High-resolution structure determination by continuous-rotation data collection in MicroED

Brent L Nannenga[1,3], Dan Shi[1,3], Andrew G W Leslie[2] & Tamir Gonen[1]

MicroED uses very small three-dimensional protein crystals and electron diffraction for structure determination. We present an improved data collection protocol for MicroED called 'continuous rotation'. Microcrystals are continuously rotated during data collection, yielding more accurate data. The method enables data processing with the crystallographic software tool MOSFLM, which resulted in improved resolution for the model protein lysozyme. These improvements are paving the way for the broad implementation and application of MicroED in structural biology.

Producing large, well-ordered crystals is a major bottleneck for protein structure determination by X-ray crystallography. Because large crystals are needed to withstand the negative effects of radiation damage during data collection, small micro- and nanocrystals have generally been cast aside as unusable. Many difficult-to-crystallize targets never reach a usable size and are generally discarded; therefore, methods that can facilitate structure determination from these small crystals would be exceedingly valuable. Advances with X-ray free-electron lasers (XFELs) have made it possible to use microcrystals to solve protein structures[1–3]. However, the current implementations require collection of XFEL diffraction patterns from thousands, or even millions, of crystals, and the instrumentation availability is limited.

Recently, we reported the development of micro-electron diffraction (MicroED)[4] as a complementary method to XFEL. In MicroED, electron diffraction data are collected from extremely small, three-dimensional (3D) protein microcrystals for protein structure determination. In our previous work[4], diffraction data of lysozyme microcrystals were taken as a tilt series of still exposures, and these data were integrated and merged using in-house–developed programs[5]. After phasing and refinement, the final structure of lysozyme was solved to 2.9 Å.

Here we report several substantial developments to MicroED. We developed an improved, continuous-rotation data collection protocol: here the electron diffraction data are collected as a movie as the crystal is continuously rotated by the microscope stage (analogously to in X-ray crystallography, in which the crystal is oscillated using the goniometer). The collected data could therefore be processed by MOSFLM[6,7], a well-established data processing program for X-ray crystallography, which further improved the process of structure determination by MicroED. These enhanced data collection and processing strategies yielded more accurate diffraction data and allowed the structure of lysozyme to be solved at 2.5-Å resolution with significantly improved refinement statistics over those in our original report.

In the original MicroED protocol[4], diffraction data were collected as a series of still exposures, each related by the tilt of the stage (0.1–1°) between exposures during data collection (**Fig. 1a**,**b**). Still exposures mainly produce partial intensities, and although we showed that a data set composed solely of still exposures is able to produce a structure, some inaccuracies resulting from the partiality will be carried through the subsequent data processing steps unless scaling is performed.

To overcome this problem and improve the quality of the MicroED data, we sought to collect diffraction data as the crystals are continuously rotated, which is more similar to how X-ray diffraction data are collected[8]. Our hypothesis was that this improved data collection method would yield more accurately measured intensities, as reciprocal space is more finely sampled (**Fig. 1c**,**d**). Such sampling would allow the use of standard X-ray data processing programs for integrating and scaling, further improving data quality and subsequent processing.

To collect diffraction data by the continuous-rotation method, lysozyme microcrystals (**Supplementary Fig. 1**) were grown, applied to an electron microscopy grid and vitrified in liquid ethane as described previously[4]. The microcrystals were wedge shaped, typically with a length and width of ~2 µm and wedge thickness of ~0.1–0.6 µm within each individual crystal. We measured the crystal thickness in real space by analyzing the same crystal at different tilts in imaging mode. The grids were then loaded, assessed and diffracted at cryogenic temperatures using a 200-kV transmission electron microscope equipped with a field emission gun. Once a crystal was selected, the CompuStage of the microscope was rotated at a constant rate of ~0.09° s^{-1}. As the crystal rotated, it was simultaneously exposed to the electron beam, and the diffraction was recorded as a movie on a complementary metal-oxide semiconductor (CMOS)-based detector in rolling-shutter mode with each frame covering an ~0.36° wedge (frame rate of 4 s per frame). Typically, a total of ~44° of data were collected per crystal before the total accumulated dose limit of 5 $e^{-}/Å^2$ was reached (see **Supplementary Results**, **Supplementary Video 1** and **Supplementary Fig. 2** for a description

[1]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, USA. [2]Medical Research Council Laboratory of Molecular Biology, Cambridge, UK. [3]These authors contributed equally to this work. Correspondence should be addressed to T.G. (gonent@janelia.hhmi.org).

RECEIVED 31 MARCH; ACCEPTED 16 JUNE; PUBLISHED ONLINE 3 AUGUST 2014; DOI:10.1038/NMETH.3043

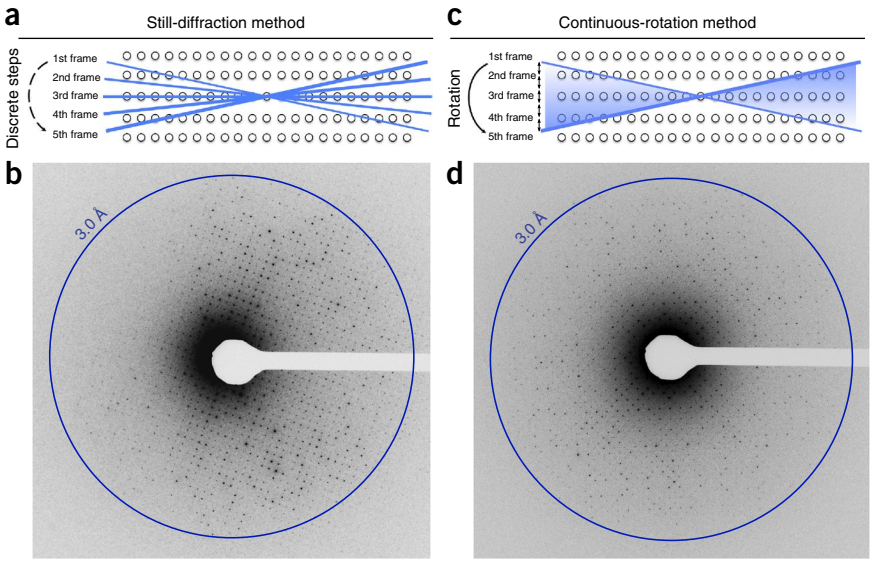

Figure 1 | Data collection strategies in MicroED. (**a**,**b**) Still diffraction[4] consists of rotating the stage in discrete steps between exposures. This provides data in the form of 2D slices through the 3D reciprocal space (**a**), and although this is sufficient for structure determination, the data are inherently incomplete because most reflections are only partially recorded (**b**). (**c**,**d**) The continuous-rotation method for MicroED samples the reciprocal space continuously as the crystal is rotated (**c**), which yields much more complete and accurate measurements of reflection intensities (**d**). In the examples shown here originating from two different crystals, the reflections on the left side of the still diffraction are missing owing to crystal orientation, whereas they are present in the continuous-rotation data because reciprocal space is being more completely sampled (**b** vs. **d**, respectively).

of radiation damage assessment). With this procedure, we were able to collect continuously rotating electron diffraction data sets with visible reflections beyond 2 Å (**Supplementary Fig. 1** and **Supplementary Video 2**) that could be merged and further processed.

One of the major factors enabling the widespread use of X-ray crystallography is the powerful and relatively user-friendly software that has been developed over many decades for efficient data processing. We sought to capitalize on the work put into such programs by processing the MicroED data with MOSFLM[6,7], a widely used program for X-ray data integration and processing. Processing the MicroED continuous-rotation data with MOSFLM was possible without making any changes to the software (**Supplementary Fig. 3**), but some modifications to the standard procedure and processing parameters were required (**Supplementary Results**).

Data from two crystals were collected and processed, with overall completeness to 2.5 Å being 80% and 45% for crystals 1 and 2, respectively. It is important to note that the total angular range collected for both crystals was the same (44°), and the higher completeness for crystal 1 is the result of crystal orientation on the grid. Ultimately, data from either one or two crystals were processed with MOSFLM and merged to provide a complete data set. The intensity data were merged, scaled and converted to structure-factor amplitudes using Pointless[9], Aimless[10] and CTruncate within the CCP4 suite[11]. The data sets were truncated at 2.5 Å

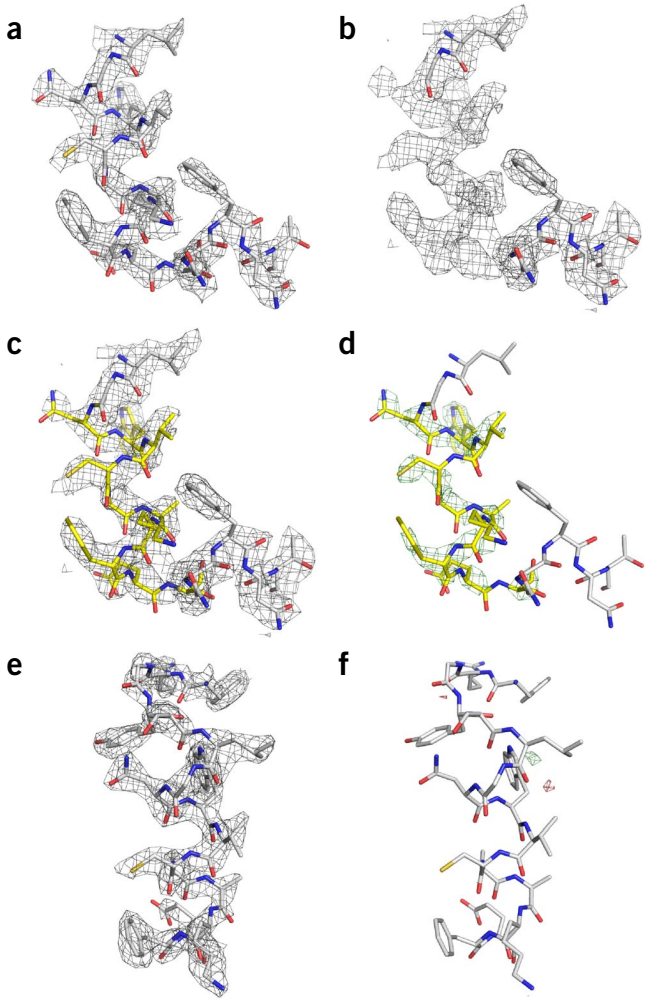

Figure 2 | Final refined structure of lysozyme at 2.5 Å from continuous-rotation MicroED data. (**a**–**d**) A representative region of the final refined structure of lysozyme originating from the two-crystal data set is shown, with the $2F_o - F_c$ density map (**a**; contoured at 1.0σ) showing well-defined density around the backbone and side chains. The final 3D structure is also shown in **Supplementary Video 3**. To test any potential model bias, we removed residues 27–36 from the final refined model, and the incomplete model was used to phase and refine the original data. The $2F_o - F_c$ map (contoured at 1.0σ) without the deleted residues (**b**) shows clearly defined density for both the backbone and side chains where the missing residues (yellow) could easily be placed (**c**). The $F_o - F_c$ map (contoured at 3σ) also shows very strong density for the deleted residues (**d**). The strong density for the missing residues in the $2F_o - F_c$ and $F_o - F_c$ maps indicates that the final map does not suffer from model bias. (**e**,**f**) Final refined structure of lysozyme at 2.5-Å resolution using data originating from a single crystal. The $2F_o - F_c$ density map (**e**; contoured at 1.0σ) around residues 20–35 shows well-defined density around both the backbone and side chains. The $F_o - F_c$ map (**f**; contoured at ±3σ) shows no clear differences between the observed data and the model.

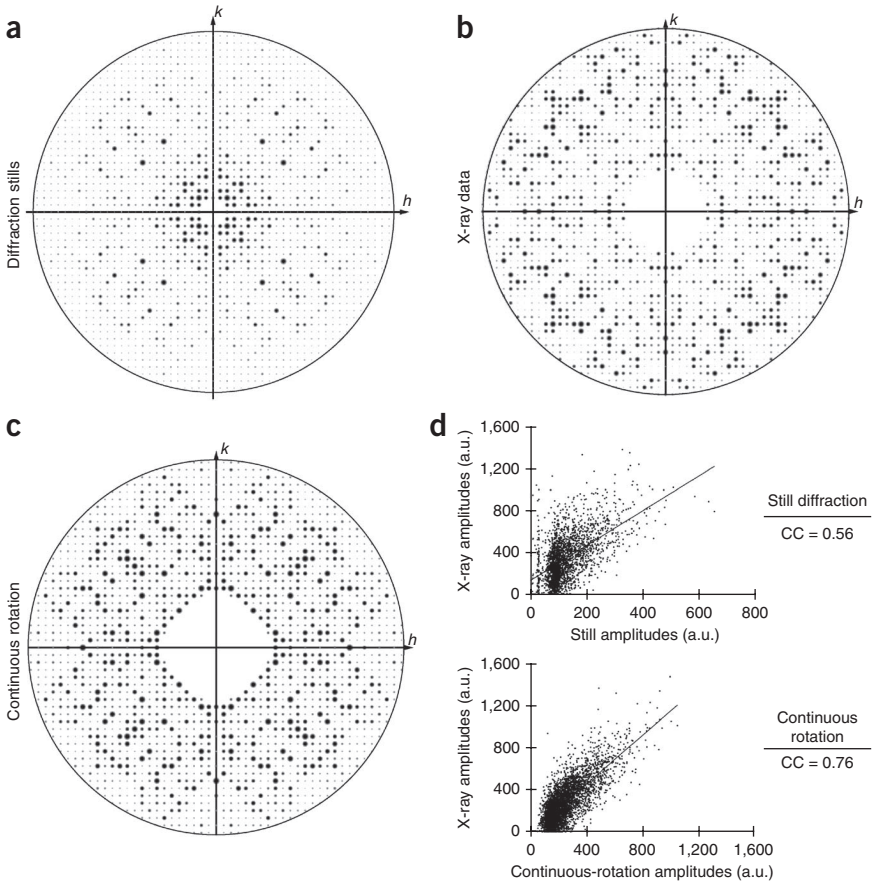

Figure 3 | Continuous rotation improves MicroED data quality. (**a**–**c**) Views of the (001) plane show that the intensities from still-diffraction data (**a**) exhibit less variation between high- and low-intensity reflections when compared to X-ray data (**b**) and the data collected from continuous rotation (**c**). Data are displayed using VIEWHKL within the CCP4 suite[11]. (**d**) A lysozyme data set collected by X-ray crystallography was compared with the previously reported MicroED data collected as diffraction stills[4] (top) and with the continuous-rotation data set (bottom). Pearson correlation coefficients (CC) are calculated for both comparisons. a.u., arbitrary units.

on the basis of the merging statistics presented in **Supplementary Table 1**.

Molecular replacement was performed using Phaser[12] in order to determine phases with lysozyme (Protein Data Bank (PDB): 1IEE)[13] as a search model. We applied refinement in PHENIX[14] using electron scattering factors to obtain models with good statistics and geometry for both data sets (**Supplementary Table 1**). The final refined structures showed excellent agreement with the density map (**Fig. 2** and **Supplementary Video 3**), with strong, unambiguous density for the peptide backbone, as well as for the side chains and some well-ordered water molecules.

In order to check for model bias in the final structure, residues 27–36 of the final model were removed, and the remaining model was used for molecular replacement with the original data. After refinement, the resulting map showed strong, interpretable density in the region where the model had been removed (**Fig. 2b**), and the correct residues could be readily fit into this density (**Fig. 2c**,**d**), indicating low levels of bias from the search model.

Overall, the quality of the MicroED data obtained by continuous rotation is better than that of the original still-diffraction data that we reported previously[4] (**Fig. 3**). When the amplitudes from still diffraction are compared with those from a lysozyme data set obtained by X-ray crystallography, the data are moderately correlated with a Pearson correlation coefficient of 0.56 for all data to 2.9 Å (n = 2,466) and 0.63 to 6.0 Å (n = 307); however, there is some spreading of the data (**Fig. 3d**). When the same comparison is made between the scaled amplitudes from continuous rotation with the X-ray data set, the correlation coefficient between the continuous-rotation data and the X-ray data is 0.76 for all data to 2.5 Å (n = 3,949) and 0.84 to 6.0 Å (n = 273; **Fig. 3d**), which represents a substantial improvement over the still-diffraction data set.

With continuous-rotation data, the effects of dynamic scattering should be diminished, as was reported for electron diffraction using precession diffraction of thicker materials[15,16]. Dynamic scattering (multiple elastic scattering events) in electron diffraction can redistribute primary reflection intensities as the primary scattered electrons scatter again elastically within the crystal, which can lead to a reduction in the accuracy of the measured intensities[17]. We previously noted that dynamic scattering in still-diffraction data of lysozyme contributed an average error of ~5% (ref. 4). This observation was based on intensity measurements for forbidden reflections in the still data set. The lysozyme crystals have $P4_32_12$ symmetry, and systematic absences are expected along the reciprocal space a^* and b^* axes at positions $(2n + 1, 0, 0)$. However, in the still data set, weak reflections were observed at the positions where absences were expected (**Supplementary Fig. 4**). We hypothesize that these reflections originate from dynamic scattering events. When we performed the same analysis as reported previously[4], it was clear that these weak forbidden reflections along the a^* and b^* axes were reduced and contributed an average error of only 2.5% (s.e.m. = 0.9%, n = 16) in the continuous-rotation data.

The continuous-rotation method described here is similar in principle to precession electron diffraction[18]. Improvements in data accuracy with precession electron diffraction have been reported in materials science and are due to the reduction in dynamic scattering effects and better intensity values as the reflections are integrated through the Bragg angle[15,16]. The effects of dynamic scattering are diminished in precession electron diffraction because at any given point along the beam's precession path, the total number of allowed secondary scattering paths is reduced[16]. This decreases the intensity redistribution due to dynamic scattering as the reflections are integrated over the complete circular path of the precession beam. Because continuous rotation is a simplified version of precession diffraction, it is not surprising that we improved the MicroED data in comparison with data collected with diffractions stills[4].

Future developments could further improve the performance and applicability of MicroED in structural biology. One possibility

is the use of detectors with improved movie mode capabilities. Although the rolling-shutter mode on the CMOS camera we use performs well, the data still suffer because of the long readout time of the sensor. The rotation of the crystal during the readout time results in an angular gradient across an image that hampers data integration at the highest-resolution shell. At the same time, signal-to-noise ratio is also an issue in rolling-shutter mode. Improvements in detector speed or inclusion of new algorithms to better handle this error in MOSFLM should help to further improve the final data. In this study, MOSFLM was used for data integration; however, there is no reason other data processing suites would not also work with MicroED data. Other processing programs could potentially improve processing if the sources of error encountered are dealt with in different ways. Also, although we have attempted to reduce the effects of multiple elastic scattering by continuously rotating the sample, we have not yet addressed the errors associated with inelastic scattering. Inelastic scattering leads to high levels of background noise in thicker samples, and the use of an energy filter would greatly reduce the noise in our diffraction patterns[19]. By collecting data on a microscope equipped with an energy filter, the signal-to-noise ratio of the MicroED data should be substantially improved. Additionally, there are other questions that need to be addressed in future work, such as collecting data from crystals with larger unit cells and determining the range of microcrystal sizes from which useful diffraction data can be collected.

The continuous-rotation method coupled with data processing using standard X-ray crystallographic software makes MicroED an accessible method to structural biologists already familiar with the suite of programs available for X-ray diffraction data. The only difference is the way in which the diffraction data are obtained: with electrons as opposed to X-rays. With such improvements and streamlining, we believe that MicroED is quickly becoming a feasible method with wide applicability for solving structures of biological materials from extremely small crystals using an electron microscope.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Protein Data Bank: Final structure factors and coordinates for the structure solved from the two-crystal data set have been deposited with accession code 3J6K.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank J. Hattne, F.E. Reyes, D. Olbris, H. Tietz and M. Stumpf for helpful discussions. Work in the Gonen lab is supported by the Howard Hughes Medical Institute. A.G.W.L. is supported by the Medical Research Council (U105184325) and Biotechnology and Biological Sciences Research Council (BBSRC; BB/F020384/1) and Collaborative Computational Project Number 4 (CCP4).

AUTHOR CONTRIBUTIONS

B.L.N. and D.S. contributed to project design, conception, data collection, data analysis, manuscript writing and figure making. A.G.W.L. contributed to data processing and analysis in MOSFLM and manuscript writing. T.G. contributed to project design, conception, data analysis and manuscript writing.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Barends, T.R. *et al. Nature* **505**, 244–247 (2014).
2. Chapman, H.N. *et al. Nature* **470**, 73–77 (2011).
3. Boutet, S. *et al. Science* **337**, 362–364 (2012).
4. Shi, D., Nannenga, B.L., Iadanza, M.G. & Gonen, T. *eLife* **2**, e01345 (2013).
5. Iadanza, M.G. & Gonen, T. *J. Appl. Crystallogr.* **47**, 1140–1145 (2014).
6. Leslie, A.G.W. & Powell, H.R. in *Evolving Methods for Macromolecular Crystallography* (eds. Read, R.J. & Sussman, J.L.) 41–51 (Springer, 2007).
7. Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R. & Leslie, A.G. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).
8. Dauter, Z. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 1703–1717 (1999).
9. Evans, P.R. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011).
10. Evans, P.R. & Murshudov, G.N. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
11. Winn, M.D. *et al. Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
12. McCoy, A.J. *et al. J. Appl. Crystallogr.* **40**, 658–674 (2007).
13. Sauter, C. *et al. Acta Crystallogr. D Biol. Crystallogr.* **57**, 1119–1126 (2001).
14. Adams, P.D. *et al. Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
15. Gjønnes, J. *et al. Acta Crystallogr. A* **54**, 306–319 (1998).
16. Oleynikov, P., Hovmöller, S. & Zou, X.D. *Ultramicroscopy* **107**, 523–533 (2007).
17. Grigorieff, N., Ceska, T.A., Downing, K.H., Baldwin, J.M. & Henderson, R. *J. Mol. Biol.* **259**, 393–421 (1996).
18. Vincent, R. & Midgley, P.A. *Ultramicroscopy* **53**, 271–282 (1994).
19. Yonekura, K., Maki-Yonekura, S. & Namba, K. *Biophys. J.* **82**, 2784–2797 (2002).

ONLINE METHODS

Collection of rotation electron diffraction data. Lysozyme microcrystals and electron microscopy (EM) samples for diffraction were prepared as described previously[4]. All electron diffraction was performed with intensity less than 0.01 $e^-/Å^2$ on a TEM (FEI) operated at 200 kV, equipped with an FEI field emission gun (FEG), and data were collected with a 4k × 4k TVIPS TemCam-F416 CMOS camera in rolling-shutter mode (15.6-µm pixel size). For continuously rotating diffraction data, the stage of the microscope was set to rotate at 0.09° s^{-1} using the microscope's standard hardware and software. Crystals were located on the grid by searching in overfocused diffraction mode. When a crystal was found, an initial diffraction pattern was recorded to judge the quality of that particular crystal. If the crystal showed strong and sharp diffraction[20], the beam was blanked and the rotation of the stage was started. Once the stage began its rotation and had achieved a constant rate, the beam was unblanked and diffraction data were recorded at a constant frame rate of 1 frame per 4 s (0.36° per frame) using the camera's rolling-shutter mode. Data sets of approximately 44° were collected for each crystal (total dose of $<5e^-/Å^2$; see **Supplementary Results** for dose rate). For the radiation damage assessment, the stage was set at 0° and oscillated back and forth between −1° and 1° during the course of the continuous dosage experiment.

Data processing and structure refinement. Raw MicroED data were converted into an SMV file format, which could be read by MOSFLM. Diffraction data were indexed, integrated, merged, scaled and prepared for refinement using MOSFLM v.7.1.0 (ref. 6) and the graphical interface iMOSFLM v.1.0.7 (ref. 7), Pointless[9] and Aimless[10]. In the scaling and merging step, the refinement of the s.d. correction factors in Aimless was unstable, and these were set to SDFAC = 1.5, SDB = 0.0, SDADD = 0.03. Phases were obtained by molecular replacement using Phaser[12] with lysozyme (PDB: 1IEE; ref. 13) as a search model (LLG = 673 and TFZ = 24.3 for the multiple-crystal data set; LLG = 821 and TFZ = 8.5 for the single-crystal data set), and the molecular replacement solutions were refined using electron scattering factors in PHENIX[14] using a 5% free data set.

20. Gonen, T. *Methods Mol. Biol.* **955**, 153–169 (2013).