



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:10 PM GMT

PDB ID : 1NS3
Title : STRUCTURE OF HCV PROTEASE (BK STRAIN)
Authors : Yan, Y.; Munshi, S.; Chen, Z.
Deposited on : 1997-04-05
Resolution : 2.80 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

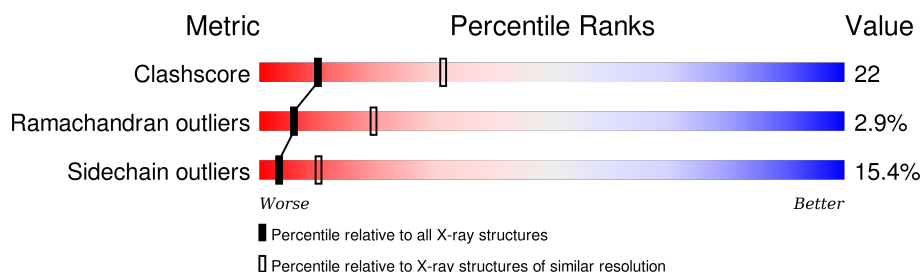
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	186	
1	B	186	
2	C	14	
2	D	14	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2814 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 NS3 PROTEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	178	Total	C	N	O	S	1	0	0
			1312	815	238	249	10			
1	B	178	Total	C	N	O	S	1	0	0
			1312	815	238	249	10			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	PRO	DELETION	UNP P26663
A	66	GLY	ALA	CONFLICT	UNP P26663
A	86	GLN	PRO	CONFLICT	UNP P26663
A	87	ALA	LYS	CONFLICT	UNP P26663
A	147	SER	PHE	CONFLICT	UNP P26663
B	?	-	PRO	DELETION	UNP P26663
B	66	GLY	ALA	CONFLICT	UNP P26663
B	86	GLN	PRO	CONFLICT	UNP P26663
B	87	ALA	LYS	CONFLICT	UNP P26663
B	147	SER	PHE	CONFLICT	UNP P26663

- Molecule 2 is a protein called NS4A PEPTIDE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	13	Total	C	N	O	9	0	0
			93	61	17	15			
2	D	13	Total	C	N	O	9	0	0
			93	61	17	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	GLY	DELETION	UNP P26663
D	?	-	GLY	DELETION	UNP P26663

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

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

- Molecule 4 is water.

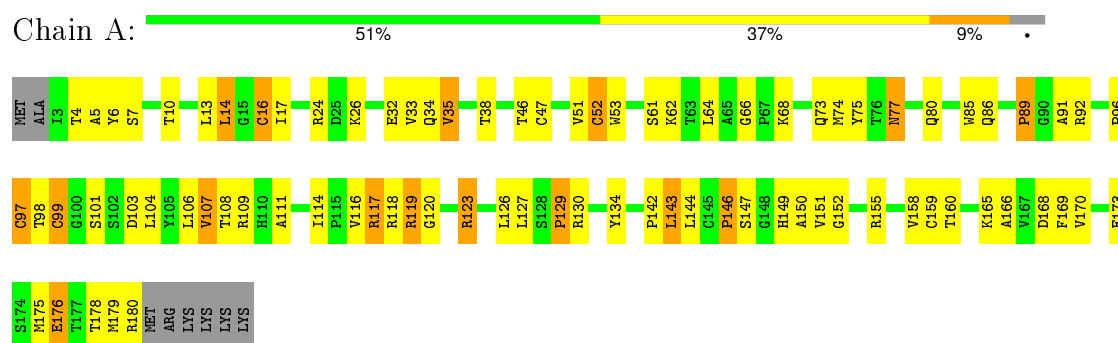
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	O 1	0	0
4	B	1	Total 1	O 1	0	0

3 Residue-property plots

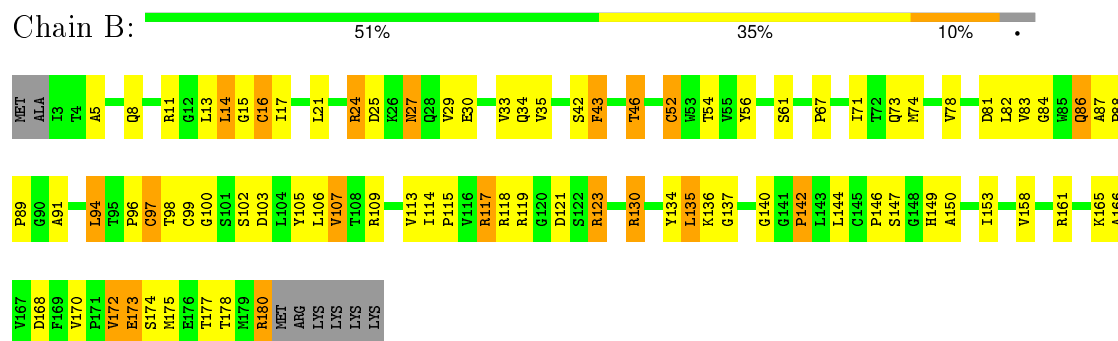
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

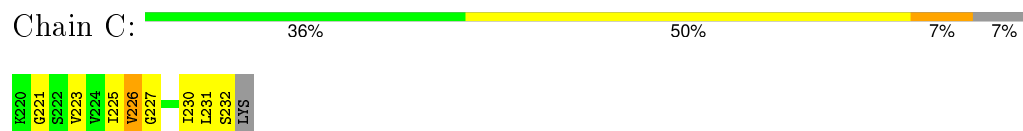
• Molecule 1: NS3 PROTEASE



• Molecule 1: NS3 PROTEASE



• Molecule 2: NS4A PEPTIDE



• Molecule 2: NS4A PEPTIDE





4 Data and refinement statistics

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

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α , β , γ	96.96 Å 96.96 Å 167.10 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	(Not available) – 2.80	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.80)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.204 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2814	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	0.56	1/1337 (0.1%)	0.83	1/1820 (0.1%)
1	B	0.58	1/1337 (0.1%)	0.82	1/1820 (0.1%)
2	C	0.58	0/92	0.85	0/122
2	D	0.60	0/92	0.92	0/122
All	All	0.57	2/2858 (0.1%)	0.83	2/3884 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	16	CYS	CB-SG	7.24	1.94	1.82
1	A	16	CYS	CB-SG	5.28	1.91	1.82

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	94	LEU	CA-CB-CG	5.73	128.48	115.30
1	A	14	LEU	CA-CB-CG	5.38	127.67	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	134	TYR	Sidechain

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	1312	0	1320	56	14
1	B	1312	0	1319	67	19
2	C	93	0	112	10	0
2	D	93	0	112	9	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
All	All	2814	0	2863	124	21

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 22.

The worst 5 of 124 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)
1:A:119:ARG:HA	1:A:119:ARG:HH11	1.37	0.89
1:B:98:THR:HG23	1:B:118:ARG:HH12	1.45	0.80
1:B:34:GLN:HE21	2:D:225:ILE:HG21	1.46	0.80
1:B:105:TYR:HB3	1:B:113:VAL:HG12	1.67	0.77
1:B:172:VAL:HG23	1:B:175:MET:HE3	1.67	0.76

The worst 5 of 21 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)
1:A:103:ASP:OD2	1:B:119:ARG:NH2[2_765]	0.50	1.70
1:B:180:ARG:NH1	1:B:180:ARG:NH2[8_675]	0.66	1.54
1:A:103:ASP:OD1	1:B:119:ARG:CZ[2_765]	0.83	1.37
1:A:103:ASP:OD1	1:B:119:ARG:NH1[2_765]	0.88	1.32

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:103:ASP:CG	1:B:119:ARG:NH2[2_765]	0.90	1.30

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	176/186 (95%)	156 (89%)	13 (7%)	7 (4%)	4	12
1	B	176/186 (95%)	156 (89%)	16 (9%)	4 (2%)	8	26
2	C	11/14 (79%)	11 (100%)	0	0	100	100
2	D	11/14 (79%)	10 (91%)	1 (9%)	0	100	100
All	All	374/400 (94%)	333 (89%)	30 (8%)	11 (3%)	6	19

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	CYS
1	A	147	SER
1	A	89	PRO
1	A	99	CYS
1	A	129	PRO

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	145/152 (95%)	123 (85%)	22 (15%)	3	10
1	B	145/152 (95%)	120 (83%)	25 (17%)	2	7
2	C	11/12 (92%)	10 (91%)	1 (9%)	12	33
2	D	11/12 (92%)	11 (100%)	0	100	100
All	All	312/328 (95%)	264 (85%)	48 (15%)	3	10

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	176	GLU
1	B	33	VAL
1	B	173	GLU
1	B	14	LEU
1	B	21	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	77	ASN
1	B	86	GLN
1	B	34	GLN
1	A	34	GLN
1	A	80	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.