



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 11, 2017 – 04:28 AM EDT

PDB ID : 5HUY  
Title : Structure of HCMV Small Terminase NLS bound to importin alpha  
Authors : Sankhala, R.S.; Lokareddy, R.K.; Cingolani, G.  
Deposited on : unknown  
Resolution : 1.98 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20029824

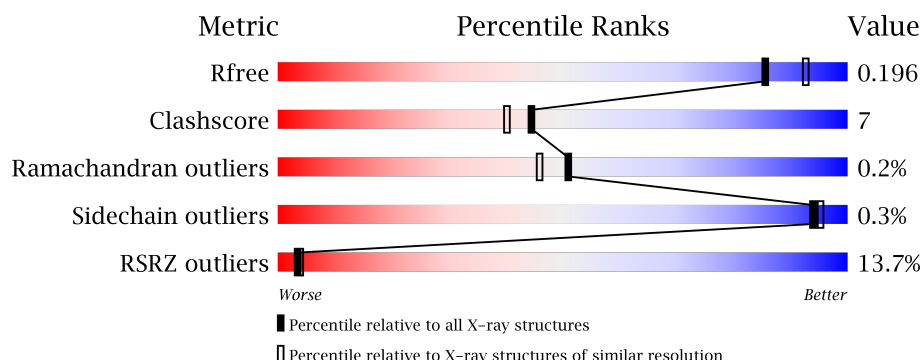
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.98 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))
$R_{free}$	100719	9293 (2.00-1.96)
Clashscore	112137	10621 (2.00-1.96)
Ramachandran outliers	110173	10502 (2.00-1.96)
Sidechain outliers	110143	10501 (2.00-1.96)
RSRZ outliers	101464	9395 (2.00-1.96)

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  $\geq 3$ , 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>10%</div> <div> <div></div> <div>71%</div> <div>9%</div> <div>20%</div> </div> </div>
2	A	16	<div> <div>13%</div> <div> <div></div> <div>56%</div> <div>44%</div> </div> </div>
2	B	16	<div> <div>19%</div> <div> <div></div> <div>31%</div> <div>13%</div> <div>56%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3560 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	0	0
			3226	2055	547	614	10			

- Molecule 2 is a protein called HCMV small terminase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	7	Total	C	N	O	0	0	0
			61	36	17	8			
2	A	9	Total	C	N	O	0	0	0
			77	45	22	10			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	185	Total	O	0	0
			185	185		
3	B	2	Total	O	0	0
			2	2		
3	A	9	Total	O	0	0
			9	9		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.25Å 90.81Å 96.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.93 – 1.98 41.06 – 1.98	Depositor EDS
% Data completeness (in resolution range)	99.7 (14.93-1.98) 94.1 (41.06-1.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 1.98Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, $R_{free}$	0.173 , 0.196 0.174 , 0.196	Depositor DCC
$R_{free}$ test set	1886 reflections (4.16%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.1	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 59.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3560	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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	0.44	0/3284	0.59	0/4475
2	A	0.91	0/77	1.05	0/100
2	B	0.31	0/61	0.61	0/79
All	All	0.45	0/3422	0.61	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	3226	0	3301	50	0
2	A	77	0	88	0	0
2	B	61	0	70	2	0
3	A	9	0	0	0	0
3	B	2	0	0	1	0
3	C	185	0	0	5	0
All	All	3560	0	3459	51	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 7.

All (51) 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:C:474:GLU:OE2	1:C:492:ILE:HD13	1.70	0.91
1:C:474:GLU:OE1	1:C:492:ILE:HD11	1.78	0.82
1:C:486:LYS:HG3	1:C:487:ALA:H	1.51	0.75
1:C:291:LYS:HE2	3:C:626:HOH:O	1.88	0.73
1:C:404:TYR:CD2	1:C:416:LEU:HD22	2.25	0.71
1:C:207:ASP:OD1	1:C:251:GLN:NE2	2.23	0.70
1:C:108:LYS:HD3	1:C:109:GLN:HG3	1.76	0.68
1:C:330:LYS:N	1:C:330:LYS:HD2	2.10	0.66
1:C:488:SER:O	1:C:492:ILE:HD12	1.96	0.65
1:C:474:GLU:OE1	1:C:492:ILE:CD1	2.45	0.65
1:C:409:THR:O	1:C:413:ILE:HD12	1.98	0.64
1:C:370:ILE:HG23	1:C:404:TYR:CE1	2.33	0.63
1:C:483:SER:O	1:C:486:LYS:HG2	1.99	0.63
1:C:287:GLU:HG2	1:C:327:GLN:NE2	2.14	0.61
1:C:486:LYS:HG3	1:C:487:ALA:N	2.14	0.61
1:C:370:ILE:CG2	1:C:404:TYR:CE1	2.84	0.60
1:C:487:ALA:HA	1:C:490:ASN:HB3	1.84	0.60
1:C:287:GLU:HG2	1:C:327:GLN:HE21	1.67	0.60
1:C:404:TYR:HD2	1:C:416:LEU:HD22	1.67	0.58
1:C:411:GLU:N	1:C:411:GLU:OE2	2.33	0.56
1:C:477:GLN:HG2	1:C:488:SER:HB3	1.88	0.55
1:C:489:LEU:O	1:C:493:GLU:HG3	2.06	0.55
1:C:197:ARG:NH1	1:C:238:ARG:O	2.38	0.54
1:C:291:LYS:NZ	3:C:606:HOH:O	2.41	0.54
1:C:457:THR:HG23	1:C:458:GLU:N	2.23	0.53
1:C:388:LYS:NZ	3:C:605:HOH:O	2.41	0.52
1:C:474:GLU:CD	1:C:492:ILE:HD13	2.29	0.52
1:C:458:GLU:O	1:C:461:SER:OG	2.20	0.51
2:B:822:ARG:HD2	3:B:901:HOH:O	2.09	0.51
1:C:404:TYR:HE2	1:C:416:LEU:HB2	1.76	0.50
1:C:404:TYR:CE2	1:C:416:LEU:HB2	2.47	0.48
1:C:370:ILE:CG2	1:C:404:TYR:HE1	2.27	0.48
1:C:474:GLU:CD	1:C:492:ILE:CD1	2.83	0.47
1:C:370:ILE:HG23	1:C:404:TYR:CD1	2.50	0.47
1:C:458:GLU:HA	1:C:461:SER:OG	2.15	0.47
1:C:238:ARG:HB3	3:C:629:HOH:O	2.15	0.46
1:C:434:THR:HG21	1:C:481:ASN:ND2	2.30	0.46
1:C:397:ALA:O	1:C:401:ILE:HG12	2.16	0.46
1:C:116:ILE:HG12	1:C:121:ILE:HD11	1.98	0.45
1:C:372:GLN:HG3	3:C:632:HOH:O	2.18	0.44
1:C:434:THR:O	1:C:438:GLN:HG3	2.19	0.43
1:C:458:GLU:C	1:C:461:SER:HG	2.18	0.43

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:457:THR:CG2	1:C:458:GLU:N	2.81	0.43
1:C:480:GLU:HG3	1:C:481:ASN:N	2.35	0.42
1:C:354:GLU:OE2	2:B:826:ARG:NH2	2.36	0.42
1:C:366:ARG:HB2	1:C:368:ASP:OD1	2.19	0.42
1:C:370:ILE:HG23	1:C:404:TYR:HE1	1.78	0.41
1:C:435:LYS:O	1:C:439:VAL:HG23	2.20	0.41
1:C:404:TYR:CE2	1:C:416:LEU:HD22	2.55	0.41
1:C:241:ASN:HA	1:C:242:PRO:HA	1.75	0.41
1:C:272:CYS:HB3	1:C:312:PRO:HB2	2.03	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	421/528 (80%)	415 (99%)	5 (1%)	1 (0%)	51	46
2	A	7/16 (44%)	7 (100%)	0	0	100	100
2	B	5/16 (31%)	5 (100%)	0	0	100	100
All	All	433/560 (77%)	427 (99%)	5 (1%)	1 (0%)	51	46

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	239	ASN

### 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	C	355/448 (79%)	354 (100%)	1 (0%)	94	95
2	A	7/14 (50%)	7 (100%)	0	100	100
2	B	6/14 (43%)	6 (100%)	0	100	100
All	All	368/476 (77%)	367 (100%)	1 (0%)	94	95

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	461	SER

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	C	423/528 (80%)	0.76	55 (13%) 4 5	30, 44, 111, 141	0
2	A	9/16 (56%)	0.67	2 (22%) 1 1	39, 49, 81, 82	0
2	B	7/16 (43%)	1.47	3 (42%) 0 0	56, 64, 90, 103	0
All	All	439/560 (78%)	0.76	60 (13%) 3 4	30, 44, 111, 141	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	496	PHE	9.6
1	C	476	LEU	9.2
1	C	485	TYR	8.7
1	C	497	SER	7.8
1	C	475	ALA	7.5
1	C	484	VAL	7.2
1	C	489	LEU	7.0
1	C	491	LEU	6.9
2	B	820	ALA	5.4
1	C	486	LYS	5.3
1	C	487	ALA	5.2
1	C	480	GLU	5.1
1	C	495	TYR	4.9
1	C	481	ASN	4.8
1	C	493	GLU	4.8
1	C	478	ARG	4.7
1	C	482	GLU	4.5
1	C	432	LYS	4.4
1	C	474	GLU	4.3
1	C	470	LEU	4.3
1	C	454	LEU	4.1
1	C	431	ALA	4.1
1	C	455	GLY	3.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	108	LYS	3.8
1	C	492	ILE	3.8
1	C	479	HIS	3.8
1	C	109	GLN	3.7
1	C	75	ASN	3.2
1	C	458	GLU	3.0
1	C	488	SER	3.0
2	A	828	ALA	3.0
1	C	159	VAL	2.8
1	C	477	GLN	2.8
1	C	490	ASN	2.8
1	C	107	GLU	2.8
1	C	164	ILE	2.5
1	C	168	ILE	2.5
1	C	200	VAL	2.5
1	C	106	ARG	2.4
2	B	822	ARG	2.4
1	C	452	GLU	2.3
1	C	419	CYS	2.3
1	C	199	LEU	2.3
1	C	238	ARG	2.3
1	C	483	SER	2.3
1	C	382	LEU	2.3
1	C	437	ILE	2.3
1	C	167	PHE	2.2
1	C	462	ILE	2.2
1	C	433	ASP	2.2
1	C	186	LEU	2.1
2	B	821	THR	2.1
1	C	472	LYS	2.1
1	C	494	LYS	2.1
1	C	163	ALA	2.1
1	C	383	VAL	2.0
2	A	820	ALA	2.0
1	C	468	GLY	2.0
1	C	450	ALA	2.0
1	C	291	LYS	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 [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.