Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Time-dependent atomic coordinates for the dissociation of carbon monoxide from myoglobin

Roman Aranda IV,[a]‡ Elena J. Levin,[b]‡ Friedrich Schotte,[c] Philip A. Anfinrud[c] and George N. Phillips Jr[b]*

[a]Department of Biomolecular Chemistry, University of Wisconsin, Madison, USA, [b]Department of Biochemistry, University of Wisconsin, Madison, USA, and [c]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA

‡ These authors contributed equally to this work.

Correspondence e-mail: phillips@biochem.wisc.edu

Received 27 February 2006
Accepted 10 May 2006

PDB References: myoglobin, crystal 1, 'laser off', 2g0r, r2g0rsf; crystal 2, 'laser off', 2g0s, r2g0ssf; 100 ps coordinate model, 2g0v, r2g0vsf; 316 ps coordinate model, 2g0x, r2g0xsf; 1 ns coordinate model, 2g0z, r2g0zsf; 3.16 ns coordinate model, 2g10, r2g10sf; 31.6 ns coordinate model, 2g11, r2g11sf; 316 ns coordinate model, 2g12, r2g12sf; 3.16 µs coordinate model, 2g14, r2g14sf.

Picosecond time-resolved crystallography was used to follow the dissociation of carbon monoxide from the heme pocket of a mutant sperm whale myoglobin and the resultant conformational changes. Electron-density maps have previously been created at various time points and used to describe amino-acid side-chain and carbon monoxide movements. In this work, difference refinement was employed to generate atomic coordinates at each time point in order to create a more explicit quantitative representation of the photo-dissociation process. After photolysis the carbon monoxide moves to a docking site, causing rearrangements in the heme-pocket residues, the coordinate changes of which can be plotted as a function of time. These include rotations of the heme-pocket phenylalanine concomitant with movement of the distal histidine toward the solvent, potentially allowing carbon monoxide movement in and out of the protein and proximal displacement of the heme iron. The degree of relaxation toward the intermediate and deoxy states was probed by analysis of the coordinate movements in the time-resolved models, revealing a non-linear progression toward the unbound state with coordinate movements that begin in the heme-pocket area and then propagate throughout the rest of the protein.

1. Introduction

Since the first determination of the three-dimensional structures of hemoglobin and myoglobin in the 1950s (Kendrew *et al.*, 1958; Perutz, 1960), X-ray crystallography has become the most widely applied technique for obtaining structural information from biological macromolecules. However, knowledge of a static structure alone is not adequate to obtain a complete understanding of how a protein performs its biological function. A complete mechanistic explanation must also include a description of its dynamic behavior. The desire to understand the dynamic properties of proteins has led to the recent development of time-resolved X-ray crystallography, which allows the researcher to supplement the static time-averaged image obtained from conventional crystallography with a series of images representing proteins in motion (Moffat, 1998*a*,*b*, 2001, 2003; Schotte *et al.*, 2003, 2004; Helliwell *et al.*, 1998, 2003).

Myoglobin (Mb) is a heme protein that reversibly binds diatomic oxygen and other small gaseous ligands such as carbon monoxide (CO) and nitric oxide and serves as a mobile oxygen buffer in muscle (Radding & Phillips, 2004; Springer *et al.*, 1994). Owing to its simple well characterized structure, the presence of a photolabile bond between the ligand and the heme iron (Gibson & Ainsworth, 1957) and the availability of a large number of mutants with altered functional and kinetic

properties, Mb is an appealing model for protein dynamics and has been thoroughly studied by various biophysical techniques (Brunori *et al.*, 2004; Bourgeois *et al.*, 2003; Nienhaus *et al.*, 2003; Srajer *et al.*, 1996). The photolabile ligand–heme bond is essential to initiating a simultaneous reaction and following the protein's dynamics in synchrony. In order to monitor the conformational changes Mb undergoes during CO dissociation, time-resolved crystallography was employed to follow CO translocation as the protein transitions from the CO-bound state to the unbound deoxy state (Schotte *et al.*, 2003).

Previously, a double-mutant L29F/D122N sperm whale Mb was used to monitor CO dissociation from 100 ps to 3.16 µs after photolysis (Schotte *et al.*, 2003). The L29F mutant was chosen for this study because it decreases the lifespan of the primary docking-site intermediate to the order of hundreds of picoseconds (∼140 ps), a time range equivalent to the current resolution limit of time-resolved crystallography, and allows detection of atomic movements on the picosecond time scale. The D122N mutation is located on the protein surface between the G and H helices; although it has no significant effect on the overall structure or binding kinetics, it causes the mutant protein to crystallize in a higher symmetry space group than native Mb (Phillips *et al.*, 1990). This alternative crystal form facilitates time-resolved crystallographic experiments by reducing the range of orientations needed to collect a full data set. Electron-density maps extrapolated to 100% photolysis revealed the path of dissociated CO migration, the relaxation of the unliganded heme and side-chain rearrangements in heme-pocket residues His64 and Phe29 (Schotte *et al.*, 2003).

Although electron-density maps can serve as useful qualitative representations for diffraction data, coordinate models are necessary for quantitative analyses of molecular geometry and atomic movements. Since small changes in structures are hard to measure accurately with independent refinement, we apply difference refinement to the measured data to generate accurate coordinate models (Terwilliger & Berendzen, 1995). This technique assumes that model errors between two highly isomorphic crystal structures will be correlated and by subtracting the model errors small differences between a well characterized native structure and an unknown variant structure are more accurately refined. Because high-quality data were collected for the CO-bound unphotolyzed Mb structure and the conformational changes between the photolyzed and liganded states are small, difference refinement is ideally suited for analysis of the time-resolved Mb diffraction data. We report the results of difference refinement on time-resolved crystallographic data and discuss how the obtained coordinate models confirm previous theoretical and experimental studies of CO dissociation in Mb (Brunori *et al.*, 2004; Srajer *et al.*, 2001; Lim *et al.*, 1993, 1995, 1997; Hummer *et al.*, 2004).

2. Methods

The methods employed to obtain the time-resolved diffraction data are described in Schotte *et al.* (2003) and are reviewed briefly here. Crystals of the L29F/D122N sperm whale Mb were grown in seeded hanging drops and were reduced with sodium dithionite and placed under a CO atmosphere with 3.2 *M* ammonium sulfate pH 9.0 in a sealed X-ray capillary. Time-resolved Laue diffraction data sets were collected on the ID09B beamline at the European Synchrotron Radiation Facility in Grenoble, France using the pump-probe method, where a photolyzing 570 nm laser pulse was followed by a polychromatic X-ray pulse 150 ps in duration with its peak intensity at a wavelength of 0.79 Å and a bandwidth of 3.5% FWHM. Time-resolved diffraction data were collected at 283 K. Complete data sets were acquired 100 ps, 316 ps, 1 ns and 3.16 ns after photolysis on one crystal and 3.16 ns, 31.6 ns, 316 ns and 3.16 µs after photolysis on a second crystal. The degree of photolysis was estimated to be 23% and was determined by quantifying the integrated electron density remaining in the CO-binding site after photolysis. For each crystal, data sets were collected without the photolysis pulse to characterize the CO-bound structure. Diffraction images were indexed with *LAUEGEN* (Campbell, 1995), integrated with *PROW* (Bourgeois *et al.*, 2000) and scaled with *LSCALE* (Arzt *et al.*, 1999) to a resolution of 1.8 Å (Schotte *et al.*, 2003).

Two CO-bound models were refined for the two unphotolyzed crystals starting from a 1.7 Å resolution L29F/D122N CO-bound myoglobin structure (Carver *et al.*, 1992). The refinements and manual fitting were carried out in the programs *Crystallography and NMR System* (*CNS*; Brünger *et al.*, 1998) and *XtalView* (McRee, 1999), respectively. Difference refinement was employed using the target function $\chi^2_{\rm diff}$ based on the method of Terwilliger & Berendzen (1995),

$$\chi^2_{\rm diff} = \sum_{hkl} [|F_{\rm diff}(hkl)| - |F_{\rm c}(hkl)_T|]^2, \qquad (1)$$

where

$$|F_{\rm diff}(hkl)| = |F_{\rm o}(hkl)_T| - [|F_{\rm o}(hkl)| - |F_{\rm c}(hkl)|]. \qquad (2)$$

$F_{\rm o}(hkl)_T$ are the experimental amplitudes for each photolyzed time point, while $F_{\rm o}(hkl)$ are the experimental amplitudes from the native structure, the CO-bound 'laser off' unphotolyzed model. $F_{\rm c}(hkl)_T$ are the structure factors calculated from the refined model for each photolyzed time point and $F_{\rm c}(hkl)$ are the structure factors calculated from the CO-bound 'laser off' unphotolyzed model. The difference structure factors, $F_{\rm diff}(hkl)$, were calculated using an in-house program and the residuals, $R_{\rm diff}$, were calculated as

$$R_{\rm diff} = \frac{\sum_{hkl} \big| |F_{\rm diff}(hkl)| - |F_{\rm c}(hkl)_T| \big|}{\sum_{hkl} |F_{\rm diff}(hkl)|} \times 100. \qquad (3)$$

5% of the reflections were set aside to calculate $R_{\rm free}$ for each time-point model.

The starting model for each time point consisted of two overlaid alternate conformations: a fixed model that corresponded to the CO-bound unphotolyzed Mb, which was refined previously using the 'laser off' data, and a variable model that corresponded to the photolyzed time point and was allowed to change to minimize the target function $\chi^2_{\rm diff}$. To create initial coordinates for the deoxy conformation that

Table 1
Refinement statistics for initial CO-bound Mb.

CO-bound Mb crystal refinement statistics for models used as the starting point in the difference refinement.

Crystal	Space group	Resolution (Å)	Unit-cell parameters (Å)	R_{std} (%)	R_{free} (%)
1	*P*6	1.95	a = 91.20, c = 45.71	17.6	20.9
2	*P*6	1.90	a = 91.20, c = 45.87	16.6	19.2

were distinct from the CO-bound model, five cycles of maximum-likelihood coordinate and temperature-factor refinement were run in *CNS* using the difference 100 ps time-point data for amplitudes and the unphotolyzed model for starting coordinates. The resulting photolyzed model was then combined with the unphotolyzed model and the occupancies set to 0.23 and 0.77, respectively. For refinement of the subsequent time points, the initial model was simply the refined structure from the previous time point. All refinements were carried out to convergence. Occupancies were kept at 0.23 and 0.77 as described above throughout all time points. Coordinates were deposited in the Protein Data Bank (Berman *et al.*, 2000) as follows: 2g0r for the 'laser off' model for crystal 1, 2g0s for the 'laser off' model for crystal 2, 2g0v for the 100 ps coordinate model, 2g0x for the 316 ps coordinate model, 2g0z for the 1 ns coordinate model, 2g10 for the 3.16 ns coordinate model, 2g11 for the 31.6 ns coordinate model, 2g12 for the 316 ns coordinate model and 2g14 for the 3.16 µs coordinate model.

To measure the progression of the photolyzed time points toward the deoxy state, vector-overlap analysis was calculated as

$$\kappa = \frac{\mathbf{X}\cdot\mathbf{Y}}{|\mathbf{Y}|^2}, \tag{4}$$

$$\cos\theta = \frac{\mathbf{X}\cdot\mathbf{Y}}{|\mathbf{X}||\mathbf{Y}|}, \tag{5}$$

where $\mathbf{X}$ is the $3N$-dimensional vector (where N is the number of C^α atoms) representing the displacement between a particular time point and the CO-bound 'laser off' model and $\mathbf{Y}$ is the analogous vector between static L29F/D122N Mb structures in the deoxy (PDB code 2spl) and CO-bound (PDB code 1moa) states. κ is the magnitude of the projection of $\mathbf{X}$ onto $\mathbf{Y}$ normalized by division by the magnitude of $\mathbf{Y}$ and is a metric that quantifies the progress of the photolyzed state towards the final deoxy structure. $\cos\theta$ is the cosine of the angle between $\mathbf{X}$ and $\mathbf{Y}$ and describes the deviation of the photolyzed trajectory from the deoxy-state trajectory. If the atoms of the photolyzed structure were randomly displaced from the CO-bound structure both κ and $\cos\theta$ would be 0; as $\mathbf{X}$ approaches $\mathbf{Y}$, both κ and $\cos\theta$ converge to 1.

C^α difference distance matrices were created and plotted in *MATLAB* (MathWorks Inc., Natick, MA, USA) using the equation

$$\Delta_{ij}^{ab} = |\mathbf{D}_{ij}^{a}| - |\mathbf{D}_{ij}^{b}|, \tag{6}$$

Table 2
Difference refinement statistics and protein movements.

Each R_{diff} was calculated using (3) and r.m.s.d. values were calculated in *VMD* (Humphrey *et al.*, 1996) for each time point *versus* the 'laser off' Mb model. Protein r.m.s.d. values encompassed all protein atoms and heme-pocket r.m.s.d. values encompassed atoms within 8 Å of the heme iron, but not including the heme. Protein r.m.s.d. values without the heme pocket were calculated by omitting the atoms within 8 Å of the heme iron.

				R.m.s.d. (Å)		
Time points	Crystal used	R_{diff} (%)	R_{free} (%)	Protein	Without heme pocket	Heme pocket
100 ps	1	5.12	5.41	0.20	0.15	0.49
316 ps	1	5.54	6.27	0.21	0.16	0.52
1 ns	1	5.82	7.04	0.19	0.17	0.31
3.16 ns	2	4.52	4.88	0.20	0.19	0.29
31.6 ns	2	5.11	5.65	0.23	0.22	0.35
316 ns	2	5.28	6.21	0.23	0.22	0.34
3.16 µs	2	5.07	5.77	0.22	0.20	0.34

where $|\mathbf{D}_{ij}^{a}|$ is the distance between C^α atoms of residues i and j of model a and $|\mathbf{D}_{ij}^{b}|$ is the distance between C^α atoms of residues i and j of model b. Coordinate uncertainties were calculated using the diffraction-component precision index with R_{free} as described by Cruickshank, which gives the average positional error of any given atom at the average temperature factor of the structure (Cruickshank, 1999). Since difference refinement more accurately refines differences between the native CO-bound structure and the variant photolyzed structure, the positional errors should be reduced in comparison to two independently refined structures.

3. Results

Unit-cell parameters and refinement statistics for the unphotolyzed CO-bound Mb models are listed in Table 1. Difference refinement was carried out as described above to produce coordinates for each photolyzed time point, with each difference residual, R_{diff}, falling in the range between 4.5 and 5.8%

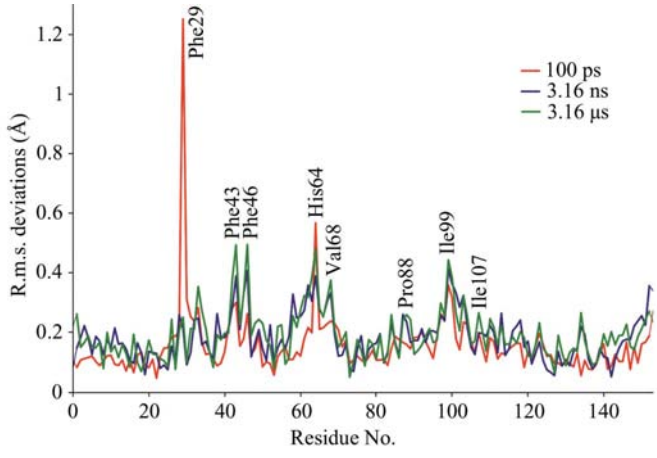

Figure 1
Residue deviations at the picosecond, nanosecond and microsecond time points. All-atom r.m.s.d. values of each residue (relative to CO-bound 'laser off' Mb) were plotted at the time points 100 ps, 3.16 ns and 3.16 µs in comparison to the CO-bound 'laser off' Mb.

(Table 2). It should be noted that $R_{\rm diff}$, as calculated with (3), differs from the R factor used in conventional independent refinement. Because the model errors are removed, $R_{\rm diff}$ should be lower than the conventional $R_{\rm std}$ by 10–15% (Terwilliger & Berendzen, 1995). Two data sets were collected on different crystals for the 3.16 ns time point and showed essentially the same coordinates, but the data from the second crystal were selected for deposition owing to a slightly lower $R_{\rm diff}$. Each $R_{\rm free}$ was less than 20% larger than the corresponding $R_{\rm diff}$ (Table 2).

Table 2 lists the root-mean-square deviations (r.m.s.d.s) from the CO-bound structures for each of the seven time points using either the entire protein or only atoms within 8 Å of the heme iron. The r.m.s.d. values for the atoms in the vicinity of the heme iron are higher at each time point than the r.m.s.d. values for the rest of the protein without the heme pocket, suggesting that the most significant side-chain rearrangements are localized to heme-pocket residues. The

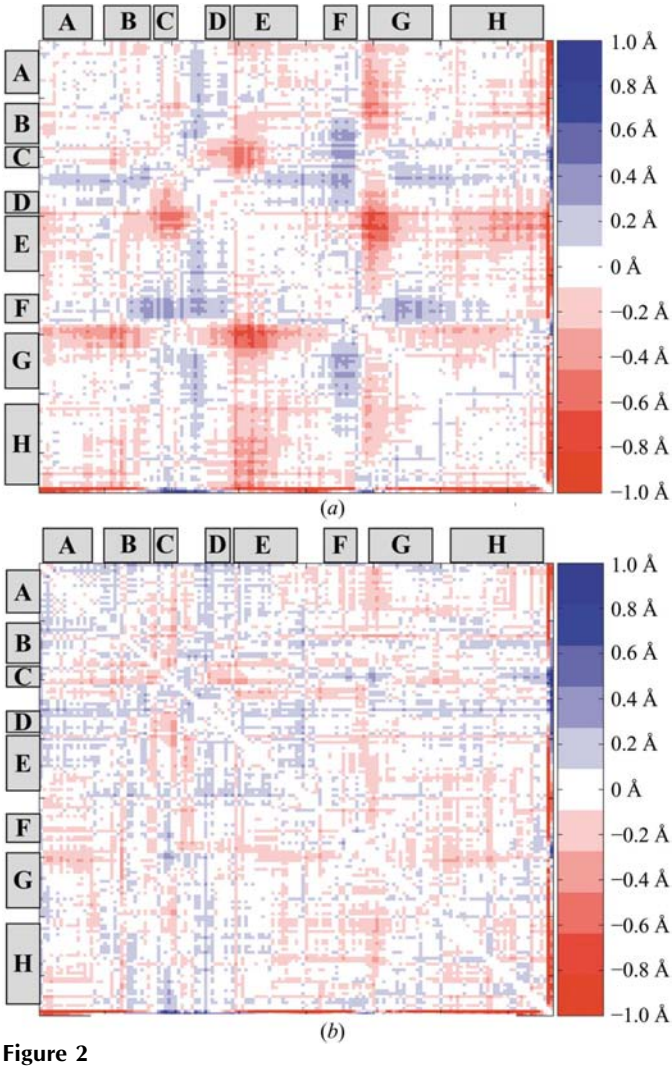

Figure 2
α-Carbon difference distance matrices comparing the CO-bound 'laser off' and 3.16 µs time-point models with deoxy-Mb. α-Carbon difference distance matrices were made by comparing the CO-bound 'laser off' and 3.16 µs model C^{α} atoms to deoxy-Mb C^{α} atoms (PDB code 1moa). The gray boxes correspond to the eight Mb helices and values are −1 Å (brightest red) to 0 Å (white) to +1 Å (brightest blue). (*a*) is the C^{α} difference distance matrix of (deoxy-Mb) − (CO-bound 'laser off' Mb). (*b*) is the C^{α} difference distance matrix of (deoxy-Mb) − (3.16 µs Mb).

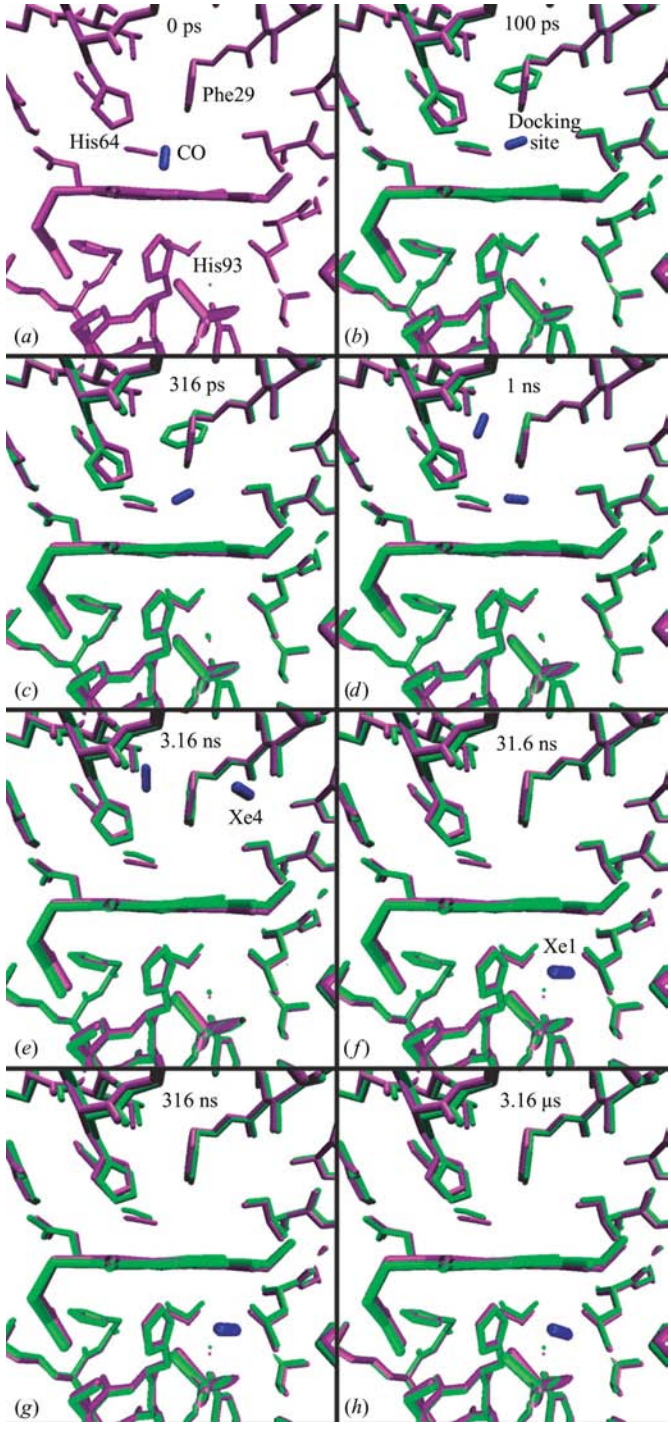

Figure 3
Time-resolved models of Mb from CO-bound to 3.16 µs after photolysis. Models were graphically depicted using *VMD* (Humphrey *et al.*, 1996) with the CO-bound 'laser off' model represented by the purple coordinates and the photolyzed time points represented by the green coordinates. The CO is in blue and the essential residues near the heme are labeled in the 'laser off' model. The orientation is the same at all time points, with the heme placed in the center and the propionates facing the solvent to the left. (*a*) CO-bound 'laser off' Mb. (*b*)–(*h*) Photolyzed models determined at 100 ps, 316 ps, 1 ns, 3.16 ns, 31.6 ns, 316 ns and 3.16 µs, respectively.

Table 3
Vector overlap of the time-point progression to the deoxy state.

As defined in (4), κ represents the fraction of the distance between the C^{α} atoms of the CO-bound and deoxy states that the photolyzed Mb has traversed by a particular time point. The angle between the time-point vector (**X**) and deoxy vector (**Y**) is described by $\cos\theta$. Random atomic movements would produce values of 0 for both κ and $\cos\theta$, whereas identical structures would produce values of 1 for both κ and $\cos\theta$. The heme-pocket calculations include C^{α} atoms from residues containing one or more atoms within 8 Å of the heme iron in the CO-bound structure.

	C^{α} all protein		C^{α} protein without heme pocket		C^{α} heme pocket	
Time point	κ	$\cos\theta$	κ	$\cos\theta$	κ	$\cos\theta$
100 ps	0.48	0.55	0.41	0.46	0.68	0.86
316 ps	0.51	0.58	0.45	0.50	0.70	0.84
1 ns	0.58	0.62	0.53	0.55	0.75	0.89
3.16 ns	0.69	0.61	0.64	0.54	0.83	0.91
31.6 ns	0.75	0.58	0.70	0.52	0.90	0.82
316 ns	0.77	0.60	0.73	0.54	0.92	0.86
3.16 µs	0.72	0.61	0.67	0.54	0.90	0.87

largest movements are found in mostly hydrophobic residues making contact with or near the heme, such as Phe29, Phe43, Phe46, His64, Val68, Ile99 and Ile107, and occur abruptly by 100 ps to accommodate doming of the heme and translocation of the CO immediately following photolysis (Fig. 1). Rearrangements of the backbone are smaller and more gradual in the later time points and are localized primarily to the C, E, F and G helices.

Although the atomic displacements between the time-point models are small, difference distance matrices and vector-overlap analysis show an overall movement toward the deoxy state and indicate that the movements are not random (Fig. 2 and Table 3). Differences between the C^{α} atoms of the CO-bound Mb and the L29F deoxy Mb (Fig. 2*a*) are localized to specific regions, notably the CD loop, the E and F helices and the FG loop. The differences between the 3.16 µs coordinate model and the deoxy-Mb C^{α} atoms (Fig. 2*b*) are not localized to any region, but rather only small dispersed adjustments are necessary to morph into the deoxy state. The vector-overlap analysis reveals a stretched-exponential progression toward the deoxy state (Table 3). As time passes, κ increases from 0.48 at 100 ps to 0.72 at 3.16 µs, indicating movement toward the deoxy state for the entire protein. The κ values for the heme-pocket residues are higher than for the protein residues excluding the heme pocket, indicating that the heme-pocket residues approach the deoxy state more quickly than the rest of the protein (Table 3). The higher $\cos\theta$ values for the heme-pocket residues than for the protein residues excluding the heme pocket result from the higher ratio of the atomic displacements to the positional uncertainties (Table 3).

Fig. 3 displays the region surrounding the heme at the seven time points, each of which is overlaid with the refined CO-bound 'laser off' Mb structure (see also the movie deposited as supplementary material[1]). After photolysis, the change in the spin state of the iron displaces the iron from the heme plane (defined by the atoms of the porphyrin ring) proximally by 0.38 Å, causing the heme to dome (Fig. 3*b*), and is consistent with the 0.35 Å seen in high-resolution deoxy-Mb crystal structures (Kachalova *et al.*, 1999). The heme reaches its equilibrium deoxy internal conformation by 100 ps after photolysis and does not change its internal structure further at subsequent time points; however, the recoil from the compression of the bond between the iron and proximal histidine, His93, causes the heme to tilt gradually throughout all the time points to a final rotation of 4.8° by 3.16 µs, coupled with a translation of 0.27 Å toward the solvent. The modest rearrangements in the Mb backbone result in part from the heme tilt exerting an apparent force in the distal direction on Ile107 that is communicated along the length of the G helix and from the iron pushing proximally on the F helix through His93 as it moves out of the heme plane.

At 100 and 316 ps after photolysis, the CO electron density is observed solely at the primary docking site between 3.9 and 4.1 Å from the distal side of the heme iron and lying roughly above the pyrrole nitrogen NC (Figs. 3*b*, 3*c* and 4*a*). This site is similar to that observed with previous cryocrystallography experiments that trapped dissociated CO and showed the photolyzed CO lying on the distal side of the heme, approximately 3–4 Å from the iron (Schlichting *et al.*, 1994; Teng *et al.*, 1997; Hartmann *et al.*, 1996). Coupled with the translocation of the CO to the primary docking site is the dislocation of the Phe29 side chain away from the 'docked' CO into a strained conformation rotated 105° around dihedral angle χ_2 at 100 ps, which increases to 110° by the 316 ps time point. Presumably, this motion provides the driving force that causes the neighboring side chain of the distal histidine side chain (His64) to swing 12° around dihedral angle χ_1 toward the solvent and away from Phe29 by the 100 ps time point (Fig. 3*b*). As His64 swings out, it encroaches on the volume formerly occupied by a crystallographic water molecule, HOH253. The water molecule undergoes an increase in its temperature factor from 33 to 61 $Å^2$ between the 'laser off' and 100 ps time points, indicating a decreased contribution to the diffraction pattern.

By 1 ns (Fig. 3*d*), the CO density at the primary docking site begins to diminish and new CO electron density appears near the solvent-facing side of the phenyl group of Phe29. As CO migrates away from the primary docking site, Phe29 returns to its original orientation. His64 lags behind Phe29 in returning towards its original conformation and instead begins to translate towards the heme iron owing to a rotation in the E helix. The timing of the disappearance of the CO at the primary docking site agrees with the transient intermediate reported in IR spectroscopy experiments of L29F Mb (Schotte *et al.*, 2003). As the distal histidine swings toward its original location and translates towards the heme iron, the temperature factor for HOH253 begins to decrease from 64 $Å^2$ at 1 ns to 47 $Å^2$ at 3.16 µs, but remains above the initial *B* factor of 33 $Å^2$ in the CO-bound Mb, thereby indicating that the water does not fully return to its original location 3.16 µs after photolysis.

By 3.16 ns (Fig. 3*e*), all CO density in the primary docking site has disappeared, while another presumed CO site appears in a cavity on the other side of the Phe29 (Figs. 3*e* and 4*b*). This

[1] Supplementary material has been depositied in the IUCr electronic archive (Reference: DZ5073). Details for accessing this material are given at the back of the journal.

site was found to bind photodissociated CO at 180 K and xenon (Xe) under high pressures and is referred to as Xe-binding site 4 (Tilton *et al.*, 1984; Ostermann *et al.*, 2000). By 31.6 ns (Fig. 3*f*), the CO disappears from the distal sites and begins to relocate below the heme to the Xe-binding site 1, as observed elsewhere (Schotte *et al.*, 2004; Nienhaus *et al.*, 2003; Srajer *et al.*, 2001). CO remains in Xe-binding site 1 throughout the 316 ns and 3.16 µs time points and steadily increases in occupancy from the 31.6 ns to 3.16 µs time points (Figs. 3*g*, 3*h* and 4*c*). At 3.16 µs, a CO electron-density peak appears below the heme and next to Ile107 near Xe-binding site 1 at 2σ in an $F_o - F_c$ map. However, the density peak at 3σ in the $F_o - F_c$ map is too small to warrant CO coordinate placement. At the final time point (Fig. 3*h*), the distance between the $N^{\varepsilon 2}$ of His64 and the heme iron is 4.7 Å, equal to the distance seen in high-resolution deoxy-Mb structures (Carver *et al.*, 1992), and His64 swings back 6° towards its pre-photolysis conformation and translates 0.4 Å towards the heme iron.

4. Discussion

Using difference refinement to produce coordinates from the time-resolved X-ray crystallographic data has facilitated quantitative analyses of the protein movements after photolysis, providing insight into the relationship between structure and function in heme proteins. Refining coordinates for time-resolved crystallographic data is difficult because the diffraction data are collected from a mixed population of CO-bound and unbound Mb and because the structural differences between the two states are small and close to typical positional uncertainties. Difference refinement is an ideal method for refining the Mb time-resolved crystallographic data because it minimizes the model errors and subsequently reduces the estimated positional uncertainties in closely related structures. To evaluate the validity of the refined structures, several methods were used to corroborate the small movements between the time-point models, including calculating the estimated coordinate uncertainties and monitoring the progression of the time-point Mb models from the CO-bound to the deoxy state.

Despite the low signal-to-noise ratio of the diffraction data, estimates of the accuracy of the refinements support our interpretation of the time-resolved structures. Ideally, errors for each of the coordinate shifts between the two structures would be calculated, but since the photolyzed time-point models contain coordinates of the CO-bound model in addition to the photolyzed coordinates, the two structures are not independent and the coordinate-shift errors are difficult to measure. Instead, estimates for the positional uncertainties of atoms with average temperature factors in the second 'laser off' model and the time-point models were estimated using the diffraction-component precision index (Cruickshank, 1999) as 0.19 and 0.05–0.07 Å, respectively. In comparison, the observed r.m.s.d.s between the time-point models and the CO-bound 'laser off' state are between 0.19 and 0.23 Å overall and between 0.29 and 0.52 Å in the heme pocket, well above the estimated uncertainties.

The distance difference matrix shows that the differences between a CO-bound and deoxy Mb are centralized to specific locations of the backbone, whereas by 3.16 µs after photolysis, only small dispersed dissimilarities remain. Comparison of the vector overlaps and angles for the heme pocket with those for the outer protein regions indicates that the structural evolution toward the deoxy state originates in the heme pocket and then propagates throughout the protein, supporting a 'proteinquake' model for conformational change, where the energy release from the breaking of the CO–iron bond is propagated from the heme outward to the rest of the protein (Ansari *et al.*, 1985). After 100 ps, the rearrangements lag in the C^{α} atoms of the protein without the heme pocket in comparison to the C^{α} atoms of the heme pocket and the lagging continues through the 3.16 µs time point. The values demonstrate that the majority of large-scale structural rearrangements occur by 3.16 ns, with minor changes continuing beyond 3.16 ns. The time-dependence of the coordinate model indicates a stretched-exponential progression from the CO-bound to the deoxy state, as observed spectroscopically by Jackson *et al.* (1994). The deoxy state is not fully realised by 3.16 µs, but comparisons between the photolyzed and deoxy-Mb coordinate models (Fig. 2*b*) show only minor structural differences.

The total photolysis is estimated to be 23%, but dissociated CO occupancies do not add up to 0.23 at every time point, presumably owing to heterogeneous diffusion of CO. Notably, this diminution occurs with the appearance of the CO on the proximal side of the heme in the 31.6 ns and later time points. Not all of the dissociated CO will reside in the binding sites during X-ray data collection and a percentage of the CO will be dispersed within the protein or released to the solvent. With the varying locations of the CO, some of its electron density also will go undetected owing to averaging over time and the unit cells of the conformational variants. Placement of the CO molecules was based on the $F_o - F_c$ maps, which showed clear difference electron density for CO placement consistent with previous experiments and simulations. The CO occupancies are difficult to quantify owing to the correlation between temperature factors and occupancy. We chose the maximum occupancies that were below 23% and yet kept the *B* factors from rising above 60 $Å^2$ and caused no negative or positive density to appear around the COs in the $F_o - F_c$ map. In general, reducing the given occupancies by more than 3–5% causes positive density to appear in both $F_o - F_c$ maps. This provides one estimate of the errors in the CO occupancies.

The large rotations in Phe29 mediate CO movement from the docking site to other cavities above the heme. The rotation is required to accommodate the CO in the primary docking site, but the strain on Phe29 shortens the lifespan of the intermediate and is relieved when CO is forced out into the other hydrophobic cavities. Similarly, the models show the 'gate-key' His64 in action: the distal histidine swings out toward the solvent and back toward its original position, perhaps providing a mechanism that allows CO to leave and enter the protein (Johnson *et al.*, 1989). Because they are less energetically favorable than their CO-bound conformations,

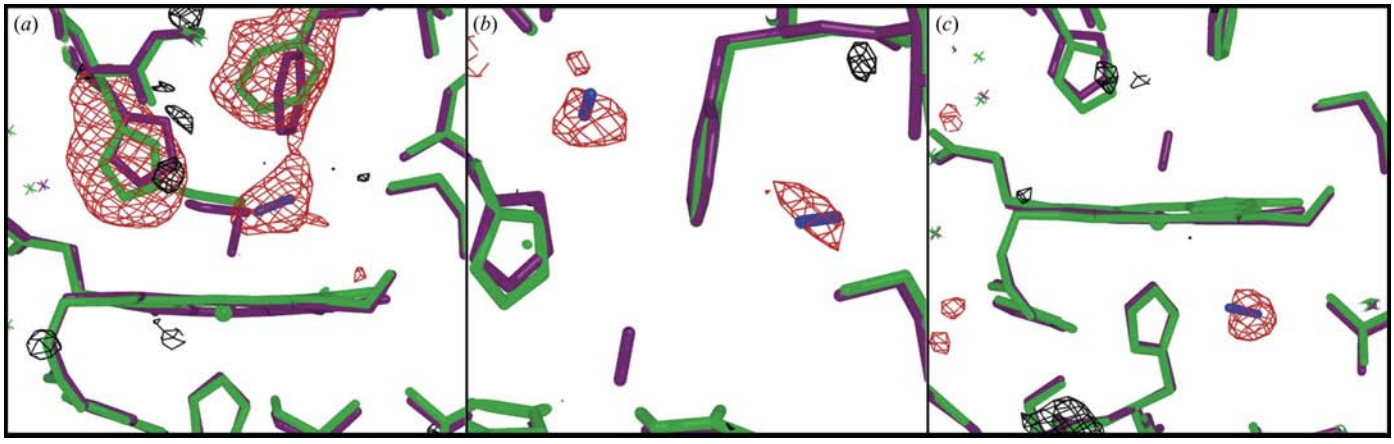

Figure 4
Electron-density maps of the various CO cavities. Electron-density maps were made by removing residues affected by CO photolysis and/or CO from the model and then calculating the σ_A-weighted $F_o - F_c$ maps. The labelled photolyzed residues and/or CO were reinserted to show that their placements fit with the $F_o - F_c$ density at 3σ. The CO-bound 'laser off' model is represented by the purple coordinates, the photolyzed time points are represented by the green coordinates and the photolyzed CO in blue. Positive density in the $F_o - F_c$ map is displayed in red and negative density is displayed in black at 3σ. (*a*) shows the 100 ps $F_o - F_c$ map with Phe29 and His64 from the photolyzed model and the primary docking site CO deleted. (*b*) shows the 316 ps $F_o - F_c$ map with only the photolyzed COs removed and (*c*) shows the 316 ns $F_o - F_c$ map with the CO removed from the proximal Xe1 site.

the intermediate rotations are short-lived: the phenyl group of Phe29 reverts back to its original position by 1 ns and, on a somewhat longer time scale, His64 swings back to within 6° of its original orientation and reduces its distance to the heme iron by the final time point.

Although the time-resolved models show the photolyzed CO binding in various hydrophobic cavities, they provide little information on the transient pathways taken by the ligand during migration between cavities. Since ligand translocation between hydrophobic cavities is expected to be asynchronous among the Mb molecules and fast compared with the residence time within those cavities, it is unlikely that any crystallographic method would be able to resolve CO transit from one cavity to another. Computational methods such as molecular-dynamics (MD) simulations may be helpful in charting possible pathways for ligand migration between internal cavities and its escape to the solvent. Recent molecular-dynamics studies hypothesize a potential route for the photolyzed CO to pass from the distal cavities to the proximal Xe-binding site 1, demonstrating the potential of MD to serve as a complement to experimental time-resolved studies (Bossa *et al.*, 2004; Hummer *et al.*, 2004). Taken together, the two techniques will expand the understanding of Mb ligand dissociation and the dynamics Mb undergoes during ligand release.

Funding was provided by NIH 5T32 GM08349, NIH T32 GM07215-29 and the University of Wisconsin Vilas Trust. This research was supported in part by the Intramural Research Program of the NIH, NIDDK. Thanks to Ryan M. Bannen for helping to develop the supplementary online movie.

References

Ansari, A., Berendzen, J., Bowne, S. F., Frauenfelder, H., Iben, I. E., Sauke, T. B., Shyamsunder, E. & Young, R. D. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 5000–5004.
Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. (1999). *J. Appl. Cryst.* **32**, 554–562.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bossa, C., Anselmi, M., Roccatano, D., Amadei, A., Vallone, B., Brunori, M. & Di Nola, A. (2004). *Biophys. J.* **86**, 3855–3862.
Bourgeois, D., Vallone, B., Schotte, F., Arcovito, A., Miele, A. E., Sciara, G., Wulff, M., Anfinrud, P. & Brunori, M. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 8704–8709.
Bourgeois, D., Wagner, U. & Wulff, M. (2000). *Acta Cryst.* D**56**, 973–985.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Brunori, M., Bourgeois, D. & Vallone, B. (2004). *J. Struct. Biol.* **147**, 223–234.
Campbell, J. W. (1995). *J. Appl. Cryst.* **28**, 228–236.
Carver, T. E., Brantley, R. E. Jr, Singleton, E. W., Arduini, R. M., Quillin, M. L., Phillips, G. N. Jr & Olson, J. S. (1992). *J. Biol. Chem.* **267**, 14443–14450.
Cruickshank, D. W. (1999). *Acta Cryst.* D**55**, 583–601.
Gibson, Q. H. & Ainsworth, S. (1957). *Nature (London)*, **180**, 1416–1417.
Hartmann, H., Zinser, S., Komninos, P., Schneider, R. T., Nienhaus, G. U. & Parak, F. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 7013–7016.
Helliwell, J. R., Nieh, Y. P., Habash, J., Faulder, P. F., Raftery, J., Cianci, M., Wulff, M. & Hadener, A. (2003). *Faraday Discuss.* **122**, 131–144.
Helliwell, J. R., Nieh, Y. P., Raftery, J., Cassetta, J., Habash, J., Carr, T., Ursby, T., Wulff, M., Thompson, A. W., Niemann, A. C. & Hadener, A. (1998). *J. Chem. Soc. Faraday Trans.* **94**, 2615–2622.
Hummer, G., Schotte, F. & Anfinrud, P. A. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 15330–15334.
Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.
Jackson, T. A., Lim, M. & Anfinrud, P. A. (1994). *Chem. Phys.* **180**, 131–140.
Johnson, K. A., Olson, J. S. & Phillips, G. N. Jr (1989). *J. Mol. Biol.* **207**, 459–463.
Kachalova, G. S., Popov, A. N. & Bartunik, H. D. (1999). *Science*, **284**, 473–476.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. (1958). *Nature (London)*, **181**, 662–666.
Lim, M., Jackson, T. A. & Anfinrud, P. A. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 5801–5804.
Lim, M., Jackson, T. A. & Anfinrud, P. A. (1995). *Science*, **269**, 962–966.
Lim, M., Jackson, T. A. & Anfinrud, P. A. (1997). *Nature Struct. Biol.* **4**, 209–214.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Moffat, K. (1998*a*). *Acta Cryst.* A**54**, 833–841.
Moffat, K. (1998*b*). *Nature Struct. Biol.* **5**, *Suppl.*, 641–643.
Moffat, K. (2001). *Chem. Rev.* **101**, 1569–1581.
Moffat, K. (2003). *Faraday Discuss.* **122**, 65–77.
Nienhaus, K., Deng, P., Kriegl, J. M. & Nienhaus, G. U. (2003). *Biochemistry*, **42**, 9647–9658.
Ostermann, A., Waschipky, R., Parak, F. G. & Nienhaus, G. U. (2000). *Nature (London)*, **404**, 205–208.
Perutz, M. F. (1960). *Brookhaven Symp. Biol.* **13**, 165–183.
Phillips, G. N. Jr, Arduini, R. M., Springer, B. A. & Sligar, S. G. (1990). *Proteins*, **7**, 358–365.
Radding, W. & Phillips, G. N. Jr (2004). *Bioessays*, **26**, 422–433.
Schlichting, I., Berendzen, J., Phillips, G. N. Jr & Sweet, R. M. (1994). *Nature (London)*, **371**, 808–812.
Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V., Soman, J., Olson, J. S., Phillips, G. N. Jr, Wulff, M. & Anfinrud, P. A. (2003). *Science*, **300**, 1944–1947.
Schotte, F., Soman, J., Olson, J. S., Wulff, M. & Anfinrud, P. A. (2004). *J. Struct. Biol.* **147**, 235–246.
Springer, B. A., Sligar, S. G., Olson, J. S. & Phillips, G. N. Jr (1994). *Chem. Rev.* **94**, 699–714.
Srajer, V., Ren, Z., Teng, T.-Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M. & Moffat, K. (2001). *Biochemistry*, **40**, 13802–13815.
Srajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M. & Moffat, K. (1996). *Science*, **274**, 1726–1729.
Teng, T.-Y., Srajer, V. & Moffat, K. (1997). *Biochemistry*, **36**, 12087–12100.
Terwilliger, T. C. & Berendzen, J. (1995). *Acta Cryst.* D**51**, 609–618.
Tilton, R. F. Jr, Kuntz, I. D. Jr & Petsko, G. A. (1984). *Biochemistry*, **23**, 2849–2857.