Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 03:25 PM EDT

PDB ID	:	6T3X
Title	:	Crystal structure of the truncated human cytomegalovirus pUL50-pUL53 complex
Authors	:	Muller, Y.A.
Deposited on	:	2019-10-11
Resolution	:	1.48 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ 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 ⓘ) were used in the production of this report:

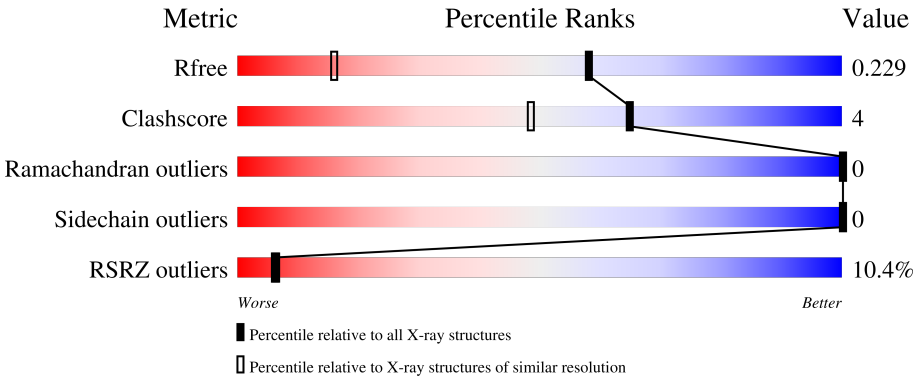
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.48 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}	130704	4690 (1.50-1.46)
Clashscore	141614	4955 (1.50-1.46)
Ramachandran outliers	138981	4846 (1.50-1.46)
Sidechain outliers	138945	4844 (1.50-1.46)
RSRZ outliers	127900	4614 (1.50-1.46)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	214	4% 87% 12%
1	C	214	15% 80% 9% 11%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6790 atoms, of which 3308 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 Nuclear egress protein 2,Nuclear egress protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	212	Total	C	H	N	O	S	0	9	0
			3456	1086	1754	285	313	18			
1	C	191	Total	C	H	N	O	S	0	2	0
			3087	977	1554	258	284	14			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP P16791
A	-4	SER	-	expression tag	UNP P16791
A	-3	HIS	-	expression tag	UNP P16791
A	-2	MET	-	expression tag	UNP P16791
A	-1	ALA	-	expression tag	UNP P16791
A	0	SER	-	expression tag	UNP P16791
A	1051	GLY	-	linker	UNP P16791
A	1052	GLY	-	linker	UNP P16791
A	1053	SER	-	linker	UNP P16791
A	1054	GLY	-	linker	UNP P16791
A	1055	SER	-	linker	UNP P16791
A	1056	GLY	-	linker	UNP P16791
A	1057	GLY	-	linker	UNP P16791
A	1058	SER	-	linker	UNP P16791
C	-5	GLY	-	expression tag	UNP P16791
C	-4	SER	-	expression tag	UNP P16791
C	-3	HIS	-	expression tag	UNP P16791
C	-2	MET	-	expression tag	UNP P16791
C	-1	ALA	-	expression tag	UNP P16791
C	0	SER	-	expression tag	UNP P16791
C	1051	GLY	-	linker	UNP P16791
C	1052	GLY	-	linker	UNP P16791
C	1053	SER	-	linker	UNP P16791
C	1054	GLY	-	linker	UNP P16791
C	1055	SER	-	linker	UNP P16791

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	1056	GLY	-	linker	UNP P16791
C	1057	GLY	-	linker	UNP P16791
C	1058	SER	-	linker	UNP P16791

- Molecule 2 is water.

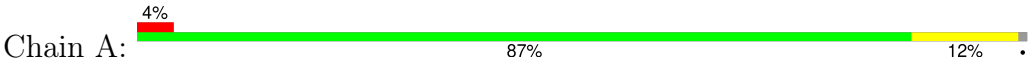
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	187	Total O 187 187	0	0
2	C	60	Total O 60 60	0	0

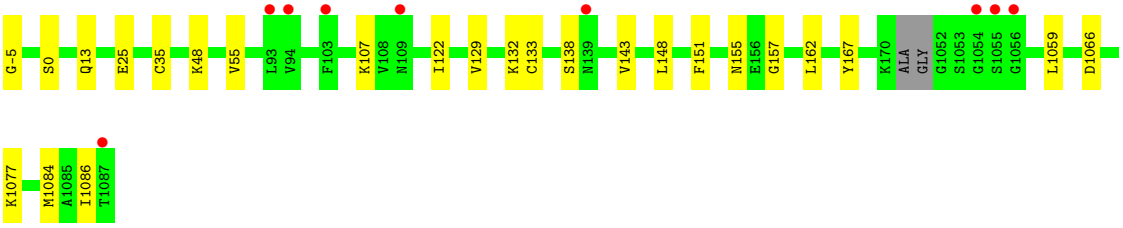
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

• Molecule 1: Nuclear egress protein 2,Nuclear egress protein 1

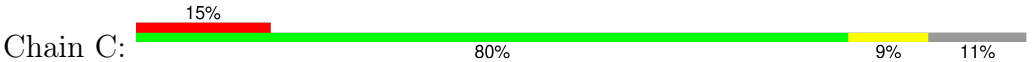
Chain A:

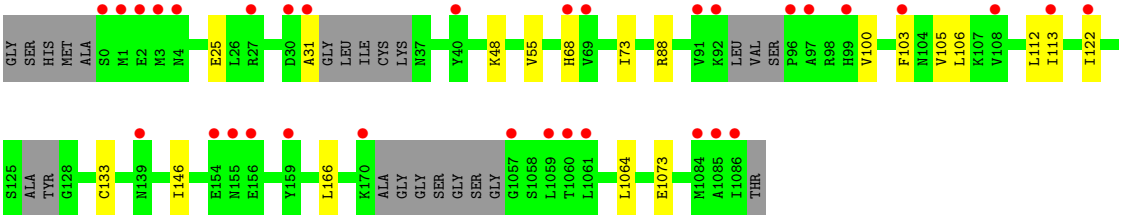




• Molecule 1: Nuclear egress protein 2,Nuclear egress protein 1

Chain C:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	37.27Å 82.57Å 63.66Å 90.00° 95.10° 90.00°	Depositor
Resolution (Å)	41.28 – 1.48 41.28 – 1.48	Depositor EDS
% Data completeness (in resolution range)	98.6 (41.28-1.48) 98.6 (41.28-1.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 1.48Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.192 , 0.225 0.195 , 0.229	Depositor DCC
R_{free} test set	5121 reflections (8.11%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.218	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 50.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6790	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.71% 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.

5 Model quality

5.1 Standard geometry

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	0.34	0/1759	0.57	0/2373
1	C	0.31	0/1564	0.49	0/2107
All	All	0.33	0/3323	0.53	0/4480

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

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	1702	1754	1754	16	0
1	C	1533	1554	1556	12	0
2	A	187	0	0	4	1
2	C	60	0	0	1	0
All	All	3482	3308	3310	27	1

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)
1:C:31:ALA:O	2:C:501:HOH:O	2.09	0.71

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:100:VAL:HG12	1:C:112:LEU:HD21	1.74	0.69
1:C:100:VAL:O	1:C:122:ILE:HD11	1.97	0.65
1:C:166:LEU:HD13	1:C:1064:LEU:HD22	1.82	0.60
1:A:107:LYS:NZ	2:A:504:HOH:O	2.30	0.59
1:A:1084:MET:HE2	2:A:580:HOH:O	2.03	0.58
1:C:73:ILE:HG12	1:C:113:ILE:HD13	1.88	0.55
1:A:162:LEU:HD13	1:A:1077:LYS:HE3	1.87	0.55
1:A:-5:GLY:N	1:A:35:CYS:SG	2.75	0.54
1:C:106:LEU:HD22	1:C:146:ILE:CD1	2.38	0.54
1:A:1066[A]:ASP:OD1	2:A:501:HOH:O	2.19	0.52
1:A:155:ASN:OD1	1:A:157:GLY:N	2.46	0.48
1:A:0:SER:OG	1:C:1073[A]:GLU:OE1	2.18	0.46
1:A:122:ILE:HD13	1:A:148[B]:LEU:HD12	1.98	0.45
1:A:25:GLU:OE2	1:A:48:LYS:HE3	2.15	0.45
1:A:138:SER:HB3	1:A:143:VAL:O	2.17	0.45
1:C:103:PHE:HE1	1:C:105:VAL:HG23	1.81	0.45
1:A:55[B]:VAL:HG21	1:A:133[B]:CYS:SG	2.58	0.44
1:A:13:GLN:HG3	2:A:637:HOH:O	2.17	0.44
1:C:103:PHE:CE1	1:C:105:VAL:HG23	2.53	0.44
1:C:55:VAL:HG21	1:C:133[A]:CYS:SG	2.57	0.44
1:A:167:TYR:CE1	1:A:1059:LEU:HG	2.54	0.43
1:A:129:VAL:HG22	1:A:1086:ILE:CD1	2.48	0.42
1:C:68:HIS:HB2	1:C:88:ARG:HG2	2.01	0.42
1:C:25:GLU:OE1	1:C:48:LYS:HE3	2.19	0.42
1:A:133[B]:CYS:SG	1:A:148[B]:LEU:CD2	3.07	0.42
1:A:132:LYS:HG3	1:A:151:PHE:CD1	2.55	0.41

All (1) 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)
2:A:518:HOH:O	2:A:647:HOH:O[2_646]	2.05	0.15

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	217/214 (101%)	214 (99%)	3 (1%)	0	100	100
1	C	183/214 (86%)	182 (100%)	1 (0%)	0	100	100
All	All	400/428 (94%)	396 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	197/188 (105%)	197 (100%)	0	100	100
1	C	176/188 (94%)	176 (100%)	0	100	100
All	All	373/376 (99%)	373 (100%)	0	100	100

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

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

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 monosaccharides 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^{th} 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	212/214 (99%)	0.36	9 (4%) 36 39	18, 32, 54, 104	0
1	C	191/214 (89%)	0.88	33 (17%) 1 1	28, 47, 90, 126	0
All	All	403/428 (94%)	0.61	42 (10%) 6 6	18, 39, 80, 126	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1061	LEU	8.9
1	C	96	PRO	8.4
1	A	94	VAL	6.8
1	C	31	ALA	6.4
1	C	97	ALA	6.1
1	C	103	PHE	5.8
1	A	1087	THR	5.5
1	C	40	TYR	5.3
1	C	1057	GLY	5.1
1	C	1	MET	4.4
1	C	99	HIS	4.4
1	C	1084	MET	3.9
1	C	108	VAL	3.7
1	C	1086	ILE	3.5
1	C	3	MET	3.3
1	C	1085	ALA	3.2
1	C	0	SER	3.2
1	C	30	ASP	3.1
1	C	1059	LEU	3.1
1	A	1055	SER	2.9
1	A	109	ASN	2.8
1	A	1054	GLY	2.8
1	C	27	ARG	2.7
1	C	2	GLU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	159	TYR	2.6
1	C	1060	THR	2.6
1	C	113	ILE	2.5
1	C	155	ASN	2.4
1	A	1056	GLY	2.4
1	C	170	LYS	2.4
1	C	139	ASN	2.2
1	C	156	GLU	2.2
1	C	91	VAL	2.2
1	C	122	ILE	2.2
1	A	139	ASN	2.2
1	C	4	ASN	2.2
1	C	69	VAL	2.2
1	C	154	GLU	2.2
1	A	103	PHE	2.1
1	A	93	LEU	2.1
1	C	92	LYS	2.0
1	C	68	HIS	2.0

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 monosaccharides 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.