SUPPORTING INFORMATION.

Joint X-ray/neutron crystallographic study of HIV-1 protease with clinical inhibitor amprenavir – insights for drug design;

Irene T. Weber, Mary Jo Waltman, Marat Mustyakimov, Matthew P. Blakeley, David A. Keen, Arun K. Ghosh, Paul Langan, and Andrey Y. Kovalevsky

Experimental section, Supporting Table 1, Supporting Figures 1 and 2.

Experimental Section.

General Information. APV was obtained through the NIH AIDS reagent program; it was dissolved in dimethylsulfoxide at the concentration of 20 mM. APV was supplied 99.7% pure as determined by HPLC (https://www.aidsreagent.org/pdfs/ds8148_002.pdf). Protein purification supplies were purchased from GE Healthcare (Piscataway, New Jersey, USA). Crystallization reagents were purchased from Hampton Research (Aliso Viejo, California, USA).

Protein expression and purification. PR used for the current study harbors five mutations Q7K, L33I, L63I, C67A and C95A to diminish autoproteolysis and prevent cysteine-thiol oxidation.[1] The expression and purification were performed in a similar way as described previously,[1,2], however instead of LB medium in H_2O we used the minimal medium made with 100 % D_2O and hydrogenous glucose ($C_6H_{12}O_6$) as the sole carbon source for *E. coli* bacteria. After induction with IPTG the cells were allowed to express PR for 6-8 hours before harvesting. PR was isolated and purified from inclusion bodies in H_2O buffers. The refolding (25mM formic acid) and buffer exchange (50 mM NaOAc, pH 5.0) of the deuterated PR was done in solutions made with 100% D_2O according to the procedures described elsewhere.[1] The level of PR perdeuteration of 85% was deduced from the mass difference of the hydrogenous and deuterated PR using mass spectrometry measurements.

Crystallization and data collection. APV was mixed with 2.3 mg/mL enzyme in a molar ratio of 10:1. The PR-APV crystals were grown in 400μL drops made by mixing the complex and the reservoir (0.1M MES, 0.8M NaCl, pH 6.0 in H_2O) solutions at a 1:1 ratio in the 9-well glass plate/sandwich box sitting drop setup. A crystal suitable for neutron diffraction measured 0.26 × 0.77 × 1.20 mm (~0.2 mm^3; Supporting Figure 1); it was mounted in a quartz capillary

containing the reservoir solution made with 100% D_2O. Another crystal from the same crystallization drop that provided the crystal for neutron data collection was mounted in the same fashion for room-temperature X-ray data collection. The labile H atoms were allowed to exchange with D by vapor for 4 weeks for both crystals before starting data collection. Quasi-Laue neutron data to 2.0 Å resolution were collected from the 0.2 mm^3 APV-PR crystal at room temperature on the LADI-III beamline at the Institut Laue-Langevin.[3] As is typical for a Laue experiment, the crystal was held stationary at a different φ setting for each exposure. In total 18 images were collected (with an average exposure time of 21.8 h per image) from 3 different crystal orientations. The neutron data were processed using the Daresbury Laboratory *LAUE* suite program *LAUEGEN*,[4] modified to account for the cylindrical geometry of the detector.[5] The program *LSCALE*[6] was used to determine the wavelength-normalization curve using the intensities of symmetry-equivalent reflections measured at different wavelengths. No explicit absorption corrections were applied. These data were then merged in *SCALA*.[7] Monochromatic X-ray diffraction data were collected using a Rigaku HomeFlux system, equipped with a MicroMax-007 HF generator and Osmic VariMax optics. The diffraction images were obtained using an RAXIS-IV++ image-plate detector. Diffraction data were collected, integrated and scaled using HKL3000 software suite.[8] The room-temperature X-ray structure of PR-APV was refined using *SHELX-97*[9] at the resolution of 2.0 Å before using it as a starting model in joint XN refinement. A summary of the experimental and refinement statistics is given in Supporting Table 1.

Joint XN structure refinement. The joint XN structure of PR-APV was determined using *nCNS*[10] and manipulated in *Coot*.[11] After initial rigid-body refinement, several macrocycles of positional, atomic displacement parameter, and occupancy refinement followed. Between each macrocycle the structure was checked, side-chain conformations were altered and water molecule orientations were built based on the F_O-F_C difference neutron scattering density map. The $2F_O$-F_C and F_O-F_C neutron scattering density maps were then examined to determine the correct orientation of hydroxyl groups, and protonation states of His and Lys residues. The protonation states of some disordered side chains could not be obtained directly, and remained ambiguous. All water molecules were refined as D_2O. Initially, water oxygen atoms were positioned according to their electron density peaks, and then were shifted slightly in accordance with the neutron scattering density maps. The level of H/D exchange at all positions was refined

because of the enzyme's 85% perdeuteration. All H positions in PR, and only labile H positions in APV, were modeled as D and then the occupancy of D was allowed to refine within the range of -0.56 to 1.00 (the scattering length of H is -0.56 times the scattering length of D). Before depositing the final structure to the PDB, a script was run that converts a record for the coordinate of D atom into two records corresponding to an H and a D atom partially occupying the same site, both with positive partial occupancies that add up to unity. The Ramachandran statistics for the PR-APV structure reported here are as follows: residues in most favored regions 99.0%, residues in additional allowed regions 1.0%. The joint XN structure of PR-APV has been deposited to the Protein Data Bank (code 4JEC).

References

(1) Wondrak, E.M., Louis, J.M. Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease. *Biochemistry* **1996**, *35*, 12957-12962.

(2) Mahalingam, B., Louis, J.M., Hung, J., Harrison, R.W., Weber, I.T. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. *Proteins* **2001**, *43*, 455-464.

(3) Blakeley, M.P., Teixeira, S.C.M., Petit-Haertlein, I., Hazemann, I., Mitschler, A., Haertlein, M, Howard, E., Podjarny, A.D. Neutron macromolecular crystallography with LADI-III. *Acta Cryst.* **2010**, *D66***,** 1198-1205.

(4) Campbell, J.W. LAUEGEN, an X-windows-based program for the processing of Laue diffraction data. *J. Appl. Cryst.* **1995**, *28*, 228-236.

(5) Campbell, J.W., Hao, Q., Harding, M.M., Nguti, N.D., Wilkinson, C. LAUEGEN version 6.0 and INTLDM. *J. Appl. Cryst.* **1998**, *31***,** 496-502.

(6) Arzt, S., Campbell, J.W., Harding, M.M., Hao, Q., Helliwell, J.R. LSCALE - the new normalization, scaling and absorption correction program in the Daresbury Laue software suite. *J. Appl. Cryst.* **1999**, *32***,** 554-562.

(7) Weiss, M.S. Global indicators of X-ray data quality. *J. Appl. Cryst.* **2001**, *34*, 130-135.

(8) Minor, W., Cymborowski, M., Otwinovski, Z., Chruszcz, M. HKL3000: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes. *Acta Cryst.* **2006**, *D62*, 859-866.

(9) Sheldrick, G.M. A short history of SHELX. *Acta Cryst.* **2008**, *A64*, 112-122.

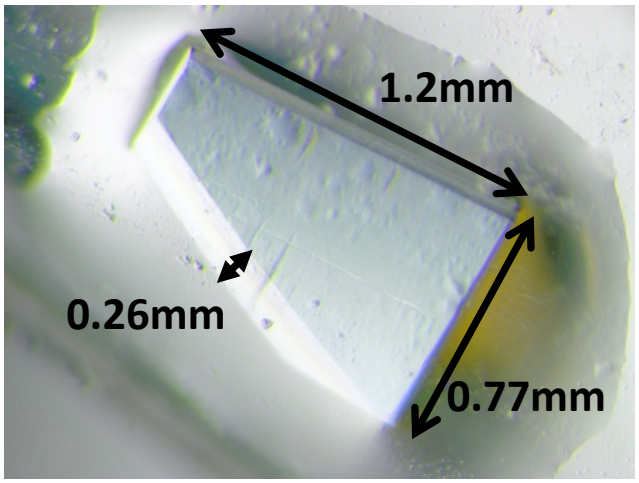
(10) Adams, P.D., Mustyalimov, M., Afonine, P.V., Langan, P. Generalized X-ray and neutron crystallographic analysis: more accurate and complete structures for biological macromolecules. *Acta Cryst.* **2009**, *D65*, 567-573.

(11) Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. Features and development of Coot. *Acta Cryst.* **2010**, *D66*, 486-501.

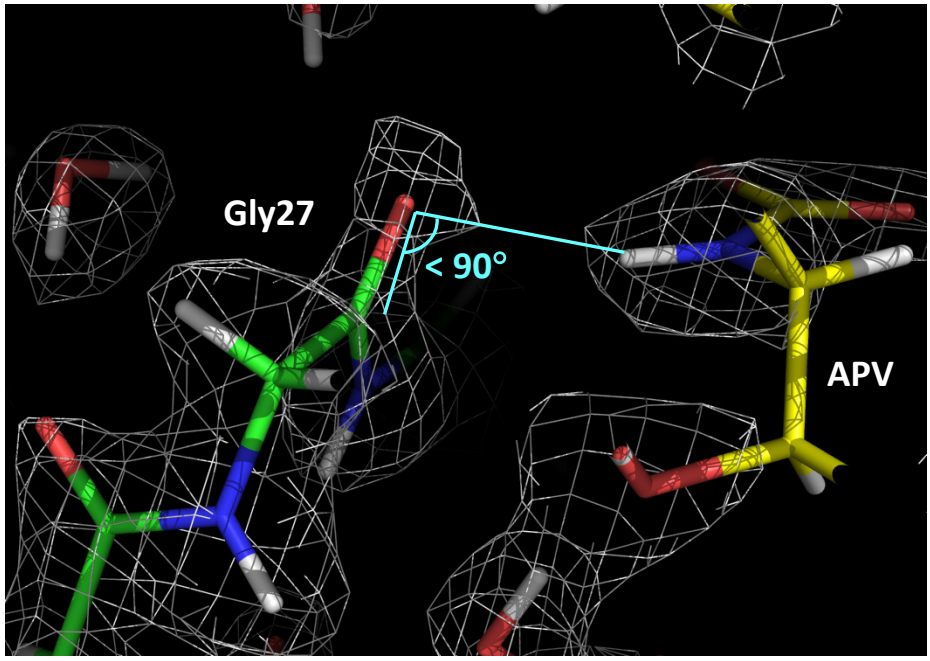
Supporting Table 1. Neutron data collection and joint XN refinement statistics.

	PR-APV
PDB ID	4JEC
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	59.19, 87.43, 46.41
α, β, γ (°)	90, 90, 90
Resolution (Å)	59.11-2.00 (2.11-2.00) *
R_{merge}	0.128 (0.269)
$\langle I / \sigma I \rangle$	7.5 (4.1)
Completeness (%)	77.3 (64.9)
Redundancy	3.2 (3.1)
No. reflections measured	41245 (4813)
No. unique reflections	12800 (1561)
Data-rejection criterion	No observation and |F| = 0
Joint XN Refinement	
Resolution (neutron) (Å)	20.00-2.00
Resolution (X-ray) (Å)	20.00-2.00
No. reflections (neutron)	12755
No. reflections (X-ray)	16017
R_{work} / R_{free} (neutron)	0.244 / 0.261
R_{work} / R_{free} (X-ray)	0.193 / 0.203
No. atoms	
Protein	3128
APV	70
Ion (Cl^-)	1
Water	393 [131 D_2O molecules]
B-factors	
Protein	28.3
APV	21.3
Ion (Cl^-)	44.9
Water	50.4
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.02

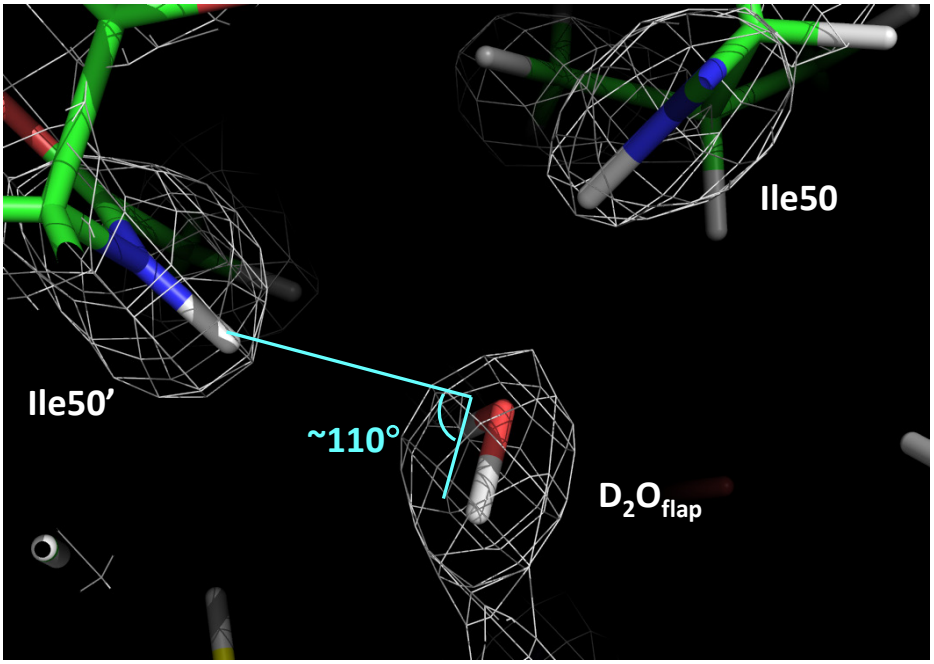
*Values in parentheses are for highest-resolution shell.



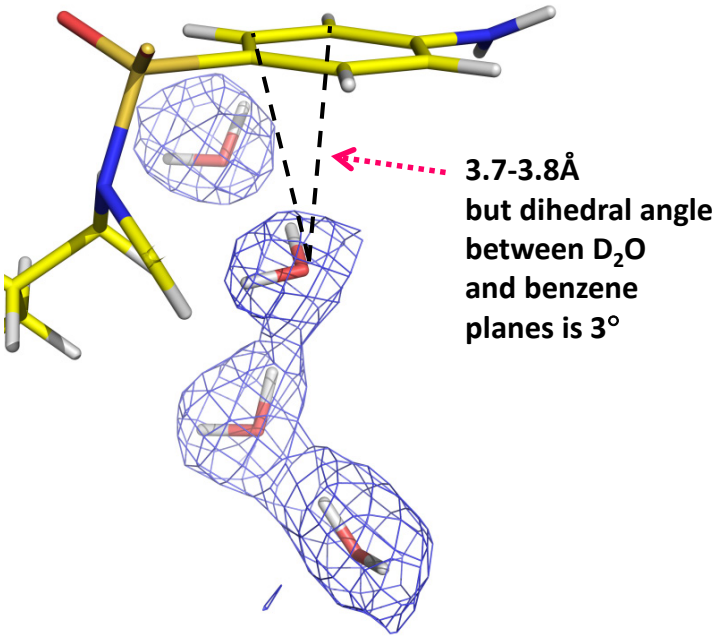
Supporting Figure 1. The 0.2mm^3 PR-APV crystal used for neutron diffraction data collection.



Supporting Figure 2a. Geometry of APV's amide interaction with the main-chain carbonyl of Gly27. The $2F_O$-F_C neutron scattering density map is contoured at 2σ level.



Supporting Figure 2b. Geometry of the flap's D_2O water molecule interaction with the main-chain amide of Ile50'. The $2F_O$-F_C neutron scattering density map is contoured at 2σ level.



Supporting Figure 2c. Geometry of the D_2O water molecules near the aromatic aniline ring, indicating no O-H…π between them. The $2F_O$-F_C neutron scattering density map is contoured at 1.2σ level.