

Protein Conformational Relaxation and Ligand Migration in Myoglobin: A Nanosecond to Millisecond Molecular Movie from Time-Resolved Laue X-ray Diffraction[†]

Vukica Šrajer,^{*,‡,§} Zhong Ren,^{‡,§,△} Tsu-Yi Teng,^{‡,§} Marius Schmidt,[‡] Thomas Ursby,^{||,⊥} Dominique Bourgeois,^{||,⊗} Claude Pradervand,^{§,#} Wilfried Schildkamp,[§] Michael Wulff,^{||} and Keith Moffat^{*,‡,§}

Department of Biochemistry and Molecular Biology, The University of Chicago, 920 East 58th Street, Chicago, Illinois 60637, Consortium for Advanced Radiation Sources, The University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, and European Synchrotron Radiation Facility, BP 220, 38043 Grenoble Cedex, France

Received April 6, 2001; Revised Manuscript Received August 20, 2001

Ⓜ This paper contains enhanced objects available on the Internet at <http://pubs.acs.org/biochemistry>.

ABSTRACT: A time-resolved Laue X-ray diffraction technique has been used to explore protein relaxation and ligand migration at room temperature following photolysis of a single crystal of carbon monoxide-myoglobin. The CO ligand is photodissociated by a 7.5 ns laser pulse, and the subsequent structural changes are probed by 150 ps or 1 μ s X-ray pulses at 14 laser/X-ray delay times, ranging from 1 ns to 1.9 ms. Very fast heme and protein relaxation involving the E and F helices is evident from the data at a 1 ns time delay. The photodissociated CO molecules are detected at two locations: at a distal pocket docking site and at the Xe 1 binding site in the proximal pocket. The population by CO of the primary, distal site peaks at a 1 ns time delay and decays to half the peak value in 70 ns. The secondary, proximal docking site reaches its highest occupancy of 20% at \sim 100 ns and has a half-life of \sim 10 μ s. At \sim 100 ns, all CO molecules are accounted for within the protein: in one of these two docking sites or bound to the heme. Thereafter, the CO molecules migrate to the solvent from which they rebind to deoxymyoglobin in a bimolecular process with a second-order rate coefficient of $4.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$. Our results also demonstrate that structural changes as small as 0.2 Å and populations of CO docking sites of 10% can be detected by time-resolved X-ray diffraction.

The structure, function, and dynamics of the small globular oxygen storage protein myoglobin (Mb)¹ have been extensively investigated over many years by a large number of experimental and computational methods. The goal is to understand in general terms the processes of ligand dissociation, rebinding, recognition, and discrimination, and to explore ligand entrance and exit pathways in the framework of protein structure and dynamics, using myoglobin as a

specific example. Comprehensive studies have been carried out over the time scale from femtoseconds to seconds, under a broad range of experimental conditions such as temperature, pressure, solvent viscosity, and pH, for several species of wild-type myoglobins and for variants of important amino acid residues (1–24).

Many of these studies exploit the fact that the carbon monoxide complex of myoglobin (MbCO) is light-sensitive. Absorption of a photon ruptures the iron–carbon monoxide bond, initiating a series of spectroscopic and structural changes, ultimately followed by carbon monoxide rebinding and regeneration of MbCO. Ultrafast spectroscopic measurements explore directly ligand photolysis and subsequent rebinding. They also provide information about rates for those structural relaxation processes that are coupled to spectral properties of the molecule. Since spectroscopic properties are quite sensitive to coupled structural changes, the time courses of those changes can be accurately determined (14, 16, 17). However, in most cases, the actual nature of the structural change is not directly revealed and can only be inferred from the spectral changes. X-ray crystallography provides direct, detailed, and global structural information. Static crystal structures of several ligation states of myoglobin are known to near-atomic resolution (25, 26). The development of ultrafast time-resolved X-ray crystallography (27–31) has provided a new tool for obtaining direct, time-resolved global structural information at physiological tem-

[†] This work was supported by NIH Grants RR07707 and GM 36452 to K.M. and by EU BIOTECH Grant CT98-0415 to M.W.

* To whom correspondence should be addressed. Telephone: (773) 702-3603 (773) 702-2116. Fax: (773) 702-0439. E-mail: vusr@midway.uchicago.edu or moffat@cars.uchicago.edu.

[‡] Department of Biochemistry and Molecular Biology, The University of Chicago.

[§] Consortium for Advanced Radiation Sources, The University of Chicago.

[△] Present address: Renz Research Inc., 1570 Wicke Ave., Des Plaines, IL 60018.

^{||} European Synchrotron Radiation Facility.

[⊥] Present address: MAX-lab, Lund University, P.O. Box 118, S-22100 Lund, Sweden.

[⊗] Present address: European Synchrotron Radiation Facility and LCCP IBS, UMR 9015, 41 Avenue Jules Horowitz, 38027 Grenoble Cedex 1, France.

[#] Present address: Swiss Light Source, Paul Scherrer Institut, 5232 Villigen PSI, Switzerland.

¹ Abbreviations: Mb, myoglobin; MbCO, carbon monoxide complex of myoglobin; ESRF, European Synchrotron Radiation Facility; Mb*, MbCO photoproduct; INS, isomorphous noise suppression; PDB, Protein Data Bank.

peratures. Using this tool, it is not necessary to trap intermediate protein conformations, either by freezing or by chemical modification, to study them by X-ray diffraction methods. Artifacts arising from the trapping process, which may be particularly evident for very short-lived intermediates (32), can then be avoided.

Myoglobin serves as a particularly challenging test with which to explore the capabilities and limitations of time-resolved X-ray diffraction. The near-atomic resolution structures of myoglobin with and without ligand (25, 26) indicate that only relatively small structural relaxation of the heme and the globin, limited in both spatial extent and magnitude, is expected to accompany ligand photolysis. Further, the entire photolysis, relaxation, and rebinding processes in the crystal are complete in less than 5 ms at room temperature, which requires that all measurements be made rapidly and with excellent time resolution. In the first nanosecond time-resolved X-ray diffraction study (28), we demonstrated that complete, redundant, and good-quality Laue crystallographic data can be collected using single, 150 ps, X-ray pulses from a third-generation synchrotron X-ray source, such as the ESRF (Grenoble, France). In these pump–probe experiments, the CO was photodissociated from a single crystal of MbCO by a 7.5 ns laser pulse and the structure of the photoproduct (Mb*) probed by a delayed X-ray pulse. At a 4 ns delay, the Mb* minus MbCO difference Fourier map revealed a large negative signal associated with the loss of ligand due to photolysis (28). This signal diminished in magnitude in a time series of five difference maps obtained at longer time delays, between 1 μ s and 1.9 ms. Heme relaxation, manifested by motion of the iron atom out of the heme plane toward the proximal histidine, was observed and a potential docking site of the photodissociated CO molecule in the heme pocket identified. Although the maps provided hints of small structural changes associated with globin relaxation, they could not be unambiguously detected and their time evolution was uncertain.

We report here results from further nanosecond time-resolved studies of MbCO photolysis. The goal of these new measurements was 3-fold: to obtain difference maps with a higher signal-to-noise (S/N) ratio that can reveal more clearly small structural changes; to examine structural changes in the time interval between 4 ns and 1 μ s, not covered by our previous data; and to examine sufficient time delays to enable the time evolution of significant structural features to be identified and modeled. An improved data collection strategy resulted in better raw data, and data processing was significantly improved. Careful global scaling of the data from different crystals and different experiments, application of an improved weighting method for the difference structure factor amplitudes, and introduction of a novel difference electron density modification technique (Z. Ren et al., unpublished results) result in a substantially enhanced signal-to-noise ratio in the final difference Fourier maps. Data at additional time delays allow us to examine the continuity of the difference electron density features in time. Identification of noise levels enables us to assess the significance of small structural changes. The improved quality of the maps and better coverage in time reveal small and subtle protein conformational changes that were undetectable in our previous studies and enable us to model the overall course of the reaction.

MATERIALS AND METHODS

Preparation of Crystals, Photolysis, and X-ray Data Collection. Sperm whale metMb crystals were grown in the $P2_1$ monoclinic form at pH 6 (33). Crystals were converted to MbCO as described by Teng et al. (34). Crystals were sealed in capillaries in air. Reoxidation to metMb was prevented by mounting the crystals quickly (in 10–20 s) and by placing a saturated solution of sodium dithionite in the capillary prior to sealing it. Once mounted, crystals were stable in the MbCO form for months as confirmed by measuring their optical spectra by a microspectrophotometer (35). Crystals were mounted 1 or 2 weeks in advance and their spectra routinely checked a few days before the X-ray data collection. Only crystals with a metMb content of $\leq 5\%$ were used for data collection. Typical crystal dimensions were $450 \mu\text{m} \times 350 \mu\text{m} \times 100 \mu\text{m}$.

Crystals were photolyzed by 7.5 ns laser pulses at 630 nm. Pulses were produced by a Nd:YAG pumped dye laser (Continuum NY61-10/ND60) with the DCM dye. Nonpolarized laser light was delivered to the sample by an optical fiber. The light was focused to a 1.5 mm diameter spot at the crystal. Since this focal spot was substantially larger than the crystal, transverse gradients of light intensity and photochemistry were minimized. A total laser pulse energy of ~ 5 mJ was used for photolysis, but less than 0.1 mJ (or 2%) of the laser pulse energy was actually absorbed by a typical crystal due to its small size and low optical density at 630 nm. The wavelength was chosen to ensure that the optical density of the crystals was < 0.2 OD and hence to minimize longitudinal gradients and photolyze uniformly through the crystals. We estimate that the absorbed laser pulse energy resulted in ~ 0.7 photons being absorbed per molecule per laser pulse and an overall temperature increase in the crystals of ~ 1 K, if we assume that half of the absorbed energy dissipates as heat (*I*).

The X-ray diffraction data were collected at the ID09 beamline at the ESRF, using the experimental setup described by Bourgeois et al. (27). Data (Table 1) were collected in experiments conducted in 1995 (labeled 1 in Table 1; see ref 28) and in 1996 (labeled 2). We used either wiggler W70 alone (1995 data) or a wiggler/undulator W70/U46 tandem (1996 data) as the X-ray source. All data were collected at 18–20 °C by bathing the crystal and capillary in a temperature-controlled gas stream. Data were collected in pump–probe style experiments in which MbCO crystals were photolyzed by a laser pulse and probed by a suitably delayed X-ray pulse. The delay time is defined as the time between the center of the laser pulse and the center of the X-ray pulse. Although delay times were as short as 1 ns, the time resolution of the experiment is determined by the 7.5 ns duration of the laser pulses. For delays shorter than 1 μ s, single 150 ps X-ray probe pulses were used in the single-bunch operation mode of the ESRF storage ring (15 mA maximum current). For all longer delays, the 1 μ s “super” X-ray pulse was used in the 1/3 filling mode (150 mA total current). For each laser/X-ray delay time, a data set that was complete and with suitable redundancy was collected. For each angular setting, the crystal was exposed to the laser/X-ray pulse pair as many times as necessary to accumulate a high-quality diffraction pattern prior to detector readout. Several exposures were typically needed: 3–30 exposures

Table 1: Crystallographic Data Collection Parameters

crystal	state	laser/X-ray time delay	laser energy (mJ)	insertion device and mode ^a	no. of images per data set	angular increment (deg)	no. of X-ray exposures per image	experiment number ^b
1	Mb*	1 ns	5.5	W70+U46/sb [§]	42	4.5	15	2
2	Mb*	4 ns	<13 ^c	W70/sb	49	4	3	1
3	Mb*	7 ns	5.5	W70+U46/sb	42	4.5	30	2
4	Mb*	7 ns	6	W70+U46/sb	42	4.5	5	2
4	Mb*	18 ns	6	W70+U46/sb	42	4.5	5	2
4	Mb*	45 ns	6	W70+U46/sb	42	4.5	5	2
5	Mb*	90 ns	6	W70+U46/sb	21	9	10	2
5	Mb*	362 ns	6	W70+U46/sb	42	4.5	10	2
6	Mb*	1.0 μ s	<13 ^c	W70/spb ^a	49	4	1	1
7	Mb*	1.0 μ s	<13 ^c	W70/spb	48	4	1	1
8	Mb*	1.9 μ s	6	W70+U46/spb	48	4	10	2
9	Mb*	7.5 μ s	<13 ^c	W70/spb	50	4	1	1
10	Mb*	25.0 μ s	6	W70+U46/spb	82	2	3	2
6	Mb*	50.5 μ s	<13 ^c	W70/spb	49	4	1	1
7	Mb*	350.0 μ s	<13 ^c	W70/spb	48	4	1	1
9	Mb*	1.9 ms	<13 ^c	W70/spb	48	4	1	1
2	MbCO	—	—	W70+U46/sb	50	4	3	1
6	MbCO	—	—	W70/spb	49	4	1	1
7	MbCO	—	—	W70/spb	47	4	1	1
9	MbCO	—	—	W70/spb	50	4	1	1
11	MbCO	—	—	W70+U46/sb	48	4	10	2

^a sb, single-bunch mode; spb, super bunch mode. ^b Experiment number refers to 1995 (1) and 1996 (2) experiments. ^c The laser pulse energy as recorded at the time of the experiment. However, we believe the energy was incorrectly measured since in all other subsequent experiments an energy of >10 mJ (with the same beam size) was too high and would damage the crystals.

in single-bunch mode and 1–10 in super pulse mode (Table 1). Consecutive laser exposures were separated by at least 5 s to allow the complete recovery of the crystal, lattice and molecules, to the initial state.

A typical MbCO Laue data set consists of 40–50 frames, 4–5° apart in crystal angular setting (Table 1). Data collection parameters for each set are given in Table 1. No significant X-ray or laser radiation damage was observed, as judged by the variation from data set to data set of the completeness and R_{merge} of the last resolution shell, in the cases where several data sets were collected on the same crystal. In the 1995 experiment, we interleaved MbCO images with the Mb* images, which provided an MbCO (dark) data set, corresponding directly to each Mb* data set, on the same crystal. We used these MbCO structure factor amplitudes for difference map calculations and thus eliminated systematic errors that could arise from comparing data collected on different crystals. This strategy also eliminates any very slow time-dependent changes (such as possible radiation damage or light- or X-ray-induced oxidation) from the Mb* minus MbCO difference maps. However, in the 1996 experiment, we concentrated on collecting as many time delays as possible. In this case, we chose one particular, high-quality MbCO data set (crystal 11 in Table 1) as a reference for all Mb* data sets collected on various crystals.

Data Reduction and Global Scaling. Data were processed using the *LaueView* software package (36). Data reduction statistics for all data sets are given in Table 2. Most data sets were integrated to 1.8 or 1.7 Å resolution, with the exception of three data sets: 7 ns (crystal 3), 1.9 μ s, and 25 μ s, which were integrated to 1.6, 2.0, and 1.9 Å, respectively. The completeness of the data is presented (Table 2) by resolution range. The overall redundancy is high, 4.6–7.7, for all data sets except for the 90 ns data set (crystal 5) where a 9° angular increment in crystal orientation was used rather than the standard 4°. Redundancy is accordingly ~2 times lower.

Each MbCO and Mb* data set was scaled to calculated MbCO structure factor amplitudes by the *LaueView* program. The MbCO structure amplitudes were calculated using the room-temperature MbCO model obtained by refining the initial, 40 K MbCO model 1AJG (34) against the room-temperature MbCO Laue data set from crystal 6. The global scaling thus places all data sets on the same, absolute scale.

Difference Maps. Difference electron density maps were calculated with time-dependent structure factor amplitudes [$\Delta F(t) = |F_{\text{Mb}^*}(t)| - |F_{\text{MbCO}}|$] obtained from the combined single and multiple reflections (Table 2). The $|F_{\text{MbCO}}|$ structure factor amplitudes are those of crystal 11 except for crystals where a corresponding MbCO data set was collected. The phases were calculated using the MbCO model derived from the crystal 6 MbCO data (see Data Reduction and Global Scaling above). The iron atom, heme nitrogen atoms, the CO ligand, and residues His64 and His93 were omitted from phase calculations to reduce the bias from the MbCO phases in the heme region. We refer to these maps as omit difference maps.

To further improve the signal-to-noise ratio of the difference maps, we calculated weighted omit difference maps using a weighting scheme similar to that employed by Šrajer et al. (28) and described by Ursby et al. (37). The weight for the difference structure factor amplitude was constructed by comparing both its magnitude [$|\Delta F(t)|$] and its estimated standard deviation [$\sigma_{|\Delta F(t)|}$] to their mean values [$\langle |\Delta F(t)| \rangle$] and $\langle \sigma_{|\Delta F(t)|} \rangle$]:

$$W_{|\Delta F(t)|} = 1/[1 + \sigma_{|\Delta F(t)|}^2 / \langle \sigma_{|\Delta F(t)|} \rangle^2 + |\Delta F(t)|^2 / \langle |\Delta F(t)| \rangle^2] \tag{1}$$

where the variance $\sigma_{|\Delta F(t)|}^2$ is the sum of the variances $\sigma_{|F(\text{Mb}^*)}|^2$ and $\sigma_{|F(\text{MbCO})}|^2$ of $|F_{\text{Mb}^*}(t)|$ and $|F_{\text{MbCO}}|$, respectively. The last term was added to reduce the contribution of abnormally large values of $|\Delta F(t)|$ that have not been properly weighted down by the first term, in those cases

Table 2: Data Reduction Statistics

crystal	delay	integration limit (Å)	single reflections					single and harmonic reflections		
			$R(F^2)$ (%)	$R(F)$ (%)	no. of observations	no. of unique reflections	redundancy	no. of unique reflections	overall completeness ^a (%)	last shell completeness (%)
1	1 ns	1.7	10.8	7.8	90519	11730	7.7	12782	87.0	59.9 (1.78–1.7 Å)
2	4 ns	1.8	10.3	6.8	34801	7289	4.8	8143	66.7	34.7 (1.99–1.9 Å)
3	7 ns	1.6	11.2	7.9	86452	12655	6.8	13913	79.5	39.9 (1.67–1.6 Å)
4 ^b	7 ns	1.7	12.4	9.2	78331	—	—	13079	89.1	66.9 (1.78–1.7 Å)
4	18 ns	1.7	12.8	9.5	73352	—	—	12987	88.4	64.9 (1.78–1.7 Å)
4	45 ns	1.7	14.2	9.8	69715	—	—	12033	82.1	44.4 (1.78–1.7 Å)
5	90 ns	1.8	10.9	7.9	22887	7902	2.9	9234	74.7	36.4 (1.88–1.8 Å)
5	362 ns	1.8	10.7	7.8	63078	10261	6.1	11020	88.8	58.9 (1.88–1.8 Å)
6	1 μs	1.8	9.8	6.8	34728	7461	4.7	8330	67.5	32.6 (1.99–1.9 Å)
7	1 μs	1.8	8.6	5.8	34241	7292	4.7	8108	66.3	33.9 (1.99–1.9 Å)
8	1.9 μs	2.0	11.1	7.4	36426	6905	5.3	7677	84.9	62.5 (2.09–2.0 Å)
9	7.5 μs	1.8	11.4	7.5	34010	7118	4.8	7932	64.7	29.1 (1.99–1.9 Å)
10	25 μs	1.9	18.0	11.9	55856	7456	7.4	8383	79.8	34.0 (1.99–1.9 Å)
6	50.5 μs	1.8	10.1	6.8	35232	7461	4.7	8280	68.1	32.4 (1.99–1.9 Å)
7	350 μs	1.8	9.2	6.0	34772	7236	4.8	8057	65.9	34.9 (1.99–1.9 Å)
9	1.9 ms	1.8	11.8	7.7	32958	7070	4.7	7871	64.2	26.7 (1.99–1.9 Å)
2	dark	1.8	10.2	6.9	33063	7160	4.6	8032	65.7	35.3 (1.99–1.9 Å)
6	dark	1.8	10.1	6.6	36824	7556	4.9	8351	67.8	33.9 (1.99–1.9 Å)
7	dark	1.8	9.1	5.9	34564	7237	4.8	8053	65.9	32.8 (1.99–1.9 Å)
9	dark	1.8	10.7	7.0	35497	7189	4.9	8058	65.7	29.1 (1.99–1.9 Å)
11	dark	1.7	7.7	5.5	90916	11889	7.6	12387	84.7	37.9 (1.78–1.7 Å)

^a Overall completeness to the integration limit. ^b The crystal 4 data sets had to be divided into two parts that were scaled separately and merged after the harmonic deconvolution. The information about unique single reflections and their redundancy is therefore not available.

where the estimate of the standard deviation $\sigma_{|\Delta F(r)|}$ may be erroneous. In data sets of somewhat poorer quality and for those with a weaker signal, the effect of weighting is to greatly increase the signal-to-noise ratio. In such cases, weighting makes a critical difference between not observing and confidently observing the signal, as illustrated in Figure 1. The figure shows a standard, nonweighted difference omit map (Figure 1a) and a weighted map (Figure 1b) of the heme region, at a time delay of 50.5 μs. Each map is contoured at the value of $\pm 3\sigma$ appropriate to that map, where σ is the rms value of the difference electron density $\Delta\rho$ across the entire asymmetric unit. This contouring procedure ensures that the signal is displayed in a consistent manner relative to the noise for all maps. As discussed below, the σ value for each map is primarily determined by the noise. The negative feature above the heme in both maps corresponds to the loss of the bound CO molecule upon photolysis. In the standard map, this feature is barely significant at -3.3σ , while in the weighted map, the same feature is quite prominent and clearly significant at -5.7σ .

In a further attempt to improve the signal-to-noise ratio in difference maps, we applied a density modification method to the difference electron density maps (Z. Ren et al., unpublished results; 38). We refer to this method as the isomorphous noise suppression (INS) method. The method modifies both the amplitudes and phases of the difference structure factors. Standard difference maps are calculated using phases derived from the known MbCO structure, since the true phases for the difference structure factors are not known. The known and true phases are in general completely uncorrelated (39). The INS method derives a new set of amplitudes and phases for the difference structure factors by modifying the initially calculated difference electron density according to the following protocol. Difference density was not modified if it was within 3 Å of any atom of residues 42–46 (CD loop), 58–70 (part of the E helix),

and 80–107 (part of the EF loop, the F helix, the FG loop, and part of the G helix) or within 5 Å of any heme, CO, or solvent SO₄ atoms. For all remaining regions of the protein, the difference density was reduced to 62% of its initial value. In the solvent region, the difference density was set to zero. The solvent region is defined as locations more than 4 Å from any protein, heme, CO, SO₄, or H₂O atoms, and from the four Xe binding sites (4). The difference density modification within the boundaries of the protein has the effect of enhancing the signal and reducing the noise in those regions where the most significant structural differences are expected and where difference density was not altered. A Fourier transform of the modified difference electron density map provides new amplitude and phase information for the difference structure factors. Difference structure factors were then further modified in several iterations according to the estimated standard deviations of the measured difference amplitudes (Z. Ren et al., unpublished results). When convergence was achieved, the final difference structure factors were Fourier transformed to yield improved difference electron density maps. We refer to this type of map as an INS difference map. Improvement in the S/N ratio as a result of the INS procedure complements the improvement resulting from the weighting procedure, and both contribute to the S/N enhancement. In the case of the 50.5 μs difference map shown in Figure 1, a major S/N improvement is accomplished by the weighting procedure (-3.3σ to -5.7σ) but only a minor further improvement by the subsequent INS procedure (-5.7σ to -6σ ; figure not shown). In the case of some other maps, however, the INS produces a more marked improvement.

The difference Fourier maps we present are identified as either weighted omit difference maps or weighted omit INS difference maps. When examining the heme and adjacent regions where a signal is likely to be observed, we used weighted omit INS difference maps to enhance the signal

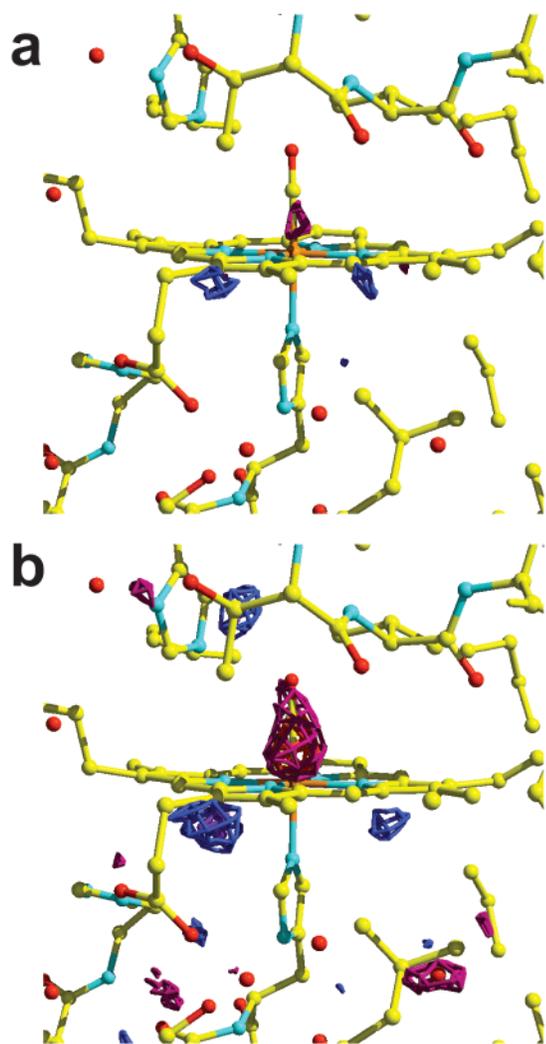


FIGURE 1: Mb* minus MbCO difference map of the heme region at 50.5 μ s calculated using standard structure factor difference amplitudes (a) and weighted (eq 1) difference amplitudes (b). Maps are calculated to 1.8 \AA resolution and contoured at $\pm 3\sigma$. Negative density (loss of electrons) is shown in red and positive density (gain of electrons) in blue, in this and all subsequent difference maps. All maps were generated using XtalView and Raster 3D (51, 52).

and determine its time course. When checking more remote protein regions for potential signal, we used weighted omit difference maps for which the original electron density has not been modified anywhere within the protein. We cannot use INS maps in this case since the weak signal, if present, would have been reduced in magnitude in the regions where electron density was modified in the INS maps.

Time Course of Structural Changes. The time course of structural changes can be described by empirical, structural rate constants. In the general case, the observed structure at any time delay represents a mixture of structurally distinct, time-independent intermediates whose concentrations vary with time (30, 40). The ultimate goal of the analysis of time-resolved data is to determine structures of the time-independent intermediates, based on these empirical rates and a proposed reaction scheme. This approach applies only if a chemical kinetic model holds (32) and is practically feasible only when there is a limited number of structural intermediates whose peak populations are relatively well-separated in time. If, however, a more complex model holds

in which structural changes are localized and many parts of the molecule change with their own, not very different rates, the hypothesis of distinct, time-independent structural intermediates breaks down. A simple chemical kinetic model implies that experimental observables will vary in time as a sum of exponentials, but a more complex model will yield a stretched-exponential variation.

The time course of structural changes can in principle be examined both in real and in reciprocal space. By the Fourier transform relationship, the contributions of each structural change are always distributed throughout reciprocal space even when they are tightly localized in real space. The overall changes in time in the structure factor amplitudes are, therefore, expected to be small and prone to be masked by noise even when the real space structural changes are relatively large but localized. On the other hand, even very small structural changes can be clearly detected in real space when they are localized, despite noise in the measured structure factor amplitudes. It is evidently preferable to follow the time evolution in real space rather than reciprocal space. We therefore extracted the empirical rates of structural changes by examining difference electron density maps rather than structure factor amplitudes.

Visual inspection of the time sequence of maps immediately indicates the regions of the molecule that exhibit major structural changes. We identify significant features in these regions, integrate their electron density, and fit the time course of their electron content with an appropriate mathematical function, typically a sum of exponentials or stretched exponentials (see the Results and Discussion). Since the fit function interpolates between data points in time, a molecular movie can be constructed. The movie is constructed by applying the rates identified from the real space fit of highly significant features to fit all structure factor amplitudes in reciprocal space, as a (continuous) function of time. Each frame of the movie is then obtained by Fourier transformation of the fitted structure factors at the desired time point.

Noise Levels and Significance of Features in the Difference Maps. The difference maps clearly contain substantial noise in addition to a structural signal. Quantitation of the magnitude of the noise, and an understanding of how it varies across the asymmetric unit and in time, is essential to proper interpretation of the maps. What features in the difference maps are significant and arise from signal rather than noise? To address this key question quantitatively, we examined the distribution of $\Delta\rho$ in the crystallographic asymmetric unit for our weighted difference maps and determined that the distribution is well described by a Gaussian whose standard deviation σ does not correlate with the magnitude of the signal (data not shown). This leads to two conclusions: the rms value of the difference electron density for the asymmetric unit is mainly determined by the noise, and the noise is random. We therefore expect noise above 3σ with a probability of 0.0027. This, however, still results in an appreciable number of noise features above 3σ in a difference map. For a complete 1.8 \AA myoglobin data set, for example, the observed ~ 12000 unique reflections result in the same number of independent difference map grid points per asymmetric unit. It can be then expected that the value of $\Delta\rho$ at 32 grid points will exceed 3σ due to random noise alone. In the more finely sampled regular grid we are using,

this could translate into 32 features above 3σ . However, there will be only about one grid point above the 4σ level due to random noise (probability of 0.000063). Although the $\Delta\rho$ distribution is well described by a Gaussian, deviations from the Gaussian distribution due to the signal clearly are present in the data and can be detected at the wings of the distribution above 3σ . As expected, these deviations diminish with time as the signal decays.

A second, important criterion for evaluating the significance of a difference electron density feature is based on whether it occurs at a chemically plausible location. For example, a negative density directly located on an atom in the MbCO structure is more likely to be signal, while positive density there is not. Positive and negative densities flanking a group of atoms are likely to be signal but, if located far from any atoms, are likely to be noise.

Finally, a further discrimination between real features and noise can be accomplished by examining the connectivity of features in time. Real features tend to persist in time over several, temporally adjacent maps. In contrast, noise features tend to be highly variable and uncorrelated from map to map. In addition to providing information about the rates of structural changes in the molecule, fitting of the time course of real-space features by a mathematical function provides smoothing in the time domain. However, systematic features that do not represent signal could also be present in temporally adjacent maps and cannot be eliminated this way.

In the case of the integrated electron content of a feature in a difference density map, additional sources of error need to be considered. A problem arises in defining the exact boundaries of a feature since they may vary with time due to the noise. In most cases, we identified the boundary and position of the integration region of a feature by examining the map where the feature is most prominent, and then kept the boundary and position constant over all maps. In more complicated cases, the location of a feature may shift with time, indicating a further structural change, or another feature may extend into the region formerly occupied by the feature of interest.

To estimate the errors associated with the integrated electron content of features we consider to be signal, we integrated regions of the same size at random locations in the unit cell and examined the distributions of the values that we obtained. We find these distributions also to be well described by a Gaussian form (data not shown). Hence, we define the error of the integrated electron content of the signal features as the standard deviation of such a distribution. This standard deviations are presented as error bars in Figures 5–7 and 8a.

RESULTS

One Nanosecond Difference Map: The Earliest Photoinduced Structural Changes. The weighted omit INS difference Fourier map of the heme region and two adjacent E (distal) and F (proximal) helices at 1 ns is shown in Figure 2a. The reference map ($\Delta F = |F_{\text{deoxyMb}}| - |F_{\text{MbCO}}|$) is shown for comparison in Figure 2b. This map was calculated to the same, 1.8 Å resolution using the high-resolution 1BZP deoxyMb and 1BZR MbCO PDB models (25) and reflections that match the observed reflections of the MbCO data set (crystal 11).

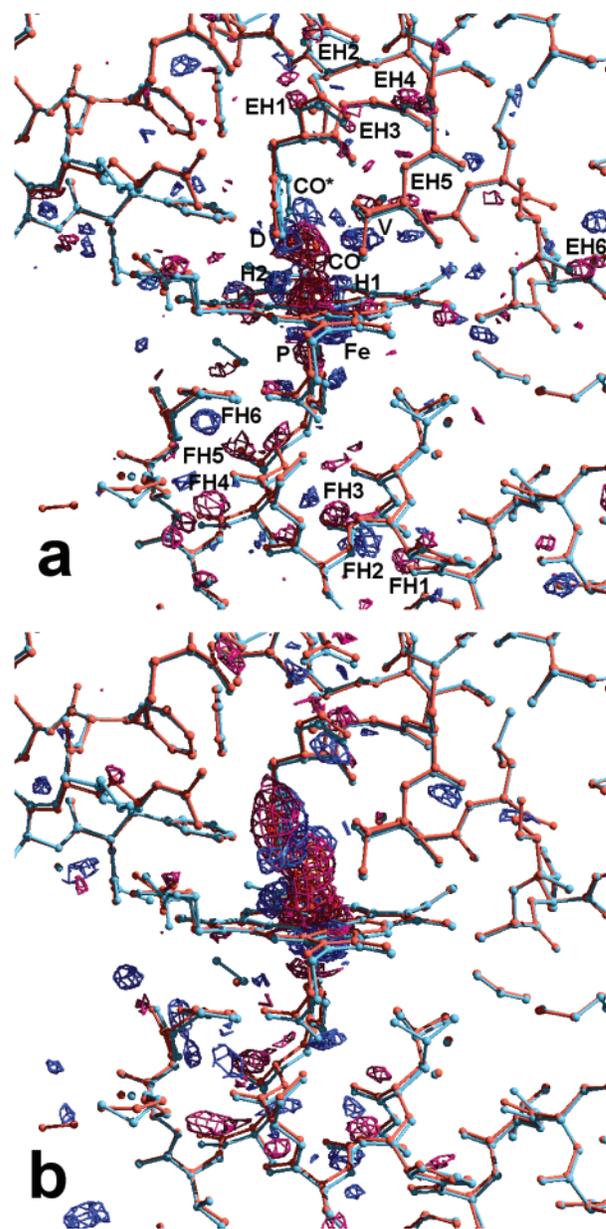


FIGURE 2: (a) Difference Fourier map of the heme region at a 1 ns laser/X-ray pulse time delay. The weighted omit INS map is calculated to 1.8 Å resolution and contoured at $\pm 3.4\sigma$. Near-atomic resolution models (1.15 Å) from PDB entries 1BZR (MbCO) and 1BZP (deoxyMb) (25), are shown in red and blue, respectively. The strongest features represent the loss of the CO-bound ligand (CO), the Fe displacement from the heme plane (Fe), the proximal histidine displacement toward the F helix (P), the swing of the distal histidine toward the bound CO location (D), and the location of the photodissociated ligand (CO*). Other smaller features indicate changes at the heme and displacement of E and F helices. These features are labeled H1–H3, EH1–EH6, and FH1–FH6, respectively. (b) Reference difference Fourier map calculated from 1BZP and 1BZR. The map is calculated to 1.8 Å resolution and contoured at $\pm 3.2\sigma$.

As in the 4 ns difference map (28), the most significant features in the 1 ns map are a large negative density at the CO-bound location, labeled CO, and a large positive density below the heme, labeled Fe. The highly significant negative density, at -14.2σ in this INS difference map, represents loss of the bound CO molecule upon photolysis. The extent of photolysis was estimated by comparing the integrated electron content of this feature with that of the corresponding

feature in the reference map (Figure 2b). Consistent with the 4 ns results (28), we find that $40 \pm 10\%$ of the MbCO molecules in the crystal are photolyzed at 1 ns.

The large positive feature labeled Fe, at $+9.7\sigma$, results from the accompanying motion of the Fe atom out of the heme plane upon photolysis. Its magnitude is also consistent with $\sim 40\%$ photolysis. Part of the negative CO-labeled feature closest to the heme also arises from the Fe motion; these positive and negative features are coupled. Notably, the positive Fe feature is not located on the heme normal that passes through the iron position in MbCO but is displaced from it, indicating the direction of the overall Fe and heme motion. Two views of the heme in MbCO are shown in panels a and b of Figure 3: looking along the Fe–His93 NE2 bond (Figure 3a) and looking along the pyrrole NB–ND axis (Figure 3b). The features in the difference map are consistent in magnitude and location with those arising from heme tilt and translation in going from the MbCO structure (shown in red) to deoxyMb (blue) (25). The motion is described as a 4° tilt of the nitrogen plane around the NB–ND axis and a 0.19 \AA translation in the direction from pyrrole ring C to A. The location of the Fe density in our difference map (Figure 3a) indicates that heme motion upon loss of ligand occurs mainly in the direction from NC to NA. The Fe density is, however, also displaced toward the pyrrole B, with the highest density located in the A/B quadrant of the heme. Several other features also indicate heme motion: two positive features, H1 and H2 (Figures 2a and 3), above the heme plane and weaker, negative densities at the rim of the heme. They are consistent with the tilt of the heme plane around the NB–ND axis, in which the C-pyrrole ring (behind the CO feature in Figure 2a) moves toward the distal side, as also more clearly shown in Figure 3b. The H1 and H2 features coincide with the extended positive density above the heme plane and behind the CO feature in the reference map (Figure 2b). In summary, heme features observed in the 1 ns map indicate that heme translation and rotation occurred by 1 ns and that these heme motions are consistent with the static structures of MbCO and deoxyMb.

Motion of the distal histidine, His64, toward the location formerly occupied by the CO ligand is represented by a positive feature labeled D (5.1σ) in Figures 2a and 3b. The integrated electron content of this positive feature as compared to the corresponding feature in the reference map (Figure 2b) also suggests $\sim 40\%$ photolysis.

A positive feature at 4.6σ , labeled CO*, is present in the distal, ligand pocket at the location identical to that of the feature noted earlier (28) and proposed to represent a docking site of the outgoing, photodissociated CO molecule. The feature (Figure 4) is located 2.18 \AA from the O atom of the bound CO molecule in the 1BZR MbCO structure (25), and $\sim 1 \text{ \AA}$ away from the site of the water molecule in the deoxyMb distal pocket (red ball, labeled O in Figure 4) and from the center of the photodissociated CO molecule at 40 K (green bar in Figure 4; 35). Side chains flanking this distal pocket docking site include Val68, Leu29, Ile107, and Phe43 (Figure 4). By comparing the magnitudes of the CO and CO* features of the 1 ns map, we conclude that at 1 ns the CO* docking site contains $\sim 45\%$ of the photodissociated CO molecules, that is, $\sim 22\%$ of the total number of CO molecules. This estimate is an upper limit which assumes that only the more distal part of the peanut-shaped CO-

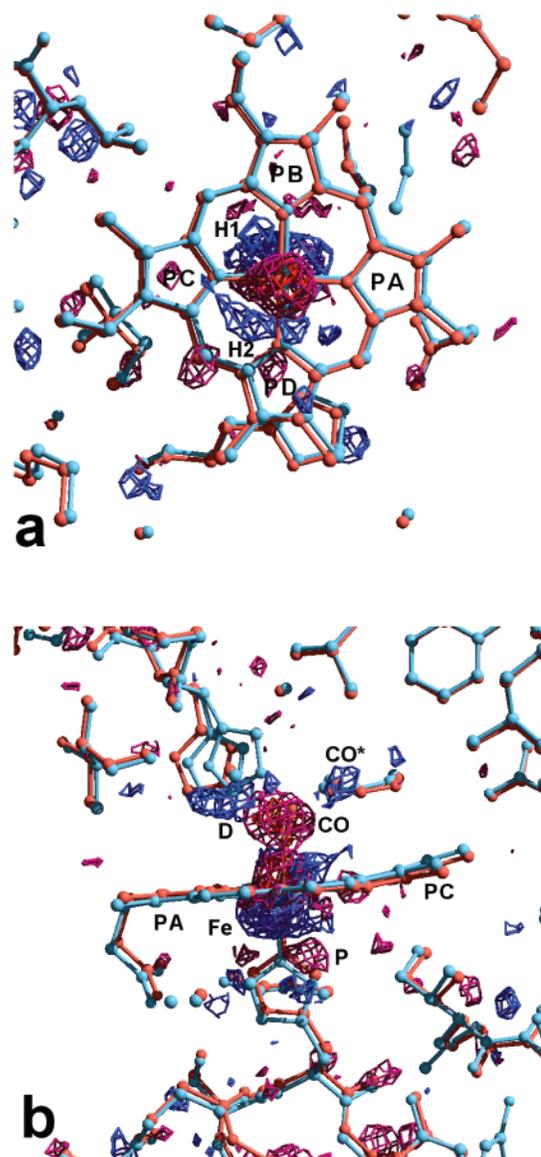


FIGURE 3: INS difference map of the heme region at 1 ns, contoured at $\pm 3.4\sigma$. The MbCO and deoxyMb models (26) are shown in red and blue, respectively. Heme pyrrole rings are labeled PA, PB, PC, and PD. (a) The map is viewed along the Fe–His93 bond. The shape of the positive, Fe-associated electron density below the heme indicates the direction of the heme translation from pyrrole C to pyrrole A, and toward pyrrole B. (b) Map viewed along the NB–ND axis. Differences between high-resolution MbCO and deoxyMb models (26) involve tilting of the heme, downward motion of the proximal His93, and swinging of the distal His64.

labeled feature arises from loss of the CO molecule, while the lower part arises from the Fe displacement.

The weighted omit INS difference electron density maps allow us to unambiguously detect more subtle but essential structural changes in the globin that were not clearly observed in the earlier maps (28). Several features, labeled FH1–FH6, with the largest feature (FH4) at -5.6σ , indicate that by 1 ns the entire F helix is already displaced, at least partially, as a result of the Fe displacement from the heme plane. The corresponding downward motion of the proximal histidine His93 is itself indicated by the negative feature, labeled P. The features indicating the F helix motion are adjacent to the main chain carbonyl oxygens of Lys87 (labeled FH1), Pro88 (FH2), Ala90 (FH3), Ala94 (FH4), and His93 (FH5).

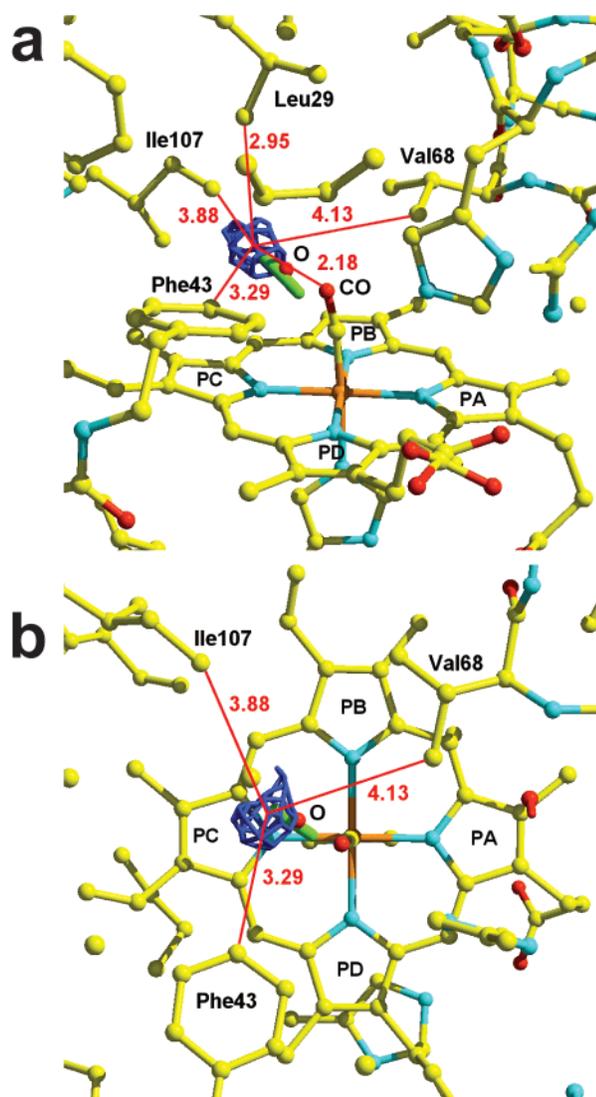


FIGURE 4: Docking site of the photodissociated CO molecule in the distal pocket. Distances (in angstroms) from the closest residues and from the O of the bound CO molecule are shown for 1BZR. The heme pocket water molecule of the deoxy structure is shown as a red ball, labeled O. The location of the photodissociated CO molecule at 40 K (35) is shown as a green bar. Two views of the site are shown in panels a and b.

Both positive and negative features exceeding 3σ are observed for Pro88 (FH2) and Ala90 (FH3), clearly indicating a downward motion of the F helix backbone, consistent with the static MbCO and deoxyMb structures (25). Several features on the distal side indicate that by 1 ns the distal E helix has also moved down, toward the heme plane. The features labeled EH1–EH6 are again adjacent to carbonyl oxygens or amide nitrogens: His64 N (labeled EH1), Lys63 N (EH2), His64 O (EH3), Val66 N (EH4), Val68 N (EH5), and Thr70 O (EH6).

Another relatively strong positive feature labeled V (Figure 2a) in the distal pocket suggests motion of the side chain of Val68 toward the heme plane, as might be expected upon loss of ligand. Such a strong feature is, however, absent in the reference map (Figure 2b). An alternative explanation is that feature V represents a second docking site for CO in the distal pocket. It is, however, located close to the Val68 side chain. To accommodate a photodissociated CO molecule at this location, the side chain of Val68 would have to be

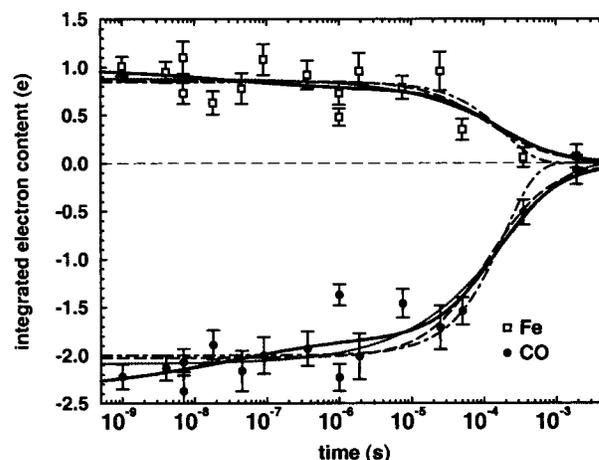


FIGURE 5: Time course of the integrated electron content of the CO and Fe-associated features from the weighted omit INS difference maps. A fit of both features by an exponential [$C_1 \exp(-k_1 t)$] is shown as dotted and dashed line, with a k_1 of $5.1 \times 10^3 \text{ s}^{-1}$. A stretched-exponential fit [$C_2 \exp(-k_2 t)^\beta$] with a k_2 of $4.6 \times 10^3 \text{ s}^{-1}$ and a β of 0.53 is shown as dotted line. A fit by a function $C_3/(1 + k_3 t)$ that represents bimolecular rebinding under conditions of comparable ligand and protein concentrations is shown as a dashed line, with a k_3 of $9.0 \times 10^3 \text{ s}^{-1}$. The fit that best describes the data is the sum of stretched-exponential ($k_2 = 2.9 \times 10^7 \text{ s}^{-1}$ and $\beta = 0.32$; 22%) and bimolecular ($k_3 = 6.7 \times 10^3 \text{ s}^{-1}$; 78%) functions and is shown as a solid line. The error bars in this figure and Figures 6, 7, and 8a were obtained as described in Materials and Methods.

displaced, but the maps do not reveal a strong signal indicating this displacement.

Evolution of Structural Changes from 1 ns to 1 ms. Unlike our previously reported data (28), the data presented here consist of a relatively complete series of difference electron density maps spanning time delays from 1 ns to 1.9 ms. We can therefore examine the time evolution of structural changes in more detail. When exploring ligand rebinding or the lifetime of the distal pocket CO* docking site, we extracted the time courses from weighted omit INS maps. For other structural changes that are more remote from the heme, we use weighted omit maps (see Materials and Methods).

Ligand Rebinding. Ligand rebinding reduces the magnitude of all features in the Mb*(t) minus MbCO difference electron density maps and has to be taken into account when the time evolution of any feature is examined. Ligand rebinding kinetics can be directly monitored through the decay of the negative feature labeled CO, associated with the loss of the bound CO molecule. As already discussed above, this feature partly arises from the Fe motion, which also generates the positive feature labeled Fe. Since it is difficult to separate the CO-labeled feature into CO and Fe contributions, we examine the time course of the integrated electron content of the entire feature and compare it with the time course of the Fe-labeled feature in Figure 5. Within the noise of the data, both features follow the same time course. To assess the time course of CO rebinding, we fit the two features jointly by three separate functions that could represent ligand rebinding kinetics in the crystal: a single exponential [$C_1 \exp(-k_1 t)$], a stretched exponential [$C_2 \exp(-k_2 t)^\beta$], and a function of the type $C_3/(1 + k_3 t)$, where k_1 – k_3 are rate coefficients and C_1 – C_3 and β are constants. A single-exponential or a stretched-exponential function would be expected to describe geminate rebinding, in which the CO

molecule does not exit into the solvent but rebinds to the same heme it left. A stretched-exponential nature of this process could result from protein fluctuations or protein relaxations that occur on the same time scale as ligand rebinding, or from multiple ligand docking sites or multiple conformations of the globin (see the Discussion). The third function represents bimolecular ligand rebinding from the solvent under conditions of comparable ligand and protein concentrations, as is the case in the crystals. The fit parameters obtained for these three cases are listed in the legend of Figure 5. The slight slope in the data at early times (<100 ns), especially evident for the stronger CO feature, suggests that a stretched function such as a stretched exponential describes the early time course of the CO rebinding in the crystal better than a single exponential. To consistently describe the overall time course of the CO feature and of two other electron density features, CO* and Xe 1 (see below), we further suggest that ligand rebinding is a two-phase process: a stretched-exponential (or multi-exponential), geminate phase followed by a bimolecular rebinding phase, prominent at longer times. The fit obtained from a combination of the stretched-exponential and bimolecular phase is shown with solid lines in Figure 5. The fit results in an amplitude of 22% for the stretched-geminate phase, with a rate coefficient k_2 of $2.9 \times 10^7 \text{ s}^{-1}$ and a β factor of 0.32. The bimolecular phase has an amplitude of 78% and a rate coefficient k_3 of $6.7 \times 10^3 \text{ s}^{-1}$. Although we choose this model since it is both chemically plausible and consistent with our data, we stress that the scatter in the data and an insufficient number of time points do not allow us at present to conclusively determine the quantitative nature of the ligand rebinding kinetics.

Given an accurate estimate (see Materials and Methods) of the random errors for the integrated electron content, represented by the error bars in Figure 5, any scatter in the data exceeding the random errors indicates the magnitude of remaining systematic error. The most likely source of this error is variation of the initial photolysis fraction from crystal to crystal and run to run, due to different crystal sizes, laser pulse energy, laser beam size, and alignment at the sample.

CO Docking Site Lifetime. The data presented here are sufficiently accurate to allow us to follow the time course of the positive electron density at the primary CO docking site, labeled CO* (Figure 2a). We compare in Figure 6 the integrated electron content of the CO* and CO features. The CO* feature diminishes with a time course similar to that of the earlier, stretched phase of the CO decay. A fit of the CO* time course by a sum of stretched-exponential and bimolecular phases, with both stretched and bimolecular parameters fixed to the values obtained from the fit of the CO time course, results in a time course in which most of the amplitude is associated with the stretched phase (Figure 6). The half-life of CO in this docking site is ~ 70 ns, in relatively good agreement with the value of 200 ns from time-resolved IR measurements in solution (19, 31).

Time Evolution of Other Structural Changes. In addition to those features directly related to heme and the CO molecule, we examined the time course of features associated with motion of the F helix, and with occupancy by CO of the four cavities associated with Xe binding (4).

To determine if the F helix features are already fully developed in the 1 ns difference map, we integrated the

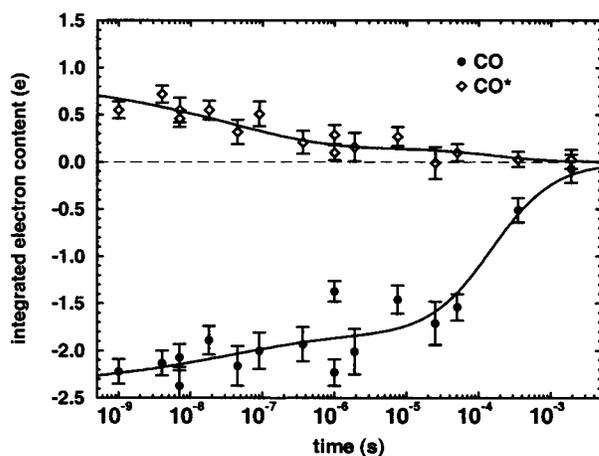


FIGURE 6: Time course of the integrated electron content of the distal pocket docking site denoted CO*, in weighted omit INS difference maps. The time course of the bound CO feature is shown for comparison.

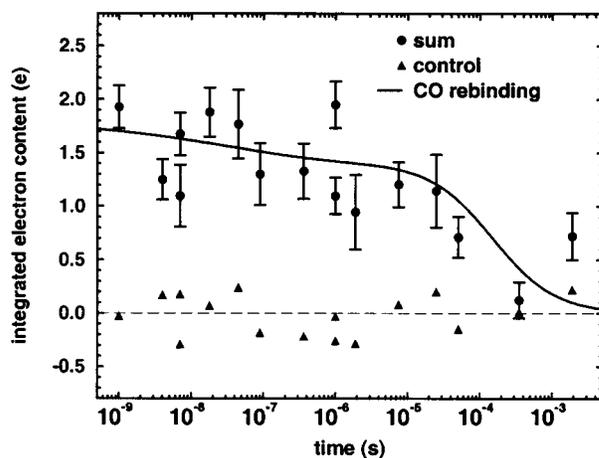


FIGURE 7: Time course of the integrated electron content of the F helix features. Five positive and five negative features, selected from the reference deoxyMb minus MbCO map, were integrated separately for all weighted omit maps. The sum of the absolute values is shown (circles), as well as the time course of the ligand rebinding (from Figure 5) for comparison (solid line). As a control, five regions with the same volume at the periphery of the myoglobin molecule were also integrated and the result is shown as triangles.

electron density in those locations of the difference maps where the most prominent features are located in the static, reference map (Figure 2b). We integrated the density within 10 spheres, each with a radius of 1 Å, five associated with positive and five with negative features. To further improve the signal-to-noise ratio, we also summed absolute values of positive and negative features. As a control, we integrated five corresponding regions at the periphery of the molecule where signal is expected to be absent. The results for weighted omit difference maps are shown in Figure 7. The summed electron content of the F helix features clearly lies above the noise and decreases monotonically with time. The substantial noise does not allow us to compare in detail the decay of the F helix features with the ligand rebinding kinetics (shown as a solid line in Figure 7). We can, however, compare the value for the summed electron content of these F helix features in the 1 ns map with the value for identical features in the reference map. The reference map value has to be scaled down to 40% of its value to account for partial

initial photolysis at the 1 ns time delay. The comparison then indicates that the amplitude of the F helix features in the 1 ns map is $65 \pm 20\%$ of the amplitude of the same features in the reference map. The possibility that some helix relaxation continues during ligand rebinding therefore cannot be ruled out.

To explore whether the CO molecule occupies any of Xe binding sites after photodissociation, we examined the weighted omit difference maps. Figure 8a shows the time dependence of the integrated difference electron content within a sphere with a 1.5 Å radius at each of the four Xe sites (4). The Xe 2, Xe 3, and Xe 4 sites do not show any positive density above the noise at any time point. However, the Xe 1 site clearly shows a positive signal whose magnitude peaks at ~ 100 ns and decays to half of its peak value in several microseconds. This site is located below the heme in the proximal pocket, flanked by residues Leu89, Ala90, His93, Leu104, Phe138, and Ile142.

The side chain of one of the neighboring residues of the Xe 1 site, Leu89, exhibits a structural rearrangement simultaneous with the appearance of the positive Xe 1 difference density. This rearrangement is indicated by both positive and negative features in the difference map as shown in Figure 8b for the 362 ns time delay. The time dependence of the magnitude of the Leu89 features exactly follows that of the positive Xe 1 feature as shown in Figure 8a. The features are consistent with the static, deoxyMb and MbCO structures (25). In MbCO, there is only one Leu89 location, but in deoxyMb, there are two, one MbCO-like (70%) and another ~ 1.9 Å away (30%). Rearrangement of the side chain of Leu89 is evidently necessary to accommodate either CO or Xe in the Xe 1 site; the site is only ~ 3 Å from the Leu89 CG atom in the MbCO location. The features represent a perfectly correlated pair of structural changes; occupancy of the Xe 1 site by CO only occurs upon displacement of Leu89.

Molecular Movie. To illustrate the time course of global structural changes, we constructed a molecular movie by using the rates we identified from the time course of the electron density features to fit the time course of the structure factor difference amplitudes in reciprocal space (see Figure 6). As described in Materials and Methods, the movie constructed this way maintains the important signal features while allowing data smoothing in the time domain. The reciprocal space fit function is a sum of a stretched-exponential phase with a rate coefficient of $2.9 \times 10^7 \text{ s}^{-1}$ and a β factor of 0.32 that account simultaneously for the fast phase of CO rebinding, decay of the CO* feature (Figure 6), and appearance of the Xe 1 feature (Figure 8a); an exponential phase with a rate coefficient of $2 \times 10^6 \text{ s}^{-1}$ that represents the fast decay of the Xe 1 feature (Figure 8a); and a bimolecular ligand rebinding phase with a rate coefficient k_3 of $6.7 \times 10^3 \text{ s}^{-1}$.

Ⓜ A molecular movie in GIF format is available.

DISCUSSION

Heme Relaxation. Results from a variety of spectroscopic experiments in solution that directly probe the Fe and heme environment suggest that although a large, initial Fe displacement and heme relaxation in Mb occur within 300 fs (3), these processes appear to be stretched in time and continue into the nanosecond and even the microsecond time

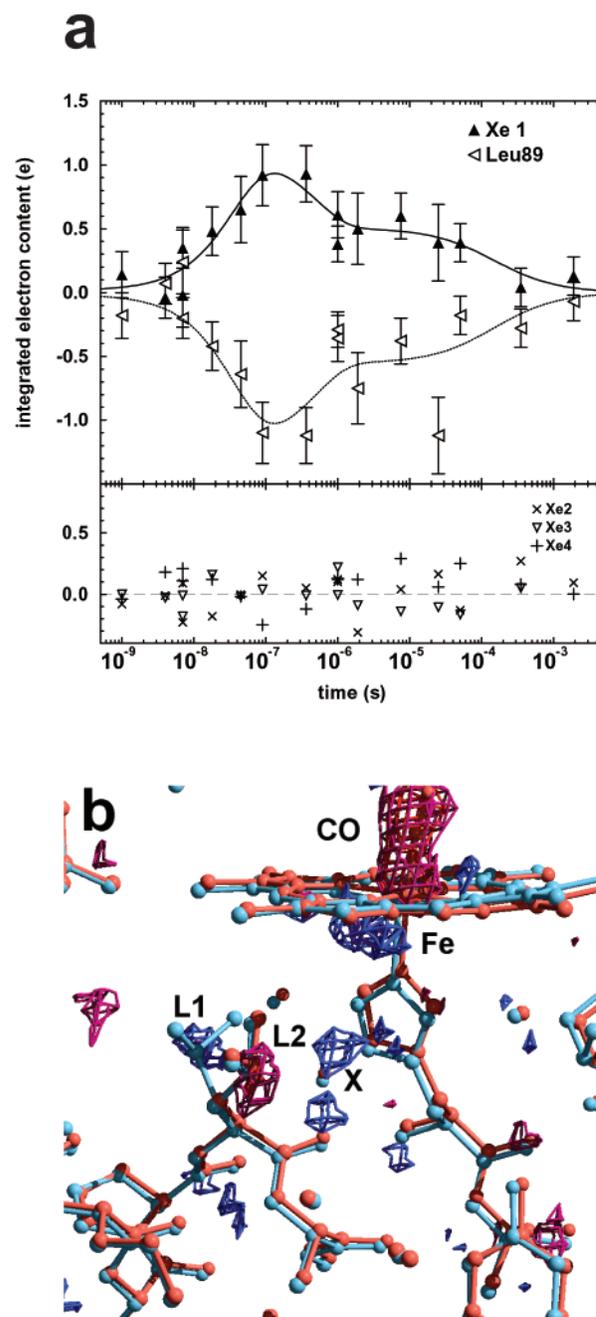


FIGURE 8: (a) Time course of the integrated electron content of the positive difference density at the Xe 1 binding site from weighted omit maps. The time courses of the other three Xe binding sites (Xe 2, Xe 3, and Xe 4) are shown for comparison. The solid line represents a fit of the time course of the Xe 1 density by two exponential phases (with rate coefficients of 2.6×10^7 and $2.0 \times 10^6 \text{ s}^{-1}$) and a bimolecular phase, fixed to that of ligand rebinding. The time course of the integrated electron content of the negative Leu89 feature is also shown. (b) Weighted difference electron density map of the Xe 1 region at 362 ns. The map is contoured at $\pm 3\sigma$. Positive density at the Xe 1 site is labeled X, while positive and negative densities indicating rearrangement of the Leu89 side chain are labeled L1 and L2, respectively. MbCO and deoxy Mb models from 1BZR and 1BZP entries are also shown.

domain (16, 17). The largest heme-associated feature that we can reliably follow in time is the positive Fe-associated density, which has already reached its largest magnitude with the 1 ns time delay, and its subsequent decay simply follows the ligand rebinding kinetics. This means that, in agreement

with the spectroscopic results, the major component of the heme relaxation does occur within 1 ns. We are not able to judge if a further Fe and/or heme displacement (albeit of small magnitude) continues and persists into the nanosecond to microsecond time domain.

Global Relaxation. Perhaps more surprising is the finding that most of the global motion of the F helix is also evident in our 1 ns map. Several spectroscopic experiments that probe more global structural changes do however suggest that at least part of the global structural changes could be very fast. A transient grating experiment (11, 41) revealed a fast (<10 ps) change in density of Mb* that corresponds to a global, photoinduced structural change. The authors suggest that collective vibrational modes are excited by rapid doming of the heme and that tertiary structure changes in the globin propagate via these modes. Time-resolved circular dichroism (CD) studies of Mb* indicate a global relaxation within 300 ps (42). It has been suggested (16) that the rapid relaxation observed in CD measurements might be correlated with a fast distal relaxation and that slower, nanosecond changes reflected by the time-dependent shift of charge-transfer band III (13, 16) is due to proximal relaxation.

Kachalova et al. (25) proposed a concerted motion of the heme and of E and F helices, based on their near-atomic resolution deoxyMb and MbCO structures. They identify the most striking structural difference between the two structures as the rigid body rotation and translation of the "V" formed by the E and F helices. The heme, covalently bound to the F helix via His93, is part of this rigid body and therefore included in the motion. The joint translation and rotation of these helices and the heme with respect to the other parts of the protein is necessary for removal of steric hindrance to CO access to the heme by the side chain of Val68. A CO molecule inserted into the deoxyMb heme pocket with its position and orientation relative to the heme nitrogens as in MbCO has its O only 2.7 Å from CG2 of Val68 (25). However, in other recent near-atomic resolution studies of deoxyMb and MbCO (26), the correlated rotational motion of the E and F helices and the heme was not observed. The reasons for this substantial disagreement in relatively precise, static structures are not clear.

Our data suggest that the rigid body helix motion in response to changes at the heme upon photolysis can be very fast. A significant fraction of the F helix displacement, ~65%, is already present at 1 ns, and the motion is in the direction consistent with the static, high-resolution structures. The time course of the amplitude of the F helix features shows only monotonic decay from the initial value at 1 ns (Figure 7). It is possible that a small subsequent increase in the F helix features is present, although any increase tends to be offset by the overall decrease due to rebinding.

Ligand Rebinding. Spectroscopic studies show that the geminate CO rebinding to Mb in solution is small in amplitude and nonexponential at room temperature (14, 17). The nonexponential nature has been attributed by Tian et al. (14) to large-scale protein fluctuations, represented for example by open and closed states of the distal pocket, that take a long time to average and indeed are not averaged on the shorter time scales of ligand rebinding and escape. When the geminate ligand rebinding fraction is small, the observed geminate rate represents the rate of ligand escape from the pocket (2, 14). The nonexponential geminate rate could

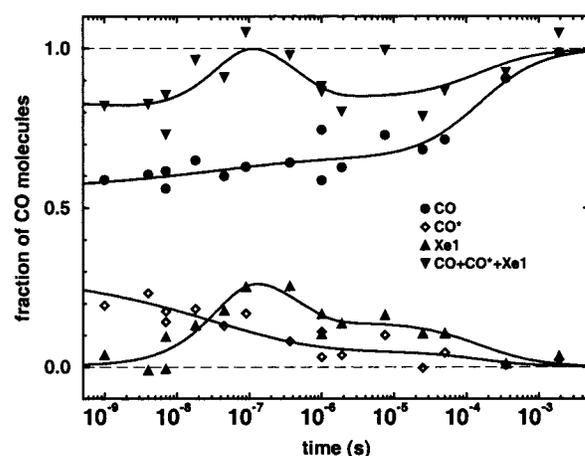


FIGURE 9: Fraction of CO molecules bound to heme (CO), at the distal site (CO*), and the proximal site (Xe 1), and the sum of these three fractions as a function of time.

therefore reflect large-scale protein fluctuations that occur on the same time scale as ligand escape.

In another comprehensive nanosecond study of ligand rebinding and protein conformational relaxation in Mb in solution as a function of temperature and solvent viscosity, Ansari et al. (17) concluded that both processes are nonexponential. Their minimal model includes two serial conformational states of the photoproduct with geminate rebinding rates that differ by a factor of 100. They propose that differences between these two states are on the proximal side of the heme, and that geminate ligand rebinding occurs on the same time scale as heme/protein relaxation from one conformation to the other. Protein relaxation seems to continue over a wide range in time, up to 10 μs as evident from changes in near-IR charge-transfer absorption band III (16).

While ligand rebinding in solution has been thoroughly investigated by numerous spectroscopic techniques, ligand rebinding in Mb crystals has not yet been measured in the same detail. Resonance Raman studies (43, 44) indicate that MbCO crystals are much more difficult to photolyze than solution. By monitoring the oxidation state marker Raman band, the authors conclude that complete photolysis of crystals could not be accomplished before irreversible crystal damage by the focused laser light occurs. Our earlier measurements of the magnitude of the optical density change at 543 nm as a function of nanosecond laser pulse energy (28) and our crystallographic data suggest that only 40–50% of molecules are photolyzed at 4 ns in crystals under our experimental conditions. Higher pulse energies damage the crystals. The number of absorbed photons per molecule per laser pulse in our experiments is estimated to be ~0.7 (see Materials and Methods). Since photolysis of the crystals is relatively uniform (OD ~ 0.2 at 630 nm), the expected fraction of photodissociated CO molecules is ~70%, larger than the observed fraction of ~40%. Both these and the resonance Raman results suggest that fast geminate rebinding from the initial contact-pair state (8) could be present in the crystal with an amplitude significantly larger than the 3% amplitude in solution.

The following picture regarding ligand rebinding and migration, and protein relaxation in the crystal (Figures 9 and 10), is consistent with our data. Ligand rebinding consists

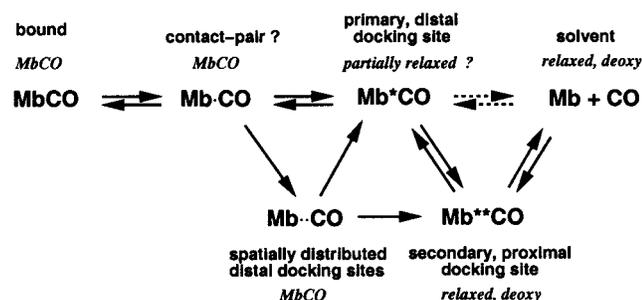


FIGURE 10: Summary of ligand and protein states following MbCO photolysis. Ligand states are labeled in bold and protein states in italic.

of two phases: a geminate phase best described by a stretched-exponential process and a bimolecular phase. The stretched nature of the initial ligand rebinding phase could be explained by fluctuations (14), protein relaxation (17), or rebinding from multiple, distinct ligand docking sites, of which the site labeled CO* is the most prominent. Our data indicate that the major part of the globin relaxation is already completed by 1 ns. However, a smaller, slower protein relaxation could still be present that cannot be directly visualized with the data presented here. This continued relaxation would affect ligand rebinding kinetics. The “unsynchronized” protein fluctuations on the time scale of ligand rebinding cannot be directly observed by time-resolved experiments, since these experiments record the structure averaged over all molecules in the crystal, at a particular laser/X-ray time delay over the duration of the X-ray pulse. The fluctuations would, if present, manifest themselves by the nonexponential nature of geminate ligand rebinding.

On the time scale of the first, stretched-exponential phase of ligand rebinding, the CO molecule begins to escape from the distal pocket docking site labeled CO* and to populate the Xe 1 proximal docking site (Figure 8a). Does the CO molecule simply pass from one site to the other, or there are other sites with occupancies too small for us to detect? Figure 9 shows the fractional occupancy of CO molecules that are located at the distal site, located at the proximal Xe 1 site, or bound to the heme, as a function of time. The sum of these three fractional occupancies is also shown. The deviation of the sum from 1.0 represents the undetected fraction of CO molecules. The conversion from the integrated electron content (Figures 5, 6, and 8a) to fractional occupancy (Figure 10) is based on two fractional occupancies determined from the 1 ns map (see the Results); at 1 ns, 40% of all MbCO molecules are photolyzed and 45% of photodissociated CO molecules are in the CO* distal pocket site. All other occupancies are established relative to these values. The fact that the sum of three fractional occupancies lies close to 1.0 in the time range around 100 ns is not a consequence of the conversion, but the result of the independently determined time evolution of three difference electron density features.

We can thus account quantitatively for the location of all CO molecules in the time range around 100 ns: 11% occupy the CO* distal docking site, 26% occupy the proximal Xe 1 site, 63% have rebound geminately (or were never photodissociated), and 0% have escaped to the solvent.

At the 1 ns time delay, we find electron density at the CO* distal docking site that can account for only ~45% of

the photodissociated CO molecules. The missing 55% of the distal site electron density (corresponding to ~20% of all CO molecules) could reflect positional disorder of the CO molecule. At low temperatures (20–40 K), a trajectory of the photodissociated CO molecule has been observed (45). The CO* site at room temperature is located at the more-distant end of the low-temperature trajectory (45). It is possible that ~20% of the CO molecules are spatially distributed along the trajectory and are simply undetectable in a crystallographic experiment. An angular and positional probability distribution of the photodissociated CO molecule has been proposed from molecular dynamics simulations (22). Despite uncertainty about their exact spatial location, it is very likely that the missing 20% of CO molecules are still located within the distal pocket at 1 ns.

In studies of the effect of Xe on oxygen kinetics, Scott and Gibson (21) identified the Xe 1 site as a possible, proximal site situated on the oxygen exit pathway. The Xe 1 site is the highest-affinity Xe binding site that is nearly fully occupied when metMb crystals are equilibrated with 7 atm of Xe gas [94% occupancy (4)]. The other sites are only ~50% occupied under those conditions. The photodissociated CO molecule was directly observed at this site in crystallographic studies of the L29W mutant of sperm whale myoglobin, and of native horse heart myoglobin after extended illumination by light at temperatures of ≥ 160 –180 K (46, 47). This temperature range marks the dynamic transition region where protein fluctuations are believed to open exit channels for ligand escape from the protein into the solvent (1, 46).

Scott and Gibson (21) considered two possibilities for the oxygen exit pathway: a sequential model, in which oxygen escapes from the distal to the proximal site and from the proximal site to the solvent; and a side-path model, in which oxygen escapes from the distal site either directly to the solvent or to the dead-end proximal site. They preferred the second, side-path model. Our structural results suggest that the decay of the CO* site has predominantly one fast phase (Figure 6) and that when the occupancy of the Xe 1 site is at a maximum, all of the CO molecules are accounted for within the protein and therefore none have escaped into the solvent. In the side-path model, fast escape from the CO* site would be directed partially toward the solvent, contrary to our observations. The sequential model, with a minor pathway for decay from the CO* site directly into the solvent, is therefore more consistent with our data. The peak occupancy of ~20% of the Xe 1 docking site is in an excellent agreement with the result of Scott and Gibson (21) for the maximum occupancy of the secondary species of 20%. The Xe 1 site density decays to half of its peak value in $\sim 10 \mu\text{s}$, consistent with the time constant of $3 \mu\text{s}$ for CO escape into the solvent obtained from time-resolved infrared spectroscopy (31).

What is the trajectory of the CO molecule between the distal and proximal docking sites? In molecular dynamic simulations (21, 48, 49), the outgoing ligand hops between the Xe-binding cavities before escaping to the solvent. Protein flexibility is certainly necessary for a ligand to reach the Xe 1 site since a photodissociated ligand is observed at this site only above the dynamic transition temperature of 180 K (46). Structural changes between ligated and unligated structures involve a concerted motion of the heme and the F

and E helices, and reorientation of the hydrophobic residues of the heme pocket, especially the residues that form a clamp around the pyrrole C ring (25). These motions are probably responsible for opening a pathway to the proximal docking site.

The final phase of CO rebinding involves bimolecular rebinding from the solvent. This phase of rebinding will appear in crystals as somewhat extended compared to a single exponential. It is described by a function $C_3/(1 + k_3t)$ under the conditions in the crystal where the free ligand concentration is equal to the unligated protein concentration. The rate k_3 represents a product $k'[\text{CO}]_{\text{bm}}$, where k' is a second-order rate coefficient and $[\text{CO}]_{\text{bm}}$ is the concentration of CO molecules, initially bound to heme, that escaped into the solvent after photolysis. Combining the value for k_3 of $6.7 \times 10^3 \text{ s}^{-1}$ from the fit and $[\text{CO}]_{\text{bm}}$ of 15 mM [for 35% of all CO molecules that rebind from the solvent (Figure 9), and assuming the total CO concentration of 45 mM] yields a k' of $4.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, consistent with the viscosity-independent bimolecular rate of $1.7 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ measured in solution (17).

CONCLUSIONS

We explored the protein response to ligand photodissociation in myoglobin with a time-resolved diffraction technique. Our results and those on photoactive yellow protein (53) illustrate the present, quantitative capability of the time-resolved technique as we were able to detect quite subtle but essential structural changes that characterize the relaxation of this molecule. The error levels are sufficiently low to permit detection of structural displacements as small as 0.2 Å such as those of the F helix, and to resolve 10% occupancy of a CO docking site, over the time range from 1 ns to a few milliseconds.

In addition to rapid structural changes around the heme, considered to be the epicenter of the "protein quake" (50), we detect early, nanosecond global structural changes involving motion of the F helix. We observe two docking sites of the photodissociated CO molecule. The primary site is in the distal pocket, close to the location of the water molecule in deoxyMb and to the site where CO is found upon photolysis at low temperatures (34, 45). The occupancy of this site is already at its maximum of ~20% in 1 ns, and the half-life for its decay is 70 ns. The secondary site is on the proximal side of the heme at the Xe 1 binding site, which attains a peak occupancy of 20% at ~100 ns before decaying with a half-life of several microseconds. The actual exit pathway of the CO molecule is not known, but it certainly involves at least one of these two docking sites since at 100 ns, all CO molecules are either bound to the heme or in one of these two sites within the protein. The missing fraction of photodissociated CO molecules at early times (<100 ns) suggests the existence of other, spatially distributed docking sites in the distal pocket that we have not been able to detect due to their spatial disorder and low occupancy ($\leq 10\%$).

ACKNOWLEDGMENT

We thank Shin-ichi Adachi for his help in data collection and processing. K.M. dedicates this paper to his mentors in crystallography and rapid reaction kinetics, Max Perutz and Quentin Gibson.

REFERENCES

- Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H., and Gunsalus, I. C. (1975) *Biochemistry* 14, 5355–5373.
- Henry, E. R., Sommer, J. H., Hofrichter, J., and Eaton, W. A. (1983) *J. Mol. Biol.* 166, 443–451.
- Martin, J. L., Migus, A., Poyart, C., Lecarpentier, Y., Astier, R., and Antonetti, A. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 173–177.
- Tilton, R. F., Jr., Kuntz, I. D., Jr., and Petsko, G. A. (1984) *Biochemistry* 23, 2849–2857.
- Findsen, E. W., Scott, T. W., Chance, M. R., Friedman, J. M., and Ondrias, M. R. (1985) *J. Am. Chem. Soc.* 107, 3355–3357.
- Gibson, Q. H., Olson, J. S., McKinnie, R. E., and Rohlfs, R. J. (1986) *J. Biol. Chem.* 261, 10228–10239.
- Janes, S. M., Dalickas, G. A., Eaton, W. A., and Hochstrasser, R. M. (1988) *Biophys. J.* 54, 545–549.
- Jongeward, K. A., Magde, D., Taube, D. J., Masters, J. C., Traylor, T. G., and Sharma, V. S. (1988) *J. Am. Chem. Soc.* 110, 380–387.
- Moore, J. N., Hansen, P. A., and Hochstrasser, R. M. (1989) *J. Am. Chem. Soc.* 111, 4563–4566.
- Carver, T. E., Rohlfs, R. J., Olson, J. S., Gibson, Q. H., Blackmore, R. S., Springer, B. A., and Sliagar, S. G. (1990) *J. Biol. Chem.* 265, 20007–20020.
- Genberg, L., Richard, L., McLendon, G., and Miller, R. J. D. (1991) *Science* 251, 1051–1054.
- Gibson, Q. H., Regan, R., Elber, R., Olson, J. S., and Carver, T. E. (1992) *J. Biol. Chem.* 267, 22022–22034.
- Nienhaus, G. U., Mourant, J. R., and Frauenfelder, H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 2902–2906.
- Tian, W. D., Sage, J. T., Šrajcar, V., and Champion, P. M. (1992) *Phys. Rev. Lett.* 68, 408–411.
- Kuczera, K., Lambry, J.-C., Martin, J.-L., and Karplus, M. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 5805–5807.
- Lim, M., Jackson, T. A., and Anfinrud, P. A. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 5801–5804.
- Ansari, A., Jones, C., Henry, E. R., Hofrichter, J., and Eaton, W. A. (1994) *Biochemistry* 33, 5128–5145.
- Franzen, S., Bohn, B., Poyart, C., and Martin, J. L. (1995) *Biochemistry* 34, 1224–1237.
- Lim, M., Jackson, T. A., and Anfinrud, P. A. (1995) *Science* 269, 962–966.
- Olson, J. S., and Phillips, G. N., Jr. (1996) *J. Biol. Chem.* 271, 17593–17596.
- Scott, E. E., and Gibson, Q. H. (1997) *Biochemistry* 36, 11909–11917.
- Vitkup, D., Petsko, G. A., and Karplus, M. (1997) *Nat. Struct. Biol.* 4, 202–208.
- Brunori, M., Vallone, B., Cutruzzola, F., Travaglini-Allocatelli, C., Berendzen, J., Chu, K., Sweet, R., and Schlichting, I. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97, 2058–2063.
- Schlichting, I., and Chu, K. (2000) *Curr. Opin. Struct. Biol.* 10, 744–752.
- Kachalova, G. S., Popov, A. N., and Bartunik, H. D. (1999) *Science* 284, 473–476.
- Vojtechovský, J., Chu, K., Berendzen, J., Sweet, R. M., and Schlichting, I. (1999) *Biophys. J.* 77, 2153–2174.
- Bourgeois, D., Ursby, T., Wulff, M., Pradervand, C., LeGrand, A., Schildkamp, W., Labouré, S., Šrajcar, V., Teng, T. Y., Roth, M., and Moffat, K. (1996) *J. Synchrotron Radiat.* 3, 65–74.
- Šrajcar, V., Teng, T.-Y., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M., and Moffat, K. (1996) *Science* 274, 1726–1729.
- Moffat, K. (1997) *Methods Enzymol.* 277, 433–447.
- Ren, Z., Bourgeois, D., Helliwell, J. R., Moffat, K., Šrajcar, V., and Stoddard, B. L. (1999) *J. Synchrotron Radiat.* 6, 891–917.
- Schotte, F., Techert, S., Anfinrud, P., Šrajcar, V., Moffat, K., and Wulff, M. (2000) in *Handbook on Synchrotron Radiation* (Mills, D., Ed.) Vol. 5, Wiley and Sons, New York.
- Moffat, K. (2001) *Chem. Rev.* 101, 1569–1581.
- Kendrew, J. C., and Parrish, R. G. (1956) *Proc. R. Soc. London, Ser. A* 238, 305–324.

34. Teng, T.-Y., Šrajer, V., and Moffat, K. (1994) *Nat. Struct. Biol.* 1, 701–705.
35. Chen, Y., Šrajer, V., Ng, K., LeGrand, A., and Moffat, K. (1994) *Rev. Sci. Instrum.* 65, 1506–1511.
36. Ren, Z., and Moffat, K. (1995) *J. Appl. Crystallogr.* 28, 461–481.
37. Ursby, T., and Bourgeois, D. (1997) *Acta Crystallogr. A* 53, 564–575.
38. Tulinsky, A. (1985) *Methods Enzymol.* 115, 77–89.
39. Henderson, R., and Moffat, K. (1971) *Acta Crystallogr. B* 27, 1414–1420.
40. Moffat, K. (1989) *Annu. Rev. Biophys. Biophys. Chem.* 18, 309–332.
41. Miller, R. J. D., Deak, J., Palese, S., Pereira, M., Richard, L., and Shilling, L. (1993) in *Ultrafast Phenomena VIII* (Martin, J.-L., Mingus, A., Mourou, G. A., and Zewail, A. H.) pp 525–528, Springer, Berlin.
42. Xie, X., and Simon, J. D. (1991) *Biochemistry* 30, 3682–3692.
43. Zhu, L., Sage, T. J., and Champion, P. M. (1993) *Biochemistry* 32, 11181–11185.
44. Morikis, D., Sage, J. T., Rizos, A. K., and Champion, P. M. (1988) *J. Am. Chem. Soc.* 110, 6341–6342.
45. Teng, T.-Y., Šrajer, V., and Moffat, K. (1997) *Biochemistry* 36, 12087–12100.
46. Ostermann, A., Waschipky, R., Parak, F. G., and Nienhaus, G. U. (2000) *Nature* 404, 205–208.
47. Chu, K., Vojtechovský, J., McMahon, B. H., Sweet, R. M., Berendzen, J., and Schlichting, I. (2000) *Nature* 403, 921–923.
48. Tilton, R. F., Jr., Singh, U. C., Kuntz, I. D., Jr., and Kollman, P. A. (1988) *J. Mol. Biol.* 199, 195–211.
49. Elber, R., and Karplus, M. (1987) *Science* 235, 318–321.
50. Ansari, A., Berendzen, J., Bowne, S. F., Frauenfelder, H., Iben, I. E. T., Sauke, T. B., Shyamsunder, E., and Young, R. D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 5000–5004.
51. McRee, D. E. (1993) in *Practical Protein Crystallography*, Academic Press, San Diego.
52. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* 277, 505–525.
53. Ren, Z., Perman, B., Šrajer, V., Teng, T.-Y., Pradervand, C., Bourgeois, D., Schotte, F., Ursby, T., Kort, R., Wulff, M., and Moffat, K. (2001) *Biochemistry* 40, xxxx–xxxx.

BI010715U