Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2023 – 05:46 AM EDT

PDB ID	:	6PTP
Title	:	Joint X-ray/neutron structure of HIV-1 protease triple mutant (V32I,I47V,V82I) with tetrahedral intermediate mimic KVS-1
Authors	:	Kovalevsky, A.; Das, A.
Deposited on	:	2019-07-16
Resolution	:	1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

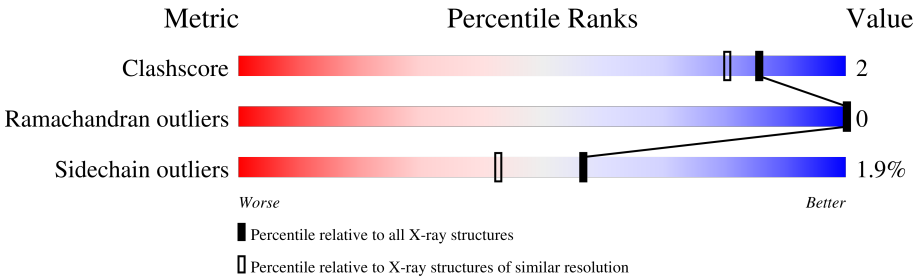
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION, NEUTRON DIFFRACTION

The reported resolution of this entry is 1.85 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	99	99% .
1	B	99	98% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4570 atoms, of which 1117 are hydrogens and 1779 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

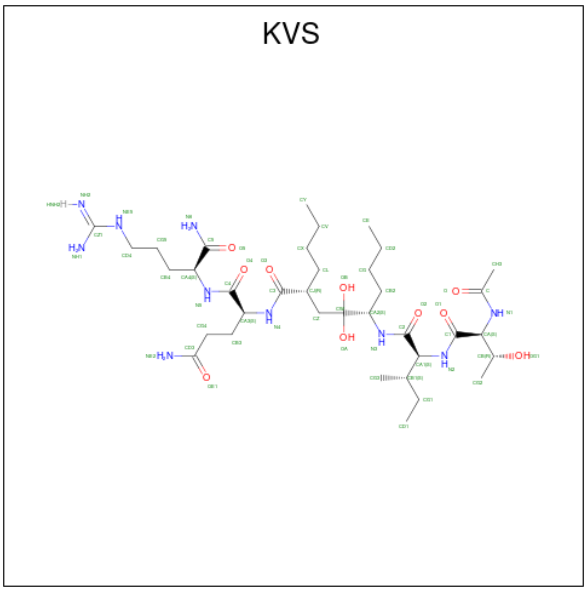
- Molecule 1 is a protein called HIV-1 Protease.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	99	Total	C	D	H	N	O	S	0	97	0
			2054	491	771	526	130	134	2			
1	B	99	Total	C	D	H	N	O	S	0	97	0
			2083	491	784	542	130	134	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	LYS	GLN	engineered mutation	UNP Q7SSI0
A	32	ILE	VAL	engineered mutation	UNP Q7SSI0
A	33	ILE	LEU	engineered mutation	UNP Q7SSI0
A	47	VAL	ILE	engineered mutation	UNP Q7SSI0
A	63	ILE	LEU	engineered mutation	UNP Q7SSI0
A	67	ALA	CYS	engineered mutation	UNP Q7SSI0
A	82	ILE	VAL	engineered mutation	UNP Q7SSI0
A	95	ALA	SER	engineered mutation	UNP Q7SSI0
B	107	LYS	GLN	engineered mutation	UNP Q7SSI0
B	132	ILE	VAL	engineered mutation	UNP Q7SSI0
B	133	ILE	LEU	engineered mutation	UNP Q7SSI0
B	147	VAL	ILE	engineered mutation	UNP Q7SSI0
B	163	ILE	LEU	engineered mutation	UNP Q7SSI0
B	167	ALA	CYS	engineered mutation	UNP Q7SSI0
B	182	ILE	VAL	engineered mutation	UNP Q7SSI0
B	195	ALA	SER	engineered mutation	UNP Q7SSI0

- Molecule 2 is N 2 -[(2R,5S)-5-({(2S,3S)-2-[(N-acetyl-L-threonyl)amino]-3-methylpent-4-enoyl}amino)-2-butyl-4,4-dihydroxynonanoyl]-L-glutaminyl-L-argininamide (three-letter code: KVS) (formula: $C_{36}H_{68}N_{10}O_{10}$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	B	1	Total	C	D	H	N	O	0	0
			121	36	16	49	10	10		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	42	Total	D	O	0	0
			126	84	42		
3	B	62	Total	D	O	0	0
			186	124	62		

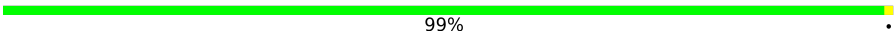
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS failed to run properly.

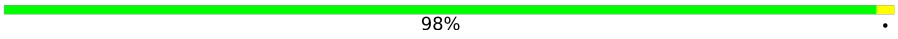
- Molecule 1: HIV-1 Protease

Chain A:



- Molecule 1: HIV-1 Protease

Chain B:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	59.50Å 87.40Å 46.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.85	Depositor
% Data completeness (in resolution range)	87.6 (40.00-1.85)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.92 (at 1.85Å)	Xtriage
Refinement program	nCNS 1.0.0	Depositor
R, R_{free}	0.192 , 0.217	Depositor
Wilson B-factor (Å^2)	19.4	Xtriage
Anisotropy	0.027	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4570	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.27% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: KVS, DOD

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.41	0/1527	0.64	0/2064
1	B	0.43	0/1532	0.66	0/2072
All	All	0.42	0/3059	0.65	0/4136

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1528	526	9	0	0
1	B	1541	542	6	0	0
2	B	72	49	65	1	0
3	A	126	0	0	0	1
3	B	186	0	0	0	1
All	All	3453	1117	80	1	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (1) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:140:DOD:O	3:B:430:DOD:D2[1_554]	1.59	0.61

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	192/99 (194%)	190 (99%)	2 (1%)	0	100	100
1	B	192/99 (194%)	188 (98%)	4 (2%)	0	100	100
All	All	384/198 (194%)	378 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	161/81 (199%)	159 (99%)	2 (1%)	71	62
1	B	162/81 (200%)	158 (98%)	4 (2%)	47	31
All	All	323/162 (199%)	317 (98%)	6 (2%)	57	43

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19[A]	LEU
1	A	19[B]	LEU
1	B	133[A]	ILE
1	B	133[B]	ILE
1	B	180[A]	THR
1	B	180[B]	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	KVS	B	301	-	53,55,55	0.86	4 (7%)	64,73,73	1.58	11 (17%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	KVS	B	301	-	-	14/76/81/81	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	301	KVS	CB2-CA2	-3.26	1.50	1.53
2	B	301	KVS	OA-CW	2.63	1.44	1.40
2	B	301	KVS	OB-CW	2.30	1.43	1.40
2	B	301	KVS	CA3-N4	2.22	1.50	1.45

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	301	KVS	O3-C3-CJ	4.38	127.82	122.12
2	B	301	KVS	CL-CJ-C3	4.00	116.33	109.53
2	B	301	KVS	CG2-CB-CA	-3.91	104.40	112.29
2	B	301	KVS	CA2-N3-C2	-3.67	117.70	123.22
2	B	301	KVS	CZ-CJ-C3	3.54	115.48	109.57
2	B	301	KVS	O3-C3-N4	-3.34	116.75	122.93
2	B	301	KVS	CB3-CA3-N4	-2.73	105.35	110.88
2	B	301	KVS	O2-C2-N3	-2.51	118.28	122.93
2	B	301	KVS	O2-C2-CA1	2.48	125.87	120.74
2	B	301	KVS	CA4-N5-C4	-2.35	116.64	121.67
2	B	301	KVS	CA3-N4-C3	-2.17	117.01	121.67

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	301	KVS	C3-CJ-CL-CX
2	B	301	KVS	NH1-CZ1-NE5-CD4
2	B	301	KVS	NH2-CZ1-NE5-CD4
2	B	301	KVS	N5-CA4-CB4-CG5
2	B	301	KVS	C5-CA4-CB4-CG5
2	B	301	KVS	CJ-CL-CX-CV
2	B	301	KVS	CA3-CB3-CG4-CD3
2	B	301	KVS	CZ-CJ-CL-CX
2	B	301	KVS	CE-CD2-CG-CB2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	B	301	KVS	O3-C3-CJ-CL
2	B	301	KVS	OA-CW-CZ-CJ
2	B	301	KVS	N3-CA2-CB2-CG
2	B	301	KVS	N4-C3-CJ-CL
2	B	301	KVS	O2-C2-CA1-N2

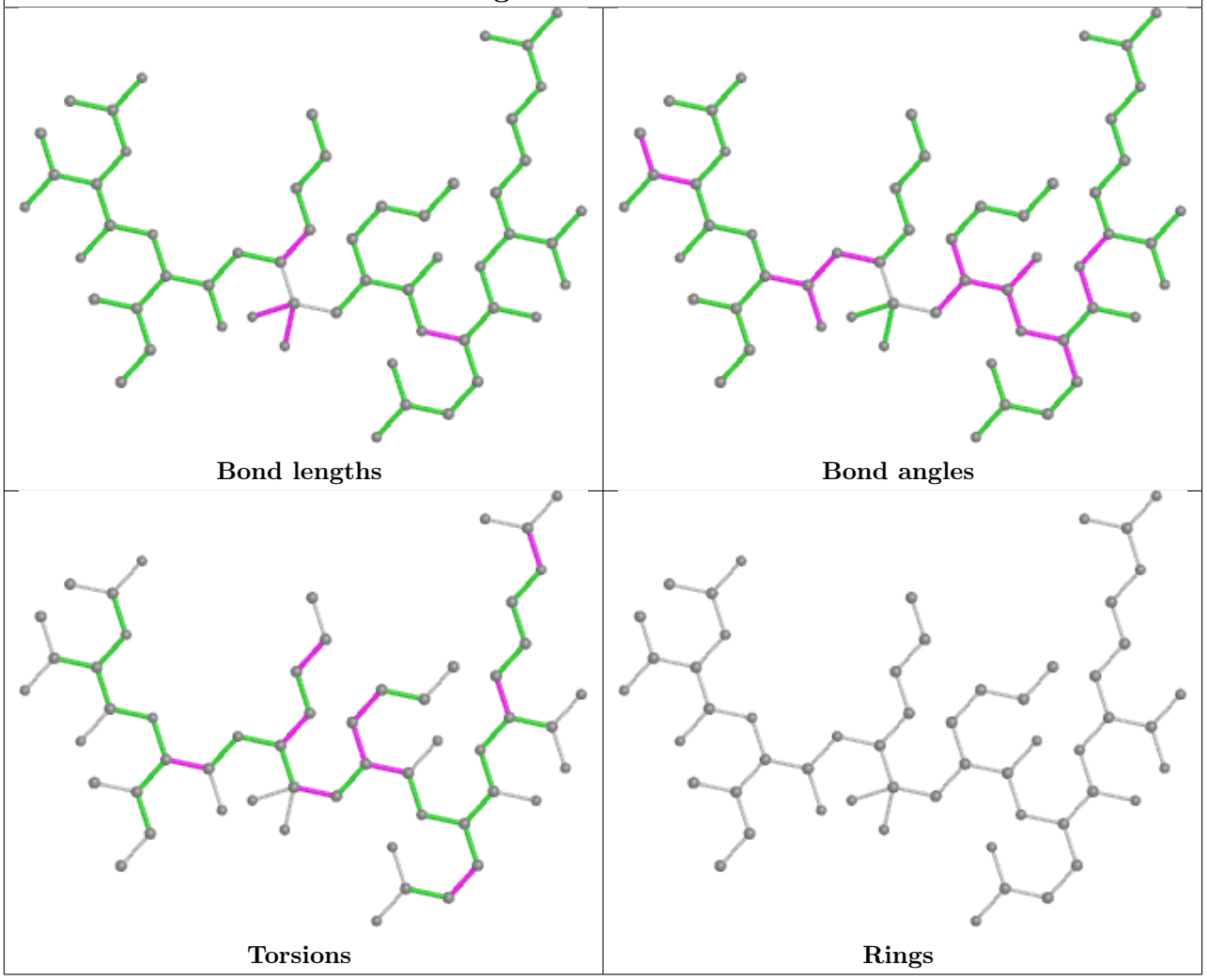
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	301	KVS	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand KVS B 301



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.