



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 7, 2018 – 12:03 pm GMT

PDB ID : 2QNP
Title : HIV-1 Protease in complex with a iodo decorated pyrrolidine-based inhibitor
Authors : Boettcher, J.; Blum, A.; Heine, A.; Diederich, W.E.; Klebe, G.
Deposited on : 2007-07-19
Resolution : 1.41 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : (not set)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

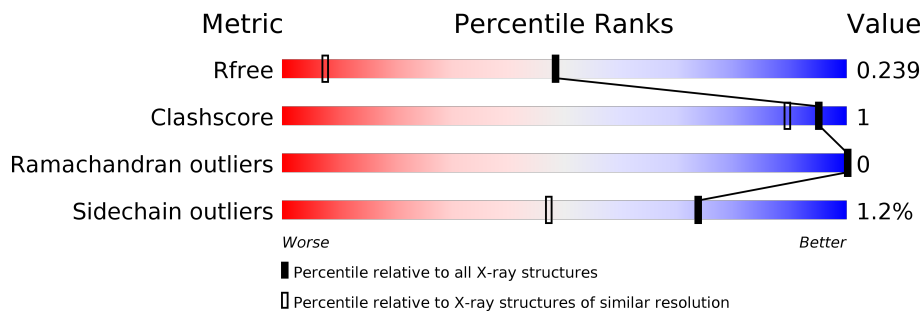
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.41 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))
R_{free}	111664	2084 (1.44-1.40)
Clashscore	122126	2190 (1.44-1.40)
Ramachandran outliers	120053	2138 (1.44-1.40)
Sidechain outliers	120020	2137 (1.44-1.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	99	
1	B	99	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1729 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.

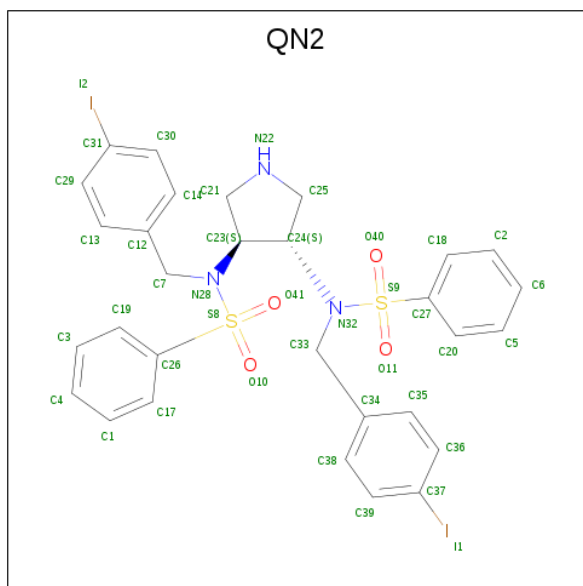
- Molecule 1 is a protein called Gag-Pol polyprotein (Pr160Gag-Pol).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	99	Total 748	C 484	N 127	O 133	S 4	0	0	0
1	B	99	Total 754	C 486	N 129	O 135	S 4	0	0	0

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total 2 Cl 2	0	0
2	A	1	Total 1 Cl 1	0	0

- Molecule 3 is N,N'-(3S,4S)-pyrrolidine-3,4-diylbis[N-(4-iodobenzyl)benzenesulfonamide] (three-letter code: QN2) (formula: C₃₀H₂₉I₂N₃O₄S₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	I	N	O			S
3	A	1	41	30	2	3	4	2	0	0

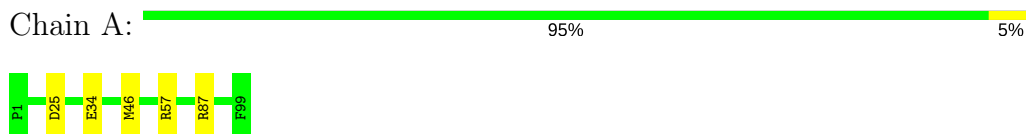
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	80	Total 80	O 80	0	0
4	B	103	Total 103	O 103	0	0

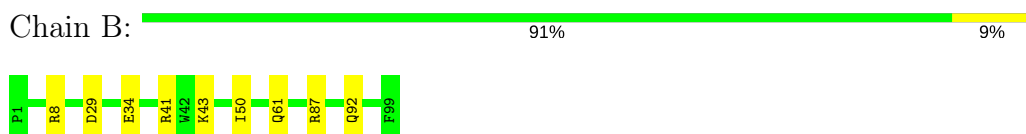
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

- Molecule 1: Gag-Pol polyprotein (Pr160Gag-Pol)



- Molecule 1: Gag-Pol polyprotein (Pr160Gag-Pol)



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	57.36Å 85.90Å 46.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.41 18.58 – 1.41	Depositor EDS
% Data completeness (in resolution range)	94.5 (10.00-1.41) 95.9 (18.58-1.41)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	5.00	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.97 (at 1.41Å)	Xtrriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.178 , 0.214 0.234 , 0.239	Depositor DCC
R_{free} test set	2240 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	1.6	Xtrriage
Anisotropy	0.468	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 57.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.30$, $\langle L^2 \rangle = 0.13$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	1729	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: QN2, CL

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	0.53	0/761	1.11	3/1031 (0.3%)
1	B	0.54	0/767	1.17	6/1039 (0.6%)
All	All	0.54	0/1528	1.14	9/2070 (0.4%)

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	41	ARG	NE-CZ-NH2	-7.92	116.34	120.30
1	B	92	GLN	OE1-CD-NE2	7.15	138.35	121.90
1	A	87	ARG	NE-CZ-NH1	6.94	123.77	120.30
1	A	57	ARG	NE-CZ-NH2	-6.30	117.15	120.30
1	B	92	GLN	CG-CD-NE2	-5.95	102.42	116.70
1	B	87	ARG	NE-CZ-NH2	5.77	123.18	120.30
1	B	34	GLU	OE1-CD-OE2	-5.47	116.73	123.30
1	A	46	MET	CG-SD-CE	5.44	108.91	100.20
1	B	29	ASP	CB-CG-OD1	5.13	122.91	118.30

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	748	0	791	2	0
1	B	754	0	795	2	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
3	A	41	0	29	2	0
4	A	80	0	0	1	0
4	B	103	0	0	1	0
All	All	1729	0	1615	4	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 1.

All (4) 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)
3:A:2501:QN2:O40	1:B:50:ILE:HD13	1.94	0.67
1:A:34:GLU:HG2	4:A:3100:HOH:O	2.00	0.61
1:A:25:ASP:OD2	3:A:2501:QN2:H332	2.00	0.61
1:B:8:ARG:HD3	4:B:3149:HOH:O	2.18	0.43

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	81/83 (98%)	81 (100%)	0	100	100
1	B	82/83 (99%)	80 (98%)	2 (2%)	52	16
All	All	163/166 (98%)	161 (99%)	2 (1%)	74	47

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	43	LYS
1	B	61	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	7	GLN
1	B	83	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

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
3	QN2	A	2501	-	41,45,45	1.77	6 (14%)	53,65,65	1.87	14 (26%)

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
3	QN2	A	2501	-	-	0/38/50/50	0/5/5/5

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2501	QN2	C7-N28	-2.53	1.44	1.48
3	A	2501	QN2	C33-C34	2.38	1.55	1.51
3	A	2501	QN2	O11-S9	2.87	1.47	1.43
3	A	2501	QN2	O40-S9	3.49	1.47	1.43
3	A	2501	QN2	S9-N32	4.04	1.69	1.63
3	A	2501	QN2	S8-N28	6.73	1.73	1.63

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2501	QN2	C26-S8-N28	-5.19	98.30	107.38
3	A	2501	QN2	C3-C19-C26	-3.30	115.50	118.95
3	A	2501	QN2	O40-S9-C27	-2.95	104.23	108.02
3	A	2501	QN2	O11-S9-O40	-2.72	115.03	119.50
3	A	2501	QN2	C18-C27-S9	-2.68	116.91	119.77
3	A	2501	QN2	C4-C1-C17	-2.57	116.67	120.20
3	A	2501	QN2	C7-C12-C14	-2.57	115.87	120.77
3	A	2501	QN2	O41-S8-C26	-2.07	105.36	108.02

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	2501	QN2	O41-S8-O10	2.01	122.81	119.50
3	A	2501	QN2	C7-C12-C13	2.07	124.73	120.77
3	A	2501	QN2	C5-C20-C27	3.05	122.14	118.95
3	A	2501	QN2	C12-C7-N28	3.84	120.41	113.55
3	A	2501	QN2	C20-C27-S9	3.96	124.00	119.77
3	A	2501	QN2	C27-S9-N32	4.26	114.84	107.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2501	QN2	2	0

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.