Published in final edited form as:

J Phys Chem Lett. 2019 November 21; 10(22): 7148–7163. doi:10.1021/acs.jpclett.9b02375.

Ultrafast Structural Changes Decomposed from Serial Crystallographic Data

Zhong Ren*

Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60607, United States;
Renz Research, Inc., Westmont, Illinois 60559, United States

Abstract

Direct visualization of electronic and molecular events during biochemical reactions is essential to mechanistic insights. This Letter presents an in-depth analysis of the serial crystallographic data sets collected by Barends and Schlichting et al. (*Science* 2015, 350, 445) that probe the ligand photodissociation in carbonmonoxy myoglobin. This analysis reveals electron density changes caused by the formation of high-spin 3d atomic orbitals of the heme iron upon photolysis and their dynamic behaviors within the first few picoseconds. The heme iron is found popping out of and recoiling back into the heme plane in succession. These findings provide long-awaited visual validations for previous works using ultrafast spectroscopy and molecular dynamics simulations. Electron density variations are also found largely in the solvent during the first period of a low-frequency oscillation. This work demonstrates the importance of the analytical methods in detecting and isolating weak, transient signals of electronic changes arising from chemical reactions.

Graphical Abstract

*Corresponding Author: zren@uic.edu.

Supporting Information

The Supporting Information is available free of charge on the [ACS Publications website](https://pubs.acs.org/doi/10.1021/acs.jpclett.9b02375) at DOI: 10.1021/acs.jpclett.9b02375.

Methods and supplementary references, tables, and figures (PDF)

Movie S1: Reconstituted difference maps showing modulation and oscillation (MP4)

Movie S2: Temperature effect on electron density map of the heme (MP4)

Movie S3: Comparison of electron density of the heme in ligated and unligated deoxy states (MP4)

The author declares the following competing financial interest(s): Z.R. is the founder of Renz Research, Inc. that currently holds the copyright of the computer software dynamiX.

One of the major applications of X-ray free electron lasers (XFELs) is to directly observe ultrafast protein dynamics at atomic resolution during photochemical reactions by serial femtosecond (fs) crystallography (SFX). Time-resolved experiments have been successfully conducted by several groups using XFELs to follow the photodissociation of carbonmonoxy myoglobin (MbCO; Figure S1), photoisomerization of the *p*-coumaric acid chromophore in photoactive yellow protein (PYP), and the hydroxybenzylidene imidazolinone chromophore in a photoswitchable fluorescent protein.¹ While the time resolution, that is, the ability to differentiate transient changes, is limited to tens of femtoseconds determined by the duration of the XFEL pulses, the reported time resolution for MbCO and PYP was around hundreds of femtoseconds due to some timing jitter between a short laser pump and an XFEL probe.

It is well-established that a ligand bound to the heme iron helps to split its five 3d orbitals into two energy levels, the triply degenerate t_{2g} at a lower energy and the doubly degenerate e_g at a higher energy, so that the six 3d electrons of Fe(II) form three pairs occupying the lower-energy orbitals and leaving the doubly degenerate empty; thus, a smaller ligated Fe(II) with zero net spin results. Upon photodissociation, the CO ligand parks at a nearby docking site. The absence of a bound ligand returns the five degenerate 3d orbitals of the iron back to approximately identical energy levels so that they are all occupied resulting in a larger, high-spin, deoxy iron, which causes the iron to pop out of the heme plane toward the proximal side. As the porphyrin structure becomes buckled, several side chains and helices of the globin move accordingly (Figure 1). These light-induced structural changes are highly reproducible and have been observed by both static and time-resolved crystallography,

repeatedly.^{2–8} Recently, they have been captured once again at a subpicosecond time scale in MbCO crystals using XFEL pulses,⁹ a major achievement in SFX. However, no significantly new structural finding is beyond those already obtained by time-resolved Laue diffraction at synchrotrons,^{10–12} which is 1000 times slower at a time resolution of 100 picosecond (ps) limited by the bunch length of the synchrotrons.^{3,13–16}

The signals of ultrafast structural events could be too weak to be observed clearly by the standard crystallographic analyses. In this work, I reexamine the time-resolved XFEL data sets of MbCO deposited in the Protein Data Bank (PDB; 5cmv, 5cn4-9, and 5cnb-g) by Barends and Schlichting et al.⁹ with a previously developed analytical method based on singular value decomposition (SVD).^{15,17–19} This SVD analysis unravels several ultrafast events that have been identified by spectroscopy and computational studies but have not been directly observed as images of electron density distribution. The new findings along with several previously known structural events upon photolysis of MbCO are summarized in a Gantt chart (Figure 1).

Resolution of heterogeneous signals is key to a clear understanding of a dynamic process. Compared to the well-established structural events of MbCO photolysis such as ligand docking and heme doming, the ultrafast signals are small in amplitude, low in population, and mixed in a multifactor interference. As a result, they do not stand out at a given time point and thus went unnoticed in the conventional analyses. As demonstrated below, these weak yet consistent signals can be manifested as major components resulting from SVD when multiple data sets collected at different time points are jointly examined. This necessity for decomposition presents both a challenge and a unique opportunity for dynamic crystallography, broadly defined as experimental and analytical approaches that tightly join crystallographic data sets and their metadata.¹⁹ We have developed a numerical process of deconvolution, which is able to identify and extract the dynamic signals from a large collection of data sets acquired at varying experimental conditions such as time points and temperatures.^{15,20–24} In essence, this strategy takes advantage of the structural heterogeneity, rather than preemptively avoiding them as in the common practice of static protein crystallography. In a nutshell, dynamic crystallography consists of two major steps: first, to experimentally introduce a functionally relevant structural heterogeneity; and second, to numerically resolve the heterogeneous conformational species as long as their compositions vary as a function of an experimental parameter.¹⁹ Here, time-resolved serial crystallography, with the variable being the ultrashort time delay, is simply a special case in the broader scope of dynamic crystallography. Therefore, the previously established methodology is largely applicable (Methods).

The ultrafast dynamics during the photodissociation of MbCO has been extensively studied with a variety of methods. The departure of the CO ligand from the iron porphyrin system has been reported within the first 15–20 fs upon photolysis.^{25–27} This signal is clear in the earliest time point of ~30 fs (nominally –0.1 ps; Figure S2). The photolyzed CO ligand is already observed in its docking site 3.5–4.5 Å from the heme iron at tens to hundreds of femtoseconds (Figures 2c and 3a,g,j), unlike the result recently suggested by quantum wavepacket dynamics that the ligand departs only 0.4 ± 0.1 Å.²⁷ A transient charge transfer from the iron to porphyrin subsequently taking place during 50–300 fs has been identified as

a key event by resonance Raman spectra.²⁸ A high-spin iron results in the charge transfer from the porphyrin back to the iron. The transient charge transfer state is not visible in the maps calculated from the data sets of Barends and Schlichting et al. perhaps because of a broad distribution of the transferred electron over the porphyrin. However, the larger, high-spin iron is clearly manifested in the earliest maps (Figures 2, 3, and S2). Vibrational signals have been observed by ultrafast spectroscopy, which is attributed to a discontinuity or “seam” at the intersection of two energy landscapes of the photoexcited MbCO and the ligand dissociated product.²⁹ Dissipation of such vibrational energy in the protein matrix and its solvent environment consists of at least two distinct components: localized, high-frequency vibration modes of the heme at approximately 1000 wavenumbers and global, low-frequency oscillations of the protein at several tens of wavenumbers.³⁰ This analysis provides direct visual evidence to validate both components and to delineate the low-frequency oscillations in greater detail.

The most observed high-frequency modes range from hundreds to more than a thousand wavenumbers.^{31–33} The corresponding vibrational periods of these high-frequency modes of 20–50 fs are comparable to the XFEL pulse duration. Hence, an individual cycle of such vibrations is beyond what can be traced out by the current XFEL data sets. Nevertheless, the averaged effect of these high-frequency vibrations is detectable (Figures 2 and 3) in the XFEL data sets of Barends and Schlichting et al. Amplitude modulation and damping of the high-frequency mode ν_7 was simulated with a period of 350 fs and a time constant of 1.3 ps.⁹ There are not sufficient time points in the existing data sets to determine these amplitude modulation and damping. However, a slower modulation of the high-frequency modes is clearly present in these data sets (Figures 2 and S4; Movie S1). These in-plane modes are thought to be coupled to out-of-plane motions of the heme and its iron,³⁴ which take place within a fraction of a picosecond and continue to develop nonexponentially for another hundreds of picoseconds according to molecular dynamics simulations.^{35,36} Recent simulations^{9,26} suggested that many transient species, including those with iron in or out of the plane, are mixed together before 1 ps. It was also suspected that the iron may recoil back into the heme plane after its initial drop out of the plane.³⁷ This analysis of the XFEL data sets determines that the out-of-plane displacement of the iron takes place simultaneously with an in-plane shift of the porphyrin (Movie S3). There is clear visual evidence to show that the first out-of-plane move peaks at a few picoseconds and that the recoil and the second drop take place afterward (Figures 2c,f and S4c; Movie S1). The signal associated with the iron displacement lingers on for many microseconds before the ligand rebinding as captured by the time-resolved Laue diffraction.⁸

At a low-frequency region, the dominant oscillations range between 25 and 45 cm^{-1} , corresponding to a period around 1 ps.^{37,38} Such a low-frequency oscillation is within the reach of the subpicosecond time resolution achieved with XFEL pulses. Therefore, each individual cycle of the low-frequency oscillation should be, and indeed is, visible in the data sets of Barends and Schlichting et al. Here I demonstrate that such an oscillation occurs more in the solvent channels than in the protein or the heme at a low frequency comparable to those detected by spectroscopy. This analysis supplies the long missing visual evidence to support such a low-frequency oscillation in a solvated protein.

Imaging the geometric shapes of atomic orbitals is highly desirable but remains debatable in theory and in practice. The shape of a d orbital of copper was first experimentally imaged by Zuo et al. in their difference electron density map,^{39,40} which was not, however, without controversy.^{41,42} Nevertheless, experimenters continued to draw parallels between captured images and atomic or molecular orbitals, sometimes with the analytical power of digital reconstructions. Molecular orbital tomography constructs images of highest occupied molecular orbitals from many projections of aligned molecules.^{43–47} Both the highest occupied and lowest unoccupied molecular orbitals of flat molecules, such as pentacene, were scanned using scanning tunneling microscopy.^{48–50} Field-emission electron microscopy and angle-resolved photoemission spectroscopy can directly capture or indirectly reconstruct orbital images or transformation of orbital configurations.^{51,52} More recently, a photoionization experiment showed the nodal structures of the atomic orbital of excited hydrogen atoms.⁵³ A charge-density analysis at an ultrahigh resolution of 0.48 Å showed iron 3d and sulfur 3p densities.⁵⁴ In all these achievements, the experimentally captured electron densities with their unique shapes are directly influenced by electrons occupying specific orbitals. Therefore, in this Letter, the images and motion pictures of difference electron densities with their characteristic non-spherical shape are attributed to a superposition of several atomic orbitals and their dynamics. However, this is not to say that the accurate image of an atomic orbital has been captured, which is not a proper task for protein crystallography, given the limit in spatial and temporal resolutions.

The chief methodology applied in this work is decompositions of difference Fourier maps and simulated annealing omit maps (SAOMs). Here, a brief introduction to the method is laid out before the results are presented below. See Methods in the Supporting Information for details. To analyze the time-resolved serial crystallographic data of horse heart (hh) MbCO, 1 dark and 12 light data sets are obtained from the PDB⁹ (Figure 1), among which three light data sets collected at the shortest time delays of 0 and ± 0.1 ps span both sides of time 0 that is centered on the excitation laser flash of 150 fs. These data sets are expected to capture ultrafast signals at delay times as short as a few femtoseconds to tens of femtoseconds. However, these ultrafast signals are also inevitably mixed with signals at longer delays as far as the X-ray pulses stretch.⁵⁵ These subpicosecond time points were taken at the same controlled time delay of 0.5 ps. However, the diffraction images were binned according to the timestamp records from an experimental timing tool,^{56,57} which produced a number of short delays before 1 ps. Several longer time points ranging from 3 to 150 ps are also included in this analysis (Figure 1) along with a deoxy hhMb data set collected from a synchrotron beamline at 7 °C, which represents the end product of the MbCO photolysis, and two hhMb structures (1dws/t) in the photolyzed state captured by cryo-trapping.⁴

Two types of electron density maps, namely, difference Fourier maps and SAOMs, are produced and subjected to the SVD analyses separately. Each type carries its own advantage (Methods). While difference Fourier maps are sensitive for detecting weak signals, they are also highly susceptible to systematic errors that often overwhelm the signal. It is critically important to choose a proper reference data set, which is preferably obtained under the identical experimental conditions, if not from the very same crystal as the light data set. An alternative is to use SAOMs, where a small portion of a protein structure is set aside from

structure refinement using the simulated annealing algorithm. In such omit maps, the electron densities within the omitted region shall not be biased by the protein model; thus, they are considered as the most authentic representation of the experimentally observed electron densities.⁵⁸ A collection of SAOMs within a common omitted region, although less sensitive to subtle changes compared to difference maps, permits a joint analysis of maps derived under very different experimental conditions, which is usually difficult with difference maps.

Difference Fourier maps are synthesized using a Fourier coefficient set of $F_{\text{light}} - F_{\text{reference}}$ combined with the best possible phase set. While $F_{\text{reference}}$ is typically the structure factor amplitudes from a dark data set (5cmv), it is important that a light data set at a given time point is also used as a reference (Table S1) to reveal changes with respect to that specific time point (Methods). Therefore, a collection of difference maps produced with the same reference data set constitutes a time series that shows continuous changes from the reference time point and on. For example, $-F_0$ series refers to the difference maps with the reference data set at the nominal 0 ps. For omit maps, the heme and several side chains within its immediate vicinity are omitted to produce SAOMs of the ligand binding site (Methods). A few static structures determined at room temperature of sperm whale (sw) Mb in the aquomet (1bzc), deoxy (1bzn), and CO-bound (1bzn) forms are also included,⁵ along with three hhMb structures⁴ (1dwr/s/t) as control data points in this joint analysis. The resolutions of these static data sets are truncated to 1.75 Å to match the XFEL data sets.

Altogether, a total of 93 difference maps (Table S1 and Figure S2) and 20 omit maps are obtained for two independent SVD analyses. The electron densities are masked around the heme site unless stated otherwise (Methods). In this work, the significant signals are identified by the top nine components from SVD (Figure 2e). A numerical procedure called rotation in the SVD space plays an important role in extracting chemically sensible components from these (difference) electron density maps.^{18,22,59,60} Proper SVD rotations reveal the correlation between the core data, i.e., electron densities, and their metadata, i.e., delay time, sample temperature, and other experimental parameters (Methods). Given their highly dynamic nature, the ultrafast signals due to multiple causes are often entangled before suitable rotations are found in the multidimensional SVD space (Methods). Using this strategy of decomposition (SVD) and deconvolution (rotation), I identify signals that reveal the electronic redistributions between high and low spin of the iron orbitals at the available spatial resolution of 1.8 Å.

The ultrafast signals within several femtoseconds to a few tens of femtoseconds are manifested in the fourth component, denoted as U_4 , of 93 difference maps and are attributed to the low-spin to high-spin crossover of the heme iron. The component U_4 exists only in the map series with the dark data set as reference (red, $-F_{\text{dark}}$ series in Figure 2). That is, all signals in this component have occurred by the time of the very first time point at ~30 fs (nominally -0.1 ps); therefore, the map series $-F_{-01}$ shows the cancellation of this component (orange series in Figure 2). Some of these signals start to decay slightly before 1 ps and diminish significantly after 3 ps (Figure 2c,d). The corresponding map of U_4 reveals the previously well-established events arising from the CO docking, out-of-plane motion of the iron, and an uneven heme doming upon photolysis of MbCO (Figure 3). The negative

densities associated with the pyrrole rings and their propionic acid side chains increase in intensity from ring A to B to C to D (Figure 3c). Positive densities are found at a cross section 1.1 Å above the heme plane toward the distal side (Figure 3b).

More remarkably, U_4 captures some hitherto unseen signals characterized by a cubic-shaped network of positive densities that cage the iron atom. To facilitate a more concise description, a cross section of $Z = 0$ is defined as the heme plane and the cross section of $X = 0$ is the dividing plane between two propionic acid side chains, from which the third orthogonal cross section of $Y = 0$ can be derived (Figure S1). The eight corners of the positive cube are oriented along the X and Y cross sections. Strong positive densities appear in the middle of four six-membered rings encircled by adjacent pyrroles, the connecting methane bridge, and the heme iron (marked 1–4 in Figure 3c). These positive features reach their four minima when they cross the heme plane. These positive densities become more prominent above and below the heme plane (Figure 3). The eight corners of this cube are located 1.7 ± 0.2 Å from the iron. This cubic-shaped network is asymmetric with a notable skew toward CD rings (Figure 3f,h). Meanwhile, the iron skews in the opposite direction toward the AB rings with an out-of-plane displacement toward the proximal side (Figure 3e). However, the major features of this cube seem more symmetric in the Y cross section.

Similar features are also evident in the seventh component U_7 from an independent SVD analysis of 20 SAOMs. It is clear that the seventh dimension after a proper rotation (Methods) represents the reaction coordinate of the ligand dissociation. Mb structures along this reaction coordinate form a lineup on the c_7 dimension regardless of sample temperatures: from the ligated structures of MbCO, to the time-resolved structures, to several photoexcited species including the static structures Mb*CO and **CO, eventually to deoxy Mb (Figure 4e). U_7 of the omit maps contains the usual signals of ligand dissociation, such as strong negative densities on the ligand indicating dissociation and a positive layer on the distal side of the heme indicating its doming (Figure 5), except that the quality of these signals are not as good as those from difference maps because the SAOMs are derived from data sets in very different experiments. Despite the systematic errors, the seventh component U_7 of the omit maps again clearly shows the cubic network of positive densities around the iron. This cubic network also skews toward the CD rings of the heme. This is a reoccurrence of the same signals shown in U_4 of the difference maps (compare Figures 3 and 5). The eight corners of the positive cube are 1.9 ± 0.2 Å from the iron.

It is clear that the positive densities arranged in the cubic network around the iron, either in the difference maps or in the omit maps, cannot be easily attributed to accidental noise or systematic errors such as thermal effect or Fourier truncation error (see discussion below). They originate from the intrinsic signals induced by light and coexist with the previously seen signals of ligand dissociation. While multiple interpretations of these distinct cube-shaped signals may apply, this positive network clearly indicates a larger, high-spin iron as an immediate overall effect upon photolysis. Here I further attribute the distinct shape of the positive feature to the instantaneous regaining of the high-spin 3d orbitals of the heme iron. However, the observed cubic densities here are not an accurate image of a probability distribution of a 3d electron. They result from a net gain of electron densities as a fully ligated iron transitions to the high-spin deoxy state when several 3d orbitals are reoccupied

by unpaired electrons. Calculations of the combined probability distributions of the textbook $3d_{xz}$ and $3d_{yz}$ with either $3d_{xy}$ or $3d_{x^2-y^2}$ orbitals show features highly similar to the observed cubic-shaped map (Figure 3l,m). The strongest electron densities are located at eight corners of the cube where the lobes of 3d orbitals intercept. It is entirely possible that the observed cubic densities result from not only a superposition of several 3d orbitals but also an average of multiple electronic configurations. Such orbital changes have not been captured previously in protein crystals by X-ray crystallography.

However, the loss of the paired low-spin electrons in the ligated state is not clear in the negative densities. Presumably, the loss of the CO ligand and the drop of the iron out of the heme plane cause the displacement of many more electrons than the low-spin electrons. Negative densities in the component U_4 are overwhelming. At the available spatial resolution, the loss of the low-spin electrons is not separated from the main events upon photolysis.

Interestingly, the signals in U_4 of difference maps decay after a few picoseconds leading up to the static structures Mb*CO, **CO, and the deoxy Mb (Figure 2c,d). Given in-plane vibration modes^{34,61} with time periods comparable to the XFEL pulse duration, the averaged electron density between vibrating atoms is expected to increase, as shown in the six-membered rings. The observed decay of U_4 can be interpreted as the damping of these vibrations (see below). These in-plane vibrations trigger the iron to pop out of the heme plane.⁶² However, it remains unclear from this observation which, or a combination of which, vibration modes play the dominant role. Importantly, this decay does not reach zero even for the static deoxy Mb (Figure 2c,d), which suggests that the above interpretation of the cubic network of positive densities as the consequence of the spin crossover largely holds, because the cubic network cannot be fully explained by high-frequency vibrations. This decay of U_4 after a few picoseconds is also evidence to support the model of iron recoiling back into the heme plane,³⁷ which has to be discussed with the top component U_1 below.

After 1 ps, the electron density differences undergo complicated development, modulation, and damping. If the cubic network of positive densities is indeed influenced by the newly formed 3d orbitals, the dynamic behaviors of these high-spin electrons should be visible in the time points shortly after the initial formation. The observed evolution of the cubic network provides the necessary validation. Along the dimension c_1 in the SVD space of 93 difference maps, all time points after 1 ps, including the deoxy Mb (5d5r), representing a time point of ∞ , are well-separated from the subpicosecond time points around $c_1 = 0$ (Figure 2a,c). Therefore, the first component U_1 of the difference maps indicates a major development after 1 ps and represents more permanent changes afterward. The cubic positive densities in U_4 have no sign in U_1 . Among the major features captured by U_1 (Figure S3), the most noticeable is a rather skewed displacement of the iron toward the AB rings of the heme. Meanwhile, the motion of His93 corroborates the skewed displacement of the iron (Figure S3a,e). It is also shown below that the entire heme moves in its plane toward the AB rings as well. Such long-lived signals in U_1 from 1 ps and on are in good agreement with a wide distribution of individual trajectories and interquartile range of iron displacement predicted by molecular dynamics simulations.^{9,35,36} However, the cryo-

trapped structures, Mb*CO and **CO, are two exceptions. They exhibit very little characteristics in U_1 (Figure 2ac), even though they presumably represent long time delays. It can be reasoned that the skewed displacement of iron requires concerted motions in the proximal His93 and helix F, which are evidently forbidden at cryogenic temperatures.

Both U_1 and U_4 feature the signal of the iron out-of-plane displacement (Figures 3e and S3c). Therefore, the large positive coefficients c_1 and c_4 occurring simultaneously around several picoseconds indicate that the iron pops out of the heme plane most at this time but only momentarily. The later decay of the fourth component U_4 records a recoiling of the iron back into the heme plane around 100 ps (Movie S1). Another major signal revealed after 1 ps is a modulation of the new formation of 3d orbitals indicated by the cubic-shaped positive densities around the iron. This is a positional modulation manifested in the sixth and eighth components of the difference maps. While the corresponding coefficients c_6 and c_8 are nearly constant at zero in the subpicosecond range, they start to ramp up and vary at a few picoseconds. This variation can be most easily modeled as a two-dimensional oscillation at a period of 220 ± 20 ps given the limited time points (Figure 2f). Additional observations in the future may update this positional oscillation. The U_6 component features a distorted network of positive densities around the heme iron (Figure S4) similar to U_4 but with an inclination toward ring B. In U_6 , the strong negative density on iron is skewed toward the proximal side in contrast to U_4 (compare Figures 3e and S4c). The positive peak on the proximal side of the iron is also extremely skewed (Figure S4c). In a linear combination sense, U_6 and U_8 can be interpreted as variations that are combined into U_4 , resulting in an oscillation of a period at a few hundreds of picoseconds (Figure 2f). Specifically, the cubic-shaped densities lean toward ring D from 0.5 to 3 ps, return to the mean position after 50 ps, then sway toward ring B in the opposite direction (Figure 2f and Movie S1). Near the end of the first oscillation, the iron again pops out of the heme plane for the second time and severely skews toward AB rings (Movie S1). These observations of the iron motions support the previous model of iron recoiling.³⁷ Such oscillating behavior suggests modulation and damping of the vibrational modes. In other words, the cubic network of positive densities is observed not only in the earliest time points. The motions of this network are captured in later time points as well. This 220 ps oscillation is largely along the BD cross section (Figure S1), while damping of the vibrational amplitude was recorded only up to 150 ps and thus is insufficient to depict a complete picture. Two other more significant components U_2 and U_3 also show an oscillation of a distinct type as discussed below (Figure 2b).

An independent SVD analysis of 20 SAOMs masked around the heme site (Methods) offers a direct comparison of the XFEL data with previous synchrotron data and reveals an instantaneous thermal effect induced by XFEL pulses. Among the top eight components of the SAOMs, the first component shows a well-defined molecular image of the heme and its surrounding (Figure S5), and the second and fourth components apparently account for differences between hhMb and swMb. It is the third component U_3 that clearly distinguishes the XFEL data sets from the synchrotron data sets in a temperature-dependent manner (Figure 4a,c,e). In other words, the cryo structures (1dwr-t), ambient temperature structures (1bz6/p/r) including the deoxy hhMb at 7 °C (5d5r), and the room-temperature XFEL structures line up in an ascending order along the dimension c_3 . Not incidentally, the photolyzed structures Mb*CO and **CO have the most negative coefficient c_3 because they

were obtained at temperatures even lower than 100 K.⁴ The corresponding decomposed map U_3 reveals the thermal effects in electron densities of the heme as a nearly symmetric shell of spherical positive densities around the iron with very little positive or negative density on the iron itself (Figure 6), which transitions into a square shape with four corners centered at the six-membered rings in the heme plane (marked 1–4 in Figure 6e). This symmetric U_3 of the omit maps shows no usual signals for ligand dissociation and heme doming and therefore contrasts with the asymmetric U_4 of the difference maps characterized by the distinct cube-shaped positive density maxima at the eight corners. Two sheets of positive densities are found at the cross sections parallel to, and 1.3 Å from, the heme plane on both sides (Figure 6df). They are connected via pillar-shaped positive densities that run through the six-membered rings (marked 1–4 in Figure 6e) and the pyrrole rings (marked 5–8 in Figure 6e) along with some other positive densities outside the heme (marked 9–20 in Figure 6e). It is noteworthy that such thermal effects do not stand out as a major component in the SVD analysis of difference maps because they are largely self-canceled during the subtraction of the reference.

The temperature dependency and the symmetry of the signals in U_3 strongly suggest that these features are related to the local temperature at which the MbCO molecules were probed, rather than artifacts arising from structural refinement or difference in spatial resolutions among data sets. All data sets used for the analysis are unified to a similar spatial resolution (Methods). The positional parameters of the models within the omitted region and the thermal parameters of the entire globin are intentionally “forgotten” in the calculation of the omit maps. The large c_3 coefficients of the XFEL data sets are due to instantaneous temperature elevation at the heme site upon exposure to each ultrashort XFEL pulse, which introduces isotropic vibrations at high frequencies to the entire tetrapyrrole cofactor as captured by U_3 . This molecular image of U_3 shows an increasing intensity from cryo to ambient temperatures and to a very high local temperature induced by XFEL pulses (Movie S2). Considering only an estimate by linear extrapolation without other validations, the instantaneous temperature of the heme may exceed 500 K during an 80 fs XFEL pulse (Figure 6i). However, it remains to be seen whether this analysis of the omit maps restricted within the heme site can be extended to show that the temperature of the globin or the crystals has a significant elevation during an XFEL pulse.

Identification of such symmetric and temperature-dependent signals is important for interpretation of the highly dynamic ultrafast signals. The fact that the symmetric U_3 and the asymmetric U_7 of the omit maps are orthogonal to each other mathematically guarantees that these coexisting components are not interchangeable and not cross-contaminated between them (Methods). The cubic shaped U_7 of the omit maps and its counterpart U_4 of the difference maps are the chemical consequence of ligand dissociation; the symmetrical and spherical shaped U_3 originates from the thermal effect. One cannot compensate for the other, and both coexist in all maps. The equivalents can be stated: The cubic positive densities influenced by the 3d orbitals are not due to thermal vibrations; the symmetrically enlarged iron and the thickened heme due to an instantaneous temperature elevation have been isolated from the chemically relevant signals. It is worth noting that the ultrafast onset of U_4 of the difference maps (Figure 2cd) is not in contradiction with U_7 of the omit maps contained not only in time-resolved XFEL data sets but also in the static structures of deoxy

Mb and photolyzed Mb*CO and **CO at different temperatures (Figure 4e). Instead, regaining of the high-spin 3d orbitals of the iron occurs before the first time point at tens of femtoseconds; this permanent change lasts throughout the dissociation reaction into the deoxy product.

The asymmetric U_7 of the omit maps also describes an important but unfamiliar structural event upon the ligand dissociation—the heme in-plane movement. This type of motion is harder to capture in difference maps because of little net gain or loss of electron density that would be caused by an in-plane movement of the heme. With respect to the protein framework, the heme moves toward its AB rings in addition to the other changes (Movie S3). Heme in-plane sliding is consistent with the strong signals of the skewed iron displacement and the proximal His93 (Figure S3) that are closely related to structural responses in the globin. Similar in-plane sliding is even more significant in tetrameric hemoglobin, ⁶³ which is the origin of the cooperative oxygen affinity.⁶⁰ However, the functional role of the heme in-plane sliding in Mb is not yet clear, and this discussion branches from the topic here on ultrafast structural changes.

A recent solution scattering experiment using XFEL pulses showed that the radius of gyration R_g first increases within 1 ps upon photodissociation of MbCO and then oscillates with a period of 3.6 ps, equivalent to 9 cm^{-1} , a frequency lower than those observed by coherence spectroscopy.⁶⁴ The molecular volume increases afterward with a phase shift. These observations indicate a wavelike motion propagating outward in the protein. Recent molecular dynamics simulations have confirmed that a pressure wave generated by photolysis reaches the molecular surface at 300 fs along the direction perpendicular to the heme plane,⁶⁵ which corresponds to a frequency of 28 cm^{-1} if the wave is reflected by the surface and oscillates. It was found by the simulations that the oscillation in the solvent would be overdamped and thus dissipates quickly.

Some SVD components derived here display strict correlation suggesting their dependency on each other. The U_2 – U_3 pair of the difference maps (Figure 2b) is evidently equivalent to the U_5 – U_8 pair of the omit maps (Figure 4b). When the difference maps of the entire globin are analyzed, the same signals emerge as the top-ranked U_1 – U_3 pair suggesting that the local signals at the heme site concur with global signals over the entire globin (Figure 7). Interestingly, this pair of components displays a nearperfect circular correlation at subpicosecond delays but is almost absent at longer delays $>1\text{ ps}$ (Figure 4a,f,g,h). Such an oscillation at a low frequency requires at least two orthogonal components to depict. The oscillating signals have spread over the entire globin at subpicosecond delays while their amplitudes decay to nearly zero after a few picoseconds. Based on least-squares fittings of the corresponding coefficients (Methods), the frequency of the oscillation is determined to be $36 \pm 1\text{ cm}^{-1}$, corresponding to a period of 0.93 ps. The equivalent oscillation frequencies determined from the difference maps and omit maps at the heme site are $42 \pm 2\text{ cm}^{-1}$ (a period of 0.8 ps) and $37 \pm 1\text{ cm}^{-1}$ (a period of 0.89 ps), respectively.

The low-frequency oscillation determined from the XFEL data sets of Barends and Schlichting et al. is highly comparable to those measured by femtosecond vibrational coherence spectroscopy.^{37,38} I postulate that this low frequency reflects the molecular size of

the solvated Mb rather than the property of the heme. First, the corresponding period of 0.93 ps is about the time that a mechanical wave takes to traverse 14 Å, that is, an average distance from the heme to the molecular surface of Mb, given a speed of sound at 1.5 km/s or nm/ps (Figure S1). Second, the molecular sizes of several heme proteins, such as cytochrome c (12 kDa), Mb (17 kDa), and cystathionine β -synthase (63 kDa), show good anticorrelation with their dominant low frequencies of 44, 40, and 25 cm⁻¹, respectively.^{37,66,67} Third, no association between the oscillating signals and specific structural features can be found, such as the heme, ligand, surrounding side chains, or helices (Figure 7i,k). These oscillating signals sharply contrast with strong, non-oscillating signals associated with the heme and various helices evidenced by the second and fourth components of the difference maps (Figure 7j,l), where helices E, G, and H move away from the heme in response to ligand dissociation.⁹ These motion signals of various structural elements are much enhanced by SVD.

The oscillating signals at low frequency seem to propagate via solvent within the protein matrix rather than the globin. Strong positive and negative peaks in U₅, U₆, and U₈ of the omit maps are consistently found in the solvent region away from the heme and protein. However, a specific description on how the solvent responds to ligand dissociation is difficult, because the solvent structure is far less well-determined than the globin structure. An overall contribution of U₅, U₆, and U₈ can be judged by a sum of the absolute values |U₅| + |U₆| + |U₈| that enhances the distribution of signals. The heme and globin are located in low intensities of these oscillating components (Figure 7b–e). These global oscillations occur only a few times before dissipation.^{37,38} However, there are not enough time points to determine the damping time constant, which is roughly estimated to be 1–4 ps. Contrary to the common belief that low-frequency oscillations are due to the properties of the heme, here I attribute the low-frequency oscillation to the global response of protein and solvent as a whole to a sudden molecular event, photodissociation of the ligand, where the vibrational energy released from the photodissociation is quickly dissipated into the environment.

It requires ultrahigh spatial resolution in a Fourier synthesis to accurately depict the probability distribution of a single electron occupying its orbital. For example, iron 3d and sulfur 3p densities in an Fe₄S₄ cluster are described with crystallographic data at a resolution of 0.48 Å with an analytical technique called charge-density analysis,⁵⁴ in which atoms are considered as aspherical and modeled by spherical harmonic functions in contrast to spherical Gaussian functions used in everyday protein crystallography.⁶⁸ Because the nonbonding inner electrons remain largely constant in a spherical distribution as a whole, the inner electrons in a majority can easily overwhelm the presence of the bonding electrons in a minority, let alone subtle changes in electron transfer. Therefore, it is very difficult, if not impossible, to detect small orbital changes upon bond rupture using regular electron density maps such as a 2Fo-Fc map. Difference Fourier technique is required as it is far more sensitive to small changes,⁶⁹ in which the constant, spherical distribution of the inner electrons is self-canceled and the remaining densities would reveal electron transfer. However, the reality is usually far less ideal because of conformational changes. Atomic displacements involved in a conformational change carry all electrons associated with the moving atoms and result in significant positive and negative densities in a difference Fourier map, that is, a spatial distribution of any gain and loss of electrons. Atomic displacements

such as a side chain rotation or the heme iron movement by a few tenths of an angstrom again conceal the subtle signal of an electron transfer from one orbital to another.

The difference maps derived here as SVD components at the available spatial resolution of 1.8 Å represent partial, variable signals in addition to the constant electron density map regardless of the bonding change. They are not accurate images of probability distributions of atomic orbitals. An ultrahigh spatial resolution is required to depict slightly aspherical features when an overwhelming spherical background is present. Here, a combined use of difference Fourier technique and decomposition eases the demand for spatial resolution to reveal subtle changes. Furthermore, the notion that an ultrahigh spatial resolution is required for observations of orbital changes contradicts the fact that there is no discontinuity in the probability distribution of any electronic orbital. A reasonably high spatial resolution shall be sufficient to analytically describe the smooth function of a gain or loss of an electron. Therefore, the small amplitude of desired changes over a much greater constant background hinders a reliable observation more than the lack of high spatial frequencies does. This work presents the effectiveness of the strategy that enables numerical deconvolution to resolve concurrent events so that the image of a partial change, such as a single electron transfer, can be isolated from other changes. At an available spatial resolution, the partial change due to a specific electronic event deconvoluted from much greater conformational changes could be the remanence of orbital densities instead of an accurate image of these orbitals.

A limited experimental spatial resolution also raises a doubt whether some unusual features are the artifacts due to a Fourier truncation error. A Fourier truncation error may occur to several static data sets included in this analysis, in which the spatial resolutions are truncated to 1.75 Å to match that of XFEL data sets. Such errors may appear as minor components after SVD (Figures 2e and 4f). When the average intensity of an experimental data set naturally decays as a function of the spatial resolution, Fourier truncation errors are damped to a minimum. Difference Fourier technique again self-cancels the remaining errors because both data sets involved in the subtraction would carry similar Fourier truncation errors if any. In addition, the applied weighting scheme^{8,19,70,71} severely reduces any large differences in Fourier synthesis (Methods), which also prevents Fourier truncation errors. After the decomposition, proper rotations in the SVD space validate that the signals under discussion, such as the positive cubic network, are not due to Fourier truncation errors because of its correlation with the metadata. An artifact due to Fourier truncation errors, if so concluded, needs to be explained why it is correlated with a physical or chemical parameter.

The theoretical radii of the 3d orbitals for an isolated Fe(II) cation are about 0.36 Å at the maximum probability of these orbitals. However, the orbital densities spread out many fold beyond the maxima.⁷² The majority of the iron 3d orbital densities in an Fe₄S₄ cluster span about 1 Å from the nucleus as measured in a protein.⁵⁴ The observed distances between the iron and the corners of the observed cubic densities are significantly larger. Therefore, the observed cubic densities here are not an accurate image of a probability distribution of the 3d electrons. However, this density distribution, given the available spatial resolution, is clearly influenced by the unpaired 3d electrons. The net gain of electron densities could be further from the nucleus compared to the maximum probability of these orbital. An accurate depiction of a 3d orbital is not a proper task in protein crystallography given the practical

limit of spatial resolution. Rather, direct observations of a specific electronic event would provide unprecedented insight into the essence of biochemistry.

Ultrashort XFEL pulses improve the time resolution to subpicosecond or even tens of femtoseconds compared to 100 ps at synchrotrons. This ultrafast time resolution, when combined with the spatial resolution of crystallography, offers an unprecedented opportunity for direct observation of transient structural changes in a photochemical reaction as detailed as individual electrons. However, because of their highly dynamic and heterogeneous nature, it is difficult to experimentally isolate short-lived structural species along the temporal dimension regardless of the ultrafast time resolution. In other words, resolving changes in the populations of chemical species from one femtosecond to the next does not mean that one can resolve two distinct structures. The strategy we proposed for resolution of structural heterogeneity¹⁹ is highly effective for separating independent electronic and atomic events that often coexist during a chemical reaction or a biological process. A major advance in data analysis demonstrated in this work is the strategy of decomposition and deconvolution. A full matrix of time-resolved data sets (Table S1) is first decomposed into orthonormal components by SVD (Methods). A multidimensional rotation in a Euclidean subspace identified by several major singular values deconvolutes the entangled effects from multiple physical and chemical parameters. The core crystallographic observations are correlated with their corresponding metadata, hence allowing the origin of small signals to be revealed. This analytical strategy demonstrates that it is possible to isolate structural changes as small as a gain or loss of individual electrons in protein crystallography.

All observations so far are the consequence, rather than the cause, of the ligand dissociation. The proposed transient event of metal-to-ligand charge transfer is not captured here, presumably because the transferred electron is delocalized all over the porphyrin.²⁸ Therefore, the fundamental questions remain for photolysis of MbCO: Upon the absorption of a quantum of energy from a green photon, what initial electronic events lead to a partitioning of the bonding electrons in two separate atoms, that is, the bond rupture? What electronic events cause spin crossover? To image these events directly, even higher time resolution is required to capture electron density changes prior to and during bond rupture.

Despite the recent developments in SFX and its applications to time-resolved studies of a few model systems, it is yet premature to declare that the current SFX platform would widely apply to other systems of more significant biological importance such as photosynthetic reaction centers, visual and microbial rhodopsins, plant and bacterial phytochromes, plant and animal cryptochromes, and light-dependent DNA repair enzymes. There are at least three major challenges. First, it is not practical to produce and inject an astronomical number of nano- or microcrystals of these proteins into an XFEL beam to support data collections of a time series. The current serial protocol used to obtain the MbCO and other data has a diminishingly small yield of useful diffraction data, based on the ratio between the number of diffracting crystals and the total number of crystals produced.⁷³ That is, the vast majority of the crystals have no chance to diffract X-rays even when they are injected into the X-ray beam. Such a yield, defined differently from the hit rate, is not sustainable for other photosensitive protein crystals. Several “fixed target” implementations of crystal delivery systems have been reported, including one that we have recently

proposed.⁷³ However, it remains to be seen how effective they are in achieving a higher yield while maintaining the diffraction quality.

Second, because the XFEL pulses usually carry insufficient bandwidths to fully integrate the Bragg reflections and the angular rotation of a crystal during an ultrashort XFEL pulse is not attainable, tens to hundreds of thousands of diffraction images are required to obtain integrated intensities for all Bragg reflections in a complete sampling of the reciprocal space. A possibility to solve this partiality problem is to use an X-ray beam with a significant convergence of about a degree,^{74,75} which will produce completely integrated intensities with a far fewer number of diffraction images, thereby alleviating the demand for a large quantity of purified proteins.

Third, data merging from a large crystal pool is a major source of error in producing difference signals for photosensitive crystals, particularly those with much lower diffraction power than Mb crystals. To address this challenge, we also proposed a protocol to obtain difference signals that bypasses the integrated intensities of Bragg reflections. We reason that integrated intensities are not necessary if the dark and light diffraction images can be obtained within a single exposure of an XFEL pulse because the percentage change in integrated intensities of a given reflection, that is, the time-resolved signal, is already captured even if this reflection is partially measured in both dark and light conditions. A plan has been laid out to implement this protocol, based on a dichromator instrument that achieves angular split and temporal delay at the same time.⁷⁶ When each of these technical hurdles is overcome, direct observations of atomic and electronic movements during ultrafast processes will be closer to broad applications for many light-sensitive crystals of great biological and medical significance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The following database and software are used in this work: CCP4 (ccp4.ac.uk), Coot (www2.mrc-lmb.cam.ac.uk/Personal/pemsley/coot), dynamX (Renz Research, Inc.), gnuplot (gnuplot.info), Mathematica (wolfram.com), PDB (rcsb.org), PHENIX (phenix-online.org), PyMOL (pymol.org), Python (python.org), and SciPy (scipy.org).

REFERENCES

- (1). Levantino M; Yorke BA; Monteiro DC; Cammarata M; Pearson AR Using Synchrotrons and XFELs for Time-Resolved X-Ray Crystallography and Solution Scattering Experiments on Biomolecules. *Curr. Opin. Struct. Biol* 2015, 35, 41–48. [PubMed: 26342489]
- (2). Bourgeois D; Vallone B; Schotte F; Arcovito A; Miele AE; Sciara G; Wulff M; Anfinrud P; Brunori M Complex Landscape of Protein Structural Dynamics Unveiled by Nanosecond Laue Crystallography. *Proc. Natl. Acad. Sci. U. S. A* 2003, 100 (15), 8704–8709. [PubMed: 12847289]
- (3). Bourgeois D; Vallone B; Arcovito A; Sciara G; Schotte F; Anfinrud PA; Brunori M Extended Subnanosecond Structural Dynamics of Myoglobin Revealed by Laue Crystallography. *Proc. Natl. Acad. Sci. U. S. A* 2006, 103 (13), 4924–4929. [PubMed: 16547137]
- (4). Chu K; Vojtechovsky J; McMahon BH; Sweet RM; Berendzen J; Schlichting I Structure of a Ligand-Binding Intermediate in Wild-Type Carbonmonoxy Myoglobin. *Nature* 2000, 403 (6772), 921–923. [PubMed: 10706294]

- (5). Kachalova GS; Popov AN; Bartunik HD A Steric Mechanism for Inhibition of CO Binding to Heme Proteins. *Science* 1999, 284 (5413), 473–476. [PubMed: 10205052]
- (6). Schotte F; Lim M; Jackson TA; Smirnov AV; Soman J; Olson JS; Phillips GN; Wulff M; Anfinrud PA Watching a Protein as It Functions with 150-Ps Time-Resolved X-Ray Crystallography. *Science* 2003, 300 (5627), 1944–1947. [PubMed: 12817148]
- (7). Šrajer V; Teng T.-y.; Ursby T; Pradervand C; Ren Z; Adachi S.-i.; Schildkamp W; Bourgeois D; Wulff M; Moffat K. Photolysis of the Carbon Monoxide Complex of Myoglobin: Nanosecond Time-Resolved Crystallography. *Science* 1996, 274 (5293), 1726–1729. [PubMed: 8939867]
- (8). Šrajer V; Ren Z; Teng T-Y; Schmidt M; Ursby T; Bourgeois D; Pradervand C; Schildkamp W; Wulff M; Moffat K Protein Conformational Relaxation and Ligand Migration in Myoglobin: A Nanosecond to Millisecond Molecular Movie from Time-Resolved Laue X-Ray Diffraction. *Biochemistry* 2001, 40 (46), 13802–13815. [PubMed: 11705369]
- (9). Barends TRM; Foucar L; Ardevol A; Nass K; Aquila A; Botha S; Doak RB; Falahati K; Hartmann E; Hilpert M; et al. Direct Observation of Ultrafast Collective Motions in CO Myoglobin upon Ligand Dissociation. *Science* 2015, 350 (6259), 445–450. [PubMed: 26359336]
- (10). Bourgeois D; Schotte F; Brunori M; Vallone B Time-Resolved Methods in Biophysics. 6. Time-Resolved Laue Crystallography as a Tool to Investigate Photo-Activated Protein Dynamics. *Photochem. Photobiol. Sci* 2007, 6 (10), 1047. [PubMed: 17914477]
- (11). Clifton IJ; Duke EMH; Wakatsuki S; Ren Z Evaluation of Laue Diffraction Patterns. *Methods Enzymol.* 1997, 277, 448–467. [PubMed: 18488319]
- (12). Ren Z; Bourgeois D; Helliwell JR; Moffat K; Šrajer V; Stoddard BL Laue Crystallography: Coming of Age. *J. Synchrotron Radiat* 1999, 6 (4), 891–917.
- (13). Aranda R; Levin EJ; Schotte F; Anfinrud PA; Phillips GN Time-Dependent Atomic Coordinates for the Dissociation of Carbon Monoxide from Myoglobin. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 2006, 62 (7), 776–783. [PubMed: 16790933]
- (14). Jung YO; Lee JH; Kim J; Schmidt M; Moffat K; Šrajer V; Ihee H Volume-Conserving Trans–Cis Isomerization Pathways in Photoactive Yellow Protein Visualized by Picosecond X-Ray Crystallography. *Nat. Chem* 2013, 5 (3), 212–220. [PubMed: 23422563]
- (15). Ren Z; Šrajer V; Knapp JE; Royer WE Cooperative Macromolecular Device Revealed by Meta-Analysis of Static and Time-Resolved Structures. *Proc. Natl. Acad. Sci. U. S.A* 2012, 109 (1), 107112.
- (16). Schotte F; Cho HS; Kaila VRI; Kamikubo H; Dashdorj N; Henry ER; Graber TJ; Henning R; Wulff M; Hummer G; et al. Watching a Signaling Protein Function in Real Time via 100-Ps Time-Resolved Laue Crystallography. *Proc. Natl. Acad. Sci. U. S. A* 2012, 109 (47), 19256–19261. [PubMed: 23132943]
- (17). Schmidt M; Rajagopal S; Ren Z; Moffat K Application of Singular Value Decomposition to the Analysis of Time-Resolved Macromolecular X-Ray Data. *Biophys. J* 2003, 84 (3), 2112–2129. [PubMed: 12609912]
- (18). Ren Z Reaction Trajectory Revealed by a Joint Analysis of Protein Data Bank. *PLoS One* 2013, 8 (11), No. e77141. [PubMed: 24244274]
- (19). Ren Z; Chan PWY; Moffat K; Pai EF; Royer WE; Šrajer V; Yang X Resolution of Structural Heterogeneity in Dynamic Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 2013, D69 (6), 946–959.
- (20). Bandara S; Ren Z; Lu L; Xiaoli Z; Shin H; Zhao K-H; Yang X Photoactivation Mechanism of a Carotenoid-Based Photoreceptor. *Proc. Natl. Acad. Sci. U. S. A* 2017, 114 (24), 6286–6291. [PubMed: 28559328]
- (21). Kim TH; Mehrabi P; Ren Z; Sljoka A; Ing C; Bezginov A; Ye L; Pomès R; Prosser RS; Pai EF The Role of Dimer Asymmetry and Protomer Dynamics in Enzyme Catalysis. *Science* 2017, 355 (6322), No. eaag2355. [PubMed: 28104837]
- (22). Ren Z Molecular Events during Translocation and Proof reading Extracted from 200 Static Structures of DNA Polymerase. *Nucleic Acids Res.* 2016, 6 (43129), 7457.
- (23). Yang X; Ren Z; Kuk J; Moffat K Temperature-Scan Cryocrystallography Reveals Reaction Intermediates in Bacteriophytochrome. *Nature* 2011, 479 (7373), 428–432. [PubMed: 22002602]

- (24). Zeng X; Ren Z; Wu Q; Fan J; Peng P-P; Tang K; Zhang R; Zhao K-H; Yang X Dynamic Crystallography Reveals Early Signaling Events in Ultraviolet Photoreceptor UVR8. *Nat. Plants* 2015, 1 (1), 14006. [PubMed: 26097745]
- (25). Dunitz BD; Dreuw A; Head-Gordon M Initial Steps of the Photodissociation of the CO Ligated Heme Group. *J. Phys. Chem. B* 2003, 107 (23), 5623–5629.
- (26). Falahati K; Tamura H; Burghardt I; Huix-Rotllant M Quantum Dynamics of Ultrafast Heme-Carbon Monoxide Photolysis and Spin-Crossover. *ArXiv1709.03402 [physics.chem-ph]* 2017.
- (27). Falahati K; Tamura H; Burghardt I; Huix-Rotllant M Ultrafast Carbon Monoxide Photolysis and Heme Spin-Crossover in Myoglobin via Nonadiabatic Quantum Dynamics. *Nat. Commun* 2018, 9 (1), 4502. [PubMed: 30374057]
- (28). Franzen S; Kiger L; Poyart C; Martin J-L Heme Photolysis Occurs by Ultrafast Excited State Metal-to-Ring Charge Transfer. *Biophys. J* 2001, 80 (5), 2372–2385. [PubMed: 11325737]
- (29). Champion PM Chemistry: Following the Flow of Energy in Biomolecules. *Science* 2005, 310 (5750), 980–982. [PubMed: 16284167]
- (30). Fujisaki H; Straub JE Vibrational Energy Relaxation in Proteins. *Proc. Natl. Acad. Sci. U. S. A* 2005, 102 (19), 6726–6731. [PubMed: 15767588]
- (31). Armstrong MR; Ogilvie JP; Cowan ML; Nagy AM; Miller RJD Observation of the Cascaded Atomic-to-Global Length Scales Driving Protein Motion. *Proc. Natl. Acad. Sci. U. S. A* 2003, 100 (9), 4990–4994. [PubMed: 12697894]
- (32). Kitagawa T; Haruta N; Mizutani Y Time-Resolved Resonance Raman Study on Ultrafast Structural Relaxation and Vibrational Cooling of Photodissociated Carbonmonoxy Myoglobin. *Biopolymers* 2002, 67 (4–5), 207–213. [PubMed: 12012433]
- (33). Mizutani Y Direct Observation of Cooling of Heme upon Photodissociation of Carbonmonoxy Myoglobin. *Science* 1997, 278 (5337), 443–446. [PubMed: 9334299]
- (34). Nagy AM; Raicu V; Miller RJD. Nonlinear Optical Studies of Heme Protein Dynamics: Implications for Proteins as Hybrid States of Matter. *Biochim. Biophys. Acta, Proteins Proteomics* 2005, 1749 (2), 148–172.
- (35). Henry ER; Levitt M; Eaton WA Molecular Dynamics Simulation of Photodissociation of Carbon Monoxide from Hemoglobin. *Proc. Natl. Acad. Sci. U. S. A* 1985, 82, 2034–2038. [PubMed: 3856881]
- (36). Kuczera K; Lambry J-C; Martin J-L; Karplus M Nonexponential Relaxation after Ligand Dissociation from Myoglobin: A Molecular Dynamics Simulation. *Proc. Natl. Acad. Sci. U. S. A* 1993, 90, 5805–5807. [PubMed: 8516332]
- (37). Gruia F; Kubo M; Ye X; Champion PM Investigations of Vibrational Coherence in the Low-Frequency Region of Ferric Heme Proteins. *Biophys. J* 2008, 94 (6), 2252–2268. [PubMed: 18065461]
- (38). Rosca F; Kumar ATN; Ionascu D; Ye X; Demidov AA; Sjodin T; Wharton D; Barrick D; Sligar SG; Yonetani T; et al. Investigations of Anharmonic Low-Frequency Oscillations in Heme Proteins. *J. Phys. Chem. A* 2002, 106 (14), 3540–3552.
- (39). Humphreys CJ Electrons Seen in Orbit. *Nature* 1999, 401 (6748), 21–22.
- (40). Zuo JM; Kim M; O’Keeffe M; Spence JCH. Direct Observation of D-Orbital Holes and Cu–Cu Bonding in Cu₂O. *Nature* 401 (6748), 49–52.
- (41). Mulder P Are Orbitals Observable? *Int. J. Philos. Chem* 2011, 17 (1), 24–35.
- (42). Scerri E R Have Orbitals Really Been Observed? *J. Chem. Educ* 2000, 77 (11), 1492.
- (43). Itatani J; Levesque J; Zeidler D; Niikura H; Pépin H; Kieffer JC; Corkum PB; Villeneuve DM Tomographic Imaging of Molecular Orbitals. *Nature* 2004, 432 (7019), 867–871. [PubMed: 15602553]
- (44). Ogilvie JF Is a Molecular Orbital Measurable by Means of Tomographic Imaging? *Found. Chem* 2011, 13 (2), 87–91.
- (45). Schwarz WHE Measuring Orbitals: Provocation or Reality? *Angew. Chem., Int. Ed* 2006, 45 (10), 1508–1517.
- (46). Stapelfeldt H Electrons Frozen in Motion. *Nature* 2004, 432 (7019), 809–810. [PubMed: 15602533]

- (47). Vozzi C; Negro M; Calegari F; Sansone G; Nisoli M; De Silvestri S; Stagira S Generalized Molecular Orbital Tomography. *Nat. Phys* 2011, 7 (10), 822–826.
- (48). Chaika AN Visualization of Electron Orbitals in Scanning Tunneling Microscopy. *JETP Lett.* 2014, 99 (12), 731–741.
- (49). Patera LL; Queck F; Scheuerer P; Repp J Mapping Orbital Changes upon Electron Transfer with Tunnelling Microscopy on Insulators. *Nature* 2019, 566 (7743), 245–248. [PubMed: 30760911]
- (50). Repp J; Meyer G; Stojkovi SM; Gourdon A; Joachim C Molecules on Insulating Films: Scanning-Tunneling Microscopy Imaging of Individual Molecular Orbitals. *Phys. Rev. Lett* 2005, 94 (2) , 026803 DOI: 10.1103/PhysRevLett.94.026803. [PubMed: 15698209]
- (51). Mikhailovskij IM; Sadanov EV; Mazilova TI; Ksenofontov VA; Velicodnaja OA Imaging the Atomic Orbitals of Carbon Atomic Chains with Field-Emission Electron Microscopy. *Phys. Rev. B: Condens. Matter Mater. Phys* 2009, 80 (16), 165404 DOI: 10.1103/PhysRevB.80.165404.
- (52). Puschnig P; Berkebille S; Fleming AJ; Koller G; Emtsev K; Seyller T; Riley JD; Ambrosch-Draxl C; Netzer FP; Ramsey MG Reconstruction of Molecular Orbital Densities from Photoemission Data. *Science* 2009, 326 (5953), 702–706. [PubMed: 19745118]
- (53). Stodolna AS; Rouzée A; Lépine F; Cohen S; Robicheaux F; Gijsbertsen A; Jungmann JH; Bordas C; Vrakking MJJ Hydrogen Atoms under Magnification: Direct Observation of the Nodal Structure of Stark States. *Phys. Rev. Lett* 2013, 110 (21), 213001 DOI: 10.1103/PhysRevLett.110.213001. [PubMed: 23745864]
- (54). Hirano Y; Takeda K; Miki K Charge-Density Analysis of an Iron–Sulfur Protein at an Ultra-High Resolution of 0.48 Å. *Nature* 2016, 534 (7606), 281–284. [PubMed: 27279229]
- (55). Haldrup K; Harlang T; Christensen M; Dohn A; van Driel TB; Kjær KS; Harrit N; Vibenholt J; Guerin L; Wulff M; et al. Bond Shortening (1.4 Å) in the Singlet and Triplet Excited States of [Ir2(Dimen)4]2+ in Solution Determined by Time-Resolved X-Ray Scattering. *Inorg. Chem* 2011, 50 (19), 9329–9336. [PubMed: 21823573]
- (56). Bionta MR; Lemke HT; Cryan JP; Glowonia JM; Bostedt C ; Cammarata M; Castagna J-C; Ding Y; Fritz DM; Fry AR; et al. Spectral Encoding of X-Ray/Optical Relative Delay. *Opt. Express* 2011, 19 (22), 21855. [PubMed: 22109037]
- (57). Bionta MR; Hartmann N; Weaver M; French D; Nicholson DJ; Cryan JP; Glowonia JM; Baker K; Bostedt C; Chollet M; et al. Spectral Encoding Method for Measuring the Relative Arrival Time between X-Ray/Optical Pulses. *Rev. Sci. Instrum* 2014, 85 (8), 083116. [PubMed: 25173255]
- (58). Hodel A; Kim SH; Brünger AT Model Bias in Macromolecular Crystal Structures. *Acta Crystallogr., Sect. A: Found. Crystallogr* 1992, 48 (6), 851–858.
- (59). Henry ER; Hofrichter J. Singular Value Decomposition: Application to Analysis of Experimental Data In Numerical Computer Methods; Academic Press, 1992; Vol. 210, pp 129–192.
- (60). Ren Z Reverse Engineering the Cooperative Machinery of Human Hemoglobin. *PLoS One* 2013, 8 (11), No. e77363. [PubMed: 24312167]
- (61). Abe M; Kitagawa T; Kyogoku Y Resonance Raman Spectra of Octaethylporphyrinato-Ni(II) and Meso-deuterated and 15N Substituted Derivatives. II. A Normal Coordinate Analysis. *J. Chem. Phys* 1978, 69 (10), 4526–4534.
- (62). Franzen S; Bohn B; Poyart C; DePillis G; Boxer SG; Martin J-L Functional Aspects of Ultra-Rapid Heme Doming in Hemoglobin, Myoglobin, and the Myoglobin Mutant H93G. *J. Biol. Chem* 1995, 270 (4), 1718–1720. [PubMed: 7829506]
- (63). Baldwin J; Chothia C Haemoglobin: The Structural Changes Related to Ligand Binding and Its Allosteric Mechanism. *J. Mol. Biol* 1979, 129 (2), 175–200. [PubMed: 39173]
- (64). Levantino M; Schirò G; Lemke HT; Cottone G; Glowonia JM; Zhu D; Chollet M; Ihee H; Cupane A; Cammarata M Ultrafast Myoglobin Structural Dynamics Observed with an X-Ray Free-Electron Laser. *Nat. Commun* 2015, 6, 6772. [PubMed: 25832715]
- (65). Brinkmann LUL; Hub JS Ultrafast Anisotropic Protein Quake Propagation after CO Photodissociation in Myoglobin. *Proc. Natl. Acad. Sci. U. S. A* 2016, 113 (38), 10565–10570. [PubMed: 27601659]
- (66). Karunakaran V; Benabbas A; Sun Y; Zhang Z; Singh S; Banerjee R; Champion PM Investigations of Low-Frequency Vibrational Dynamics and Ligand Binding Kinetics of Cystathionine β -Synthase. *J. Phys. Chem. B* 2010, 114 (9), 3294–3306. [PubMed: 20155941]

- (67). Karunakaran V; Sun Y; Benabbas A; Champion PM Investigations of the Low Frequency Modes of Ferric Cytochrome c Using Vibrational Coherence Spectroscopy. *J. Phys. Chem. B* 2014, 118 (23), 6062–6070. [PubMed: 24823442]
- (68). Coppens P Charge-Density Analysis at the Turn of the Century. *Acta Crystallogr., Sect. A: Found. Crystallogr* 1998, 54 (6), 779–788.
- (69). Henderson R; Moffat JK The Difference Fourier Technique in Protein Crystallography: Errors and Their Treatment. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem* 1971, B27 (7), 1414–1420.
- (70). Ren Z; Perman B; Srajer V; Teng T-Y; Pradervand C; Bourgeois D; Schotte F; Ursby T; Kort R; Wulff M; et al. A Molecular Movie at 1.8 Å Resolution Displays the Photocycle of Photoactive Yellow Protein, a Eubacterial Blue-Light Receptor, from Nanoseconds to Seconds. *Biochemistry* 2001, 40 (46), 13788–13801. [PubMed: 11705368]
- (71). Ursby T; Bourgeois D Improved Estimation of Structure-Factor Difference Amplitudes from Poorly Accurate Data. *Acta Crystallogr., Sect. A: Found. Crystallogr* 1997, 53 (5), 564–575.
- (72). Waber JT; Cromer DT Orbital Radii of Atoms and Ions. *J. Chem. Phys* 1965, 42 (12), 4116–4123.
- (73). Ren Z; Ayhan M; Bandara S; Bowatte K; Kumarapperuma I; Gunawardana S; Shin H; Wang C; Zeng X; Yang X Crystal-on-Crystal Chips for in Situ Serial Diffraction at Room Temperature. *Lab Chip* 2018, 18, 2246–2256. [PubMed: 29952383]
- (74). Ho JX; Ruble JR; McInnis TR; Carter DC; Huang H; Gibson WM Convergent-Beam Method in Macromolecular Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr* 2002, 58 (12), 2087–2095. [PubMed: 12454469]
- (75). Spence JCH; Zatsepin NA; Li C Coherent Convergent-Beam Time-Resolved X-Ray Diffraction. *Philos. Trans. R. Soc., B* 2014, 369 (1647), 20130325–20130325.
- (76). Ren Z; Yang X Angular-Split/Temporal-Delay Approach to Ultrafast Protein Dynamics at XFELs. *Acta Crystallogr. Sect. Struct. Biol* 2016, 72 (7), 871–882.

Figure 1.

Gantt chart summarizing structural events during photodissociation of MbCO. Previously known structural events are in pink. The excitation laser pulse and XFEL pulses are in yellow. The newly observed structural events from decomposition of the serial crystallographic data sets are in dark blue.

Figure 2.

SVD of 93 difference maps around the heme site. These difference maps are divided into several series. Each series consists of the difference maps produced from subtracting the same reference (Methods and Table S1 in the Supporting Information). Each of the 93 difference maps can be expressed as a linear combination of many components $c_1U_1 + c_2U_2 + \dots$ after the decomposition. Among a common set of basis components U_k ($k = 1, 2, \dots$) shared by all difference maps, also known as left singular vectors in linear algebra, the first few components are the most significant. Each of the difference maps requires a coefficient

set c_k for the linear combination. This coefficient set varies from one map to the other; that is, a function of the delay time or other metadata associate with the data sets. Each coefficient c_k indicates how much of the corresponding component U_k exists in the linear combination. Each coefficient set is derived from a right singular vector as traditionally known in linear algebra (Methods). These coefficients produced by SVD should not be confused with the Fourier coefficients used to synthesize each map. (a–d) Orthographic projections of coefficients in some of the significant dimensions. Each difference map is represented by a dot at the location that marks the coefficients associated with the major components U_1 through U_4 . Difference maps in a series are connected by lines in a corresponding color. The multidimensional space is projected to various planes and tiled together so that any two adjacent panels from a to d, horizontal or vertical, can be imagined erecting a three-dimensional subspace. (e) Singular values indicating the significance of the components are sorted in a descending order before a rotation. A multidimensional rotation alters their values. (f) Coefficients c_6 and c_8 of the $-F_{-0,1}$ series in yellow are least-squares fitted with a two-dimensional oscillation (Methods).

Figure 3.

Ultrafast formation of high-spin 3d orbitals of the heme iron. Each component resulting from SVD, that is, a left singular vector, can also be presented as a difference map. The fourth component U_4 of the difference maps captures a cubic shaped positive density network. It is rendered as a difference map in two styles: two-dimensional cross sections and three-dimensional contours. The corresponding coefficient c_4 is plotted in Figure 2c,d. (a–h) Images of two-dimensional cross sections through the difference map of the component U_4 . Cross sections defined in Figure S1 are labeled in the top right corner of each panel. Noncentric cross sections are marked to indicate small offsets from the centric cross

sections. The structure of the heme, CO ligand in both bound and docked positions, and some side chains are also projected onto the images. C, N, O, and Fe atoms are in black, blue, red, and rusty dots, respectively. The imidazole rings of the distal His64 and proximal His93 are approximately coplanar with X and AC cross sections, respectively. Two propionic acid side chains of AD rings extend to the left in panels b–d, but they are not coplanar with these cross sections. (i–k) Three-dimensional mesh contours superimposed over a structure are traditionally rendered to display electron density maps. Green and red meshes are positive and negative contours at $\pm 7\sigma$, respectively. (l and m) Theoretical contours of probability distributions of electrons in 3d orbitals of iron. The combined probability distributions of $3d_{xz}$ and $3d_{yz}$ with $3d_{xy}$ (l) or with $3d_{x^2-y^2}$ (m) are both in the shape of a cube with the largest probability at the eight corners. A 1/16 cutaway in each contour shows the cross sections. This characteristic shape of the electron distribution is observed in U_4 of the difference maps.

Coefficients c_5 and c_8 show a damped two-dimensional oscillation as marked by the curved arrow. (f) Singular values indicating the significance of the components are sorted in a descending order before a multidimensional rotation.

U_7 of SAOMs at heme site

Figure 5.

Seventh component of SAOMs correlated with photolysis at all temperatures. The coefficient corresponding to U_7 of SAOMs increases as the reaction of ligand dissociation proceeds from MbCO to deoxy Mb in static, cryo-trapping, and time-resolved XFEL experiments (Figure 4e). The hypothetical transition from MbCO to deoxy Mb at zero K is shown in Movie S3. (a–h) Cross sections through the seventh component U_7 . A cubic network of positive densities is marked 1–4 in panel c. (i–k) Three-dimensional meshes of this component in green and red are contoured at $\pm 15\sigma$, respectively. See Figure 3 legend for more detail.

Figure 6.

Thermal vibrations isolated from the chemical signals of photolysis. The third component of the SAOMs U_3 is highly correlated with the sample temperature (Figure 4a,c,e). A linear extrapolation based on the data sets available at cryo and ambient temperatures predicts that the instantaneous local temperature jump of the heme may have exceeded 500 K in the XFEL data sets (i). (a and b) Linear combinations of $900U_1 - 586U_3$ and $900U_1 + 476U_3$ reconstitute the omit maps of the heme at zero and 600 K, respectively, excluding all chemical signals related to photolysis, where the coefficient $c_1 = 900$ of U_1 is nearly

constant for all omit maps. The transition from zero to 600 K free of the chemical signals from photolysis is shown in Movie S2. (c–h) Cross sections through the third component U_3 . Positive densities in 20 pillars connecting two sheets of positive densities above and below the heme in the cross sections of $Z \pm 1.3 \text{ \AA}$ are marked in panel e. (j–l) Three-dimensional meshes of this component in green and red are contoured at $\pm 15\sigma$, respectively. See Figure 3 legend for more detail.

Figure 7.

Low-frequency oscillations in the solvent. (a) SVD analysis of the difference maps in the entire globin produces the first and third components circularly oscillating before 1 ps and quickly diminishing during the first few picoseconds. Similar oscillations are also observed only around the heme site in both the difference maps (Figure 2b) and the omit maps (Figure 4b). (b-e) The average of both the positive and negative halves of U_5 , U_6 , and U_8 of the SAOMs. Low densities in white and light green are located consistently on the heme and protein, which suggests that the oscillating signals are more associated with the solvent

rather than the globin. (f) The circularly oscillating coefficients c_1 and c_3 are plotted as a function of delay time in many series. The offset sinusoidal curves indicate a circular oscillation. Each map series is derived from subtracting a different reference data set as indicated by the color bar below. (g and h) c_1 and c_3 of the $-F_{\text{dark}}$ series are least-squares fitted with a two-dimensional damped oscillation (Methods). (i-l) Comparison of several components of difference maps in the entire globin. U_1 and U_3 contain widespread signals, as they are ranked the most significant components. However, no signal is concentrated on the heme or the protein. Instead, U_2 and U_4 clearly contain strong signals associated with the heme and secondary structures of the protein.