

# Hydrogen properties in an organic molecule revealed by XFEL and electron crystallography

**Authors:** Kiyofumi Takaba<sup>1, †</sup>, Saori Maki-Yonekura<sup>1, †</sup>, Ichiro Inoue<sup>1</sup>, Kensuke Tono<sup>1, 2</sup>, Tasuku Hamaguchi<sup>1</sup>, Keisuke Kawakami<sup>1</sup>, Hisashi Naitow<sup>1</sup>, Tetsuya Ishikawa<sup>1</sup>, Makina Yabashi<sup>1, 2</sup> and Koji Yonekura<sup>1, 3, 4, \*</sup>

## **Affiliations:**

<sup>1</sup> RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

<sup>2</sup> Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo, Hyogo, 679-5198, Japan

<sup>3</sup> Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, 2-1-1

Katahira, Aoba-ku, Sendai 980-8577, Japan

<sup>4</sup> Advanced Electron Microscope Development Unit, RIKEN-JEOL Collaboration Center, RIKEN Baton Zone Program, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

<sup>†</sup> These authors contributed equally.

\* Corresponding author. E-mail: [yone@spring8.or.jp](mailto:yone@spring8.or.jp). Phone: +81-791-58-2837.

## **Summary**

Structure analysis of small crystals is important in synthetic organic chemistry, pharmaceutical and material sciences, and related areas, as the conformation of these molecules may differ in large and small crystals, thus affecting the interpretation of their functional properties and drug efficacy. From small crystals, X-ray and electron beams could furnish electron densities and Coulomb potentials of target molecules, respectively.

The two beams provide distinctly different information, and this potential has not been fully explored. Here we present the detailed structure of an organic molecule, rhodamine-6g by X-ray free-electron laser (XFEL) and electron crystallography from the same sample batch of microcrystals. This is the first organic molecular structure determined using XFEL at subatomic resolution. Direct comparison between the electron-density and the Coulomb-potential maps together with theoretical models based on Poisson's equation shows that the position of hydrogen atoms depends on bond type and charge distribution. The combined approach could lead to better insights into their chemical and/or binding properties for a broad range of organic molecules.

## **Introduction**

X-rays are scattered by the electrons around atoms, and electrons by charges in their path. Thus, the two beams yield an electron-density map and a Coulomb-potential map of an object, respectively, and each map should exhibit distinct features, in a similar way that neutron diffraction (ND) complements X-ray diffraction (XRD). It is well known that the atomic cross-section that defines usable sample sizes for structure determination differ substantially between an electron beam and the other two (neutron:  $\sim 10^1$  barns; X-ray:  $10^0 \sim 10^1$  barns; electron:  $\sim 10^5$  barns)<sup>1</sup>. Also, electron scattering factors vary considerably between neutral and charged atoms particularly at low resolution<sup>2,3</sup>. This feature has been exploited to measure charge-distributions in inorganic materials from convergent-beam electron diffraction (CBED) patterns<sup>2,4,5</sup>. For protein molecules, charged states at functional sites have been investigated by rotation 3D electron diffraction with a parallel beam (3D ED / microED)<sup>6</sup> and single-particle cryo-electron microscopy (cryo-EM)<sup>7</sup>. Hydrogen densities in small organic molecules and protein structures revealed by 3D ED and high-resolution single-particle analysis appear further away from their parent atoms than those in X-ray crystal structures<sup>8,9</sup>. However, such observations are still limited and qualitative. This is due to an inability to directly compare atom densities and charge distribution because X-ray and electron analyses are not performed on the same sample.

The strong scattering power of electrons limits the sample thickness such that only small and thin crystals are suitable for electron diffraction (ED) studies, while X-ray free-electron lasers (XFELs) need small but thicker crystals thanks to the intense beam. It seems possible that a single sample could yield a range of crystal sizes that would be useful for both techniques. In practice, small organic compounds may be the most promising for simultaneous growth of small and larger crystals. 3D ED can collect rotational diffraction patterns from such

undersized crystals<sup>6,10,11</sup>, and can resolve hydrogen atoms in organic molecular crystals<sup>8,12</sup>. While, serial crystallography with XFEL (SX / SXFEL) records still diffraction frames with exposure of ultrafast X-ray pulses<sup>13</sup>, and has been applied mainly to protein crystals so far.

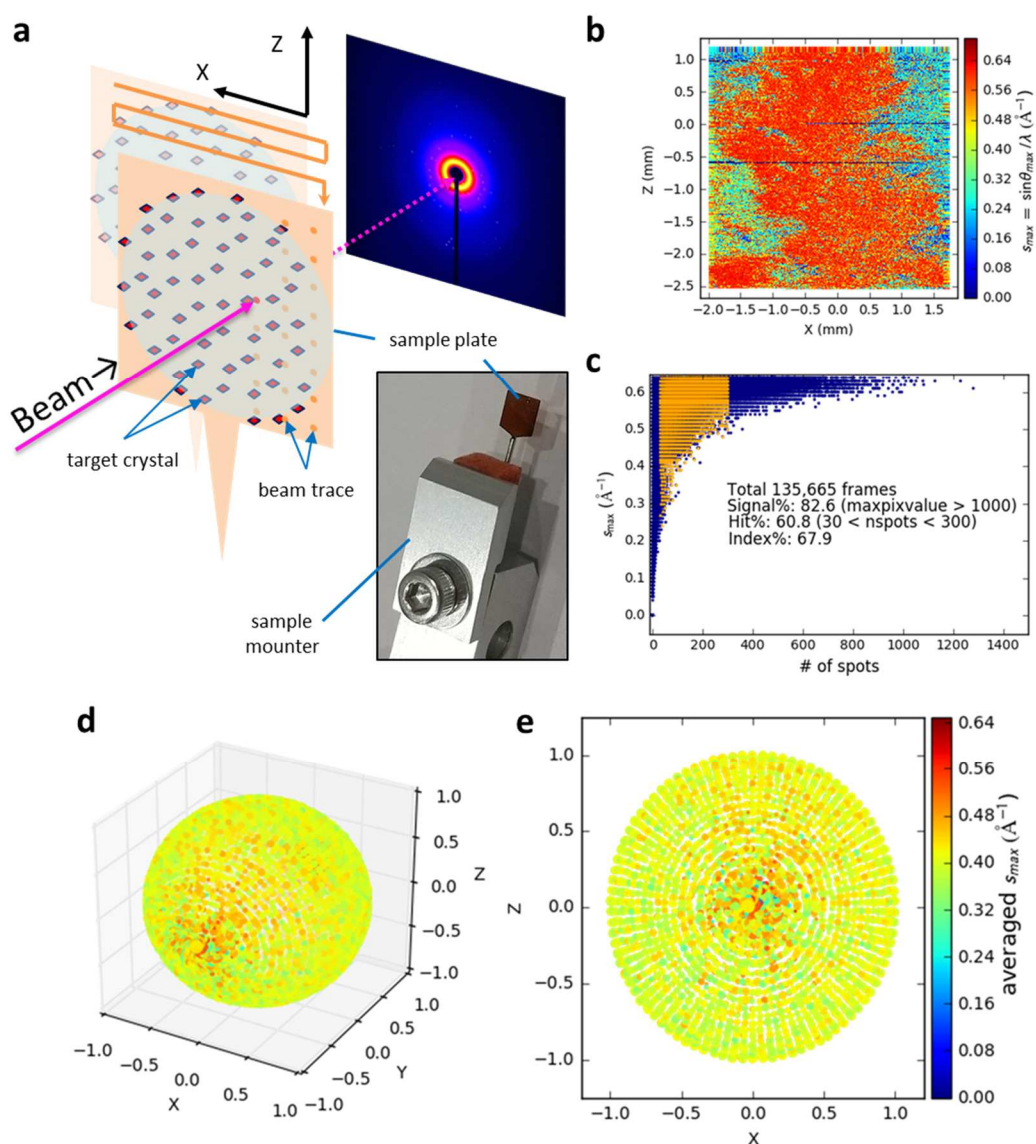
Moreover, the conformations of organic molecules may differ in large and small crystals, which affects interpretation of functional properties and drug efficacy<sup>14,15</sup>. Thus, structure analysis of small crystals is important in synthetic organic chemistry, pharmaceutical and material sciences and related areas.

Here, we apply both electron and serial XFEL crystallography to an organic molecule, rhodamine-6g (9-[2-(ethoxycarbonyl)phenyl]-3,6-bis-(ethylamino)-2,7-dimethylxanthylium), obtained from the same sample batch containing many tiny crystals. For use of XFEL to organic molecular crystals, we have introduced a new scheme using a simplified sample delivery for efficient data collection and combining lattice parameters obtained from ED for data processing. To our knowledge, this is the first organic-molecular structure determination by SX and ED, an approach that could find broad application in the structural science of organic molecules for a more realistic portrayal of their structures.

## **Structure determination at subatomic resolution by SX**

Microcrystals of a chloride salt of rhodamine-6g were attached on a custom-designed flat polyimide plate ( $4 \times 4 \text{ mm}^2$  square) with liquid paraffin. SX diffraction patterns were recorded at 30 Hz by 2D scanning of the sample plate (Fig. 1a), and X-ray pulses with high energy (15 keV) constantly revealed spots beyond 1 Å resolution, which were still limited by the camera size and camera distance (Fig. 1b). More than ~130,000 patterns were collected from one plate in ~1.5 h. Compared with SX macromolecular crystallography, the number of observed diffraction spots (~20 to ~1000) per frame appeared to be smaller (Fig. 1c), owing to the smaller

lattice lengths and lower mosaicity in the crystals. Thus, assignment of lattice indices to diffraction spots in three dimensions was difficult without prior knowledge of lattice parameters. We supplied the parameters obtained from an ED experiment of the same sample (see below), and this allowed for processing of the XFEL data including partially recorded reflections in the same way as for SX of macromolecules. The success rate of indexing was 69.2%, which is comparable to the rate for protein crystals reported previously<sup>16,17</sup>. Then, extracted intensities were integrated using the Monte Carlo method (Extended Data Table 1). Orientations of the crystals on the sample plate were random (Fig. 1d, e), and the completeness reached 100%. The crystal structure was solved by the direct-method and a Fourier difference map obtained by subtraction of calculated structure factors omitting hydrogen atoms from the experimental data clearly resolved densities for hydrogen atoms (Fig. 2a, Extended Data Fig. 1a, b). Residual densities between atoms in the difference map may be attributed to covalent electrons (Extended Data Fig. 1b). To our knowledge, this is the first crystal structure of an organic compound determined by XFEL indicating that neither bonded hydrogen atoms nor bonding electrons disperse away from the molecule on exposure to a hard XFEL pulse without helium-ambiance<sup>16</sup> or cryogenic cooling<sup>17</sup>.



**Fig. 1 | SX of rhodamine-6g crystals.** (a) Experimental setup. (b) A typical map of the sample plate showing data-taking positions from rhodamine-6g microcrystals. The  $s_{\max}$  of identified diffraction spots per frame is represented by a color dot. (c) Plot of the spot number vs  $s_{\max}$  in one frame. Frames used for indexing were shown in orange. (d) 3D distribution of crystal orientations over the indexed frames. (e) Projection of (d) along incident X-rays. The terminals of  $a^*$  vectors are shown with colored points. The size of each point and color represent the number of frames contributing to the corresponding orientation and the averaged  $s_{\max}$ , respectively.

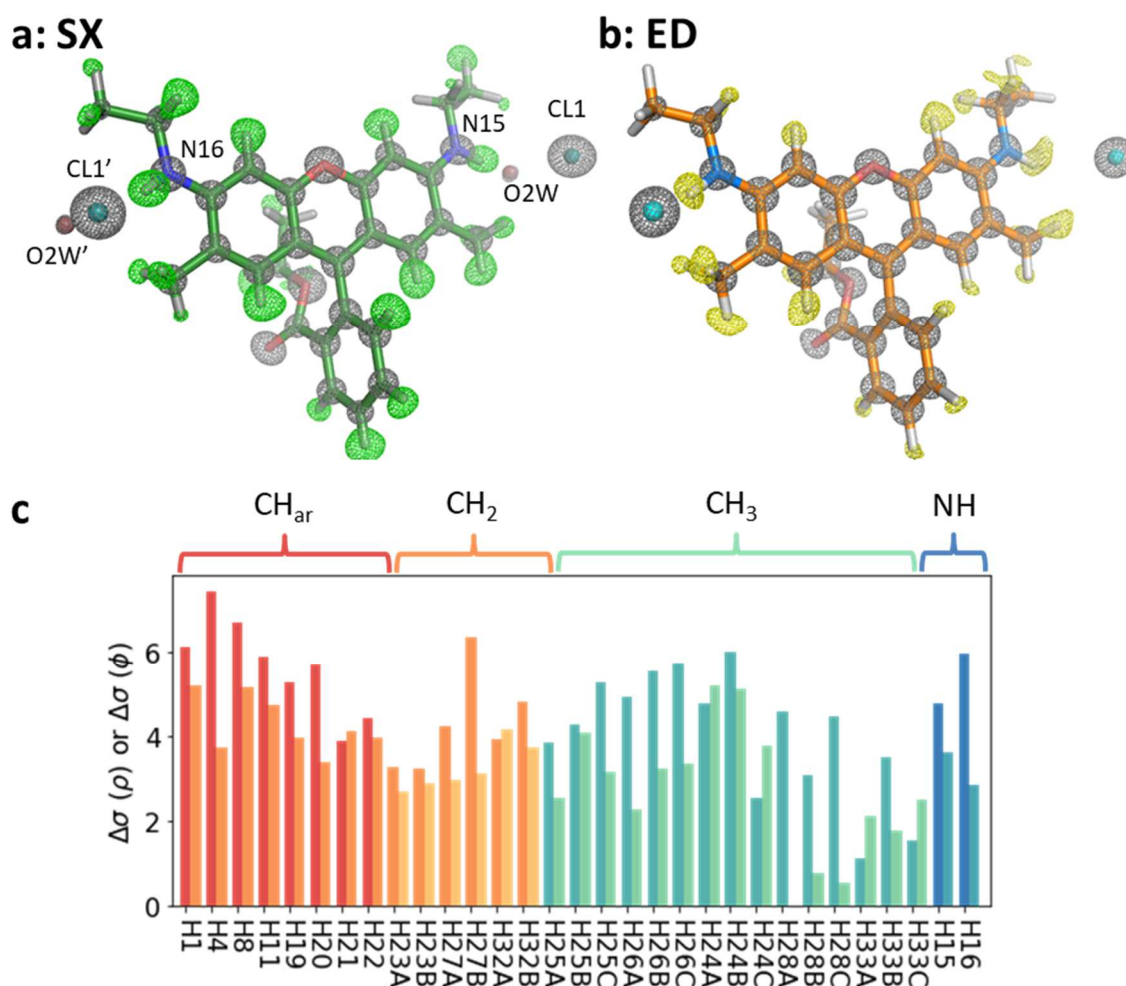
## **Comparison of SX and ED structures**

ED patterns were collected with a parallel beam of 300 kV electrons at room (rt-ED) and cryogenic (cryo-ED) temperatures from the same sample batch (see Methods). The crystal structures were solved from merged datasets by the direct-method, as described previously<sup>18</sup>. Hydrogen atoms are also resolved in the Fourier difference map (Fig. 2b, Extended Data Fig. 1, c-f).

The SX and ED structures are almost identical, but a solvent-accessible site near the amides and the chloride is partially occupied by water in the SX and cryo-ED structures, and empty in the rt-ED structure (Extended Data Fig. 2, a-e). The rt-ED structure likely lost the water molecules in the high vacuum of the electron microscope column. Compared with the SX structure, the cell lengths along the *c* axis shrink in both the ED structures but the cell length along the *b* axis shrinks only in the cryo-ED structure likely upon cooling with liquid nitrogen (Extended Data Table 1). Two molecules of this aromatic dye compose the H-type dimer through the planes of the xanthene rings (C1 to C12 and O13; Extended Data Fig. 2d). The configuration of the dimers in this orthorhombic crystal is not identical to the triclinic crystal structure with chlorine determined previously by XRD from a single crystal (Supplementary Discussion; Extended Data Fig. 2, f-h, Extended Data Table 2). We then proceeded to re-crystallize the same sample following the described condition<sup>19</sup>, and obtained triclinic crystals, and determined the structure by cryo-ED. The structure (triclinic-cryo-ED) is identical to the reported triclinic crystal structure (Extended Data Fig. 2, f, g, Extended Data Tables 1, 2 and see Methods). We confirm that the SX diffraction patterns could not be processed with triclinic cell parameters.

The rt-ED structure yields better data statistics and less deviations in bond geometry and resolves hydrogen densities at higher  $\sigma$  levels than that of the cryo-ED (Extended Data Table 1 and Supplementary Tables 1, 2). Radiation damage would be sufficiently small to maintain structural details at least to this resolution in the rt-ED structure. As mentioned, the cell volume changes the most in the cryo-ED structure and is reduced by  $\sim 7.4\%$  of that of the SX, and by  $\sim 3.2\%$  in the rt-ED structure. The large cell change of the former may detract from the quality and/or homogeneity of the crystals, and thus, we adopted the rt-ED structure primarily for further comparison with the SX structure. The rt-ED structure indeed shows significantly better correlations with the SX structure in bond lengths and atomic displacement parameters (Extended Data Fig. 3). It is generally observed that the  $R$  factor, the discrepancy between experimental data and model, for an ED structure is worse than that for an XRD structure, and reflecting this, geometry errors in ED structures have been reported to be greater<sup>8</sup>. Here, these tendencies are revealed in the SX and ED structures. The estimated standard uncertainties in the bond lengths in the ED crystal structures are  $\sim 2$ – $5$  times worse than those in the SX structure (Supplementary Table 1). Hydrogen densities, particularly in alkyl groups, are also resolved at lower  $\sigma$  levels (Fig. 2c). The higher errors may be attributed to suboptimal assignment of electron scattering factors, effects of dynamical scattering, smaller numbers of units in crystals in the final datasets and so on.

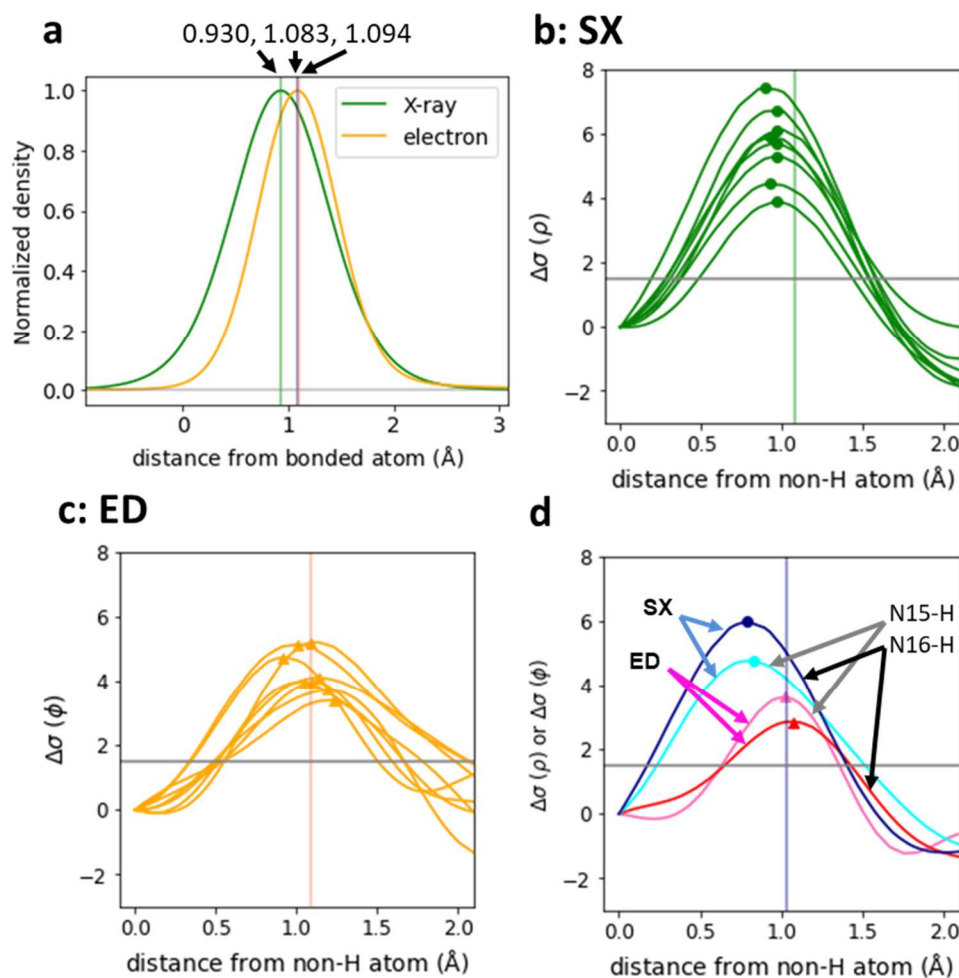




**Fig. 2 | Structures of rhodamine-6g determined by SX and ED.** (a) The SX structure containing chloride ions (CL1 and CL1') and water molecules (O2W and O2W'). The prime symbols represent atoms related by the crystallographic symmetry. (b) The rt-ED structure. Gray nets were calculated from the observed amplitudes ( $F_o$ ). Hydrogen-omitting difference maps ( $mF_o - DF_o$ ) are shown in green in (a) and yellow in (b). Display contour levels are  $3.0\sigma$  for all the nets. (c) Bar plot of heights of density peaks in the hydrogen omit maps along individual hydrogen atoms. Bars are grouped by bond types, and displayed in dark colors for the SX data ( $\rho$ ) and in light colors for the rt-ED data ( $\phi$ ).

## Hydrogen densities

Comparison of the topology of hydrogen densities in the electron-density and Coulomb-potential maps showed differences in peak positions (Fig. 2a, b). Electrons in hydrogen atoms are attracted toward bonded non-hydrogen atoms, yielding peak shifts toward the bonded atoms in the electron-density maps<sup>20,21</sup>. The shifts are larger in polar bonds, as the attractive force is stronger. This is in marked contrast to the nuclear density observed with ND<sup>22</sup>. In ED experiments, incident electrons are affected by both electrons and nuclear charges, the latter of which is localized in the nucleus and dominant in the Coulomb potential. Therefore, the peak positions of hydrogen atoms are closer to the nuclei but not identical to those in the nuclear density by ND. In our theoretical calculation based on Poisson's equation (Fig. 3a), the peak locations in the Coulomb-potential map appear slightly longer than those in ND and exhibit the opposite shift compared to that in XRD. In fact we observed that distances from the parent atom to the density peak of hydrogen appear shorter in the SX structure than in ED (Fig. 3, b-d). Differences in the distances,  $\Delta X-H_{\text{peak}}$ , reveal a significant tendency of the more distant positions of hydrogen density peaks in aromatic C-H bonds, whereas those in methyl and methylene C-H bonds are comparable to the standard deviation level of the peak positions (Table 1 and Extended Data Fig. 4). This observation is consistent with the hydrogen atoms of aromatic C-H bonds modeled with more of a shift toward its parent atom in XRD structures than those of methyl and methylene C-H bonds (Supplementary Table 3)<sup>23</sup>. These density features fundamentally arise from the polarity of C-H bonds, but also reflect the flexibility of hydrogen atoms. Motions of the hydrogen atom are more restricted along the bonding directions in the  $sp^2$ -hybridized C-H bond than in the other  $C(sp^3)$ -H bonds, which are revealed due to increased hydrogen visibility (Fig. 2c). Larger  $\Delta X-H_{\text{peak}}$  values are measured for the amides owing to the stronger polarity in the  $N(sp^2)$ -H bonds. This theme is further discussed in Supplementary Discussion.



**Fig. 3 | Peak positions of hydrogen densities.** (a) Theoretical curves of hydrogen density along C-H bond. The Coulomb-potential curve in green was calculated from the electron-density curve in yellow (see Methods). The peak heights are adjusted for clarity. Vertical lines represent the center position of the curves and hydrogen nucleus (1.083 Å)<sup>22</sup>. (b, c) Observed hydrogen density along aromatic C-H bonds in the hydrogen-omitted map in the SX (b) and rt-ED (c) structures. Small circles and triangles on the plots represent peak locations, and gray horizontal lines refer to a density threshold level of  $1.5\sigma$ . (d) The same plots as in (b) and (c) but for the N-H bonds.

**Table 1. | Summary of hydrogen densities**

	$N_{\text{all}}^{\text{a}}$	$N_{\text{obs}}^{\text{SX}}$	$X - H_{\text{peak}}^{\text{SX}} (\text{\AA})^{\text{b}}$	$N_{\text{obs}}^{\text{ED}}$	$X - H_{\text{peak}}^{\text{ED}} (\text{\AA})$	$N_{\text{obs}}^{\text{SXandED}}$	$\Delta X - H_{\text{peak}} (\text{\AA})$
C-H <sub>3</sub>	15	14	1.074 (103) <sup>c</sup>	10	1.152 (118)	9	0.095 (175)
C-H <sub>2</sub>	6	6	1.064 (89)	5	1.128 (100)	5	0.077 (76)
C-H <sub>ar</sub>	8	8	0.949 (28)	8	1.088 (95)	8	0.139 (100)
N15-H	1	1	0.824	1	1.027	1	0.203
N16-H	1	1	0.788	1	1.070	1	0.282

<sup>a</sup>  $N_{\text{all}}$ ,  $N_{\text{obs}}^{\text{SX}}$ ,  $N_{\text{obs}}^{\text{ED}}$ ,  $N_{\text{obs}}^{\text{SXandED}}$ : the numbers of all and observed hydrogen atoms in the SX and rt-ED densities and both of the two, respectively.

<sup>b</sup>  $X - H_{\text{peak}}^{\text{SX}}$ ,  $X - H_{\text{peak}}^{\text{ED}}$  and  $\Delta X - H_{\text{peak}}$ : averages of observed peak positions and their differences, respectively.

<sup>c</sup> The standard deviations in the same type of bonds are in the parentheses.

We then refined the rt-ED structure with the riding hydrogen model constrained with the longer bond lengths, as the resolution of the data was not sufficiently high to treat hydrogen atoms as unrestrained<sup>8</sup>. This yielded a drop of  $R$  values from 0.167 for  $F_o > 4\sigma$  and 0.206 for all  $F_o$  to 0.164 and 0.203, respectively.

## **Analysis of charge**

In theory, ED could provide charge information directly. One rhodamine-6g molecule forms a salt with a chloride ion of negative charge, and positive charge should reside around the two nitrogen atoms, N15 and N16. The  $\Delta X - H_{\text{peak}}$  values are much larger in these amides than those in carbohydrates (Fig. 3). This indicates that both hydrogen atoms, H15 and H16, are polarized to some extent. The distances between the nitrogen atom (N15 or N16) in the amide and the chloride ion (CL1) are longer in the SX structure (3.529 (3) Å between N15 and CL1; and

3.433 (3) Å between N16 and CL1', [Supplementary Table 4](#)) than in the rt-ED structure (3.354 (10) Å; and 3.408 (10) Å). These changes would correlate with the existence of a water molecule (O2W) and the difference in the cell length along the *c* axis. Thus, from the geometries around both the chloride-binding sites, either of the hydrogen atoms H15 and H16 would not be exclusively charged in the rt-ED structure. Indeed, *R* values are not improved when assigning a scattering factor of fully positive charge to one of the hydrogen atoms H15 or H16: 0.167 and 0.205 for  $F_o > 4\sigma$  and all  $F_o$ , respectively, when H15 given +1.0, and 0.169 and 0.207 when H16 given +1.0 ([Supplementary Table 5](#) and [Extended Data Fig. 5a](#)). We then changed the partial charge values by a step of 0.1 in H15, H16, and CL1 as described<sup>7</sup>, and obtained a refined structure giving *R* values of 0.162 and 0.201 when assigning +0.4 to H15, +0.2 to H16, and -0.9 to CL1 ([Extended Data Fig. 5, d-i and see Methods](#)). The atomic coordinates were not much changed by refinement with and without assignment of charges. The same approach was tested for the SX data, but *R* values do not improve ([Extended Data Fig. 6](#)) due to considerably less sensitivity of X-ray scattering factors to charges<sup>3</sup>.

Charge is evidently not localized to a particular half of the molecule and this is reasonable for the observed configuration, where the phenolic ring at the center of the xanthene ring is approximately perpendicular to the xanthene plane ([Extended Data Fig. 2f](#)), making both moieties nearly identical. The flipped oxycarbonyl group also appears to be consistent with this conformation, as the ether oxygen (O31) is located close to the center of the xanthene ring ([Extended Data Fig. 7 and Supplementary Table 1](#)). This configuration contrasts to that of the triclinic cryo-ED / XRD structure ([Extended Data Fig. 2](#)).

Residual densities in the difference maps appear noisier for the ED data ([Extended Data Fig. 1b](#)), which may represent that assignment of electron scattering factors are suboptimal as mentioned. We then converted the rt-ED experimental structure factors to X-ray structure

factors by the Mott–Bethe formula<sup>24,25</sup> instead of view in the Coulomb-potential dimension. A difference Fourier electron density map between the converted data and calculation with neutral charges now reveals interpretable features, excess of electron on an oxygen atom but deficiency on carbon and some hydrogen atoms (Extended Data Fig. 8). This must reflect electronegativity of each atom. In contrast, it is known to be difficult to express electron scattering by target molecules including charges with simple theoretical scattering curves in CBED, and so structure factors are optimized against experimental CBED patterns independently of the theoretical curves<sup>2,4,26</sup>. This kind of treatment will need further investigation for application to 3D ED.

## **Conclusion**

Compared with X-rays, electrons reveal unique properties of hydrogen, not only due to the relatively higher scattering power<sup>27</sup> but also to the distal location of the density and to their higher sensitivity to charges. ED can provide more accurate information on the hydrogen nucleus than XRD, despite larger coordinate errors. Ultra-high resolution XRD, ND, and combination of the two have been used to obtain information beyond atomic coordinates of target molecules<sup>28</sup>. These studies, particularly by ND, have been severely limited by the requirement for large crystals. The use of tiny crystals in the sample here is more practical and no special treatment was needed for either SX or ED measurements. Combining the techniques clearly reveals their distinct complementary features, and should be applicable for investigation of more complex structures and their chemical properties.

## **References**

1. Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193 (1995).
2. Zuo, J. M., Kim, M., O’Keeffe, M. & Spence, J. C. H. Direct observation of d-orbital holes and Cu-Cu bonding in Cu<sub>2</sub>O. *Nature* **401**, 49–52 (1999).
3. Yonekura, K. et al. Ionic scattering factors of atoms that compose biological molecules. *IUCrJ* **5**, 348–353 (2018).
4. Nakashima, P. N. H., Smith, A. E., Etheridge, J. & Muddle, B. C. The bonding electron density in aluminum. *Science* **331**, 1583–1586 (2011).
5. Gao, W. et al. Real-space charge-density imaging with sub-ångström resolution by four-dimensional electron microscopy. *Nature* **575**, 480–484 (2019).
6. Yonekura, K., Kato, K., Ogasawara, M., Tomita, M. & Toyoshima, C. Electron crystallography of ultrathin 3D protein crystals: Atomic model with charges. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3368–3373 (2015).
7. Yonekura, K. & Maki-Yonekura, S. Refinement of cryo-EM structures using scattering factors of charged atoms. *J. Appl. Crystallogr.* **49**, 1517–1523 (2016).
8. Clabbers, M. T. B., Gruene, T., van Genderen, E. & Abrahams, J. P. Reducing dynamical electron scattering reveals hydrogen atoms. *Acta Crystallogr. Sect. A Found. Adv.* **75**, 82–93 (2019).
9. Nakane, T. et al. Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).
10. Zhang, D., Oleynikov, P., Hovmöller, S. & Zou, X. Collecting 3D electron diffraction data by the rotation method. *Z. Krist.* **225**, 94–102 (2010).

- 283 11. Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. High-resolution structure  
284 determination by continuous-rotation data collection in MicroED. *Nat. Methods* **11**, 927–  
285 930 (2014).
- 286 12. Palatinus, L. et al. Hydrogen positions in single nanocrystals revealed by electron  
287 diffraction. *Science* **355**, 166–169 (2017).
- 288 13. Chapman, H. N., et al. Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–  
289 78 (2011).
- 290 14. Datta, S. & Grant, D. J. W. Crystal structures of drugs: Advances in determination,  
291 prediction and engineering. *Nat. Rev. Drug Discov.* **3**, 42–57 (2004).
- 292 15. Liu, H., et al. Phosphorene: An unexplored 2D semiconductor with a high hole mobility.  
293 *ACS Nano*. **8**, 4033–4041 (2014).
- 294 16. Masuda, T., et al. Atomic resolution structure of serine protease proteinase K at ambient  
295 temperature. *Sci. Rep.* **7**, 45604 (2017).
- 296 17. Kwon, H., et al. XFEL crystal structures of peroxidase Compound II. *Angew. Chem. Int.*  
297 *Ed. Engl.* **60**, 14578–14585 (2021).
- 298 18. Takaba, K., Maki-Yonekura, S., Inoue, S., Hasegawa, T. & Yonekura, K. Protein and  
299 organic-molecular crystallography with 300kV electrons on a direct electron detector.  
300 *Front. Mol. Biosci.* **7**, 612226 (2020).
- 301 19. Adhikesavalu D. N., Mastropaolo, D., Camerman, A. & Camerman, N. Two rhodamine  
302 derivatives: 9-[2-(ethoxycarbonyl)phenyl]-3,6-bis(ethylamino)-2,7- dimethylxanthylum  
303 chloride monohydrate and 3,6-diamino-9-[2-(methoxycarbonyl)phenyl]xanthylum  
304 chloride trihydrate. *Acta Crystallogr. C* **57**, 657–659 (2001).
- 305 20. A. Ø. Madsen, Modeling and Analysis of Hydrogen Atoms, in. *Structure and bonding*  
306 (Springer, Berlin, 2012), 21–52



21. Wońska, M., Grabowsky, S., Dominiak, P. M., Woźniak, K. & Jayatilaka, D. Hydrogen atoms can be located accurately and precisely by x-ray crystallography. *Sci. Adv.* **2**, e1600192 (2016).
22. Allen, F. H. & Bruno, I. J. Bond lengths in organic and metal-organic compounds revisited: X - H bond lengths from neutron diffraction data. *Acta Crystallogr. B Struct. Sci.* **66**, 380–386 (2010).
23. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **71**, 3–8 (2015).
24. Mott, N. F. The scattering of electrons by atoms. *Proc. R. Soc. London A*, **127**, 658–665 (1930).
25. Bethe, H. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Ann. Phys.* **397**, 325–400 (1930).
26. Ogata, Y., Tsuda, K., Tanaka, M. Determination of the electrostatic potential and electron density of silicon using convergent-beam electron diffraction. *Acta Crystallogr. Sect. A Found. Adv.* **64**, (2008).
27. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).
28. Afonine, P. V., et al. Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 1153–1163 (2010).

## **Methods**

### Serial X-ray crystallography

We designed a sample support composed of a polyimide flat-faced plate with a size of  $4 \times 4$  mm<sup>2</sup> and a thickness of 20 µm (Protein Wave Corporation) and incorporated a metal pin

attached to a tail of the plate (right bottom in Fig. 1a). Microcrystal powder of rhodamine-6g (9-[2-(ethoxycarbonyl)phenyl]-3,6-bis-(ethylamino)-2,7-dimethylxanthylium) – chloride (Tokyo Chemical Industry) was mixed with low-viscosity liquid paraffin (Nacalai Tesque), and spread over the polyimide plate. The sample was sandwiched with another plate, and held between the plates.

A prepared sample was fixed through the pin onto a specially designed sample-pin mounter and vertically placed on a sample stage of beamline 2 at the SACLA XFEL facility<sup>29,30</sup>. The photon energy of XFEL and the beam size at the sample plane were adjusted to 15.0 keV and  $\sim 1 \mu\text{m}$ , respectively. The duration and the repetition rate of the pulses were 7 fs and 30 Hz, respectively. The pulse energy was  $\sim 160 \mu\text{J/pulse}$ . The sample plate was scanned in 2D directions over the whole plate plane, while the XFEL beam was exposed every  $10 \mu\text{m}$  and diffraction patterns recorded on an MX300-HS CCD detector (Rayonix), placed at 100 mm downward from the sample plane. All data collection was performed at room temperature.

CCD frames showing Bragg spots were first identified using a diffraction data processing program DIALS<sup>31</sup>. The highest  $s$ ,  $s_{\text{max}} = \frac{\sin \theta}{\lambda}$  (where  $\theta$  is half the scattering angle and  $\lambda$  is the wavelength of X-rays) of identified diffraction spots per frame was reconstructed on the sample plate with their data-taking position (Fig. 1b). Hit frames were processed using the CrystFEL suite<sup>32</sup>. Frames with no pixel value of  $> 1000$  or having few ( $< 30$  spots) or too many ( $> 300$  spots) spots were excluded, as spot indexing for these frames was less successful. Some frames taken during the beam was down (represented in blue horizontal lines in Fig. 1b) were also excluded. Diffraction spots were indexed with XGANDALF<sup>33</sup>, Dirax<sup>34</sup> and Mosflm<sup>35</sup> based on lattice parameters obtained from rotational ED patterns (see below). Total 46,272 of  $\sim 13,600$  frames were selected and indexed from a plate. The unit cell orientations are shown to cover all the directions without missing directions (Fig. 1d, e). Integrated intensities from

many frames were merged by process\_hkl in the CrystFEL suite. An initial structure was solved by direct phasing with SHELXT<sup>36</sup>. Datasets measured from two plates with success indexing rates of 67.8% and 71.4% were merged and used for the structure determination. The structure was refined and the estimated standard uncertainties of bonds and angles were calculated by SHELXL<sup>23</sup>. Hydrogen atoms were generated during the refinement using a riding model implemented in SHELXL, and the other restraints were removed at the end.

Distribution of crystal size was measured and estimated using an optical digital microscope (KEYENCE VHX-7000) and a function of particle analysis in OpenCV<sup>37</sup>. The crystal particle diameters were  $2.53 \pm 3.72 \mu\text{m}$  (N=377).

#### Electron crystallography

Rhodamine-6g powder crystals suspended in Novec7100 (3M), an inert solvent of a hydrofluorocarbon, were spread over a 200 mesh copper grid (Maxtaform HF-34) covered with holey carbon film (Quantifoil R1.2/1.3, R1/4). This hydrofluorocarbon solvent is effective for yielding a good distribution of crystals over the grid. After drying the solution, the grid was directly cooled with liquid nitrogen, and transferred into a CRYO ARM 300 electron microscope (JEOL) operated at an accelerating voltage of 300 kV and maintained at a specimen temperature of  $\sim 300$  or 98 K. A semi-automated data collection of rotational ED patterns was carried out by combined use of SerialEM<sup>38</sup> and Parallem<sup>39,40</sup> as previously described<sup>18,41</sup>. Crystals were illuminated with a parallel electron beam of  $\sim 5.1 \mu\text{m}$  diameter at an electron dose rate of  $\sim 0.02 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$ . Sequential diffraction frames per crystal were collected by continuously rotating the sample stage from  $-68^\circ$  to  $68^\circ$  at a goniometer rotation speed of  $1^\circ \text{ s}^{-1}$ , and patterns were recorded on a hardware-binned  $4\text{k} \times 4\text{k}$  pixel array of a DE64 direct

detection detector (Direct Electron). The nominal camera length was set to 800 mm and subsequently calibrated from gold sputtered on a carbon film at the end of the session.

Total 48 and 38 rotation series for the rt and cryo datasets, respectively (numbers of datasets are written in the same manner below), were first  $\times 2$  binned, and processed with DIALS for indexing and integrating of diffraction spots. We found that the crystal form was different from that previously reported<sup>19</sup>. Reduced datasets were grouped and sorted by KAMO<sup>42</sup>, which carried out scaling and merging by using Pointless<sup>43</sup>, XSCALE<sup>44</sup> and BLEND<sup>45</sup>. All merged clusters were examined for direct phasing by SHELXT and SHELXD<sup>46</sup> and for following initial refinement by SHELXL. A merged cluster derived from 23 and 27 datasets was selected by *R*1 criteria and number of contributed reflections. An initial structure was then refined using a model with elongated riding hydrogen bond and partially charged atoms as described briefly in the main text and in detail below. Scaling between the experimental data and the model was adjusted as previously described<sup>47</sup>. The geometrical restraints were finally removed, while *B*-factors were kept restrained with commands SIMU and ISOR.

The triclinic form of rhodamine-6g crystals was obtained by recrystallization as described<sup>21</sup>. Briefly, the original powder of tiny crystals was dissolved in a mixed solution of methanol, ethanol and water. Crystals were obtained by drying the solution over a few days, and ED data were collected at cryogenic temperature in the manner described above. Total 30 rotation series yielded the best cluster from 17 datasets.

#### Conversion of electron density to Coulomb potential

The electron-density model of an independent hydrogen atom is expressed at a radius *r* from the center of the electron density as,

$$\rho(\mathbf{r}) = \frac{1}{\pi a_0^3} \exp\left(-\frac{2r}{a_0}\right) \quad (1)$$

where  $a_0$  is to the Bohr radius. The electron density is smeared by thermal motion as

$$\langle \rho(\mathbf{r}) \rangle = \rho(\mathbf{r}) * P(\mathbf{u}, \mathbf{U}) \quad (2)$$

where  $*$  denotes convolution,  $\mathbf{u}$  defines a displacement vector from the mean position, and  $\mathbf{U}$  is the displacement amplitude. A probability distribution function,  $P$ , represents the Fourier transform of the Debye-Waller factor. The electron density can be converted to the Coulomb potential using an integrated form of the Poisson's equation:

$$\phi(\mathbf{r}) = \frac{Z_{\text{nuclear}}}{|\mathbf{r} - \mathbf{r}_{\text{nuclear}}|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (3)$$

where the nuclear charge ( $Z_{\text{nuclear}}$ ) is one for the hydrogen atom. The Coulomb potential is also converted to the smeared form, as shown in Eq. (2).

A one-dimensional electron-density curve, shown in Fig. 3a, was plotted based on the model in Eq. (1), with a peak position of 0.930 Å, derived from SHELXL<sup>23</sup>. Coulomb-potential curves and peak positions in Supplementary Table 3 were calculated using Eqs. (1–3) by adopting the nuclear positions of hydrogen atoms from an ND study<sup>22</sup>.

#### Measure of peak heights in hydrogen densities

The peak height for hydrogen atoms in the difference density maps,  $\Delta\sigma$ , was measured from the estimated base density level at the corresponding bonded non-hydrogen atom in the map (Fig. 2c, Fig. 3, b-e and Supplementary Table 2).

#### Atomic charges around ion-binding sites

Charges in pairs of amide-hydrogen and chloride atoms in rhodamine-6g were examined through a grid-search over all combinations, in which hydrogen is positive and  $\leq +1.0$ , and

chloride is negative and  $\geq -1.0$ , with a step size of 0.1. Scattering factors of partially charged atoms were calculated by linear combinations of those of neutral and ionized atoms, as in the previous study<sup>3</sup>. All the coordinates and *B*-factors were refined again for all combinations of the scattering factors. *R* values in the lowest resolution shell ( $s < 0.2 \text{ \AA}^{-1}$ ), in the other remaining shells ( $s \geq 0.2 \text{ \AA}^{-1}$ ), and including all the shells were monitored, as diffraction data at lower resolutions have higher sensitivity to charges. A model having charges of +0.4 for H15, +0.2 for H16 and -0.9 for CL1 provided the best *R* values for the rt-ED data in the lowest-resolution shell (Extended Data Fig. 5), while no significant improvement in *R* values was observed for the SX data. Bond lengths and *B*-factors in the charged rt-ED and neutral rt-ED models were plotted for model validation (data not shown), which shows no deviation between the models. Additional trials with assignment of positive or negative charge on the amide nitrogen atoms (N15 and N16) did not improve *R* values.

#### Conversion of Coulomb potential to electron density

The rt-ED experimental scattering factors were converted to the X-ray structure factors by the Mott–Bethe formula<sup>24,25</sup> with the obtained model structure using the GEMMI library (<https://github.com/project-gemmi/gemmi>). A difference Fourier map was calculated between the converted and model structure factors with neutral charges.

**Data and materials availability:** Crystallographic data was deposited at the Cambridge Crystallographic Data Centre, under a deposition number CCDC 2119567 (SX).

#### Methods references

- 448 29. Ishikawa, T., et al. A compact X-ray free-electron laser emitting in the sub-ångström region.  
449 *Nat. Photonics* **6**, 540–544 (2012).
- 450 30. Yabashi, M., Tanaka, H. & Ishikawa, T. Overview of the SACLA facility. *J. Synchrotron*  
451 *Radiat.* **22**, 477–484 (2015).
- 452 31. Clabbers, M. T. B., Gruene, T., Parkhurst, J. M., Abrahams, J. P. & Waterman, D. G.  
453 Electron diffraction data processing with DIALS. *Acta Crystallogr. Sect. D Struct. Biol.* **74**,  
454 506–518 (2018).
- 455 32. White, T. A., et al. Recent developments in CrystFEL. *J. Appl. Cryst.* **49**, 680–689 (2016).
- 456 33. Gevorkov, Y., et al. XGANDALF – extended gradient descent algorithm for lattice finding.  
457 *Acta Crystallogr. Sect. A Found. Adv.* **75**, 694–704 (2019).
- 458 34. Duisenberg, A. J. M. Indexing in single-crystal diffractometry with an obstinate list of  
459 reflections. *J. Appl. Crystallogr.* **25**, 92–96 (1992).
- 460 35. Leslie, A. G. W. & Powell, H. R. Processing diffraction data with mosflm, in *Evolving*  
461 *Methods for Macromolecular Crystallography* (Springer, Dordrecht, 2007), 41–51.
- 462 36. Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination.  
463 *Acta Crystallogr. Sect. A Found. Crystallogr.* **71**, 3–8 (2015).
- 464 37. Bradski, G. The OpenCV library, *Dr Dobbs J. Softw. Tools* (2000).
- 465 38. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of  
466 specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
- 467 39. Yonekura, K., Ishikawa, T. & Maki-Yonekura, S. A new cryo-EM system for electron 3D  
468 crystallography by eEFD. *J. Struct. Biol.* **206**, 243–253 (2019).
- 469 40. Hamaguchi, T., et al. A new cryo-EM system for single particle analysis. *J. Struct. Biol.*  
470 **207**, 40–48 (2019).

41. Takaba, K., Maki-Yonekura, S. & Yonekura, K. Collecting large datasets of rotational electron diffraction with ParallelEM and SerialEM. *J. Struct. Biol.* **211**, 107549 (2020).
42. Yamashita, K., Hirata, K. & Yamamoto, M. KAMO: towards automated data processing for microcrystals. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 441–449 (2018).
43. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 282–292 (2011).
44. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 133–144 (2010).
45. Foadi, J., et al. Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1617–1632 (2013).
46. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 479–485 (2010).
47. Kato, K., et al. Double-helix supramolecular nanofibers assembled from negatively curved nanographenes. *J. Am. Chem. Soc.* **143**, 5465–5469 (2021).
48. Koide, Y., Urano, Y., Hanaoka, K., Terai, T. & Nagano, T. Evolution of Group 14 rhodamines as platforms for near-infrared fluorescence probes utilizing photoinduced electron transfer. *ACS Chem. Biol.* **6**, 600–608 (2011).
49. Kjær, C., et al. Gas-phase action and fluorescence spectroscopy of mass-selected fluorescein monoanions and two derivatives. *Phys. Chem. Chem. Phys.* **22**, 9210–9215 (2020).
50. Bujdák, J., Iyi, N. & Sasai, R. Spectral properties, formation of dye molecular aggregates, and reactions in rhodamine 6G/layered silicate dispersions. *J. Phys. Chem. B* **108**, 4470–4477 (2004).



- 495 51. Chapman, M. & Euler, W. B. Rhodamine 6G structural changes in water/ethanol mixed  
496 solvent. *J. Fluoresc.* **28**, 1431–1437 (2018).
- 497 52. Madsen, A. Ø., Mason, S., & Larsen, S. A neutron diffraction study of xylitol: Derivation  
498 of mean square internal vibrations for H atoms from a rigid-body description. *Acta*  
499 *Crystallogr. Sect. B Struct. Sci.* **59**, 653–663 (2003).
- 500

**Acknowledgements:** We thank Kazunori Hata for designing a special sample-pin mounter, Yuko Kageyama for support of sample preparation, Kunio Hirata for advice on SX data analysis at the beginning of this study, and David B. McIntosh for help in improving the manuscript. This work was partly supported by JSPS KAKENHI (Grant Number 20K15764 to Ki.T.), JST-Mirai Program (Grant Number JPMJMI20G5 to K.Y.), JST CREST (Grant Number JPMJCR18J2 to K.Y., S.M-Y., Ki.T.) and the Cyclic Innovation for Clinical Empowerment (CiCLE) from the Japan Agency for Medical Research and Development, AMED (to K.Y.).

**Author contributions:**

Ki.T., S.M-Y., I.I., Ke.T., T.I., M.Y. and K.Y. conceived the project. Ki.T. and S.M-Y. prepared target specimens for SX and ED experiments. I.I., Ke.T. and M.Y. set up XFEL beamline for the measurement. K.Y. set up cryo-electron microscope for the measurement. Ki.T., S.M-Y., K.Y. collected SX data and Ki.T. and S.M-Y collected ED data. T.H., K.K., and H.N. supported SX data collection. Ki.T. processed the raw-data, solved structures and analysed them. Ki.T., S.M-Y., I.I., Ke.T., K.Y. discussed the results. Ki.T. and K.Y. wrote the manuscript. All authors joined discussion of the manuscript.

**Competing interests:** The authors declare no competing interests.

**Supplementary information** is available for this paper.

Supplementary Discussion  
Supplementary Tables 1 to 5.

523 **Extended Data Table 1** | Summary of data taking and crystallographic data statistics.

524 <sup>a</sup>  $R_{\text{split}}$  is shown for SX data, while  $R_{\text{merge}}$  is for ED data.

525 <sup>b</sup> estimated standard uncertainties calculated from full-matrix refinement with SHELXL<sup>23</sup>.

	SX	rt-ED	cryo-ED	triclinic-cryo-ED
Crystal	Microcrystal	Microcrystal	Microcrystal	Re-crystallized microcrystal
Temperature	R.T.	R.T.	~98 K	~98 K
Wavelength (Å)	0.833	0.0197	0.0197	0.0197
# of collected images	265,624	6,480	5,130	3,630
# of hit images	112,871	-	-	-
# of indexed images	78,106	5,952	3,737	2,420
# of images used for refinement	27,908	2,852	2,727	2,057
Spacegroup	<i>Pbca</i>	<i>Pbca</i>	<i>Pbca</i>	<i>P</i> -1
Unit Cell (Å, deg.)	14.88/15.11/23.31	14.74/15.05/22.88	14.63/14.53/22.83	9.14/11.16/13.32 /95.9/91.1/104.6
Resolution (Å)	9.65-0.82 (0.84-0.82)	11.44-0.90 (0.92-0.90)	22.83-0.90 (0.92-0.90)	13.40-0.90 (0.92-0.90)
# of unique reflections	5652 (556)	3651 (243)	3483 (217)	3748 (240)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	99.9 (99.2)
Multiplicity	2688 (993.7)	81.6 (84.5)	70.8 (72.1)	15.4 (15.5)
$R_{\text{split}}$ or $R_{\text{merge}}$ (%) <sup>a</sup>	8.97 (22.77)	62.0 (372.4)	108.1 (475.3)	98.1 (293.5)
$CC_{1/2}$ (%)	98.4 (93.8)	99.6 (56.0)	99.3 (27.9)	91.7 (27.5)
$\langle I/\sigma(I) \rangle$	12.36 (3.81)	8.97 (0.83)	5.04 (0.81)	2.22 (0.55)
Resolution (Å)	9.65-0.82	11.44-0.90	11.41-0.90	13.40-0.90
$R1$ ( $F > 4\sigma(F)$ ), $R1$ (all $F$ )	0.110, 0.110	0.161, 0.200	0.206, 0.234	0.181/0.250
$Z$	8	8	8	2
$D_{\text{calc}}$ (g/cm <sup>3</sup> )	1.235	1.254	1.333	1.249
# of atoms, C/N/O/Cl/H	28/2/3.5/1/31	28/2/3/1/31	28/2/3.5/1/31	28/2/4/1/31
# of parameters	322	310	325	337
$\langle B \rangle$ for non-H atoms (Å <sup>2</sup> )	8.36	42.5	5.82	5.12
$\langle \text{e.s.u.} \rangle^b$ for bond lengths (Å)	0.003	0.011	0.016	0.019
$\langle \text{e.s.u.} \rangle^b$ for bond angles (deg.)	0.194	0.792	1.44	1.34
CCDC #	2119567	-	-	-

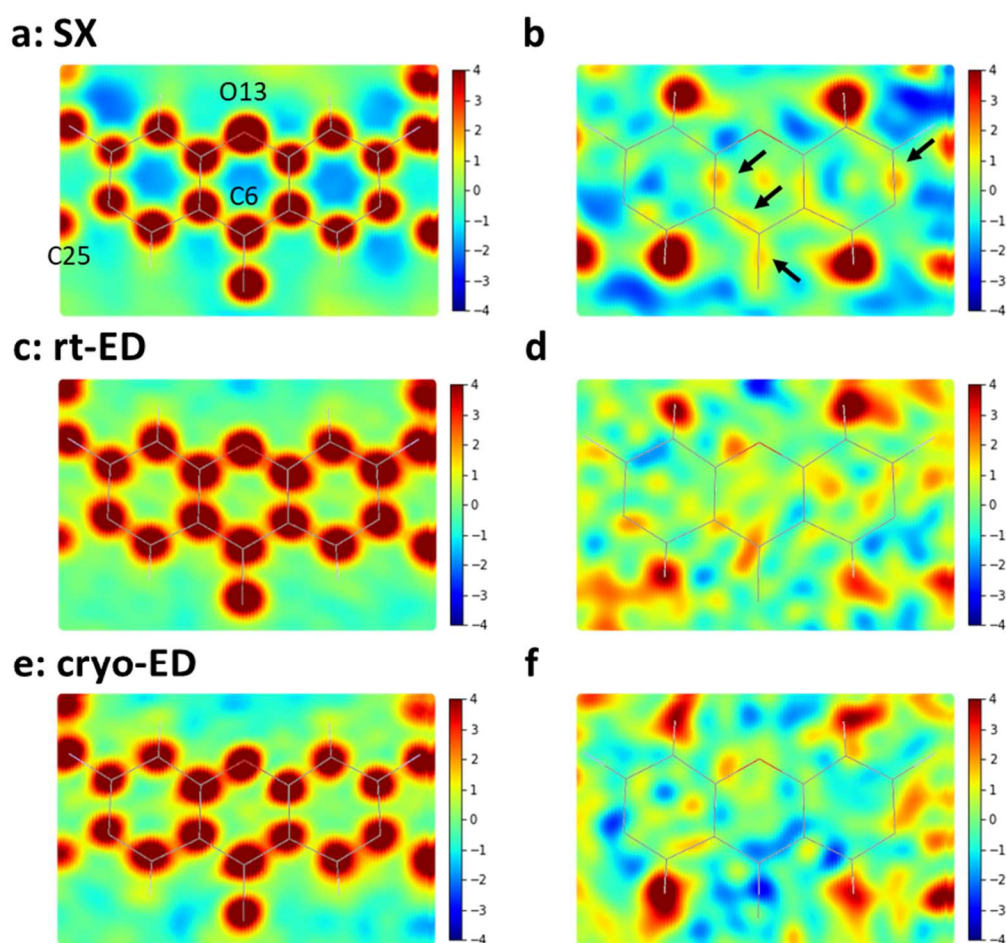
526

527

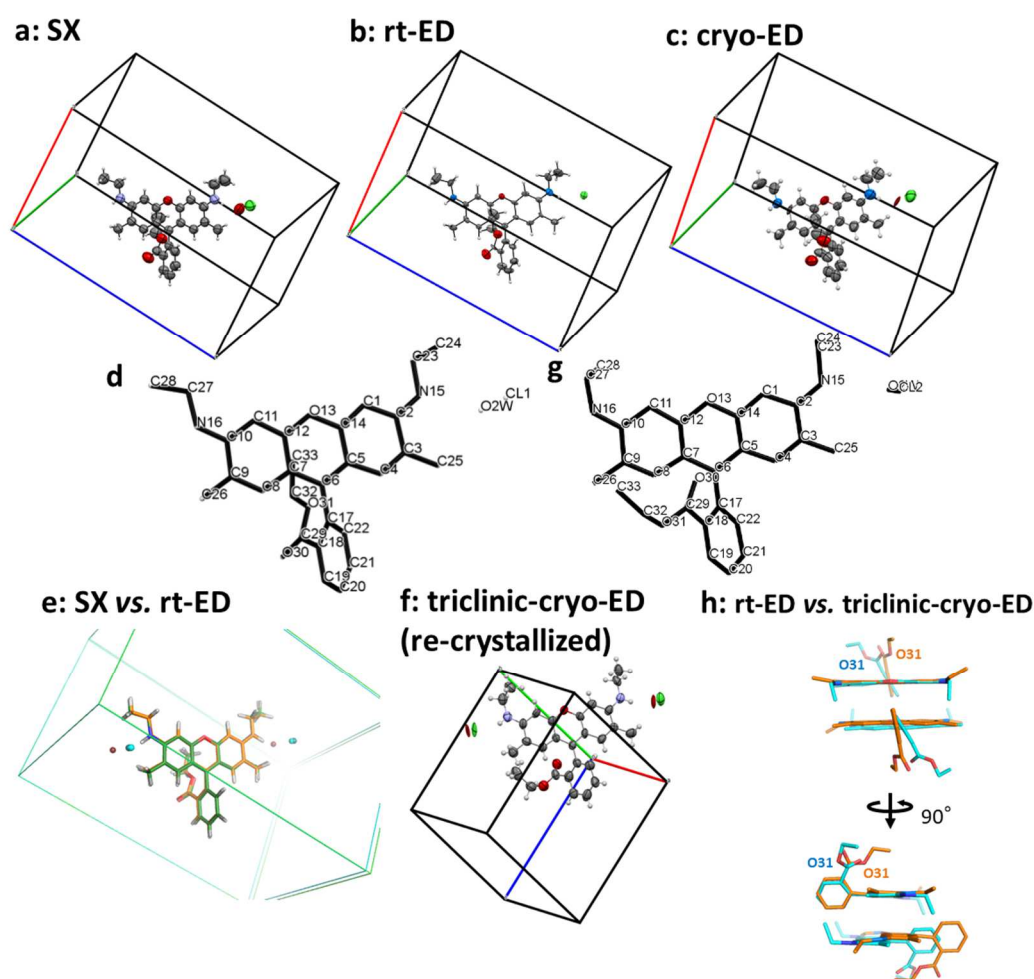
**Extended Data Table 2 | Crystal systems of rhodamine-6g compounds in Cambridge Structural Database and in this work.**

<sup>a</sup> Atomic composition is shown for one rhodamine-6g molecule.

Unit formula <sup>a</sup>	Water	Ions, solvents	System	Ref.
C <sub>28</sub> H <sub>32</sub> ClN <sub>2</sub> O <sub>3.5</sub>	0.5	Cl	Orthorhombic	This work ('SX' , 'cryo-ED')
C <sub>28</sub> H <sub>31</sub> ClN <sub>2</sub> O <sub>3</sub>	0	Cl	Orthorhombic	This work ('rt-ED')
C <sub>28</sub> H <sub>33</sub> ClN <sub>2</sub> O <sub>4</sub>	1	Cl	Triclinic	CCDC#164693, this work ('triclinic-cryo-ED')
C <sub>28</sub> H <sub>33</sub> IN <sub>2</sub> O <sub>4</sub>	1	I	Orthorhombic	CCDC#129329
C <sub>29</sub> H <sub>34</sub> Cl <sub>2</sub> Zn <sub>0.5</sub> N <sub>2</sub> O <sub>3.5</sub>	0	Cl, Zn, C <sub>2</sub> H <sub>5</sub> OH	Triclinic	CCDC#1298021
C <sub>29</sub> H <sub>32</sub> Cl <sub>3</sub> Sn <sub>0.5</sub> N <sub>3</sub> O <sub>3</sub>	0	Cl, Sn, CH <sub>3</sub> CN	Triclinic	CCDC#674356



**Extended Data Fig. 1 | 2D slices of electron density and Coulomb potential maps at the plane of the xanthene ring.** (a) A slice of the  $F_o$  map calculated from the SX data and overlaid with the model. Some atoms are labelled. (b) A slice of the hydrogen-omitted map ( $mF_o - DF_c$ ) from the SX. (c, d) The same slices as in (a) and (b) but calculated from the rt-ED data. (e, f) The same as in (c) and (d) but from the cryo-ED data. Arrows in (b) indicate residual densities likely representing covalent electrons. The color display is gradually changed from  $-4\sigma$  to  $+4\sigma$  as in a gradient bar at the right of each map.



543

544 **Extended Data Fig. 2 | Atomic models of rhodamine-6g obtained in this study.** (a-c) The

545 SX model determined from the orthorhombic crystal placed in the unit cell (a), the rt-ED (b)

546 and the cryo-ED (c) models from the orthorhombic crystal. (d) The atomic model in (a) with

547 labels of non-hydrogen atoms. (e) Overlay of the models in (a) and (b) along with the unit cells.

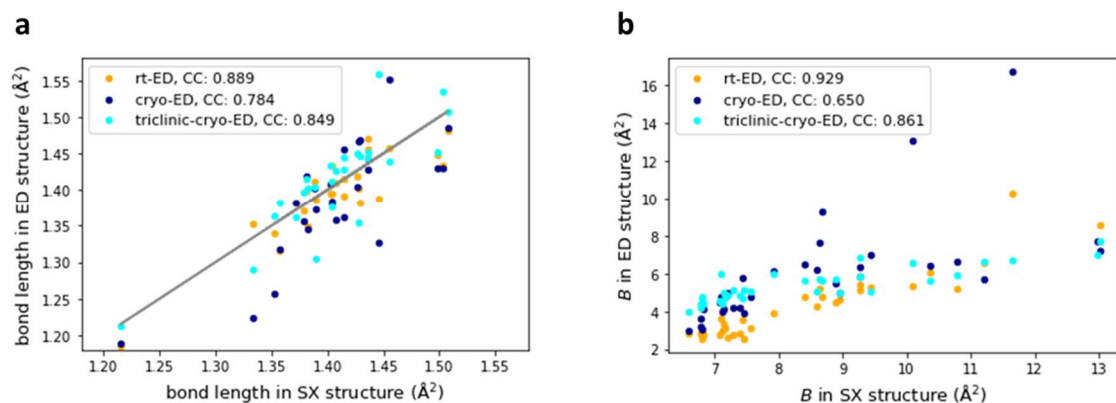
548 (f) The triclinic cryo-ED model from the recrystallized crystal. (g) The model in (f) with labels

549 of non-hydrogen atoms. (h) Rhodamine-6g dimers in the orthorhombic (orange) and triclinic

550 crystals (cyan). The rt-ED (b) and triclinic-cryo-ED (f) models are overlaid by adjusting the

551 monomers at the top onto each other.

552

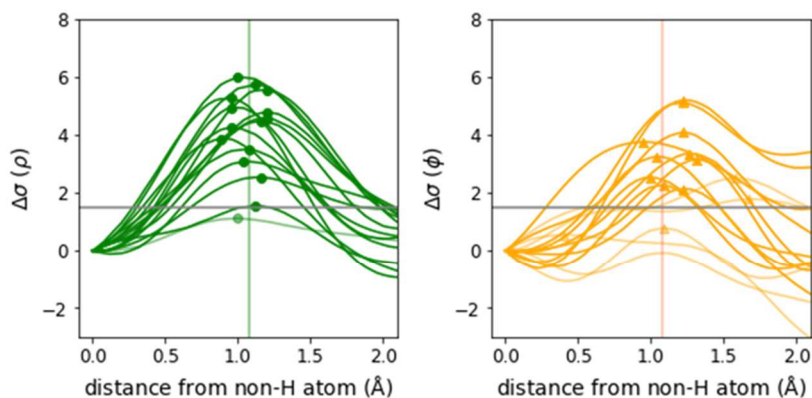


**Extended Data Fig. 3 | Similarities among the atomic models of the SX and ED structures.**

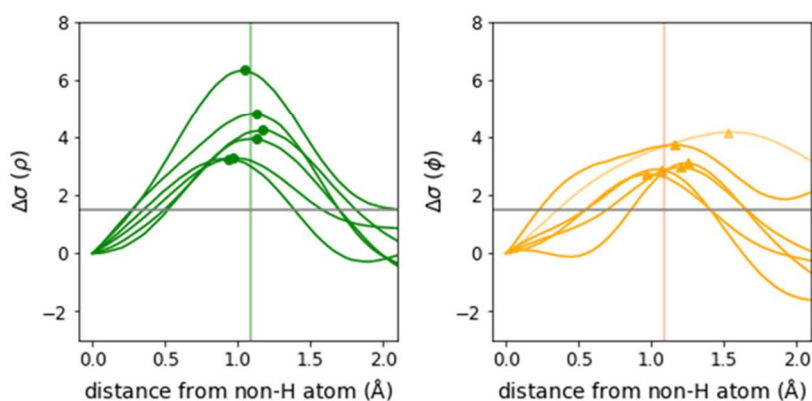
**(a, b)** Relation plots between the SX and ED structures in bond lengths (a) and in  $B$ -factors (b).

Correlation coefficients between the SX and ED models are also shown in the graphs.

### a: C-H<sub>3</sub> (methyl)



### b: C-H<sub>2</sub> (methylene)



558

559 **Extended Data Fig. 4 | Peak positions of hydrogen densities in methyl and methylene**

560 **groups. (a)** The same density plots as in Fig. 3b and C but in C-H<sub>3</sub> (methyl) bonds from the

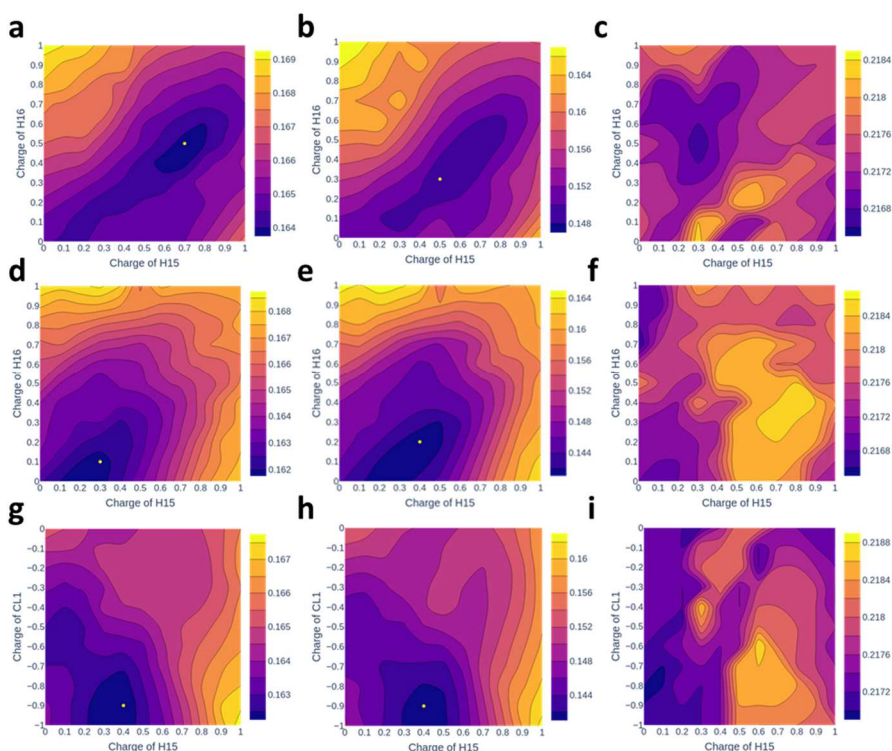
561 SX and rt-ED structures in green and yellow, respectively. **(b)** The same plots as in (a) but in

562 C-H<sub>2</sub> (methylene) bonds. Gray horizontal lines refer to a density level of  $1.5\sigma$ , and vertical

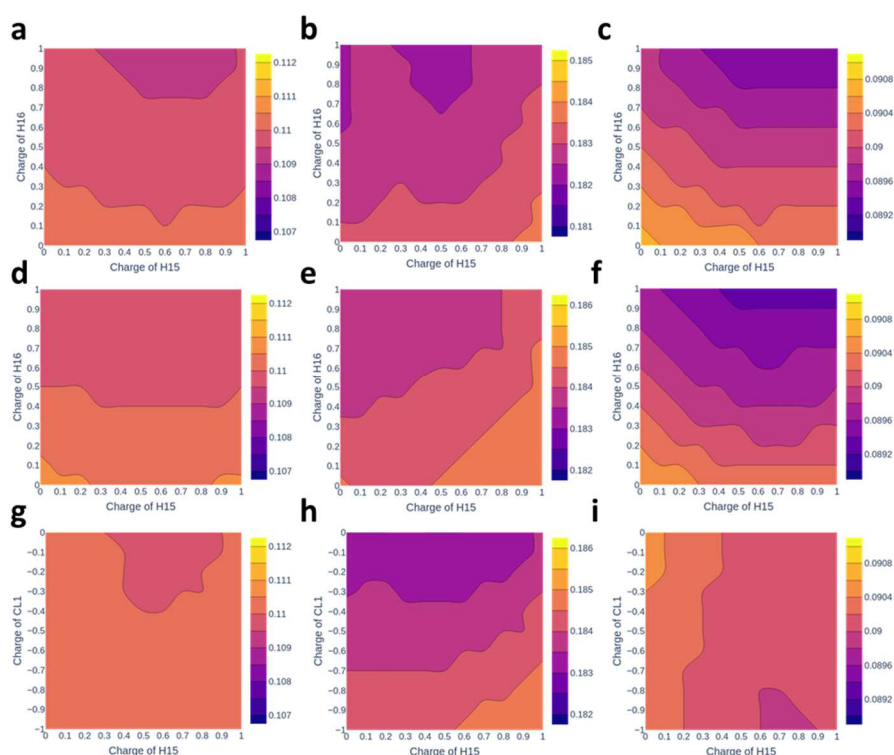
563 lines represent the positions of the hydrogen nuclei, obtained by ND studies<sup>22</sup>.

564

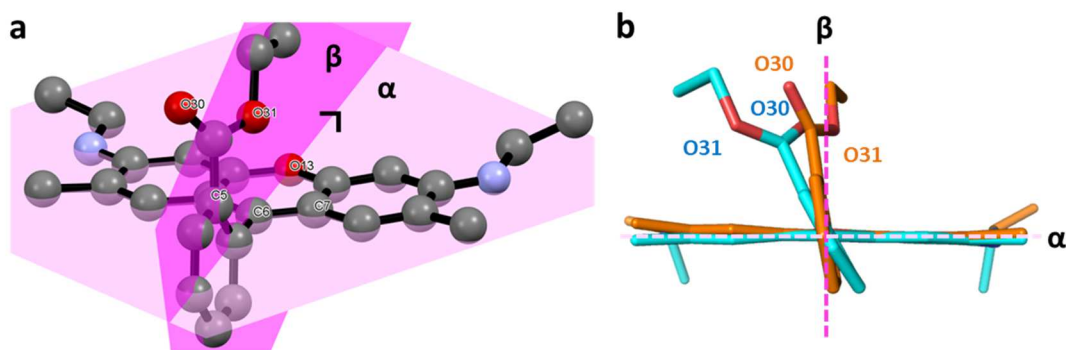




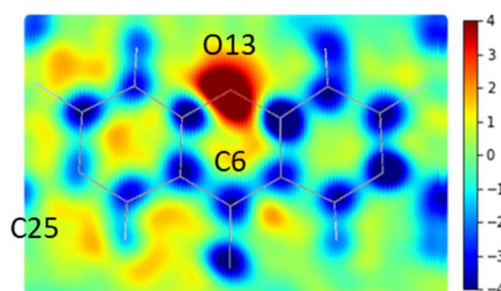
**Extended Data Fig. 5 | Diagrams of  $R$  value variations for refinement of the rt-ED structure with the charged-atom model.** (a – i) Charge values were varied for specific atoms shown in the horizontal and vertical axes, and  $R$  values are represented in gradient colors according to the gradient bar on the right in each diagram.  $R$  values for the data with  $F_o > 4\sigma$  in the whole resolution shells (a, d, g), in the lowest resolution shell of  $s < 0.2 \text{ \AA}^{-1}$  (b, e, h) and in the other remaining shells ( $s \geq 0.2 \text{ \AA}^{-1}$ ) (c, f, i). Amide-hydrogen atoms (H15 and H16) were exclusively charged with given values along the axes in (a – c). Given a charge of -0.9 to a chloride atom (CL1), H15 and H16 were positively charged in (d – f). Given +0.2 to H16, H15 was positively and CL1 was negatively charged in (g – i). The lowest  $R$  value is shown with yellow markers in d, e, g and h. All possible combinations of charges for H15, H16 and CL1 were examined as in [Methods](#).



**Extended Data Fig. 6 | Diagrams of  $R$  value variations for refinement of the SX structure with the charged-atom model. (a – i)** Done in the same way as in [Extended Data Fig. 5](#), but for the SX data and model. No grid points giving  $R$  value minima are depicted and variations in  $R$  values are smaller than those for the ED data (c.f. [Extended Data Fig. 5](#)).



**Extended Data Fig. 7 | Configuration of the xanthene ring and the ethoxycarbonyl tail in the SX structure of rhodamine-6g.** (a) The plane including the xanthene ring shown in pale pink is labeled ‘ $\alpha$ ’. The plane ‘ $\beta$ ’ in magenta is defined as a cross-section cutting through the center of the xanthene ring along O13 – C6 and perpendicular to the plane  $\alpha$ . (b) A side view of (a). The ether oxygen (O31) in the ethoxycarbonyl tail faces to the xanthene-ring side and is close to the central plane  $\beta$ : only 0.163 Å and 0.038 Å apart from the plane in the SX- and cryo-ED structures, respectively. The distances of which are also listed in [Supplementary Table 1](#).



**Extended Data Fig. 8 | A 2D slice of a difference Fourier electron density map converted from rt-ED.** Calculated between the converted  $F_o$  from the rt-ED data and  $F_c$  for X-rays with neutral changes. Displayed as in [Extended Data Fig. 1](#).

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21

# Hydrogen properties in an organic molecule revealed by XFEL and electron crystallography

Kiyofumi Takaba, Saori Maki-Yonekura, Ichiro Inoue, Kensuke Tono, Tasuku Hamaguchi, Keisuke Kawakami, Hisashi Naitow, Tetsuya Ishikawa, Makina Yabashi, and Koji Yonekura

Correspondence to: [yone@spring8.or.jp](mailto:yone@spring8.or.jp)

**This document file includes:**

## Supplementary Discussion

### Supplementary Tables 1 to 5

## Supplementary Discussion

### Comparison of rhodamine-6g structures

In the SX, rt-ED, cryo-ED structures, the plane of the stacked rings is approximately perpendicular to the plane of the phenolic ring (C17 to C22), which would impede resonance between rings and hinder electron transfer<sup>48,49</sup>. The configuration of the dimers in this orthorhombic crystal is nearly identical to that in an iodide salt crystal (CCDC#129329; [Extended Data Table 2](#)), but different from the triclinic crystal structure with chlorine determined previously by XRD from a single crystal, in which the planes of the two rings are oblique to each other rather than normal<sup>19</sup> ([Extended Data Fig. 2, f-h](#)). This particular dye exhibits a tendency to aggregate at higher concentrations, and hence, the effects of dimerization have been studied with respect to its photochemical properties<sup>50,51</sup>. A similar orientation of the H-type dimer to the SX and ED orthorhombic crystal structures was also proposed in a recent theoretical calculation<sup>51</sup>. However, the oxycarboxyl plane is flipped, and the configuration is different from that in the experimentally obtained structures here. The monomer conformation of the SX, rt-ED and cryo-ED structures likewise differs from the organic-solvent and metal complex structures of rhodamine-6g or its derivative (CCDC#1298021; CCDC#674356; [Extended Data Table 2](#)).

### Further consideration on variations in hydrogen densities

In addition to the polarity of bonding and the motion of hydrogen atoms indicated in the main text, other factors such as smearing by truncation of information in reciprocal space, isotropic and anisotropic atomic displacement parameters, and anharmonicity along the bonding direction<sup>20</sup> should also be considered for quantification of the chemical characteristics at each site through electron-density and Coulomb-potential maps. The requirement for deconvolution

of these factors, however, is beyond the precision limits of the current SX and ED experiments. Practically, observations like those in this study could follow hybrid-analyses of XRD, ND and/or theoretical calculations<sup>20,21,52</sup>. Well-resolved hydrogen atoms, as seen in aromatic and amide groups, would be suitable candidates for further analyses.

### Supplementary Table 1 | Geometries of the atomic models of rhodamine-6g

Values in parentheses are estimated standard uncertainties derived from the full-matrix least squares refinement.

<sup>a</sup> Defined as a cross-section cutting through the center of the xanthene ring along O13 – C6 and perpendicular to the plane of the xanthene ring (see main text and [Extended Data Fig. 7](#)).

		$d_{SX}$ (Å)	$d_{fit-ED}$ (Å)	$d_{cryo-ED}$ (Å)	$d_{triclinic-cryo-ED}$ (Å)			$d_{SX}$ (Å)	$d_{fit-ED}$ (Å)	$d_{cryo-ED}$ (Å)	$d_{triclinic-cryo-ED}$ (Å)
C1	C2	1.427 (3)	1.419 (11)	1.405 (16)	1.450 (18)	C17	C18	1.428 (3)	1.404 (10)	1.440 (14)	1.477 (18)
C1	C14	1.382 (3)	1.411 (10)	1.402 (14)	1.404 (18)	C17	C22	1.403 (3)	1.434 (10)	1.467 (15)	1.354 (21)
C2	C3	1.464 (3)	1.458 (10)	1.551 (14)	1.440 (19)	C18	C19	1.407 (3)	1.407 (11)	1.358 (16)	1.428 (21)
C2	N15	1.372 (3)	1.349 (10)	1.471 (17)	1.462 (20)	C18	C29	1.503 (3)	1.433 (10)	1.430 (15)	1.535 (20)
C3	C4	1.507 (3)	1.500 (12)	1.345 (14)	1.403 (19)	C19	C20	1.402 (4)	1.397 (12)	1.409 (17)	1.434 (22)
C3	C25	1.436 (3)	1.457 (11)	1.429 (16)	1.453 (18)	C20	C21	1.389 (5)	1.385 (12)	1.374 (19)	1.304 (23)
C4	C5	1.415 (2)	1.391 (10)	1.424 (14)	1.417 (18)	C21	C22	1.405 (4)	1.395 (12)	1.383 (17)	1.411 (21)
C5	C6	1.438 (3)	1.397 (10)	1.361 (14)	1.444 (17)	C23	C24	1.499 (5)	1.480 (14)	1.486 (19)	1.507 (24)
C5	C14	1.414 (3)	1.416 (11)	1.455 (15)	1.428 (17)	C27	C28	1.508 (5)	1.452 (16)	1.376 (18)	1.457 (21)
C6	C7	1.510 (2)	1.478 (10)	1.361 (14)	1.444 (17)	C29	O30	1.214 (3)	1.185 (9)	1.225 (13)	1.290 (18)
C6	C17	1.437 (3)	1.469 (10)	1.358 (13)	1.406 (18)	C29	O31	1.334 (3)	1.354 (10)	1.188 (14)	1.214 (17)
C7	C8	1.433 (3)	1.396 (10)	1.443 (15)	1.442 (18)	O31	C32	1.446 (6)	1.387 (15)	1.327 (24)	1.558 (27)
C7	C12	1.382 (3)	1.382 (11)	1.443 (15)	1.442 (18)	C32	C33	1.460 (3)	1.410 (11)	1.225 (13)	1.290 (18)
C8	C9	1.517 (3)	1.498 (11)	1.468 (18)	1.506 (20)						
C9	C26	1.456 (3)	1.456 (10)	1.451 (15)	1.470 (19)			$d_{SX}$ (Å)	$d_{fit-ED}$ (Å)	$d_{cryo-ED}$ (Å)	$d_{triclinic-cryo-ED}$ (Å)
C9	C10	1.358 (3)	1.338 (11)	1.468 (15)	1.447 (19)	O30	Plane $\beta^a$	2.701	2.712	2.676	4.114
C10	N16	1.426 (3)	1.383 (10)	1.256 (14)	1.363 (18)	O31	Plane $\beta$	0.163	0.038	0.512	2.253
C10	C11	1.389 (3)	1.397 (10)	1.468 (15)	1.447 (19)						
C11	C12	1.370 (2)	1.371 (9)	1.357 (13)	1.396 (17)						
C12	O13	1.379 (2)	1.361 (10)	1.304 (13)	1.357 (16)						
O13	C14	1.463 (4)	1.458 (12)	1.317 (14)	1.381 (18)						
N15	C23	1.462 (4)	1.435 (12)	1.376 (18)	1.457 (21)						
N16	C27	1.427 (3)	1.419 (11)	1.405 (16)	1.450 (18)						

59 **Supplementary Table 2 | A list of hydrogen peak positions from bonded atoms. See also**  
60 **Fig. 3.**

	X-H <sub>peak</sub> , SX (Å)	peak height, SX (Δσ)	X-H <sub>peak</sub> , rt-ED (Å)	peak height, rt-ED (Δσ)	ΔX-H <sub>peak</sub> (Å)
H1	1.047	4.997	1.219	2.983	0.172
H4	0.891	4.795	1.080	2.599	0.189
H8	0.930	4.588	0.982	4.599	0.052
H11	0.930	3.810	1.125	3.688	0.195
H19	1.007	3.968	1.041	3.279	0.033
H20	1.007	2.918	0.997	1.415	-0.010
H21	0.891	2.889	1.040	2.860	0.148
H22	0.892	3.526	0.993	2.357	0.102
H25A	0.919	1.255	2.840	1.738	1.921
H25B	1.120	1.905	1.203	3.783	0.083
H25C	0.961	3.665	1.058	4.422	0.098
H26A	1.000	3.403	0.914	2.356	-0.086
H26B	1.200	3.503	0.770	3.084	-0.430
H26C	1.199	3.944	0.962	1.936	-0.237
H24A	0.880	2.878	1.251	1.502	0.371
H24B	0.960	0.697	0.048	0.028	-0.912
H24C	0.200	0.061	1.252	1.397	1.052
H28A	1.080	3.444	1.348	0.096	0.268
H28B	1.160	1.663	1.011	1.112	-0.149
H28C	1.240	2.541	1.443	1.090	0.202
H33A	1.039	2.413	2.309	0.507	1.270
H33B	0.800	1.396	0.000	0.000	-0.800
H33C	1.239	0.761	1.396	1.453	0.157
H23A	0.848	2.564	1.076	2.209	0.228
H23B	0.970	1.865	1.271	0.299	0.301
H27A	1.212	3.910	0.440	1.001	-0.772
H27B	1.092	5.039	0.978	0.437	-0.114
H32A	0.970	2.684	0.830	1.906	-0.139
H32B	1.132	2.008	0.684	3.503	-0.448
H15	0.789	4.096	0.775	2.355	-0.014
H16	0.752	3.516	1.047	4.484	0.295

61

**Supplementary Table 3 | Peak positions of the hydrogen atom from the bonded atom for X-ray, neutron and electron beams.**

$X - H_{IAM}^{XRD}$  refers to the peak position of the hydrogen atom from the bonded non-hydrogen atom for XRD under IAM, and values in the column are adopted from ref 23; and  $X - H_{IAM}^{ND}$  is the same as  $X - H_{IAM}^{XRD}$  but determined by ND<sup>22</sup>.  $\Delta_{XRD-ND}$  refers to difference between  $X - H_{IAM}^{XRD}$  and  $X - H_{IAM}^{ND}$ . Values in  $X - H_{IAM}^{ED}$  were converted from those in  $X - H_{IAM}^{XRD}$  using the Poisson's equation for ED under IAM (see the supplementary materials).  $\Delta_{ED-ND}$  and  $\Delta_{ED-XRD}$  are differences between peak locations in ED ( $X - H_{IAM}^{ED}$ ) and XRD ( $X - H_{IAM}^{ND}$ ) and in ED and ND ( $X - H_{IAM}^{XRD}$ ).

	$X - H_{IAM}^{XRD}(\text{\AA})$	$X - H_{IAM}^{ND}(\text{\AA})$	$\Delta_{XRD-ND}(\text{\AA})$	$X - H_{IAM}^{ED}(\text{\AA})$	$\Delta_{ED-ND}(\text{\AA})$	$\Delta_{ED-XRD}(\text{\AA})$
C-H <sub>3</sub>	0.960	1.077	-0.117	1.084	0.007	0.124
C-H <sub>2</sub>	0.970	1.092	-0.122	1.109	0.017	0.139
C-H <sub>aromatic</sub>	0.930	1.083	-0.153	1.094	0.011	0.164
N-H (CCNH)	0.860	1.027	-0.167	1.049	0.022	0.189



72 **Supplementary Table 4 | Geometry of hydrogen bonding around amides**

73 <sup>a</sup> The disposition of the three atoms is not allowed for hydrogen-bond formation.

	D – H ... A	D – H (Å)	D ... A (Å)	H ... A (Å)	D – H ... A (deg.)
SX	N15 – H15 ... CL1	0.86	3.529 (3)	2.82 (5)	141 (4)
	N16 – H16 ... CL1'	0.86	3.433 (3)	2.70 (4)	144 (3)
	N15 – H15 ... O2W	0.86	3.413 (7)	2.73 (5)	138 (4)
	N16 – H16 ... O2W' <sup>a</sup>	0.86	[3.696 (8)]	[3.31 (4)]	[110 (3)]
rt-ED	N15 – H15 ... CL1	1.049	3.35 (1)	2.48 (4)	140 (3)
	N16 – H16 ... CL1'	1.049	3.41 (1)	2.58 (5)	136 (3)
	N15 – H15 ... O2W	1.049	-	-	-
	N16 – H16 ... O2W'	1.049	-	-	-
cryo-ED	N15 – H15 ... CL1	1.049	3.78 (2)	3.03 (7)	129 (5)
	N16 – H16 ... CL1'	1.049	3.32 (1)	2.43 (6)	142 (4)
	N15 – H15 ... O2W	1.049	3.17 (3)	2.30 (8)	139 (5)
	N16 – H16 ... O2W' <sup>a</sup>	1.049	[3.78 (3)]	[3.44 (7)]	[102 (3)]
triclinic-cryo-ED	N15 – H15 ... CL2	1.049	3.22 (2)	2.41 (8)	157 (7)
	N16 – H16 ... CL1	1.049	2.92 (2)	2.13 (8)	152 (7)
	N15 – H15 ... O2W	1.049	2.82 (3)	2.02 (8)	155 (7)
	N16 – H16 ... O1W	1.049	3.44 (2)	2.66 (8)	151 (6)

74

75 **Supplementary Table 5 | Parameterization of scattering factors of partially charged**  
76 **atoms for structure refinement with SHELXL<sup>23</sup>**

Element	$a_1$	$b_1$	$a_2$	$b_2$	$a_3$	$b_3$	$a_4$	$b_4$	$c$
Cl <sup>0.1-</sup>	1.412	2.804	-1.793	474.3	3.026	16.62	-5.803	1899	0.1936
Cl <sup>0.2-</sup>	1.343	2.66	-2.472	348	3.039	15.85	-7.947	1321	0.1888
Cl <sup>0.3-</sup>	1.298	2.567	-3.234	304.5	3.045	15.36	-11.34	1198	0.1856
Cl <sup>0.4-</sup>	1.262	2.493	-4.025	282.7	3.049	14.97	-14.82	1144	0.1829
Cl <sup>0.5-</sup>	1.231	2.428	-4.828	269.4	3.053	14.63	-18.33	1113	0.1805
Cl <sup>0.6-</sup>	1.203	2.369	-5.635	260.5	3.057	14.33	-21.84	1093	0.1783
Cl <sup>0.7-</sup>	1.176	2.313	-6.444	254.1	3.06	14.06	-25.36	1078	0.1761
Cl <sup>0.8-</sup>	1.151	2.26	-7.254	249.2	3.064	13.81	-28.87	1067	0.174
Cl <sup>0.9-</sup>	1.127	2.209	-8.063	245.3	3.068	13.57	-32.39	1059	0.1718
Cl <sup>1.0-</sup>	1.104	2.16	-8.871	242.1	3.071	13.34	-35.9	1052	0.1697
H <sup>0.1+</sup>	0.1983	3.845	0.8256	193.8	0.4125	24.17	3.422	958.4	0.01586
H <sup>0.2+</sup>	0.2464	4.74	1.762	224.9	0.5602	32.87	7.059	1022	0.01798
H <sup>0.3+</sup>	0.2903	5.538	2.775	247.7	0.7641	41.21	10.78	1068	0.0198
H <sup>0.4+</sup>	0.329	6.195	4.54	1101	1.007	47.99	3.833	264.8	0.02127
H <sup>0.5+</sup>	0.364	6.742	4.913	277.5	1.273	53.13	18.33	1125	0.02246
H <sup>0.6+</sup>	0.397	7.219	6.005	287	1.55	57.07	22.13	1143	0.02349
H <sup>0.7+</sup>	0.4294	7.651	7.104	294.4	1.835	60.18	25.94	1157	0.02442
H <sup>0.8+</sup>	0.4618	8.053	8.209	300.4	2.126	62.71	29.75	1169	0.02527
H <sup>0.9+</sup>	0.4945	8.431	9.319	305.5	2.422	64.82	33.57	1178	0.02607
H <sup>1.0+</sup>	0.5277	8.792	10.43	309.7	2.721	66.63	37.4	1186	0.02683

77  
78