



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 11:45 PM GMT

PDB ID : 1YKS
Title : Crystal structure of yellow fever virus NS3 helicase
Authors : Wu, J.; Bera, A.K.; Kuhn, R.J.; Smith, J.L.
Deposited on : 2005-01-18
Resolution : 1.80 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
<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) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

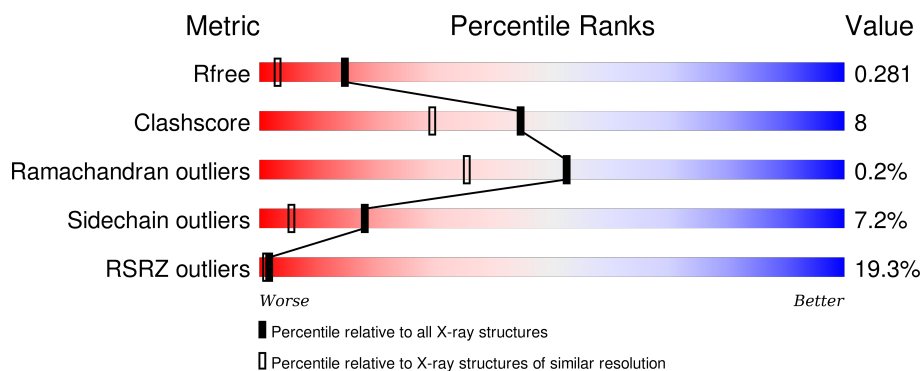
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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))
R_{free}	91344	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

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.

Mol	Chain	Length	Quality of chain
1	A	440	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3643 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 Genome polyprotein [contains: Flavivirin protease NS3 catalytic subunit].

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	431	Total	C	N	O	S	0	0	0
			3382	2115	626	619	22			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	184	GLY	-	CLONING ARTIFACT	UNP P19901
A	185	SER	-	CLONING ARTIFACT	UNP P19901
A	186	HIS	-	CLONING ARTIFACT	UNP P19901

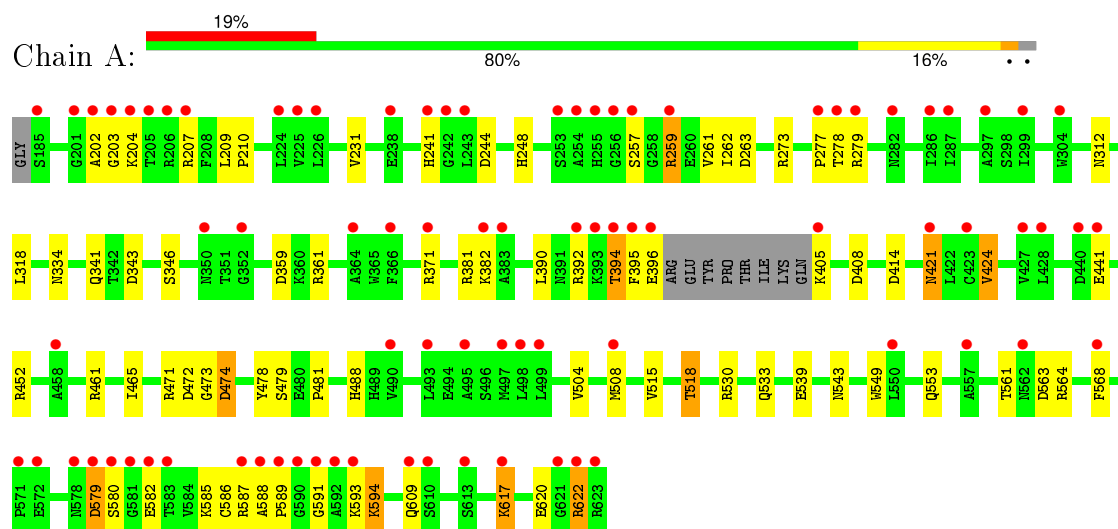
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	261	Total	O	0	0
			261	261		

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 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 ($\text{RSRZ} > 2$). 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.

- Molecule 1: Genome polypeptide [contains: Flavivirin protease NS3 catalytic subunit]



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.02Å 43.80Å 93.07Å 90.00° 104.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.80 20.87 – 1.79	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-1.80) 97.3 (20.87-1.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.28 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.193 , 0.238 0.251 , 0.281	Depositor DCC
R_{free} test set	1985 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	23.9	Xtriage
Anisotropy	0.170	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.48 , 70.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 39561 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3643	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.43% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.57	0/3458	0.77	8/4678 (0.2%)

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	563	ASP	CB-CG-OD2	6.14	123.83	118.30
1	A	579	ASP	CB-CG-OD2	5.82	123.54	118.30
1	A	472	ASP	CB-CG-OD2	5.35	123.11	118.30
1	A	343	ASP	CB-CG-OD2	5.23	123.01	118.30
1	A	461	ARG	NE-CZ-NH2	-5.16	117.72	120.30
1	A	474	ASP	CB-CG-OD2	5.10	122.89	118.30
1	A	414	ASP	CB-CG-OD2	5.10	122.89	118.30
1	A	408	ASP	CB-CG-OD2	5.07	122.87	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	3382	0	3357	53	0
2	A	261	0	0	12	0
All	All	3643	0	3357	53	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 8.

All (53) 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:508:MET:SD	2:A:839:HOH:O	2.35	0.84
1:A:204:LYS:HD2	2:A:809:HOH:O	1.82	0.80
1:A:564:ARG:HH22	1:A:622:ARG:HH21	1.27	0.78
1:A:568:PHE:CE1	1:A:588:ALA:HA	2.19	0.78
1:A:568:PHE:CZ	1:A:588:ALA:HA	2.19	0.77
1:A:203:GLY:O	1:A:207:ARG:HB2	1.85	0.77
1:A:479:SER:HB3	2:A:802:HOH:O	1.90	0.70
1:A:421:ASN:H	1:A:421:ASN:HD22	1.40	0.69
1:A:504:VAL:HG21	2:A:875:HOH:O	1.94	0.68
1:A:568:PHE:CZ	1:A:588:ALA:CA	2.80	0.64
1:A:539:GLU:HG2	1:A:561:THR:HG21	1.79	0.63
1:A:346:SER:HB2	2:A:807:HOH:O	1.98	0.62
1:A:361:ARG:NH2	1:A:473:GLY:O	2.32	0.62
1:A:568:PHE:CE2	1:A:588:ALA:HB2	2.36	0.61
1:A:202:ALA:HB2	1:A:334:ASN:HD21	1.66	0.60
1:A:488:HIS:CD2	1:A:488:HIS:H	2.19	0.60
1:A:394:THR:HG22	1:A:395:PHE:HD2	1.66	0.58
1:A:421:ASN:H	1:A:421:ASN:ND2	2.02	0.58
1:A:381:ARG:HD3	2:A:674:HOH:O	2.08	0.52
1:A:515:VAL:O	1:A:518:THR:HB	2.10	0.52
1:A:568:PHE:CD2	1:A:588:ALA:HB2	2.46	0.51
1:A:424:VAL:HG12	1:A:465:ILE:CG2	2.41	0.51
1:A:371:ARG:NE	2:A:769:HOH:O	2.44	0.50
1:A:568:PHE:CD2	1:A:594:LYS:HB2	2.47	0.50
1:A:549:TRP:O	1:A:553:GLN:HG2	2.12	0.49
1:A:278:THR:OG1	2:A:866:HOH:O	2.20	0.49
1:A:471:ARG:HD2	1:A:474:ASP:OD1	2.14	0.48
1:A:277:PRO:HD3	2:A:832:HOH:O	2.13	0.48
1:A:539:GLU:CD	1:A:543:ASN:HD22	2.17	0.48
1:A:209:LEU:HB3	1:A:210:PRO:HD3	1.95	0.47
1:A:588:ALA:HB1	1:A:589:PRO:CD	2.45	0.47
1:A:568:PHE:CZ	1:A:588:ALA:HB2	2.50	0.47
1:A:568:PHE:CE1	1:A:588:ALA:CA	2.95	0.46
1:A:248:HIS:HE1	1:A:263:ASP:OD1	1.98	0.46
1:A:359:ASP:OD1	1:A:361:ARG:HD3	2.16	0.46
1:A:587:ARG:HG3	1:A:591:GLY:HA2	1.98	0.45
1:A:530:ARG:HG2	1:A:533:GLN:OE1	2.16	0.45
1:A:568:PHE:CE1	1:A:588:ALA:CB	3.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:261:VAL:HG13	1:A:262:ILE:HG13	2.00	0.44
1:A:588:ALA:HB1	1:A:589:PRO:HD2	2.00	0.43
1:A:341:GLN:HG3	1:A:478:TYR:CE1	2.54	0.42
1:A:568:PHE:CZ	1:A:588:ALA:CB	3.03	0.42
1:A:394:THR:HG22	1:A:395:PHE:CD2	2.52	0.42
1:A:488:HIS:CD2	2:A:877:HOH:O	2.73	0.41
1:A:203:GLY:O	1:A:207:ARG:CB	2.64	0.41
1:A:587:ARG:CG	1:A:591:GLY:HA2	2.50	0.41
1:A:568:PHE:CE1	1:A:589:PRO:HD3	2.55	0.41
1:A:564:ARG:HA	1:A:564:ARG:HD3	1.89	0.41
1:A:452:ARG:CZ	1:A:481:PRO:HG2	2.50	0.41
1:A:617:LYS:HD2	1:A:622:ARG:O	2.22	0.40
1:A:586:CYS:HB2	1:A:620:GLU:OE1	2.20	0.40
1:A:259:ARG:NH1	2:A:811:HOH:O	2.40	0.40
1:A:231:VAL:HG23	2:A:797:HOH:O	2.20	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	427/440 (97%)	406 (95%)	20 (5%)	1 (0%)	52	35

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	257	SER

5.3.2 Protein sidechains ⓘ

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	359/368 (98%)	333 (93%)	26 (7%)	18 5

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	241	HIS
1	A	244	ASP
1	A	259	ARG
1	A	273	ARG
1	A	279	ARG
1	A	312	ASN
1	A	318	LEU
1	A	382	LYS
1	A	390	LEU
1	A	392	ARG
1	A	394	THR
1	A	396	GLU
1	A	405	LYS
1	A	421	ASN
1	A	424	VAL
1	A	441	GLU
1	A	518	THR
1	A	579	ASP
1	A	580	SER
1	A	582	GLU
1	A	585	LYS
1	A	593	LYS
1	A	594	LYS
1	A	609	GLN
1	A	617	LYS
1	A	622	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	199	HIS
1	A	248	HIS
1	A	334	ASN
1	A	341	GLN
1	A	391	ASN
1	A	421	ASN
1	A	488	HIS

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](#)

There are no ligands in this entry.

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 ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	431/440 (97%)	1.08	83 (19%) 2 1	5, 17, 38, 69	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	395	PHE	13.7
1	A	394	THR	9.0
1	A	255	HIS	7.7
1	A	396	GLU	7.4
1	A	242	GLY	6.7
1	A	623	ARG	6.5
1	A	393	LYS	6.0
1	A	277	PRO	5.7
1	A	592	ALA	5.5
1	A	256	GLY	5.3
1	A	591	GLY	5.0
1	A	392	ARG	4.7
1	A	254	ALA	4.6
1	A	241	HIS	4.6
1	A	279	ARG	4.5
1	A	257	SER	4.4
1	A	582	GLU	4.3
1	A	405	LYS	4.2
1	A	580	SER	4.1
1	A	621	GLY	4.1
1	A	590	GLY	4.0
1	A	202	ALA	3.8
1	A	286	ILE	3.8
1	A	205	THR	3.7
1	A	589	PRO	3.6
1	A	579	ASP	3.6
1	A	622	ARG	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	287	ILE	3.5
1	A	581	GLY	3.4
1	A	562	ASN	3.4
1	A	382	LYS	3.3
1	A	499	LEU	3.3
1	A	204	LYS	3.3
1	A	498	LEU	3.2
1	A	225	VAL	3.2
1	A	428	LEU	3.2
1	A	617	LYS	3.2
1	A	203	GLY	3.1
1	A	224	LEU	3.1
1	A	259	ARG	3.0
1	A	568	PHE	3.0
1	A	383	ALA	2.9
1	A	587	ARG	2.9
1	A	495	ALA	2.9
1	A	366	PHE	2.9
1	A	497	MET	2.8
1	A	207	ARG	2.8
1	A	243	LEU	2.8
1	A	226	LEU	2.8
1	A	364	ALA	2.7
1	A	352	GLY	2.7
1	A	572	GLU	2.7
1	A	253	SER	2.6
1	A	458	ALA	2.6
1	A	371	ARG	2.6
1	A	350	ASN	2.5
1	A	282	ASN	2.5
1	A	440	ASP	2.5
1	A	588	ALA	2.5
1	A	238	GLU	2.5
1	A	508	MET	2.4
1	A	427	VAL	2.4
1	A	613	SER	2.4
1	A	278	THR	2.4
1	A	441	GLU	2.4
1	A	550	LEU	2.4
1	A	593	LYS	2.4
1	A	185	SER	2.4
1	A	610	SER	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	423	CYS	2.3
1	A	206	ARG	2.3
1	A	297	ALA	2.3
1	A	490	VAL	2.2
1	A	493	LEU	2.2
1	A	609	GLN	2.2
1	A	583	THR	2.2
1	A	304	TRP	2.1
1	A	201	GLY	2.1
1	A	299	ILE	2.1
1	A	571	PRO	2.1
1	A	421	ASN	2.0
1	A	557	ALA	2.0
1	A	578	ASN	2.0

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.