SI Appendix

Electron crystallography of ultrathin 3D protein crystals: Atomic model with charges

Koji Yonekura, Kazuyuki Kato, Mitsuo Ogasawara, Masahiro Tomita, and Chikashi Toyoshima

SI Discussion

In the main paper, we show that electron crystallography can be a tool for building atomic-models from ultrathin 3D protein crystals and discerning charged states of amino acid residues and metals at functional sites of proteins. In the discussion following, we examine the potential and the limits of the method developed here and compare it with others.

Crystal thickness. The mean free path of inelastically scattered electrons by material consisting of light atoms is measured as ~ 4,000 Å for 300-keV electrons (1) and that of elastically scattered electrons as much longer (2). Energy filtration can effectively remove multiply scattered electrons, as they are likely to be inelastically scattered at least once. Unlike perfect crystals of strong scatterers, protein crystals are highly mosaic and scatter electrons only weakly, yielding broad and weak diffraction spots. If multiple scattering is significant, systematic absence of diffraction spots due to crystal symmetry is unlikely. In fact, the systematic absence of such spots is a prominent feature of the electron diffraction patterns of catalase (Fig. 2*A*; Movie S1) and Ca^{2+}-ATPase (Fig. S4*A* and *B*). Hence, inelastic and dynamic scatterings are unlikely to cause a serious problem in analyzing ultrathin protein 3D crystals (3, 4, 5). 3D crystals consisting of just a few layers are not rare (*e.g.* 6, 7). It is also very likely that such thin crystals have been abandoned or left unnoticed, as they are too

thin for examination by optical microscopy. Thus, there may already be many examples of proteins that yield crystals suitable for electron crystallography.

Phasing. We have so far solved structures only by molecular replacement based on known structures. In X-ray crystallography, heavy atom derivatives are used very effectively for acquiring phase information by isomorphous replacement and anomalous scattering. In electron crystallography, it has been thought that isomorphous replacement would not work, as, unlike with X-rays, the scattering factor for electrons does not increase proportionally to atomic number (8). Nevertheless, according to a later report, phasing by anomalous scattering and/or isomorphous replacement may still be feasible (5). We are now pursuing these possibilities by making heavy atom derivatives of Ca^{2+}-ATPase (9).

Data statistics. Since radiation damage prevents collection of a full diffraction dataset from a single crystal, partial datasets from several crystals need to be merged. To minimize deviations from isomorphism, we applied post-refinement of cell parameters and chose only datasets showing cell parameters close to those determined by X-ray crystallography [Ca^{2+}-ATPase (9)] or powder diffraction [catalase (10)]. Thus, $CC_{1/2}$ and CC^* values, now used as objective measures of crystallographic data quality (11), produce good statistics (Fig. S5*A*). These figures also indicate that variation in crystal thickness is not a serious problem, even when datasets from different crystals composed of a few to ten layers are merged. However, refinement statistics (R_{work} / R_{free} in Table S1) appear rather poor compared to those common in X-ray crystallography, presumably due to inaccurate atomic scattering factors. The missing cone might adversely affect the *R*-values, but poorer *R*-values have often been observed in structure analysis of 2D crystals including data at higher tilt angles (12, 13). Accurate estimations of atomic scattering factors and partial charges for all atoms contained in protein crystals including solvent molecules are a future challenge.

Comparison with other methods. A precession technique of tilting and rotating the electron beam with specially designed deflection coils has been successfully applied to structure determination with 3D nano crystals of inorganic compounds such as hemimorphite and mayenite (14). They are almost perfect crystals that contain many unit cells consisting of atoms with relatively large scattering cross-sections. In contrast, protein crystals are mosaics of small crystalline blocks consisting of only weakly scattering atoms. With such specimens, tilting the beam away from the optical axis of the electron microscope tends not to yield good diffraction patterns. Also a relatively high electron dose needs to be applied during the precession, making beam-sensitive specimens unsuitable.

Serial femto second X-ray crystallography with X-ray free electron laser has recently become available for structure determination with sub μm to a few μm protein crystals (*e.g.* 15). In this technique, a huge number of randomly-oriented diffraction patterns are collected by exposing a high-intensity X-ray beam to a liquid jet of crystal solution. Electron crystallography may be applied to even smaller and thinner protein crystals (16), whose structures cannot be solved by other methods.

1. Yonekura K, Braunfeld MB, Maki-Yonekura S, Agard DA (2006) Electron energy filtering significantly improves amplitude contrast of frozen-hydrated protein at 300 kV. *J Struct Biol* 156(3):524–536.
2. Angert I, Burmester C, Dinges C, Rose H, Schröder RR (1996) Elastic and inelastic scattering cross-sections of amorphous layers of carbon and vitrified ice. *Ultramicroscopy* 63(3-4):181–192.
3. Toyoshima C, Ogawa H, Tani K (1998) Energy-filtering electron crystallography of proteins. *Hitachi Instrument News, the 33rd Electron Microscopy Edition* (Hitachi Ltd., Tokyo), pp 3–7.
4. Yonekura K, Maki-Yonekura S, Namba K (2002) Quantitative comparison of zero-loss and conventional electron diffraction from two-dimensional and thin three-dimensional protein crystals. *Biophys J* 82(5):2784–2797.
5. Burmester C, Schröder RR (1997) Solving the phase problem in protein electron crystallography:

Multiple isomorphous replacement and anomalous dispersion as alternatives to imaging. *Scanning Microsc* 11:323–334.

6. Jeng TW, Chiu W (1983) Low dose electron microscopy of the crotoxin complex thin crystal. *J Mol Biol* 164(2):329–346.
7. Stuart MCA, Koning RI, Oostergetel GT, Brisson A (2004) Mechanism of formation of multilayered 2D crystals of the enzyme IIC-mannitol transporter. *Biochim Biophys Acta*, 1663(1–2):108–116.
8. Ceska TA, Henderson R (1990) Analysis of high resolution electron diffraction patterns from purple membrane labelled with heavy atoms. *J Mol Biol* 213(3):539–560.
9. Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* 405(6787):647–655.
10. Unwin PN (1975) Beef liver catalase structure: Interpretation of electron micrographs. *J Mol Biol* 98(1):235–242.
11. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.
12. Gonen T et al. (2005) Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438(7068):633–638.
13. Mitsuoka K et al. (1999) The structure of bacteriorhodopsin at 3.0 Å resolution based on electron crystallography: implication of the char*ge* distribution. *J Mol Biol* 286(3):861–882.
14. Gemmi M et al. (2013) Structure determination of nano-crystals with precession 3D electron diffraction tomography in the transmission electron microscope. *Microscopy and Analysis* 27(2):24–29.
15. Boutet S et al. (2012) High-resolution protein structure determination by serial femtosecond crystallography. *Science* 337(6092):362–364.
16. Pedrini B et al. (2014) 7 Å resolution in protein two-dimensional-crystal X-ray diffraction at Linac Coherent Light Source. *Philos Trans R Soc Lond B Biol Sci*, 369(1647):20130500.
17. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D*53(Pt 3):240–255.
18. Shi D, Nannenga BL, Iadanza MG, Gonen T (2013) Three-dimensional electron crystallography of protein microcrystals. *Elife* 2:e01345.
19. Rossmann MG (1979) Processing oscillation diffraction data for very large unit cells with an automatic convolution technique and profile fitting. *J Appl Crystallogr* 12(Pt 2):225–238.
20. Brünger AT et al. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D*54(Pt 5):905–921.

Table S1. Data collection and refinement statistics for Ca^{2+}-ATPase and catalase crystals

Specimen	Ca^{2+}-ATPase	catalase
Data collection		
Space group	*C*2	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	166.3, 64.4, 147.3	69.0, 173.5, 206.0
α, *β*, *γ* (º)	90, 98.3, 90	90, 90, 90
Resolution (Å)	145.8 – 3.4	132.7 – 3.2
Completeness (%)	67.5	72.9
R_{merge}	0.364 (0.314) *	0.360 (0.346) ‡
R_{meas}	0.375 (0.323) *	0.368 (0.353) ‡
R_{pim}	0.083 (0.070) *	0.071 (0.067) ‡
I/σ	8.44 (6.03) *	11.42 (9.59) ‡
Redundancy	15.8 (15.5) *	20.8 (20.9) ‡
Number of crystals	99	58
Refinement		
Resolution (Å)	8.0 - 3.4	8.0 - 3.2
Completeness (%)	67.5 (65.7) †	73.0 (72.8) §
R_{work}	0.277 (0.356) †	0.272 (0.361) §
R_{free} ¶	0.315 (0.392) †	0.317 (0.387) §
R.m.s deviations		
Bond lengths (Å)	0.01	0.01
Bond angles (º)	1.03	1.04
Ramachandran plot (%)		
Most favoured	82.5	77.1
Additionally allowed	16.8	21.7
Generously allowed	0.6	1.1
Disallowed	0.1	0.2

* Numbers in brackets refer to the highest resolution shell 3.52 to 3.40 Å.
† Numbers in brackets refer to the highest resolution shell 3.47 to 3.40 Å.
‡ Numbers in brackets refer to the highest resolution shell 3.31 to 3.20 Å.
§ Numbers in brackets refer to the highest resolution shell 3.27 to 3.20 Å.
¶ R_{free} is calculated for a randomly selected subset containing 3% of reflections.

Table S2. Difference peaks[*] near Tyr-357 OH in $|F_{obs}| - |F_{calc}|$ maps[†] calculated with assumed charge states around the heme binding site of catalase

charge state	O^-		OH [‡]		$O^{-0.4}H^{+0.4}$ [§]	
	Fe [¶]	Tyr [‖]	Fe	Tyr	Fe	Tyr
Fe	3.6	3.9	2.9	2.2	3.0	2.0
Fe + Fe(II) [**]	2.5	3.4	2.4	2.4	2.2	2.1
Fe(II)	-2.4	3.5	-2.2	2.3	-2.0	2.0
Fe(III)	-3.9	4.0	-3.5	2.5	-3.4	2.1

* Represented in σ.
† σ_A-weighted (m $|F_{obs}|$ – D $|F_{calc}|$).
‡ Both O and H are treated as neutral.
§ Both O and H are treated as partially charged as in Refmac (17).
¶ Near the iron atom in the heme (corresponding to the red peak in Fig. 4*C*).
‖ Near the hydroxyl group of Tyr-357 (corresponding to the green peak in Fig. 4*C*).
** Equal contributions of Fe and Fe (II) are assumed.

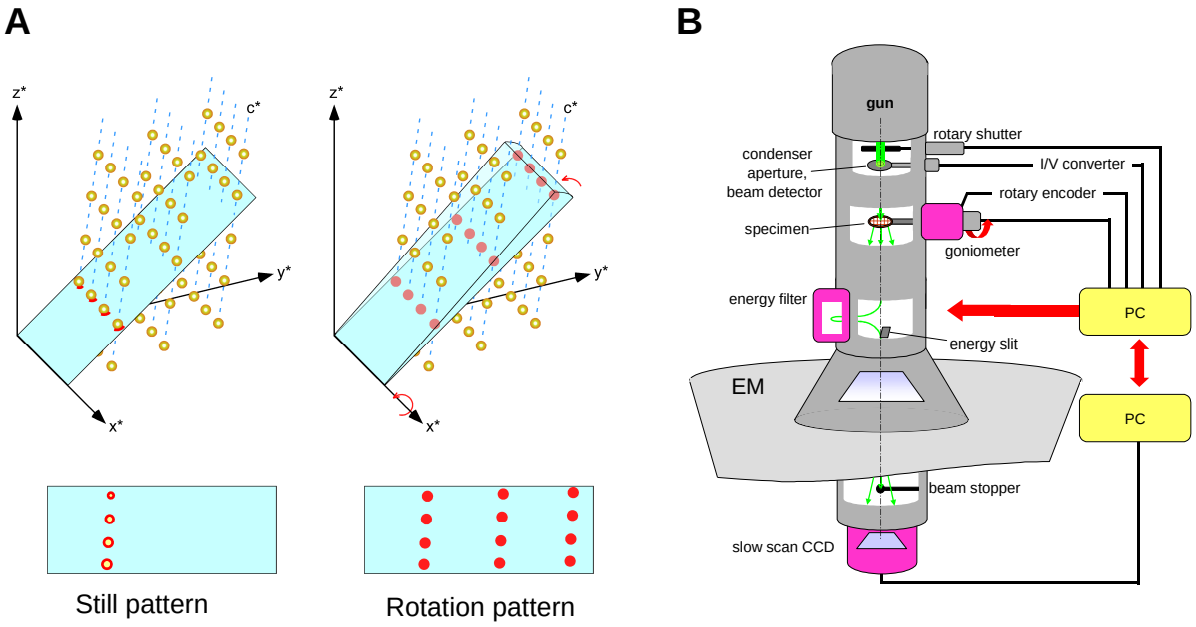

Fig. S1. Data collection from 3D crystals and design of an electron diffractometer. (*A*) Data collection from a single 3D crystal. The "still" picture records a cross-section of 3D Fourier space, thereby recording only limited numbers of "partial" diffraction spots on the CCD (*Bottom*; 18). By rotating the crystal during exposure, it is possible to sweep the fan-shaped volume and record integrated intensities of diffraction spots inside the fan-shaped wedge. (*B*) Design of an electron diffractometer for electron crystallography of ultrathin 3D protein crystals. A high-precision rotary encoder installed in the goniometer and a simple Faraday cup integrated into the condenser lens aperture continuously monitor the stage rotation (tilt) angle and incident beam current, respectively, and control the opening and closing of the shutter. Also installed in the diffractometer is an energy filter that allows passage of only unscattered and elastically scattered electrons. A beam stop placed just above the detector is effective for recoding lower-resolution reflections correctly without parasitic scattering.

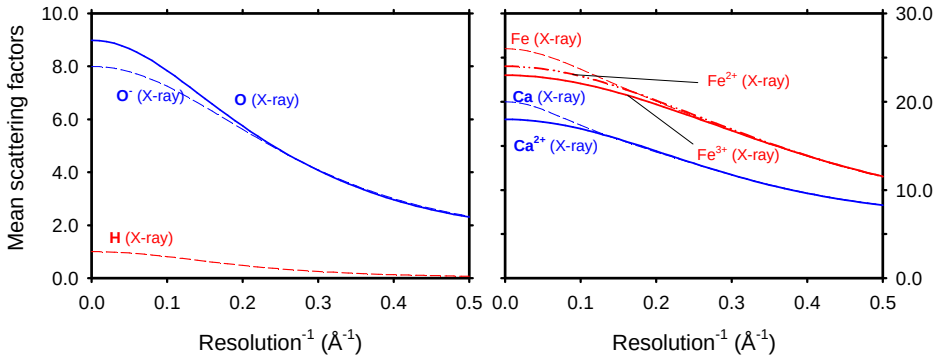

Fig. S2. Atomic scattering factors for X-rays. Scattering factors for X-rays are much less affected by the charged state compared with those for electrons. For instance, the difference between O and O^- at 0.2 $Å^{-1}$ is only 2% for X-rays but more than 460% for electrons. Similarly, the difference between Fe (II) and Fe (III) is 2% for X-rays but 26% for electrons at 0.2 $Å^{-1}$.

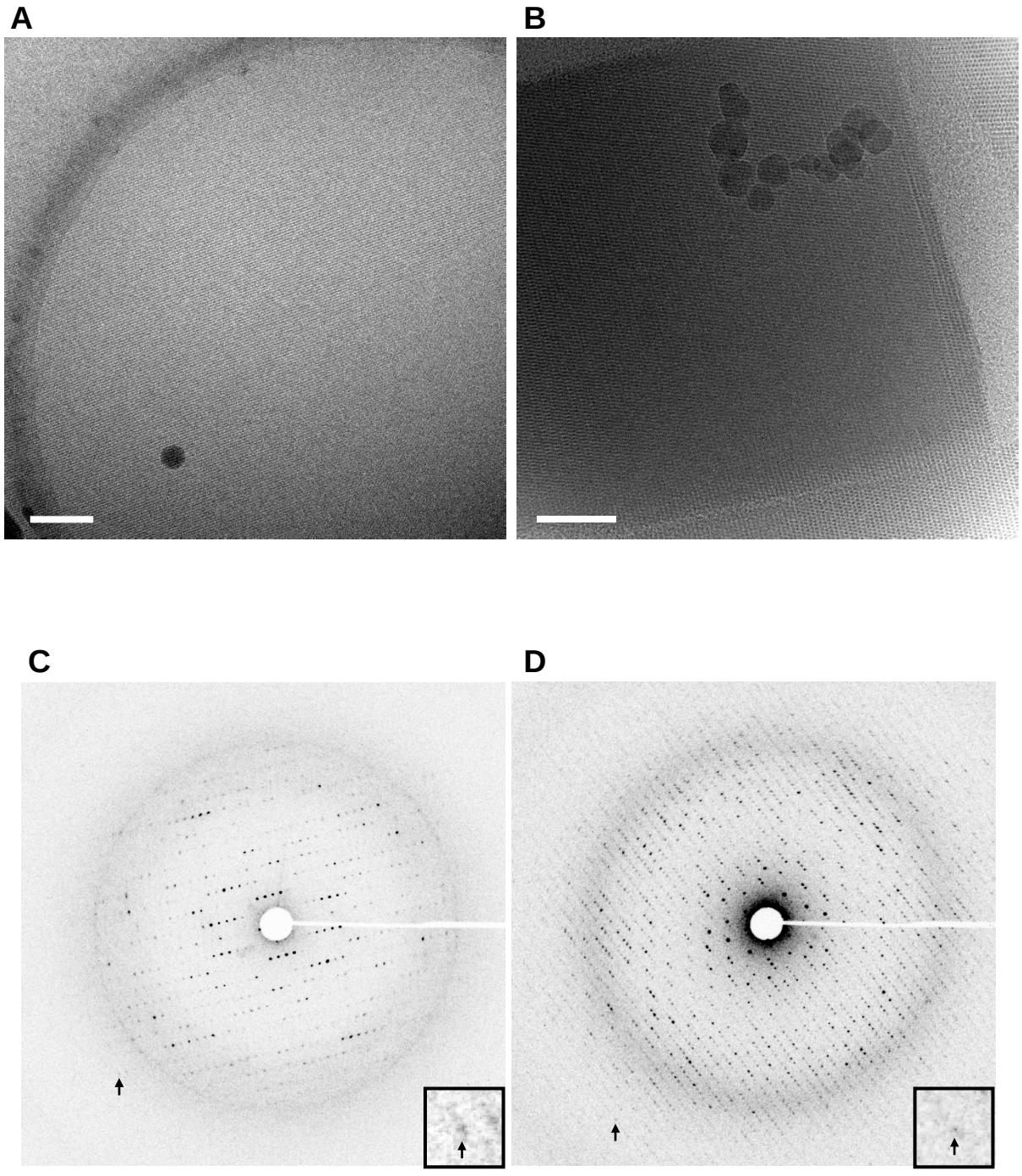

Fig. S3. Electron micrographs of frozen-hydrated ultrathin 3D crystals and electron diffraction patterns covering higher rotational (tilt) angles. (*A*) An electron micrograph of a thin 3D Ca^{2+}-ATPase crystal. (*B*) An electron micrograph of a thin 3D catalase crystal. Bars indicate 150 nm. (*C*) A frame for rotation angles from 53.0 to 54.0º, recorded from a Ca^{2+}-ATPase crystal. (*D*) A frame for rotation angles from -50.0 to -49.5º, recorded from a

catalase crystal. The exposure times of the frames are 2 s, corresponding to ~0.03 electrons per Å^2 (Ca^{2+}-ATPase) and 3 s, corresponding to ~0.04 electrons per Å^2 (catalase). Arrows indicate the reflection (49, -7, 3) at 3.3 Å (Ca^{2+}-ATPase) and (23, -6, 32) at 2.6 Å (catalase). *Insets* show a zoom-up of the reflections.

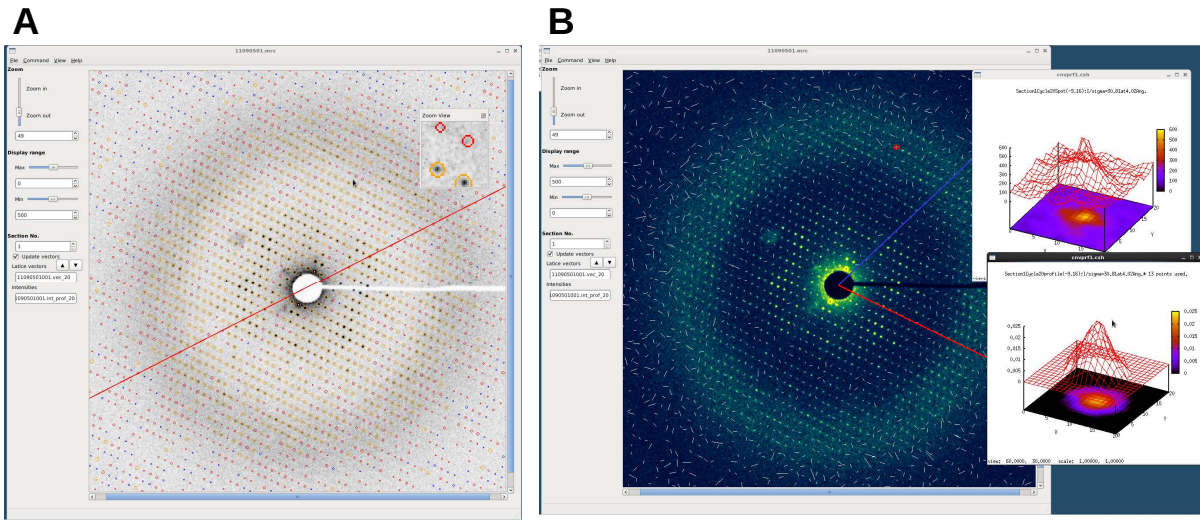

Fig. S4. Processing of electron diffraction patterns. (*A*) Indexing of diffraction spots from a Ca^{2+}-ATPase crystal with a program XQED. Multicolor circles identify diffraction spots and their quality (I/σ_I): orange, > 10; red, > 0; and blue, < 0. *Inset* shows a zoom-up of a region specified by the mouse cursor (arrow). The red line represents the rotation (tilt) axis. (*B*) Deviations of individual diffraction spots from the ideal lattice after 2D lattice refinement. White bars represent the direction and dimension of deviations (× 5 magnified). The program also displays the 2D intensity distribution of a specified diffraction spot (red circle; *Right Upper* panel) and the averaged profile around that spot used for 2D profile fitting (19) (*Right Lower* panel).

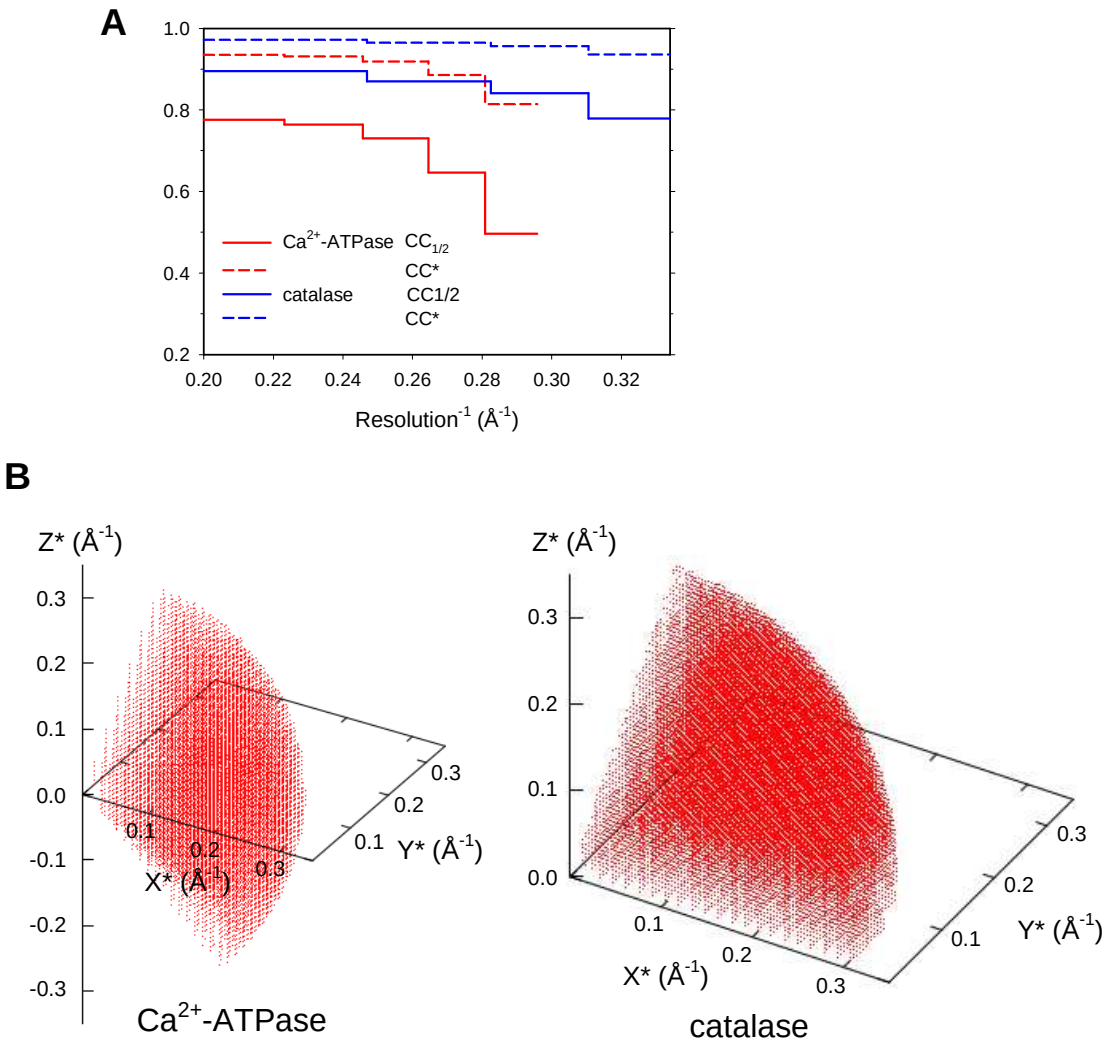

Fig. S5. Correlation and distribution of electron diffraction data. (*A*) $CC_{1/2}$ and *CC** (11). $CC_{1/2}$ is the Pearson correlation coefficient in resolution bins calculated between random halves of intensity measurements of unique reflections. *CC** is defined as,

$$CC^* = \sqrt{\frac{2 \cdot CC_{1/2}}{1 + CC_{1/2}}} \quad .$$

(*B*) Distribution of the merged diffraction spots used for structure analyses in this study. Completeness: Ca^{2+}-ATPase, 67.5%; catalase, 72.9%.

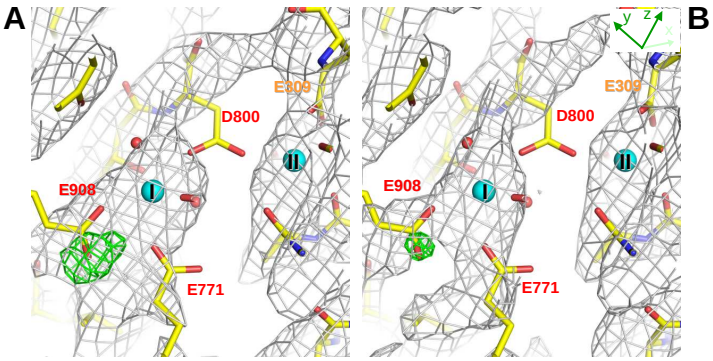

Fig. S6. Coulomb potential maps of Ca^{2+}-ATPase calculated from two independent datasets. (*A* and *B*) σ_A-weighted 2 $|F_{obs}| - |F_{calc}|$ maps (1σ; grey nets) were calculated from 8- to 3.4-Å resolution data derived from two data subsets (49 and 50 datasets) randomly selected from the full dataset. The σ_A-weighted $|F_{obs}| - |F_{calc}|$ difference map (3.0σ; green nets) shows peaks of 3.9σ (*A*) and 3.2σ (*B*), indicating the presence of a proton.

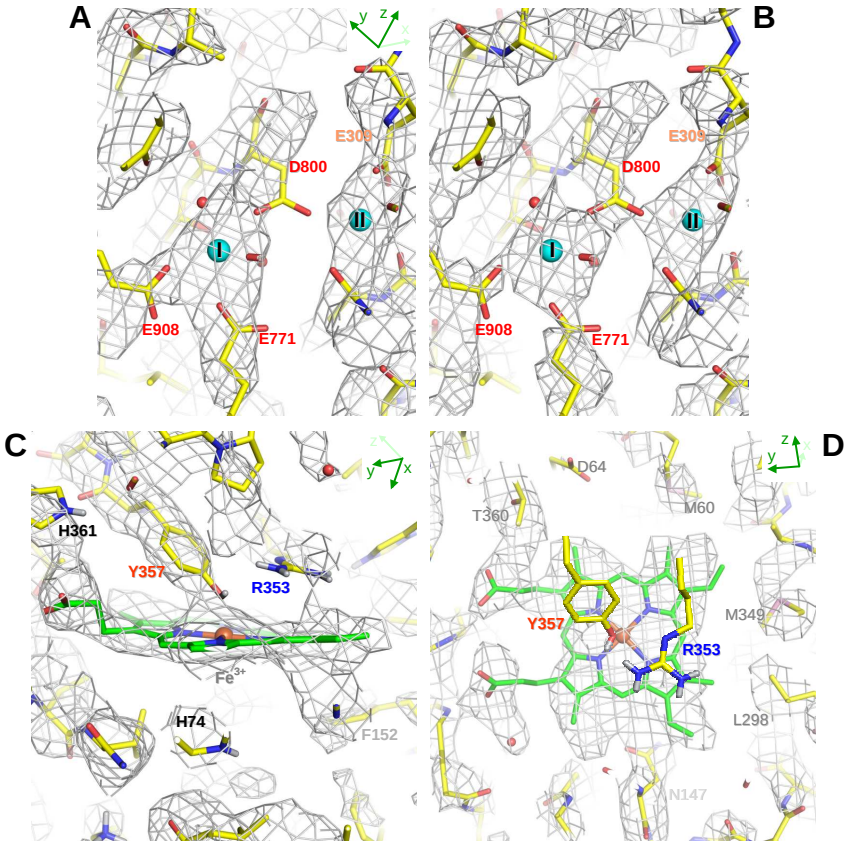

Fig. S7. Composite omit maps. (*A* and *B*) Around the Ca^{2+}-binding site of Ca^{2+}-ATPase. Calculated from 8- to 3.4-Å (*A*) or 5- to 3.4-Å (*B*) resolution data. (*C* and *D*) Around the heme binding site of catalase. Without NCS averaging. (*C*) Viewed as in Fig. 5*C* and *D*. (*D*) Viewed as in Fig. 5*B*. Calculated with CNS (20) and contoured at 1.1σ in (*A*) and (*B*) and at 1.3σ in (*C*) and (*D*). Overlaid with the atomic models refined in this study.

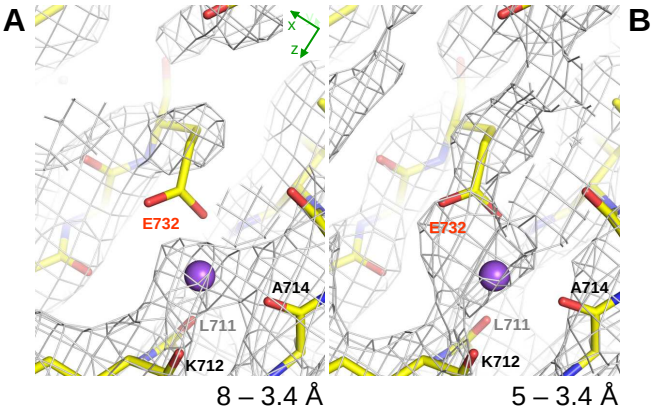

Fig. S8. Coulomb potential maps around the Na^+-binding site of Ca^{2+}-ATPase. (*A* and *B*) σ_A-weighted 2 $|F_{obs}| - |F_{calc}|$ maps (1σ; grey nets). The density for the carboxyl group of Glu-732 is absent in the 8- to 3.4-Å resolution map (*A*), but present in the 5- to 3.4-Å resolution map (*B*). Purple spheres represent one bound Na^+.

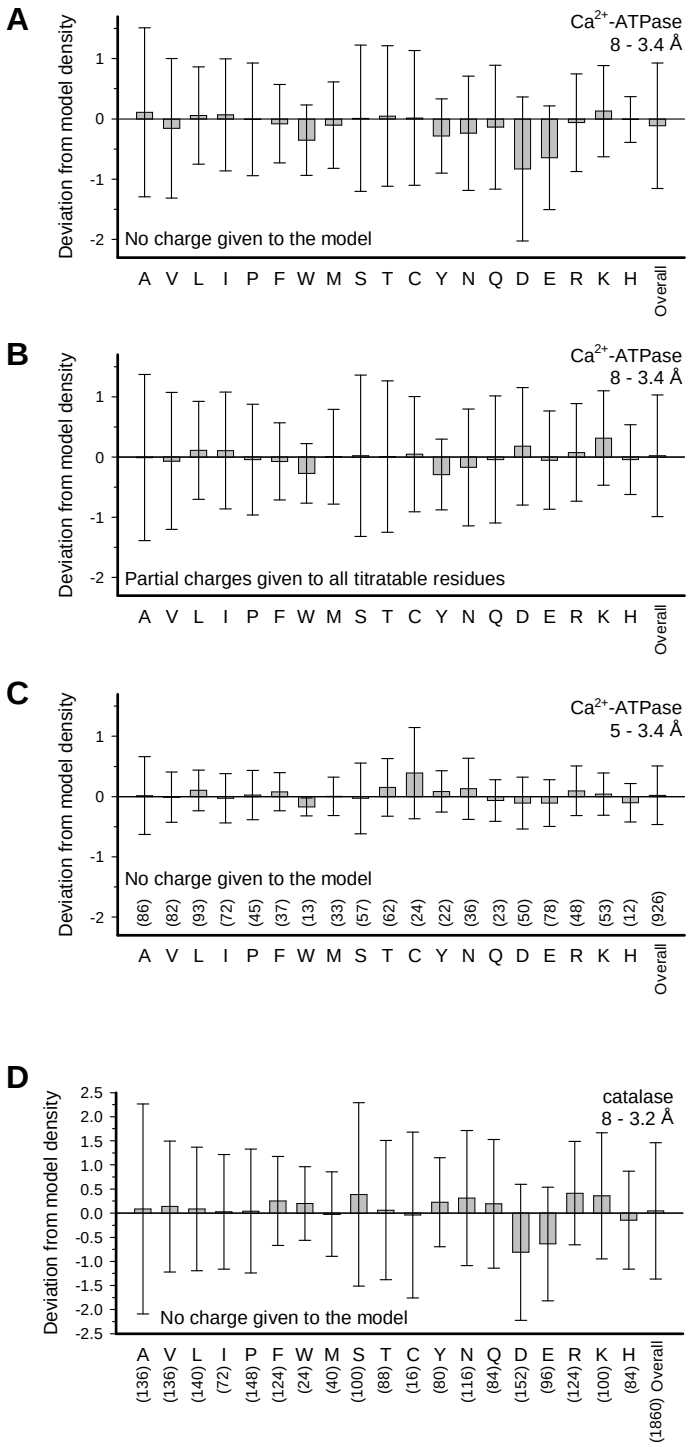

Fig. S9. Deviation from model density of the side chains in the difference (σ_A-weighted $|F_{obs}| - |F_{calc}|$) Coulomb potential maps of Ca^{2+}-ATPase and catalase. Deviations are calculated as described in the legend of Fig. 4. (*A*–*C*) Ca^{2+}-ATPase. (*D*) Catalase (without NCS averaging). No charge is given to all amino-acid residues (*A*, *C* and *D*) or partial charges

are given to all titratable residues (*B*) in the atomic models used for F_{calc}. Resolution ranges are 8 to 3.4 Å (*A* and *B*), 5 to 3.4 Å (*C*) and 8 to 3.2 Å (*D*). The bar charts in *A* and *D* show large negative values for Asp and Glu residues, reflecting loss or weaker densities of negatively-charged residues. These deviations become smaller when standard charges are assigned to the acidic side chains (*B*) or lower-resolution data are excluded (*C*). R_{work} / R_{free} values are 0.285 / 0.320 (*A*) and 0.288 / 0.318 (*D*) (cf. R_{work} / R_{free} in Table S1). Standard deviations are shown as error bars and the vertical axis in arbitrary units. The numbers in brackets correspond to those of amino acid residues.

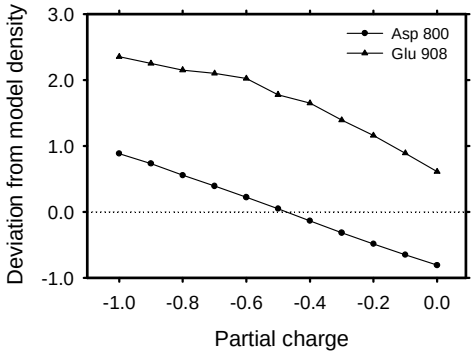

Fig. S10. Deviation from model density of the side chains of Asp-800 and Glu-908 in the difference Coulomb potential maps of Ca^{2+}-ATPase. Plotted against partial charges assigned to the carboxyl oxygen atoms. Calculated from 8- to 3.4-Å resolution data as described in the legend to Fig. 4.

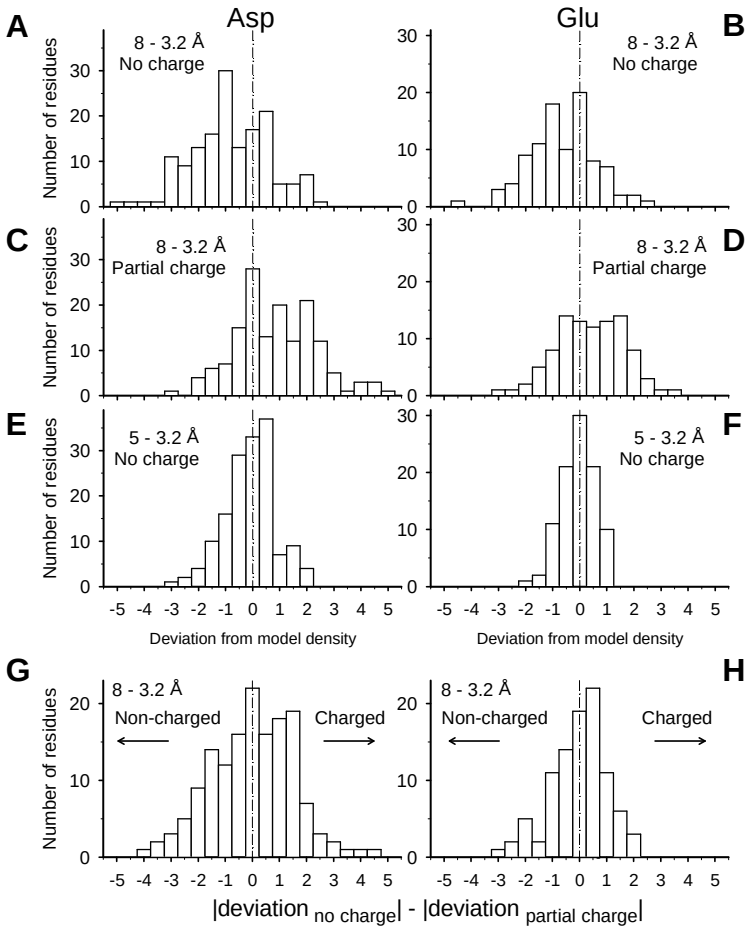

Fig. S11. Deviation from model density of the side chains of acidic residues in the difference (σ_A-weighted $|F_{obs}| - |F_{calc}|$) Coulomb potential map of catalase. Deviations are calculated as in the legend of Fig. 4 without NCS averaging. (*A*–*F*) Histograms of the deviation for Asp (*A*, *C* and *E*) and Glu (*B*, *D* and *F*). All residues are treated as not charged (*A*, *B*, *E* and *F*), or partial charges are assigned to all titratable residues (*C* and *D*) in the atomic models used for F_{calc}. Resolution ranges are 8 to 3.2 Å (*A*–*D*) and 5 to 3.2 Å (*E* and *F*). (*G* and *H*) Histograms of difference in deviation (absolute number) from model density between the two $|F_{obs}| - |F_{calc}|$ maps, in which no or partial charges are given to all titratable residues [i.e. difference between *A* and *C* (*G*) or *B* and *D* (*H*)]; calculated for the Asp and Glu side chains using the data from 8- to 3.2-Å resolution.

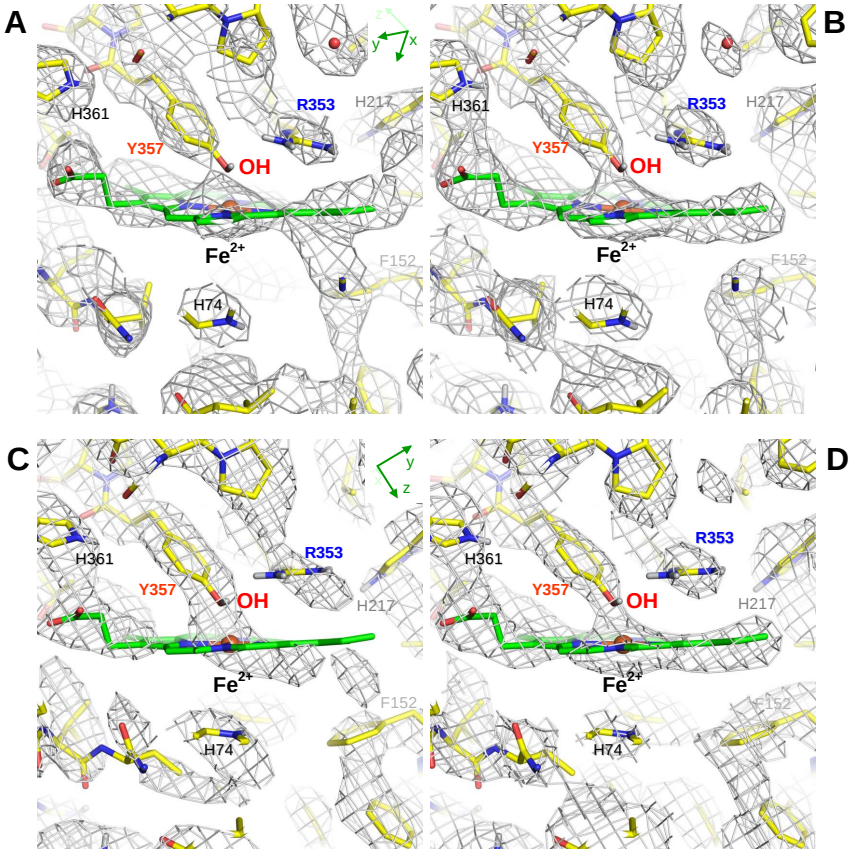

Fig. S12. Improvement of Coulomb potential maps of catalase by NCS averaging. Before (*A* and *C*) and after (*B* and *D*) NCS averaging of 4 subunits (*A* and *B*, subunit A; *C* and *D*, subunit C). Viewed as in Fig. 5*C* and *D* and contoured at 1.3σ and overlaid with the atomic model refined in this study. Note that NCS averaging improved the appearance of the map but was more effective in the case of the poorer map (subunit C).