Modeling difference x-ray scattering observations from an integral membrane protein within a detergent micelle

Cite as: Struct. Dyn. 9, 054102 (2022); doi: 10.1063/4.0000157
Submitted: 25 April 2022 · Accepted: 15 September 2022 · Published Online: 31 October 2022

Daniel Sarabi,1 Lucija Ostojic,1 Robert Bosman,1 Adams Vallejos,1 Johanna-Barbara Linse,1 Michael Wulff,2 Matteo Levantino,2 and Richard Neutze1,a)

AFFILIATIONS
1Department of Chemistry and Molecular Biology, University of Gothenburg, Box 462, 40530 Gothenburg, Sweden
2European Synchrotron Radiation Facility, 38043 Grenoble Cedex 9, France

a)Author to whom correspondence should be addressed: richard.neutze@chem.gu.se

ABSTRACT

Time-resolved x-ray solution scattering (TR-XSS) is a sub-field of structural biology, which observes secondary structural changes in proteins as they evolve along their functional pathways. While the number of distinct conformational states and their rise and decay can be extracted directly from TR-XSS experimental data recorded from light-sensitive systems, structural modeling is more challenging. This step often builds from complementary structural information, including secondary structural changes extracted from crystallographic studies or molecular dynamics simulations. When working with integral membrane proteins, another challenge arises because x-ray scattering from the protein and the surrounding detergent micelle interfere and these effects should be considered during structural modeling. Here, we utilize molecular dynamics simulations to explicitly incorporate the x-ray scattering cross term between a membrane protein and its surrounding detergent micelle when modeling TR-XSS data from photoactivated samples of detergent solubilized bacteriorhodopsin. This analysis provides theoretical foundations in support of our earlier approach to structural modeling that did not explicitly incorporate this cross term and improves agreement between experimental data and theoretical predictions at lower x-ray scattering angles.

INTRODUCTION

An accurate measurement of protein conformational changes as they evolve with time has long been a goal of structural biology.1 Serial crystallography approaches, first implemented at x-ray free electron lasers2 (XFELs), have revolutionized the field of time-resolved crystallography since numerous technical challenges that limited the sphere of application of time-resolved Laue diffraction were overcome.1,3 The critical benefit of time-resolved crystallography is that a successful study delivers high-resolution structural insight into the chemical details of an evolving reaction from sub-picoseconds to minutes. Recent developments in single-particle cryo-electron microscopy can also provide structural insight into the progress of a biological reaction at high-resolution,3 and real-time NMR studies of protein dynamics have yielded insights into protein unfolding and refolding.4 As methods for characterizing protein conformational changes become increasingly mainstream, a full appreciation of the relationship between structure and function must increasingly incorporate structural dynamics.

Although time-resolved crystallography is undergoing a period of rapid growth,1 a fundamental limitation of time-resolved crystallographic techniques is that samples must both crystallize and remain in a highly ordered arrangement as the reaction evolves. It is not always possible to recover well diffracting crystals for any chosen system of study and a reaction of interest may be slowed in the crystalline state.4,9,10 Moreover, crystallographic methods cannot visualize protein conformational changes that are incompatible with the crystal packing lattice since they are either suppressed or they disrupt the crystalline order. Time-resolved x-ray solution scattering (TR-XSS) studies of protein structural changes have been developed to address these limitations.11–29 The key advantage of TR-XSS is that it provides a relatively direct measurement of the timescale and number of species with distinct secondary structural characteristics. Conversely, since the sample is a randomly oriented ensemble of protein molecules in solution, the measured information is only one-dimensional. As such, TR-XSS is not as information rich as time-resolved x-ray diffraction and...
structural modeling relies upon low-resolution shape reconstructions \(^{25,26,27}\) or leans heavily upon additional information that may be extracted from crystallographic structures \(^{11,12,15,18,25}\) or molecular dynamics trajectories.\(^ {13–15,23,24,27–29}\)

Structural interpretation of TR-XSS data recorded from integral membrane proteins \(^{11,12,15,18,25}\) is further complicated by the fact that a detergent micelle, which surrounds the integral membrane protein, \(^{31}\) frequently contains more atoms than the protein itself and is highly disordered. Consequently, difference x-ray scattering data extracted from TR-XSS experiments rely upon structural fluctuations within the detergent micelle averaging out due to the very large number of molecules, which are exposed to an x-ray beam in any experiment. For example, at a concentration of 0.5 mM, approximately \(10^{12}\) protein molecules are found in a 3.6 nl sample volume, which is the volume calculated from an x-ray beam \(60 \times 120 \mu \text{m}^2\) passing through a \(500 \mu \text{m}^2\) capillary.\(^ {12}\) It is, however, not computationally feasible to perform simulations of \(10^{12}\) membrane-protein/micelle systems, and therefore, approximations are required. In our earlier modelling,\(^ {12,18,24,25}\) we argued that the influence of the membrane protein-micelle cross term of atom \(j\) is proportional to the magnitude squared of the structure factor, \(F_j\), with the spatial location of each atom relative to an arbitrary origin,\(^ {33}\)

\[
F_j(q) = \sum_i f_i \exp(i \cdot r_i) ,
\]

where \(q\) is a 3D vector in reciprocal space corresponding to the change in momentum of the scattered x-ray and has a magnitude \(q = 4\pi \sin \theta /\lambda\); \(\lambda\) is the x-ray wavelength; \(\theta\) is half the angle through which the x-ray beam is deflected; \(F_j(q)\) is a complex number called a structure factor that is specified for all values of \(q\); \(f_i\) is the atomic scattering factor of atom \(j\) of the sample; and \(r_i\) is the 3D coordinate in real space of atom \(j\). The intensity of the scattered x-ray beam for any vector \(q\) is proportional to the magnitude squared of the structure factor, \(S(q) = |F(q)|^2\),

\[
S(q) = |F(q)|^2 = \sum_{n=1}^{N} |F_n(q)|^2 ,
\]

If samples are assumed to be monodisperse in solution, it is possible to average over all random orientations of molecules to recover,

\[
S(q) = N \langle |F(q)|^2 \rangle ,
\]

where \(\langle \cdots \rangle\) denotes rotational averaging in space and yields an isotropic x-ray scattering distribution (i.e., depends only upon the magnitude of \(q\)). When summing over all \(N\) molecule in the sample, we have made the assumption that the interference between separate molecules may be neglected.\(^ {33}\) Several packages have been written which build upon this formula to analyze x-ray scattering data. In this work, we utilize the package CRY SOL\(^ {24}\) for x-ray scattering calculations. CRY SOL evaluates Eq. (3) using the atomic coordinates of macromolecules as input and utilizes a multipole expansion to calculate the spherically averaged scattering pattern.

When working with integral membrane proteins surrounded by a detergent micelle, we can divide the protein and micelle contributions of Eq. (1) to become

\[
F_{pm}(q) = F_p(q) + F_m(q) ,
\]

for which \(F_p(q)\) represents the structure factor calculated from the protein alone; \(F_m(q)\) represents the structure factor calculated from the detergent micelle alone; and \(F_{pm}(q)\) is the structure factor calculated from both combined. Substituting Eq. (4) into Eq. (3) and performing rotational averaging over samples in solution, and noting that \(\langle |F(q)|^2 \rangle\) is a real number, gives

\[
\langle |F_{pm}(q)|^2 \rangle = \langle |F_p(q)|^2 + |F_m(q)|^2 \rangle = \langle |F_p(q)|^2 \rangle + \langle |F_m(q)|^2 \rangle + 2\cos \Phi_{pm}(q) \sqrt{\langle |F_p(q)|^2 \rangle \langle |F_m(q)|^2 \rangle} ,
\]

where this expansion is used to define a new variable, \(\Phi_{pm}(q)\), which we introduce to in order to separate the protein and detergent micelle scattering terms within this formalism. As such, \(\cos \Phi_{pm}(q)\) can be evaluated by inverting Eq. (5) and utilizing Eq. (3) to recover,

\[
\cos \Phi_{pm}(q) = \frac{S_{pm}(q) - S_p(q) - S_m(q)}{2\sqrt{S_p(q)S_m(q)}} ,
\]

where \(S_p(q)\) is the isotropic solution scattering intensity calculated from the protein alone, \(S_m(q)\) is the isotropic solution scattering intensity calculated from the detergent micelle alone, and \(S_{pm}(q)\) is the isotropic solution scattering intensity calculated from both protein and micelle combined. In practice, we calculate \(\cos \Phi_{pm}(q)\) by writing separate pdb files containing all of the atoms of the system (Fig. 1(a)), another containing the atoms of the protein alone, and another containing the atoms of the detergent micelle alone. We then calculate the x-ray solution scattering intensities from the corresponding pdb files using CRY SOL. A representative example of \(\cos \Phi_{pm}(q)\) calculated from bacteriorhodopsin in the detergent \(\beta\)-octyl glucoside [Fig. 1(a)] is given in Fig. 1(b). It can be seen that \(\cos \Phi_{pm}(q)\) varies smoothly with \(q\) between the values \(1\) and \(−1\); \(\cos \Phi_{pm}(0)\) is unity due to the fact that all scattering in the forward direction is in phase; and at higher scattering angles, \(\cos \Phi_{pm}(q)\) varies smoothly with \(q\) with the values \(1\) and \(−1\).

In TR-XSS studies, the most important experimental observation is the change in x-ray scattering as a function of x-ray scattering angle and time, \(\Delta S_{exp}(q, t)\). When modeling protein conformational changes, we are, therefore, concerned with how changes in the structure of the integral membrane protein lead to changes in the measured x-ray scattering amplitudes. From Eq. (5), and using the notation \(S_p(q)\) in favor \(N\langle |F_p(q)|^2 \rangle\), etc., we find that if structure of the membrane protein or detergent micelle changes then corresponding change in x-ray scattering intensities will be
Structural Dynamics

\[
\Delta S_{pm}(q) = \Delta S_p(q) + \Delta S_m(q) + 2 \cos \Phi_{pm}(q) \left( \sqrt{S_p(q)} \Delta S_m(q) \right)
\]

where \( \Delta S_{pm}(q) \) is the difference between the x-ray scattering predictions generated from one protein structural model (the candidate structure) and the initial structural model (the resting conformation) and \( \Delta S_m(q) \) is the difference between the x-ray scattering predictions generated from one micelle structural model and another. Likewise, \( \Delta \left[ \cos \Phi_{pm}(q) \right] \) is calculated as the change in \( \cos \Phi_{pm}(q) \) predicted by Eq. (6) when comparing two different molecular conformations of the entire system. We also used the chain rule to make the approximation \( \Delta \sqrt{S_p(q)} \approx \Delta S_p(q)/[2 \sqrt{S_p(q)}] \) and \( \Delta \sqrt{S_m(q)} \approx \Delta S_m(q)/[2 \sqrt{S_m(q)}] \), which hold if \( \Delta \sqrt{S_p(q)} \ll S_p(q) \) and \( \Delta S_m(q) \ll S_m(q) \), both of which are reasonable assumptions for any TR-XSS study of protein structural changes.

The three terms of Eq. (7) illustrate the challenge of analyzing TR-XSS studies of integral membrane proteins. The first term, \( \Delta S_p(q) \left[ 1 + \frac{\Delta S_m(q)}{\sqrt{S_p(q)}} \cos \Phi_{pm}(q) \right] \), is proportional to x-ray scattering.
changes arising from conformational changes in the protein but is modulated by the scattering cross term between the protein and membrane. Likewise the second term, \( \Delta S_{\text{tot}}(q) \left[ 1 + \frac{\sqrt{S_{\text{pm}}(q)}}{S_{\text{solv}}(q)} \cos \phi_{\text{pm}}(q) \right] \), is proportional to the x-ray scattering changes associated with structural changes in the detergent micelle, which is also modulated due to the scattering cross term between the protein and membrane. The final term, \( 2\sqrt{S_{\text{pm}}(q)} / S_{\text{solv}}(q) \Delta \cos \phi_{\text{pm}}(q) \), reflects how changes in the positions of the atoms of the protein and the micelle influence the effective phase coherence of x-ray scattering cross term between the protein and solute.

In our previous analyses\textsuperscript{12,18,24} we effectively assumed that both \( \Delta S_{\text{tot}}(q) \) and \( \Delta \cos \phi_{\text{pm}}(q) \) were vanishingly small outside the small-angle scattering domain and, therefore, only the first term of Eq. (7) needed to be considered during structural fitting. In this work we use the analysis of multiple conformations sampled in molecular dynamics trajectories of the protein within a detergent micelle to assess the validity of these assumptions and, consequently, to extend the modeling formalism as appropriate.

When comparing theoretical predictions with experimental data, it is also essential to determine the influence of the solvent volume displaced by both the protein and micelle, which is usually referred to as the solvent-excluded volume.\textsuperscript{12} When predicting x-ray scattering changes as an integral membrane molecular dynamics trajectories of the protein within a detergent micelle, we can write the total x-ray scattering after correcting for the solvent-excluded volume, \( F_{\text{tot}}(q) \), as

\[
|F_{\text{tot}}(q)|^2 = |F_{\text{pm}}(q) - F_{\text{solv}}(q)|^2, \quad (8)
\]

where the negative sign preceding \( F_{\text{solv}}(q) \) arises from the fact that the x-ray scattering from the solvent-excluded volume due to the presence of the protein and detergent micelle then we can write the total x-ray scattering after correcting for the solvent-excluded volume, \( F_{\text{tot}}(q) \), as

\[
S_{\text{tot}}(q) = S_{\text{pm}}(q) + S_{\text{solv}}(q) - 2 \cos \phi_{\text{solv}} \sqrt{S_{\text{pm}}(q)} \sqrt{S_{\text{solv}}(q)}, \quad (9)
\]

which, in turn, defines

\[
\cos \phi_{\text{solv}}(q) = \frac{S_{\text{pm}}(q) + S_{\text{solv}}(q) - S_{\text{tot}}(q)}{2 \sqrt{S_{\text{pm}}(q)} \sqrt{S_{\text{solv}}(q)}}. \quad (10)
\]

When predicting x-ray scattering changes as an integral membrane protein changes conformation, we cannot reliably predict the changes in structure of the solvent-excluded volume. We instead assume that both the resting and altered conformations are equally modulated by solvent-exclusion effects and Eq. (9) can then be used to calculate the ratio \( S_{\text{tot}}(q) / S_{\text{pm}}(q) \) for all values of \( q \). With this rationale, we make the approximation that the theoretical predictions, \( \Delta S_{\text{pm}}(q) \), calculated from Eq. (7) should be scaled according to

\[
\Delta S_{\text{theory}}(q) = \Delta S_{\text{pm}}(q) \times \frac{S_{\text{tot}}(q)}{S_{\text{pm}}(q)}, \quad (11)
\]

in order to compare theoretical predictions against experimental difference x-ray scattering data.

**RESULTS**

**TR-XSS data collection**

Detergent solubilized samples of bR were prepared as previously described.\textsuperscript{12} TR-XSS data (Fig. 2) were recorded at the beamline ID09 of the ESRF using polychromatic x-rays following photo-excitation with a green (532 nm) laser pulse 4 ns in duration. TR-XSS data were recorded for the time delays \( \Delta t = 5.25, 13.8, 36.2, 95.2, 250, 657 \) \( \mu s \), 1.725, 4.54 12.0, and 31.4 ms between the arrival of the 532 nm photo-activation laser pulse and the x-ray probe. Results from these measurements were compared with earlier TR-XSS measurements from bR\textsuperscript{12} also recorded using a 532 nm laser pulse for photoactivation but 100 ns in duration, and for the time-delays \( \Delta t = 360, 800 \) ns, 2, 20, 65, 200, 650 \( \mu s \), 2, 20, and 100 ms. In our previous TR-XSS studies, we used a physical translation of the goniometer to translate samples between each laser flash,\textsuperscript{12} whereas in these studies, we flowed the sample through a capillary and thereby replaced the sample between each and every laser flash. In this work, we further recorded XSS data during the photo-excitation of bR samples using continuous illumination with an infrared (IR) laser with a wavelength of 1470 nm, which allowed samples of solubilized bR to be heated without initiating the bR photocycle. In this manner, we recorded directly the influence of sample heating on different XSS spectra [Fig. 2(a)].

In our earlier analysis,\textsuperscript{12} we assumed that the time-delay \( \Delta t = 100 \) ms represented the difference spectrum, \( \Delta S(q) \), arising purely from the effects of heating. This assumption, however, was an approximation since a superposition of the difference scattering data from 2009 for the 100 ms time-point, \( \Delta S_{2009}(q, \Delta t = 100 \) ms), with the difference scattering recorded here using IR excitation, \( \Delta S_{2017}(q, \text{IR}) \), shows a significant discrepancy for \( q < 0.3 \) Å\(^{-1} \) [Fig. 2(a)]. We, therefore, removed laser-induced heating from the difference XSS data by subtracting \( \Delta S_{2017}(q, \text{IR}) \) from both the 2017 and 2009 data, where the scaling factor for \( \Delta \) is chosen to minimize the difference x-ray scattering \( \sum (|\Delta S_{\text{IR}}(q, \Delta t) - \Delta S_{\text{IR}}(q, \Delta t)|^2) \) over the domain 0.51 Å\(^{-1} \) \( \leq q \leq 1.45 \) Å\(^{-1} \). The resulting heat-corrected time-dependent difference XSS data, \( \Delta S_{\text{HE}}(q, \Delta t) \), are shown in Figs. 2(c) and 2(d), respectively. Linear decomposition of the 2009 data after the removal of heating produced two basis spectra [Fig. 2(b)] with the populations of components 1 and 2 crossing at \( \Delta t \approx 35 \) ms [Fig. 2(e)]. The limited time-sampling of these data, this value is close to the timescale for the crossing of two populations (\( \Delta t \approx 13 \) ms) in a time-resolved serial femtosecond crystallography (TR-SEK) study using microcrystals of bR\textsuperscript{12} and which correlated well with the L-to-M spectral transition. Unexpectedly, the amplitude of the difference XSS signal after the removal of heat from the 2017 data was only 40% of the amplitude achieved in the 2009 data. As such, linear decomposition of the TR-XSS difference data from 2017 failed to produce a convincing difference signal for the first component, but the difference XSS signal for the second component was in close agreement (Pearson correlation coefficient of 97%) with that recovered for component 2 from the 2009 data [Fig. 2(b)]. This reproducibility demonstrates that different sample preparations, which show variations in detergent composition as well as other experimental parameters, do not unduly influence the experimental difference XSS signal.

The choice of an appropriate pump laser fluence in time-resolved diffraction and x-ray scattering studies has received considerable attention.\textsuperscript{4,5,15,36} In these studies, the pump laser fluence and protein concentration differed between the 2009 and 2017 data collection runs. The earlier studies passed 250 pJ/pulse through a spot focused to a FWHM (full width half maximum) of 265 \( \times \) 265 \( \mu m^2 \), amounting to an average pump laser fluence across the FWHM of 227 mJ/cm\(^2\). Using the procedure described in Ref. 24, we estimate a laser-induced temperature.
jump $\Delta T = 0.5^\circ$C from the observed difference XSS signal. Since the protein concentration was 0.57 mM, this heat jump equates to approximately 15 photons being absorbed per bR molecule, which is 14% of the dimensionless product $\pi / C_1 F / h / C_23$, where $\pi$ is the resting-state cross section of bR, $F$ is the average fluence through the FWHM, $h$ is Planck’s constant, and $\nu$ is the frequency of the absorbed photon. In the 2017 data-collection, 500 $\mu l$ pulse were focused into a laser spot FWHM of 310 $\mu m \times 1700 \mu m$, which equates to an average pump laser fluence of 60 mJ/cm$^2$ yielding $\sigma - F / h \nu = 28$. Given that the protein concentration was 1.3 mM, the measured laser induced heating of 0.25$^\circ$C equates to three photons being absorbed per bR molecule, or 11% of $\sigma - F / h \nu$. That these heating estimates are 14% and 11% of the value anticipates from $\sigma - F / h \nu$, respectively, may be due to the optical density of the sample through a 1 mm capillary (OD $\approx 1.3$ over 0.5 mm, $10^{-1.3} = 0.05$, 2009 data) and a 500 $\mu m$ capillary (OD $\approx 1.5$ over 0.25 mm, $10^{-1.5} = 0.02$, 2017 data) causing the laser fluence throughout.
the probed sample volume to be much lower than at the surface of the capillary. It is, however, difficult to explain why the amplitude of the difference signal extracted from the 2017 study [Fig. 2(b)] is only 34% of the difference signal observed in 2009, given that the protein concentration in the latter experiment was more than twice that of the earlier study, and therefore, the occupancy of the excited state population recorded in 2017 was only 15% of that achieved in the 2009 data. It is possible that the difference photo-kinetics of the photo-stationary phase (data were collected in 2009 using a 100 ns pump-laser; data were collected in 2017 using a 5 ns pump-laser) may underpin this result. A Guinier plot calculated for the absolute x-ray scattering at two protein concentrations [Fig. 2(b)] also shows a departure from linearity, which may indicate sample aggregation. It may, therefore, be counterproductive to push the protein concentration to the point where the optical density of the sample across the flow-cell becomes too large. We, thus, recommend that, in addition to optimizing the spatial overlap between the laser focus and the x-ray beam, time should also be taken to optimize other parameters such as the protein concentration, the flow-cell thickness, the pump laser fluence, and (whenever possible) the pump-laser duration before proceeding to collect TR-XSS data.

**Influence of the detergent micelle on difference x-ray scattering**

When modeling difference x-ray scattering basis spectra [Fig. 2(b)] extracted from heat-corrected TR-XSS data [Figs. 2(c) and 2(d)], it is necessary to predict x-ray scattering changes for a set of candidate secondary structural conformational changes. Disagreement between these predictions and the experimental difference spectra is then minimized by sampling various amplitudes and combinations of the chosen motions. From the theoretical developments above, Eq. (7) highlights in addition to the influence of protein conformational changes on the x-ray scattering, $\Delta_S(q)$, it is important to also model the influence of changes within the micelle on the difference x-ray scattering, $\Delta_{Sm}(q)$, and changes in the phase relationships between the x-ray scattering from the protein and from the detergent micelle, $\Delta \{\cos \Phi_{Sm}(q)\}$. We approach this problem by performing molecular dynamics simulations of bacteriorhodopsin placed within a detergent micelle of $\beta$-octyl glucoside [Fig. 1(a)] using the package GROMACS. Candidate motions of $\alpha$-helices were extracted from published time-resolved serial crystallography studies of bacteriorhodopsin, and energy restraints of 1000 kJ/nm$^2$ were used to drive the protein conformational states of bacteriorhodopsin toward the $\alpha$ backbone coordinates describing these candidate motions. Specifically, a vector describing $\alpha$ movements within helix C, $(\Delta C_{\alpha}^C)$, was extracted from pdb entry 5B6Z minus 5B6V using $\alpha$ coordinates from Pro77 to Pro91; a vector describing $\alpha$ movements within helices E and F, $(\Delta C_{\alpha}^{EF})$, was extracted from pdb entry 6RPH minus 6RQP using $\alpha$ coordinates from Leu152 to Pro186; and a perturbation of $\alpha$ atoms from Leu211 to Phe219 of helix G associated with retinal isomerization was also incorporated $(\Delta C_{\alpha}^{G})$. Moreover, electron density for the EF loop for entry 6RPH was inspected, and for those residues where the electron density was not well defined (residues 158, 159, and 161–167), the $\Delta C_{\alpha}^{EF}$ displacements were determined from a weighted average of the coordinate displacements of the adjacent residues. The amplitudes of helices E and F and the helix C motions were then scaled by the variables $\gamma$ and $\delta$, respectively, and the $\alpha$ atoms of bacteriorhodopsin were driven toward the coordinates,

$$C_{\alpha} \rightarrow C_{\alpha} + \gamma \times \Delta C_{\alpha}^{EF} + \delta \times \Delta C_{\alpha}^{G},$$

(12)

using energy restraints of 1000 kJ/nm$^2$ during molecular dynamics simulations [Fig. 1(c)] of 20 ns (preparative analysis) or 5 ns (fitting runs) in duration using a time step of 2 fs, with pdb files written out for x-ray scattering calculations every 10 ps. Each of these simulations contained 3565 atoms within the protein of which 1816 atoms are hydrogen atoms and 1749 are non-hydrogen atoms; and the micelle contained 190 molecules of $\beta$-octyl glucoside with 9120 atoms of which 5320 are hydrogen atoms and 3800 are non-hydrogen atoms. Because the retinal is isomerized for all activated state trajectories, we did not scale $\Delta C_{\alpha}^{G}$ by a variable parameter but rather we kept the amplitude of this motion equal to the change observed using time-resolved x-ray crystallography. Although this assumption could be challenged, the inclusion of too many free parameters during refinement against the 1D difference x-ray scattering data also risks overfitting.

Since the x-ray scattering contribution is proportional to $\langle |F_i(q)|^2 \rangle$, the scattering contribution of the detergent micelle is approximately five times larger than that of the protein. Moreover, of the 231 residues within the Br model, only 59 residues (15 in helix C + 35 in helices EF + 9 in helix G) had their backbone $\alpha$ atoms displaced (i.e., 26% of the protein) according to Eq. (12). As such, we are aiming to predict the impact on the x-ray scattering of C$\alpha$ movements associated with approximately 26% of the protein, or 8% of the protein plus micelle system. Moreover, whereas molecular dynamics simulations restrain the C$\alpha$ atoms of the protein, there were no such restraints associated with the atoms of the detergent micelle. As such, to utilize the predictions of Eq. (7), it was essential to develop a strategy for which x-ray scattering changes associated with random fluctuations of the detergent micelle did not completely dominate the predicted changes.

X-ray scattering changes (without solvent correction) arising from CRYSTOL calculations from the protein in isolation, $\Delta S_i(q)$ of Eq. (7), are shown in Fig. 3(a) as $\gamma$ is varied stepwise from 0 to 1.5 in ten steps of $\frac{1}{4}$ and $\delta$ is varied from 0 to $\frac{1}{2}$ in five steps of $\frac{1}{4}$. Despite fluctuations associated with individual conformations along these C$\alpha$-restrained molecular dynamics trajectories, when the principal SVD components of the difference x-ray scattering intensities are extracted from the set of 2001 difference x-ray scattering curves (calculated pairwise, by sampling a perturbed and resting conformation every 100 fs of their respective trajectories), there is a smoothly varying change in the predicted difference x-ray scattering as $\gamma$ and $\delta$ are varied. This demonstrates that the structural bias imposed by Eq. (12) during C$\alpha$-restrained molecular dynamics simulations is sufficiently powerful to allow quantitative estimates of $\gamma$ and $\delta$ values when fitting against the experimental difference x-ray scattering spectra. As such the use of energy restraints to drive C$\alpha$ atoms toward a given target structure extends our earlier modeling of these difference x-ray scattering data using rigid motions of $\alpha$-helices. Similarly, conformational ensembles extracted from molecular dynamics simulations have been used to fit TR-XSS data from visual rhodopsin and several other approaches using molecular dynamics have also been reported.

In contrast with the x-ray scattering predictions deriving from the protein in isolation, difference x-ray scattering predictions extracted from the same set of molecular dynamics trajectories but focusing upon changes in x-ray scattering resulting from the micelle...
alone [the term proportional to $\Delta S_m(q)$ in Eq. (7)] do not predict smoothly evolving changes that correlate with $\gamma$ and $\delta$. Whereas there are similarities in the period of oscillations in the predicted difference x-ray scattering curves which evolve smoothly as $\gamma$ and $\delta$ are varied, these curves are calculated from 2000 snapshots from a trajectory 20 ns in duration. (b)–(e) Difference x-ray scattering predictions extracted from the same molecular dynamics trajectories from the micelle alone. (b) is calculated from the first 500 snapshots of the trajectory; (c) is calculated from the next 500 snapshots; (d) is calculated from the next 500 snapshots; and (e) is calculated from the last 500 snapshots of the 20 ns trajectory. These results show little correlation of $\Delta S_m(q)$ with the values of $\gamma$ and $\delta$ and change significantly as the simulation evolves, reflecting the random nature of structural changes within the micelle. (f) Difference x-ray scattering predictions from the protein plus micelle system, $\Delta S_{pm}(q)$, calculated using Eq. (13). As with (a), the interchange the detergent micelle coordinates extracted from parallel simulations [Figs. 1(h) and 1(g)] yields x-ray scattering changes that evolve smoothly as $\gamma$ and $\delta$ are varied.

In addition to structural changes within the protein and within the detergent micelle, we must also consider changes in the x-ray scattering cross term between the protein and the detergent micelle as the protein changes structure, summarized in Eq. (7) as the term involving $\Delta \{\cos \Phi_{pm}(q)\}$. This cross term is influenced by structural changes in both the protein and the micelle as $\gamma$ and $\delta$ are varied [Fig. 1(c)]. To side-step the issue of structural fluctuations within the detergent micelle [Figs. 3(b)–3(e)], we developed the following protocol to extract the influence of $\Delta \{\cos \Phi_{pm}(q)\}$ on $\Delta S_{pm}(q)$ arising from conformational changes within the protein alone. As illustrated in Fig. 1, we first generate coordinate files extracted from molecular dynamics trajectories of the protein in its resting conformation ($p_{rest}$), within a detergent micelle ($m_{rest}$) [Fig. 1(d)], and generate x-ray scattering predictions, $\Delta S_{pm_{rest}}(q)$ using CRY SOL. Similarly, we also extract the coordinates from molecular dynamics trajectories of the protein in a modified conformation ($p_{exc}$), within another detergent micelle ($m_{exc}$), and again use CRY SOL to predict $\Delta S_{pm_{exc}}(q)$ for every value of $\gamma$ and $\delta$ [Fig. 1(e)]. To ensure that the effects of structural fluctuations within the detergent micelle cancel, we next exchanged the coordinates of the detergent micelle to create a set of complementary pdb files that were written with the sampled resting conformations of the protein, $p_{rest}$, but placed within the sampled micelle conformations of the excited trajectory, $m_{exc}$ [Fig. 1(f)]; and conversely for the sampled excited conformation of the protein, $p_{exc}$, placed within the sampled micelle conformations of the resting trajectory, $m_{rest}$ [Fig. 1(g)]. From these complementary sets of pdb files, $S_{pm_{exc}}(q)$ and $S_{pm_{rest}}(q)$ were calculated and the expression,

$$\Delta S_{pm}(q) = \frac{1}{2} \left( S_{pm_{exc}}(q) - S_{pm_{rest}}(q) + S_{pm_{exc}}(q) + S_{pm_{rest}}(q) - S_{pm_{exc}}(q) - S_{pm_{rest}}(q) \right),$$

was utilized to ensure that the term proportional to $\Delta S_m(q)$ of Eq. (7) is zero. Figure 3(f) shows the results of these calculations, revealing a continuous evolution of $\Delta S_{pm}(q)$ as $\gamma$ and $\delta$ are varied, and is similar to the smooth variation in x-ray scattering seen for the protein alone [Fig. 3(a)]. In this manner we established a protocol that allows the cross term between the changing protein structure and a detergent micelle to be evaluated from molecular dynamics trajectories, while avoiding the central problem that these predictions become completely dominated by random fluctuations within the detergent micelle. This procedure is therefore utilized for recovering theoretical fits (i.e., best estimates of $\gamma$ and $\delta$ their respective uncertainties) against the
Influence of different sized detergent micelles on difference x-ray scattering

Equation (13) provides a working protocol for predicting x-ray scattering changes from a membrane protein within a detergent micelle without these predictions being dominated by random structural fluctuations of detergent molecules. In fitting the difference x-ray scattering data, we used a detergent micelle consisting of 190 molecules of β-octyl glucoside. This number of detergent molecules was chosen since it formed a stable micelle which completely surrounded a single molecule of bacteriorhodopsin without excess molecules breaking off during simulations. This choice, however, is somewhat arbitrary and there will always be a distribution of micelle sizes within any experimental sample. To account for the influence of micelle-size variations, predictions from Eq. (13) were repeated for a selected movement of helices E and F \( [\gamma = 1; \delta = 0, \text{Eq. (12)}] \), but with the size of the detergent micelle varying from 170 to 230 β-octyl glucoside molecules, increasing in steps of five detergent molecules additional for each trajectory. Figure 4(a) shows that changes in the detergent micelle size cause variation in the predicted \( \Delta S_{\text{pm}}(q) \), but nevertheless the major features of the difference x-ray scattering are consistent for all micelle sizes. Singular value decomposition using Matlab extracted the mean [principle component, Fig. 4(b), blue line] and variation about the mean [second SVD component, Fig. 4(b), mustard line, scaled 100 fold for visibility] from the x-ray scattering predictions shown in Fig. 4(a). What is striking is that the major variation about the mean in \( \Delta S_{\text{pm}}(q) \) reflected by the second SVD component is highly correlated [Fig. 4(b), red line; Pearson correlation coefficient for 96% for \( q \geq 0.1 \text{ Å}^{-1} \)] with the principal SVD component of micelle-only fluctuations during the last 5 ns of a 20 ns long trajectory [Fig. 3(e)]. Thus, for all practical purposes, if one wishes to apply corrections to the x-ray scattering due to micelle fluctuations [the term of Eq. (7) involving \( \Delta S_{\text{pm}}(q) \), x-ray scattering data represented in Fig. 3(e)] or due to fluctuations associated with changes in the size of the detergent-micelle (extracted as the second SVD component when using Eq. (13) applied to a sequence of trajectories with varying micelle size, Fig. 4), the resulting correction curves are almost indistinguishable for \( q \geq 0.1 \text{ Å}^{-1} \). This observation is presumably due to x-ray scattering changes reflecting the intrinsic length-scale of the detergent micelle in both cases. The inclusion of both terms during minimization would therefore risk a nonphysical situation of both terms having large amplitudes but opposite signs and therefore almost canceling each other. For this reason, we used only one of these two corrections during structural fitting [that extracted as the micelle size was varied, mustard line Fig. 4(b)], with the consequence that it is not possible distinguish corrections due to the micelle size fluctuations or those due to micelle structural fluctuations.

Correction for the solvent excluded volume

Predictions for light-induced changes in the x-ray scattering from the protein and detergent micelle must be further corrected due to the influence of the solvent excluded volume [Eq. (11)]. CRYOSOL predicts the total x-ray scattering after correcting for the solvent excluded volume by incorporating empirical corrections to the x-ray scattering function for individual atoms and these corrections were developed against a wealth of experimental data from soluble proteins. A detergent micelle typically has a much lower electron density than proteins and it was unclear if the empirical solvent excluded volume correction developed for soluble proteins should also be applied to integral membrane proteins within a detergent micelle. We therefore again utilized molecular dynamics simulations in GROMACS to extract the solvent excluded volume correction for the protein plus micelle system. Specifically, in addition to simulations of the protein and micelle within a box filled with water, we ran parallel unrestrained simulations of a box with water alone. We then superimposed the protein plus micelle coordinates upon the simulation with water alone, selected only those water molecules that were within a sphere of 2 Å radius from an atom of the protein or micelle, and wrote out a pdb file of water molecules alone that fully encompassed the protein and detergent [illustrated in Fig. 5(a)]. From Eq. (9), the total scattering after the excluded volume was calculated whereby the term \( S_{\text{excl}}(q) \) was extracted from pdb files of only water molecules selected above; and \( \cos \theta_{\text{wtr}} \) was calculated according to Eq. (10). When cutting out spheres of water molecules on all atoms in the protein-micelle complex, there will be an over-counting due to some water atoms which protrude beyond the protein’s surface. We, therefore, compensated for over counting the selected water molecules by varying the sphere radius between 1 and 4 Å, and counting the number of water molecules at each radii. From this result we created a linear regression to
Convolution with the undulator spectrum

Since polychromatic radiation was used to probe light-induced structural changes in bR, it was essential to account for the effect of the undulator spectrum on the x-ray scattering predictions. This effect was incorporated into the analysis by interpolating all x-ray scattering predictions onto a variable q-axis which varied inversely with the x-ray energy, and summing these interpolated x-ray scattering predictions according to the amplitude of the undulator spectrum for any given x-ray energy.

Structural fitting to the experimental data

Having developed the above tools and approximations, we searched for a best fit to the first and second experimental basis spectra [Fig. 2(b)] by varying the Cz perturbations according to the parameters γ and δ of Eq. (12). As in our earlier approaches, structural refinement minimized the function,

\[
R = \frac{\sum w(q) \cdot [D_{\text{theory}}(q) - D_{\text{exp}}(q)]^2}{\sum w(q) \cdot D_{\text{exp}}(q)^2},
\]

(14)

for each of the two basis spectra [states 1 and 2, Fig. 1(b)]; w(q) is a weighting factor introduced to emphasize regions where the structural content of the difference x-ray scattering data are strongest; and \(D_{\text{exp}}(q)\) is the experimental data for each of the two states. The term \(\Delta D_{\text{theory}}(q)\) is extracted for each value of γ and δ [Eq. (10)] using the following expression:

\[
\Delta D_{\text{theory}}(q) = \frac{S_{\text{sol}}(q)}{S_{\text{pm}}(q)} \cdot \exp \left( -\frac{B_2 q^2}{16 \pi} \right) \cdot \left[ \Delta S_{p}(q) + 1 - \text{erf} \left( \frac{q_{pm}}{2} \right) \right]
\]

\times \left( \Delta S_{\text{pm}}(q) - \Delta S_{\text{mp}}(q) + A_2 \cdot \Delta S_{\text{mc}}(q) \right)
\]

(15)

where the ratio \(S_{\text{sol}}(q)/S_{\text{pm}}(q)\) is used to correct for the influence of the solvent excluded volume [Fig. 3(c), Eq. (11)]; where \(\Delta S_{p}(q)\) is calculated directly from the atomic coordinates for the protein extracted from restrained molecular dynamics trajectories 5 ns in duration [Fig. 3(a)]; \(\Delta S_{\text{pm}}(q)\) is extracted from the atomic coordinates of the protein and micelle using Eq. (13) [Fig. 3(o)]; \(B_2\) is introduced in analogy to the crystallographic B-factor\(^2\) and reflects a loss of scattering power at higher angle arising from fluctuations about the mean atomic positions of the atoms of the protein; erf is the error function and the parameter \(q_{pm}\) was varied during structural refinement and \(\Delta q_{pm}\) was set as 25% of \(q_{pm}\), and this parameter quantifies the x-ray scattering angle above which the change in x-ray scattering from the micelle may become negligible due to the highly disordered nature of the detergent micelle; and \(\Delta S_{\text{mc}}(q)\) is a micelle correction term introduced to account for either low-angle corrections due to micelle fluctuations [Fig. 3(o)] or first order-corrections to the protein-micelle scattering term due to variations in the number of detergent molecules per micelle [Fig. 4(b)], used here throughout structural fitting) and the amplitude, \(A_2\), was also varied during refinement. For any given 5 ns restrained molecular dynamics trajectory, 100 structural pairs (of 501) were selected with the highest Pearson-correlation scores calculated relative to the experimental data over the domain \(q \geq 0.20 \text{ Å}^{-1}\), and Cz movements associated with.

count the number of water molecules at 0 Å [Fig. 5(b)] and the amplitude of \(S_{\text{sol}}(q)/S_{\text{pm}}(q)\) was sensitive to the details of this procedure, both the GROMACS and the CRYSTAL protocols yielded very similar shaped corrections curves for \(q \geq 0.15 \text{ Å}^{-1}\) [Fig. 5(c)] and consequently structural fitting (which treats the overall scaling as a parameter to be optimized) is not very sensitive to those details. Below \(q = 0.15 \text{ Å}^{-1}\), the GROMACS approach predicts a smaller contrast between the protein and micelle and, as such, the approach developed here provides a low-angle correction when considering a protein-micelle complex.

FIG. 5. Influence of the solvent excluded volume of the protein and detergent micelle on the predicted difference x-ray scattering curves. (a) Cartoon illustrating how the surface calculated from selected water molecules (transparent surface) used for solvent exclusion calculations fully surrounds the protein and detergent micelle. (c) Comparison of the solvent exclusion correction expressed as the ratio \(S_{\text{sol}}(q)/S_{\text{pm}}(q)\) [see Eq. (11)] when calculated directly for the protein plus detergent micelle coordinates using CRYSTAL (blue line); and from Eqs. (10) and (11) (red line) where CRYSTAL is used to calculate \(S_{\text{pm}}(q)\) from the protein plus micelle, and \(S_{\text{sol}}(q)\) is calculated from water molecules selected in (a) (blue line is scaled by a factor of 2.1 to best map to the red line). (b) Plot of the number of water molecules selected (blue circles) when using sphere radii from 1.0 to 4.0 Å, whereas 4331 water molecules were selected when using a 2 Å radius, the relative scaling between \(S_{\text{sol}}(q)\) and \(S_{\text{pm}}(q)\) in Eqs. (10) and (11) was determined by projecting this plot back to an infinitesimal radius (red line, 0 Å corresponds to 3215 water molecules).
with these 100 pairs were used in quantifying the most probable conformational changes. Overall, six parameters were varied during refinement ($\gamma$, $\delta$, $A_1$, $A_2$, $B_p$, $q_{pm}$), with $A_1$ optimized for both states simultaneously and, as noted below, it was not necessary to optimize $B_p$.

Results of fitting to the experimental data using this procedure are illustrated in Figs. 6(a) and 6(b) as a contour plot showing the changes in R-factor [Eq. (12)], calculated with the weighting function $w(q)$ set to unity for all $q$ as $\gamma$ and $\delta$ are varied for states 1 and 2, respectively. The minima in these contour plots are identified for state 1 at $\gamma = 0.5 \pm 0.2$ and $\delta = 3 \pm 2.5$, where the uncertainty is estimated as the change needed to increase the R-factor by 25% of R-minimum on the contour plot [Figs. 6(c) and 6(d)]. For state 2, the corresponding values were $\gamma = 1.0 \pm 0.2$ and $\delta = 3 \pm 1.5$. Whereas the parameter $\delta$ describing the movement of helix C seems large and very poorly defined, from the selected sub-set of optimal structures we find that the mean $Cz$ displacements on helix C were only $0.4 \pm 0.3$ Å and $0.7 \pm 0.3$ Å for state 1 and state 2, respectively, when calculated from residues 81 to 89 of helix C. By comparison, the crystallographic structures 5B6Z vs 5B6V gives a mean $Cz$ displacement of 0.46 Å for these residues of helix C. Similarly, the mean $Cz$ displacement from residue 152 of helix E to residue 182 of helix F were $1.6 \pm 0.6$ Å and $3.0 \pm 0.6$ Å (Fig. 7) when calculated for the sub-set of optimal structures for state 1 and state 2, respectively. For comparison, pdb entry 6RPH vs 6RQP gives a mean $Cz$ displacement of 3.4 Å (after adjustments for disordered regions described above) for the selected residues of helices E and F.

The micelle damping parameter $q_{pm}$ was refined as 0.129 Å$^{-1}$ as this value led to the lowest R-factor [Eq. (14)], and this value which can be equated with a B-factor of 8000 Å$^2$. Using $B = 8\pi^2(\mu^2)$, where $\mu$ is the amplitude of mean atomic displacement about the median position, this estimates random atomic motions within the membrane as the order of 10 Å, which is compatible with the motions of detergent micelles during the course of a simulated trajectory. By contrast the best fit was recovered when $B_p$, the damping parameter for the protein alone, was made arbitrary small and it was not optimized further. We,

FIG. 6. Results of structural fitting against difference x-ray scattering data when minimizing the R-factor [Eq. (14)] using the model summarized in Eq. (15). Protein conformations were sampled from restrained molecular dynamics simulations of 5 ns in duration, where Eq. (12) defines $\gamma$ and $\delta$. (a) Contour surface showing the R-factor recovered for a given displacement of helices EF ($\gamma$) and helix C ($\delta$) [Eq. (12)] when fitted against the state 1 basis spectrum [Fig. 1(b), black line]. (b) Contour surface showing the R-factor recovered for a given displacement of helices EF ($\gamma$) and helix C ($\delta$) when fitted against the state 2 basis spectrum [Fig. 1(b), black line]. All fits in panels (a) and (b) apply the same (optimal) scaling ratio, $A_1$, defined in Eq. (14). (c) Best-fit x-ray scattering prediction (red line) superimposed upon the experimental data (black circles) for state 1. The blue line resulted when the fitting procedure was repeated without micelle and cross term corrections. (d) Best-fit x-ray scattering prediction (red line) superimposed upon the experimental data (black circles) for state 2. Experimental data were originally presented in Ref. 12. The blue line resulted when the fitting procedure was repeated without micelle and cross term corrections.
However, the signal-to-noise ratio of state 2 is using a weighting of unity over the 33.5% and 20.4% being recovered for states 1 and 2, respectively, when fitting using rigid-motions.12

We observe that the structural fluctuations within the protein when fitting using rigid-motions.12

FIG. 7. Light-induced conformational changes in bR predicted by fitting against the difference x-ray scattering basis spectra. (a) Structural results from the fitting protocol developed here. The BR resting conformation is illustrated in purple; the best-fit model for state 1 is illustrated in pink; and the best-fit model for state 2 is illustrated in gray/blue. (b) Structural results from the fitting protocol previously developed in Ref. 12. The BR resting conformation is illustrated in purple; the best-fit model for the intermediate conformational state in Ref. 12 is illustrated in pink; and the best-fit model for the late conformational state in Ref. 12 is illustrated in gray/blue. (c) Light-induced displacements of C in C-atoms for the two modeling protocols. The amplitude of the atomic displacement of C-atoms is plotted vs residue number for state 1 (blue) and state 2 (purple) using the fitting model developed here; and the intermediate (mustard) and late (red) conformational states of Ref. 12.

therefore, conclude that the structural fluctuations within the protein were quite well described by structural variations sampled during the C2-restrained molecular dynamics simulations and the fitting routine did not benefit from the parameter B0, which was otherwise essential when fitting using rigid-motions.12

Figures 6(c) and 6(d) superimpose the best-fits from this procedure onto the experimental basis-spectra. Optimal predictions for ΔSmicel( q) from Eq. (15) [red line, Figs. 6(c) and 6(d)] appear to capture all aspects of the experimental data ΔSexp(q) for both basis spectra for state 1 and 2. However, the signal-to-noise ratio for state 2 is better than that of state 1, and this reflects in the R-factor [Eq. (15)] of 33.5% and 20.4% being recovered for states 1 and 2, respectively, when using a weighting of unity over the q-domain 0.15 Å ≤ q ≤ 0.9 Å over which the structural fitting was performed. If we remove the micelle-correction terms deriving from Eq. (13) [Fig. 3(f)] and that emerging from variations in the detergent micelle structure [Fig. 4(b)], then the resulting fits at lower x-ray scattering angle are compromised [red line, Figs. 6(c) and 6(d)] with the R-factors increasing to 52.5% for state 1 and 29.6% for state 2. For comparison, the R-factors reported in Ref. 12 for the intermediate and late states were 32% and 28% respectively, but these were calculated over the smaller q-domain 0.19 Å ≤ q ≤ 0.7 Å.

DISCUSSION

Over the last decade, TR-XSS studies of membrane protein structural changes have become a mature experimental method.12,18,24,28,29 Despite variations in detergent concentrations, variations in photo-excitation levels, difference in signal-to-noise ratios, and other factors associated with sample preparation or the execution of an experiment, the measured difference XSS basis spectra from light-sensitive integral membrane proteins are highly reproducible [Fig. 2(b)].

A sequence of time-dependent difference-XSS measurements from detergent solubilized membrane proteins [Figs. 2(c) and 2(d)] reveal if secondary structural changes occur and it is usually quite straightforward to extract the rise and decay times for a sequence of structural states. While structural models have been presented from this body of work, the theoretical foundations underpinning the modeling of different XSS data from integral membrane proteins have been somewhat glossed over due to challenges associated with detergent micelles that both fluctuate dramatically and often scatter x-rays more strongly than the protein of interest.

This analysis explicitly addresses this challenge by creating a framework that allows an atomistic description of the protein and detergent micelle using the molecular dynamics package GROMACS.38,41 As previously, the first step is to identify candidate motions of secondary structural elements. In this analysis, these motions were extracted from recent time-resolved serial crystallography studies of bacteriorhodopsin,12,34 which makes bacteriorhodopsin a particularly favorable case. It is, however, possible to identify candidate motions from molecular dynamics trajectories13–15,21,23,24,27–29 or potentially building approximate models using experimental results from other biophysical methods.42 Using selected candidate motions to guide molecular dynamics simulations, we identified how structural fluctuations within the detergent micelle influenced the different x-ray scattering curves [Figs. 3(b)–3(e)] and how changes in the detergent micelle size influenced the difference x-ray scattering (Fig. 4). Since both corrections were similar, they were incorporated as a single micelle correction to the predicted x-ray scattering [ΔSmicel( q)] of Eq. (15). The most important conceptual advance of the new approach was to interchange the atomic coordinates of the protein micelle when comparing molecular dynamics trajectories for resting and photo-activated structures [Figs. 1(d)–1(g)]. This allowed a formalism to be established in which the protein-micelle x-ray scattering cross term could be determined without being dominated by structural fluctuations within the micelle [Eq. (13), Fig. 3(f)]. Damping corrections were introduced for the protein-micelle systems [Eq. (15)], solvent contrast corrections were required (Fig. 5), and the x-ray spectrum of the pink-beam undulator had to be explicitly incorporated into the analysis, all of which are consistent with our earlier formalism.12,18,24,25

Structural conclusions drawn from this approach to molecular fitting of difference x-ray scattering data [Fig. 6] are largely consistent with those previously reported (Fig. 7). The mean Cα displacement helix C was found to be 0.4 ± 0.3 Å and 0.7 ± 0.3 Å for state 1 and
state 2 in this structural analysis, whereas our previous fitting against TR-XSS data\textsuperscript{12} yielded the values 0.9 and 1.3 \AA\ for the intermediate and late conformational states, respectively. It seems plausible that the uncertainties in the amplitude of this displacement are large since the Cz coordinate fluctuations throughout the protein within our restrained molecular dynamics simulations are the order of 0.2 \AA\ (Fig. 6). For the movement of the cytoplasmic portions of helix E and F, we recovered a mean Cz displacement of 1.6 \pm 0.6 \AA\ and 3.0 \pm 0.6 \AA\ for state 1 and state 2, respectively, whereas we previously concluded that these amplitudes were 2.0 and 2.6 \AA\ for the intermediate and late conformational states of bR.\textsuperscript{12} We suggest that agreement within uncertainties in the displacement associated with state 2 and the earlier late conformational state is acceptable.

Time-resolved serial crystallography\textsuperscript{4-6} is a rapidly growing field of research and TR-XSS has been established for more than a decade.\textsuperscript{11} The theoretical approach presented here underpins earlier assumptions when working with integral membrane proteins\textsuperscript{4,5,12,13,15-20} that the presence of detergent micelles does not necessarily affect structural fitting protocols that focus attention upon data in domain $q > 0.2 \text{ \AA}^{-1}$. Moreover, this analysis improves the fit to low scattering angle data [Figs. 5(e) and 5(f)] and justifies the assumption that this region is dominated by the influence of the detergent micelle. With the widespread use of rapid x-ray detectors at synchrotron radiation sources\textsuperscript{44,45} and the development of appropriate microradiographic mixing technologies for reaction initiation, it should become possible to extend the field of different XSS studies of integral membrane proteins to incorporate protein and substrate or reagent mixing studies to initiate time-resolved studies. Since caged compounds have also been successfully used for such studies,\textsuperscript{4,5,20} the sub-field of TR-XSS has the potential to grow beyond the study of naturally light-driven systems. Because all future experimental developments require solid theoretical foundations and because the detergent micelle may itself be influenced by mixing, the analysis tools and fitting protocols developed here should provide a framework that allows the field to look to the future with confidence.

MATERIALS AND METHODS
Sample preparation

bR was solubilized in \(\beta\)-octylglucoside and concentrated to 1.3 mM, as judged from the bR absorption peak at 570 nm, as described in more detail earlier in studies.\textsuperscript{12}

TR-XSS data acquisition and processing

TR-XSS data were collected at beamline ID09 of the European Synchrotron Radiation Facility (ESRF). Samples bR concentrated to 35 mg/ml were delivered across a polychromatic x-ray beam produced from an x-ray undulator (18 keV, pink beam, \(\Delta E/E \approx 4\%\)) through a quartz capillary (0.5 mm diameter, Hampton Research). Samples were driven using a motorized syringe pump (neMESYS) with a flow rate of 3 \mu\text{L}/s. X-ray pulse trains of 1 or 5 \mu\text{s} in duration were isolated from a continuous filling of the storage ring using an x-ray chopper. X-ray scattering images were recorded on a Rayonix MX170-HS detector (1920 \times 1920 pixels, pixels binned 2 \times 2 to a pixel-size of 88.5 \mu\text{m}) located 350 mm from the sample position. X-ray scattering data were collected at room temperature (\textdegreeC 20\textdegreeC). Data were merged over the domain 0.025 \text{ \AA}^{-1} < q < 3.0 \text{ \AA}^{-1}.

Samples of bacteriorhodopsin were photo-activated using a 532 nm laser pulse 4 ns in duration with 500 \mu\text{J}/pulse focused into a spot size of 0.31 \times 1.7 mm\textsuperscript{2} (full width half maximum). The triggering and timing of the green laser pulse and the x-ray detector were integrated into the beamline control system. X-ray scattering data were integrated in concentric rings\textsuperscript{12} prior to normalization about the isobestic points at $q = 1.6 \text{ \AA}^{-1}$. Outlier rejection was performed in two steps: the ring-integrated absolute scattering curves were rejected if they deviated more than 10\% from the median value in the q-range 2.0–2.5 \text{ \AA}^{-1}, followed by an outlier rejection scheme where difference scattering curves (ON minus OFF) where rejected if they deviated by more than three standard deviations from the mean of a given run. Typically, 4\%–10\% of all the data were rejected, depending upon the experimental run. The influence of laser-induced heating was measured and removed using data recorded using an IR laser (\(\lambda = 1470\) nm) to heat the sample. At this wavelength heat was deposited, the IR light did not photo-isomerize the retinal chromophore. Heating data were collected and were processed in the same fashion as described for the photo-activated datasets. Heat free light-induced difference scattering data were then computed by scaling and removing the thermal signal from each time-point, where the scaling factor is chosen to minimize the difference x-ray scattering differences over the domain 0.51 \text{ \AA}^{-1} \leq q \leq 1.45 \text{ \AA}^{-1}. Data after heat-removal are presented in Figs. 2(c) and 2(d) and were analyzed using two-component fitting based upon an exponential decay of the first component and a complementary rise of the second component. Experimental details for the 2009 studies performed at the same beamline have been described previously.\textsuperscript{12}

Molecular dynamics simulations using GROMACS

Monomeric bacteriorhodopsin protein/micelle complexes were built into a \(\beta\)-octylglucoside membrane using CHARMM-GUI Micelle Builder.\textsuperscript{44,45} A range of protein/micelle complexes were built, including 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, and 230 \beta\-octylglucoside molecules, with simulations using 190 \(\beta\)-octylglucoside molecules being preferred for most of the analysis. Lipid ring and protein surface penetration was checked by CHARMM as part of the system building options. The system was built within a solvent box consisting of 20 \AA\ in all dimensions, using three-site CHARMM modified TIP3P water model and a concentration of 0.15 M of NaCl. GROMACS input files were created by CHARMM.\textsuperscript{46,47} GROMACS 2019.1\textsuperscript{17-28} was used to run restrained molecular dynamics simulations of bR using the CHARMM36 force-field.\textsuperscript{48} Energy minimization was carried out first using LINCS constraints algorithm with constraints on hydrogen bonds. The second stage was equilibration of the system in six steps with energy and dihedral restraints on the \(\beta\)-octylglucoside molecules. No movement of the protein backbone was allowed during the equilibration and production steps. Molecular dynamics trajectories of 20 ns in duration (preparative runs) or 5 ns (fitting runs) were used to produced 2000 or 500 pdb files, respectively, each pdb file separated by 10 ps. These pdb files were used to calculate the theoretical scattering terms for protein, membrane, and protein plus membrane. Un-restrained molecular dynamics simulations involving only water were also performed using standard settings in GROMACS 2019.1.
Theoretical difference x-ray scattering curves were calculated using CRYSOL.\textsuperscript{24} Data and in-house code are available at Ref. 49.

ACKNOWLEDGMENTS
R.N. acknowledges funding from the Swedish Research Council (Grant No. 2015-00560) and the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No. 789030). We thank Giorgia Ortolani, Swagatha Ghosh, and Robert Dods for assistance during data collection and Anna Reymer, David Van der Spoel, and Magnus Andersson for advice when using GROMACS. We acknowledge support from the Partnership for Soft Condensed Matter (PSCM) during experiments at the ESRF.

AUTHOR DECLARATIONS
Conflict of Interest
The authors have no conflicts to disclose.

Author Contributions
Daniel Sarabi: Data curation (lead); Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (supporting); Writing – review & editing (supporting). Lucija Ostojic: Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). Robert Bosman: Data curation (supporting); Formal analysis (supporting); Methodology (supporting); Software (supporting). Adams Vallejos: Data curation (supporting); Software (supporting). Johanna-Barline Linse: Methodology (supporting); Software (supporting). Michael Wulff: Data curation (supporting); Methodology (supporting). Matteo Levantino: Data curation (supporting); Methodology (supporting). Richard Neutze: Conceptualization (lead); Data curation (supporting); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (lead); Project administration (lead); Resources (lead); Software (equal); Supervision (lead); Validation (equal); Visualization (equal); Writing – original draft (lead); Writing – review & editing (lead).

DATA AVAILABILITY
The data and in-house code that support the findings of this study are openly available in GitHub at https://github.com/Neutze-lab/TR_XSS, Ref. 49.

REFERENCES
1. K. Moffat, “Laue diffraction and time-resolved crystallography: A personal history,” Philos. Trans. R. Soc., A 377, 20180243 (2019).
2. H. N. Chapman et al., “Femtosecond x-ray protein nanocrystallography,” Nature 470, 73–77 (2011).
3. Tsenboer et al., “Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein,” Science 346, 1242–1246 (2014).
4. G. Branden and R. Neutze, “Advances and challenges in time-resolved macromolecular crystallography,” Science 373, eaba9954 (2021).
5. H. Poddar et al., “A guide to time-resolved structural analysis of light-activated proteins,” FEBS J. 280(3), 576–595 (2021).
6. J. Frank, “Time-resolved cryo-electron microscopy: Recent progress,” J. Struct. Biol. 200, 303–306 (2017).
7. J. Frank and A. Ourmazd, “Continuous changes in structure mapped by manifolds embedding of single-particle data in cryo-EM,” Methods 100, 61–67 (2016).
8. G. Pintet et al., “Real-time nuclear magnetic resonance spectroscopy in the study of biomolecular kinetics and dynamics,” Magn. Reson. 2, 291–320 (2021).
9. P. Haji et al., “Catalysis in the crystal: Synchrotron radiation studies with glycogen phosphorylase b,” EMBO J. 6, 539–546 (1987).
10. I. Ishigami et al., “Snapshot of an oxygen intermediate in the catalytic reaction of cytochrome c oxidase,” Proc. Natl. Acad. Sci. U. S. A. 116, 3572–3577 (2019).
11. M. Cammarata et al., “Tracking the structural dynamics of proteins in solution using time-resolved wide-angle x-ray scattering,” Nat. Methods 5, 881–886 (2008).
12. M. Andersson et al., “Structural dynamics of light-driven proton pumps,” Structure 17, 1265–1275 (2009).
13. S. Ahn, K. H. Kim, Y. Kim, J. Kim, and H. Ihee, “Protein tertiary structural changes visualized by time-resolved x-ray solution scattering,” J. Phys. Chem. B 113, 13131–13133 (2009).
14. Orfadd, F. et al., “Tracking the ATP-binding response in adenylate kinase in real time,” Sci Adv 7, eabi5514 (2021).
15. J. H. Lee and H. Ihee, “Advantages of time-resolved difference x-ray solution scattering curves in analyzing solute molecular structure,” Struct. Chem. 21, 37–42 (2010).
16. H. S. Cho et al., “Protein structural dynamics in solution unveiled via 100-ps time-resolved x-ray scattering,” Proc. Natl. Acad. Sci. U. S. A. 107, 7281–7286 (2010).
17. S. Westenhoff et al., “Rapid readout detector captures protein time-resolved WAXS,” Nat. Methods 7, 775–776 (2010).
18. E. Malmerberg et al., “Time-resolved WAXS reveals accelerated conformational changes in iodoreductin-stabilized proteorhodopsin,” Biophys. J. 101, 1345–1353 (2011).
19. K. H. Kim et al., “Direct observation of myoglobin structural dynamics from 100 picoseconds to 1 microsecond with picosecond x-ray solution scattering,” Chem. Commun. 47, 289–291 (2011).
20. I. Ramachandran et al., “The short-lived signaling state of the photoactive yellow protein photoreceptor revealed by combined structural probes,” J. Am. Chem. Soc. 133, 9395–9404 (2011).
21. K. H. Kim et al., “Direct observation of cooperative protein structural dynamics of homodimeric hemoglobin from 100 ps to 10 ms with pump-probe x-ray solution scattering,” J. Am. Chem. Soc. 134, 7901–7908 (2012).
22. T. W. Kim et al., “Protein structural dynamics of photoactive yellow protein in solution revealed by pump-probe x-ray solution scattering,” J. Am. Chem. Soc. 134, 3145–3153 (2012).
23. H. Takala et al., “Signal amplification and transduction in phytochrome photoreceptors,” Nature 509, 245–248 (2014).
24. D. Arnould et al., “Visualizing a protein quake with time-resolved x-ray scattering at a free electron laser,” Nat. Methods 11, 923–926 (2014).
25. E. Malmerberg et al., “Conformational activation of visual rhodopsin in native disc membranes,” Sci. Signaling 8, ra26 (2015).
26. T. W. Kim et al., “Combined probes of x-ray scattering and optical spectroscopy reveal how global conformational change is temporally and spatially linked to local structural perturbation in photoactive yellow protein,” Phys. Chem. Chem. Phys. 18, 8911–8919 (2016).
27. A. Björling et al., “Structural photoactivation of a full-length bacterial phytochrome,” Sci. Adv. 2, e1600920 (2016).
28. F. Oradl and M. Andersson, “Tracking membrane protein dynamics in real time,” J. Membr. Biol. 254, 51–64 (2021).
29. H. Ravishankar et al., “Tracking Ca\textsuperscript{2+} ATPase intermediates in real time by x-ray solution scattering,” Sci. Adv. 6, eaa09811 (2020).
30. S. A. Kynde et al., “Small-angle scattering gives direct structural information about a membrane protein inside a lipid environment,” Acta Crystallogr., Sect. D 70, 371–383 (2014).
31. E. Nango et al., “A three-dimensional movie of structural changes in bacteriorhodopsin,” Science 354, 1552–1557 (2016).
32 T. Weinert et al., "Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography," Science 365, 61–65 (2019).
33 I. N. Serdyuk, N. R. Zaccai, and J. Zaccai, Methods in Molecular Biophysics: Structure, Dynamics, Function (Cambridge University Press, Cambridge, 2007).
34 D. Svergun, C. Barberato, and M. H. J. Koch, "Crysol: A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates," J. Appl. Crystallogr. 28, 768–773 (1995).
35 R. J. D. Miller, O. Pare-Labrosse, A. Sarracini, and J. E. Besaw, "Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multiphoton regime and biological relevance," Nat Commun. 11, 1240 (2020).
36 M. L. Grunbein et al., "Illumination guidelines for ultrafast pump-probe experiments by serial femtosecond crystallography," Nat. Methods 17, 681–684 (2020).
37 D. Van Der Spoel et al., "GROMACS: Fast, flexible, and free," J. Comput. Chem. 26, 1701–1718 (2005).
38 S. Pronk et al., "GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit," Bioinformatics 29, 845–854 (2013).
39 J. Lipfert, L. Columbus, V. B. Chu, S. A. Lesley, and S. Doniach, "Size and shape of detergent micelles determined by small-angle x-ray scattering," J. Phys. Chem. B 111, 12427–12438 (2007).
40 H. Fischer, I. Polikarpov, and A. F. Craievich, "Average protein density is a molecular-weight-dependent function," Protein Sci. 13, 2825–2828 (2004).
41 M. J. Abraham et al., "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," SoftwareX 1–2, 19–25 (2015).
42 R. Biswas, H. Kuhne, G. W. Brudvig, and V. Gopalan, "Use of EPR spectroscopy to study macromolecular structure and function," Sci. Prog. 84, 45–67 (2001).
43 J. Johnson et al., "Elger: A single-photon counting x-ray detector," J. Instrum. 9, C05032 (2014).
44 X. Cheng, S. Jo, H. S. Lee, J. B. Klauda, and W. Im, "CHARMM-GUI: micelle builder for pure/mixed micelle and protein/micelle complex systems," J. Chem. Inf. Model. 53, 2171–2180 (2013).
45 S. Jo, T. Kim, V. G. Iyer, and W. Im, "CHARMM-GUI: A web-based graphical user interface for CHARMM," J. Comput. Chem. 29, 1859–1865 (2008).
46 R. R. Brooks et al., "CHARMM: The biomolecular simulation program," J. Comput. Chem. 30, 1545–1614 (2009).
47 J. Lee et al., "CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field," J. Chem. Theory Comput. 12, 405–413 (2016).
48 J. Huang and A. D. MacKerell, Jr., "CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data," J. Comput. Chem. 34, 2135–2145 (2013).
49 D. Sarabi, L. Ostojic, A. Vallejos, and R. Neutze, see https://github.com/Neutze-lab/Tr_XSS for the data and inhouse code that support the findings of this study, 2022.
50 D. S. Molodenskiy, H. D. T. Mertens, and D. I. Svergun, "An automated data processing and analysis pipeline for transmembrane proteins in detergent solutions," Sci. Rep. 10, 8081 (2020).