

## BIOPHYSICS

# Determination of functional collective motions in a protein at atomic resolution using coherent neutron scattering

Liang Hong,<sup>1\*</sup> Nitin Jain,<sup>2</sup> Xiaolin Cheng,<sup>3,4</sup> Ana Bernal,<sup>2</sup> Madhusudan Tyagi,<sup>5,6</sup> Jeremy C. Smith<sup>3,4\*</sup>

2016 © The Authors,  
some rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

Protein function often depends on global, collective internal motions. However, the simultaneous quantitative experimental determination of the forms, amplitudes, and time scales of these motions has remained elusive. We demonstrate that a complete description of these large-scale dynamic modes can be obtained using coherent neutron-scattering experiments on perdeuterated samples. With this approach, a microscopic relationship between the structure, dynamics, and function in a protein, cytochrome P450cam, is established. The approach developed here should be of general applicability to protein systems.

## INTRODUCTION

Large-amplitude, collective atomic motions in proteins play a crucial role in many functional processes, including the entry of substrates into catalytic sites (1), allosteric conformational change (2), and enzymatic reactions (3). These motions have been extensively investigated using molecular dynamics (MD) simulations (4, 5). However, the corresponding quantitative experimental characterization of their amplitudes, forms, and time dependences remains challenging.

Complete, atomic-detail information on collective internal protein motions requires the simultaneous direct determination of time and space correlations in the interatomic displacements. Generally speaking, diffraction techniques provide no information on time correlations, whereas techniques deriving motional spectra lack the spatial component (6). However, in principle, both can be derived from inelastic coherent neutron scattering (7). Although inelastic coherent neutron scattering has been widely applied to characterize collective atomic motions in soft matter systems, including small-molecule liquids (8), polymers (9), and glasses (10), its application to protein dynamics has been rare (7, 11, 12).

Recently, neutron spin echo (NSE) spectroscopy has been used to measure inelastic coherent neutron-scattering signals from protein motions (11, 12). Although this technique has been shown to successfully detect collective internal protein motion in aqueous solution, the global rotation and translation of the macromolecule dominate the neutron signal measured and render the quantitative characterization of the protein internal dynamics difficult (11, 13). To overcome this problem, we performed high-flux backscattering experiments on perdeuterated protein powders, where the global motion of the protein molecules is shown to contribute negligibly to the neutron signals measured. The experiments reveal that in-phase collective dynamic modes, that is, protein residues that move synchronously in time with similar amplitudes, extend through several nanometers across the protein molecule on a time scale of

~100 ps, and the findings are consistent with corresponding MD simulations. Moreover, it is found that the wave vector dependence of the fluctuation rate of interatomic motion measured experimentally can be used to derive the most probable forms of the large-scale protein motions at atomic resolution while also furnishing the explicit amplitudes of the motions sampled on the resolution time of the instrument. Hence, coherent neutron-scattering signals can be used to characterize large-scale functional modes in a protein.

## RESULTS

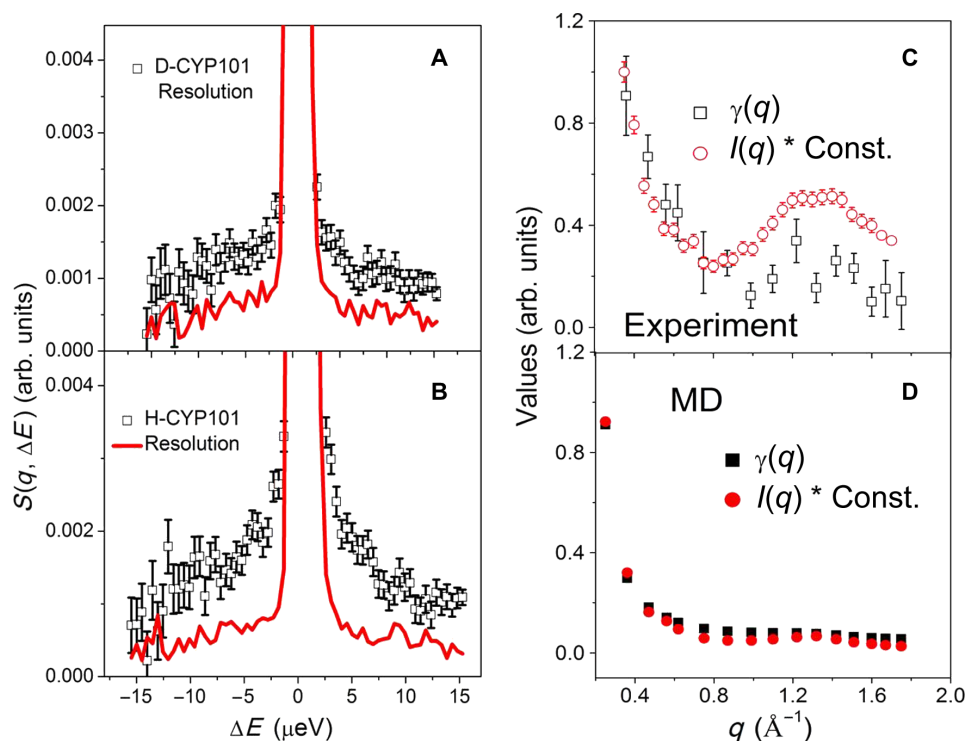
### Neutron scattering

The model protein studied is cytochrome P450cam (CYP101). P450s are an important enzyme family that catalyzes a variety of biochemical reactions involved in carcinogenesis, drug metabolism, lipid and steroid biosynthesis, and degradation of pollutants in higher organisms (14). CYP101 from *Pseudomonas putida* has long served as a model system for studying the dynamic characteristics of substrate binding in P450s (15). Here, neutron-scattering experiments were conducted on both camphor-bound CYP101 and its perdeuterated counterpart using the high-flux backscattering spectrometer (HFBS) at the National Institute of Standards and Technology (NIST). For simplicity, these two samples are referred to as H-CYP101 and D-CYP101. Details of the experimental conditions and MD protocols are provided in Materials and Methods. The experimentally measured quantity is the dynamic structure factor (eqs. S4 and S5),  $S(q, \Delta E)$ , which is the time Fourier transform of the intermediate scattering function (eqs. S1 and S2),  $I(q, t)$ , and furnishes the distribution of the dynamic modes in the sample over frequency at a given scattering wave vector,  $q$ . Typical spectra of  $S(q, \Delta E)$  measured at 285 K for D-CYP101 and H-CYP101 are presented in Fig. 1 (A and B).  $S(q, \Delta E)$  contains incoherent and coherent components:  $S_{\text{inc}}(q, \Delta E)$  and  $S_{\text{coh}}(q, \Delta E)$ .  $S_{\text{inc}}(q, \Delta E)$  characterizes self-correlations in atomic motions, whereas  $S_{\text{coh}}(q, \Delta E)$  probes mostly cross-correlations, that is, interatomic fluctuations. As we performed neutron-scattering experiments on both D-CYP101 and H-CYP101, the contributions of coherent and incoherent scattering to the neutron signals measured on these two samples can be explicitly separated; the procedures followed are presented in the Supplementary Materials.

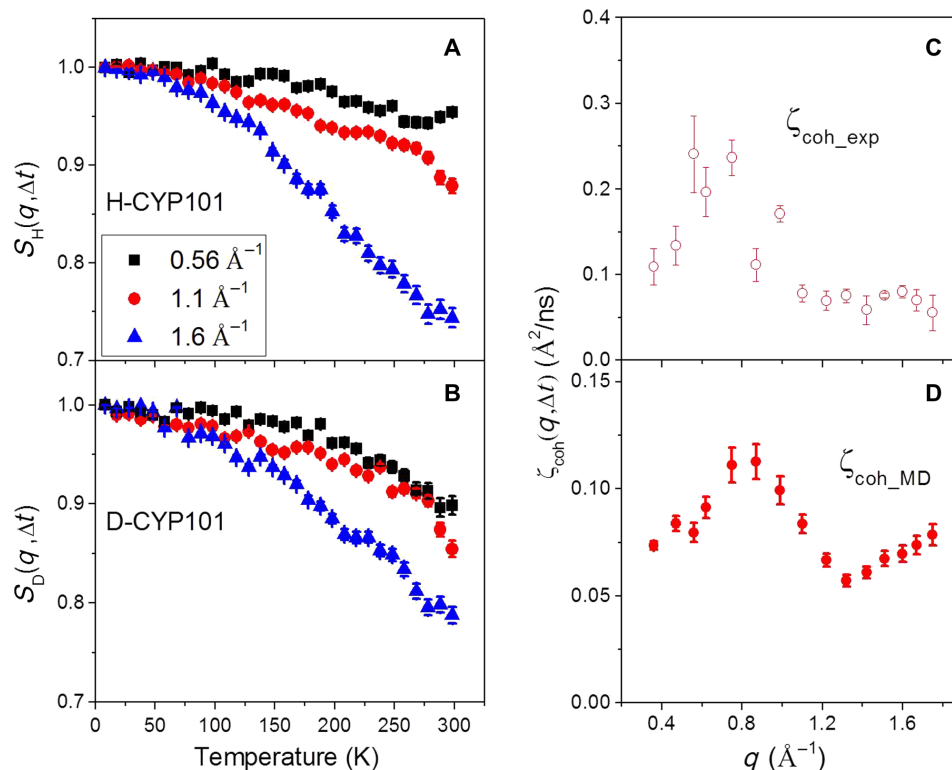
A system undergoing in-phase collective motions, that is, constituting units moving synchronously in time with similar amplitudes,

<sup>1</sup>Institute of Natural Sciences and Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China. <sup>2</sup>Department of Biochemistry & Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA. <sup>3</sup>Center for Molecular Biophysics, Oak Ridge National Laboratory, TN 37831, USA. <sup>4</sup>Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA. <sup>5</sup>NIST Center for Neutron Research, National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA. <sup>6</sup>Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742, USA.

\*Corresponding author. Email: smithjc@ornl.gov (J.C.S.); hongl3liang@sjtu.edu.cn (L.H.)



**Fig. 1. Quasi-elastic neutron spectra of CYP101.** Dynamic structure factor,  $S(q, \Delta E)$ , averaged over the  $q$  windows from 0.35 to 1.75  $\text{\AA}^{-1}$ , measured on (A) D-CYP101 and (B) H-CYP101 at 285 K. The red curves are the resolution function measured on vanadium. The raw data have been smoothed by averaging five neighboring data points. Comparison of the  $q$  dependence of  $\gamma(q)$  with that of  $I(q)$  obtained (C) experimentally and (D) from simulation.  $\gamma(q)$  is the ratio of the coherent to incoherent  $S(q, \Delta E)$  intensity with both integrated over the energy window from 5 to 10  $\mu\text{eV}$  at a given  $q$  at 285 K. Error bars throughout the text represent 1 standard deviation (SD). arb. units, arbitrary units; Const., constant.



**Fig. 2. Intensity of the elastic peak of the dynamic structure factor and the derived decay of the intermediate scattering function.** Experimental temperature dependence of  $S(q, \Delta t)$  for (A) H-CYP101 and (B) D-CYP101 at three different  $q$  values: 0.56, 1.1, and 1.6  $\text{\AA}^{-1}$ , normalized by the values at 4 K. The data points have been averaged in every 10-K window. The decay of the coherent intermediate scattering function,  $\zeta_{\text{coh}}(q, \Delta t)$ , derived (C) experimentally and (D) from simulation at 285 K.

will exhibit  $S_{\text{coh}}(q, \Delta E) \propto I(q)S_{\text{inc}}(q, \Delta E)$ , whereas conversely  $S_{\text{coh}}(q, \Delta E) \propto S_{\text{inc}}(q, \Delta E)$  is found if the motion is completely random and uncorrelated (7, 9, 16). Here,  $I(q)$  is the overall static structure factor, characterizing the atomic structure of the system (eq. S3).

Figure 1C presents  $\gamma(q)$ , the ratio between the integral of  $S_{\text{coh}}(q, \Delta E)$  in the energy window from 5 to 10  $\mu\text{eV}$ , which corresponds to  $\sim 100$  to 200 ps, and the corresponding integral for  $S_{\text{inc}}(q, \Delta E)$  at 285 K. As seen in Fig. 1C, the  $q$  dependence of  $\gamma(q)$  resembles that of  $I(q)$ .

To verify the above result, all-atom MD simulation was performed on CYP101 in the same conditions as the experiment, and the same analysis as above was performed using the MD-derived scattering functions (Fig. 1D). Again, the MD-derived  $\gamma(q)$  resembles the  $q$  dependence of  $I(q)$ . Hence, as shown by both experiment and simulation (Fig. 1, C and D),  $\gamma(q)$  exhibits a  $q$  dependence similar to  $I(q)$  down to  $0.35 \text{ \AA}^{-1}$  ( $\sim 2 \text{ nm}$ ). This demonstrates the existence of collective in-phase atomic motions across several nanometers in the protein on the corresponding time scale,  $\sim 100$  to 200 ps. A closer examination of the experimental data (Fig. 1C) reveals that  $\gamma(q)$  resembles  $I(q)$  more closely in the low- $q$  region ( $< 0.8 \text{ \AA}^{-1}$ ) than at high  $q$ . This arises from the fact that large-scale collective modes spanning a large portion of the protein molecule dominate the low- $q$  neutron signal, whereas local modes, which are less correlated, smear the behavior at high  $q$  (13).

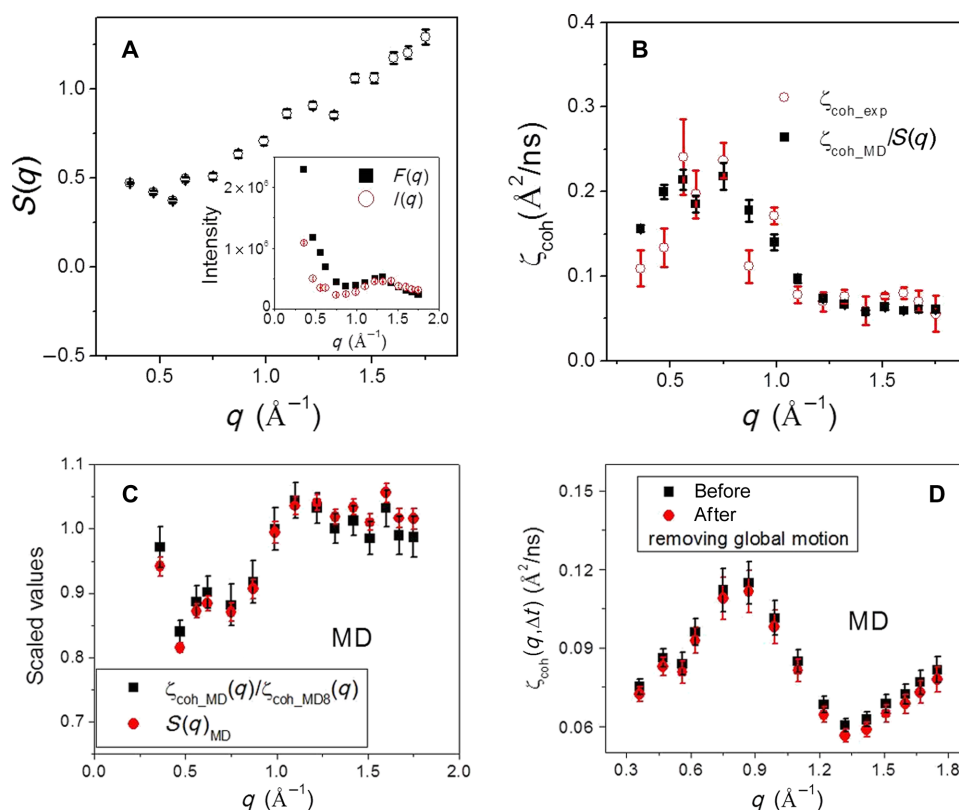
Earlier inelastic coherent neutron-scattering work on glasses, fully deuterated green fluorescent protein, and polymers using

time-of-flight neutron spectroscopy demonstrated the collective, in-phase nature of vibrational modes on subpicosecond time scales in these systems (7, 9, 10). The present study reveals that the relaxation processes in protein molecules on much longer time scales, that is, 100 to 200 ps, can also be collective, in-phase and can span nanometer length scales.

To quantitatively characterize the spatial and temporal features of the interatomic motion in CYP101, the decay of the coherent intermediate scattering function at instrument resolution was derived as

$$\zeta_{\text{coh}}(q, \Delta t) = -\frac{\ln(I_{\text{coh}}(q, \Delta t)/I(q))}{q^2 \Delta t} \approx -\frac{\ln S_{\text{coh}}(q, \Delta t)}{q^2 \Delta t} \quad (1)$$

The experimental  $\zeta_{\text{coh}}(q, \Delta t)$  was calculated using the intensity of the elastic peak of the coherent dynamic structure factor,  $S_{\text{coh}}(q, \Delta t)$ , which is derived from the elastic intensity measured on both D-CYP101 and H-CYP101 (see eq. S8 and Fig. 2, A and B) at 285 K. The simulation-derived  $\zeta_{\text{coh}}(q, \Delta t)$  was estimated using the coherent intermediate scattering function,  $I_{\text{coh}}(q, \Delta t)$ , computed directly from the MD trajectories, and  $\Delta t$  is the instrument resolution  $\sim 1 \text{ ns}$ .  $\zeta_{\text{coh}}(q, \Delta t)$  thus defined furnishes a quantitative measure of how much the intermediate scattering function decays at  $\Delta t$  and has the dimensions of a diffusion coefficient.  $\zeta_{\text{coh}}(q, \Delta t)$  characterizes



**Fig. 3. Interprotein structure factor and its correction to obtain single-protein dynamics.** (A) Experimental interprotein structure factor,  $S(q) = I(q)/F(q)$ . To obtain  $S(q)$  on an absolute scale,  $I(q)$  was multiplied by a constant to match the values of  $F(q)$  at  $q > 1 \text{ \AA}^{-1}$  (inset), because the interprotein scattering at high  $q$  is much smaller than at low  $q$  (32). (B) Comparison of  $\zeta_{\text{coh\_exp}}(q, \Delta t)$  with  $\zeta_{\text{coh\_MD}}(q, \Delta t)$  corrected by experimental  $S(q)$ . (C) Comparison of MD-derived  $S(q)_{\text{MD}}$  with the ratio between the single-protein  $\zeta_{\text{coh\_MD}}(q)$  and the overall value,  $\zeta_{\text{coh\_MD8}}(q)$ .  $\zeta_{\text{coh\_MD8}}(q)$  is the average over the values calculated from each of the eight protein molecules in the simulation box, whereas  $\zeta_{\text{coh\_MD8}}(q)$  is estimated by taking the eight protein molecules as a whole solution.  $S(q)_{\text{MD}}$  arises from the interprotein packing in the MD box. (D) Comparison of  $\zeta_{\text{coh\_MD}}(q, \Delta t)$  before and after removing the global motion of the protein molecule. Removal of the global motion was performed by fitting the MD trajectory of each protein molecule to its structure in the first frame.

the fluctuation rate of the interatomic motion at a given distance ( $\sim 2\pi/q$ ) on the experimental time scale and thus contains the characteristic spatiotemporal information on the prevalent collective dynamic modes in the protein (17). Experimentally,  $\zeta_{\text{coh}}(q, \Delta t)$  exhibits a prominent peak at  $\sim 0.7 \text{ \AA}^{-1}$  ( $\sim 9 \text{ \AA}$ ) (Fig. 2C). The corresponding MD-derived  $\zeta_{\text{coh}}(q, \Delta t)$  qualitatively resembles the experiment, with a peak at  $\sim 0.8 \text{ \AA}^{-1}$  and a minimum at  $\sim 1.3 \text{ \AA}^{-1}$  (Fig. 2D), indicating that the interatomic motion reaches maximum and minimum fluctuation rates at corresponding distances ( $\sim 8$  and  $\sim 5 \text{ \AA}$ ) (17).

When comparing  $\zeta_{\text{coh}}(q, \Delta t)$  (Fig. 2, C and D) with  $I(q)$  (Fig. 1, C and D), it can be seen that the maxima in  $\zeta_{\text{coh}}(q, \Delta t)$  occurs roughly at the minima in  $I(q)$ . This observation can be explained by the theory of “de Gennes narrowing” such that the system concerned exhibits the fastest fluctuations at the least-favored configurations (17–19).

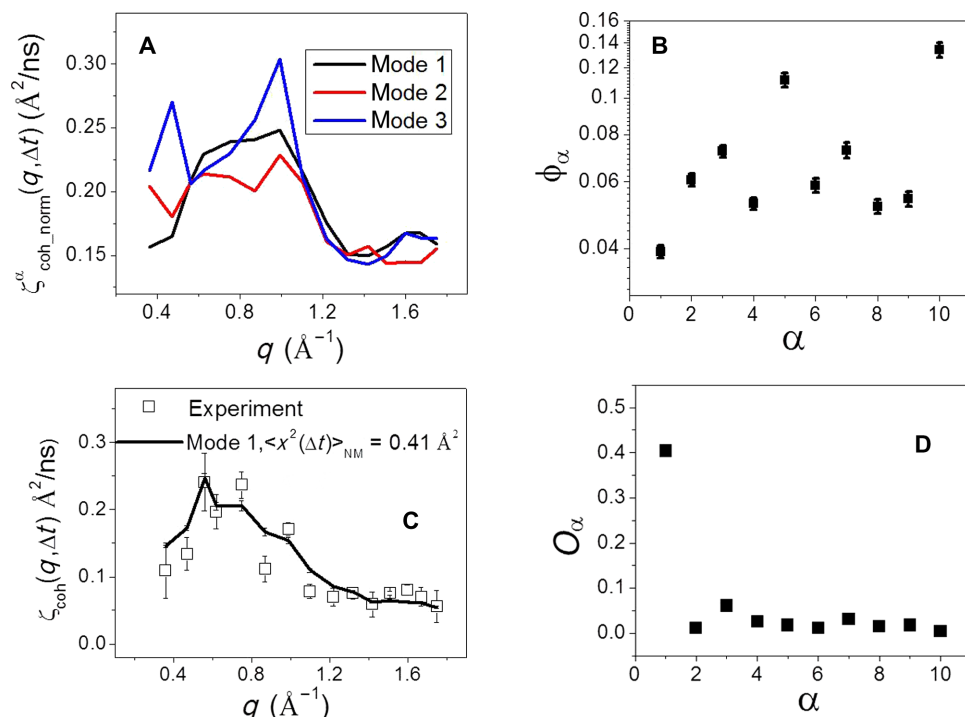
A close examination of Fig. 2 (C and D) shows that  $\zeta_{\text{coh\_exp}}(q, \Delta t)$  differs quantitatively from  $\zeta_{\text{coh\_MD}}(q, \Delta t)$ . This is largely due to the experimental data having been collected on protein powders, which involve interprotein scattering, whereas the simulation-derived values are averages over the results calculated from each of the eight protein molecules in the simulation box. The interprotein scattering effect can be evaluated using the interprotein structure factor,  $S(q)$ , defined as the ratio between the overall static scattering intensity  $I(q)$  measured on the powder sample and the form factor  $F(q)$  of a single protein molecule (Fig. 3A) (11, 18). As seen in Fig. 3B, after correction for  $S(q)$ , the simulation-derived  $\zeta_{\text{coh}}(q, \Delta t)$  agrees quantitatively with the experimental values. This correction procedure is further

supported by the MD results presented in Fig. 3C. Moreover, unlike the earlier NSE experiments performed on protein solutions, where the neutron signals measured are dominated by the global rotation and translation of the biomacromolecules (11, 13), rendering the quantitative characterization of the protein internal dynamics difficult, we used powder samples here, where the dynamic neutron signals come almost entirely from protein internal motions (see Fig. 3D).

### Normal mode analysis

To determine the forms of the large-scale collective in-phase protein motions probed experimentally, we performed a normal mode analysis using a commonly used elastic network model (20), which assumes that all protein residues move synchronously along the eigenvectors of the normal modes. Because low-frequency normal modes are those most likely to be global, only the 10 lowest-frequency normal modes are considered.  $\zeta_{\text{coh}}(q, \Delta t)$  derived individually from the first three lowest-frequency normal modes is presented in Fig. 4A. The  $q$  dependence of  $\zeta_{\text{coh}}(q, \Delta t)$  differs significantly for these modes, especially in the low- $q$  region,  $< 1 \text{ \AA}^{-1}$ . To determine the normal mode that matches best the experimental  $\zeta_{\text{coh}}(q, \Delta t)$ , a dimensionless scoring function is defined as below

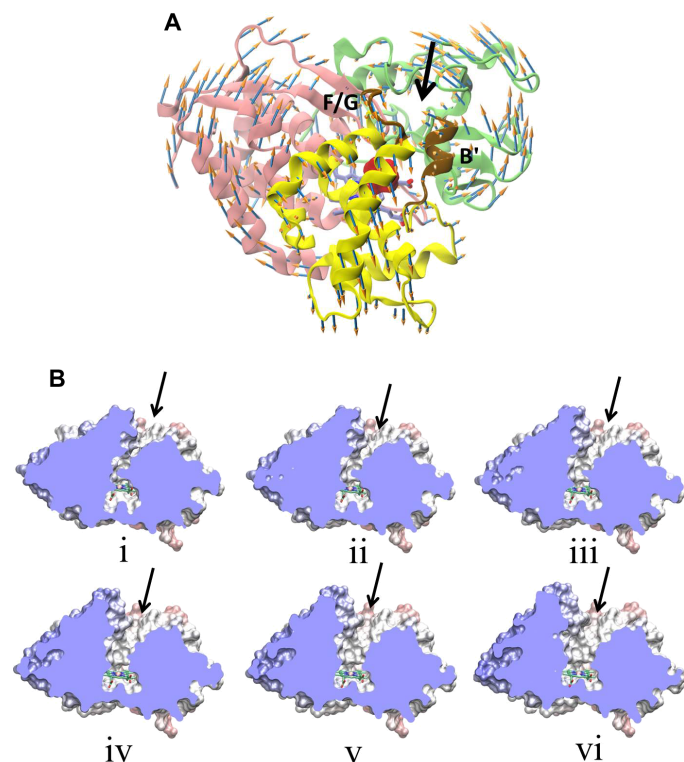
$$\phi_{\alpha} = \frac{\sum_q (\langle x^2(\Delta t) \rangle_{\text{NM}}^{\alpha} \zeta_{\text{coh\_norm}}^{\alpha}(q, \Delta t) / S(q) - \zeta_{\text{coh\_MD}}(q, \Delta t))^2}{\sum_q (\zeta_{\text{coh\_MD}}(q, \Delta t))^2} \quad (2)$$



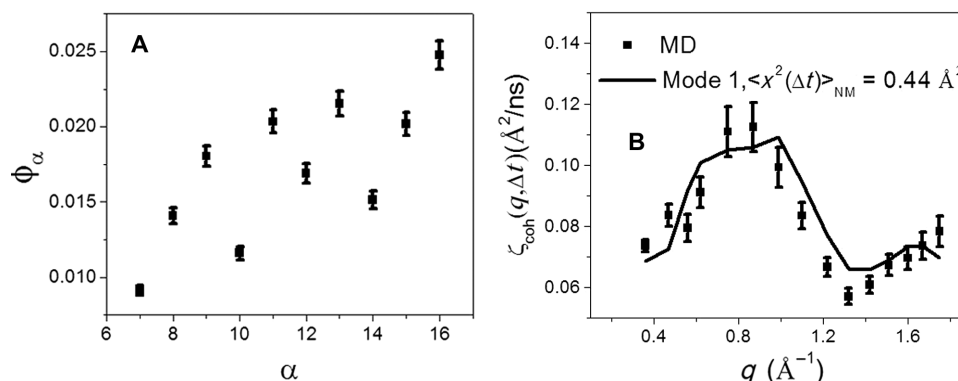
**Fig. 4. Normal mode analysis.** (A)  $\zeta_{\text{coh\_norm}}^{\alpha}(q, \Delta t)$  derived from the first three lowest-frequency normal modes (see eq. S11 and the corresponding text in the Supplementary Materials). (B) Scoring function  $\phi_{\alpha}$  for different normal modes (Eq. 2). (C) Comparison of  $\zeta_{\text{coh}}(q, \Delta t)$  measured experimentally and that derived from the first normal mode, the latter being corrected by the interprotein structure factor (Fig. 3A) and multiplied by the amplitude ( $\langle x^2 \rangle_{\text{NM}} = 0.41 \text{ \AA}^2$ ) required to minimize  $\phi_{\alpha}$ . (D) Overlap parameter,  $O_{\alpha}$ , between the eigenvector of a given normal mode and the direction of the conformational change between the open, substrate-free and the closed, substrate-bound forms of CYP101 (Eq. 4).



where  $\langle x^2(\Delta t) \rangle_{\text{NM}}$  is the mean-squared atomic displacement of mode  $\alpha$  (eq. S10), quantitatively characterizing the amplitude of the mode, and serves as a  $q$ -independent fitting parameter adjusted to minimize  $\phi_\alpha$ .  $S(q)$  is the experimental interprotein structure factor (Fig. 3A), and  $\zeta_{\text{coh\_norm}}^\alpha(q, \Delta t)$  is the decay of the coherent intermediate scattering function resulting from mode  $\alpha$  when its amplitude is normalized, that is,  $\langle x^2(\Delta t) \rangle_{\text{NM}} = 1 \text{ \AA}^2$  (Fig. 4A and eq. S11). The so-obtained  $\phi_\alpha$  characterizes how far the dynamic form of mode  $\alpha$  differs from that measured experimentally. The results (Fig. 4B) identify



**Fig. 5. Schematic pictures of the first normal mode.** (A) Three parts (marked as green, yellow, and pink) in the protein rotate against each other (shown by the colored arrows). The region marked red is the catalytic site. The entrance of the substrate access channel formed by the B' helices and F/G loops is marked brown. (B) Cross-sectional view schematically illustrating the opening of the substrate access channel step by step via mode 1. The black arrow points to the entrance of the channel.



**Fig. 6. Testing the method using MD-derived results.** (A) Scoring function  $\phi_\alpha$  (Eq. 3) for different normal modes with reference to the MD-derived  $\zeta_{\text{coh}}(q, \Delta t)$ . (B) Comparison of  $\zeta_{\text{coh}}(q, \Delta t)$  derived from MD with that calculated from the first normal mode with  $\langle x^2 \rangle_{\text{NM}} = 0.44 \text{ \AA}^2$  required to minimize  $\phi_\alpha$ .

mode 1 as best matching the experiment. A schematic picture of mode 1 is presented in Fig. 5A, showing a relative rotational motion of three domains in the protein.

$\zeta_{\text{coh}}(q, \Delta t)$  derived from mode 1 with the minimized value of  $\phi_\alpha$  is shown in Fig. 4C, and this explains the experimental results well. The minimization of  $\phi_\alpha$  is achieved when the amplitude of the normal mode  $\langle x^2(\Delta t) \rangle_{\text{NM}} = 0.41 \text{ \AA}^2$  (Eq. 2). This value is only slightly smaller than that of the average mean-squared atomic displacement,  $0.57 \text{ \AA}^2$ , estimated experimentally from the  $q$  dependence of the intensity of the elastic peak in the incoherent dynamic structure factor,  $S_{\text{inc}}(q, \Delta t)$  (see fig. S4). The small difference is consistent with the fact that the coherent scattering is primarily sensitive to correlated interatomic motions, whereas the mean-squared atomic displacement estimated from  $S_{\text{inc}}(q, \Delta t)$  includes both correlated and uncorrelated dynamics.

The catalytic site of CYP101 is deeply buried, and the substrate enters the site via a channel (Fig. 5), the entrance of which is formed by the B' helix and the F/G loop (Fig. 5A). Setting the amplitude of mode 1 as  $\langle x^2 \rangle_{\text{NM}} = 0.41 \text{ \AA}^2$ , the gate of the substrate access channel, approximated here as the distance between Gly<sup>189</sup> in the F/G loop and Arg<sup>89</sup> in the B' helix, varies by  $\sim 0.5 \text{ \AA}$  on an experimental time scale of  $\sim 1 \text{ ns}$  (eq. S10).

To test the above method, we also calculated the scoring function for different normal modes with respect to the MD results using Eq. 3

$$\phi_\alpha = \frac{\sum_q (\langle x^2(\Delta t) \rangle_{\text{NM}} \zeta_{\text{coh\_norm}}^\alpha(q, \Delta t) - \zeta_{\text{coh\_MD}}(q, \Delta t))^2}{\sum_q (\zeta_{\text{coh\_MD}}(q, \Delta t))^2} \quad (3)$$

Unlike Eq. 2, Eq. 3 does not contain potential contamination from the interprotein structure factor,  $S(q)$ , because the MD-derived result is just an average over the values calculated from the individual protein molecules in the MD box.  $\phi_\alpha$  thus obtained for different normal modes is presented in Fig. 6A. Again, mode 1 matches best the MD results, with the amplitude of the mode at  $\langle x^2 \rangle_{\text{NM}} = 0.44 \text{ \AA}^2$  (see Fig. 6B). This amplitude leads to the fluctuation of the gate for the substrate access channel again being  $\sim 0.5 \text{ \AA}$ , in good agreement with the value of  $0.4 \text{ \AA}$  calculated directly from the MD trajectory, quantitatively validating our method.

## DISCUSSION

The opening and closing of the substrate access channel in CYP101 (Fig. 5) regulate the access of the substrate into the catalytic site (21) and are thus crucial for the protein function. To quantify the relevance of the dynamical pathway in each normal mode for the opening and closing of the channel, we calculated the overlap parameter,  $O_\alpha$ , between the direction vector of a given normal mode,  $\alpha$ , and that for the conformational change between the open, substrate-free and the closed, substrate-bound forms of CYP101.  $O_\alpha$  is a dot product defined as

$$O_\alpha = \frac{\left| \sum_i^Z \vec{d}_i^\alpha \cdot \Delta r_i \right|}{\left[ \sum_i^Z |\vec{d}_i^\alpha|^2 \sum_i^Z |\Delta r_i|^2 \right]^{1/2}} \quad (4)$$

where  $\vec{d}_i^\alpha$  is the displacement of atom  $i$  in normal mode  $\alpha$  and  $\Delta r_i$  is the displacement of that atom from the open, substrate-free structure [Protein Data Bank (PDB): 3L62] to the closed, substrate-bound form (PDB: 3L63) of CYP101 (21).  $O_\alpha$  for mode 1 is about one order of magnitude larger than that for other modes (Fig. 4D). Therefore, mode 1, identified through the coherent neutron-scattering experiment in the present work, is also the mode most likely to determine the opening and closing of the substrate access channel in CYP101. Figure 5B schematically illustrates the opening-closing process of the substrate access channel via mode 1 viewed from a cross section of the protein molecule. The present work demonstrates that a complete characterization of large-scale functional protein dynamics, that is, atomic-detail information on the forms, time scales, and amplitudes of modes, can be obtained from backscattering spectroscopic experiments that measure the inelastic coherent neutron-scattering signal from a perdeuterated protein powder.

We found that collective, in-phase dynamic modes, that is, residues moving synchronously in time with similar amplitudes, span several nanometer length scales across the P450 molecule on a ~100-ps time scale. Moreover, the  $q$  dependence of the decay of the coherent intermediate scattering function obtained experimentally and from simulation is shown to be able to determine the most probable pathway for the collective motion in the protein molecule at atomic resolution and, furthermore, to derive the explicit amplitude of the motion on the instrumentally accessible time scale. Both the experimental and simulation results presented here indicate that CYP101 uses a specific collective mode in which three domains rotate against each other to create access to the catalytic site for the substrate, a mechanism crucial for the functionality of the enzyme. The experimental method developed here for the quantitative characterization of functional modes is generally applicable to protein systems.

## MATERIALS AND METHODS

### Backscattering

Dynamic neutron-scattering experiments were performed on camphor-bound CYP101 (H-CYP101) and its perdeuterated counterpart (D-CYP101) using the NG2 HFBS at the NIST Center for Neutron Research. To suppress the contributions of water and of global rotational and translational motion of the protein molecule to the scattering signal, lyophilized powder samples were used (22–24). The

dynamic neutron signal thus obtained arose almost entirely from protein internal motions (Fig. 3D). However, the collective motion observed in the present work may exhibit a dependence on the hydration level of the protein, and this deserves further experimental exploration. Details of the sample preparation can be found in the study of Miao *et al.* (25). The quasi-elastic dynamic structure factor,  $S(q, \Delta E)$ , was measured at 285 K. Higher temperatures than this can lead the protein to denature over the lengthy experimental time (~2 days). Typical experimental spectra are presented in Fig. 1 (A and B).

In addition to quasi-elastic scattering, the intensity of the elastic peak in the dynamic structure factor,  $S(q, \Delta t)$ , at an instrumental resolution of HFBS ~ 0.8  $\mu\text{eV}$ , which corresponds to  $\Delta t \sim 1$  ns, was also measured over the  $q$  range of 0.35 to 1.75  $\text{\AA}^{-1}$  in the temperature window from 4 to 298 K. The temperature dependence of  $S(q, \Delta t)$  at three different  $q$  values, normalized by the values at 4 K, is presented in Fig. 2 (A and B, for H-CYP101 and D-CYP101, respectively).

### Neutron spin echo

The overall static structure factor  $I(q)$  (Fig. 1C) was measured on fully deuterated CYP101 powder using an NSE spectrometer at the NIST Center for Neutron Research (Gaithersburg, MD). A standard annular sample can was used in which the sample was loaded into a thin aluminum pouch under helium to avoid any water intake. The wavelength of the incoming neutron beam was set to 5  $\text{\AA}$  with a  $\Delta\lambda/\lambda \sim 0.2$ . With this setup, measurements were made over the  $q$  range of 0.20 to 1.70  $\text{\AA}^{-1}$  with a  $q$  resolution of 0.05  $\text{\AA}^{-1}$ . To obtain good statistics, each point at any given  $q$  was measured for about 5 min. Background measurements were made in an empty annular sample can and were subtracted.

### Simulations

All-atom MD simulations of camphor-bound CYP101 were carried out using NAMD22 (26) with the 3L63 crystal structure [1.5  $\text{\AA}$  resolution (21)]. The simulation box contained eight proteins, obtained by replicating the unit cell (four proteins) along the  $Z$  direction, and 943 water molecules, corresponding to the experimental hydration level, 0.05 g of protein per gram of water. MD simulation at this low hydration level is often used to mimic the neutron-scattering experiments performed on dry protein powders (23, 24). Seventy-two  $\text{K}^+$  ions were added to neutralize the system. The CHARMM22 force field (27) was used for the protein, and TIP3P (28) was used for water molecules. The particle mesh Ewald method (29) was applied for the electrostatics with a real space cutoff of 12  $\text{\AA}$  and a Fourier grid spacing of 1  $\text{\AA}$ . The systems were energy-minimized with 2000 steepest-descent steps. Constant pressure and temperature (NPT) equilibration using velocity-rescaling temperature (30) and Parrinello-Rahman pressure coupling (31) was carried out for 200 ns at  $T = 285$  K and  $P = 1$  atm with coupling time constants of  $\tau = 0.1$  ps and  $\tau = 0.4$  ps, respectively. The last 100 ns were used for the analysis, and the coordinates were written out every 1 ps. More details are presented in the Supplementary Materials of Miao *et al.* (25). The coherent neutron-scattering profiles were calculated using the same MD trajectories used for deriving the incoherent scattering results but by replacing the scattering length of hydrogen atoms by that of deuterium atoms to mimic the experimental conditions. All atoms in the protein were taken into account to derive both the coherent and incoherent scattering profiles. The incoherent signal derived from H-CYP101 was dominated by the signal of hydrogen atoms (>95%), whereas the coherent signal derived from D-CYP101 arose about equally from deuterium atoms and from other elements (carbon, nitrogen, and oxygen).

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/2/10/e1600886/DC1>

Supplementary Text

fig. S1. Coherent and incoherent contribution into the neutron-scattering signal arising from the hydrogenated and deuterated protein samples.

fig. S2. Comparison of  $\frac{I_{\text{D}}(q)P_{\text{coh,D}}(q)}{I_{\text{H}}(q)P_{\text{coh,H}}(q)}$  and  $\frac{I_{\text{D}}(q)}{I_{\text{H}}(q)}$ .

fig. S3. Comparison of  $S_{\text{coh}}(q,\Delta t)$  with  $S_{\text{D}}(q,\Delta t)$ .

fig. S4.  $q^2$  dependence of the intensity of the elastic peak in the incoherent dynamic structure factor,  $S_{\text{inc}}(q,\Delta t)$  (eq. S9), at 285 K.

References (33–36)

## REFERENCES AND NOTES

1. A. Johs, I. M. Harwood, J. M. Parks, R. E. Nauss, J. C. Smith, L. Liang, S. M. Miller, Structural characterization of intramolecular  $\text{Hg}^{2+}$  transfer between flexibly linked domains of mercuric ion reductase. *J. Mol. Biol.* **413**, 639–656 (2011).
2. G. S. Martin, The hunting of the Src. *Nat. Rev. Mol. Cell Biol.* **2**, 467–475 (2001).
3. R. D. Banks, C. C. F. Blake, P. R. Evans, R. Haser, D. W. Rice, G. W. Hardy, M. Merrett, A. W. Phillips, Sequence, structure and activity of phosphoglycerate kinase: A possible hinge-bending enzyme. *Nature* **279**, 773–777 (1979).
4. K. A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M. A. Wilson, G. A. Petsko, M. Karplus, C. G. Hübner, D. Kern, Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838–844 (2007).
5. S. Y. Noskov, S. Bernèche, B. Roux, Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* **431**, 830–834 (2004).
6. K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
7. J. D. Nickels, S. Peticola, H. O'Neill, Q. Zhang, G. Ehlers, A. P. Sokolov, Coherent neutron scattering and collective dynamics in the protein, GFP. *Biophys. J.* **105**, 2182–2187 (2013).
8. G. Venkataraman, B. A. Dasannacharya, K. R. Rao, Cold neutron scattering from liquid  $\text{Cd}_4$ . *Phys. Lett.* **23**, 226–227 (1966).
9. U. Buchenau, A. Wischnewski, D. Richter, B. Frick, Is the fast process at the glass transition mainly due to long wavelength excitations? *Phys. Rev. Lett.* **77**, 4035 (1996).
10. A. P. Sokolov, Vibrations at the boson peak: Random- and coherent-phase contributions. *J. Phys. Condens. Matter* **11**, A213 (1999).
11. R. Biehl, B. Hoffman, M. Monkenbusch, P. Falus, S. Préost, R. Merkel, D. Richter, Direct observation of correlated interdomain motion in alcohol dehydrogenase. *Phys. Rev. Lett.* **101**, 138102 (2008).
12. Z. Bu, R. Biehl, M. Monkenbusch, D. Richter, D. J. E. Callaway, Coupled protein domain motion in *Taq* polymerase revealed by neutron spin-echo spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17646 (2005).
13. L. Hong, M. A. Sharp, S. Poblete, R. Biehl, M. Zamponi, N. Szekely, M.-S. Appavou, R. G. Winkler, R. E. Nauss, A. Johs, J. M. Parks, Z. Yi, X. Cheng, L. Liang, M. Ohl, S. M. Miller, D. Richter, G. Gompper, J. C. Smith, Structure and dynamics of a compact state of a multidomain protein, the mercuric ion reductase. *Biophys. J.* **107**, 393–400 (2014).
14. S. G. Sligar, Glimpsing the critical intermediate in cytochrome P450 oxidations. *Science* **330**, 924–925 (2010).
15. I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, The catalytic pathway of cytochrome p450cam at atomic resolution. *Science* **287**, 1615–1622 (2000).
16. J. M. Carpenter, C. A. Pelizzari, Inelastic neutron scattering from amorphous solids. I. Calculation of the scattering law for model structures. *Phys. Rev. B* **12**, 2391 (1975).
17. L. Hong, N. Smolin, J. C. Smith, de Gennes narrowing describes the relative motion of protein domains. *Phys. Rev. Lett.* **112**, 158102 (2014).
18. B. J. Ackerson, P. N. Pusey, R. J. A. Tough, Interpretation of the intermediate scattering function at short times. *J. Chem. Phys.* **76**, 1279 (1982).
19. P. G. De Gennes, Liquid dynamics and inelastic scattering of neutrons. *Phys. Chem. Chem. Phys.* **25**, 825–839 (1959).
20. K. Suhre, Y.-H. Sanejouand, *ElNémo*: A normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res.* **32**, W610–W614 (2004).
21. Y.-T. Lee, R. F. Wilson, I. Rupniewski, D. B. Goodin, P450cam visits an open conformation in the absence of substrate. *Biochemistry* **49**, 3412–3419 (2010).
22. J. A. Rupley, G. Careri, Protein hydration and function. *Adv. Protein Chem.* **41**, 37–172 (1991).
23. L. Hong, X. Cheng, D. C. Glass, J. C. Smith, Surface hydration amplifies single-well protein atom diffusion propagating into the macromolecular core. *Phys. Rev. Lett.* **108**, 238102 (2012).
24. Y. L. Miao, Z. Yi, D. C. Glass, L. Hong, M. Tyagi, J. Baudry, N. Jain, J. C. Smith, Temperature-dependent dynamical transitions of different classes of amino acid residue in a globular protein. *J. Am. Chem. Soc.* **134**, 19576–19579 (2012).
25. Y. L. Miao, Z. Yi, C. Cantrell, D. C. Glass, J. Baudry, N. Jain, J. C. Smith, Coupled flexibility change in cytochrome P450cam substrate binding determined by neutron scattering, NMR, and molecular dynamics simulation. *Biophys. J.* **103**, 2167–2176 (2012).
26. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
27. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkerich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
28. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
29. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
30. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
31. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
32. R. Inoue, R. Biehl, T. Rosenkranz, J. Fitter, M. Monkenbusch, A. Radulescu, B. Farago, D. Richter, Large domain fluctuations on 50-ns timescale enable catalytic activity in phosphoglycerate kinase. *Biophys. J.* **99**, 2309–2317 (2010).
33. A. M. Gaspar, S. Busch, M.-S. Appavou, W. Haeussler, R. Georgiib, Y. Su, W. Dostera, Using polarization analysis to separate the coherent and incoherent scattering from protein samples. *Biochim. Biophys. Acta* **1804**, 76–82 (2010).
34. U. Lehnert, V. Réat, M. Weik, G. Zaccai, C. Pfister, Thermal motions in bacteriorhodopsin at different hydration levels studied by neutron scattering: Correlation with kinetics and light-induced conformational changes. *Biophys. J.* **75**, 1945–1952 (1998).
35. M. M. Tirion, Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* **77**, 1905–1908 (1996).
36. F. Tama, F. X. Gadea, O. Marques, Y.-H. Sanejouand, Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins* **41**, 1–7 (2000).

**Acknowledgments:** We thank Y. Miao for providing simulation trajectories and A. Faraone for the NSE measurement. **Funding:** L.H. acknowledges support from the High-Performance Computer Facility at Shanghai Jiao Tong University and NSF China (grant 11504231). The Oak Ridge National Laboratory (ORNL) group acknowledges support from the ORNL Adaptive Biosystems Imaging project funded by the Biological and Environmental Research Directorate of the Office of Science of the U.S. Department of Energy. This work used facilities supported in part by the NSF (agreement no. DM-1508249). **Author contributions:** L.H. conducted the experiment and simulation, designed the work, and wrote the paper. X.C. and J.C.S. contributed to the writing and the design of the work. M.T. participated in the neutron experiment. A.B. and N.J. participated in the experiment and made the samples. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 24 April 2016

Accepted 2 September 2016

Published 14 October 2016

10.1126/sciadv.1600886

**Citation:** L. Hong, N. Jain, X. Cheng, A. Bernal, M. Tyagi, J. C. Smith, Determination of functional collective motions in a protein at atomic resolution using coherent neutron scattering. *Sci. Adv.* **2**, e1600886 (2016).

This article is published under a Creative Commons license. The specific license under which this article is published is noted on the first page.

For articles published under [CC BY](#) licenses, you may freely distribute, adapt, or reuse the article, including for commercial purposes, provided you give proper attribution.

For articles published under [CC BY-NC](#) licenses, you may distribute, adapt, or reuse the article for non-commercial purposes. Commercial use requires prior permission from the American Association for the Advancement of Science (AAAS). You may request permission by clicking [here](#).

***The following resources related to this article are available online at <http://advances.sciencemag.org>. (This information is current as of May 23, 2017):***

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://advances.sciencemag.org/content/2/10/e1600886.full>

**Supporting Online Material** can be found at:

<http://advances.sciencemag.org/content/suppl/2016/10/11/2.10.e1600886.DC1>

This article **cites 36 articles**, 4 of which you can access for free at:

<http://advances.sciencemag.org/content/2/10/e1600886#BIBL>

*Science Advances* (ISSN 2375-2548) publishes new articles weekly. The journal is published by the American Association for the Advancement of Science (AAAS), 1200 New York Avenue NW, Washington, DC 20005. Copyright is held by the Authors unless stated otherwise. AAAS is the exclusive licensee. The title *Science Advances* is a registered trademark of AAAS