CHEMICAL STRUCTURE

# Electron diffraction determines molecular absolute configuration in a pharmaceutical nanocrystal

**Petr Brázda[1]*, Lukáš Palatinus[1], Martin Babor[2,3]**

**Determination of the absolute configuration of organic molecules is essential in drug development and the subsequent approval process. We show that this determination is possible through electron diffraction using nanocrystalline material. Ab initio structure determination by electron diffraction has so far been limited to compounds that maintain their crystallinity after a dose of one electron per square angstrom or more. We present a complete structure analysis of a pharmaceutical cocrystal of sofosbuvir and L-proline, which is about one order of magnitude less stable. Data collection on multiple positions of a crystal and an advanced-intensity extraction procedure enabled us to solve the structure ab initio. We further show that dynamical diffraction effects are strong enough to permit unambiguous determination of the absolute structure of material composed of light scatterers.**

About 20% of compounds crystallize in noncentrosymmetric space groups. They often possess unusual physical properties such as second harmonic generation, multiferroic properties, or show specific interactions with biological systems. Structure solution of these materials using diffraction techniques employs algorithms, which cannot determine the absolute structure of crystals or, in particular, the absolute configuration of chiral molecules (*1*). Therefore, each of two structure solutions related by inversion is obtained with 50% probability for noncentrosymmetric materials. However, determination of the absolute structure by diffraction techniques plays an irreplaceable role in characterization of these materials. In single-crystal x-ray diffraction, the absolute structure is determined owing to the resonant scattering effect, which breaks the inversion symmetry of the intensities of $hkl$ and $\bar{h}\bar{k}\bar{l}$ reflections, called Bijvoet pairs (*2*). This effect is very weak for light atoms so excellent data quality is required for the determination of absolute structure of organic compounds, often necessitating the use of synchrotron radiation. Electron diffraction data permit, in principle, determination of the absolute structure of the crystals in a more robust manner than x-ray diffraction. The inversion symmetry in diffraction space is broken because of dynamical diffraction effects (multiple scattering of diffracted electrons). These strongly depend on crystal orientation, i.e., on the reflections that are at the same time close to the Bragg condition. These effects are about one order of magnitude larger than the resonant scattering of x-rays using Cu Kα radiation (Fig. 1). The possibility of absolute structure determination by electron diffraction has been demonstrated on stable, inorganic materials (*3*–*7*), but beam-sensitive compounds containing only light atoms remain a challenge. In the diffraction data of these materials, two phenomena impeding determination of the absolute structure combine: (i) weak scatterers showing weak dynamical diffraction effects and (ii) limited quality of diffraction data, i.e., low resolution, loss of crystallinity due to electron beam damage, and, in the case of serial electron diffraction (*8*, *9*), different crystal thicknesses and unknown crystal orientations. Crystals of soft materials such as organics or metal–organic frameworks also exhibit high mosaicity (*10*), which further complicates the analysis of the diffraction data.

Pharmaceutical substances are an important group of materials, which often form only micro- or nanocrystals. Important breakthroughs in the structure solution of these materials have been made recently (*11*–*15*) by the use of electron diffraction tomography (EDT) techniques (*16*–*21*). However, a key part of a complete analysis of chiral pharmaceuticals (*22*) by electron diffraction—absolute structure determination—has not been addressed.

Pharmaceutical compounds often form crystals that are sensitive to an electron beam, and radiation damage (*23*–*25*) is a limiting factor for structure determination of organic molecular crystals by EDT. Structures of organic compounds determined by EDT have been limited to crystals stable enough to allow the collection of complete or almost complete diffraction data by irradiation of the same crystal spot (*11*, *12*, *14*, *15*, *24*). This work, however, shows that compounds exist in which crystallinity is seriously damaged at doses well below 1 $e^-$ Å$^{-2}$. In such cases, a complete dataset cannot be collected on a single part of the crystal. Serial electron crystallography, which is analogous to serial femtosecond crystallography (*26*–*28*), allows for construction of a single diffraction dataset from a large number of crystals.

In this work on a pharmaceutical cocrystal (*29*) of enantiopure sofosbuvir and L-proline (*30*), we show that the structure solution is possible even for materials with stability lower than 0.2 $e^-$ Å$^{-2}$, and that dynamical diffraction effects are sufficiently strong even for highly beam-sensitive light-atom structures (chemical formula sum $C_{27}H_{38}F_1N_4O_{11}P_1$) to allow an unambiguous determination of their absolute structure.

Data measurements showed that it was possible to collect only up to three diffraction patterns on a particular crystal spot, which corresponded to a dose smaller than 0.2 $e^-$ Å$^{-2}$. The ribbon-shaped crystals allowed scanning of the crystal spot by spot. This strategy permitted collection of more-complete datasets from a single crystal (Fig. 2A) at the cost of sparsely sampled reciprocal space. The crystals were strongly bent and twisted and thus the orientation of the frames was not known precisely despite being collected on a single crystal. Moreover, the scale factor of each frame could vary. These data thus resembled the data obtained by serial electron crystallography. Because of sparse sampling, crystal orientation uncertainty (crystals twisted up to 10°), and unknown diffraction pattern (frame) scales, standard data processing for the intensity extraction necessary for the structure solution step could not be used.

The key ingredient of the successful structure analysis was the use of the precession technique (*31*) during electron diffraction tomography (PEDT) data acquisition. Because of the application of this technique, (i) the number of reflections on each frame is substantially larger than without precession, providing more data for the determination of frame orientation; (ii) most reflections are recorded on more than one frame, allowing a more robust frame scaling and intensity extraction; and (iii) the intensities are integrated and therefore more suitable for subsequent structure refinement (*6*, *32*).

Each dataset was processed separately. The orientation of each frame was determined independently by least-squares fitting of the simulated kinematical intensities and their positions based on the unit cell geometry to the experimental frame in PETS software (*33*). As a result, it was possible to determine the excitation error of each recorded intensity maximum and estimate an average rocking-curve profile. Excitation error is the distance of the recorded intensity from the Bragg condition in reciprocal space and rocking-curve profile describes the reflection intensity as a function of the excitation error. The rocking curve for a PEDT experiment is a two-peaked curve. The crystal-related variables influencing the rocking-curve profile are reflection full width at half maximum (FWHM) and mosaicity (see the supplementary materials). The experimental rocking curves of measured reflections were averaged in the reciprocal space shells with 0.1-Å$^{-1}$ thicknesses. These averaged profiles were then fitted by calculated rocking-curve profiles using the reflection FWHM and mosaicity as adjustable parameters. An approximation of isotropic crystal mosaicity was used in this procedure. The

[1]Institute of Physics of the Czech Academy of Sciences, Na Slovance 2, 18200 Prague 8, Czech Republic. [2]University of Chemistry and Technology, Technická 3, 16628 Prague 6, Czech Republic. [3]Zentiva, U Kabelovny 130, 10237 Prague 10, Czech Republic.
*Corresponding author. Email: brazda@fzu.cz

experimental rocking-curve plot with the fitted profiles is shown in Fig. 2B and fig. S8. The specific and well-defined rocking curve of the PEDT data allows efficient intensity estimation by fitting the rocking-curve profile to the experimental intensities. This method allows a meaningful estimation of the reflection intensity even if the reflection is present on a single frame (Fig. 2C). During the intensity extraction, we also optimized the scales of individual frames (fig. S9).

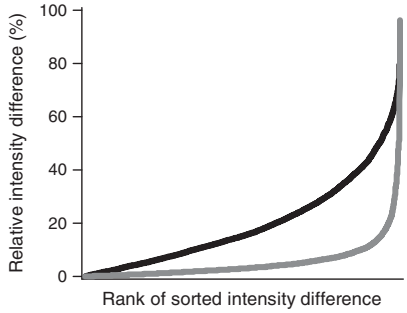
Combination of datasets from several crystals was necessary for structure solution because of low completeness of the individual datasets. Four datasets with the best completeness of frames with resolution $d_{min}$ better than 1.3 Å were selected with 128 diffraction frames. The resulting completeness reached 84%. The structure was solved ab initio using direct methods as implemented in SIR2014 (*34*). All atoms were found in the solution (space group $P2_12_12_1$, 44 independent nonhydrogen atoms, unit cell volume 3127 Å$^3$).

The structure solution was followed by kinematical refinement (see supplementary materials). To improve the structure model, we used refinement taking into account the dynamical diffraction effects (dynamical refinement) (*32*). Only 86 frames obtained from the first or second exposure of the particular crystal spot were used. The structure model (nonhydrogen atoms without restraints and H-atoms restrained to geometrical positions) was refined against data from four crystals separated into four independent intensity files, with combined refinement (*R*) values *R*(obs) and weighted *R*(all) [*wR*(all)] of 11.78 and 11.82%, respectively. After the application of the frame orientation refinement, which is a part of the standard dynamical refinement procedure (*32*) and which was applied in addition to the geometrical orientation refinement described above, the average absolute tilt correction was 0.11° and the combined *R*-factors decreased to 9.86 and 9.62%, respectively.

In addition to improving the structure model and fitting indicators, dynamical refinement is the only approach that allows the determination of absolute structure, because it takes into account the inversion symmetry violation due to multiple scattering. To determine the absolute structure, we refined an inverted structure model of the cocrystal of sofosbuvir and L-proline in Jana2006 (*35*) using the same procedure as the correct model including frame orientation refinement. The resulting combined *R*-factors increased by ~2.6 percentage points [*R*(obs) 12.34%, *wR*(all) 12.43%]. All four crystals (26, 27, 14, and 19 frames) showed a similar increase of the *R*(obs) for the wrong absolute structure model (2.9, 2.1, 2.2, and 2.5 percentage points, respectively). The large difference between *R*-factors shows unambiguously that the expected structure model is the correct one. The average *R*-factor increase due to the wrong absolute structure model for frames obtained during the second exposure of the crystals spot reached only ~75% of the average increase for the first measurement frames [2.8 percentage points first measurement, 2.0 percentage points second measurement for *R*(obs)]. The differences [*R*(obs)inverted – *R*(obs)] for each frame are shown in Fig. 3. There are 10 frames (out of 86) in which the difference is negative, thus the *R*(obs) is lower for the wrong absolute structure. Eight out of these 10 frames were obtained during the second measurement of a crystal spot. The average *R*(obs) increase and the number of negative *R*(obs) differences illustrate how beam damage and the resulting lower resolution complicate determination of correct absolute structure. Successful determination of the absolute structure of highly beam-sensitive light-atom structures from nonideal crystals thus cannot rely on a single frame, but rather requires a combination of a larger number of frames.

The complete structure analysis procedure described here was designed to be directly transferable to serial electron crystallography data. First, each frame is treated strictly individually and the procedure does not make any use of the knowledge that the frames were collected in a consecutive manner. The exception to this rule is the definition of the orientation matrix (see the supplementary materials). Second, the frame-scaling procedure does not make use of the correlation between consecutive frames and thus can be used directly for data from different crystals. Third, the intensity extraction by fitting of the rocking curve is especially suitable for a large number of mutually randomly oriented diffraction patterns and can provide accurate intensity

**Fig. 1. Comparison of intensity differences $|(I_{hkl} - I^{inv}_{hkl})|/(I_{hkl} + I^{inv}_{hkl})$ for simulated electron (black line) and x-ray Cu Kα (gray line) diffraction data of a cocrystal of sofosbuvir and L-proline.** The simulated data for electrons are equal to $I_{calc}$ values from dynamical refinement. The x-ray data were simulated with the "simulation of single crystal data" tool in Jana2006 using the same structure model.



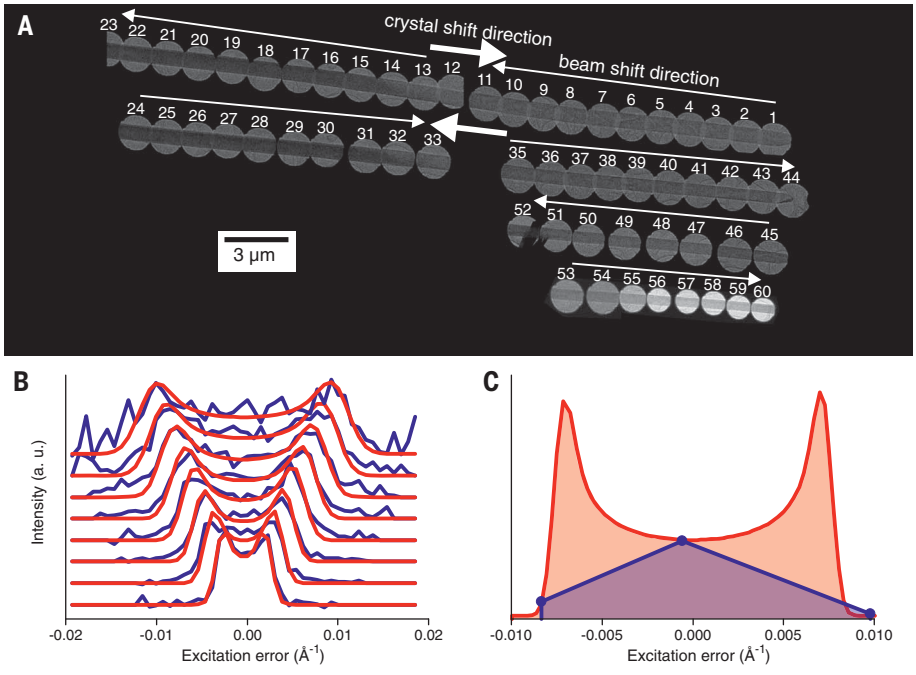




**Fig. 2. Diffraction data collection and advanced diffracted intensity extraction.** (**A**) Scanning of crystal during data collection. Beam and crystal shifts are indicated by white arrows. (**B**) Plot of the rocking-curve profiles of the experimental precession electron diffraction data collected on one of the four crystals used for determination of rocking-curve profile parameters. The lowest blue curve is the averaged rocking curve in the range of 0.2 to 0.3 Å$^{-1}$ and the highest blue curve is the averaged rocking curve in the range of 0.9 to 1.0 Å$^{-1}$. The precession angle is 0.65°. The red curves correspond to the fitted rocking-curve profiles with the FWHM of the interference function equal to 0.0005 Å$^{-1}$ and an apparent mosaicity of 0.08°. (**C**) Comparison of intensity integration in case of sparse sampling of reciprocal space. Experimental points (blue) are fitted with rocking-curve profile (red line) and the resulting intensity corresponds to the red area. Blue area corresponds to the area under experimental points.

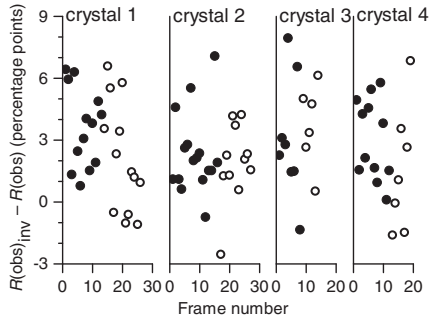



**Fig. 3. *R*(obs) increase due to wrong absolute structure for individual frames shown for each crystal separately.** Filled and empty circles correspond to the frames obtained during the first and second measurement of a crystal spot, respectively.

information even from a single measurement of a reflection.

## REFERENCES AND NOTES


1. For the definition of absolute structure and its distinction from absolute configuration, see *Online Dictionary of Crystallography* (http://reference.iucr.org/dictionary/Absolute_structure).
2. H. D. Flack, G. Bernardinelli, *J. Appl. Cryst.* **33**, 1143–1148 (2000).
3. J. Jansen, H. W. Zandbergen, *Ultramicroscopy* **90**, 291–300 (2002).
4. M. Tanaka, H. Takayoshi, M. Ishida, Y. Endoh, *J. Phys. Soc. Jpn.* **54**, 2975–2982 (1985).
5. A. W. S. Johnson, A. R. Preston, *Ultramicroscopy* **55**, 348–355 (1994).
6. L. Palatinus *et al.*, *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater.* **71**, 740–751 (2015).
7. Y. Ma, P. Oleynikov, O. Terasaki, *Nat. Mater.* **16**, 755–759 (2017).
8. J. Hadermann *et al.*, *Chem. Mater.* **24**, 3401–3405 (2012).
9. S. Smeets, X. Zou, W. Wan, *J. Appl. Crystallogr.* **51**, 1262–1273 (2018).
10. J. Hynek, P. Brázda, J. Rohlíček, M. G. S. Londesborough, J. Demel, *Angew. Chem. Int. Ed. Engl.* **57**, 5016–5019 (2018).
11. E. van Genderen *et al.*, *Acta Crystallogr. A Found Adv.* **72**, 236–242 (2016).
12. L. Palatinus *et al.*, *Science* **355**, 166–169 (2017).
13. P. P. Das *et al.*, *Org. Process Res. Dev.* **22**, 1365–1372 (2018).
14. T. Gruene *et al.*, *Angew. Chem. Int. Ed. Engl.* **57**, 16313–16317 (2018).
15. C. G. Jones *et al.*, *ACS Cent. Sci.* **4**, 1587–1592 (2018).
16. U. Kolb, T. Gorelik, C. Kübel, M. T. Otten, D. Hubert, *Ultramicroscopy* **107**, 507–513 (2007).
17. D. Zhang, P. Oleynikov, S. Hovmöller, X. Zou, *Z. Kristallogr.* **225**, 94–102 (2010).
18. E. Mugnaioli, T. Gorelik, U. Kolb, *Ultramicroscopy* **109**, 758–765 (2009).
19. I. Nederlof, E. van Genderen, Y.-W. Li, J. P. Abrahams, *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1223–1230 (2013).
20. B. L. Nannenga, D. Shi, A. G. W. Leslie, T. Gonen, *Nat. Methods* **11**, 927–930 (2014).
21. M. Gemmi, M. G. I. La Placa, A. S. Galanis, E. F. Rauch, E. F. S. Nicolopoulos, *J. Appl. Cryst.* **48**, 718–727 (2015).
22. A. J. Hutt, J. O'Grady, *J. Antimicrob. Chemother.* **37**, 7–32 (1996).
23. R. Henderson, *Sci.* **241**, 6–8 (1990).
24. U. Kolb, T. Gorelik, E. Mugnaioli, E. Stewart, *Polym. Rev. (Phila. Pa.)* **50**, 385–409 (2010).
25. J. Hattne *et al.*, *Structure* **26**, 759–766.e4 (2018).
26. J. C. H. Spence, P. W. Hawkes, *Ultramicroscopy* **108**, 1502–1503 (2008).
27. K. Hirata *et al.*, *Nat. Methods* **11**, 734–736 (2014).
28. J. C. H. Spence, *IUCrJ* **4**, 322–339 (2017).
29. D. J. Berry, J. W. Steed, *Adv. Drug Deliv. Rev.* **117**, 3–24 (2017).
30. W. Albrecht, J. Geier, D. Perez Palacios, International patent WO2017029408A1 (2017).
31. R. Vincent, P. A. Midgley, *Ultramicroscopy* **53**, 271–282 (1994).
32. L. Palatinus, V. Petříček, C. A. Corrêa, *Acta Crystallogr. A Found Adv.* **71**, 235–244 (2015).
33. L. Palatinus, *PETS: Program for Analysis of Electron Diffraction Data* (Institute of Physics of the Czech Academy of Sciences, 2011).
34. M. C. Burla *et al.*, *J. Appl. Cryst.* **48**, 306–309 (2015).
35. V. Petříček, M. Dušek, L. Palatinus, *Z. Kristallogr.* **229**, 345–352 (2014).


## ACKNOWLEDGMENTS

We thank Zentiva for providing sofosbuvir. **Funding:** This research was supported by the Czech Science Foundation (project no. 16-10035S). The experiments were conducted in part using instruments of the ASTRA laboratory established within the operational program Prague Competitiveness (project no. CZ.2.16/3.1.00/24510. **Author contributions:** M.B and P.B. prepared crystals of sofosbuvir L-proline. P.B. collected PEDT data. L.P. and P.B. developed algorithms for raw data processing. L.P. coded the software. P.B. processed the raw data. P.B and L.P. determined the structure of the cocrystal. P. B. determined the absolute configuration. P.B. and L.P. wrote the manuscript. All authors contributed to revision of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Crystallographic data were deposited in the Crystallography Open Database (COD) under the entry 3000217. All other data are available in the text or in the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6441/667/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S4
References (*36–41*)

5 December 2018; accepted 3 April 2019
10.1126/science.aaw2560

Science

# Electron diffraction determines molecular absolute configuration in a pharmaceutical nanocrystal

Petr Brázda, Lukás Palatinus and Martin Babor

*Science* **364** (6441), 667-669.
DOI: 10.1126/science.aaw2560

**Dynamical refinement spots a difference**

For chiral molecules used in drugs, one isomer can have beneficial bioactivity, whereas the others are useless or even harmful. Determining the absolute configuration of molecules with chiral centers is often achieved through x-ray crystallography, but this requires relatively large crystals and high-quality data. Brázda *et al.* used electron diffraction to determine the absolute structure of an extremely radiation-sensitive crystal with micrometer dimensions (see the Perspective by Xu and Zou). In a strategy analogous to serial crystallography methods, many frames were combined to generate a complete dataset. Refinement incorporating dynamical effects differentiated the correct and incorrect molecular configuration.

*Science*, this issue p. 667; see also p. 632



*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 