



Full wwPDB X-ray Structure Validation Report ⓘ

May 30, 2016 – 10:52 AM EDT

PDB ID : 5HUW
Title : Structure of HSV-1 Large Terminase NLS bound to importin alpha
Authors : Sankhala, R.S.; Lokareddy, R.K.; Cingolani, G.
Deposited on : 2016-01-27
Resolution : 1.95 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 : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027674
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) : rb-20027674

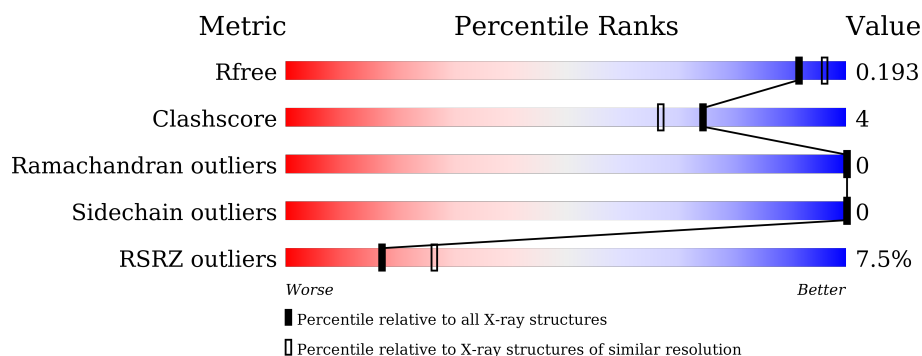
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	C	528	<div> <div>5%</div> <div>74%</div> <div>6%</div> <div>20%</div> </div>
2	A	12	<div> <div>8%</div> <div>75%</div> <div>8%</div> <div>17%</div> </div>
2	B	12	<div> <div>33%</div> <div>33%</div> <div>17%</div> <div>50%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3691 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 Importin subunit alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	423	Total	C	N	O	S	0	1	0
			3236	2061	550	615	10			

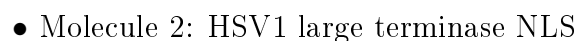
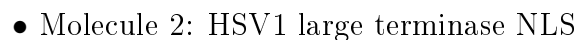
- Molecule 2 is a protein called HSV1 large terminase NLS.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	10	Total	C	N	O	0	0	0
			76	48	16	12			
2	B	6	Total	C	N	O	0	0	0
			48	31	11	6			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	317	Total	O	0	0
			317	317		
3	A	13	Total	O	0	0
			13	13		
3	B	1	Total	O	0	0
			1	1		

- Molecule 1: Importin subunit alpha-1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	78.26 Å 90.73 Å 97.66 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.94 – 1.95 45.36 – 1.94	Depositor EDS
% Data completeness (in resolution range)	99.7 (14.94-1.95) 93.4 (45.36-1.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.27 (at 1.94 Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.173 , 0.195 0.171 , 0.193	Depositor DCC
R_{free} test set	1865 reflections (3.86%)	DCC
Wilson B-factor (Å ²)	30.4	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3691	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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	C	0.40	0/3295	0.55	0/4490
2	A	0.36	0/77	0.56	0/101
2	B	0.33	0/49	0.65	0/63
All	All	0.39	0/3421	0.55	0/4654

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	C	3236	0	3307	25	0
2	A	76	0	86	3	0
2	B	48	0	58	2	0
3	A	13	0	0	0	0
3	B	1	0	0	1	0
3	C	317	0	0	4	0
All	All	3691	0	3451	27	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 4.

All (27) 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:B:186:ARG:NH1	3:B:201:HOH:O	2.11	0.82
1:C:366:ARG:NH2	3:C:601:HOH:O	2.12	0.82
1:C:263:ASN:HB3	3:C:638:HOH:O	1.93	0.68
1:C:469:GLY:O	1:C:473:ILE:HG13	1.96	0.64
1:C:479:HIS:HB3	1:C:484:VAL:HG21	1.80	0.64
1:C:238:ARG:NE	2:A:184:LYS:HE3	2.16	0.60
1:C:475:ALA:O	1:C:478:ARG:HG3	2.06	0.56
1:C:386:LEU:HD21	1:C:425:LEU:HD13	1.89	0.55
1:C:238:ARG:CZ	2:A:184:LYS:HE3	2.37	0.54
1:C:479:HIS:HE1	1:C:481:ASN:HB3	1.73	0.53
3:C:819:HOH:O	2:B:187:ALA:HA	2.08	0.53
1:C:479:HIS:CE1	1:C:481:ASN:HB3	2.47	0.50
1:C:488:SER:O	1:C:492:ILE:HG12	2.13	0.49
1:C:282:PRO:HD2	1:C:285:ARG:HG3	1.96	0.48
1:C:239:ASN:OD1	2:A:184:LYS:NZ	2.47	0.48
1:C:438:GLN:NE2	1:C:483:SER:HB3	2.29	0.47
1:C:475:ALA:O	1:C:478:ARG:CG	2.62	0.47
1:C:207:ASP:OD1	1:C:251:GLN:NE2	2.48	0.47
1:C:285:ARG:HD2	3:C:714:HOH:O	2.16	0.45
1:C:272:CYS:HB3	1:C:312:PRO:HB2	1.99	0.44
1:C:116:ILE:HG12	1:C:121:ILE:HD11	1.99	0.43
1:C:253:LEU:O	1:C:257:VAL:HG23	2.20	0.42
1:C:273:TRP:CD2	1:C:312:PRO:HB3	2.53	0.42
1:C:474:GLU:OE1	1:C:474:GLU:HA	2.19	0.42
1:C:470:LEU:O	1:C:474:GLU:HG2	2.20	0.41
1:C:480:GLU:OE1	1:C:481:ASN:HA	2.21	0.41
1:C:470:LEU:HD13	1:C:496:PHE:CD1	2.57	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	C	422/528 (80%)	412 (98%)	10 (2%)	0	100	100
2	A	8/12 (67%)	8 (100%)	0	0	100	100
2	B	4/12 (33%)	3 (75%)	1 (25%)	0	100	100
All	All	434/552 (79%)	423 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	C	356/448 (80%)	356 (100%)	0	100	100
2	A	8/9 (89%)	8 (100%)	0	100	100
2	B	5/9 (56%)	5 (100%)	0	100	100
All	All	369/466 (79%)	369 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

There are no ligands in this entry.

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

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	C	423/528 (80%)	0.20	28 (6%) 22 31	23, 34, 93, 144	0
2	A	10/12 (83%)	0.06	1 (10%) 9 15	30, 43, 60, 88	0
2	B	6/12 (50%)	4.38	4 (66%) 0 0	76, 88, 116, 118	0
All	All	439/552 (79%)	0.26	33 (7%) 17 26	23, 34, 94, 144	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	183	PRO	8.6
2	B	182	PRO	6.7
1	C	485	TYR	6.1
1	C	478	ARG	5.7
1	C	483	SER	5.7
1	C	486	LYS	5.5
1	C	474	GLU	5.5
1	C	476	LEU	5.3
1	C	481	ASN	5.1
1	C	482	GLU	5.0
1	C	475	ALA	4.9
2	B	187	ALA	4.6
1	C	108	LYS	4.6
1	C	489	LEU	4.2
1	C	496	PHE	4.1
1	C	479	HIS	4.1
1	C	432	LYS	3.9
1	C	484	VAL	3.6
1	C	493	GLU	3.6
1	C	480	GLU	3.5
2	B	186	ARG	3.4
1	C	434	THR	3.3
1	C	477	GLN	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	487	ALA	3.0
1	C	437	ILE	2.9
1	C	491	LEU	2.9
2	A	190	ASP	2.6
1	C	473	ILE	2.5
1	C	107	GLU	2.4
1	C	436	ILE	2.3
1	C	75	ASN	2.2
1	C	490	ASN	2.1
1	C	471	ASP	2.1

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.