wwPDB X-ray Structure Validation Summary Report (i)

Oct 2, 2023 – 02:56 PM EDT

PDB ID	:	6NZT
Title	:	Crystal structure of HCV NS3/4A protease in complex with voxilaprevir
Authors	:	Appleby, T.C.; Taylor, J.G.
Deposited on	:	2019-02-14
Resolution	:	1.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	FAILED
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)	:	2.35.1

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 1.40 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	L9P	A	1201	X	-	-	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 1613 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 HCV NS3/4A protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	191	1356	849	238	263	6	0	2	0

There are 35 discrepancies between the modelled and reference sequences:

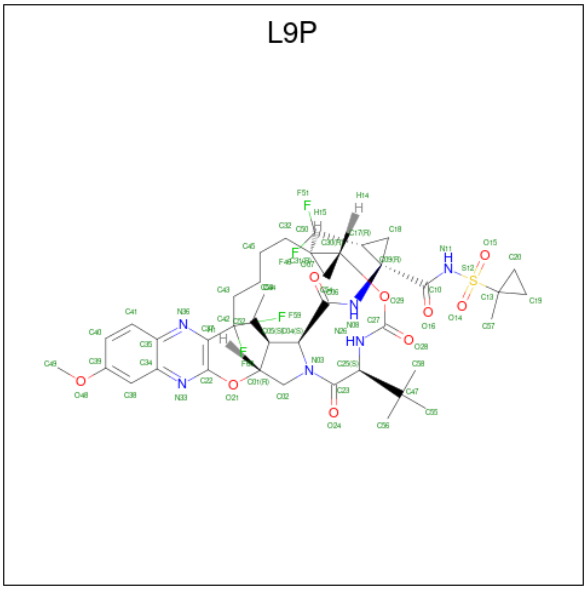
Chain	Residue	Modelled	Actual	Comment	Reference
A	980	GLY	-	expression tag	UNP S4UY05
A	981	SER	-	expression tag	UNP S4UY05
A	982	HIS	-	expression tag	UNP S4UY05
A	983	MET	-	expression tag	UNP S4UY05
A	984	ALA	-	expression tag	UNP S4UY05
A	985	SER	-	expression tag	UNP S4UY05
A	986	MET	-	expression tag	UNP S4UY05
A	987	LYS	-	expression tag	UNP S4UY05
A	988	LYS	-	expression tag	UNP S4UY05
A	989	LYS	-	expression tag	UNP S4UY05
A	990	GLY	-	expression tag	UNP S4UY05
A	991	SER	-	expression tag	UNP S4UY05
A	992	VAL	-	expression tag	UNP S4UY05
A	993	VAL	-	expression tag	UNP S4UY05
A	994	ILE	-	expression tag	UNP S4UY05
A	995	VAL	-	expression tag	UNP S4UY05
A	996	GLY	-	expression tag	UNP S4UY05
A	997	ARG	-	expression tag	UNP S4UY05
A	998	ILE	-	expression tag	UNP S4UY05
A	999	ASN	-	expression tag	UNP S4UY05
A	1000	LEU	-	expression tag	UNP S4UY05
A	1001	SER	-	expression tag	UNP S4UY05
A	1002	GLY	-	expression tag	UNP S4UY05
A	1003	ASP	-	expression tag	UNP S4UY05
A	1013	GLU	LEU	variant	UNP S4UY05
A	1014	GLU	LEU	variant	UNP S4UY05
A	1017	GLN	ILE	variant	UNP S4UY05

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1018	GLU	VAL	variant	UNP S4UY05
A	1021	GLN	LEU	variant	UNP S4UY05
A	1047	SER	CYS	variant	UNP S4UY05
A	1052	LEU	CYS	variant	UNP S4UY05
A	1072	THR	VAL	variant	UNP S4UY05
A	1086	GLN	PRO	variant	UNP S4UY05
A	1159	SER	CYS	variant	UNP S4UY05
A	1168	GLN	ASP	engineered mutation	UNP S4UY05

- Molecule 2 is Voxilaprevir (three-letter code: L9P) (formula: $C_{40}H_{52}F_4N_6O_9S$) (labeled as "Ligand of Interest" by depositor).

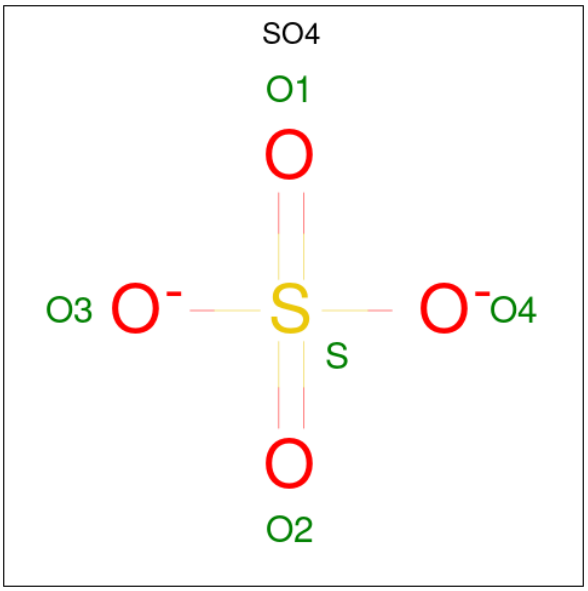


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	F	N	O	S		
2	A	1	60	40	4	6	9	1	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	1	1	1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
4	A	1	5	4	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	191	191	191	0	0

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

3 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.19Å 59.01Å 60.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.49 – 1.40	Depositor
% Data completeness (in resolution range)	98.9 (33.49-1.40)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.26 (at 1.40Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.176 , 0.197	Depositor
Wilson B-factor (Å^2)	10.3	Xtriage
Anisotropy	0.322	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.015 for -h,l,k	Xtriage
Total number of atoms	1613	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.07% 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.

4 Model quality (i)

4.1 Standard geometry (i)

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

4.2 Too-close contacts (i)

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

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

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

4.3.2 Protein sidechains (i)

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

4.3.3 RNA (i)

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

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

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	$\#\|Z\| > 2$
4	SO4	A	1203	-	4,4,4	0.11	0	6,6,6	0.33	0
2	L9P	A	1201	-	59,66,66	1.37	7 (11%)	78,105,105	1.90	12 (15%)

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	L9P	A	1201	-	1/1/18/21	4/64/110/110	0/6/7/7

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1201	L9P	C37-N36	4.01	1.36	1.30
2	A	1201	L9P	S12-N11	3.84	1.66	1.60
2	A	1201	L9P	C09-C10	-3.20	1.48	1.53
2	A	1201	L9P	C04-N03	2.57	1.49	1.46
2	A	1201	L9P	C38-C39	2.47	1.41	1.37

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1201	L9P	C10-C09-N08	7.89	124.54	116.06
2	A	1201	L9P	C10-N11-S12	-7.72	111.60	124.08
2	A	1201	L9P	C57-C13-C20	5.12	131.16	117.35
2	A	1201	L9P	C57-C13-C19	4.35	129.08	117.35
2	A	1201	L9P	O15-S12-O14	-3.62	113.22	120.57

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	1201	L9P	N03

All (4) torsion outliers are listed below:

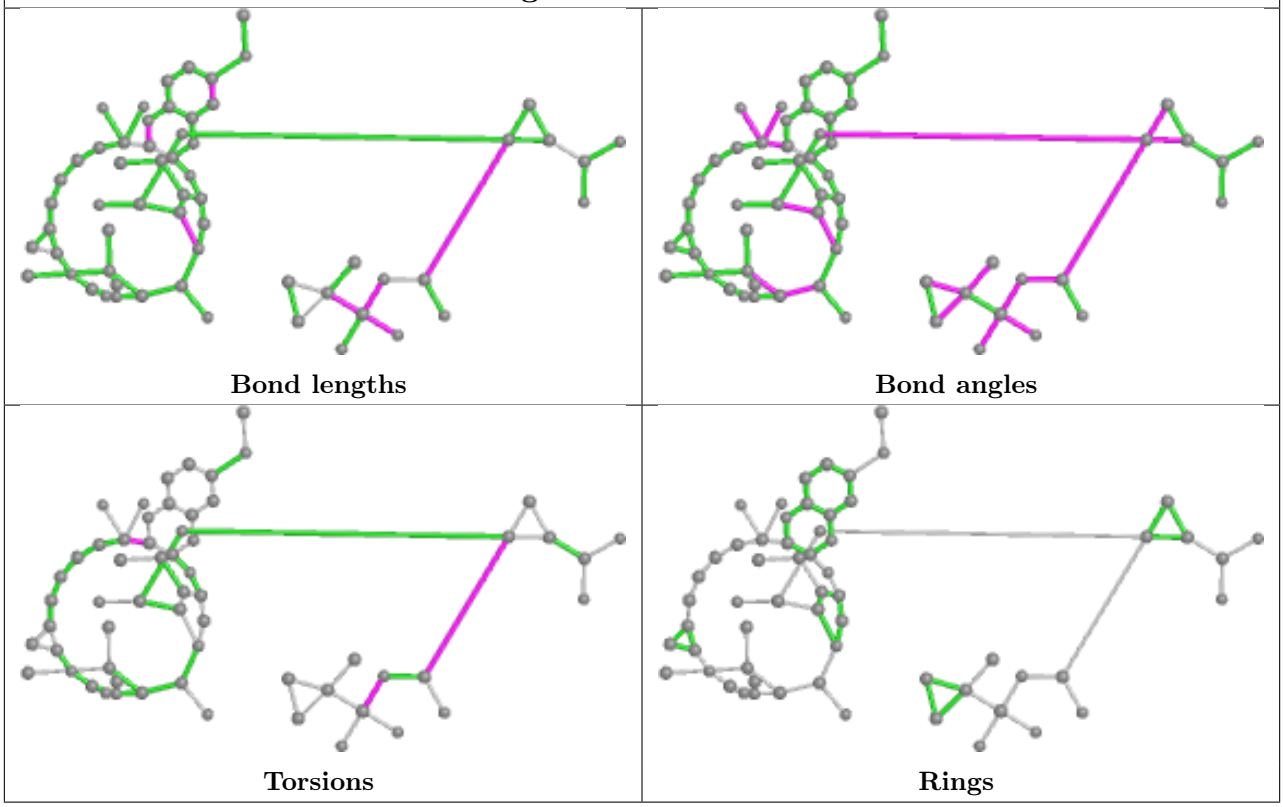
Mol	Chain	Res	Type	Atoms
2	A	1201	L9P	N36-C37-C42-F60
2	A	1201	L9P	C10-N11-S12-C13
2	A	1201	L9P	C22-C37-C42-F60
2	A	1201	L9P	N08-C09-C10-O16

There are no ring outliers.

No monomer is involved in short contacts.

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 L9P A 1201



4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 Protein, DNA and RNA chains (i)

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

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

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

5.3 Carbohydrates (i)

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

5.4 Ligands (i)

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

5.5 Other polymers (i)

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