Full wwPDB X-ray Structure Validation Report (i)

Feb 25, 2024 – 05:51 AM EST

PDB ID	:	5E5J
Title	:	Joint X-ray/neutron structure of HIV-1 protease triple mutant (V32I,I47V,V82I) with darunavir at pH 6.0
Authors	:	Kovalevsky, A.Y.; Gerlits, O.O.
Deposited on	:	2015-10-08
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 (i) 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 (1)) were used in the production of this report:

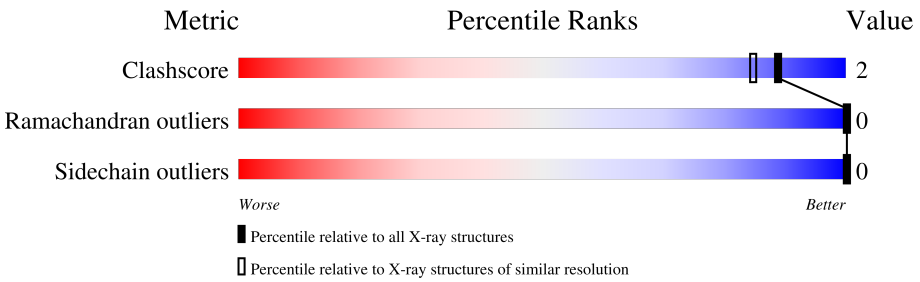
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson 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	100%
1	B	99	100%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4828 atoms, of which 1300 are hydrogens and 1860 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.

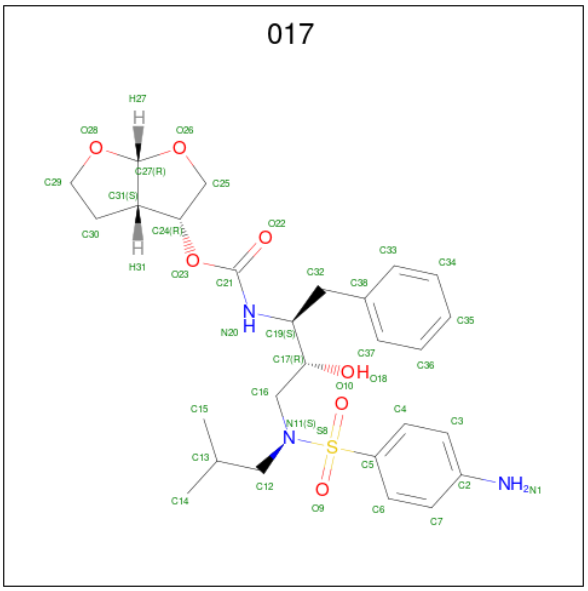
- Molecule 1 is a protein called Protease.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	D	H	N	O	S			
1	A	99	2216	491	813	646	130	134	2	0	98	0
1	B	99	2189	491	811	621	130	134	2	0	99	0

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	7	LYS	GLN	engineered mutation	UNP Q7SSI0
B	32	ILE	VAL	engineered mutation	UNP Q7SSI0
B	33	ILE	LEU	engineered mutation	UNP Q7SSI0
B	47	VAL	ILE	engineered mutation	UNP Q7SSI0
B	63	ILE	LEU	engineered mutation	UNP Q7SSI0
B	67	ALA	CYS	engineered mutation	UNP Q7SSI0
B	82	ILE	VAL	engineered mutation	UNP Q7SSI0
B	95	ALA	SER	engineered mutation	UNP Q7SSI0

- Molecule 2 is (3R,3AS,6AR)-HEXAHYDROFURO[2,3-B]FURAN-3-YL(1S,2R)-3-[[(4-AMINOPHENYL)SULFONYL](ISOBUTYL)AMINO]-1-BENZYL-2-HYDROXYPROPYLCARBAMATE (three-letter code: 017) (formula: $C_{27}H_{37}N_3O_7S$).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
			Total	C	D	H	N	O	S		
2	A	1	75	27	4	33	3	7	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	D	O		
3	A	60	180	120	60	0	0
3	B	56	Total	D	O	0	0
			168	112	56		

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.

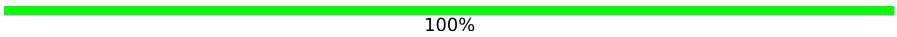
Note EDS failed to run properly.

- Molecule 1: Protease

Chain A:

There are no outlier residues recorded for this chain.

- Molecule 1: Protease

Chain B: 

There are no outlier residues recorded for this chain.

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	59.72Å 87.26Å 46.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.85	Depositor
% Data completeness (in resolution range)	87.9 (20.00-1.85)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.91 (at 1.85Å)	Xtriage
Refinement program	nCNS 1.0.0	Depositor
R, R_{free}	0.194 , 0.201	Depositor
Wilson B-factor (Å^2)	20.1	Xtriage
Anisotropy	0.039	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4828	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 6.49% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: DOD, 017

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	0/1536	0.65	0/2077
1	B	0.38	0/1540	0.64	0/2082
All	All	0.40	0/3076	0.64	0/4159

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 (i)

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	1570	646	3	0	0
1	B	1568	621	0	0	0
2	A	42	33	37	1	0
3	A	180	0	0	2	0
3	B	168	0	0	2	0
All	All	3528	1300	40	5	0

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 (5) 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)

There are no symmetry-related clashes.

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	193/99 (195%)	191 (99%)	2 (1%)	0	100	100
1	B	194/99 (196%)	194 (100%)	0	0	100	100
All	All	387/198 (196%)	385 (100%)	2 (0%)	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	162/81 (200%)	162 (100%)	0	100	100
1	B	162/81 (200%)	162 (100%)	0	100	100
All	All	324/162 (200%)	324 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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	017	A	201	-	41,41,41	2.60	16 (39%)	52,58,58	2.20	13 (25%)

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	017	A	201	-	-	4/36/55/55	0/4/4/4

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	201	017	O9-S8	8.98	1.53	1.43
2	A	201	017	O10-S8	7.65	1.52	1.43
2	A	201	017	C32-C38	-4.05	1.41	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	201	017	C6-C5	3.92	1.45	1.38
2	A	201	017	S8-N11	3.24	1.68	1.63
2	A	201	017	C33-C38	3.02	1.45	1.38
2	A	201	017	O28-C27	2.92	1.47	1.41
2	A	201	017	C36-C37	2.87	1.45	1.38
2	A	201	017	C7-C6	2.71	1.43	1.38
2	A	201	017	C36-C35	2.62	1.45	1.38
2	A	201	017	C37-C38	2.62	1.44	1.38
2	A	201	017	C35-C34	2.58	1.44	1.38
2	A	201	017	C34-C33	2.39	1.44	1.38
2	A	201	017	C7-C2	2.26	1.44	1.40
2	A	201	017	C4-C3	2.20	1.42	1.38
2	A	201	017	C4-C5	2.15	1.42	1.38

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	201	017	O10-S8-O9	-7.35	107.62	119.52
2	A	201	017	O23-C21-N20	5.49	117.44	110.32
2	A	201	017	C5-S8-N11	5.11	113.39	107.30
2	A	201	017	C24-O23-C21	4.57	122.98	117.03
2	A	201	017	C19-N20-C21	3.97	129.81	122.37
2	A	201	017	C13-C12-N11	3.76	116.77	112.41
2	A	201	017	O22-C21-N20	-3.31	119.42	124.85
2	A	201	017	C32-C19-C17	3.28	117.14	111.65
2	A	201	017	O9-S8-N11	3.25	109.65	106.69
2	A	201	017	C16-N11-S8	-3.24	110.41	117.52
2	A	201	017	C38-C32-C19	2.47	117.62	113.33
2	A	201	017	C25-O26-C27	-2.15	103.98	106.86
2	A	201	017	C7-C2-C3	2.15	121.48	118.15

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	201	017	C12-N11-S8-O10
2	A	201	017	C16-C17-C19-N20
2	A	201	017	C25-C24-O23-C21
2	A	201	017	C31-C24-O23-C21

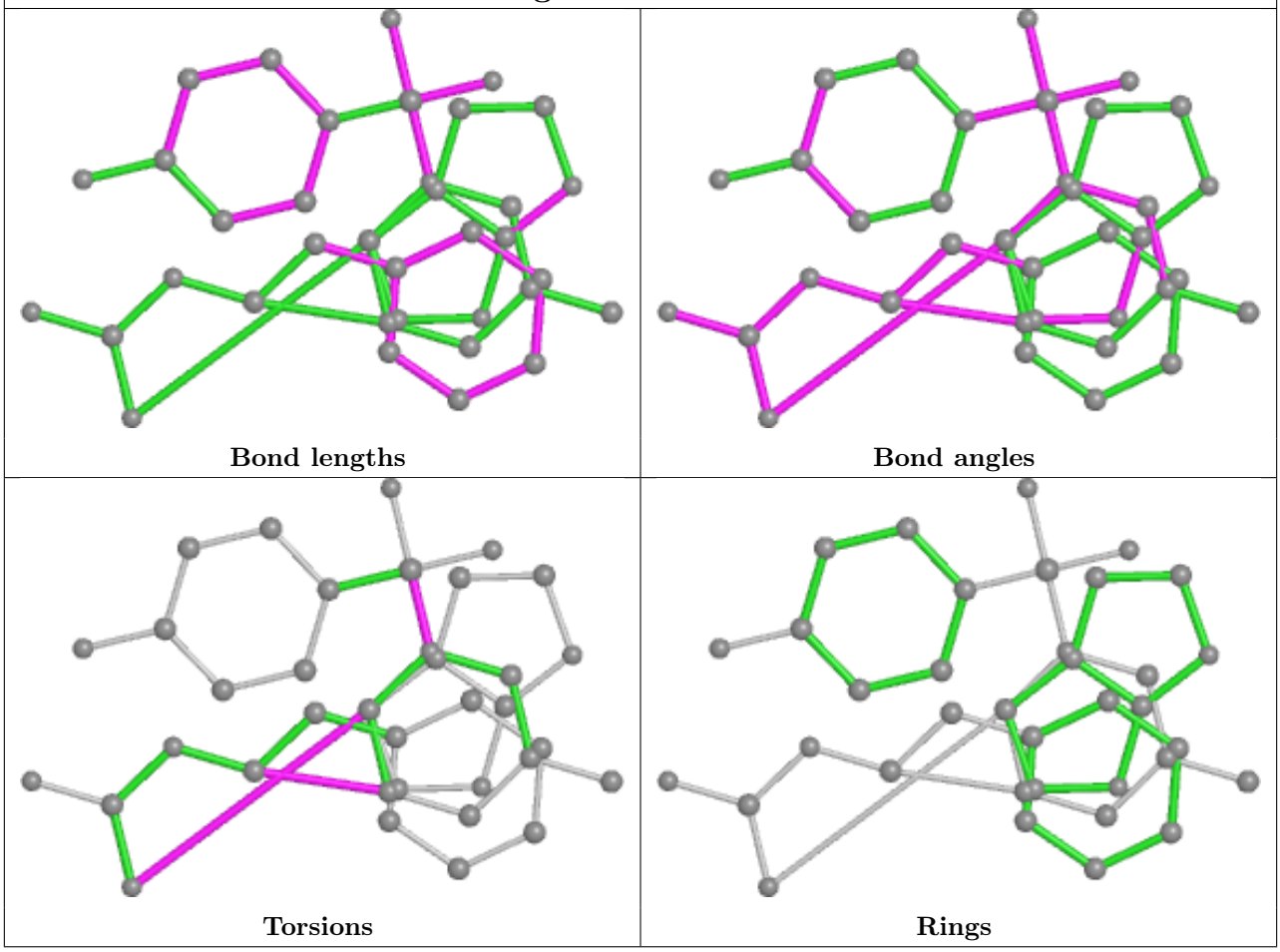
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	017	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 017 A 201



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 (i)

6.1 Protein, DNA and RNA chains (i)

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

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

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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