Full wwPDB X-ray Structure Validation Report (i)

Oct 23, 2021 – 10:27 AM EDT

PDB ID	:	1D4Y
Title	:	HIV-1 PROTEASE TRIPLE MUTANT/TIPRANAVIR COMPLEX
Authors	:	Watenpaugh, K.D.; Janakiraman, M.N.
Deposited on	:	1999-10-06
Resolution	:	1.97 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 following versions of software and data (see references (1)) were used in the production of this report:

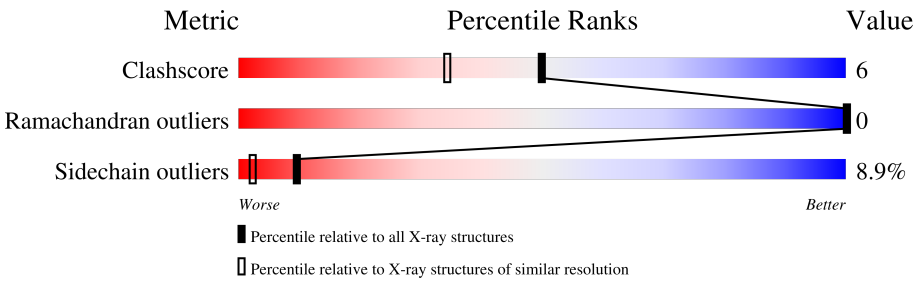
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.97 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	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)

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 was not executed.

Mol	Chain	Length	Quality of chain
1	A	99	82% 18%
1	B	99	86% 11% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1737 atoms, of which 0 are hydrogens and 0 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.

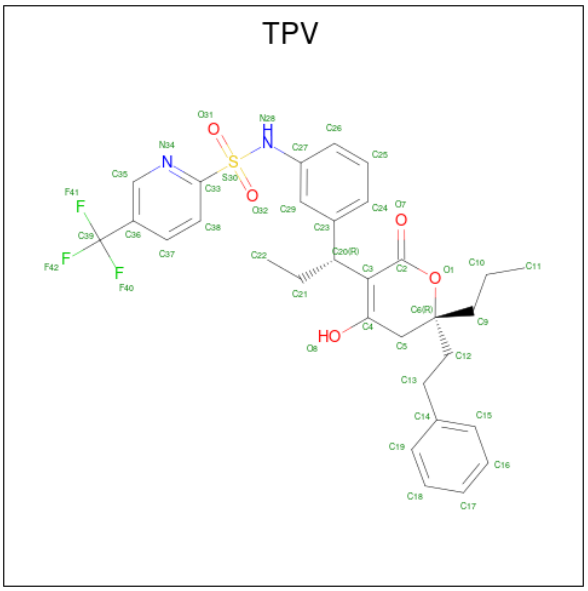
- Molecule 1 is a protein called PROTEIN (HIV-1 PROTEASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	99	741	479	124	134	4	0	0	0
1	B	99	747	484	126	133	4	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	LYS	GLN	engineered mutation	UNP P03367
A	33	ILE	LEU	engineered mutation	UNP P03367
A	63	ILE	LEU	engineered mutation	UNP P03367
B	7	LYS	GLN	engineered mutation	UNP P03367
B	33	ILE	LEU	engineered mutation	UNP P03367
B	63	ILE	LEU	engineered mutation	UNP P03367

- Molecule 2 is N-(3-{(1R)-1-[(6R)-4-HYDROXY-2-OXO-6-PHENETHYL-6-PROPYL-5,6-DIHYDRO-2H-PYRAN-3-YL]PROPYL}PHENYL)-5-(TRIFLUOROMETHYL)-2-PYRIDINESULFONAMIDE (three-letter code: TPV) (formula: $C_{31}H_{33}F_3N_2O_5S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	F	N	O	S		
2	A	1	42	31	3	2	5	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	95	95	95	0	0
			Total	O		
3	B	112	112	112	0	0

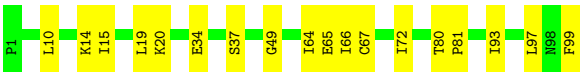
3 Residue-property plots (i)

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 was not executed.

- Molecule 1: PROTEIN (HIV-1 PROTEASE)

Chain A:



- Molecule 1: PROTEIN (HIV-1 PROTEASE)

Chain B:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	59.78Å 87.50Å 46.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.97	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-1.97)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.175 , 0.216	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1737	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

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: TPV

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	0.68	0/754	1.02	0/1023
1	B	0.72	0/760	1.10	0/1030
All	All	0.70	0/1514	1.06	0/2053

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	741	0	770	11	0
1	B	747	0	782	10	0
2	A	42	0	29	3	0
3	A	95	0	0	1	0
3	B	112	0	0	1	0
All	All	1737	0	1581	18	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 6.

All (18) 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)
2:A:501:TPV:N28	2:A:501:TPV:C27	1.80	1.43
1:A:81:PRO:HD2	1:B:50:ILE:HD12	1.67	0.74
2:A:501:TPV:C27	2:A:501:TPV:S30	2.75	0.74
1:B:61:GLN:HG3	3:B:159:HOH:O	1.95	0.67
1:A:80:THR:OG1	1:B:50:ILE:HD11	2.01	0.61
1:A:14:LYS:HB3	1:A:65:GLU:HB2	1.83	0.59
1:B:65:GLU:OE1	1:B:68:GLY:HA2	2.04	0.58
1:A:20:LYS:HD3	1:A:34:GLU:OE2	2.06	0.55
1:A:99:PHE:CE1	1:B:1:PRO:HD2	2.47	0.50
1:A:49:GLY:HA3	1:B:50:ILE:HG22	1.94	0.49
2:A:501:TPV:H121	1:B:49:GLY:HA2	1.95	0.48
1:A:81:PRO:HD2	1:B:50:ILE:CD1	2.39	0.46
1:A:10:LEU:HD12	3:A:550:HOH:O	2.18	0.44
1:A:93:ILE:HD13	1:A:93:ILE:HG21	1.87	0.43
1:A:15:ILE:HD12	1:A:64:ILE:HD11	2.00	0.43
1:B:50:ILE:HD12	1:B:50:ILE:HA	1.74	0.43
1:B:42:TRP:CE2	1:B:57:ARG:HD2	2.57	0.40
1:A:66:ILE:HD13	1:A:66:ILE:HG21	1.95	0.40

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	97/99 (98%)	97 (100%)	0	0	100	100
1	B	97/99 (98%)	96 (99%)	1 (1%)	0	100	100
All	All	194/198 (98%)	193 (100%)	1 (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	79/83 (95%)	74 (94%)	5 (6%)	18 7
1	B	79/83 (95%)	70 (89%)	9 (11%)	5 1
All	All	158/166 (95%)	144 (91%)	14 (9%)	9 2

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	LEU
1	A	37	SER
1	A	67	CYS
1	A	72	ILE
1	A	97	LEU
1	B	37	SER
1	B	46	MET
1	B	50	ILE
1	B	61	GLN
1	B	64	ILE
1	B	65	GLU
1	B	67	CYS
1	B	82	VAL
1	B	97	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	GLN

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	TPV	A	501	-	43,45,45	5.81	15 (34%)	56,66,66	2.94	23 (41%)

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	TPV	A	501	-	-	9/37/55/55	0/4/4/4

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	TPV	C5-C6	-27.06	1.27	1.53
2	A	501	TPV	C27-N28	22.82	1.80	1.43
2	A	501	TPV	C9-C10	-6.79	1.31	1.52
2	A	501	TPV	O31-S30	6.62	1.51	1.43
2	A	501	TPV	C12-C13	-3.19	1.46	1.53
2	A	501	TPV	O1-C2	-3.18	1.32	1.36
2	A	501	TPV	C5-C4	2.79	1.54	1.50
2	A	501	TPV	F40-C39	2.72	1.42	1.32

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	TPV	F42-C39	2.63	1.42	1.32
2	A	501	TPV	C13-C14	-2.40	1.44	1.51
2	A	501	TPV	C2-C3	2.25	1.49	1.44
2	A	501	TPV	C3-C4	2.20	1.41	1.37
2	A	501	TPV	C12-C6	-2.18	1.50	1.53
2	A	501	TPV	C33-N34	2.13	1.36	1.33
2	A	501	TPV	C19-C14	2.13	1.43	1.38

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	TPV	O1-C6-C12	-10.86	80.08	105.53
2	A	501	TPV	O1-C6-C5	9.18	120.98	109.64
2	A	501	TPV	C5-C4-C3	-6.08	116.41	123.08
2	A	501	TPV	C12-C13-C14	5.61	127.54	112.85
2	A	501	TPV	C11-C10-C9	5.08	123.54	112.19
2	A	501	TPV	C27-N28-S30	-4.94	107.84	123.24
2	A	501	TPV	C12-C6-C5	3.93	120.42	111.76
2	A	501	TPV	F40-C39-C36	-3.86	104.46	112.93
2	A	501	TPV	F42-C39-C36	-3.80	104.58	112.93
2	A	501	TPV	C22-C21-C20	3.38	119.97	112.70
2	A	501	TPV	C23-C20-C3	-3.26	106.41	112.37
2	A	501	TPV	O7-C2-C3	-2.92	121.61	125.50
2	A	501	TPV	F41-C39-C36	-2.84	106.69	112.93
2	A	501	TPV	C26-C27-N28	-2.55	114.57	120.09
2	A	501	TPV	F42-C39-F41	2.40	114.54	105.72
2	A	501	TPV	O32-S30-O31	-2.32	116.70	119.55
2	A	501	TPV	C9-C6-C5	-2.31	106.67	111.76
2	A	501	TPV	C13-C12-C6	-2.27	109.81	115.34
2	A	501	TPV	C13-C14-C19	-2.26	115.50	121.23
2	A	501	TPV	C38-C33-N34	-2.24	122.82	125.28
2	A	501	TPV	C38-C37-C36	-2.21	118.27	121.22
2	A	501	TPV	F41-C39-F40	2.14	113.58	105.72
2	A	501	TPV	C29-C27-N28	2.10	125.32	119.93

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	TPV	C5-C6-C9-C10
2	A	501	TPV	C12-C6-C9-C10
2	A	501	TPV	C6-C12-C13-C14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	501	TPV	C3-C20-C23-C29
2	A	501	TPV	C3-C20-C23-C24
2	A	501	TPV	N34-C33-S30-N28
2	A	501	TPV	C12-C13-C14-C15
2	A	501	TPV	N34-C33-S30-O32
2	A	501	TPV	C12-C13-C14-C19

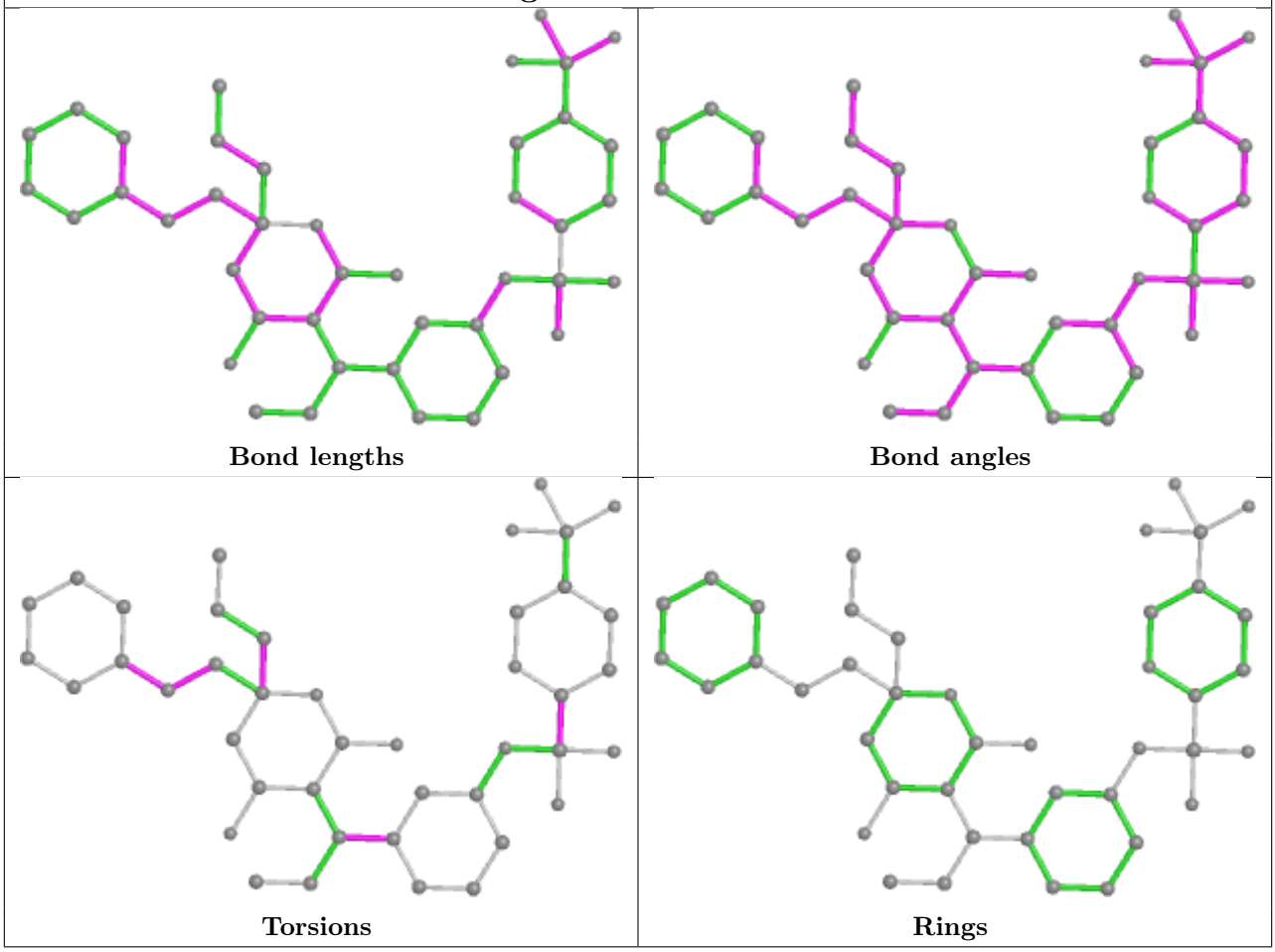
There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	TPV	3	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 TPV A 501



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 was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.