Supplemental Information

Structure of the Angiotensin Receptor Revealed by Serial Femtosecond Crystallography

Haitao Zhang,[1] Hamiyet Unal,[2] Cornelius Gati,[3] Gye Won Han,[4] Wei Liu,[5] Nadia A. Zatsepin,[6] Daniel James,[6] Dingjie Wang,[6] Garrett Nelson,[6] Uwe Weierstall,[6] Michael R. Sawaya,[7] Qingping Xu,[8] Marc Messerschmidt,[9] Garth J. Williams,[10] Sébastien Boutet,[10] Oleksandr M. Yefanov,[3] Thomas A. White,[3] Chong Wang,[11] Andrii Ishchenko,[4] Kalyan C. Tirupula,[2] Russell Desnoyer,[2] Jesse Coe,[5] Chelsie E. Conrad,[5] Petra Fromme,[5] Raymond C. Stevens,[1,4,12] Vsevolod Katritch,[1] Sadashiva S. Karnik,[2] and Vadim Cherezov[4,*]

[1]Department of Biological Sciences, Bridge Institute, University of Southern California, Los Angeles, CA 90089, USA

[2]Department of Molecular Cardiology, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA

[3]Center for Free Electron Laser Science, Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany

[4]Department of Chemistry, Bridge Institute, University of Southern California, Los Angeles, CA 90089 USA

[5]Department of Chemistry and Biochemistry, Center for Applied Structural Discovery at the Biodesign Institute, Arizona State University, Tempe, AZ 85287, USA

[6]Department of Physics, Arizona State University, Tempe, AZ 85287, USA

[7]Department of Chemistry and Biochemistry, the UCLA-DOE Institute for Genomics and Proteomics, University of California, Los Angeles, CA 90095, USA

[8]Joint Center for Structural Genomics, Stanford Synchrotron Radiation Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

[9]BioXFEL Science and Technology Center, Buffalo, NY 14203, USA

[10]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

[11]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

[12]iHuman Institute, ShanghaiTech University, Shanghai, 201210 China

***Correspondence:** cherezov@usc.edu (V.C.)

EXTENDED EXPERIMENTAL PROCEDURES

Protein engineering for structural studies

The sequence of the human AT_1R gene was optimized for insect cells expression system and synthesized by GenScript. A thermostabilized apocytochrome *b*562 RIL (BRIL) from *E. coli* (M7W, H102I, R106L) was fused to the N-terminus of the human AT_1R, using overlapping PCR. The construct has truncations of the AT_1R residues 1, 7-16 and 320-359. The resulting BRIL-AT_1R chimera sequence was subcloned into a modified pFastBac1 vector (Invitrogen), which contains a haemagglutinin (HA) signal sequence, a FLAG tag and 10×His tag, followed by a tobacco etch virus (TEV) protease cleavage site, before the N-terminus of the chimera sequence.

Protein expression and purification

The BRIL-AT_1R protein was expressed in *Spodoptera frugiperda* (Sf9) insect cells using the Bac-to-Bac baculovirus expression system (Invitrogen). Cells with a density of $2\text{-}3\times10^6$ cells per ml were infected with baculovirus at 27 °C, and at 48 hours after infection, cells were collected by centrifugation and stored at -80 °C until use.

Cells were thawed and lysed by repeated washing and centrifugation in the hypotonic buffer of 10 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 20 mM KCl, and the high osmotic buffer of 10 mM HEPES, pH 7.5, 1.0 M NaCl, 10 mM $MgCl_2$, 20 mM KCl, with EDTA-free complete protease inhibitor cocktail tablets (Roche). The washed membranes were suspended into the hypotonic buffer with 100 μM ZD7155 (Tocris Bioscience), and incubated at 4 °C for 1 hour. The membranes were then solubilized in the buffer containing 50 mM HEPES, pH 7.5, 500 mM NaCl, 1% (w/v) n-dodecyl-beta-D-maltopyranoside (DDM, Anatrace) and 0.2% (w/v) cholesterol hemisucinate (CHS, Sigma-Aldrich), and 20% (v/v) glycerol, at 4 °C for 4 hours. The supernatants containing the solubilized AT_1R proteins were isolated by high-speed centrifugation, and then incubated with TALON IMAC resin (Clontech) and 20 mM imidazole, at 4 °C overnight. The resin was washed with 10 column volumes of washing buffer I containing 50 mM HEPES, pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 0.1% (w/v) DDM, 0.02% (w/v) CHS, 20 mM imidazole and 20 μM ZD7155, and 10 column volumes of washing buffer II containing 50 mM HEPES, pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 50 mM imidazole and 20 μM ZD7155. The proteins were eluted by 3 column volumes of eluting buffer containing 50 mM HEPES, pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 0.02% (w/v) DDM, 0.004% (w/v) CHS, 300 mM imidazole and 100 μM ZD7155. PD MiniTrap G-25 column (GE Healthcare) was used to remove imidazole. The protein was then treated overnight with His-tagged TEV protease to cleave the N-terminal FLAG/His tags from the proteins. The cleaved FLAG/His tags and TEV protease were removed by TALON IMAC resin. The protein was not treated with PNGase F and therefore remained fully glycosylated. Finally, the purified protein was concentrated to 30 mg/ml with a 100 kDa cutoff concentrator (Vivaspin) and used in crystallization trials. The protein yield and monodispersity were tested by analytical size exclusion chromatography (aSEC).

Lipidic cubic phase crystallization

For the initial crystallization setup, purified BRIL-AT_1R in complex with ZD7155 was reconstituted into lipidic cubic phase (LCP) by mixing with molten lipid (90% (w/w) monoolein and 10% (w/w) cholesterol) at a protein/lipid ratio of 2:3 (v/v) using a mechanical syringe mixer (Caffrey and Cherezov, 2009). LCP crystallization trials were performed using an NT8-LCP crystallization robot (Formulatrix). 96-well glass sandwich plates (Marienfeld) were incubated and imaged at 20 °C using an automatic incubator/imager (RockImager 1000, Formulatrix). The crystals grew in the condition of 100 mM sodium citrate, pH 5.0-6.0, 300-600 mM $NH_4H_2PO_4$, 20-30% (v/v) PEG400 and 2-8% (v/v) DMSO. The crystals were harvested using micromounts (MiTeGen) and flash-frozen in liquid nitrogen for data collection at a synchrotron source. However, these crystals only diffracted to about 4 Å resolution, even after extensive optimization of crystallization conditions.

Microcrystals for SFX data collection at an XFEL source were prepared in syringes (Hamilton) as previously described (Liu et al., 2014b). Briefly, 5 µl protein-laden LCP aliquots were injected into 100 µL gas-tight syringes filled with 60 µl precipitant solution (100 mM sodium citrate, pH 5.0, 450 mM $NH_4H_2PO_4$, 28% (v/v) PEG400 and 4% (v/v) DMSO). The syringes were incubated at 20 °C for 3 days. Excess precipitant solution was carefully removed from the syringes, and the remaining LCP with embedded crystals was consolidated together. Then 3 µL of 7.9 MAG was added and mixed with LCP, to absorb the residual precipitant solution and prevent formation of lamellar crystalline phase due to rapid evaporative cooling when injecting LCP into vacuum (Weierstall et al., 2014).

X-ray free electron laser data collection

Data collection was performed at the Coherent X-ray Imaging (CXI) end station of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, using a micro-focus setup (Boutet and Williams, 2010). The X-ray beam was focused to a size of 1.5×1.5 μm^2 full width at half maximum using X-ray optics in Kirkpatrick-Baez geometry. The XFEL pulse duration was estimated to be 36 fs with a repetition rate of 120 Hz. A photon energy of 7.9 keV, an average pulse energy of 2.7 mJ and a transmission level of 16% resulted in a maximum dose of 75 MGy at the sample.

A batch of the sample (typically 20 µl) was manually loaded into the LCP injection system (Weierstall et al., 2014). By applying a pressure of ~70 psi using an HPLC system (corresponding to a pressure of ~2,400 psi on the sample due to the pressure amplification in the injector), the crystal-containing LCP matrix was delivered into the interaction region through a 50 µm diameter nozzle with a flow rate of roughly 170 nl per minute. Diffraction patterns were collected on a Cornell-SLAC Pixel array detector (CSPAD - version 1.5) (Hart et al., 2012) at a rate of 120 Hz.

With a total sample volume of 65 µl, a total of 2,764,739 diffraction frames were collected within 6.4 hours. Initial frames were corrected and filtered using the software package Cheetah

(Barty et al., 2014). A crystal 'hit' was defined as an image containing a minimum of 15 peaks (with a signal to noise ratio above 4) per image. A total of 457,275 positive 'hits' (average hit rate 17%) were further processed using the CrystFEL software suite (version 0.5.3) (White et al., 2013; White et al., 2012).

The detector geometry was refined using an automated algorithm designed to match found and predicted peaks to sub-pixel accuracy. By further refinement of parameters (peak detection, prediction and integration), a total of 73,130 images were indexed (indexing rate 16%), integrated and merged into a final dataset. To reduce noise and outliers and thus improve data quality we have applied two data rejection criteria: 1) per pattern resolution cutoff, and 2) rejection of patterns based on a Pearson correlation coefficient threshold. A conservative resolution limit for each crystal was estimated using the spots found by the initial peak search, by taking the 98th percentile of the scattering angles of the peaks which could be explained by the lattice for the crystal. A spot was considered to be explained by the lattice if its calculated Miller indices, assuming a thin Ewald sphere, were all within 0.25 of integers. Predicted spots were integrated for each crystal up to a maximum resolution 0.12 $Å^{-1}$ higher than this conservative limit. An initial merging pass was performed without scaling factors and including all crystals. A second pass was then performed, in which crystals were rejected if the Pearson correlation coefficients between their intensities and the merged data set were less than 0.2. The final resolution cutoff was estimated to be 2.9 Å using a combination of *CC** (Karplus and Diederichs, 2012) and other parameters (**Figure S3**). The final dataset had overall R_{split}=9.8%, and *CC**=0.872 in the highest resolution shell.

Structure determination

Initial attempts to find a molecular replacement (MR) solution with known class A GPCR structures as search models using Phaser (McCoy et al., 2007) did not generate any reliable results. Subsequently, we utilized a systematic MR approach to locate the receptor and BRIL, similar to the method that we used to solve the human glucagon class B GPCR (Siu et al., 2013). First, we selected six GPCR structures with the highest sequence homology to AT_1R using HHpred (Soding et al., 2005). Further, for each structural model, a hybrid model was generated using phenix.mr_model_preparation (Adams et al., 2010). All hybrid models were then superimposed to each other and the poorly structurally conserved regions were manually trimmed. Either an original or trimmed hybrid model, together with the BRIL structure (Protein Data Bank (PDB) code: 1M6T), were used as two independent templates for MR using Phaser (McCoy et al., 2007). A script running on a Linux cluster was used to perform multiple MR jobs at different resolution cut-offs and other parameters, followed by rigid body refinement using Refmac5 (Murshudov et al., 1997), and restrained refinement using Refmac5 and autoBUSTER (Bricogne et al., 2009). Overall, 180 MR jobs were executed and resulted in sorted solutions by TFZ scores and R factors, with a maximum TFZ score of 9.0 (R/R_{free}=0.320/0.369). The electron density maps unambiguously showed new interpretable features that were not present in the search models, indicating the correct MR solutions were found. The solution with the lowest R_{free} value in autoBUSTER (R/R_{free}=0.316/0.345), corresponding to a search model based on the

trimmed hybrid model of CCR5 (PDB code: 4MBS), was used for further refinement. Refinements and model completion were performed by repetitive cycling between Refmac5 and autoBUSTER, followed by manual examination and rebuilding of the refined coordinates in Coot (Emsley et al., 2010), using both 2mFo-DFc and mFo-DFc maps, as well as omit maps calculated using Bhat's procedure (Bhat, 1988). During the refinement, data processing was further optimized to improve the electron densities for both the protein and ligand. The following residues had non-interpretable electron density and were not included in the model: Glu173-Asn176 and Ser186-Ser189 in ECL2, Ala225-Arg234 in ICL3 and Lys318-Tyr319 in C-terminus. The extracellular part of helix V (residues Ile193-Lys199) contains two Glycines and apparently highly dynamic. This part of the structure has high B-factors and low electron density. The data collection and refinement statistics are shown in **Table S1**.

Docking of ARBs into AT_1R ligand-binding pocket

Representative ARBs were docked into the AT_1R crystal structure using an energy-based docking protocol implemented in ICM molecular modeling software suite (Molsoft). The initial receptor docking model was generated by adding missing side chains and hydrogen atoms, and optimizing their conformations, followed by generation of soft potential maps of the receptor in a large box (30×30×30 $Å^3$) covering the extracellular half of the receptor. Molecular models of compounds were generated from two-dimensional representations and their 3D geometry was optimized using MMFF-94 force field (Halgren, 1995). Molecular docking employed biased probability Monte Carlo (BPMC) optimization of the ligand internal coordinates in the grid potentials of the receptor (Totrov and Abagyan, 1997). To assure convergence of the docking procedure, at least five independent docking runs were performed for each ligand starting from a random conformation; Monte Carlo sampling and optimization was performed at high thoroughness set to 30. The objective energy function included the ligand internal strain and a weighted sum of the grid map values in ligand atom centers. Note that the $Lys199^{5.42}$ side chain atoms lack electron density and are missing in the crystal structure of AT_1R-ZD7155 complex. Therefore we treated the $Lys199^{5.42}$ side chain as an explicit flexible group in receptor model, allowing comprehensive sampling of the side chain rotamers with each Monte Carlo iteration of the docking procedure. Up to 30 alternative complex conformations of the ligand-receptor complex were generated and rescored using ICM Binding Score function (Bursulaya et al., 2003; Schapira et al., 1999) calculated as:

$$S_{bind} = E_{int} + T\Delta S_{Tor} + E_{vw} + \alpha_1 \times E_{el} + \alpha_2 \times E_{hb} + \alpha_3 \times E_{hp} + \alpha_4 \times E_{sf}$$

where Evw, Eel, Ehb, Ehp, and Esf are Van der Waals, electrostatic, hydrogen bonding, non-polar, and polar atom solvation energy differences between bound and unbound states, Eint is the ligand internal strain, ΔSTor is its conformational entropy loss upon binding, T = 300 K, and αi are ligand- and receptor-independent constants. The results of individual docking runs for each ligand were considered consistent if at least three of the five docking runs produced similar ligand conformations (RMSD < 2.0 Å) and Binding Score < -20.0 kJ/mol. All calculations were performed using twelve core Intel Xeon 2.67Ghz workstation running Linux Fedora OS, taking about 10 min per ligand. The unbiased docking procedure did not use distance restraints or any

other *a priori* derived information for the ligand-receptor interactions.

Construction of AT_1R mutants and cell transfection for ligand-binding assays

The complementary DNA (cDNA) encoding the human AT_1R with N-terminal HA-tag was originally cloned into the expression vector pMT3 at the EcoRI and NotI sites. The single mutants were constructed by a PCR-based site directed mutagenesis strategy as previously described (Unal et al., 2010). COS-1 cells were grown in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher Scientific) and 100 IU penicillin/streptomycin (Sigma-Aldrich). Cells were seeded onto poly-D-lysine treated cell culture plates at a density of 3×10^6 cells per 10 cm diameter plate. After overnight culture, the cells were transiently transfected with wild-type or mutated AT_1R DNA using FuGENE 6 transfection reagent (Roche Applied Science) according to the manufacturer's instructions.

Membrane preparation for ligand-binding assays

Ligand binding was analyzed using total membranes prepared from COS-1 transiently expressing wild-type HA-AT_1R, ΔBRIL-AT_1R, and BRIL-AT_1R constructs. Transfected/infected cells were harvested in osmotic lysis buffer (25 mM Tris-HCl, pH 7.5 and 5 mM EDTA, pH 8.0) with protease inhibitor cocktail (Sigma-Aldrich) homogenized by a dounce homogenizer. The homogenate was incubated by rotating for 10 min at 4°C and centrifuged for 5 min at 200×g. The supernatant was then centrifuged at 37,000×g for 30 min at 4°C. The precipitate containing the total membranes was suspended in membrane binding buffer (140 mM NaCl, 5.4 mM KCl, 1 mM EDTA, 0.006% BSA, 25 mM HEPES, pH 7.4). Protein concentration was determined by Bio-Rad Protein Assay (Bio-Rad). For both saturation and competition binding assays, 10 µg of homogenous cell membrane was used per well as described above.

^{125}I-[Sar^{1}Ile^8] AngII binding assay. Saturation binding assays were performed under equilibrium conditions, with ^{125}I-[Sar^{1}Ile^8] Ang II (Dr. Robert Speth, Nova Southeastern University) concentrations ranging between 0.125 and 12 nM (specific activity, 2176 Ci/mmol) as triplicates in 96-well plates for 1 h at room temperature as previously described (Unal et al, 2013). Nonspecific binding was measured in the presence of 10^{-5} M ^{125}I-[Sar^{1}Ile^8] AngII (Bachem).Competition binding assays were performed under equilibrium conditions, with 2 nM ^{125}I-[Sar^{1}Ile^8] Ang II and concentrations of the competing ligand (AngII, Losartan, Candesartan, TRV120027) ranging between 0.5 nM and 1000 nM. The cells were harvested by filtering the binding mixture through Whatman GF/C glass fiber filters (102×256 mm), which were extensively washed with washing buffer (20 mM sodium phosphate, 100 mM NaCl, 10 mM MgCl2, 1 mM EGTA, pH 7.2). The bound ligand fraction was determined as the cpm (counts per minute) using a scintillation counter (MicroBeta2 Plate Counter, PerkinElmer). The binding kinetics were analyzed by nonlinear curve-fitting program LigandR, which yields the mean ± S.D. for the *Kd* and *B*max values.

^{3}H-candesartan binding assays

Saturation binding assays with ^{3}H-candesartan were performed under equilibrium conditions, with ^{3}H-candesartan (Amersham Pharmacia Biotech) concentrations ranging between 0.125 and 12 nM (specific activity, 16 Ci/mmol) as duplicates in 96-well plates for 1h at room temperature. Nonspecific binding was measured in the presence of 10 µM candesartan (gift from AstraZeneca). The cells were harvested by filtering the binding mixture through Whatman GF/C glass fiber filters (102×256 mm), which were extensively washed with washing buffer (20 mM sodium phosphate, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA, pH 7.2). The filter membranes were soaked with 5 ml of Ecoscint A scintillation fluid (National Diagnostics) and incubated for 1h at room temperature. The bound ligand fraction was determined as the dpm (disintegrations per minute) using Beckman LS 6000 Liquid Scintillation Counter (Global Medical Instrumentation). The binding kinetics was analyzed by nonlinear curve-fitting program GraphPad Prism 5, which yields the mean ± S.D. for the K_d and B_{max} values.

ZD7155 competition binding assays

Competition binding assays were performed under equilibrium conditions, with 2 nM ^{3}H-candesartan and various concentrations of the ZD7155 ranging between 0.04 and 1000 nM. The binding kinetics was analyzed by nonlinear curve-fitting program GraphPad Prism 5, which yields the mean ± S.D. for the IC_{50} values.

FLIPR-based intracellular calcium levels in cells

Calcium levels were measured using **a** Fluorescent Imaging Plate Reader (FLIPR®) Calcium 5 assay kit (Molecular Devices). One day before the calcium experiments, wild-type and mutated AT_1R transfected cells were seeded at a density of 100,000 cells/well in 100 µl medium onto a 96-well clear bottom black cell culture plate that was pre-coated with poly-L-lysine. The plate was maintained in a cell culture incubator for 26–28 h. On the day of experiment, cells were initially serum-starved for 2h by replacing the medium with 80 µl of serum-free DMEM. Following serum starvation, 100 µl of calcium-sensitive dye along with 2× Probenecid (2.5 mM final concentration, Life Technologies) was added to the cells. During this step, for antagonist dose-response curves, 20 µl of various concentrations of ZD7155 from a 10× stock prepared in D-PBS (1.47 mM KH_2PO_4, 138 mM NaCl, 2.67 mM KCl, 8.1 mM Na_2HPO_4, pH 7.3) were added to the cells. For all other wells, 20 µl of D-PBS was added. The cells were maintained for half an hour in the cell culture incubator and another half an hour at room temperature. During the incubation, AngII at 5× the desired final concentration in D-PBS were prepared in a U-bottom 96-well plate. Both the cells and ligand containing 96-well plates were loaded on to a FlexStation 3 instrument (Molecular Devices). The instrument was programmed in FLEX mode to add ligands (50 µl at 5× concentration) to the cells and to monitor the fluorescence before and after adding the ligands. In case of agonist dose response, AngII was added at various concentrations. While, 100 nM AngII was added for antagonist dose-response curves wherein the cells were already pre-treated with desired concentrations of ZD7155 during the calcium dye loading step.

ELISA-based estimation of wild-type and mutated AT_1R in cells

24h after transfection, cells were split into 12-well plates at a density of 500,000 cells per well. After additional 24h, cells were washed twice with HBSS (Hank's Balanced Salt Solution: 0.44 mM KH_2PO_4, 0.34 mM Na_2HPO_4, 137 mM NaCl, 5.36 mM KCl, 1.26 mM $CaCl_2$, 0.81 mM $MgSO_4$, 0.5 mM $MgCl_2$, 4.17 mM $NaHCO_3$, 5.55 mM D-Glucose, pH 7.3) with 1% BSA (bovine serum albumin), and HA-hAT_1R were labeled with anti-HA antibody (1µg/ml, Sigma-Aldrich) in HBSS/1% BSA for 1h at 4°C. The cells were then washed twice with HBSS/1% BSA and cells were fixed with 4% paraformaldehyde in HBSS for 15 min. All steps prior to cell fixing were carried out on ice to prevent endocytosis of AT_1R during processing of the samples. The cells were washed twice with HBSS/0.5% BSA and then incubated for another 1h at room temperature in HBSS/0.5% BSA supplemented with horseradish peroxidase-conjugated anti-mouse IgG (1:1000, Sigma-Aldrich) secondary antibody. The cells were washed twice with HBSS/0.5% BSA. Finally, the cells were incubated with 400 µl per well of o-Phenylenediamine dihydrochloride substrate (0.4 mg/ml) prepared in 0.05 M phosphate-citrate buffer (pH 5.0), containing 0.03% sodium perborate (Sigma-Aldrich) for 10 min in the dark at room temperature. The reaction was stopped with 100 µl per well of 3 N HCl and the absorbance was read at 492 nm using an ELISA plate reader (Molecular Devices). For detection of total protein, 0.1% Triton X-100 was included in all the buffers to promote permeabilization of cell membrane and the cells were fixed with 4% paraformaldehyde before adding the primary antibody. Control experiments were performed with mock transfected cells and wild-type transfected cells to which the primary antibody was not added.

Statistical Analysis

Results were presented as mean±SEM. Changes in specific radio-labelled ligand binding and cell surface expression of AT_1R constructs were normalized to those measured with wild-type AT_1R control (100%). IC_{50} values in binding assays were determined by non-linear regression analysis using the Prism software (GraphPad Software).

SUPLEMENTAL FIGURES

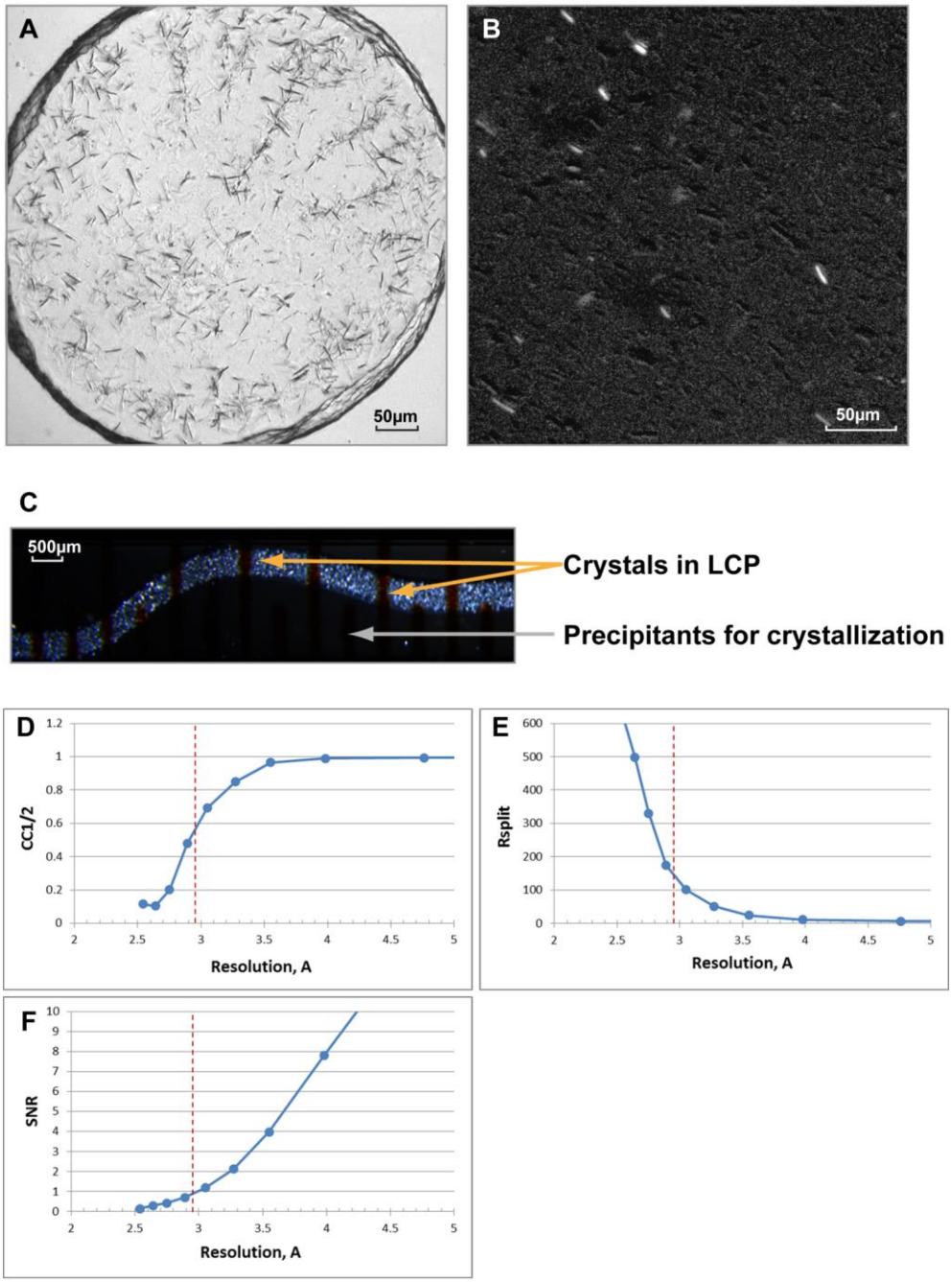

Figure S1. Crystals of AT_1R-ZD7155 and XFEL data processing statistics, Related to Figure 2.

(A) Crystals obtained in glass sandwich plates and used for synchrotron data collection.

(B), (C) Pictures of microcrystals used for LCP-SFX data collection, taken using cross-polarizers. (B) Sample is squeezed between two glass slides. (C) Sample is inside of a syringe.

(D) Pearson correlation coefficient calculated between two randomly split parts of the dataset, $CC_{1/2}$.

(E) R-factor calculated between two randomly split parts of the dataset, R_{split}.

(F) Signal-to-noise ratio, SNR, or I/σ(I), where σ(I) values were estimated as the standard deviations of the means of the intensity measurements (White et al, 2012). This calculation is different from the one used in traditional crystallography, and therefore the values of I/σ(I) reported here are not directly comparable to those reported for structures solved by traditional methods.

Red dash line in (D)-(F) indicates the applied resolution cut-off.

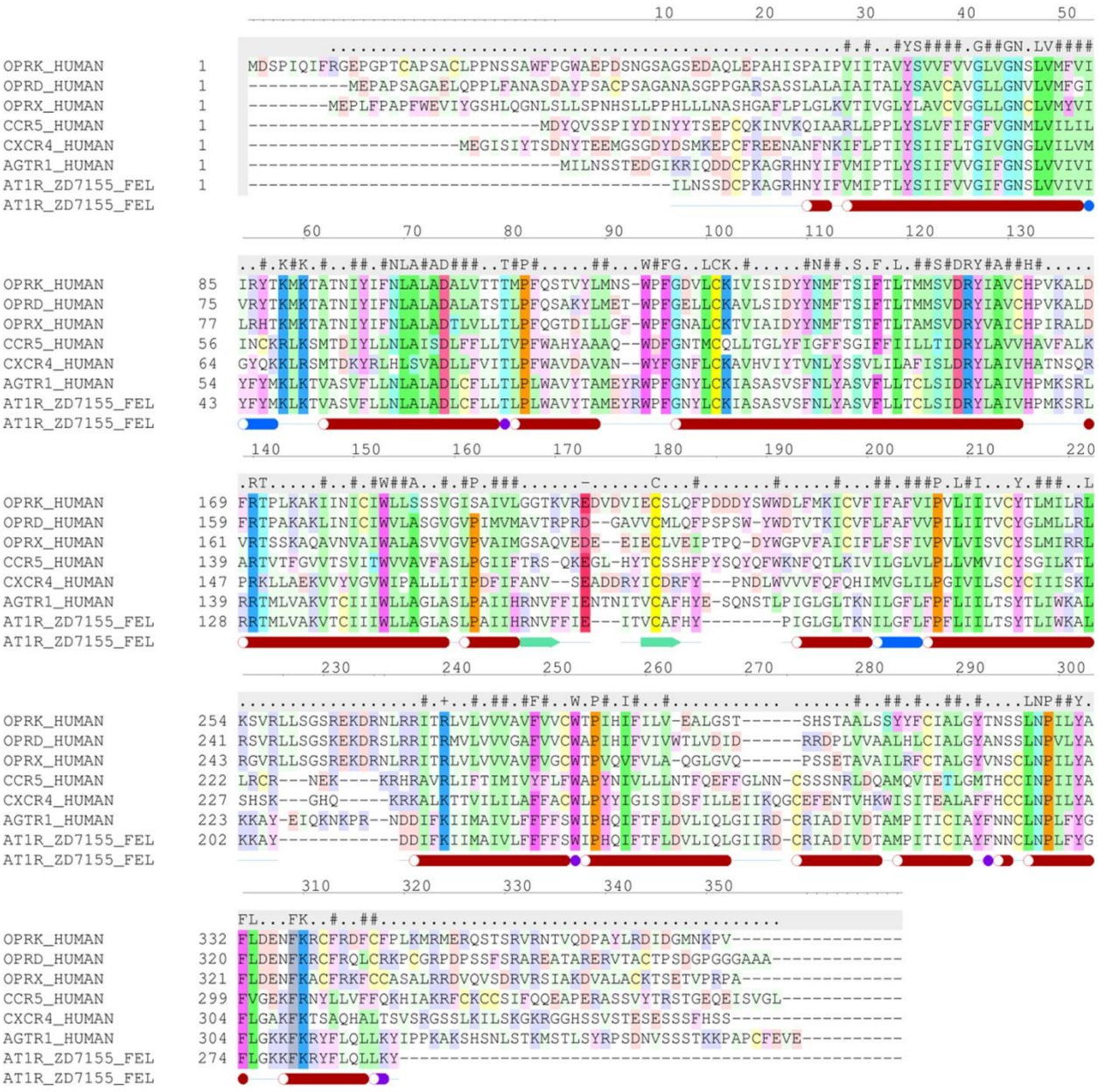

Figure S2. Structure-based sequence alignment of human AT_1R with chemokine and opioid receptors, Related to Figures 2 and 3.

Structure-based sequence alignment of human AT_1R (AGTR1) with CXCR4, CCR5, κ-OR (OPRK), δ-OR (OPRD), and NOP (OPRX) is shown, along with the secondary structure content of AT_1R_ZD7155 crystal structure displayed below alignment (red: α-helix, blue: π-helix, green: β-strand). Uncolored residues show lack of conservation, while coloring highlights conserved amino acids with specific functional properties (green: hydrophobic, blue: basic, red: acidic, magenta: aromatic, cyan: small polar, orange: proline, yellow: cysteine). The graphics were prepared using ICM molecular modeling package (Molsoft LLC).

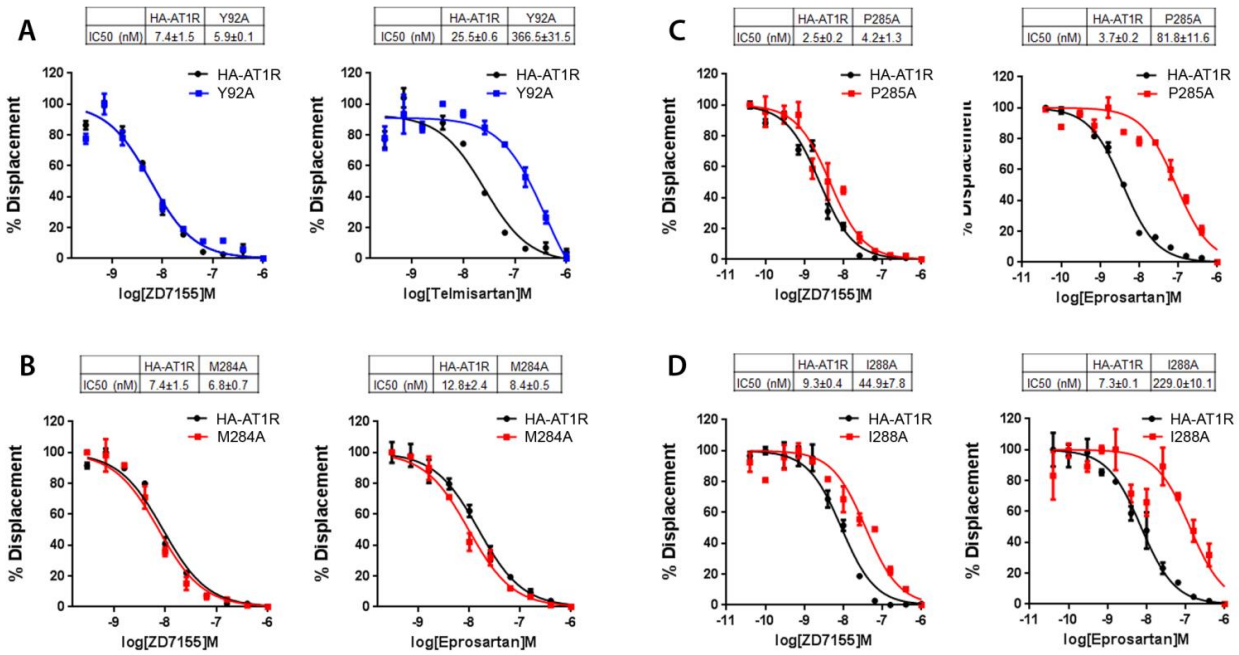

Figure S3. Experimental validation of docking results for telmisartan and eprosartan. Related to Figure 4.

(A) Displacement of ^{3}H-candesartan bound to the wild-type receptor (HA-AT$_1$R) and to a Tyr92^{ECL1}Ala mutant by competing ligands, ZD7155 and telmisartan. The binding affinities of ^{3}H-candesartan for the wild-type receptor, Kd = 2.4 ± 0.1 nM and Tyr92 ^{ECL1}Ala mutant , Kd = 4.4 ± 0.5 nM (data not shown).

(B) Displacement of ^{3}H-candesartan bound to the wild-type receptor (HA-AT$_1$R) and to a Met284$^{7.35}$Ala mutant by competing ligands, ZD7155 and eprosartan. The binding affinities of ^{3}H-candesartan for the wild-type receptor, Kd = 2.4 ± 0.1 nM and for Met284$^{7.35}$Ala mutant Kd = 3.1 ± 0.2 nM (data not shown).

(C) Displacement of ^{3}H-candesartan bound to the wild-type receptor (HA-AT$_1$R) and to a Pro285$^{7.36}$Ala mutant by competing ligands, ZD7155 and eprosartan. The binding affinities of ^{3}H-candesartan for the wild-type receptor, Kd = 2.4 ± 0.1 nM and for Pro285$^{7.36}$Ala mutant Kd = 4.9 ± 0.5 nM (data not shown).

(D) Displacement of ^{3}H-candesartan bound to the wild-type receptor (HA-AT$_1$R) and to a Ile288$^{7.39}$Ala mutant by competing ligands, ZD7155 and eprosartan. The binding affinities of ^{3}H-candesartan for the wild-type receptor, Kd = 2.4 ± 0.1 nM and for Ile288$^{7.39}$Ala mutant Kd = 4.5 ± 2.2 nM (data not shown).

Binding studies were performed using isolated membranes from transiently transfected COS-1 cells. Competition binding curves for non-peptide antagonist ZD7155 and telmisartan or

eprosartan in the presence of 2 nM ^{3}H-candesartan were generated and the corresponding mean ± SEM for IC50 values were measured. The curves from a representative experiment wherein measurements were made in triplicate are shown.

SUPPLEMENTAL TABLES

Table S1. Data collection and refinement statistics, Related to Figure 2.

AT_1R-ZD7155-XFEL	
Data collection	
Temperature (K)	294
Wavelength (Å)	1.56
Beam size (μm^2)	1.5×1.5
Average crystal size (μm^3)	10×2×2
Flux (ph/pulse) / Pulse duration (fs)	$1 \cdot 10^{11}$ / 36
Max dose per crystal (MGy)	75
Space group	C 1 2 1
Unit cell parameters *a*,*b*,*c* (Å); β (°)	72.8, 41.0, 167.7; 99.4
Number of collected frames	2,764,739
Number of hits / indexed images	457,275 / 73,130
Number of total / unique reflections	14,415,424 / 11,190
Resolution (Å) [a]	32.64 – 2.90 (3.00 - 2.90)
Completeness (%)	100.0 (100.0)
Multiplicity	1,288 (215)
$I/\sigma(I)$	8.2 (0.84)
CC^* [b]	0.999 (0.872)
R_{split} (%) [c]	9.8 (140)
Refinement	
Resolution (Å)	32.64 – 2.90
Number of reflections / test set	11,167 / 576
R_{work} / R_{free} (%)	22.8 / 27.4
Number of atoms	
Receptor / BRIL	3,077
Ligand	33
Wilson B-factors (Å^2)	76.1
Mean overall B value (Å^2)	
Receptor	90.0
BRIL	111.7
Ligand	61.0
R.m.s bonds (Å) / angles (°)	0.010/0.95
Ramachandran plot stats (%) [d]	
Favored	96.1
Allowed	3.9
Disallowed	0

[a] Highest resolution shell is shown in parentheses.

[b] $CC^* = \sqrt{\frac{2CC_{1/2}}{1+CC_{1/2}}}$

[c] $R_{split} = 1/\sqrt{2}\,\frac{\sum_{hkl}|I_{even} - I_{odd}|}{1/2\sum_{hkl}|I_{even} + I_{odd}|}$

[d] As defined in MolProbity (Chen et al., 2010).

Table S2. Ligand binding affinities of AT_1R mutants, Related to Figure 3.

Residue	Mutation	B&W number	Total expression	K_d (nM) ([Sar[1], Ile[8]]-AngII)	ΔKd (fold change)	K_d (nM) (cande sartan)	ΔKd (fold change)	K_i (nM) (ZD7155)	ΔKi (fold change)
	WT		**100.0 ± 3.7**	**5.6 ± 1.7**	**1**	**1.4 ± 0.2**	**1**	**3 ± 0.4**	**1**
Y35[a]	A	1.39	110.4 ± 6.6	NB	CBD	NB	CBD	CBD	CBD
D74	N	2.50	129.8 ± 4.1	13.9 ± 0.9	2.5 ± 0.8	not tested	not tested	7.6 ± 0.3	2.5 ± 0.4
D74	A	2.50	116.9 ± 3.3	9.7 ± 1.0	1.7 ± 0.5	not tested	not tested	0.6 ± 0.3	0.2 ± 0.1
W84[a]	A	2.60	116.1 ± 2.8	NB	CBD	NB	CBD	CBD	CBD
S109	A	3.33	112.7 ± 3.6	5.8 ± 0.7	1.0 ± 0.3	4.6 ± 1.7	3.3 ± 1.3	3.8 ± 0.5	1.3 ± 0.2
N111	G	3.35	109.1 ± 8.1	1.8 ± 0.1	0.3 ± 0.1	not tested	not tested	4.1 ± 1.8	1.4 ± 0.6
N111	A	3.35	139.2 ± 8.7	2.0 ± 0.3	0.4 ± 0.1	not tested	not tested	2.4 ± 1.4	0.8 ± 0.5
L112	A	3.36	103.2 ± 8.9	20.1 ± 1.9	3.6 ± 1.1	3.8 ± 0.7	2.7 ± 0.6	6.6 ± 1.3	2.2 ± 0.5
R167[a]	K	ECL2	128.6 ± 1.0	37.1 ± 1.7	6.6 ± 2.0	2.8 ± 1.1	2.0 ± 0.8	7.4 ± 1.2	2.5 ± 0.5
R167[a]	A	ECL2	95.9 ± 6.7	NB	CBD	>100	>70	CBD	CBD
F182[a,b]	A	ECL2	116.7 ± 8.4	>100	>20	3.0 ± 1.3	2.1 ± 1.0	CBD	CBD
K199[a]	Q	5.42	123.5 ± 5.4	NB	CBD	>100	>70	CBD	CBD
K199[a]	A	5.42	85.4 ± 3.1	17.6 ± 1.0	3.1 ± 1.0	13.5 ± 0.6	9.6 ± 1.4	15.9 ± 2	5.3 ± 1.0
N200	A	5.43	92.9 ± 2.9	8.1 ± 2.9	1.4 ± 0.7	not tested	not tested	2.9 ± 1.7	1.0 ± 0.6
Q257	A	6.52	113.3 ± 4.6	19.7 ± 1.9	3.5 ± 1.1	not tested	not tested	6.9 ± 0.7	2.3 ± 0.4
D281	A	7.32	98.7 ± 9.4	20.5 ± 0.8	3.7 ± 1.1	not tested	not tested	6.7 ± 0.6	2.2 ± 0.4
I288[a,b]	A	7.39	100.5 ± 3.3	NB	CBD	4.9 ± 2.2	3.5 ± 1.6	CBD	CBD
Y292	A	7.43	86.2 ± 4.2	10.6 ± 1.7	1.9 ± 0.6	6.4 ± 1.9	4.6 ± 1.5	2 ± 0.3	0.7± 0.1
N294	A	7.45	105.6 ± 2.0	1.7 ± 0.4	0.3 ± 0.1	not tested	not tested	3.1 ± 0.4	1.0 ± 0.2
N295	A	7.46	116.8 ± 2.7	15.7 ± 1.9	2.8 ± 0.9	not tested	not tested	4.4 ± 1.9	1.5 ± 0.7
N295	S	7.46	123.0 ± 1.1	12.7 ± 0.5	2.3 ± 0.7	not tested	not tested	2.7 ± 0.4	0.9 ± 0.2

NB: no binding
CBD: cannot be determined
Not tested: Because the residue is not in the binding pocket

[a] Size and chemical characteristics of these residues are critical for ligand binding.
[b] Residues that discriminate between peptide and non-peptide ligands (these mutants do not bind [Sar[1],Ile[8]]AngII but still bind candesartan).

Results presented as mean ± SEM of six replicates from two independent experiments, calculated using the Prism software (GraphPad Software).

Table S3. Interactions of ARBs with AT_1R determined by crystal structure and docking, Related to Figure 4.

Residue	ZD7155	candesartan	azilsartan	olmesartan	losartan	EXP3174	telmisartan	eprosartan	irbesartan	valsartan
$A21^{N\text{-term}}$	0	0	0	0	0	0	18	0	0	0
$R23^{N\text{-term}}$	0	0	0	0	0	0	1	0	0	0
$I31^{1.35}$	8	0	0	2	0	0	8	0	6	0
$Y35^{1.39}$	29	16	13	22	14	21	28	26	34	16
$F77^{2.53}$	8	9	12	13	9	12	7	14	12	9
$L81^{2.57}$	3	0	6	10	4	7	9	13	11	0
$W84^{2.60}$	46	46	42	48	45	40	35	42	47	40
$Y87^{2.63}$	16	15	12	16	18	14	15	16	10	18
$T88^{2.64}$	12	2	1	7	0	0	3	0	12	0
$Y92^{ECL1}$	4	7	8	9	8	6	43	4	8	7
$S105^{3.29}$	17	16	14	16	17	16	15	14	19	17
$V108^{3.32}$	43	44	43	43	45	42	42	34	43	42
$S109^{3.33}$	26	30	30	27	27	28	23	0	25	29
$L112^{3.36}$	12	15	16	13	12	14	11	0	10	12
$Y113^{3.37}$	2	3	3	3	2	2	1	0	0	4
$A163^{4.60}$	14	14	13	14	14	13	2	0	14	12
$R167^{ECL2}$	40	35	39	32	39	37	21	33	24	34
$C180^{ECL2}$	2	3	3	0	5	0	0	8	0	0
$F182^{ECL2}$	8	7	10	8	8	8	0	0	6	13
$K199^{5.42}$	11	27	13	8	10	13	21	0	4	12
$W253^{6.48}$	3	3	4	3	3	3	3	0	4	0
$H256^{6.51}$	0	0	0	0	0	0	0	4	0	0
$D281^{7.32}$	0	0	0	0	0	0	4	0	0	0
$M284^{7.35}$	0	0	0	0	0	0	3	14	0	0
$P285^{7.36}$	8	0	0	7	0	0	31	9	12	0
$I288^{7.39}$	42	30	23	30	33	35	43	55	50	48
$C289^{7.40}$	0	0	0	0	0	3	0	5	4	0
$Y292^{7.43}$	19	22	13	10	22	23	12	15	24	20

Numbers show the contact area (%) of the residue surface participating in the ligand interactions; the cells are colored according to the contact area, from blue (no direct contact) to red (maximum contact area).

SUPLEMENTAL REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr Section D, Biol Crystallogr *66*, 213-221.

Bhat, T.N. (1988). Calculation of an Omit Map. J Appl Crystallogr *21*, 279-281.

Boutet, S., and Williams, G.J. (2010). The Coherent X-ray Imaging (CXI) instrument at the Linac Coherent Light Source (LCLS). New J Phys *12*, 035024.

Bursulaya, B.D., Totrov, M., Abagyan, R., and Brooks, C.L., 3rd (2003). Comparative study of several algorithms for flexible ligand docking. J Comput Aided Mol Des *17*, 755-763.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr Section D, Biol Crystallogr *66*, 12-21.

Schapira, M., Totrov, M., and Abagyan, R. (1999). Prediction of the binding energy for small molecules, peptides and proteins. J Mol Recognit *12*, 177-190.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res *33*, W244-248.

White, T.A., Barty, A., Stellato, F., Holton, J.M., Kirian, R.A., Zatsepin, N.A., and Chapman, H.N. (2013). Crystallographic data processing for free-electron laser sources. Acta Crystallogr Section D, Biol Crystallogr *69*, 1231-1240.