



Correction of rhodopsin serial crystallography diffraction intensities for a lattice-translocation defect

Matthew J. Rodrigues,^{a,*} ‡ Cecilia M. Casadei,^{a,b} ‡ Tobias Weinert,^a Valerie Panneels^a and Gebhard F. X. Schertler^{a,b,*}

Received 1 December 2022

Accepted 1 February 2023

Edited by R. J. Read, University of Cambridge, United Kingdom

‡ These authors contributed equally.

Keywords: serial crystallography; lattice-translocation defects; membrane proteins; lipidic cubic phase; G protein-coupled receptors.

PDB references: bovine rhodopsin in lipidic cubic phase, dark state, SACL A, 7zbc; dark state, SwissFEL, 7zbe; 1 ps light-activated, 8a6c

Supporting information: this article has supporting information at journals.iucr.org/d

^aLaboratory of Biomolecular Research, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland, and ^bDepartment of Biology, ETH-Zurich, Zurich, Switzerland. *Correspondence e-mail: matthew.rodrigues@psi.ch, gebhard.schertler@psi.ch

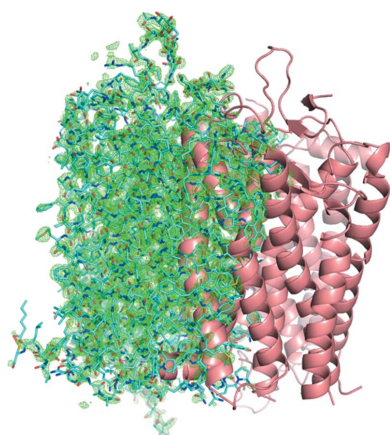
Rhodopsin is a G-protein-coupled receptor that detects light and initiates the intracellular signalling cascades that underpin vertebrate vision. Light sensitivity is achieved by covalent linkage to 11-*cis* retinal, which isomerizes upon photo-absorption. Serial femtosecond crystallography data collected from rhodopsin microcrystals grown in the lipidic cubic phase were used to solve the room-temperature structure of the receptor. Although the diffraction data showed high completeness and good consistency to 1.8 Å resolution, prominent electron-density features remained unaccounted for throughout the unit cell after model building and refinement. A deeper analysis of the diffraction intensities uncovered the presence of a lattice-translocation defect (LTD) within the crystals. The procedure followed to correct the diffraction intensities for this pathology enabled the building of an improved resting-state model. The correction was essential to both confidently model the structure of the unilluminated state and interpret the light-activated data collected after photo-excitation of the crystals. It is expected that similar cases of LTD will be observed in other serial crystallography experiments and that correction will be required in a variety of systems.

1. Introduction

The primary event in mammalian vision is the absorption of a photon by 11-*cis* retinal, which is covalently linked to a lysine side chain in rhodopsin via a protonated Schiff base (Nakanishi, 1991). Upon light absorption, retinal isomerizes to the all-*trans* form and the receptor transitions through a number of spectroscopically distinct intermediate states (Lewis & Kligler, 1992; Mathies & Lugtenburg, 2000). Once in the active state, the G protein-coupled receptor (GPCR) catalyses the exchange of GDP for GTP in the transducin G protein to initiate intracellular signalling cascades that result in neuronal signalling (Bennett *et al.*, 1982; Emeis *et al.*, 1982).

Structural and spectroscopic studies have exploited the abundance of rhodopsin in bovine retina to investigate how it performs its critical role in vision. The initial resting-state crystal structure of rhodopsin was followed by further crystal and cryo-EM structures of the receptor in the active state, in complex with signalling partners and in cryo-trapped intermediate states following light activation (Palczewski *et al.*, 2000; Standfuss *et al.*, 2011; Tsai *et al.*, 2018, 2019; Nakamichi & Okada, 2006*a,b*). However, until recently it has not been possible to determine how the structure of the protein changes as a function of time at physiological temperatures.

In the past decade, time-resolved serial femtosecond crystallography (TR-SFX) experiments have shone light on the photochemistry driving the activity of light-sensitive proteins



OPEN ACCESS

Published under a CC BY 4.0 licence

Table 1

Merging statistics for dark-state data sets collected at SwissFEL and SACLA (Gruhl *et al.*, 2022).

Values in parentheses are for the highest resolution shell.

Data set	SwissFEL	SACLA
Resolution range (Å)	16.10–1.80 (1.86–1.80)	10.47–1.80 (1.86–1.80)
<i>a</i> , <i>b</i> , <i>c</i> (Å)	61.51, 91.01, 151.11	61.29, 90.81, 150.51
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Space group	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 22 ₁ 2 ₁
Measured reflections	65870940 (4371475)	105018485 (7436654)
Unique reflections	79305 (7852)	78209 (7715)
Multiplicity	830.6 (556.7)	1342.8 (963.9)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
$\langle I/\sigma(I) \rangle$	7.62 (0.95)	8.87 (1.32)
<i>R</i> _{split} (%)	8.21 (109.06)	6.75 (77.70)
CC*	0.9982 (0.8947)	0.9994 (0.9219)
CC _{1/2}	0.9926 (0.6672)	0.9977 (0.7391)
Translation vector (t_d)	0.245	0.243
Translated fraction	$\alpha_0 = 0.01$, $\alpha_1 = 0.19$	$\alpha_0 = 0.00$, $\alpha_1 = 0.15$

(Poddar *et al.*, 2022). In particular, the molecular mechanisms of several retinal-dependent microbial opsins, including the proton pump bacteriorhodopsin (Nango *et al.*, 2016; Nogly *et al.*, 2018; Nass Kovacs *et al.*, 2019; Weinert *et al.*, 2019), the sodium pump KR2 (Skopintsev *et al.*, 2020), the chloride pump NmHR (Yun *et al.*, 2021; Mous *et al.*, 2022) and the C1C2 channelrhodopsin cation channel (Oda *et al.*, 2021), have been deciphered. TR-SFX experiments with rhodopsins typically rely on the delivery of microcrystals embedded in a viscous medium to the interaction region (Nogly *et al.*, 2016; James *et al.*, 2019), where they are illuminated by an optical pulse. After a controlled time delay, the crystals are probed by an X-ray pulse from a free-electron laser (FEL). As only one, still diffraction image can be measured from an individual crystal before it is destroyed by the FEL X-ray pulse (Chapman *et al.*, 2014), a single data set is typically composed of images collected from tens of thousands of randomly oriented crystals.

While TR-SFX experiments are uniquely capable of visualizing protein conformational changes with atomic spatial and subpicosecond temporal resolution, the sample requirements are often challenging to fulfil. Crystallization in the lipidic cubic phase (LCP) has been the method of choice for most membrane proteins prepared for TR-SFX experiments (Landau & Rosenbusch, 1996), and LCP can also be used as a medium to deliver crystals to the X-ray beam via a high-viscosity sample injector (Weinert & Panneels, 2020).

Optimization of the bulk purification and LCP crystallization protocols for bovine rhodopsin yielded large quantities of well diffracting crystals that were suitable for time-resolved experiments at FELs (Wu *et al.*, 2015; Weinert & Panneels, 2020; Gruhl *et al.*, 2022).

It was possible to collect diffraction data sets from ‘dark-state’ crystals, which were not illuminated, and also from crystals 1, 10 and 100 ps after light activation at the Swiss Free Electron Laser (SwissFEL) and the SPring-8 Angstrom Compact Free-Electron Laser (SACLA) (Gruhl *et al.*, 2022; Table 1). An initial model of the dark-state structure could be obtained by molecular replacement, after which several

iterations of model building and refinement were carried out. Inspection of the electron-density maps revealed significant density features in the solvent channels that could not be successfully modelled by placing solvent molecules.

After in-depth analysis of the electron-density maps and diffraction intensities, we were able to identify the presence of a lattice-translocation defect within the rhodopsin crystals. Procedures to correct rotation data collected from single crystals exhibiting such defects have been described previously (Wang, Kamtekar *et al.*, 2005; Wang, Rho *et al.*, 2005). Here, we describe the procedure performed to detect, characterize and correct serial crystallography data sets for this pathology. The correction was essential to confidently model both the dark and light-activated structures of rhodopsin, and therefore to elucidate the first molecular rearrangements in vertebrate vision.

2. Data collection and initial processing

2.1. Data collection

Bovine rhodopsin was purified from the native source and crystallized in LCP as has been described in detail (Gruhl *et al.*, 2022). The microcrystals embedded in LCP were delivered to the X-ray beam by a high-viscosity sample injector (Weierstall *et al.*, 2014) and diffraction data were collected from crystals either without illumination by the pump laser (dark sample) or at a defined time delay after illumination. Data were collected as still images at SwissFEL and SACLA, with separate dark-state data sets measured at both FELs. Experimental parameters are detailed in Gruhl *et al.* (2022).

2.2. Initial processing

As described in Gruhl *et al.* (2022), all data collected at SwissFEL (SF dark, 1 ps, 10 ps) were indexed using *indexamajig* (White *et al.*, 2012) with the *XGANDALF* algorithm (Gevorkov *et al.*, 2019). The *MOSFLM*, *DirAx* and *XGANDALF* methods (Powell, 1999; Duisenberg, 1992; Gevorkov *et al.*, 2019) were used for the data collected at SACLA (SACLA dark, 100 ps). Diffraction images collected at SwissFEL were indexed in space group *P*22₁2₁ with unit-cell parameters *a* = 61.51, *b* = 91.01, *c* = 151.11 Å and data collected at SACLA were indexed in the same space group with unit-cell parameters *a* = 61.29, *b* = 90.81, *c* = 150.51 Å.

All data were integrated using *indexamajig* in *CrystFEL* (White *et al.*, 2012). The integration radius was set to two pixels for SwissFEL data and three pixels for SACLA data, while the background annulus was set to between four and six pixels for SwissFEL data and between four and seven pixels for data collected at SACLA. The crystal-to-detector distance was optimized by sampling detector distances between 91.5 and 97.5 mm at SwissFEL and 47.5 and 53.5 mm at SACLA first in 200 µm increments and then in 20 µm increments to determine the detector distance at which the standard deviations of the unit-cell parameters were minimized.

Initial phasing by molecular replacement was performed with *Phaser* (McCoy *et al.*, 2007) using a deposited structure of

Table 2

Refinement statistics for the SwissFEL dark-state model refined against the original and the corrected diffraction intensities.

	Original intensities	Corrected intensities
$R_{\text{work}}/R_{\text{free}}$ (%)	26.65/28.27	21.43/24.61
No. of atoms		
Protein	4971	4971
Ligand	526	526
Water	174	174
B factors (\AA^2)		
Protein	31.60	32.20
Ligand	46.77	49.08
Water	38.20	40.23
Ramachandran statistics		
Favoured (%)	96.48	96.81
Allowed (%)	3.52	3.19
Outliers (%)	0.00	0.00
R.m.s. deviations		
Bond angles ($^\circ$)	0.009	0.007
Bond lengths (\AA)	0.928	0.847

Table 3

Refinement statistics for the SACLA dark-state model refined against the original and the corrected diffraction intensities.

	Original intensities	Corrected intensities
$R_{\text{work}}/R_{\text{free}}$ (%)	22.38/23.68	19.50/22.11
No. of atoms		
Protein	4970	4970
Ligand	539	539
Water	168	168
B factors (\AA^2)		
Protein	30.48	31.16
Ligand	47.53	49.74
Water	38.13	39.88
Ramachandran statistics		
Favoured (%)	96.64	97.14
Allowed (%)	3.36	2.86
Outliers (%)	0.00	0.00
R.m.s. deviations		
Bond angles ($^\circ$)	0.005	0.009
Bond lengths (\AA)	0.738	0.977

rhodopsin at cryogenic temperature (PDB entry 1u19; Okada *et al.*, 2004), with solvent and ligand molecules removed, as a search model. The asymmetric unit contains two rhodopsin molecules arranged as an antiparallel dimer. Rotation of this antiparallel dimer around the crystallographic screw axes generates translational noncrystallographic symmetry (tNCS), which was detected as a peak with a magnitude 65% of the origin peak height at position $0.5\mathbf{a} + 0.378\mathbf{b} + 0.5\mathbf{c}$ in the native Patterson map with *phenix.xtriage* (Liebschner *et al.*, 2019). Multiple iterations of model building in *Coot* (Casañal *et al.*, 2020) and refinement with *phenix.refine* (Liebschner *et al.*, 2019) were carried out until convergence of the R_{free} statistics. Most metrics of model quality were reasonable at this stage (see Tables 2 and 3). In particular, model refinement against dark-state data collected at SwissFEL yielded R_{work} and R_{free} values of 26.65% and 28.27%, respectively, in the resolution range 16.1–1.8 Å, showing that the modelled structure explained the experimental observations well.

3. Indications of lattice-translocation disorder in real space

Although the R -factor statistics, which quantify the agreement between the observed and model structure factors, converged to values which are commonly considered to be acceptable in serial crystallography, inspection of the electron-density maps indicated that our model did not appropriately describe the underlying data. As is common practice, we calculated σ_A -weighted electron-density difference maps, where the Fourier coefficients $(mF_o - DF_c)\exp(i\varphi_c)$, with observed structure-factor amplitudes F_o and model amplitudes and phases F_c , φ_c , include weighting factors m and D to account for the coordinate errors within the model and the partiality of the modelled structure (Read, 1986). These maps are less affected by model bias compared with $(F_o - F_c)\exp(i\varphi_c)$ maps and are particularly useful to identify errors in the model and aid model building. In particular, positive difference peaks correspond to features of the electron density which are unaccounted for by the model.

Significant positive difference electron-density features were visible both in the solvent channels between rhodopsin molecules and overlapping with the rhodopsin dimer (Fig. 1). This difference density could not be sensibly modelled by the placement of solvent molecules. Furthermore, some of the observed features resembled amino-acid side chains (Fig. 2*a*). However, there was insufficient space between the protein molecules to accommodate an additional rhodopsin chain.

We were able to model a single tryptophan residue into a particularly prominent positive difference electron-density feature in the solvent channel (Fig. 2*a*) and then to least-squares align tryptophan-centred tripeptides from the rhodopsin structure to the placed tryptophan residue in *Coot* (Fig. 2*b*). In this way, we could assign this unexplained density feature to Trp265. Least-squares alignment of the entire rhodopsin dimer to this tryptophan residue resulted in the model overlapping well with the unexplained density (Fig. 2*c*). This second copy of the dimer was related to the original structure by a translation of 22.5 Å along the unit-cell \mathbf{b} axis, corresponding to approximately one-quarter of the length of the \mathbf{b} axis. As the original dimer and the translated copy spatially overlap, it is not possible for both copies to be present in the same unit cell.

This observation was reminiscent of several previous observations of lattice-translocation disorder (LTD), namely the presence of translation-related domains in the crystals, in a number of systems (Bragg & Howells, 1954; Pickersgill, 1987; Wang, Rho *et al.*, 2005; Hare *et al.*, 2009; Tsai *et al.*, 2009; Ponnusamy *et al.*, 2014; Li *et al.*, 2020). We therefore sought to determine whether this was another such case and whether the data could be corrected as previously described in Wang, Kamtekar *et al.* (2005).

As a first step, a Fourier transform of the merged intensities was performed to produce a Patterson map, which represents the autocorrelation function of the electron density in real space and exhibits prominent peaks at positions corresponding to highly frequent interatomic and intermolecular vectors in the structure. By inspecting this map, we observed a prominent peak at the position $0.5\mathbf{a} + 0.378\mathbf{b} + 0.5\mathbf{c}$, where

a, **b**, **c** are the unit-cell axes, which was accounted for by the noncrystallographic symmetry operation relating the two copies of the molecule in the asymmetric unit. We also observed a second prominent peak, with a magnitude of 18% of the origin peak for the SACLA data and 25% of the origin peak for the SwissFEL data, at $0.245\mathbf{b}$ (Fig. 3). We attributed this peak to the vector \mathbf{t}_d relating the two translation-related domains.

4. Indications of lattice-translocation disorder in reciprocal space

We assume that terms from translation-related domains corresponding to integer multiples of the fundamental translation \mathbf{t}_d contribute to the structure factors of an LTD-affected crystal. Following and generalizing the treatment of LTD presented in Wang, Kamtekar *et al.* (2005), we model such structure factors with the weighted sum of an infinite number of translation-related terms,

$$\mathbf{F}_o(\mathbf{h}) = \left\{ \sum_{n=0}^{\infty} \alpha_n \exp[i2\pi(n\mathbf{t}_d)\mathbf{h}] \right\} \mathbf{F}_d(\mathbf{h}), \quad (1)$$

where \mathbf{h} denotes the reciprocal-lattice point $h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$ with integer h, k, l and reciprocal-lattice basis vectors $\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*$. Real-space domain translations are given by the set $\{n\mathbf{t}_d\}$ of integer multiples of the fundamental translation \mathbf{t}_d , and the weights α_n represent the unit-cell fractions pertaining to each domain. The single-domain structure factor \mathbf{F}_d is given by the usual sum over contributions from all atoms within one unit cell,

$$\mathbf{F}_d(\mathbf{h}) = \sum_j f_j \exp(i2\pi\mathbf{x}_j\mathbf{h}), \quad (2)$$

where f_j is the atomic form factor and \mathbf{x}_j is the position of atom j in the unit cell. The structure factor of a translated domain is

$$\mathbf{F}_d^{\text{tr}} = \sum_j f_j \exp[i2\pi(\mathbf{x}_j + n\mathbf{t}_d)\mathbf{h}] = \exp(i2\pi n\mathbf{t}_d\mathbf{h}) \mathbf{F}_d(\mathbf{h}). \quad (3)$$

Inspection of the Patterson map provides the estimate $\mathbf{t}_d = t_d\mathbf{b} = 0.245\mathbf{b}$. Because high-order terms are progressively less relevant, we truncate the summation in equation (1) to the term with $n = 2$, obtaining

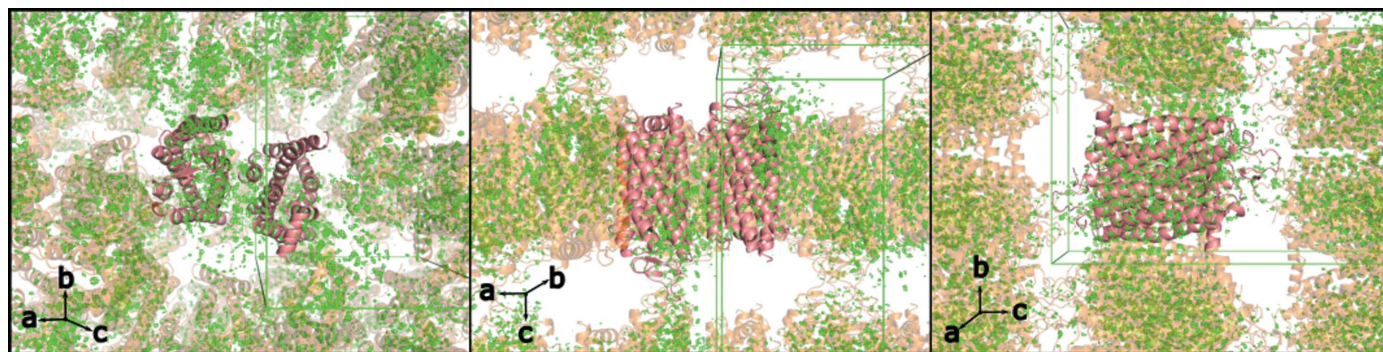


Figure 1 Rhodopsin dimer (pink) and symmetry-related rhodopsin molecules (pale orange) shown in cartoon format with σ_A -weighted $F_o - F_c$ difference density contoured at 4.0σ overlaid (green). The unit cell is shown as a green box.

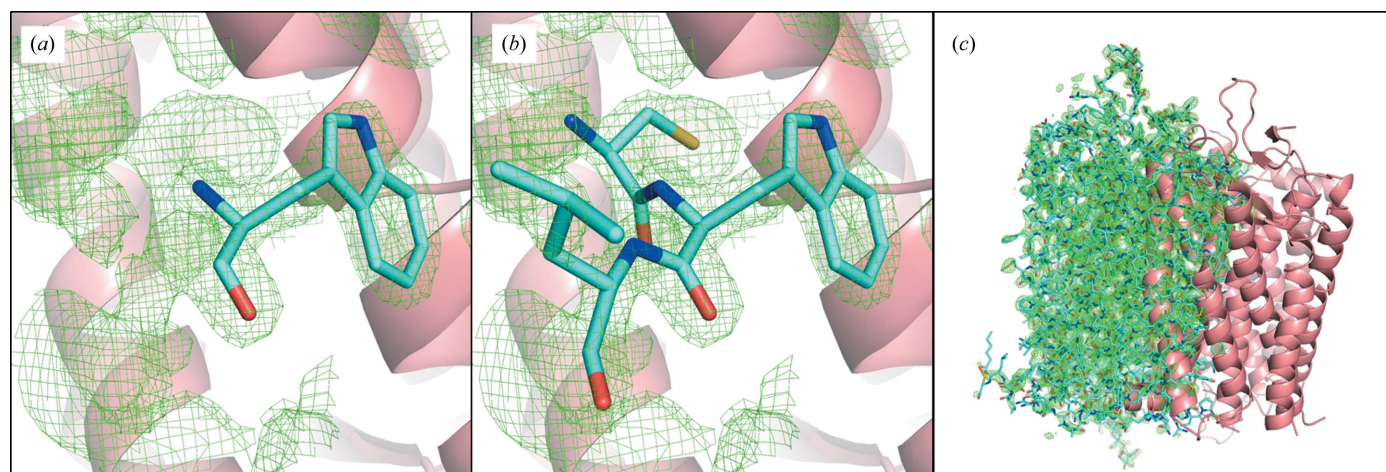


Figure 2 Rhodopsin dimer in cartoon format (pink) overlaid with σ_A -weighted $F_o - F_c$ difference density contoured at 2.0σ (green mesh). (a) Trp265 (cyan) fitted into σ_A -weighted $F_o - F_c$ difference density. (b) Cys264, Trp265 and Leu266 (cyan) aligned to the fitted Trp265. Transmembrane helices 5 (left) and 6 (right) are visible behind the translated peptide. (c) A translated rhodopsin dimer (cyan stick format) aligned with Trp265 overlaid with the main rhodopsin dimer.

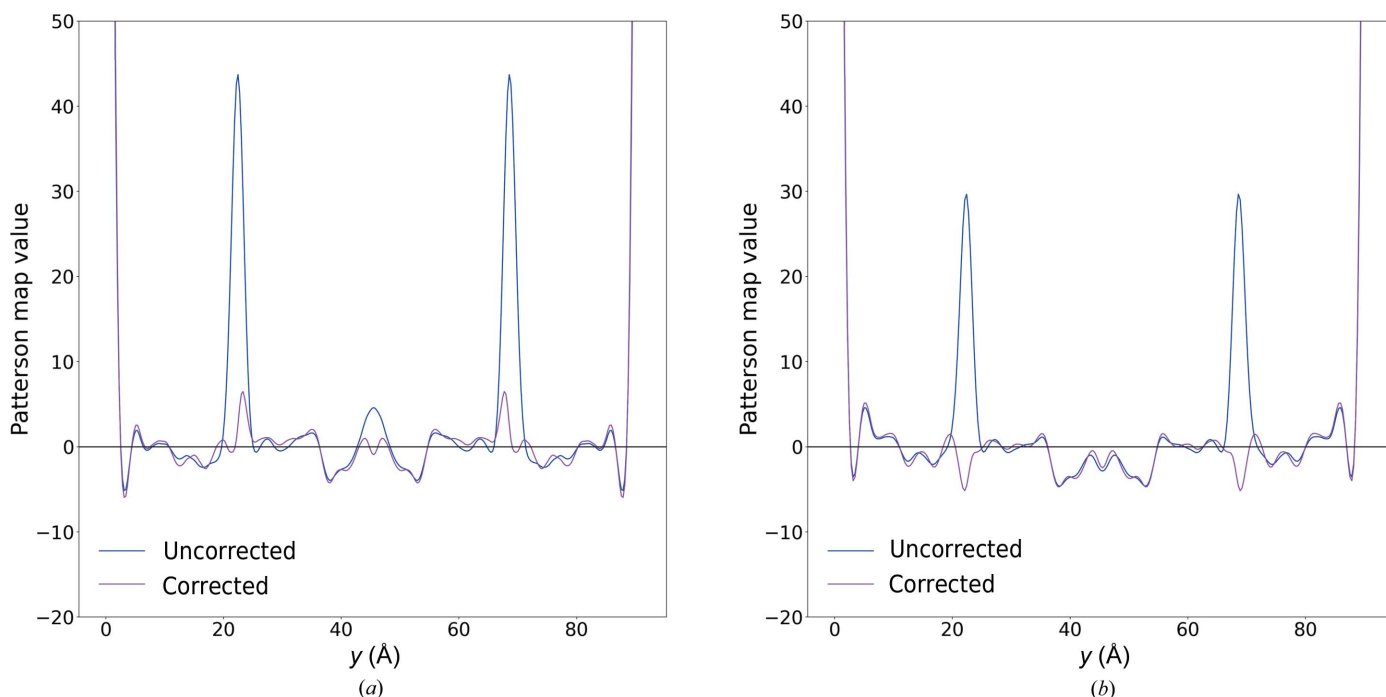


Figure 3 Patterson map values for SwissFEL data (a) and SACLA data (b) along the direction $0\hat{\mathbf{a}} + y\hat{\mathbf{b}} + 0\hat{\mathbf{c}}$, where $\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}}$ are unit vectors aligned with the corresponding unit-cell basis vectors before and after correction of the intensities with $\alpha_0 = 0.01, \alpha_1 = 0.19, \alpha_2 = 0.80$ (SwissFEL) and $\alpha_0 = 0.00, \alpha_1 = 0.15, \alpha_2 = 0.85$ (SACLA). The peak height at $0.245\mathbf{b}$ is equivalent to 25% and 18% of the origin peak heights for uncorrected SwissFEL and SACLA data, respectively.

$$\begin{aligned} \mathbf{F}_o(\mathbf{h}) &\simeq [\alpha_0 + \alpha_1 \exp(i2\pi t_d \mathbf{h}) + \alpha_2 \exp(i4\pi t_d \mathbf{h})] \mathbf{F}_d(\mathbf{h}) \\ &= [\alpha_0 + \alpha_1 \exp(i2\pi t_d k) + \alpha_2 \exp(i4\pi t_d k)] \mathbf{F}_d(\mathbf{h}). \end{aligned} \quad (4)$$

As a result of squaring the complex structure factors to obtain diffraction intensities $I_o = \mathbf{F}_o \mathbf{F}_o^*$, interference terms $(\alpha_0 \alpha_1 + \alpha_1 \alpha_2) [\exp(i2\pi t_d k) + \exp(-i2\pi t_d k)]$ and $\alpha_0 \alpha_2 [\exp(i4\pi t_d k) + \exp(-i4\pi t_d k)]$ appear. This gives rise to a characteristic oscillatory behaviour of the observed intensities as a function of reciprocal space index k ,

$$\begin{aligned} I_o(\mathbf{h}) &\simeq [\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + 2(\alpha_0 \alpha_1 + \alpha_1 \alpha_2) \cos(2\pi t_d k) \\ &\quad + 2\alpha_0 \alpha_2 \cos(4\pi t_d k)] I_d(\mathbf{h}), \end{aligned} \quad (5)$$

with $I_d = \mathbf{F}_d \mathbf{F}_d^*$. This oscillatory behaviour can be recognized in Fig. 4, where measured intensities, averaged over reciprocal space indices (h, l) , are shown as a function of index k .

5. Correction of diffraction intensities

The inspection and analysis of the electron-density maps, the Patterson map and the merged intensities supported the diagnosis of LTD within the rhodopsin crystals. Tentative domain fractions α_0, α_1 and α_2 were sampled with steps of 0.01 in the range $[0, 1]$. For each triplet $(\alpha_0, \alpha_1, \alpha_2)$ satisfying

$$\alpha_2 = 1 - \alpha_0 - \alpha_1 \quad (6)$$

with

$$0 \leq \alpha_0 \leq \alpha_1 \leq \alpha_2 \leq 1, \quad (7)$$

we estimated individual domain intensities by inversion of equation (5) as follows:

$$\begin{aligned} I_d(\mathbf{h}; \alpha_0, \alpha_1, \alpha_2) &\simeq [\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + 2(\alpha_0 \alpha_1 + \alpha_1 \alpha_2) \cos(2\pi t_d k) \\ &\quad + 2\alpha_0 \alpha_2 \cos(4\pi t_d k)]^{-1} I_o(\mathbf{h}). \end{aligned} \quad (8)$$

For each set of corrected intensities $I_d(\mathbf{h}; \alpha_0, \alpha_1, \alpha_2)$, we calculated the Patterson map value at positions $t_d \mathbf{b}$ and $2t_d \mathbf{b}$. The values of $|\tilde{I}_d(0, t_d b, 0)|$ and $|\tilde{I}_d(0, 2t_d b, 0)|$, where \tilde{I}_d is the Fourier transform of I_d , are reported in Figs. 5 and 6 for the SwissFEL and SACLA data, respectively, as a function of (α_0, α_1) . We optimized the domain fractions by selecting the

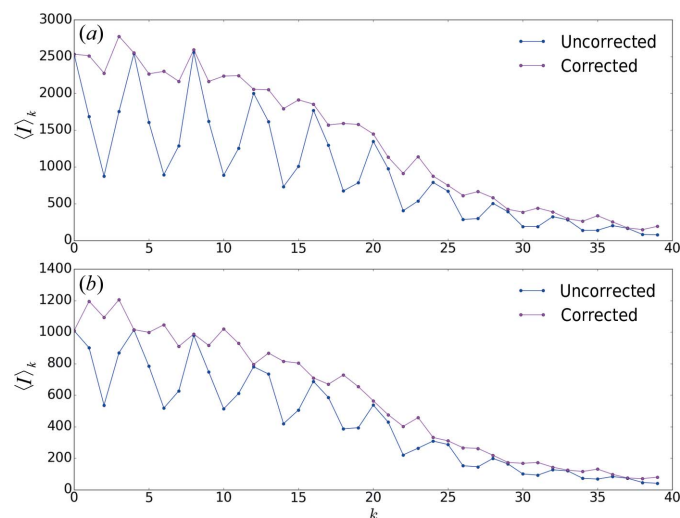


Figure 4 Diffraction intensities from SwissFEL (a) and SACLA (b), averaged over reciprocal-space indices h and l , as a function of index k before and after correction with $\alpha_0 = 0.01, \alpha_1 = 0.19, \alpha_2 = 0.80$ (SwissFEL) and $\alpha_0 = 0.00, \alpha_1 = 0.15, \alpha_2 = 0.85$ (SACLA).

triplet resulting in small Patterson peaks at both examined positions. We obtained $\alpha_0 = 0.01$, $\alpha_1 = 0.19$ and $\alpha_2 = 0.80$ for the SwissFEL data set and $\alpha_0 = 0.00$, $\alpha_1 = 0.15$ and $\alpha_2 = 0.85$ for the SACLA data set. Fig. 3 shows a comparison of the Patterson map values before and after correction along the direction of the unit-cell axis **b**. Upon correction of the intensities using the optimized values, a dampening of the oscillations of (h, l) -averaged values as a function of index k was observed (Fig. 4).

With the definitions

$$\beta_0 = \alpha_0^2 + \alpha_1^2 + \alpha_2^2, \quad (9)$$

$$\beta_1 = 2(\alpha_0\alpha_1 + \alpha_1\alpha_2), \quad (10)$$

and

$$\beta_2 = 2\alpha_0\alpha_2, \quad (11)$$

equation (5) can be rewritten

$$I_o(\mathbf{h}) \simeq [\beta_0 + \beta_1 \cos(2\pi t_d k) + \beta_2 \cos(4\pi t_d k)] I_d(\mathbf{h}), \quad (12)$$

with weights $\beta_0 = 0.6762$, $\beta_1 = 0.3078$, $\beta_2 = 0.016$ for the SwissFEL data and $\beta_0 = 0.745$, $\beta_1 = 0.255$, $\beta_2 = 0$ for the SACLA data. The ratios β_1/β_0 (0.46 and 0.34 for the SwissFEL and SACLA data, respectively) and β_2/β_0 (0.02 and 0 for the SwissFEL and SACLA data, respectively) express the relative importance of interference terms for increasing relative displacement and show that it is well justified to neglect terms with $n > 2$ in equation (1).

Translation correction factors were determined separately for each of the data sets. Refinement of the rhodopsin dark-state atomic model against the LTD-corrected intensities from SwissFEL yielded an immediate improvement in both the R_{work} and R_{free} statistics (21.43% and 24.61%, respectively) and the model geometry. More importantly, the amount of unexplained $mF_o - DF_c$ difference density was much reduced

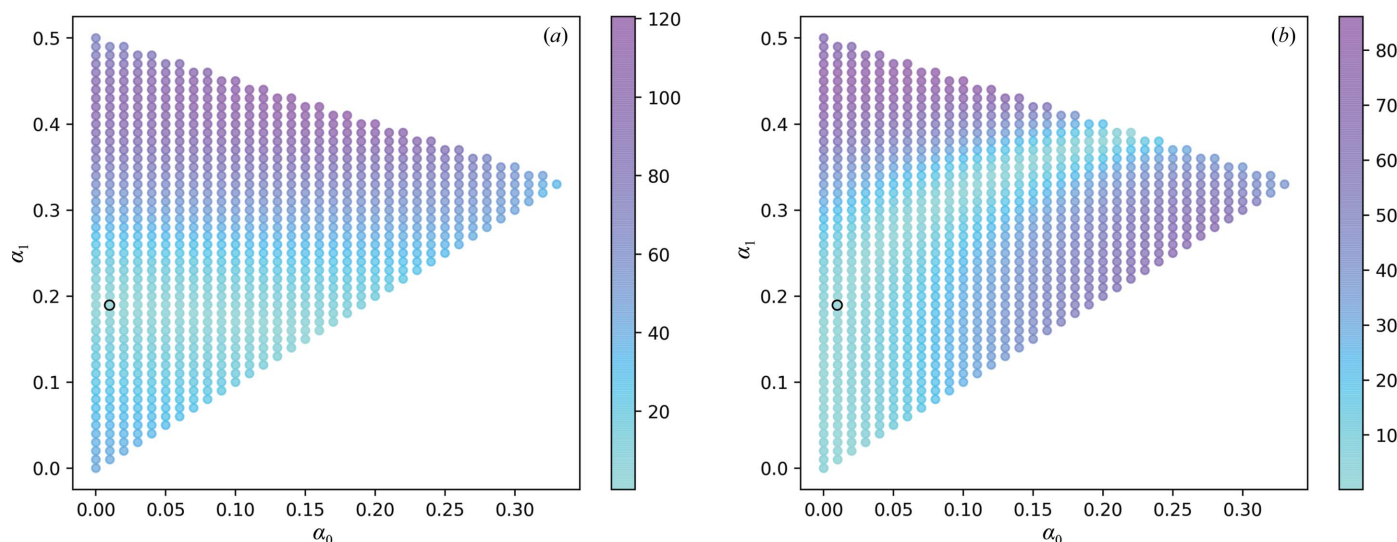


Figure 5 Absolute values of the Patterson function of corrected SwissFEL intensities with varying α_0 , α_1 and $\alpha_2 = 1 - \alpha_0 - \alpha_1$ at real-space positions $t_d\mathbf{b}$ (a) and $2t_d\mathbf{b}$ (b). Correction of intensities with $\alpha_0 = 0.01$, $\alpha_1 = 0.19$ (black circle) dampens the Patterson peaks at both positions simultaneously.

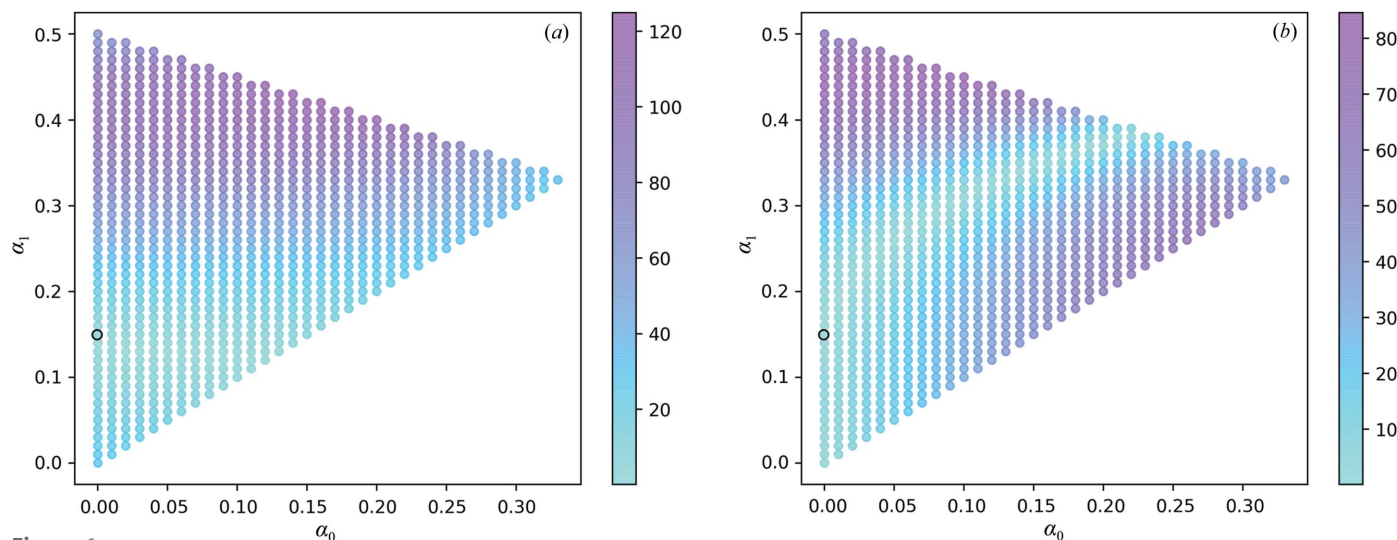


Figure 6 Absolute values of the Patterson function of corrected SACLA intensities with varying α_0 , α_1 and $\alpha_2 = 1 - \alpha_0 - \alpha_1$ at real-space positions $t_d\mathbf{b}$ (a) and $2t_d\mathbf{b}$ (b). Correction of intensities with $\alpha_0 = 0.00$, $\alpha_1 = 0.15$ (black circle) dampens the Patterson peaks at both positions simultaneously.

in the resulting electron-density maps. This allowed further model building to remove inappropriately placed solvent molecules, fix distorted amino-acid side-chain rotamers and build more of the flexible loops in the protein. The dark-state final models therefore accounted better for the experimental data. Improved model phases were available for calculation of the light-to-dark electron-density difference maps, with Fourier coefficients $(F_o^{\text{light}} - F_o^{\text{dark}}) \exp(i\varphi_c^{\text{dark}})$, derived from the observed amplitudes F_o^{dark} and F_o^{light} of the dark and light-activated state, respectively, and dark model phases φ_c^{dark} . These maps are essential to interpret the conformational changes following light activation.

There is a stark contrast between the $mF_o - DF_c$ maps produced after refinement of the final dark-state model against the diffraction intensities with and without correction, demonstrating the importance of the correction in obtaining an accurate dark-state model (Figs. 7 and 8). As the dark-state model serves as the foundation for interpreting light-induced structural changes, the correction described was required to confidently derive biological insights into the very first processes in receptor activation (Gruhl *et al.*, 2022).

6. Discussion

Several examples of LTD in crystals of soluble proteins have been published since the first observation in 1954 (Bragg & Howells, 1954). Furthermore, occurrences of LTD may be underreported, as this pathology is difficult to detect and may often be ignored when the translated fraction is small (Lovelace & Borgstahl, 2020). LTD has also been identified in LCP-grown crystals of the rice silicon transporter Lsi1 (Saitoh

et al., 2021). Similar to our observation of LTD in rhodopsin crystals, the Lsi1 lattice translation appears to occur parallel to the plane of the membrane. Indeed, lattice translations perpendicular to the plane of the LCP bilayer seem improbable as the hydrophobic transmembrane region of the protein would be displaced from the hydrophobic lipid environment to the aqueous environment between the lipid bilayers. Instead, type I LCP crystals, which assemble as stacks of 2D crystals, appear to be susceptible to translations in the relative positions of these 2D layers, which result in translations parallel to the plane of the membrane.

In the previously observed cases of LTD the data were collected using the rotation method (Dauter, 1999) and it was possible to diagnose the disorder by direct inspection of the diffraction images. In Wang, Kamtekar *et al.* (2005), for example, a translation of 0.5 in the *c* axis in *Bacillus* phage φ 29 DNA polymerase crystals resulted in the appearance of weak and streaky spots for reflections with odd values of *l* and strong spots for even values of *l*, caused by the main lattice and the translated lattice scattering out of phase and in phase with odd and even values of *l*, respectively.

Such an analysis is not possible with serial crystallography data, as the measured intensities from individual frames are partial and the intensity of a given reflection cannot be determined by inspection of a single diffraction image. Instead, it is necessary to search for non-origin peaks in the Patterson map that are too close to the origin to be explained by tNCS.

Despite working with a single protein purification and crystallization protocol, we observed some variability in the determined values of the translated domain fractions

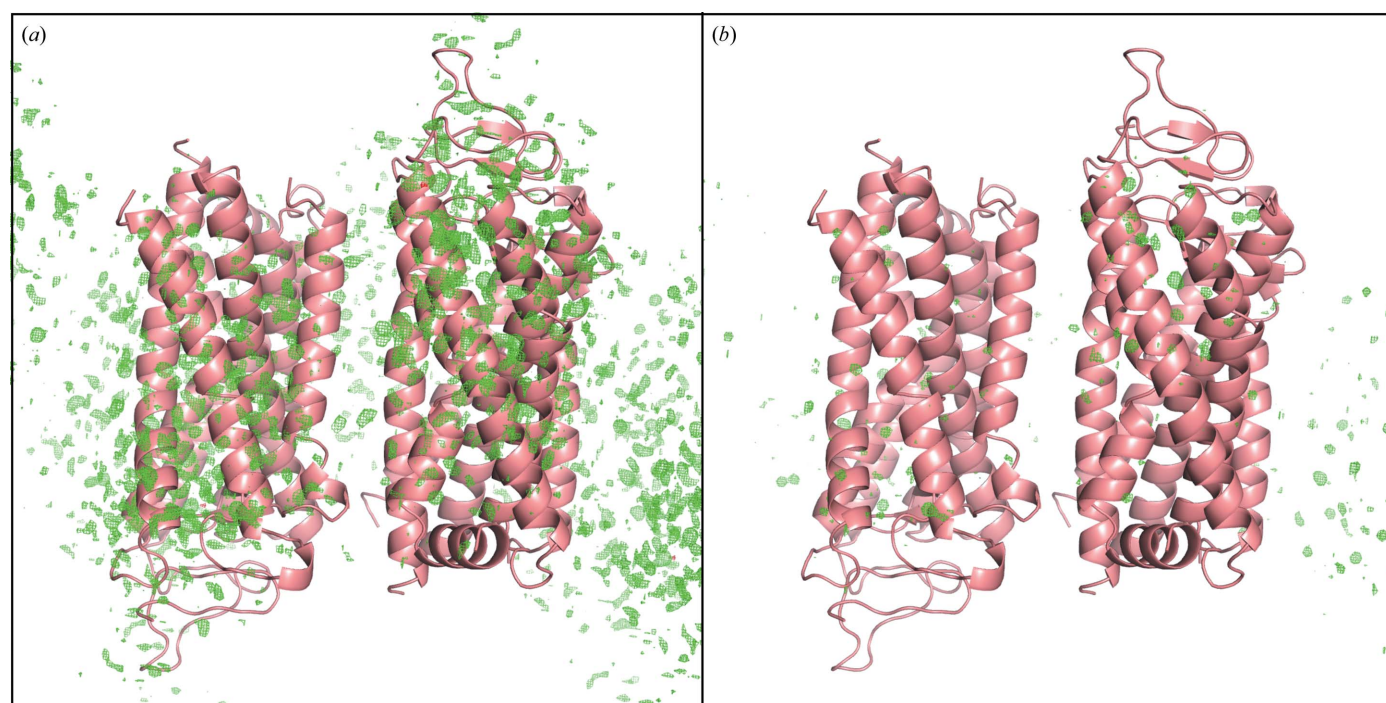


Figure 7 Map showing $mF_o - DF_c$ difference electron density after refinement of the SwissFEL dark-state model against (a) the original diffraction intensities (contoured at $0.514 \text{ e}^- \text{ \AA}^{-3}$, equivalent to 4.00σ) and (b) the corrected intensities (contoured at $0.514 \text{ e}^- \text{ \AA}^{-3}$, equivalent to 4.96σ).

depending on the examined data set. While data sets from the SwissFEL beamtime showed α_0 and α_1 values close to 0.01 and 0.19, respectively, we found that data sets collected during a previous beam time at SACLA had lower translated domain fractions, with values of $\alpha_0 = 0.00$ and $\alpha_1 = 0.15$. The unexplained difference density was therefore less prominent after model refinement using uncorrected intensities from the SACLA data set. Although the crystalline sample was prepared using the same protocol for both beam times, subtle differences in the sample preparation such as the crystallization temperature or precipitant composition may affect the probability of a lattice translocation occurring during crystal growth. Such a variation in the translated fraction has previously been observed in crystals of lentiviral integrase in complex with LEDGF and was ascribed to differences in crystal-growth conditions (Hare *et al.*, 2009). It is also possible that the smaller beam size used at SACLA, $1 \times 1 \mu\text{m}$ compared with $5 \times 5 \mu\text{m}$ at SwissFEL, reduced the chance of collecting data from a region of the crystal containing a lattice-translocation defect.

The translated domain fractions were found to be similar for all data sets collected during the same beam time. This is to be expected as both dark and light data sets were collected from the same sample batches and the translated domain fractions are the result of an average over all diffracting crystals contributing to each data set.

We attempted the refinement of a composite atomic model containing the main dimer at occupancy $(1 - \kappa)$ and a second model of the dimer with occupancy κ , translated from the main model by \mathbf{t}_d , against the uncorrected data, as was previously

successful for the 1918 H1N1 neuraminidase (Zhu *et al.*, 2008). The introduction of an additional copy of the dimer during refinement with our serial crystallography rhodopsin data necessitates the use of strict NCS constraints to avoid significantly worsening the data-to-parameter ratio. However, the refinement of a composite model against uncorrected data was deemed impractical with bovine rhodopsin, as interpretation of the $2mF_o - DF_c$ electron-density maps was challenging in the region where the two models overlapped. As previously observed by Zhu *et al.* (2008), we found that correction of the data for LTD was required to model solvent molecules in the overlapped region accurately. The positions of ordered water molecules are of key interest in studies of receptor dynamics as they often stabilize hydrogen-bond networks in the transitions between conformational states.

We did not observe dramatic changes in the time-resolved $F_o^{\text{light}} - F_o^{\text{dark}}$ difference maps as a result of the correction (Fig. 9). This is because the light and dark data sets were collected from the same sample using an interleaved sequence of visible light laser pulses to pump the sample and X-rays to probe the structure (Nango *et al.*, 2019). The systematic error due to LTD is therefore very similar in both data sets, usually with less than 1% difference in the translated populations. Instead, the most significant effect of LTD was to increase the number of poorly modelled regions in the dark-state models. As the dark-state model was taken as a starting point for the refinement of light-state structures into the extrapolated maps generated from the light-activated data sets, correction for LTD improved our models of both the dark and light-activated rhodopsin structures.

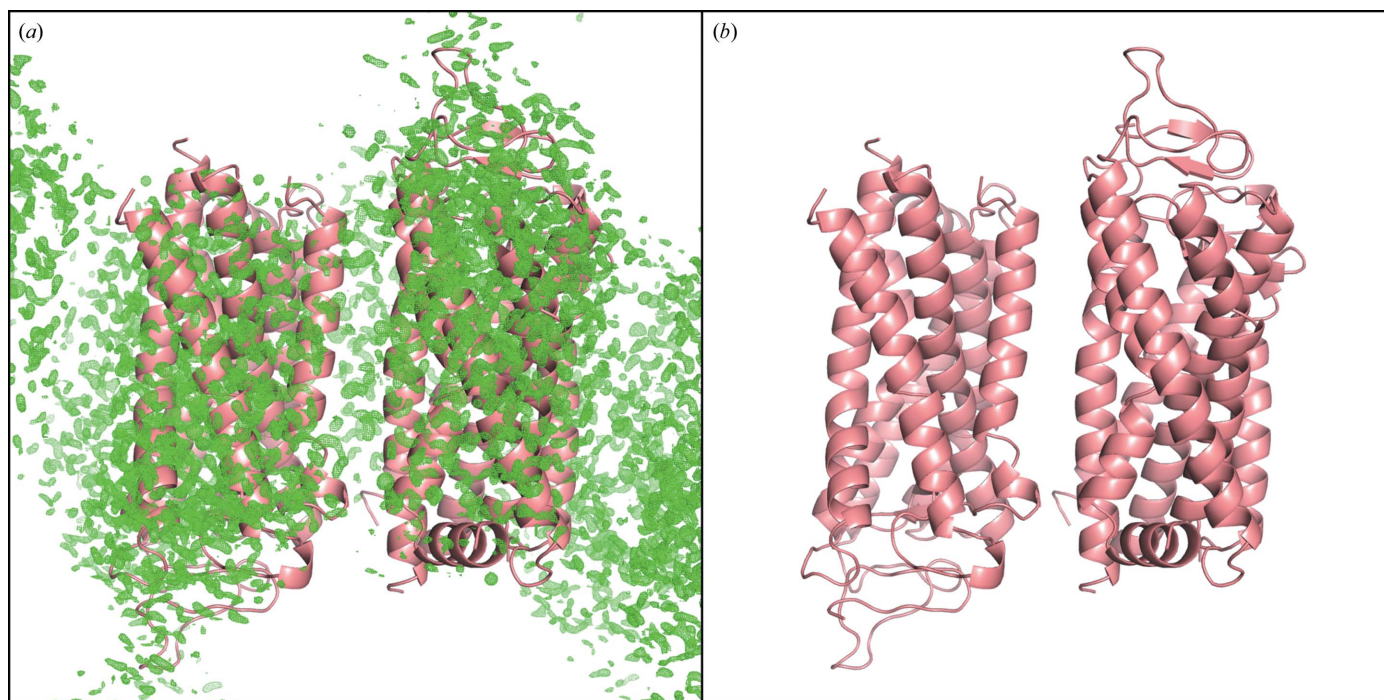


Figure 8

Map showing $mF_o - DF_c$ difference electron density after refinement of the SACLA dark-state model against (a) the original diffraction intensities (contoured at $1.071 \text{ e}^- \text{ \AA}^{-3}$, equivalent to 4.00σ) and (b) the corrected intensities (contoured at $1.071 \text{ e}^- \text{ \AA}^{-3}$, equivalent to 11.04σ).

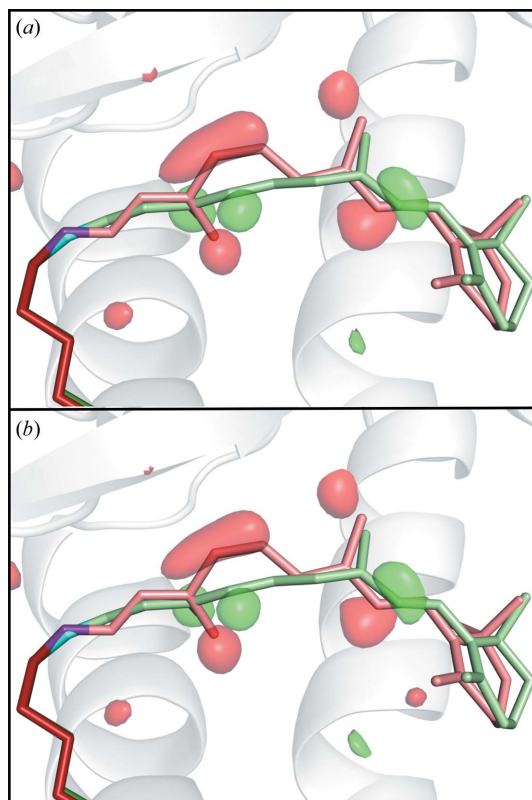


Figure 9
Isomorphous $F_o^{\text{light}} - F_o^{\text{dark}}$ difference electron density within 3.2 Å of the retinal chromophore 1 ps after light activation. Dark-state (SwissFEL) C atoms of Lys296 are shown in red, N atoms in purple and retinal atoms in pink. 1 ps light-activated (SwissFEL) C atoms of Lys296 are shown in dark green, N atoms in cyan and retinal atoms in light green. (a) Original diffraction intensities (contoured at 4.0σ , equivalent to $0.075 \text{ e}^- \text{ \AA}^{-3}$) and (b) corrected intensities (contoured at 4.0σ , equivalent to $0.075 \text{ e}^- \text{ \AA}^{-3}$).

7. Conclusions

TR-SFX experiments provide unprecedented opportunities to understand ultrafast biological processes, including the photoisomerization of retinal in rhodopsin, which is the first event in mammalian vision. Having overcome many of the challenging barriers to performing a TR-SFX experiment with rhodopsin, we identified the LTD in the crystals at the data-processing stage. The described correction facilitated the interpretation of the data collected during these time- and resource-intensive experiments (Gruhl *et al.*, 2022). Given the demanding pressures on sample preparation and beam time, it is often not possible to repeat an experiment after optimization of the crystals to avoid the LTD.

Here, we show the identification and characterization of LTD in a serial crystallography experiment. As these experiments, including time-resolved studies, become more widespread, it is likely that the present method of LTD correction will be of choice.

Acknowledgements

Open access funding provided by ETH-Bereich Forschungsanstalten.

Funding information

M. J. Rodrigues received funding from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No. 701647. This research project was funded by Swiss National Science Foundation Grant No. 192760 to G. F. X. Schertler.

References

- Bennett, N., Michel-Villaz, M. & Kühn, H. (1982). *Eur. J. Biochem.* **127**, 97–103.
- Bragg, W. L. & Howells, E. R. (1954). *Acta Cryst.* **7**, 409–411.
- Casañal, A., Lohkamp, B. & Emsley, P. (2020). *Protein Sci.* **29**, 1069–1078.
- Chapman, H. N., Caleman, C. & Timneanu, N. (2014). *Phil. Trans. R. Soc. B.* **369**, 20130313.
- Dauter, Z. (1999). *Acta Cryst.* **D55**, 1703–1717.
- Duisenberg, A. J. M. (1992). *J. Appl. Cryst.* **25**, 92–96.
- Emeis, D., Kühn, H., Reichert, J. & Hofmann, K. (1982). *FEBS Lett.* **143**, 29–34.
- Gevorgov, Y., Yefanov, O., Barty, A., White, T. A., Mariani, V., Brehm, W., Tolstikova, A., Grigat, R.-R. & Chapman, H. N. (2019). *Acta Cryst.* **A75**, 694–704.
- Gruhl, T., Weinert, T., Rodrigues, M., Milne, C. J., Ortolani, G., Nass, K., Nango, E., Sen, S., Johnson, P. J. M., Cirelli, C., Furrer, A., Mous, S., Skopintsev, P., James, D., Dworkowski, F., Båth, P., Kekilli, D., Ozerov, D., Tanaka, R., Glover, H., Bacellar, C., Brünle, S., Casadei, C. M., Diethelm, A. D., Gashi, D., Gotthard, G., Guixà-González, R., Joti, Y., Kabanova, V., Knopp, G., Lesca, E., Ma, P., Martiel, I., Mühle, J., Owada, S., Pamula, F., Sarabi, D., Tejero, O., Tsai, C.-J., Varma, N., Wach, A., Boutet, S., Tono, K., Nogly, P., Deupi, X., Iwata, S., Neutze, R., Standfuss, J., Schertler, G. F. & Panneels, V. (2022). *bioRxiv*, 2022.10.14.511948.
- Hare, S., Cherepanov, P. & Wang, J. (2009). *Acta Cryst.* **D65**, 966–973.
- James, D., Weinert, T., Skopintsev, P., Furrer, A., Gashi, D., Tanaka, T., Nango, E., Nogly, P. & Standfuss, J. (2019). *J. Vis. Exp.*, e59087.
- Landau, E. & Rosenbusch, J. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 14532–14535.
- Lewis, J. W. & Kliger, D. S. (1992). *J. Bioenerg. Biomembr.* **24**, 201–210.
- Li, L., Dai, S., Gao, G. F. & Wang, J. (2020). *Acta Cryst.* **D76**, 1057–1064.
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* **D75**, 861–877.
- Lovelace, J. J. & Borgstahl, G. E. O. (2020). *Crystallogr. Rev.* **26**, 3–50.
- Mathies, R. & Lugtenburg, J. (2000). *Molecular Mechanisms in Visual Transduction*, edited by D. Stavenga, W. DeGrip & E. Pugh, pp. 55–90. Amsterdam: North-Holland.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Mous, S., Gotthard, G., Ehrenberg, D., Sen, S., Weinert, T., Johnson, P. J. M., James, D., Nass, K., Furrer, A., Kekilli, D., Ma, P., Brünle, S., Casadei, C. M., Martiel, I., Dworkowski, F., Gashi, D., Skopintsev, P., Wrani, M., Knopp, G., Panepucci, E., Panneels, V., Cirelli, C., Ozerov, D., Schertler, G. F. X., Wang, M., Milne, C., Standfuss, J., Schapiro, I., Heberle, J. & Nogly, P. (2022). *Science*, **375**, 845–851.
- Nakamichi, H. & Okada, T. (2006a). *Angew. Chem. Int. Ed.* **45**, 4270–4273.
- Nakamichi, H. & Okada, T. (2006b). *Proc. Natl Acad. Sci. USA*, **103**, 12729–12734.
- Nakanishi, K. (1991). *Am. Zool.* **31**, 479–489.

- Nango, E., Kubo, M., Tono, K. & Iwata, S. (2019). *Appl. Sci.* **9**, 5505.
- Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A. N., Standfuss, J., Neutze, R. & Iwata, S. (2016). *Science*, **354**, 1552–1557.
- Nass Kovacs, G., Colletier, J.-P., Grünbein, M. L., Yang, Y., Stensitzki, T., Batyuk, A., Carbajo, S., Doak, R. B., Ehrenberg, D., Foucar, L., Gasper, R., Gorel, A., Hilpert, M., Kloos, M., Koglin, J. E., Reinstein, J., Roome, C. M., Schlesinger, R., Seaberg, M., Shoeman, R. L., Stricker, M., Boutet, S., Haacke, S., Heberle, J., Heyne, K., Domratcheva, T., Barends, T. R. M. & Schlichting, I. (2019). *Nat. Commun.* **10**, 3177.
- Nogly, P., Panneels, V., Nelson, G., Gati, C., Kimura, T., Milne, C., Milathianaki, D., Kubo, M., Wu, W., Conrad, C., Coe, J., Bean, R., Zhao, Y., Båth, P., Dods, R., Harimoorthy, R., Beyerlein, K. R., Rheinberger, J., James, D., DePonte, D., Li, C., Sala, L., Williams, G. J., Hunter, M. S., Koglin, J. E., Berntsen, P., Nango, E., Iwata, S., Chapman, H. N., Fromme, P., Frank, M., Abela, R., Boutet, S., Barty, A., White, T. A., Weierstall, U., Spence, J., Neutze, R., Schertler, G. & Standfuss, J. (2016). *Nat. Commun.* **7**, 12314.
- Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D., Furrer, A., Gashi, D., Borin, V., Skopintsev, P., Jaeger, K., Nass, K., Bath, P., Bosman, R., Koglin, J., Seaberg, M., Lane, T., Kekilli, D., Brunle, S., Tanaka, T., Wu, W., Milne, C., White, T., Barty, A., Weierstall, U., Panneels, V., Nango, E., Iwata, S., Hunter, M., Schapiro, I., Schertler, G., Neutze, R. & Standfuss, J. (2018). *Science*, **361**, eaat0094.
- Oda, K., Nomura, T., Nakane, T., Yamashita, K., Inoue, K., Ito, S., Vierock, J., Hirata, K., Maturana, A. D., Katayama, K., Ikuta, T., Ishigami, I., Izume, T., Umeda, R., Eguma, R., Oishi, S., Kasuya, G., Kato, T., Kusakizako, T., Shihoya, W., Shimada, H., Takatsuji, T., Takemoto, M., Taniguchi, R., Tomita, A., Nakamura, R., Fukuda, M., Miyauchi, H., Lee, Y., Nango, E., Tanaka, R., Tanaka, T., Sugahara, M., Kimura, T., Shimamura, T., Fujiwara, T., Yamanaka, Y., Owada, S., Joti, Y., Tono, K., Ishitani, R., Hayashi, S., Kandori, H., Hegemann, P., Iwata, S., Kubo, M., Nishizawa, T. & Nureki, O. (2021). *eLife*, **10**, e62389.
- Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P. & Buss, V. (2004). *J. Mol. Biol.* **342**, 571–583.
- Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M. & Miyano, M. (2000). *Science*, **289**, 739–745.
- Pickersgill, R. W. (1987). *Acta Cryst.* **A43**, 502–506.
- Poddar, H., Heyes, D. J., Schirò, G., Weik, M., Leys, D. & Scrutton, N. S. (2022). *FEBS J.* **289**, 576–595.
- Ponnusamy, R., Lebedev, A. A., Pahlow, S. & Lohkamp, B. (2014). *Acta Cryst.* **D70**, 1680–1694.
- Powell, H. R. (1999). *Acta Cryst.* **D55**, 1690–1695.
- Read, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
- Saitoh, Y., Mitani-Ueno, N., Saito, K., Matsuki, K., Huang, S., Yang, L., Yamaji, N., Ishikita, H., Shen, J., Ma, J. F. & Suga, M. (2021). *Nat. Commun.* **12**, 6236.
- Skopintsev, P., Ehrenberg, D., Weinert, T., James, D., Kar, R. K., Johnson, P. J. M., Ozerov, D., Furrer, A., Martiel, I., Dworkowski, F., Nass, K., Knopp, G., Cirelli, C., Arrell, C., Gashi, D., Mous, S., Wraniak, M., Gruhl, T., Kekilli, D., Brünle, S., Deupi, X., Schertler, G. F. X., Benoit, R. M., Panneels, V., Nogly, P., Schapiro, I., Milne, C., Heberle, J. & Standfuss, J. (2020). *Nature*, **583**, 314–318.
- Standfuss, J., Edwards, P. C., D'Antona, A., Fransen, M., Xie, G., Oprian, D. D. & Schertler, G. F. X. (2011). *Nature*, **471**, 656–660.
- Tsai, C.-J., Marino, J., Adaixo, R., Pamula, F., Muehle, J., Maeda, S., Flock, T., Taylor, N. M., Mohammed, I., Matile, H., Dawson, R. J. P., Deupi, X., Stahlberg, H. & Schertler, G. (2019). *eLife*, **8**, e46041.
- Tsai, C.-J., Pamula, F., Nehmé, R., Mühle, J., Weinert, T., Flock, T., Nogly, P., Edwards, P. C., Carpenter, B., Gruhl, T., Ma, P., Deupi, X., Standfuss, J., Tate, C. G. & Schertler, G. F. X. (2018). *Sci. Adv.* **4**, eaat7052.
- Tsai, Y., Sawaya, M. R. & Yeates, T. O. (2009). *Acta Cryst.* **D65**, 980–988.
- Wang, J., Kamtekar, S., Berman, A. J. & Steitz, T. A. (2005). *Acta Cryst.* **D61**, 67–74.
- Wang, J., Rho, S.-H., Park, H. H. & Eom, S. H. (2005). *Acta Cryst.* **D61**, 932–941.
- Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Doak, R. B., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Han, G. W., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G. J., Koglin, J. E., Seibert, M. M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T. A., Howe, N., Caffrey, M. & Cherezov, V. (2014). *Nat. Commun.* **5**, 3309.
- Weinert, T. & Panneels, V. (2020). *Methods Mol. Biol.* **2127**, 321–338.
- Weinert, T., Skopintsev, P., James, D., Dworkowski, F., Panepucci, E., Kekilli, D., Furrer, A., Brünle, S., Mous, S., Ozerov, D., Nogly, P., Wang, M. & Standfuss, J. (2019). *Science*, **365**, 61–65.
- White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.
- Wu, W., Nogly, P., Rheinberger, J., Kick, L. M., Gati, C., Nelson, G., Deupi, X., Standfuss, J., Schertler, G. & Panneels, V. (2015). *Acta Cryst.* **F71**, 856–860.
- Yun, J.-H., Li, X., Yue, J., Park, J.-H., Jin, Z., Li, C., Hu, H., Shi, Y., Pandey, S., Carbajo, S., Boutet, S., Hunter, M. S., Liang, M., Sierra, R. G., Lane, T. J., Zhou, L., Weierstall, U., Zatsepin, N. A., Ohki, M., Tame, J. R. H., Park, S.-Y., Spence, J. C. H., Zhang, W., Schmidt, M., Lee, W. & Liu, H. (2021). *Proc. Natl Acad. Sci. USA*, **118**, e2020486118.
- Zhu, X., Xu, X. & Wilson, I. A. (2008). *Acta Cryst.* **D64**, 843–850.