensities polarized parallel and perpendicular to the polarization of the incident beam at time t. This can be expressed (following Ichiye and Karplus, 1983) in a way amenable to molecular dynamics simulations by using the language of correlation functions:

$$\alpha(t) = \frac{2}{5} e^{-t/\tau_m} \langle P_2[\hat{\mu}_A(0) \cdot \hat{\mu}_E(t)] \rangle \tag{2}$$

where τ_m is the overall correlation time for rotation of the whole molecule (usually tens of nanoseconds), and $\hat{\mu}_A$ and $\hat{\mu}_E$ are the unit transition dipole vectors (absorption and emission, respectively) with components defined in a local coordinate system fixed in the molecule. P_2 is a second-order legendre polynomial. The angle brackets represent an ensemble average.

Fig. 6 presents the behavior of the correlation function $\langle P_2(\vec{r_{ij}}(t) \cdot \vec{r_{ij}}(0)) \rangle$, where i and j represent two atoms in the system. The correlation function was computed following the method described by Palmer and Case (1992). The thick line corresponds to a vector that joins atoms $C_{\delta 1}$ and $C_{\zeta 3}$ in W106, and should be considered as representative of

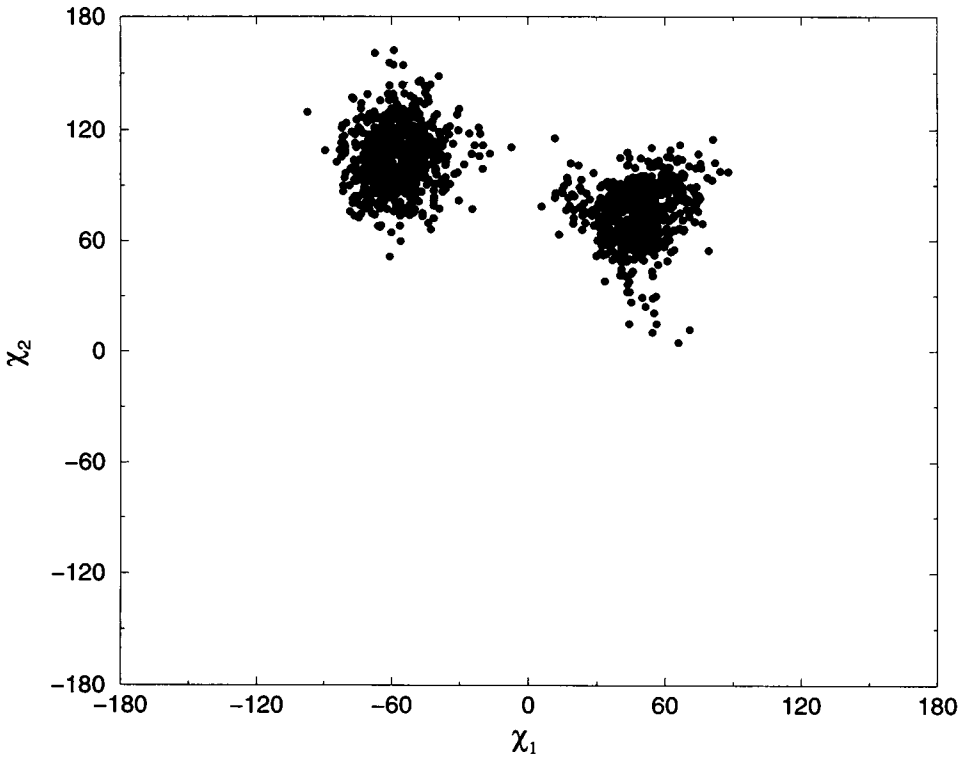

FIGURE 5 Values of χ_1 and χ_2 for W106 during the dynamics runs. Note the existence of two distinct clusters.

the indole ring rigid rotation. The dashed line represents the correlation function for a vector joining the N and H atoms in the backbone of W106. Although there is no fluorescence associated with the backbone, it serves as a comparison between the movements of the side chain with respect to the main chain in the protein. The backbone plot reaches a plateau by 200 ps, whereas the signal corresponding to the side chain never stops decaying in this time scale. The

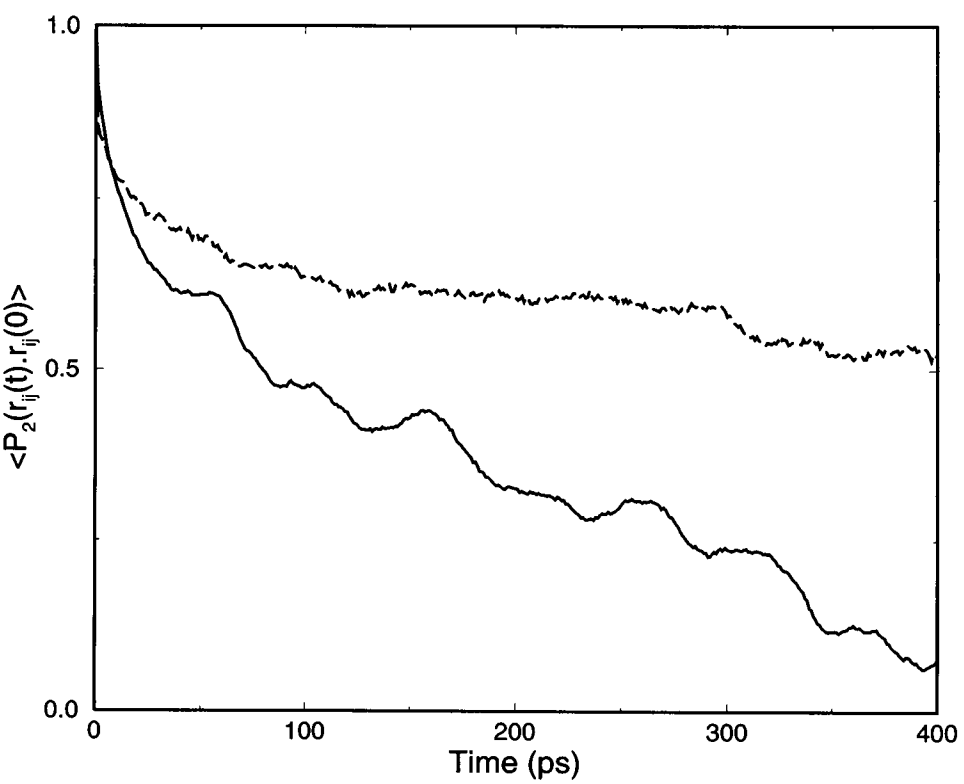

FIGURE 6 Correlation function defined as $\langle P_2(\vec{r_{ij}}(t) \cdot \vec{r_{ij}}(0))\rangle$ as a function of time. The thick line corresponds to a vector that joins atoms $C_{\delta 1}$ and $C_{\zeta 3}$ in W106. The dashed line represents the correlation function for a vector joining the N and H atoms in the backbone of W106.

experimental value for the anisotropy of the W106 ring has been reported as 0.04 in a steady-state measurement. The time scale for such an experiment is on the order of a few nanoseconds, and it indicates that the indole ring is effectively free to rotate within that time scale. This is completely consistent with the results shown in Fig. 6, where the indole ring shows extreme dephasing with respect to its initial conditions in a nanosecond time scale.

Pdx-P450cam binding

Recently Pochapsky has proposed a plausible structure for the Pdx-P450cam complex based on modeling techniques (Pochapsky et al., 1996). According to that model, there are mainly three ion pairs, some favorable hydrogen bonds, and an aromatic-aromatic interaction responsible for the binding. The three ion pairs proposed are between (first residue in Pdx and second in P450cam) D38-R112, D34-R109, and COOH W106-R79. The MD runs have been analyzed to search for distance distributions for these residue pairs. Fig. 7 presents the histograms for the interatomic distances for selected residues in Pdx (C_γ D34–C W106 (*top*), C_γ D34–C_γ D38 (*middle*), and C_γ D38–C W106 (*bottom*)). The fluctuations of these solvent-exposed charged residues are very large. The distances between these atom have to match, on average, with the corresponding residues in P450cam. The matching distances in the X-ray structure of P450cam (Poulos et al., 1987) provide a tight control for the binding hypothesis. Table 2 presents the distances (average distances with errors measured as a single standard deviations) between relevant residues (as measured between C_γ for D, C_ζ for R, and the carboxylate C for W). The distances from the MD runs match almost perfectly with the corresponding distances in P450cam, providing very strong support to Pochapsky's hypothesis regarding the residues involved in binding. It is important to mention that the distances obtained from the PDB structure for Pdx do not match as well, mostly because of a change in conformation of W106 and its carboxylate group. Fig. 8 shows residues D34, D38, and W106, for both the MD trajectory (*solid line*) and PDB (*dashed line*) structure. The difference in the conformation for residue W106 appears in every MD run performed, and happens within 20 ps of dynamics. The overall result of this change in conformation is to increase the distance between the carboxylate group in W106 and the carboxylate ends of D34 and D38, toward values that match the associated triangle on P450cam, as shown in Table 2.

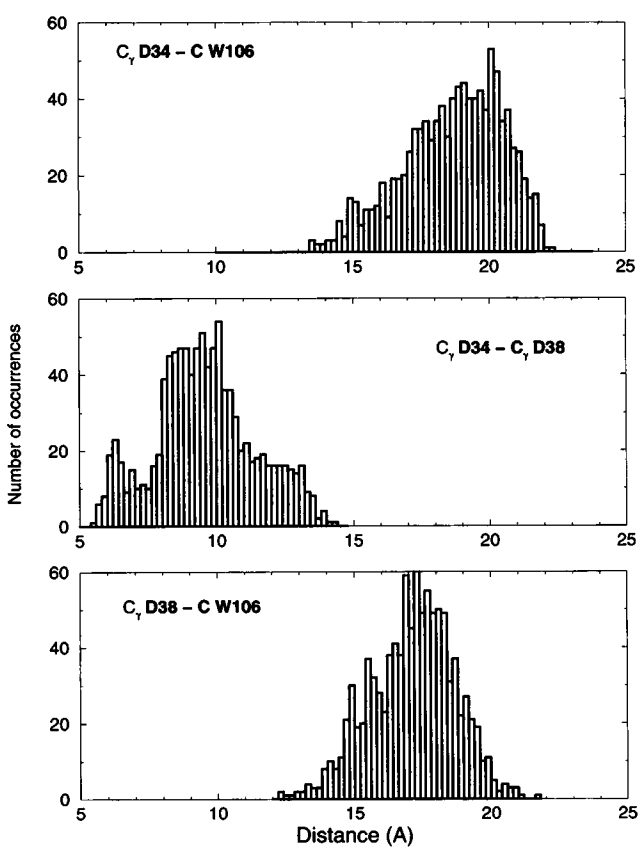

FIGURE 7 Histograms showing distance distribution for C_γ D34–C W106 (*top*), C_γ D34–C_γ D38 (*middle*), and C_γ D38–C W106 (*bottom*).

Solvent accessibility of W106

One of the findings of Stayton and Sligar (1991) is the fact that the fluorescence emission spectrum for W106 is representative of a fully solvent-exposed indole ring. This model has been tested in the simulations performed for this article.

Accessible surface areas (ASAs) are calculated by computing the van der Waals surface for the protein and then rolling a sphere of 1.4 Å (representing a water molecule) around it. The area of the contact points is then the ASA. These areas have been computed for the indole ring in W106 in four different systems. The first corresponds to the structures deposited in the PDB data base (expressed as an average for the 12 structures), the second to the MD runs presented here (also as an average over the dynamics), and the third and fourth are capped tryptophan residues, in α-helix and β-sheet conformations, respectively. The results are presented in Table 3 and are given in $Å^2$.

The most important result is that there is a statistically significant difference between the ASAs of the MD structures and those deposited on the PDB data base. Moreover, if we consider the ASA values for the α-helix and β-sheet conformations as upper limits on solvent exposure, we can see than the MD structures are substantially closer to those values than those deposited in the PDB data base. This point is important because, as mentioned above, experimental evidence suggests a fully solvent-exposed indole ring for W106.

Locally enhanced sampling

A more detailed modeling of the C-terminal region of the protein was carried out using a technique known as locally

TABLE 2 Distances between residues proposed to interact during P450cam-Pdx binding

Pairs (P450cam)	Distance (Å)	Pairs (Pdx)	Distance (Å) (MD)	Distance (Å) (PDB)
R112-R109	11.0	D38-D34	9.6 ± 1.8	10.1 ± 2.0
R112-R79	16.9	D38-W106	17.1 ± 1.5	7.7 ± 1.6
R109-R79	15.3	D34-W106	18.7 ± 1.8	12.8 ± 2.3

Errors are expressed as single standard deviations.

enhanced sampling (LES). LES has been designed to oversample regions of a system that are of interest, without spending large amounts of computer time in parts of the system that are not as important. In the present case, one would like to search for low-energy configurations of the C-terminal region, provided we know the structure of the rest of the protein. LES functions by making copies of the interesting region, while keeping the uninteresting parts intact. These copies do not interact with each other, while interacting with the rest of the system in a mean field way. This means that the rest of the system feels a force that comes from the average of the copies. This method has been used previously to study ligand diffusion in proteins (Elber and Karplus, 1990), side-chain placement (Roitberg and Elber, 1992; Koehl and DeLarue, 1994), protein loop modeling (Zheng and Kyle, 1994), small molecules (Huber et al., 1996), peptides (Simmerling and Elber, 1994, 1995), and ligand docking (Miranker and Karplus, 1991; Rosenfeld et al., 1993). It has been shown (Roitberg and Elber, 1992) that the global minimum of the new system with the copies is exactly the same as the global minimum of the real system, and the barriers between different local minima are reduced substantially. This two facts combine to ensure that a global optimization of the new system provides the proper answer for the original ensemble, but that the optimization proceeds in a much easier way because of the lower barriers. Once the LES ensemble is set up, the system is usually optimized by using a simulated annealing algorithm (Kirkpatrick et al., 1983). The LES protocol provides a self control for finding the global minimum. As mentioned above, the global minimum for the real and modified systems coincides. This happens when all of the copies in the LES system manage to converge to the same position. It is possible that, after careful convergence, copies are still found in different positions. This happens when either the system has not been cooled slowly enough to reach the global minimum, or said minimum simply does not exist, and the system has a number of degenerate multiple minima. The first possibility can be discarded by running the annealing at a much slower pace. For all LES simulations presented in this article, convergence has been checked by increasing the annealing times to much longer values. In all cases, results for the shortest simulations that maintain the overall results are shown.

In the present case, residues 1–104 plus the Fe_2S_2 cluster were kept fixed at their MD equilibrated positions. Residues Q105 and W106 were the LES part of the system. The backbone for those two residues, taken as a whole, was copied twice. Then the side chains for the same residues were extended to three copies. This means that a total of two copies for the two C-terminal residue backbones, plus six copies for the side chains in those residues were studied. These copies, plus all water molecules in the box described in the Methods section, were allowed to move. A 0.5-fs time step was used, and the moving part of the system was heated to 1600 K. This set was equilibrated for 200 ps at that high temperature. Improper dihedral constraints were added to preserve chirality at this temperature. The dihedral angles were completely randomized. A sphere of 10 Å was set up around the two last residues, and all of the waters molecules within this zone were allowed to move for the next part of the simulation. The new system consisted then of a moving section containing 210 protein atoms and 317 water mole-

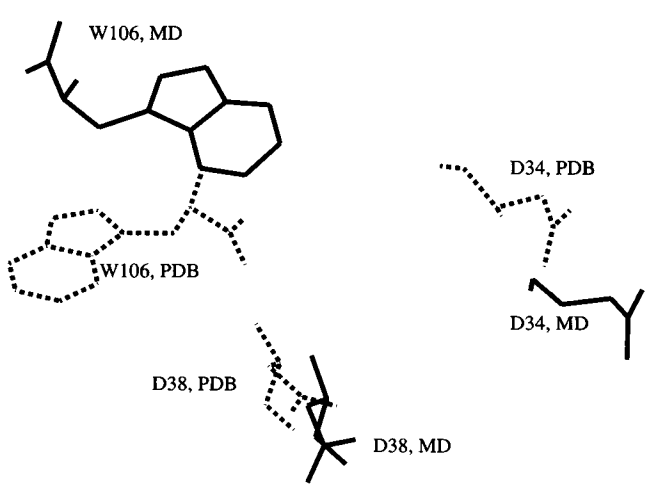

FIGURE 8 View of the three residues involved in binding between Pdx and P450cam. The thick lines correspond to the MD structures and the dashed lines to the PDB structure.

TABLE 3 Accessible surface areas for the structures deposited in the PDB data base expressed as an average for 12 structures (PDB), for the MD runs presented here (also as an average over the dynamics, MD) and for capped tryptophan residues, in an α-helix and β-sheet conformations, respectively

Structures	Accessible surface area (Å^2)
PDB	78.1 ± 9.7
MD	111.8 ± 4.8
α-helix	120.16
β-sheet	113.51

Errors are expressed as single standard deviations.

cules, while the rest of the original system remained fixed. This new system was equilibrated at 1600 K for another 100 ps. Each simulation consisted of a very slow annealing schedule. The temperature was lowered from 1600 K to 100 K over a 1-ns period, with a linear scaling in temperature. Finally, 20 ps of MD was performed at 10 K to produce structures that are minimized.

Five different annealings were performed, each one following the same protocol but starting from different initial conditions. All results will be presented as the relevant dihedral angles for Q105 and W106, that is, ϕ_{105}, ψ_{105}, ϕ_{106}, $\chi_{1,106}$, and $\chi_{2,106}$. The most important finding of these simulations is the fact that the C-terminal region seems to be highly disordered, with *many* different minima having essentially the same energy. Fig. 9 shows the behavior of the dihedral angles for W106 during the annealing for one of the LES runs. At high temperatures (1600 K), the dihedral angles can take most values. Then, as the annealing proceeds, the system tries to find the global minimum. For instance, ϕ for W106 stops undergoing transitions at temperatures around 400 K, whereas χ_1 does so only when the temperature drops below 200 K. χ_2 is still disordered, even at 100 K.

The issue of the final values for the relevant dihedral angles is slightly more complicated. It was found that in almost all cases, the results from different simulations show the angles to have converged to different values. Fig. 10 shows the final annealed structure for two of the five LES runs.

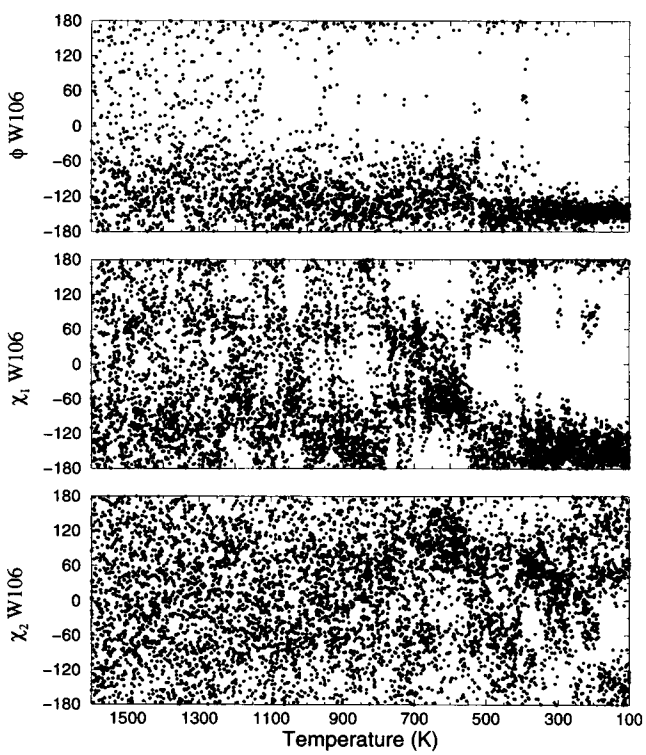

FIGURE 9 Dihedral angle behavior for residue W106 during LES annealing. The total annealing time was 1 ns.

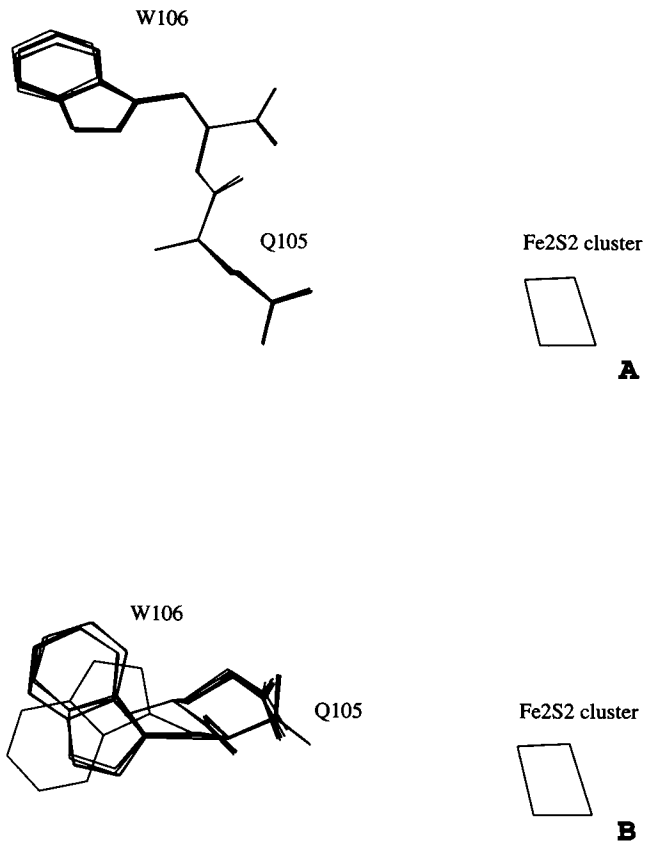

FIGURE 10 Final structures for two different LES annealings. Note that *A* (*top*) converged to a single minimum for all copies, whereas *B* (*bottom*) has each copy in a different conformation.

The results of run A (shown on top) are consistent with a single deep minimum for residues Q105 and W106 (note the six copies for the side chain essentially superimposed). However, run B shows a very different picture; not only have the copies within this run not found a single minimum (five copies of the side chain for W106 are in a particular conformation, whereas the other is in a different one), but the main conformer for this run is very different from the one found in run A. These discrepancies are found over the five simulations. Because all of them are converged in a simulated annealing sense (each one has found a "good" minimum), we can only conclude that the system is highly disordered.

It is important to note that these LES simulations do not offer any answers to the questions of barriers between possible minima, or the times scale required for them. It is only relevant to the existence of a large number of almost degenerate minima. We know, based on the regular MD discussed above, that at least two of these conformers can be visited in the nanosecond regime.

CONCLUSIONS

The work presented here provides some details and results from molecular dynamics simulations of putidaredoxin.

One of the main results is the fact that the structure of Pdx, as deposited in the PDB data base and obtained by NMR techniques, seems to hold up well when it is subjected to long, unrestrained MD with full solvent. A number of small differences are present between the PDB and MD structures. In all of these cases, the differences could be traced back to the high mobility of some regions of the protein. In particular, the C-terminal region, that has been suggested to be important for the interaction between Pdx and P450cam, is highly mobile. In accordance with experimental results, the side chain for W106 has been found in at least two conformations. The structures obtained from the dynamics show extreme angular movement in the nanosecond time scale for the indole ring in W106, consistent with the experimental evidence from Sligar's group, obtained by fluorescence anisotropy techniques.

The accessible surface areas seen during the dynamics are significantly larger than those of the structures deposited in the PDB data base. These high values for the ASAs are consistent with the experimental evidence that the fluorescence spectrum for W106 in Pdx has a maximum corresponding to extreme solvent exposure.

It is then important, when dealing with proteins that have high mobility, to use techniques such as molecular dynamics to sample some of the conformational space available to the molecule, especially solvent and full electrostatic effects.

I am grateful to Tom Cheatham III at UCSF for his help in setting up a successful PME calculation. I am also thankful to Carlos Simmerling at UCSF for his help and code for the implementation of LES into AMBER 4.1. I also thank the staff of the Bioprocess Engineering Group at NIST for very helpful discussions, and for introducing me to the fascinating world of the P450cam system. Certain commercial instruments, reagents, or materials are identified in this paper to adequately specify the experimental procedures. Such identification does not imply recommendation by the National Institute of Standards and Technology, nor does it imply that the materials are necessarily the best available for the purpose.

This work was done under a contract with Geo-Centers, Inc., with funding provided by the Bioprocess Engineering Group in the Biotechnology Division at the National Institute of Standards and Technology.

REFERENCES

Bernstein, F., T. Koetzle, G. Williams, E. Meyer, M. Brice, J. Rogers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Chae, Y., and J. Markley. 1995. Analysis of the hyperfine-shifted N-15 resonances of the oxidized form of anabaena 7210 heterocyst ferredoxin. *Biochemistry.* 34:188–193.

Collins, J., and G. Loew. 1988. Theoretical study of the product specificity in the hydroxylation of camphor, norcamphor, 5,5-difluorocamphor and pericyclocamphanone by cytochrome P450cam. *J. Biol. Chem.* 263: 3164–3170.

Cornell, W., P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell, and P. Kollman. 1995. A second generation force-field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an n log(n) method for Ewald sums in large systems. *J. Chem. Phys.* 98: 10089–10092.

Davies, M., and S. Sligar. 1992. Genetic variants in the putidaredoxin-cytochrome P450cam electron-transfer complex: identification of the residue responsible for redox-state-dependent conformers. *Biochemistry.* 31:11383–11389.

Elber, R., and M. Karplus. 1990. Enhanced sampling in molecular dynamics: use of the time-dependent hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. *J. Am. Chem. Soc.* 112:9161.

Fruetel, J., Y. Chang, J. Collins, G. Loew, and P. Ortiz de Montellano. 1994. Thioanisole sulfoxidation by cytochrome P450cam(CYP101): experimental and calculated absolute stereochemistries. *J. Am. Chem. Soc.* 116:11643–11648.

Fruetel, J., J. Collins, D. Camper, G. Loew, and P. Ortiz de Montellano. 1992. Calculated and experimental absolute stereochemistry of the styrene and beta-methylstyrene epoxides formed by cytochrome P450cam. *J. Am. Chem. Soc.* 114:6987–6993.

Fu, W., P. Drozdzewski, M. Davies, S. Sligar, and M. Johnson. 1992. Resonance Raman and magnetic circular dichroism studies of reduced [2Fe-2S] proteins. *J. Biol. Chem.* 267:15502–15510.

Gelb, M., O. Malkonen, and S. Sligar. 1982. Stereochemistry and deuterium isotope effects in camphor hydroxylation by the cytochrome P450 cam monoxygenase system. *Biochem. Biophys. Res. Commun.* 104: 853–858.

Gerber, N., T. Horiuchi, T. Koga, and S. Sligar. 1990. *Biochem. Biophys. Res. Commun.* 169:1016.

Grayson, D., Y. Tewari, M. Mayhew, V. Vilker, and R. Goldberg. 1996. Tetralin as a substrate for camphor (cytochrome p450) 5-monooxygenase. *Arch. Biochem. Biophys.* 332:239–247.

Han, S., R. Czernuszewicz, and R. Spiro. 1989. Vibrational spectra and normal mode analysis for [2Fe-2S] protein analogues using ^{34}S, ^{54}Fe and ^{2}H substitution: coupling of the Fe-S stretching with S-C-C bending modes. *J. Am. Chem. Soc.* 111:3496–3504.

Holden, M., M. Mayhew, D. Bunk, A. Roitberg, and V. Vilker. 1997. Probing the interactions of putidaredoxin with redox partners in camphor P450 5-monooxygenase by mutagenesis of surface residues. *J. Biol. Chem.* 272:21720–21725.

Huber, T., A. E. Torda, and W. F. van Gunsteren. 1996. Optimization methods for conformational sampling using a Boltzmann-weighted mean field approach. *Biopolymers.* 39:103–114.

Ichiye, T., and M. Karplus. 1983. Fluorescence depolarization of tryptophan residues in proteins: a molecular dynamics study. *Biochemistry.* 22:2884–2893.

Lakowicz, J. 1983. *Principles of Fluorescence Spectroscopy.* Plenum Press, New York.

Jones, J., W. Trager, and T. Carlson. 1993. The binding and regioselectivity of reaction of (R)- and (S)-nicotine with cytochrome P-450cam: parallel experimental and theoretical studies. *J. Am. Chem. Soc.* 115: 381–387.

Kirkpatrick, S., C. Gelatt, and M. Vecchi. 1983. Optimization by simulated annealing. *Science.* 220:671–680.

Koe, G., and V. Vilker. 1993. Dehalogenation by cytochrome-p-450(cam)—effect of oxygen level on the decomposition of 1,2-dibromo-3-chloropropane. *Biotechnol. Prog.* 9:608–614.

Koehl, P., and M. DeLaruc. 1994. Application of a self-consistent mean-field theory to predict protein side-chains conformation and estimate their conformational entropy. *J. Mol. Biol.* 239:249–275.

Kollman, P., T. Cheatham, III, and J. Miller. 1995. Molecular-dynamics simulations on solvated biomolecular systems—the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* 117(14):4193–4194.

Li, S., and L. Wackett. 1993. Reductive dehalogenation by cytochrome p450(cam)—substrate-binding and catalysis. *Biochemistry.* 32: 9355–9361.

Loida, P., S. Sligar, M. Paulsen, G. Arnold, and R. Ornstein. 1995. Stereoselective hydroxylation of norcamphor by cytochrome p450(cam)—experimental verification of molecular dynamics simulations. *J. Biol. Chem.* 270:5326–5330.

Lyons, T., G. Ratnaswamy, and T. Pochapsky. 1996. Redox-dependent dynamics of putidaredoxin characterized by amide proton exchange. *Protein Sci.* 5:627–639.

Mei, X., T. Pochapsky, and S. Pochapsky. 1992. ^{1}H NMR sequential assignments and identification of secondary structural elements in oxidized putidaredoxin, an electron-transfer protein from *Pseudomonas*. *Biochemistry*. 31:1961–1968.

Miranker, A., and M. Karplus. 1991. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins*. 11:2934.

Mouesca, J.-M., J. Chen, L. Noodleman, D. Bashford, and D. Case. 1994. Density functional/poisson-Boltzmann calculations of redox potentials for iron-sulfur clusters. *J. Am. Chem. Soc.* 116:11898–11914.

Palmer, A., and D. Case. 1992. Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide. *J. Am. Chem. Soc.* 114:9059–9067.

Pearlman, D., D. Case, J. Caldwell, W. Ross, T. Cheatham, III, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman. 1995. Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Comm.* 91:1–41.

Pochapsky, T., T. Lyons, K. Kazanis, T. Arakaki, and G. Ratnaswamy. 1996. A structure-based model for cytochrome P450(cam)-putidaredoxin interactions. *Biochimie*. 78(8–9):723–733.

Pochapsky, T., and X. Mei. 1991. ^{1}H NMR identification of a beta-sheet structure and description of folding topology in putidaredoxin. *Biochemistry*. 30:3850–3856.

Pochapsky, T., X. Mei, G. Ratnaswamy, and T. Lyons. 1994. An NMR-derived model for the solution structure of oxidized putidaredoxin, a 2-Fe, 2-S ferredoxin from *Pseudomonas*. *Biochemistry*. 33:6424–6432.

Poe, M., W. Phillips, J. Glickson, C. McDonald, and A. San Pietro. 1971. Proton magnetic resonance studies of the ferredoxins from spinach and parsley. *Proc. Natl. Acad. Sci. USA*. 68:68–71.

Poulos, T., B. Finzel, and A. Howard. 1987. High resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* 195:687–700.

Roitberg, A., and R. Elber. 1992. Modeling side-chains in peptides and proteins: application of the locally enhanced sampling and the simulated annealing methods to find minimum energy conformations. *J. Chem. Phys.* 95:9277–9287.

Rosenfeld, R., Q. Zheng, S. Vajda, and C. DeLisi. 1993. Computing the structure of bound peptides—application to antigen recognition by class-i major histocompatibility complex receptors. *J. Mol. Biol.* 234: 515–521.

Ryckaert, J.-P., G. Ciccotti, and H. Berendsen. 1997. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Phys.* 23:327–341.

Simmerling, C. L., and R. Elber. 1994. Hydrophobic collapse in a cyclic hexapeptide—computer-simulations of CHDLFC and CAAAAC in water. *J. Am. Chem. Soc.* 116:2534–2547.

Simmerling, C. L., and R. Elber. 1995. Computer determination of peptide conformations in water—different roads to structure. *Proc. Natl. Acad. Sci. USA*. 92:3190–3193.

Simmerling, C., R. Elber, and J. Zhang. 1995. Moil-view, a program for visualization of structure and dynamics of biomolecules, and sto, a program for computing stochastic paths. *In Modelling of Biomolecular Structure and Mechanisms*. A. Pullman, editor. Kluwer, Dordrecht, the Netherlands. 241–265.

Sligar, S., P. Debrunner, J. Lipscomb, M. Namtvedt, and I. Gunsalus. 1974. A role of the putidaredoxin COOH terminus in P450cam (cytochrome m°) hydroxilation. *Proc. Natl. Acad. Sci. USA*. 71:3906–3910.

Stayton, P., and S. Sligar. 1991. Structural microheterogeneity of a tryptophan residue required for efficient biological electron transfer between putidaredoxin and cytochrome P450cam. *Biochemistry*. 30:1845–1851.

Zheng, Q., and D. J. Kyle. 1994. Multiple copy sampling—rigid versus flexible protein. *Proteins*. 19:324–329.