nature | methods

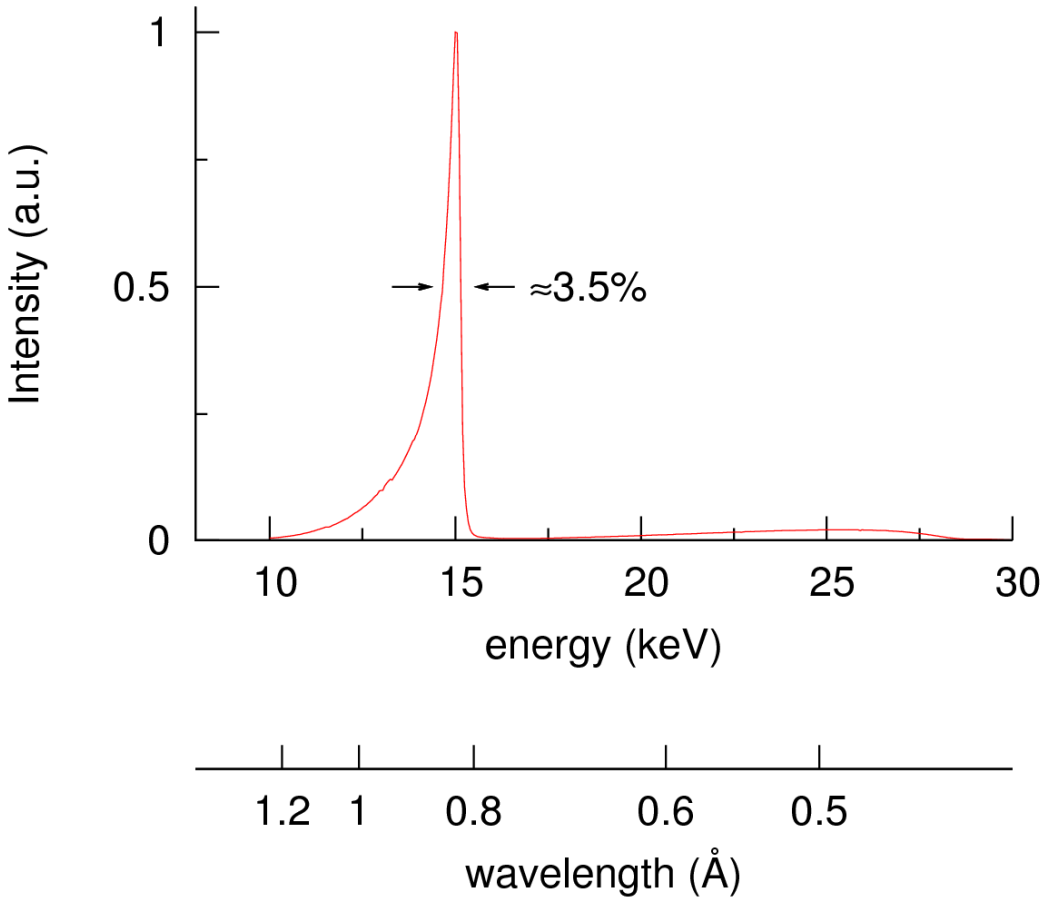
Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering

Marco Cammarata, Matteo Levantino, Friedrich Schotte, Philip A Anfinrud, Friederike Ewald, Jungkweon Choi, Antonio Cupane, Michael Wulff & Hyotcherl Ihee

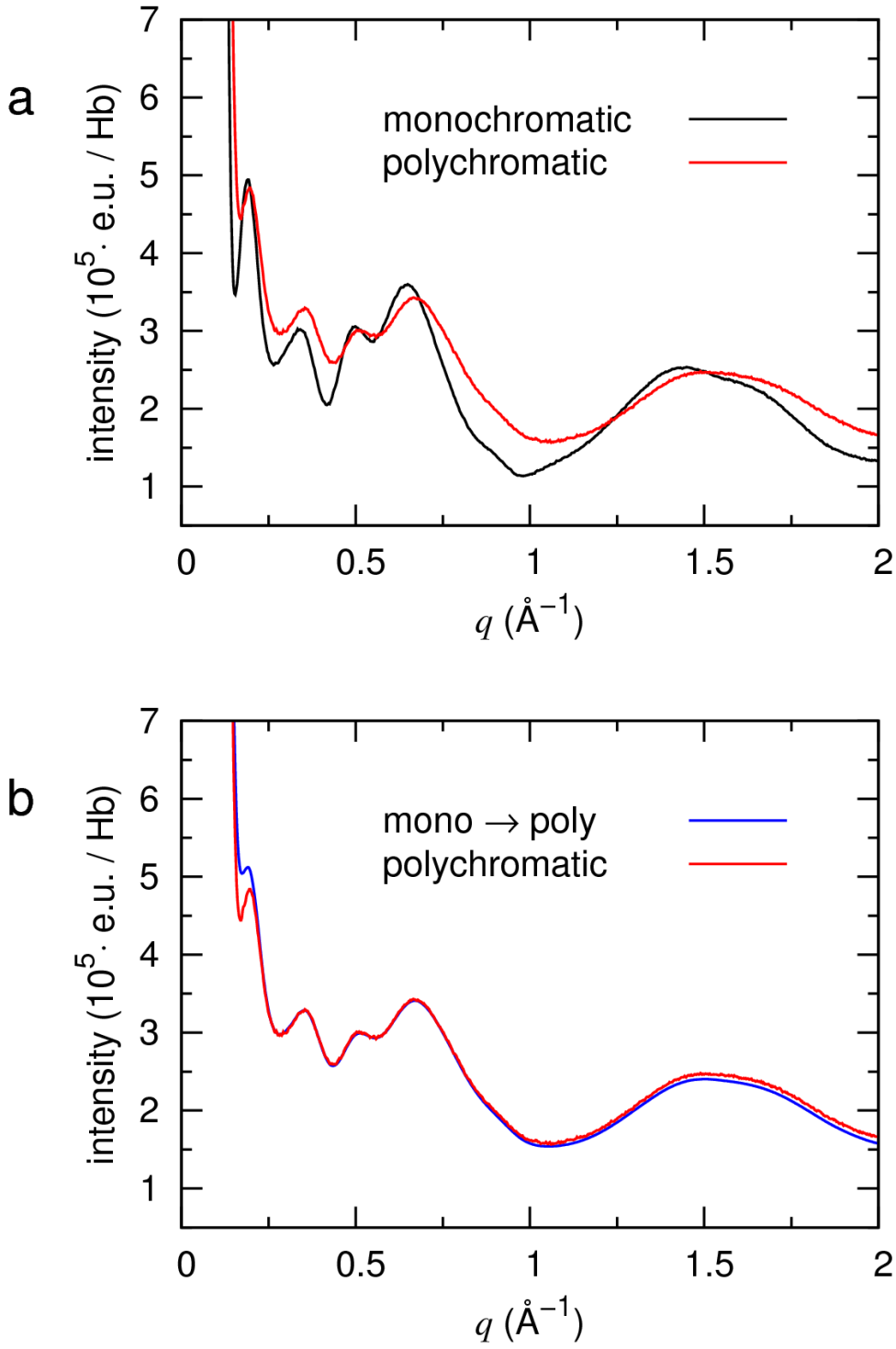
Supplementary figures and text:

Supplementary Figure 1	Allosteric kinetic model used to fit the TR-WAXS data on hemoglobin.
Supplementary Figure 2	Results of the analysis of TR-WAXS data on hemoglobin in terms of the allosteric kinetic model.
Supplementary Figure 3	Results of a "control experiment" on a 2 mM hemoglobin sample.
Supplementary Figure 4	Comparison of experimental scattering differences with those calculated starting from PDB structures.
Supplementary Figure 5	Spectrum of the X-ray beam produced by the U17 undulator operated with a 6 mm gap.
Supplementary Figure 6	Effect of the X-ray beam spectrum on the scattering patterns.
Supplementary Table 1	Parameter definitions for the allosteric kinetic model.
Supplementary Table 2	Fitting parameters resulting from the kinetic analysis of the TR-WAXS data on hemoglobin reported in the main text.
Supplementary Table 3	Fitting parameters resulting from the kinetic analysis of the TR-WAXS data on a 2 mM hemoglobin sample ("control experiment").
Supplementary Table 4	Details on the data gathering protocol.
Supplementary Methods	

Supplementary Figure 1. Spectrum of the X-ray beam produced by the U17 undulator operated with a 6 mm gap.

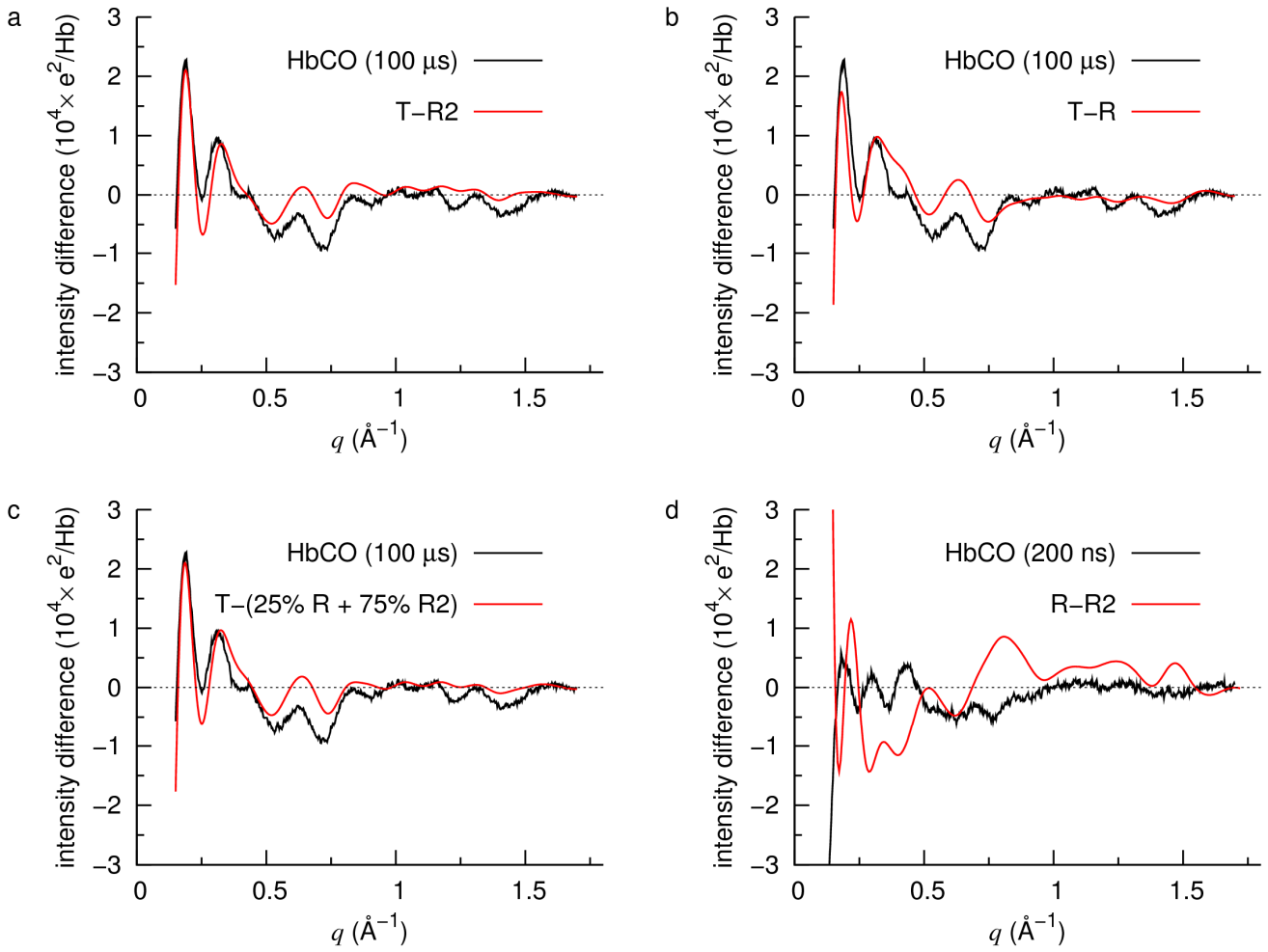


Supplementary Figure 2. Effect of the X-ray beam spectrum on the scattering patterns.



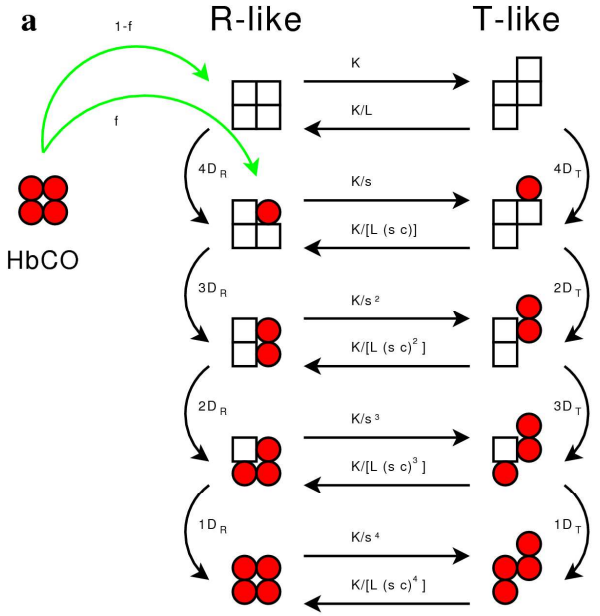
(**a**) Static WAXS HbCO patterns recorded with a monochromatic (black) and polychromatic (red) X-ray beam. The scattered intensity is expressed in units of electrons^2-per-protein-molecule. To highlight the protein contribution to the scattering, the background and buffer contributions were subtracted. (**b**) The blue curve (mono → poly) is the convolution of the monochromatic curve of **panel a** with the X-ray beam spectrum shown in **Supplementary Figure 1**. It reproduces the polychromatic scattering curve (in red) with high fidelity over most of the usable q range.

Supplementary Figure 3. Comparison of experimental scattering differences with those calculated starting from PDB structures.



The scattering difference pattern measured at 100 µs from photolysis of HbCO is compared with the T-R2 (**panel a**) and T-R (**panel b**) differences calculated from crystallographic structures (see **Supplementary Methods**). The better agreement of the 100 µs data with the T-R2 difference is evident. (**c**) Taking a linear combination of these two curves, the best fit is obtained when R2 is about 75% of the total population and R about 25%. (**d**) The scattering difference pattern measured at 200 ns from photolysis of HbCO is compared with the R-R2 difference calculated from crystallographic structures, showing that the experimental data cannot be simply interpreted as the effect of an R-to-R2 transition.

Supplementary Figure 4. Allosteric kinetic model used to fit the TR-WAXS data on hemoglobin.



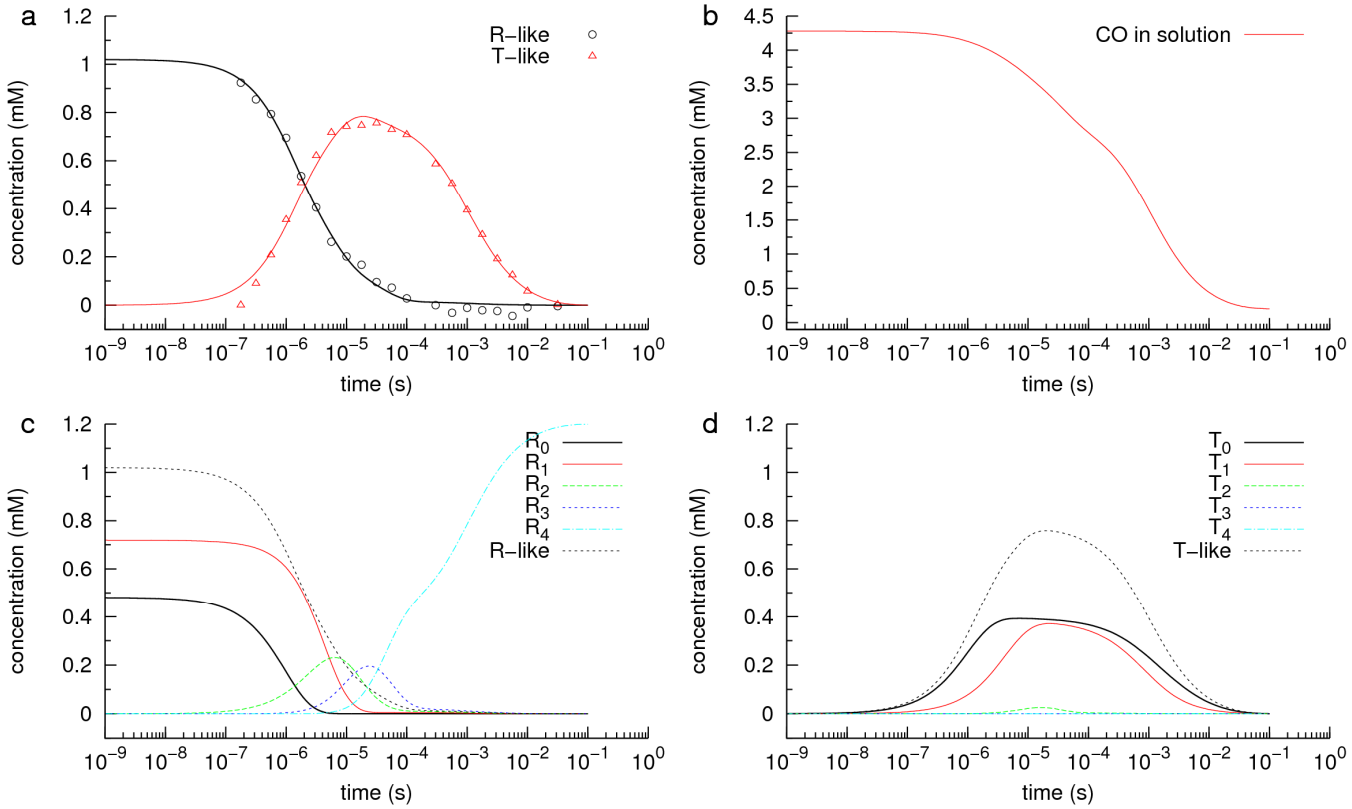
b

$$\partial[CO]/\partial t = -\sum_{i=0}^{3} -i\, D_R[CO][R_i] - \sum_{i=0}^{3} -i\, D_T[CO][T_i]$$

$$\begin{aligned}
\partial[R_0]/\partial t &= -K[R_0] && -4\,D_R[CO][R_0] && +(K/L)\,[T_0] \\
\partial[R_1]/\partial t &= -\left(K/s^1\right)[R_1] && +4\,D_R[CO][R_0] && -3\,D_R[CO][R_1] && +(K/L)\cdot\left(1/(s\,c)^1\right)[T_1] \\
\partial[R_2]/\partial t &= -\left(K/s^2\right)[R_2] && +3\,D_R[CO][R_1] && -2\,D_R[CO][R_2] && +(K/L)\cdot\left(1/(s\,c)^2\right)[T_2] \\
\partial[R_3]/\partial t &= -\left(K/s^3\right)[R_3] && +2\,D_R[CO][R_2] && -1\,D_R[CO][R_3] && +(K/L)\cdot\left(1/(s\,c)^3\right)[T_3] \\
\partial[R_4]/\partial t &= -\left(K/s^4\right)[R_4] && +1\,D_R[CO][R_3] && && +(K/L)\cdot\left(1/(s\,c)^4\right)[T_4] \\
\partial[T_0]/\partial t &= +K[R_0] && && -4\,D_T[CO][T_0] && -(K/L)\,[T_0] \\
\partial[T_1]/\partial t &= +\left(K/s^1\right)[R_1] && +4\,D_T[CO][T_0] && -3\,D_T[CO][T_1] && -(K/L)\cdot\left(1/(s\,c)^1\right)[T_1] \\
\partial[T_2]/\partial t &= +\left(K/s^2\right)[R_2] && +3\,D_T[CO][T_1] && -2\,D_T[CO][T_2] && -(K/L)\cdot\left(1/(s\,c)^2\right)[T_2] \\
\partial[T_3]/\partial t &= +\left(K/s^3\right)[R_3] && +2\,D_T[CO][T_2] && -1\,D_T[CO][T_3] && -(K/L)\cdot\left(1/(s\,c)^3\right)[T_3] \\
\partial[T_4]/\partial t &= +\left(K/s^4\right)[R_4] && +1\,D_T[CO][T_3] && && -(K/L)\cdot\left(1/(s\,c)^4\right)[T_4]
\end{aligned}$$

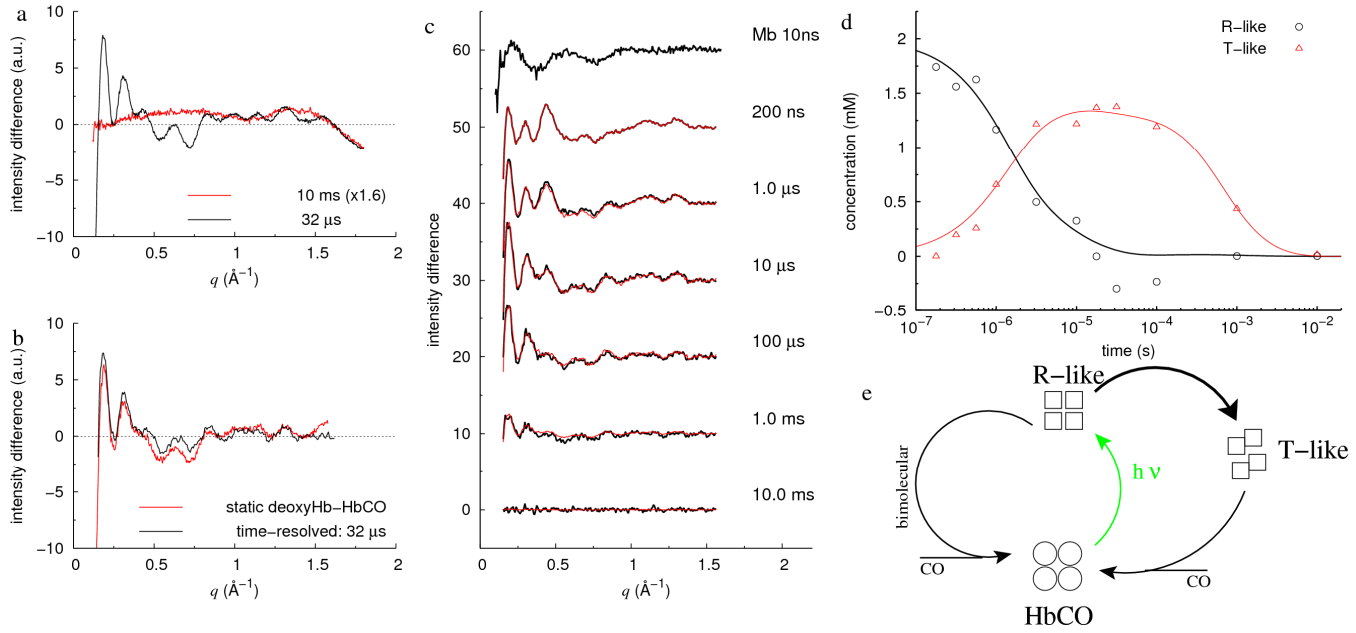
(**a**) Graphical sketch of the allosteric kinetic model taking into account the different species generated after photolysis of HbCO in our experimental conditions. Filled circles represent Hb subunits bound to CO, while empty squares unbound Hb subunits. The relative disposition of subunits within a given species indicates the quaternary structure ("R-like" vs. "T-like"). K is the transition rate for going from R state to T state Hb when no ligand is bound to the protein, L is the allosteric constant. Other symbols are defined in **Supplementary Table 2**. (**b**) Set of equations governing the population dynamics of the species generated in our experiment. $[R_i]$ and $[T_i]$ are the concentrations of Hb molecules in R or T quaternary structure, respectively, and having *i* ligand molecules bound to them. The 150 ns long laser pulse is assumed to generate a mixture of R_0 and R_1 due to multiple excitations of the subunits that undergo fast ("geminate") rebinding. Thus, at time t=0 in our experiment, a fraction *f* of Hb molecules has populated the R_1 state, the remaining fraction 1-*f* has populated the R_0 state, while the other remaining states have not been significantly populated.

Supplementary Figure 5. Results of the analysis of TR-WAXS data on hemoglobin in terms of the allosteric kinetic model.



Time evolution of (**a**) "R-like" and "T-like" populations (see **Supplementary Methods**), (**b**) carbon monoxide concentration in solution, (**c**) population of R state species, (**d**) and population of T state species calculated from the allosteric kinetic model.

Supplementary Figure 6. Results of the "control experiment" on hemoglobin.



(**a**) Subtraction of the solvent response signal. After 10 ms from photolysis, the dissociated CO molecules have recombined to Hb and the scattering difference signal is due essentially to a temperature rise in the solvent (red curve). This curve is used to subtract the solvent response at any time delay. (**b**) The solvent-response-subtracted time-resolved data at 32 µs after photolysis (in black) is compared with the static equilibrium scattering difference (in red) between a deoxyHb sample and an HbCO one. This comparison shows that the data reduction procedure and the subtraction of the solvent response are valid; more importantly, it confirms the structural sensitivity of TR-WAXS to the R-T transition. (**c**) Laser-induced changes of the scattering patterns for several selected time-delays after subtraction of the solvent response (black curves). Data at all time-delays have been fitted as a linear combination of the difference patterns at 200 ns and 32 µs (red curves). The reconstructed curves follow with high fidelity the experimental curves at all time-delays, thus indicating the absence of any relevantly populated structural intermediate between the initial intermediate and the final T structure. In the same panel data on Mb are shown to demonstrate the sensitivity to local tertiary structural changes. (**d**) Time evolution of the "R-like" (circles) and "T-like" (triangles) populations estimated from TR-WAXS data; continuous lines have been obtained by fitting the experimental data in terms of an allosteric kinetic model (see **Supplementary Fig. 4**). (**e**) Schematic representation of hemoglobin structural dynamics between approximately 100 ns and 100 ms from photolysis.

Supplementary Table 1. Details on the data gathering protocol.

	1st step	**2nd step**	**3rd step**
Time delays	-50μs, 200ns, 320ns, 560ns, 1μs, 1.8μs, 3.2μs, 5.6μs, 10μs, 100μs	-50μs, 10μs, 18μs, 32μs, 56μs, 100μs, 300μs	-50μs, 100μs, 300μs, 560μs, 1ms, 1.8ms, 3.2ms, 5.6ms, 10ms, 32ms
Pump-probe pulse pairs	1024	300	100
X-ray pulses per chopper opening	1	13	47
X-ray pulse duration	100 ps	2 μs (thirteen 100 ps pulses separated by 176 ns)	8 μs (forty-seven 100 ps pulses separated by 176 ns)
Repetition rate	10 Hz	10 Hz	5 Hz

Supplementary Table 2. Parameter definitions for the allosteric kinetic model.

Parameter	Description
K	$R_0 \rightarrow T_0$ transition rate
L	Allosteric constant = $[T_0]/[R_0]$
c	Ratio of T state ligand affinity over R state one
s	Scale factor for R→T rates with increasing number of bound ligands
f	$[R_1]/([R_0]+[R_1])$ at t=0
D_R	bimolecular rebinding rate to R
D_T	bimolecular rebinding rate to T
W_R	scale factor for "R-like" population*
W_T	scale factor for "T-like" population*
C_0	Equilibrium CO concentration

*See **Supplementary Methods**

Supplementary Table 3. Fitting parameters resulting from the kinetic analysis of the TR-WAXS on hemoglobin reported in the main text.

Parameter	Starting value	Best fit value	Range of values
1/K	2.5 μs	1.19 μs	1.12–1.3 μs
f	0.25	0.6	0.57–0.63
s	8	7.2	6.6–7.7
c	4×10^{-3}	--	--
L	10^4	--	--
D_R	1.1×10^4 mM^{-1} s^{-1}	--	--
D_T	100 mM^{-1} s^{-1}	60 mM^{-1} s^{-1}	50–70 mM^{-1} s^{-1}
W_R	0.9	0.77	0.76–0.78
W_T	0.5	0.49	0.48–0.50
C_0	0.2 mM	--	--
[Hb]	1.2 mM	--	--

The double dash "--" indicates that the parameter has been kept fixed during the fitting procedure. Values reported in the "Range of values" column correspond to ± one standard deviation.

Supplementary Table 4. Fitting parameters resulting from the kinetic analysis of the TR-WAXS on hemoglobin at high concentration ("control experiment").

Parameter	Starting value	Best fit value	Range of values
1/K	2.5 μs	2.17 μs	2.11–2.22 μs
f	0	--	--
s	8	5.8	4.9–6.6
c	4×10^{-3}	--	--
L	10^4	--	--
D_R	1.1×10^4 mM^{-1} s^{-1}	--	--
D_T	130 mM^{-1} s^{-1}	110 mM^{-1} s^{-1}	110–120 mM^{-1} s^{-1}
W_R	0.9	0.87	0.86–0.88
W_T	0.5	0.68	0.67–0.7
C_0	1.0 mM	--	--
[Hb]	2.0 mM	--	--

The double dash "--" indicates that the parameter has been kept fixed during the fitting procedure. Values reported in the "Range of values" column correspond to ± one standard deviation.

Supplementary Methods

X-ray source and mechanical isolation of X-ray pulses

The European Synchrotron Radiation Facility (ESRF) operates its electron storage ring at 6 GeV. With this electron energy, a U17 in-vacuum undulator (17 mm period; SmCo magnets) set to 6 mm gap produces an X-ray spectrum that is sharply peaked at 15.0 keV, but has a long wavelength tail (**Supplementary Fig. 1**). The spectrum is quasi-monoenergetic with a relative bandwidth of 3.5%. In order to reduce the acquisition time, we took advantage of the full X-ray spectrum which is 250 times more intense compared to a standard monochromatic beam. Note that the loss of information in relaxing the bandwidth to 3.5% is minor due to the orientational disorder of molecules in solution (see **Supplementary Fig. 2**).

A Pt-coated toroidal mirror, set at the incidence angle 2.668 mrad, focuses the radiation onto the sample. To minimize distortion of the toroidal mirror from the heat load from the undulator (~100 W), a synchronous heat-load chopper was positioned upstream of the mirror; it reduces the average heat load on the mirror and high-speed chopper by a factor 20 without affecting the pulses used in the pump probe experiment. When operated in 16-bunch mode, the ID09B beamline optics focuses $\approx 1.0\times10^{9}$ X-ray photons in a single pulse down to an elliptical spot as small as 0.06×0.1 mm^2 (vertical×horizontal, FWHM) at the sample position.

Quasi-monochromatic X-ray pulses (100 ps long) can be extracted from the high frequency pulse train when the storage ring is filled with either 4 or 16 electron bunches by using a high speed synchronized rotor spinning at ~1 kHz (360th sub-harmonic of the time needed for a single electron bunch to complete the storage ring orbit).

To be able to measure time-delays longer than 1 ms, and to allow our sample positioning system to move the sample into a fresh position at every shot, the repetition rate is further reduced by using a "millisecond" shutter. This lowers the repetition rate from 1 kHz down to 0 to 20 Hz.

Laser photolysis

The protein sample was photolyzed with a circularly-polarized, 527 nm, 0.5 mJ laser pulse (DM-50, Laser Photonics). At this excitation wavelength, half the incident flux is absorbed after penetrating 0.09 mm into the 1.2 mM HbCO protein solution. To maximize the overlap between the pump- and probe-illuminated volumes, orthogonal pump-probe geometry was employed. The X-ray beam penetrated the 1 mm diameter capillary 0.22 mm below its top edge, its path length through the capillary was 0.83 mm. The photolysis beam was focused with cylindrical lenses to an elliptical spot approximately 1.0×0.22 mm^2 (FWHM) with the long axis oriented parallel to the X-ray beam path. Laser power density on the surface of the capillary was ~2.5 mJ mm^{-2}. The sample was maintained at 15 ºC with a cold nitrogen stream (Oxford Cryostream). To dilute any X-ray radiation damage over a large sample volume, the sample was translated back and forth along its long axis over a 20 mm range. To ensure that successive pulses in the 5-10 Hz pulse train excite adjacent but spatially-separated sample volumes in the capillary, it was translated by 0.2 mm after each probe pulse.

Time-resolved data collection protocol

Though the pink-beam flux generated by the U17 in-vacuum undulator is very intense, approximately 10^3 pump-probe cycles are required to generate a high-dynamic range scattering image. When repeated at 5 or 10 Hz, the integration time is less than 2 minutes per image. "laser off" images were also acquired with laser pulses arriving 50 μs after the X-ray pulse in order to probe the ground state while assuring the same "average temperature" of the solution. These "laser off" images were used to compute the TR-WAXS differences. Usually a *laser off* image was collected after every 4-5 *laser on* images to compensate for slow drifts in the beamline. To improve the signal-to-noise ratio, up to 30 images were acquired at each time delay. The data collection was divided in three parts as detailed in **Supplementary Table 1**.

Comparison of monochromatic and polychromatic static WAXS patterns

The ~250-fold reduction in flux that accompanies the insertion of a monochromator into the

polychromatic beam would require, in the case of time-resolved experiments, integration times that are incompatible with the limited beam time available, or with the limited lifetime of biological samples. Consequently, the TR-WAXS data were acquired with the full polychromatic beam (**Supplementary Fig. 1**). To assess the effect of the polychromatic beam on the WAXS patterns, static scattering data were collected for several different samples with both monochromatic and polychromatic X-ray beams. Although the polychromatic spectrum smears out the peaks in the HbCO WAXS spectrum (**Supplementary Fig. 2a**), the effect is small and can be reproduced by convolving the monochromatic WAXS spectrum with the X-ray beam spectrum (**Supplementary Fig. 2b**). The convolution of the monochromatic curve with the X-ray beam spectrum reproduces the polychromatic scattering curve with high fidelity over most of the usable q-range. The discrepancies near the low and high q limits are a consequence of truncation effects at the boundaries of the q-range explored in monochromatic mode.

Estimation of scattering from proteins

Solution scattering patterns for given protein solutions were calculated from crystallographic three-dimensional atomic coordinates deposited in the Protein Data Bank (PDB) using the software CRYSOL (version 2.6, http://www.embl-hamburg.de/ExternalInfo/Research/Sax/crysol.html). All calculations have been performed using 50 spherical harmonics, an average solvent electron density of 0.34 electrons $Å^{-3}$, a shell contrast of 0.015 electrons $Å^{-3}$, and a Fibonacci grid order equal to 18. No attempt was made to fit the calculated patterns to the experimental curves.

Simulation of clamshell motion

The clamshell motion is a tertiary structure modification that allows the escape of CO from the protein, the doming of the heme, the relaxation of the proximal histidine, and the rotation/translation of the so-called "E-F clamshell" (Guallar et al., *J. Am. Chem. Soc.*, 128, 5427-5435, 2006; Kachalova et al., *Science*, 284, 473-476, 1999). In order to simulate the effect of the clamshell motion on the heme pockets of human Hb, we have first aligned one subunit of the crystallographic T structure (PDB code: 2HHB) onto the corresponding subunit of the R2 crystallographic structure (PDB code: 1BBB) with respect to the A, B, D (only for beta chains) and H helices. The E and F helices, the FG corner and the heme of the T subunit have been replaced with those of the R2 corresponding subunit. The whole procedure has been repeated for all the four subunits of the protein.

Information content of the scattering difference curves

We have estimated the information content of our scattering difference curves using two different approaches:

1) we have estimated the number of Shannon channels (Svergun et al., *Acta Cryst.*, A52, 419-426, 1996) of our scattering patterns;
2) we have fitted our data either with a sum of Gaussian functions or with a polynomial function.

The number of Shannon channels is defined as $N_{SC}=D_{max}(q_{max}-q_{min})/\pi$, where D_{max} is the maximum particle size, while q_{max} and q_{min} are the maximum and minimum scattering vector respectively. Plugging the numbers (D_{max} ~ 66 Å for Hb) we get N_{SC}=35. On the other hand, approximately 14 Gaussians (i.e. 42 free parameters), or a polynomial function of order 30-35, are required to reproduce the 100 µs experimental curve (after subtracting the heating contribution). Although the above estimations are not identical, they indicate that approximately 30-40 free parameters are needed to reproduce the information content of the scattering difference curves.

Calculation of the expected signal-to-noise ratio

The detection noise in scattering experiments is typically dominated by photon noise. In the hypothesis that fluctuations in the signal measured at each pixel follow the Poisson statistics, the uncertainty, $\delta I(q)$, on the measured intensity averaged over the annular region of the CCD detector at the scattering vector q can be estimated as follows:

$$\delta I(q) = \frac{\sqrt{I_{tot}(q)}}{N_{pixel}(q)} \tag{S1}$$

where $N_{pixel}(q)$ is the number of pixels contained in the annular region and $I_{tot}(q)$ is the total scattered intensity measured in the region:

$$I_{tot}(q) = \varepsilon\, N_{pixel}(q) \frac{dN(q)}{d\Omega} \delta\Omega(q) \tag{S2}$$

where ε is the detector efficiency for the X-ray radiation at the given wavelength, $dN(q)$ is the total number of X-ray photons scattered in the solid angle $d\Omega$, and $\delta\Omega(q)$ is the solid angle corresponding to a single pixel at the scattering vector q.

The above quantities can be further expressed in terms of physical parameters characterizing the incident X-ray beam, the sample, and the CCD detector:

$$N_{pixel}(q) = \frac{2\pi}{\delta} d \tan(\theta) \tag{S3}$$

$$\frac{dN(q)}{d\Omega} = \frac{N_0}{A} r_e^2 N_A \sum A\, c_i t_i S_i(q) \tag{S4}$$

$$\delta\Omega = \frac{\delta^2}{[d\cos(\theta)]^2} \frac{1}{\cos(\theta)} \tag{S5}$$

where:

- δ is the detector pixel size;
- d is the sample-to-detector distance;
- θ is the scattering angle corresponding to the scattering vector q;
- N_0 is the number of photons impinging on the sample;
- A is the size of the X-ray beam at the sample;
- r_e is the classical electron radius (=2.81710×10^{-15} m);
- N_A is the Avogadro number (=6.022×10^{23});
- c_i is the concentration (in moles per cubic meter) of the molecules of kind i (e.g. protein, water, and silica of the capillary);
- t_i is the thickness of the volume occupied by molecules of kind i;
- $S_i(q)$ is the intensity scattered by a molecule of kind i expressed in electron^2/molecule.

The standard deviations of the data reported in **Figure 1b** of the paper have been calculated using typical values of the physical parameters in our experimental conditions (ε at 15 keV is approximately 60%, δ = 0.06 mm, d = 0.25 m, A cancels out in Eq. S4, the protein concentration is ~1 mM while the water concentration ~52 M) and the $S_i(q)$ derived from three-dimensional atomic coordinates. The background contribution, $c_{bkg}t_{bkg}$, which arise from X-ray photons scattered mainly by air and the capillary surfaces, have been estimated from the experimental data. **Figure 1b** reports the results of calculations performed for three different values of N_0 (the line thickness is equal to the standard deviation on the signal).

Interpretation of the Hb population kinetic

Although, as discussed in the main text, more experiments will be needed to pin down the structural dynamics of Hb, we have used an allosteric kinetic model, similar to that of Sawicki and Gibson (*J. Biol.*

Chem., 251, 1533-1542, 1976), to fit the population kinetics of the different Hb binding/structural states.

Essentially 10 different configurations are possible R_i {i=0,4} and T_i {i=0,4} where the subscript indicates the number of bound CO ligands. A schematic representation of the model and the corresponding set of equations are reported in **Supplementary Figure 4**. The model depends on the parameters shown in **Supplementary Table 2**.

The geminate rebinding has been indirectly taken into account by having also the R_1 state (apart from the R_0 one) populated at time t=0. The parameters *c* and *L* (that control the $T \leftrightarrow R$ equilibrium) have been taken from Goldbeck et al. (*Biochemistry*, 35, 8628-8639, 1996).
In particular (see **Supplementary Table 3**) the initial conditions at $t = 0$ were set to:

$$\begin{aligned} [R_0] &= [Hb]\cdot(1-f) \\ [R_1] &= [Hb]\cdot(f) \\ [R_{i=2,4}] &= 0 \\ [T_{i=0,4}] &= 0 \\ [CO] &= C_0 + 4\cdot[R_0] + 3\cdot[R_1] \end{aligned} \tag{S6}$$

The system of equations has been numerically integrated using the Hindmarsh's ordinary differential equation solver *LSODE* using the stiff backward differentiation formula method. Solving the system of differential equations, we obtain (for a given set of reaction parameters) the population of the different species as a function of time.

From these populations, the "R-like" and "T-like" behavior were constructed using the following relationships:

$$\begin{aligned} \text{[R-like]}(t) &= [R_0](t) + 3/4\cdot[R_1](t) + 2/4\cdot[R_2](t) + 1/4\cdot[R_3](t) \\ \text{[T-like]}(t) &= [T_0](t) + [T_1](t) + [T_2](t) + [T_3](t) \end{aligned} \tag{S7}$$

where $[R_i](t)$ and $[T_i](t)$ denote the concentration of Hb molecules with *i* CO molecules bound either to R state or T state Hb and are the solutions of the set of differential equations discussed above. Note that tertiary differences are confined within individual subunits and therefore their weight is expected to scale according to the number of unbound subunits.

The difference curve at each time delay, $S(t_i)$, has been written as linear combination of the 178 ns and the 100 μs time delay signals, i.e.

$$S(t_i) = W_{178\,ns}(t_i)\cdot S(178\,\text{ns}) + W_{100\,\mu s}(t_i)\cdot S(100\,\mu s) \tag{S8}$$

Using a least-squares fit we obtained the two weight vectors $W_{178ns}(t_i)$ and $W_{100\mu s}(t_i)$. The comparison between the model and the data has been done using the fitting package *Minuit* (ver. 1.7) minimizing the χ^2 defined as follows:

$$\chi^2 = \sum_{t_i}\left[\frac{W_R\cdot W_{178\text{ns}}(t_i) - \text{[R-like]}(t_i)}{\sigma_i}\right]^2 + \sum_{t_i}\left[\frac{W_T\cdot W_{100\mu s}(t_i) - \text{[T-like]}(t_i)}{\sigma_i}\right]^2 \tag{S9}$$

where W_R and W_T are scale factors needed to convert the weight vectors $W_{178ns}(t_i)$ and $W_{100\mu s}(t_i)$ into "R-like" or "T-like" populations, respectively. The parameters obtained after the minimization procedure are shown in **Supplementary Table 2**. The corresponding populations are shown in the **Supplementary Figure 5**.

The timescale of the R to T transition is found to be 1.2 μs. Using different models or different fitting procedures (for example leaving all the parameters unconstrained) values up to 3 μs were found. A fit of the data imposing an R-T transition timescale equal to the "classical" 20 μs value, while leaving all the other parameters free to vary, resulted in a very poor fit and unphysical best fit values for some of the fitting parameters (for example the bimolecular rebinding to T was found to be faster than that to R).

Control experiment on hemoglobin

Sample preparation: The human Hb used in the TR-WAXS study was kindly prepared by Jayashree Soman and John Olson using standard procedures. The final Hb concentration was 8 mM on a heme basis in 20 mM Tris buffer at pH 8.9.

Protein laser photolysis: The energy per laser pulse was about 0.8 mJ on the sample and the focus spot was 650 μm × 150 μm (8.2 mJ mm^{-2}) to maximize the overlap with the X-ray. Increasing the pump energy to 1.3 mJ per pulse did not increase the signal significantly, indicating that the photolysis extent was close to 100%. Note that, since the laser pulse duration is longer than the characteristic rotational time of the Hb molecule, no photoselection effect is expected. To maximize the overlap between the pump- and probe-illuminated volumes, the orthogonal pump-probe geometry was employed. The laser illuminates a 1 mm diameter X-ray capillary from the top (penetration depth ~100 μm), while the X-ray beam was slicing the capillary 0.1 mm below its top edge.

Experimental details: the capillary was surrounded by a temperature-stabilized nitrogen stream (15 ºC, Oxford Cryostream). To dilute X-ray radiation damage over a larger sample volume, the sample was translated back and forth along its long axis over a 20 mm range. To ensure that successive pulses in the 5 Hz pulse train excite adjacent but spatially-separated sample volumes in the capillary, it was translated by 0.2 mm after each probe X-ray pulse. The laser to X-ray time delay was defined as zero when the X-ray pulse was in the center of mass of the laser pulse in time.

Results: The main results of the control experiment on Hb are summarized in **Supplementary Figure 6** where the two-component decomposition and the fit with the allosteric kinetic model presented above are shown. **Supplementary Table 4** shows the best fit parameters obtained from the analysis of the control experiment. Note that due to the higher laser power used in this "control experiment", the fraction f of R_1 populated at t=0 is much smaller than in the case of the experiment reported in the main text. Moreover, the signal-to-noise ratio of the data is worse in this case. Since it was not possible to estimate reliably f, for the sake of simplicity, we have assumed f=0 in the analysis of the data in terms of the allosteric kinetic model. This is justified by the high laser fluence used in the experiment.