XFEL structure of carbonic anhydrase II: a comparative study of XFEL, NMR, X-ray and neutron structures

Joshua A. Hull,[a]‡ Cheol Lee,[b]‡ Jin Kyun Kim,[b]§ Seon Woo Lim,[b] Jaehyun Park,[c,d] Sehan Park,[c] Sang Jae Lee,[c] Gisu Park,[c] Intae Eom,[c] Minseok Kim,[c] HyoJung Hyun,[c] Jacob E. Combs,[a] Jacob T. Andring,[a] Carrie Lomelino,[a] Chae Un Kim[b]* and Robert McKenna[a]*

[a]Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Gainesville, FL 32610, USA, [b]Department of Physics, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea, [c]Pohang Accelerator Laboratory, POSTECH, Pohang 37673, Republic of Korea, and [d]Department of Chemical Engineering, POSTECH, Pohang 37673, Republic of Korea. *Correspondence e-mail: cukim@unist.ac.kr, rmckenna@ufl.edu

The combination of X-ray free-electron lasers (XFELs) with serial femtosecond crystallography represents cutting-edge technology in structural biology, allowing the study of enzyme reactions and dynamics in real time through the generation of 'molecular movies'. This technology combines short and precise high-energy X-ray exposure to a stream of protein microcrystals. Here, the XFEL structure of carbonic anhydrase II, a ubiquitous enzyme responsible for the interconversion of CO_2 and bicarbonate, is reported, and is compared with previously reported NMR and synchrotron X-ray and neutron single-crystal structures.

1. Introduction

In recent years, the use of X-ray free-electron lasers (XFELs) to determine the structural dynamics of enzymatic reactions and track the movement of substrates and resultant products has become a cutting-edge technique in structural biology, enabling the study of the molecular mechanism of enzymes at rates on the femtosecond timescale (Lomelino et al., 2018).

One enzyme for which XFEL structural methods could be utilized in order to study its mechanism is the zinc metalloprotein carbonic anhydrase II (CAII), which is responsible for the rapid reversible hydration/dehydration of carbon dioxide (CO_2)/bicarbonate (HCO_3^-) and has been at the leading edge of understanding enzymatic reactions since its discovery (Meldrum & Roughton, 1933; Stadie & O'Brien, 1933; Silverman & Lindskog, 1988; Fisher et al., 2011). It has also been a key system for the study and development of structural biology methods, with the determination of its X-ray (Kannan et al., 1975; Sippel et al., 2009), neutron (Fisher et al., 2011; Sippel et al., 2009) and NMR (Singh et al., 2019) structures and subsequently its X-ray structure with the CO_2 substrate bound (Domsic et al., 2008). In addition, CAII also has important physiological roles in gas exchange, ion transport and extracellular and intracellular pH regulation (Frost & McKenna, 2014).

CAII is a 30 kDa protein with a solvent-accessible active-site cavity; the zinc is at the base of the cavity, tetrahedrally coordinated by three histidines (His94, His96 and His119) and a solvent molecule (Kannan et al., 1975; Sippel et al., 2009). The active-site cavity is furthermore subdivided into distinct hydrophobic and hydrophilic sides. The hydrophobic side (Ile91, Val121, Phe131, Val135, Leu141, Val143, Leu198,
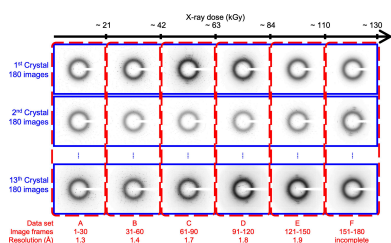
Pro202, Leu204, Val207 and Trp209) stabilizes the CO_2 substrate, while the hydrophilic side (Asn62, His64, Asn67, Gln92, Thr199 and Thr200) orders and regulates the solvent [W1, W2, W3a, W3b and Dw (deep water)] required for rapid catalytic turnover (Frost & McKenna, 2014). Of special importance is His64, which modulates between an 'in' and an 'out' conformation (referring to its direction relative to the active site) and is known to be important in proton transfer (Fisher *et al.*, 2011).

The role of CAII in the hydration/dehydration of CO_2/HCO_3^- has been extensively studied structurally and biophysically. The reaction is a two-step, ping-pong mechanism. In the hydration direction, the first step is the nucleophilic attack on CO_2 of a zinc-bound hydroxyl that results in the formation of HCO_3^-, which is displaced by a water molecule (Domsic *et al.*, 2008; Kim *et al.*, 2020). The second step of the reaction is the transfer of a proton from the zinc-bound water to the bulk solvent via the well defined solvent network and His64 (Silverman & Lindskog, 1988; Kim *et al.*, 2018). Regeneration of the zinc-bound hydroxyl permits the catalytic reaction cycle. The k_{cat}/K_m of the reaction is 120 M^{-1} μs^{-1}, which means that Zn-CAII has evolved to catalytic near-perfection for the hydration/dehydration of CO_2/HCO_3^-, as it is diffusion rate-limited (Maupin *et al.*, 2009).

In this study, we use CAII to characterize XFEL serial crystallography as a structure-determination method in comparison to other available methods. As described previously, the relative radiation damage and thermal motion, as indicated by the B factors, are elevated in the XFEL structures (Lomb *et al.*, 2011; Nass, 2019); however, the quality and completeness of the CAII structure appear to be unaffected. Differences in the NMR-determined CAII structure compared with synchrotron and neutron diffraction crystallographic structures highlight the amino-terminal domain (NTD_{1-30}) and the amino-acid region proximal to it ($Core_{197-204}$) as displaying high flexibility and B factors in the NMR structure. However, the B factor in NMR is not directly analogous to that in crystallography (Reinknecht *et al.*, 2021). In comparison to these methods, in terms of B factors the XFEL CAII structure was more analogous to the previously determined synchrotron and neutron diffraction crystallographic structures, with the only exception being elevated B factors and an apparent slight reduction in data quality associated with this observation. Additionally, the NMR method resulted in a lack of information on the ordered water network in the deposited structure, while the neutron diffraction crystallographic structure provided the most complete detailed water network, with H atoms being observed; in the determined XFEL structure the placement of the water O atoms of the network was observed, but with no indication of the hydrogen-bond network stabilizing the solvent. Taken together, these results imply that despite the high B factors for the XFEL-determined structure, the overall results can be taken as analogous to traditional synchrotron crystallography, therefore confirming that XFEL serial crystallography is an excellent tool to address time-resolved questions regarding the catalytic mechanism of CAII.

2. Materials and methods

2.1. Production and purification of human CAII

Human CAII was expressed and purified according to previously published protocols (Pinard *et al.*, 2013; Tanhauser *et al.*, 1992). Briefly, a CAII gene-containing plasmid under the control of a T7 promoter was transformed into competent *Escherichia coli* BL21 cells via a standard transformation protocol. The cells were grown to an optical density of 0.6 at 600 nm in the presence of antibiotics. The cells were then induced by the addition of isopropyl β-D-1-thiogalacto-pyranoside (IPTG) and supplementary zinc sulfate and incubated for an additional 3 h. The cells were pelleted via centrifugation and subsequently lysed using a microfluidizer (Microfluidics model LM-10) set to 124 MPa. CAII was purified from the cell lysate by affinity chromatography using a *p*-aminomethylbenzenesulfonamide affinity column. The final protein stock was buffer-exchanged with storage buffer (50 mM Tris pH 7.8) using a centrifugal filter. The purity was determined by SDS–PAGE and the protein concentration was determined from the optical density at 280 nm.

2.2. Crystallization for room-temperature multi-crystal crystallography

Crystals of CAII were obtained using the hanging-drop vapor-diffusion method (Díaz-Torres *et al.*, 2015). A 10 µl drop consisting of equal volumes of protein solution (5 µl) and well solution (5 µl) was equilibrated against 500 µl precipitant solution (1.6 M sodium citrate, 50 mM Tris–HCl pH 7.8) at room temperature ($\sim$20°C; Lomelino *et al.*, 2018). Crystals grew to approximate dimensions of 200 $\times$ 50 $\times$ 50 µm in a few days.

2.3. Crystallization for serial femtosecond crystallography

Upon visual inspection, a single high-quality crystal of CAII was transferred from the crystal drop into a new 10 µl drop of precipitant solution and crushed using a needle. The needle was then dipped into a secondary 10 ml droplet of precipitant solution to dilute the crystals and create a CAII seed stock.

The CAII microcrystals for serial femtosecond crystallography were grown at room temperature ($\sim$20°C) utilizing a combination of seeding and batch crystallization methods by adding CAII seed stock and purified protein directly to the precipitant solution. In a 24-well culture plate, 5 µl seed stock, 300 µl CAII (30 mg ml^{-1}) and 1.2 ml precipitant solution were added to each well as described previously (Lomelino *et al.*, 2018). Microcrystal growth was observed after 12 h. The microcrystal suspension was diluted in precipitant solution (1:4 ratio) and syringe-filtered through a metal filter, removing crystals of greater than 100 µm in length. After five successive filtrations, the microcrystal suspension was concentrated by centrifugation at $\sim$840g for 5 min. The microcrystals were then mixed with monoolein (100%, Hampton Research) in a 1:1 ratio in gas-tight syringes and transferred into a lipid cubic phase (LCP) injector for data collection.

2.4. Data collection for room-temperature multi-crystal crystallography

Diffraction data were collected on beamline 5C at Pohang Light Source II (wavelength 0.9794 Å, X-ray energy 12.659 keV, beam size ∼100 µm) at room temperature (∼20°C). Data were collected using the oscillation method in intervals of 1° on an ADSC Quantum 315r CCD detector (Area Detector Systems Corporation, USA) with 3072×3072 pixel resolution (2×2 binning, pixel size 51.2 µm) and a crystal-to-detector distance of 150 mm. A total of 180 images were collected from each of the CAII crystals. Data collections were performed at room temperature from 13 fresh crystals coated with mineral oil to prevent dehydration.

To observe the effect of radiation damage, six different data sets (A–F) were obtained, scaled and merged from the 13 crystal data sets. Images 1–30 from the 13 crystals were merged into data set A with an X-ray dose of ∼2.1×10^4 Gy, images 31–60 from the 13 crystals were merged into data set B with an X-ray dose of ∼4.2×10^4 Gy, images 61–90 from the 13 crystals were merged into data set C with an X-ray dose of ∼6.3×10^4 Gy, images 91–120 from the 13 crystals were merged into data set D with an X-ray dose of ∼8.4×10^4 Gy, images 121–150 from the 13 crystals were merged into data set E with an X-ray dose of ∼1.1×10^5 Gy and images 151–180 from the 13 crystals were merged into data set F with an X-ray dose of ∼1.3×10^5 Gy (Fig. 1). However, the absorbed X-ray dose for data set F was close to the X-ray dose limit of 4.0×10^5 Gy at room temperature (Barker et al., 2009; Warkentin et al., 2012; Fischer, 2021; de la Mora et al., 2020) and the data quality was too poor to perform refinement. Indexing, integration and scaling were performed using HKL-2000 (Otwinowski & Minor, 1997). The 'expected crystal lifetime calculator' (https://bl831.als.lbl.gov/xtallife.html) was used to determine the dose applied to the crystals. The data-processing statistics are given in Table 1.

2.5. Data collection for serial femtosecond crystallography

Diffraction data were collected at the Coherent X-ray Imaging (CXI) station at Pohang Accelerator Laboratory XFEL (Yu et al., 2014) [wavelength 1.278 Å, X-ray energy 9.7 keV, bandwidth 26 eV, beam size ∼3×4 µm (horizontal × vertical), 50 fs pulse width, 30 Hz repetition rate] at room temperature. Data were collected on a Rayonix MX225-HS CCD detector with 1440×1440 pixel resolution (4×4 binning, pixel size 156 µm) and a crystal-to-detector distance of 108 mm using serial femtosecond crystallography, as described previously (Lomb et al., 2011). The suspension of the microcrystals in monoolein allowed sample injection using an isocratic flow mode at a flow rate of 470 nl min⁻¹ with an injector nozzle diameter of 75 µm. Approximately 0.7 million
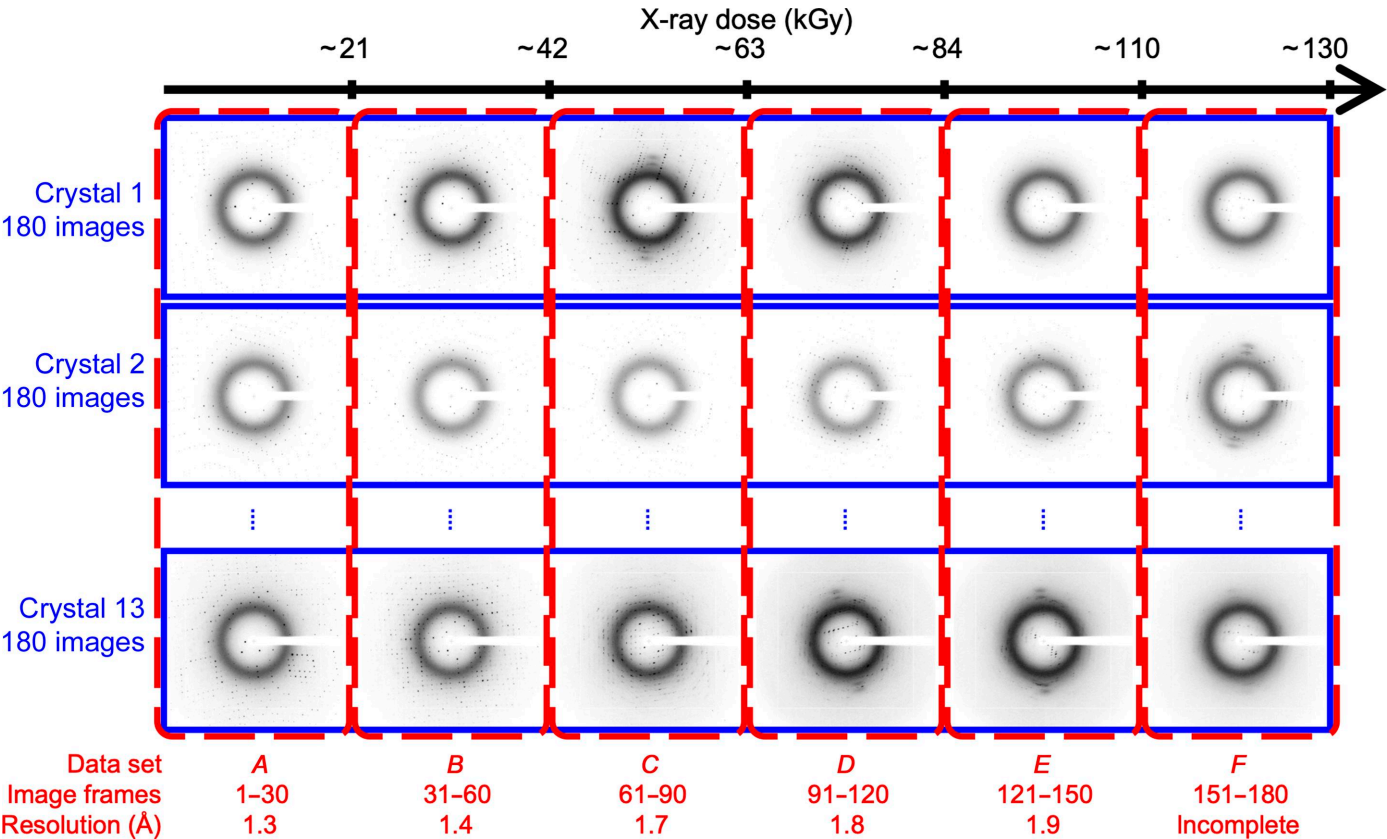

Figure 1
Room-temperature multi-crystal crystallography to observe radiation damage at a synchrotron source. The blue boxes represent each CAII crystal sample and 180 diffraction pattern images were collected with the oscillation method in 1° intervals. The red dotted boxes represent each data set; 30° wedges from the 13 crystals were scaled and merged to obtains the CAII structure. The X-ray dose was increased from data sets A–F (black arrow).

Table 1
Data-collection and refinement statistics for serial femtosecond crystallography (SFX) and room-temperature (RT) multi-crystal crystallography.

Values in parentheses are for the highest resolution shell.

Data set	XFEL (SFX)	RT A	RT B	RT C	RT D	RT E
Wavelength (Å)	1.278	0.9794				
X-ray focus (µm)	$\sim3 \times 4$	$\sim100 \times 100$				
Pulse energy (µJ)	190	n.a.				
Fluence at sample	1.2×10^{11} photons per pulse	1.4×10^{10} photons s^{-1}				
Dose per crystal (MGy)	8.3	0.021	0.042	0.063	0.084	0.11
Dose rate (Gy s^{-1})	1.7×10^{20}	7.0×10^{2}				
Space group	$P2_1$					
a, b, c (Å)	42.49, 41.30, 72.15	42.59, 41.32, 72.50	42.57, 41.33, 72.36	42.71, 41.58, 72.71	42.71, 41.54, 72.75	42.73, 41.59, 72.77
α, β, γ (°)	90, 104.63, 90	90, 104.37, 90	90, 104.28, 90	90, 104.44, 90	90, 104.48, 90	90, 104.90, 90
Oscillation range (°)	Still experiment	1.0				
Exposure time	50 fs†	1.0 s				
No. of images	729169	30×13				
No. of hits	153936	n.a.				
No. indexed	57713	30×13				
Resolution (Å)	30.0–1.70 (1.73–1.70)	30.0–1.30 (1.32–1.30)	30.0–1.40 (1.42–1.40)	30.0–1.70 (1.73–1.70)	30.0–1.80 (1.83–1.80)	30.0–1.90 (1.93–1.90)
No. of reflections	3403288	353369	274033	149193	113670	83106
No. unique	25565	59893	48076	27126	23193	19327
Completeness (%)	100.0 (100.0)	99.0 (98.1)	99.2 (98.9)	99.8 (100.0)	99.2 (99.0)	98.7 (99.0)
Multiplicity (%)	130.3 (86.6)	5.9 (5.3)	5.7 (5.2)	5.5 (5.3)	4.9 (4.6)	4.3 (4.3)
$\langle I/\sigma(I) \rangle$	5.51 (1.75)	16.00 (2.05)	16.05 (2.15)	16.96 (4.48)	14.48 (4.06)	13.62 (4.09)
R_{split}‡ (%)	16.5	n.a.	n.a.	n.a.	n.a.	n.a.
R_{merge} (%)	n.a.	12.2 (65.9)	12.1 (96.5)	18.3 (125.2)	19.2 (93.0)	15.7 (76.0)
B factor (Å^2)	49.2	22.7	32.1	27.4	24.0	23.9
R_{work} (%)	0.2179 (6.4776)	0.1403 (0.2035)	0.1374 (0.2175)	0.1315 (0.1557)	0.1359 (0.1629)	0.1359 (0.1573)
R_{free} (%)	0.2490 (7.0682)	0.1492 (0.2250)	0.1535 (0.2395)	0.1528 (0.1923)	0.1595 (0.1688)	0.1633 (0.2039)
R.m.s.d., bond lengths (Å)	0.017	0.009	0.009	0.008	0.023	0.009
R.m.s.d., angles (°)	2.05	1.38	1.38	1.04	2.01	1.16
PDB code	8sf1	8sd1	8sd6	8sd7	8sd8	8sd9

† Electron bunch length. ‡ $R_{\mathrm{split}} = 2^{1/2}(\sum_{hkl} |I^{\mathrm{even}}_{hkl} - I^{\mathrm{odd}}_{hkl}|)/(\sum_{hkl} |I^{\mathrm{even}}_{hkl} + I^{\mathrm{odd}}_{hkl}|)$.

still diffraction patterns were collected. Of these, $\sim20\%$ were classified as hit images, defined as an image that has more than five peaks, with each peak comprising more than two pixels that have a signal-to-noise ratio of >4 on the detector. Of these, 37% of hit images were indexed and integrated with *CrystFEL* (White *et al.*, 2016). Details are given in Table 1.

2.6. Data refinement

The diffraction images were indexed and integrated using *XDS* and then merged and scaled in space group $P2_1$ using *AIMLESS* via the *CCP*4 suite (Kabsch, 2010; Evans & Murshudov, 2013; Agirre *et al.*, 2023). The diffraction data were phased using standard molecular-replacement methods in the *Phenix* package (Liebschner *et al.*, 2019) using PDB entry 3ks3 (Avvaru *et al.*, 2010) as the search model. Coordinate refinements were performed using *Phenix*, while *Coot* was utilized to add solvent molecules and to perform individual real-space refinements of each residue when needed (Liebschner *et al.*, 2019; Emsley *et al.*, 2010). Data statistics can be found in Table 1. Figures were generated using the *UCSF Chimera* molecular-graphics software (Pettersen *et al.*, 2004).

3. Results and discussion

3.1. Crystallization and data collection

High-quality CAII crystals were grown at room temperature (RT) and prepared for data collection at a synchrotron or for XFEL data collection as described in Section 2. The RT crystals exposed to synchrotron radiation were binned into six data sets based on the angular rotation of the crystals (Fig. 1). We presume that the unit cells of the crystals submitted to the synchrotron are the same (due to the binning of the data) and they had unit-cell parameters $a = 42.7 \pm 0.1$, $b = 41.6 \pm 0.3$, $c = 72.8 \pm 0.4$ Å, $\beta = 104.4 \pm 0.1°$ (Table 1). The XFEL-determined structure had similar unit-cell parameters $a = 42.9$, $b = 42.0$, $c = 73.4$ Å, $\beta = 104.6°$ (Table 1). The final resolutions of the binned data groups were 1.3, 1.4, 1.7, 1.8, 1.9 and 1.7 Å for synchrotron data sets A, B, C, D and E and the XFEL data set, respectively.

For the synchrotron data, split by radiation dose, a steady decrease in the number of observed reflections (both in total and in the highest resolution shells) was seen (Supplementary Fig. S1). This result is most likely to be a consequence of consistent radiation damage to the crystal lattice with increased X-ray exposure time. In contrast, for the XFEL data set the entire crystal was vaporized, with each image being collected in a single X-ray hit.

For the synchrotron data, the global B factor increased in data set B but then continued at the baseline observed in data set A (Table 2), whereas the global B factors for the XFEL data set were seen to be significantly higher when compared with any of the dose-split synchrotron data sets. This suggests that radiation damage does not deteriorate the certainty of atomic positions in the synchrotron data to the same extent as in the XFEL data, only the number of reflections observed,
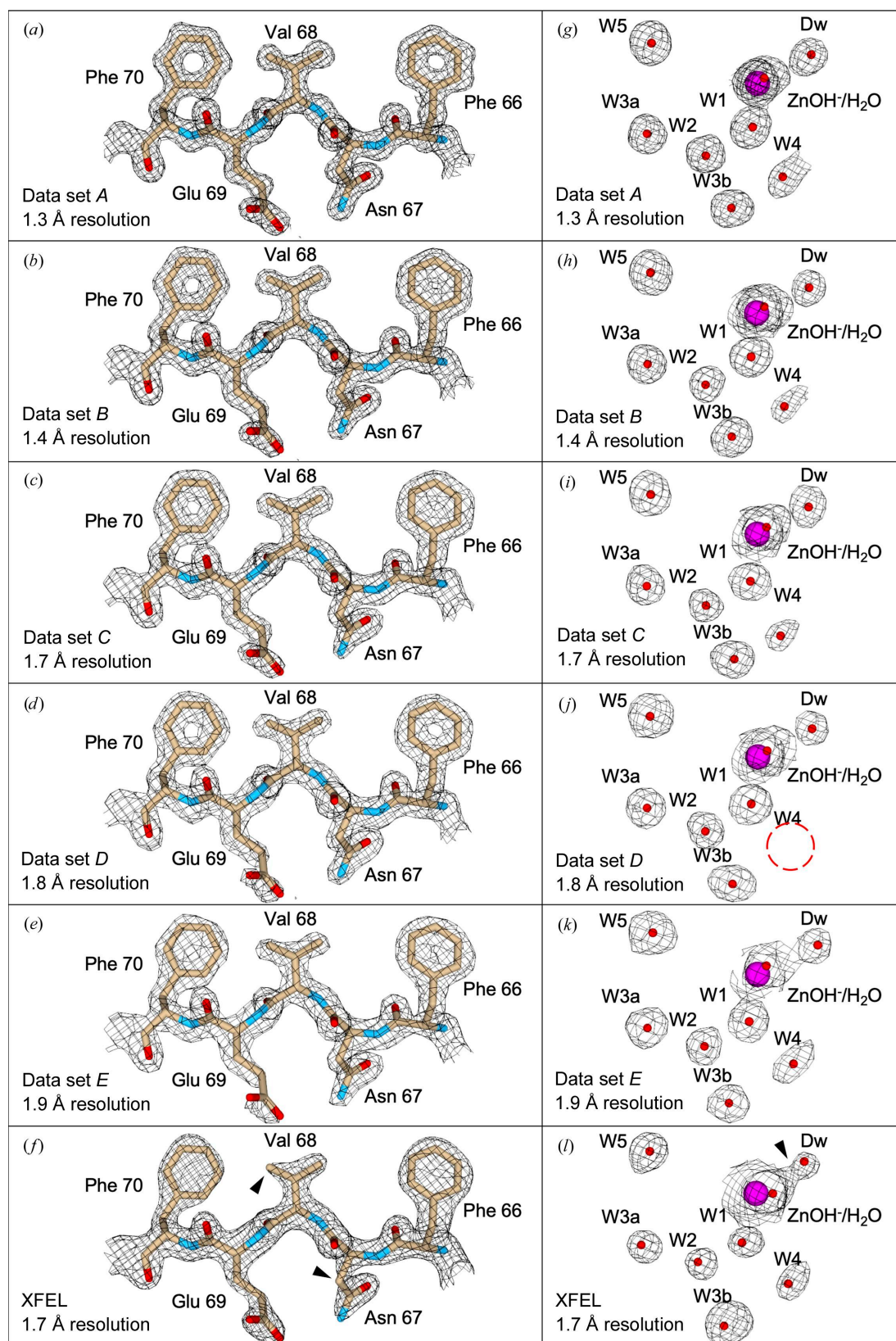
Figure 2
Quality assessment of structures. (*a–f*) Density for the *β*-strand region, residues 66–70, is shown for data sets *A* (*a*), *B* (*b*), *C* (*c*), *D* (*d*), *E* (*e*) and the XFEL data (*f*). (*g–l*) Density for the active-site ordered water networks is shown for data sets *A* (*g*), *B* (*h*), *C* (*i*), *D* (*j*), *E* (*k*) and the XFEL data (*l*). In (*f*), regions of poor map quality in the XFEL structure are indicated for Asn67 and Val68 by a black arrow. The aromatic residues Phe66 and Phe70 are also noted to have no aromatic 'donut holes' as observed for data sets *A–D* (*a–d*). W4 was not resolved in data set *D* (*j*). In (*l*), regions of poor map quality in the XFEL structure are apparent from the merged density between ZnOH⁻/H₂O and Dw, which is indicated by a black arrow.

while for the XFEL data the atomic certainty would be reduced without a corresponding loss of diffraction. Assessment of the structure quality within equivalent regions of density demonstrates that this appears to be the case (Fig. 2). For data sets A–E with reduced resolution, the side-chain quality was reduced as expected with resolution (Figs. 2a–2e). For comparison, the XFEL data quality was similar to or worse than data set E, with no apparent aromatic 'donut-like' holes and less well defined side chains (evident for Asn67 and Val68 in the figure) than in data sets A–E. Within the ordered

water network in the active site, all waters were observed in the structures except for data set D, implying that increasing radiation dose has a detrimental impact on the certainty with which water assignments can be made (Figs. 2f–2l). Again, the XFEL data quality was similar to data set E or worse, which was most evident in the blurring of density between $ZnOH^-$/H_2O and Dw. Therefore, while the XFEL crystals do diffract to high resolution, the apparent quality of the data is lower than would be expected from a crystal structure at similar resolution and shows signs of radiation damage, apparent in
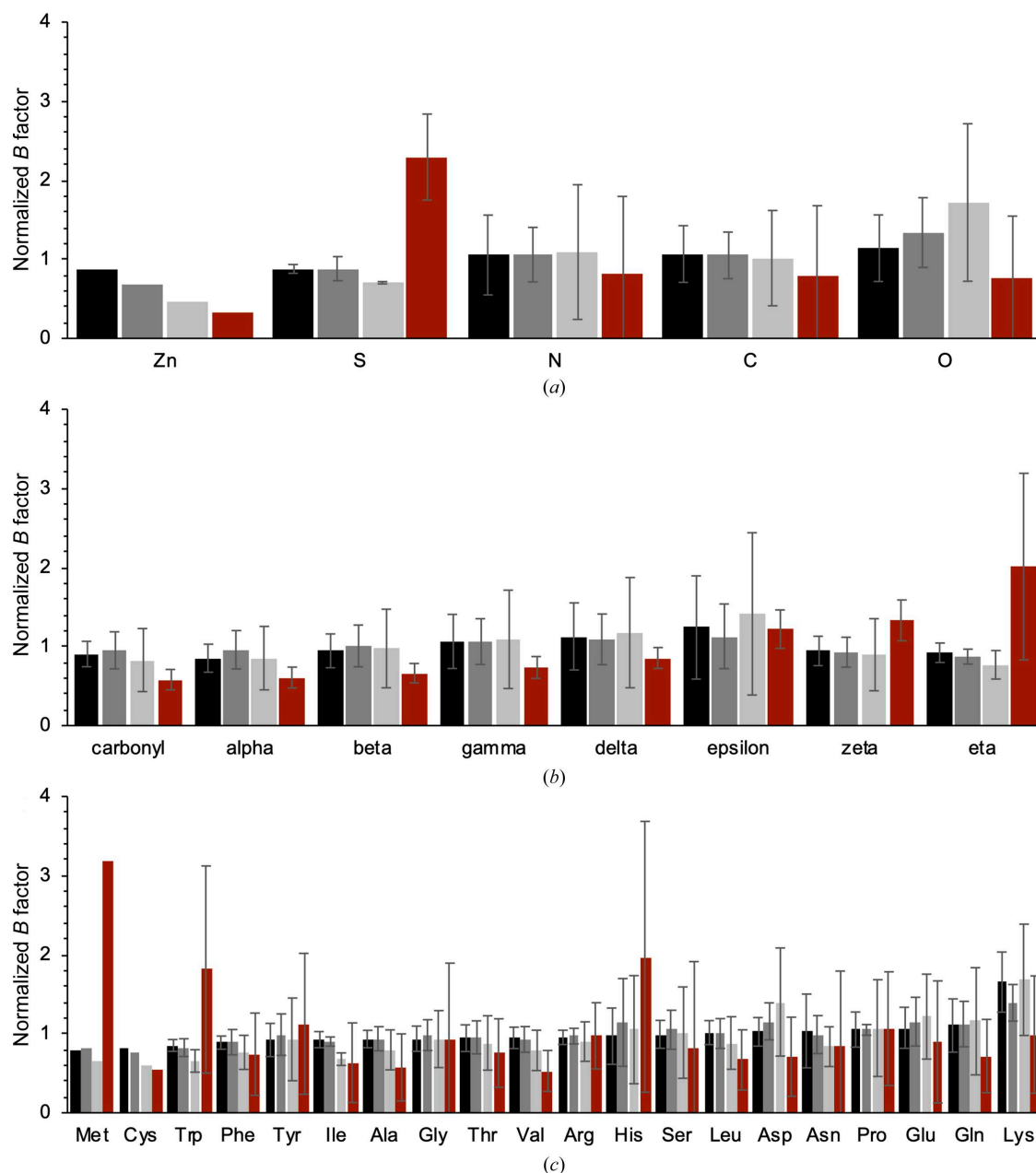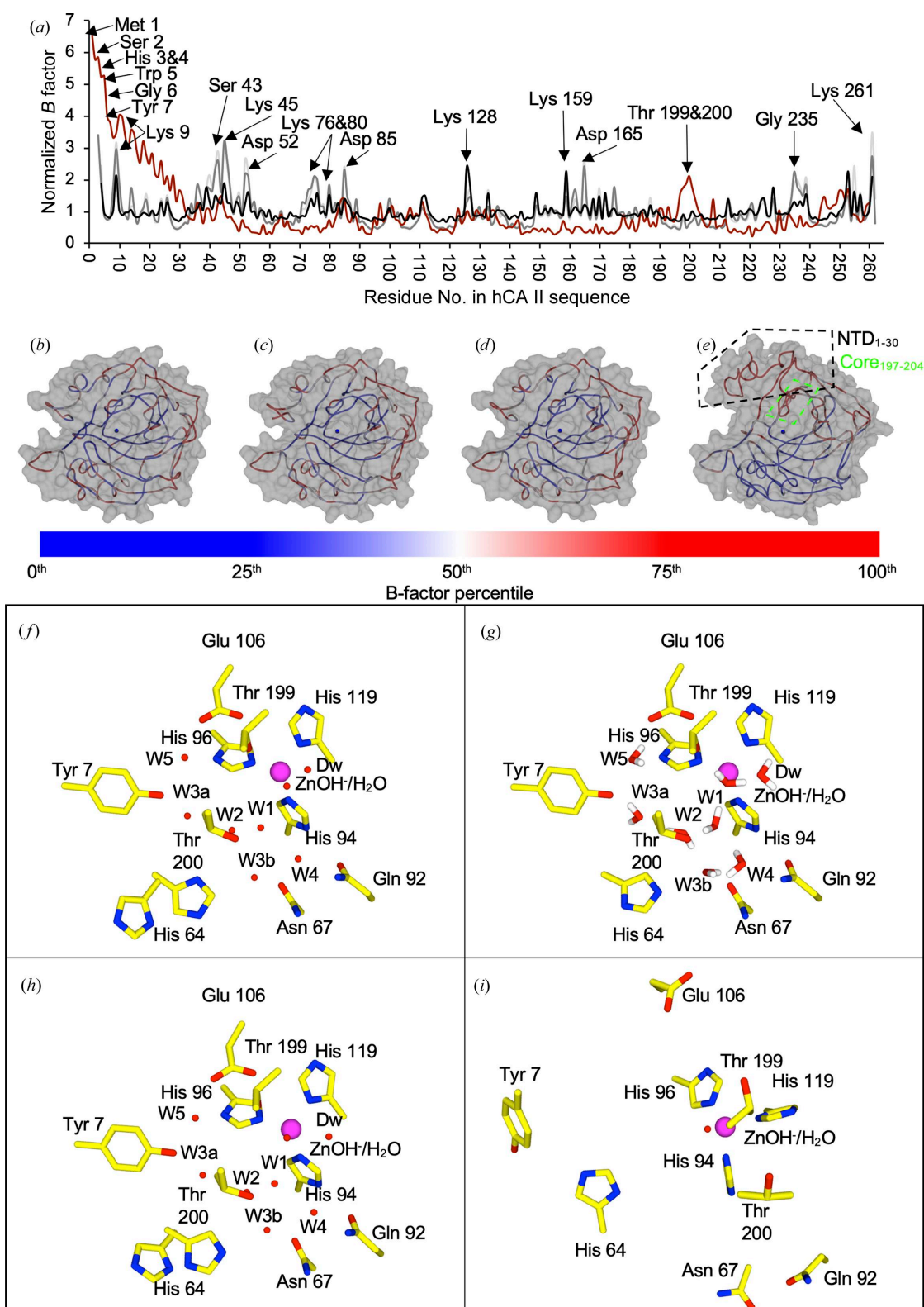

Figure 3
Comparison of per-atom B factors for XFEL, neutron diffraction, X-ray crystallographic and solid-state NMR structures. XFEL data are shown in black, the neutron diffraction structure (PDB entry 3kkx) is in gray, the crystal structure (PDB entry 3ks3) is in light gray and the solid-state NMR structure (PDB entry 6hd2) is in dark red. Standard error is shown for data where more than three unique entries are present. (a) B factors by atom are shown in ascending order of average B factor in the XFEL data. (b) B factors of C atoms are shown ordered by increasing distance from the carbonyl C atom. (c) Average B factors are shown per residue arranged by increasing B factor in the XFEL data.

Figure 4
Comparison of B factors for XFEL, neutron diffraction, X-ray crystallographic and solid-state NMR structures in a structural context. (a) Normalized B factor per residue. XFEL data are shown in black, the neutron diffraction structure (PDB entry 3kkx) is in gray, the crystal structure (PDB entry 3ks3) is in light gray and the solid-state NMR structure (PDB entry 6hd2) is in dark red. XFEL (b), neutron diffraction (c), X-ray crystallographic (d) and NMR (model 1 as a representative) (e) structures are depicted as ribbon diagrams with the zinc shown. All three are colored by average B-factor percentile on the relative scale shown below. In (e) $NTD_{1–30}$ was divergent in models 1–20; this region and the adjacent residues (indicated as $Core_{197–204}$) in the structure are identified and have systemically higher B factors. The active sites of the XFEL (f), neutron diffraction (g), X-ray crystallographic (h) and solid-state NMR (model 1 as a representative) (i) structures are depicted and labeled. Waters are identified by the letter W, with deep water indicated by Dw. In (i) Tyr7, Thr199, and Thr200 were all part of the variable NTD or adjacent regions of the structure.

Table 2
Mean atomic *B* factors.

The error is the standard deviation from the mean.

Data set	Mean atomic *B* factor (Å^2)
A	22.7 ± 16.2
B	32.1 ± 19.8
C	27.4 ± 17.0
D	24.0 ± 17.1
E	23.9 ± 17.7
XFEL	49.2 ± 20.6
Neutron crystallography	25.9 ± 8.5
X-ray crystallography	15.0 ± 10.2
Solid-state NMR	2.6 ± 2.7

the *B* factors, similar to what has been previously reported (Lomb *et al.*, 2011).

3.2. Comparison of the XFEL *B* factor with those from the highest resolution NMR, synchrotron and neutron diffraction crystallographic structures

Structural studies have extensively been used to characterize CAII. Methods used to date include XFEL, neutron, synchrotron diffraction crystallography and NMR. To assess the quality of the XFEL structure, we compared it with the best available structures from each of these methods. We note that since there are large differences in *B* factors between XFEL, neutron diffraction, synchrotron and NMR structures (Table 2), for analysis all data sets were normalized to the mean *B*-factor value in the respective structures.

On a per-atom-type basis (Fig. 3*a*), the active-site zinc consistently has the lowest *B* factor in all structures, while all structures had oxygen with the highest *B* factor, apart from the NMR structure, which had sulfur with the highest *B* factor. When considering the C atoms extending out from the peptide backbone (Fig. 3*b*), their *B* factors generally increase until the ε carbon and then decrease, except for the NMR structure, in which this decrease was not observed. For the XFEL, neutron and synchrotron crystallographic structures, this is likely to be a result of the fact that only phenylalanine, tyrosine and tryptophan have ζ and η carbons, and these amino acids are overwhelmingly biased towards the protein core, with relatively low *B* factors. By amino acid (Fig. 3*c*), hydrophobic residues expected near the protein core generally tend to have lower *B* factors, while charged residues expected near the protein exterior are biased towards higher *B* factors, with the exception again being the NMR structure, where methionine, tryptophan, tyrosine and histidine have some of the highest *B* factors. Of the four structures, the NMR structure is the least comparable to those obtained using the other three diffraction techniques when analyzed by atom, by side-chain radial distance or on a per-residue basis.

Within the context of the primary sequence, the XFEL, neutron diffraction and synchrotron crystallographic structures exhibit *B* factors that consistently peak at similar amino acids in the primary sequence, while the values for the NMR structure again deviate (Fig. 4*a*). Within the tertiary structure, the XFEL, neutron diffraction and synchrotron crystallo-

graphic structures exhibit lower *B* factors in the protein interior and higher *B* factors in the protein exterior in regions analogous to each other (Figs. 4*b*–4*d*). By contrast, in the NMR structure the deviations are mostly observed in NTD$_{1-30}$. This region is near the surface of CAII and many of these residues also had elevated *B* factors in the other structures, but the extent of elevation was considerably higher in the NMR structure. Unlike the three other structures, the NMR structure starts at residue 1 (Met), while the others start at residue 3 or 4 (both His). Normalization of the NMR *B* factors to the residue mean demonstrates that residue 1 has a sevenfold higher *B* factor than the mean, and the following ∼30 amino acids, along with the core residues 197–204, which are proximal in the tertiary structure, possess the majority of normalized *B* factors of >1 in the NMR structure. We previously noted elevated *B* factors for S atoms, residues at extended atomic distances from the backbone and for methionine, tryptophan, tyrosine and histidine (Fig. 3). The first seven amino acids of the sequence are enriched in these residues and atoms, which is likely to explain the previously observed deviations. Therefore, a general observation for the NMR structure is that the deviation per residue appears to be much higher, but that the variation is localized to a single region of the tertiary structure (Fig. 4*e*). This is likely to imply that NMR is more sensitive to domain flexibility than the crystallographic methods assessed here. This is in agreement with past reports that the first 24 amino acids can both fold and unfold independently and can be deleted with apparent functional conservation of CAII (Aronsson *et al.*, 1995).

Within the active site, the structure remains largely invariant, with the XFEL, neutron and synchrotron crystallographic structures having nearly identical amino-acid placements (Figs. 4*f*–4*h*). Waters possess H atoms in the neutron structure due to the unique features of this technique, allowing study of the exact placement and hydrogen-bonding assignment within the active site. The placements of residues 7, 199 and 200 in the NMR structure deviate and these residues are located in regions that were previously noted to have systemically high *B* factors (∼3.7-fold, 1.7-fold and 2.1-fold higher than the total mean *B* factor, respectively). Waters were not built into the NMR structure due to methodological limitations. Overall, the XFEL method appears to produce results that are analogous to synchrotron crystallography but with systemically higher *B* factors and slightly lower quality resolution. However, the decrease in quality in the XFEL structure does not seem to impede accurate model determination or building of the protein structure, and therefore it is an excellent tool for specialized crystallographic techniques where time-resolved questions can be studied.

References

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. É., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. & Yamashita, K. (2023). *Acta Cryst.* D**79**, 449–461.

Aronsson, G., Mårtensson, L. G., Carlsson, U. & Jonsson, B.-H. (1995). *Biochemistry*, **34**, 2153–2162.

Avvaru, B. S., Kim, C. U., Sippel, K. H., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N. & McKenna, R. (2010). *Biochemistry*, **49**, 249–251.

Barker, A. I., Southworth-Davies, R. J., Paithankar, K. S., Carmichael, I. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 205–216.

de la Mora, E., Coquelle, N., Bury, C. S., Rosenthal, M., Holton, J. M., Carmichael, I., Garman, E. F., Burghammer, M., Colletier, J. P. & Weik, M. (2020). *Proc. Natl Acad. Sci. USA*, **117**, 4142–4151.

Díaz-Torres, N. A., Mahon, B. P., Boone, C. D., Pinard, M. A., Tu, C., Ng, R., Agbandje-McKenna, M., Silverman, D., Scott, K. & McKenna, R. (2015). *Acta Cryst.* D**71**, 1745–1756.

Domsic, J. F., Avvaru, B. S., Kim, C. U., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N. & McKenna, R. (2008). *J. Biol. Chem.* **283**, 30766–30771.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.

Fischer, M. (2021). *Q. Rev. Biophys.* **54**, e1.

Fisher, Z., Kovalevsky, A. Y., Mustyakimov, M., Silverman, D. N., McKenna, R. & Langan, P. (2011). *Biochemistry*, **50**, 9421–9423.

Frost, S. & McKenna, R. (2014). Editors. *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*. Dordrecht: Springer.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kannan, K. K., Notstrand, B., Fridborg, K., Lövgren, S., Ohlsson, A. & Petef, M. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 51–55.

Kim, J. K., Lee, C., Lim, S. W., Adhikari, A., Andring, J. T., McKenna, R., Ghim, C.-M. & Kim, C. U. (2020). *Nat. Commun.* **11**, 4557.

Kim, J. K., Lomelino, C. L., Avvaru, B. S., Mahon, B. P., McKenna, R., Park, S. & Kim, C. U. (2018). *IUCrJ*, **5**, 93–102.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.

Lomb, L., Barends, T. R. M., Kassemeyer, S., Aquila, A., Epp, S. W., Erk, B., Foucar, L., Hartmann, R., Rudek, B., Rolles, D., Rudenko, A., Shoeman, R. L., Andreasson, J., Bajt, S., Barthelmess, M., Barty, A., Bogan, M. J., Bostedt, C., Bozek, J. D., Caleman, C., Coffee, R., Coppola, N., DePonte, D. P., Doak, R. B., Ekeberg, T., Fleckenstein, H., Fromme, P., Gebhardt, M., Graafsma, H., Gumprecht, L., Hampton, C. Y., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Holton, J. M., Hunter, M. S., Kabsch, W., Kimmel, N., Kirian, R. A., Liang, M., Maia, F. R. N. C., Meinhart, A., Marchesini, S., Martin, A. V., Nass, K., Reich, C., Schulz, J., Seibert, M. M., Sierra, R., Soltau, H., Spence, J. C. H., Steinbrener, J., Stellato, F., Stern, S., Timneanu, N., Wang, X., Weidenspointner, G., Weierstall, U., White, T. A., Wunderer, C., Chapman, H. N., Ullrich, J., Strüder, L. & Schlichting, I. (2011). *Phys. Rev. B*, **84**, 214111.

Lomelino, C. L., Kim, J. K., Lee, C., Lim, S. W., Andring, J. T., Mahon, B. P., Chung, M., Kim, C. U. & McKenna, R. (2018). *Acta Cryst.* F**74**, 327–330.

Maupin, C. M., McKenna, R., Silverman, D. N. & Voth, G. A. (2009). *J. Am. Chem. Soc.* **131**, 7598–7608.

Meldrum, N. U. & Roughton, F. J. W. (1933). *J. Physiol.* **80**, 113–142.

Nass, K. (2019). *Acta Cryst.* D**75**, 211–218.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.

Pinard, M. A., Boone, C. D., Rife, B. D., Supuran, C. T. & McKenna, R. (2013). *Bioorg. Med. Chem.* **21**, 7210–7215.

Reinknecht, C., Riga, A., Rivera, J. & Snyder, D. A. (2021). *Molecules*, **26**, 1484.

Silverman, D. N. & Lindskog, S. (1988). *Acc. Chem. Res.* **21**, 30–36.

Singh, H., Vasa, S. K., Jangra, H., Rovó, P., Päslack, C., Das, C. K., Zipse, H., Schäfer, L. V. & Linser, R. (2019). *J. Am. Chem. Soc.* **141**, 19276–19288.

Sippel, K. H., Robbins, A. H., Domsic, J., Genis, C., Agbandje-McKenna, M. & McKenna, R. (2009). *Acta Cryst.* F**65**, 992–995.

Stadie, W. C. & O'Brien, H. (1933). *J. Biol. Chem.* **103**, 521–529.

Tanhauser, S. M., Jewell, D. A., Tu, C. K., Silverman, D. N. & Laipis, P. J. (1992). *Gene*, **117**, 113–117.

Warkentin, M., Badeau, R., Hopkins, J. B., Mulichak, A. M., Keefe, L. J. & Thorne, R. E. (2012). *Acta Cryst.* D**68**, 124–133.

White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 680–689.

Yu, C.-J., Lee, H. C., Kim, C., Cha, W., Carnis, J., Kim, Y., Noh, D. Y. & Kim, H. (2014). *J. Synchrotron Rad.* **21**, 264–267.