



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 22, 2024 – 12:15 pm GMT

PDB ID : 8QDO  
Title : Crystal structure of the tegument protein UL82 (pp71) from Human Cytomegalovirus  
Authors : Bresch, I.P.; Eberhage, J.; Reubold, T.F.; Eschenburg, S.  
Deposited on : 2023-08-30  
Resolution : 2.70 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

<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  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

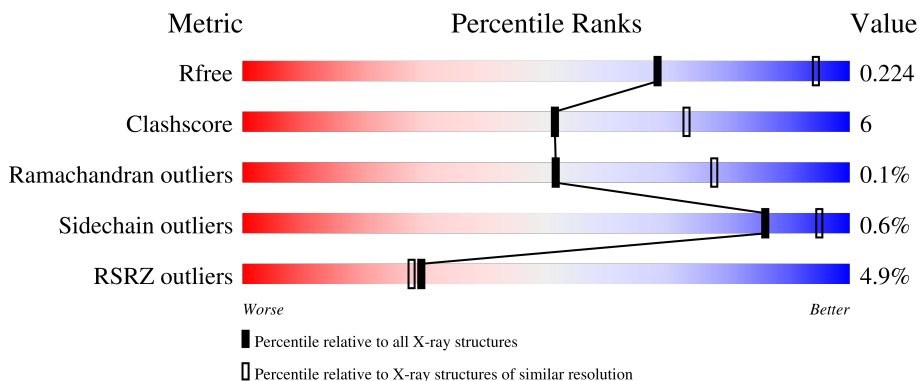
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	564	 3% 68% 9% 22%
1	B	564	 4% 69% 10% 21%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7075 atoms, of which 18 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 Protein pp71.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	438	Total	C	N	O	S	0	0	0
			3469	2189	647	625	8			
1	B	445	Total	C	N	O	S	0	0	0
			3532	2225	661	637	9			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP P06726
A	-3	HIS	-	expression tag	UNP P06726
A	-2	MET	-	expression tag	UNP P06726
A	-1	GLY	-	expression tag	UNP P06726
A	0	SER	-	expression tag	UNP P06726
B	-4	GLY	-	expression tag	UNP P06726
B	-3	HIS	-	expression tag	UNP P06726
B	-2	MET	-	expression tag	UNP P06726
B	-1	GLY	-	expression tag	UNP P06726
B	0	SER	-	expression tag	UNP P06726

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	31	8	18	5	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	26	Total	O	0	0
			26	26		
3	B	17	Total	O	0	0
			17	17		



SER  
THR  
PRO  
ARG  
SER  
ARG  
PRO  
ARG  
ILE

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.61Å 137.44Å 275.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.61 – 2.70 48.61 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.61-2.70) 100.0 (48.61-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.207 , 0.231 0.200 , 0.224	Depositor DCC
$R_{free}$ test set	2380 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	85.3	Xtrriage
Anisotropy	0.237	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 53.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7075	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PG4

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	0.32	0/3554	0.60	0/4840
1	B	0.29	0/3617	0.58	0/4924
All	All	0.30	0/7171	0.59	0/9764

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	A	3469	0	3464	43	0
1	B	3532	0	3525	35	0
2	A	13	18	18	3	0
3	A	26	0	0	0	0
3	B	17	0	0	0	0
All	All	7057	18	7007	78	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 6.

All (78) 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:B:154:ARG:HB3	1:B:155:PRO:HD2	1.40	1.01
1:A:147:ARG:HH22	1:A:155:PRO:HD3	1.31	0.95
1:B:85:LEU:HD12	1:B:85:LEU:H	1.43	0.84
1:B:43:VAL:HG21	1:B:114:LEU:HD21	1.60	0.82
1:B:83:VAL:HG21	1:B:382:ARG:NH1	1.97	0.80
1:A:370:TRP:HB3	1:A:490:VAL:HG21	1.70	0.73
1:A:147:ARG:NH2	1:A:155:PRO:HD3	2.05	0.72
1:A:516:ARG:HB2	1:A:525:THR:HB	1.72	0.71
1:A:34:CYS:HB3	1:A:524:ARG:HG2	1.76	0.66
1:B:154:ARG:HB3	1:B:155:PRO:CD	2.23	0.64
1:A:137:VAL:HG21	1:A:283:PHE:CZ	2.33	0.64
1:A:287:ILE:HD11	1:A:353:LEU:HD23	1.80	0.64
1:B:43:VAL:CG2	1:B:114:LEU:HD21	2.29	0.63
1:B:213:LEU:CD2	1:B:258:VAL:HG22	2.30	0.61
1:A:317:ILE:HD13	1:A:365:PHE:CD1	2.36	0.61
1:A:137:VAL:HG12	1:A:265:VAL:HB	1.84	0.59
1:A:155:PRO:O	1:A:156:ARG:HG2	2.02	0.59
1:A:142:LEU:HD21	1:A:213:LEU:HD11	1.86	0.58
1:B:293:LEU:HD11	1:B:343:VAL:HG12	1.85	0.58
1:A:118:SER:O	1:A:119:GLU:HG2	2.05	0.57
1:A:293:LEU:HD11	1:A:343:VAL:HG12	1.87	0.57
1:A:370:TRP:HB3	1:A:490:VAL:CG2	2.35	0.57
1:B:510:ARG:HB3	1:B:511:PRO:HD2	1.87	0.57
1:A:137:VAL:HG21	1:A:283:PHE:CE1	2.41	0.55
1:B:287:ILE:HD11	1:B:359:LEU:HD13	1.88	0.55
1:A:147:ARG:HD2	1:A:236:TRP:HZ2	1.72	0.54
1:B:537:PRO:HB2	1:B:539:PRO:HD2	1.91	0.53
1:A:301:ILE:H	2:A:601:PG4:C8	2.21	0.53
1:A:43:VAL:CG2	1:A:114:LEU:HD21	2.38	0.53
1:A:48:LEU:HG	1:A:108:ASN:HB2	1.90	0.52
1:B:117:GLY:O	1:B:118:SER:HB3	2.10	0.52
1:B:213:LEU:HD22	1:B:258:VAL:HG22	1.92	0.52
1:A:49:GLU:HB2	1:A:52:ARG:HG3	1.93	0.51
1:B:61:ASN:OD1	1:B:101:ARG:HG2	2.11	0.51
1:B:374:PRO:HG2	1:B:397:VAL:HA	1.91	0.51
1:A:39:LEU:HB2	1:A:125:VAL:HG12	1.93	0.51
1:B:156:ARG:NH2	1:B:158:GLU:HG3	2.26	0.51
1:B:34:CYS:HB3	1:B:524:ARG:HG2	1.92	0.50
1:A:243:HIS:CE1	1:A:249:PRO:HD3	2.47	0.50
1:A:59:ARG:HD3	1:A:508:ALA:HB2	1.94	0.49
1:A:137:VAL:HG23	1:A:277:ARG:HH22	1.77	0.49
1:B:382:ARG:HH11	1:B:382:ARG:HG3	1.77	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:32:LEU:HD11	1:B:524:ARG:NH1	2.28	0.48
1:B:317:ILE:HD13	1:B:332:PRO:HA	1.95	0.48
1:A:393:PRO:O	1:A:395:SER:N	2.47	0.46
1:A:137:VAL:CG2	1:A:277:ARG:HH12	2.29	0.46
1:B:58:LEU:HD23	1:B:102:LEU:HB3	1.99	0.45
1:B:376:LYS:HG2	1:B:489:HIS:CD2	2.52	0.45
1:B:372:LEU:HD12	1:B:489:HIS:O	2.17	0.45
1:A:156:ARG:O	1:A:156:ARG:HG3	2.17	0.44
1:B:178:PRO:HG2	1:B:180:TRP:NE1	2.32	0.44
1:A:219:SER:HB3	1:A:253:PHE:HB3	1.99	0.44
1:A:279:PRO:O	1:A:280:ASP:HB2	2.17	0.44
1:B:327:SER:HB2	1:B:385:VAL:HG13	1.99	0.44
1:A:301:ILE:H	2:A:601:PG4:H82	1.82	0.44
1:A:180:TRP:CD1	1:A:248:PRO:HB2	2.52	0.44
1:B:49:GLU:OE1	1:B:113:ARG:HG3	2.16	0.44
1:A:137:VAL:HG22	1:A:277:ARG:HH12	1.82	0.44
1:B:138:SER:HB2	1:B:266:VAL:HG12	1.99	0.43
1:B:295:ARG:HH11	1:B:295:ARG:HG2	1.82	0.43
1:A:194:PHE:HB2	1:A:243:HIS:HB3	2.00	0.43
1:A:43:VAL:HG21	1:A:114:LEU:HD21	1.99	0.43
1:A:142:LEU:HD22	1:A:230:THR:HG23	2.00	0.43
1:A:142:LEU:HD22	1:A:230:THR:CG2	2.49	0.42
1:A:318:PHE:CZ	1:A:360:GLY:HA3	2.54	0.42
1:B:488:TRP:HB2	1:B:490:VAL:HG22	2.00	0.42
1:A:49:GLU:HB2	1:A:52:ARG:CG	2.48	0.42
1:B:59:ARG:NH1	1:B:505:GLN:OE1	2.53	0.42
1:A:61:ASN:OD1	1:A:101:ARG:HG2	2.18	0.42
1:B:62:ILE:O	1:B:99:GLN:HA	2.20	0.42
1:B:32:LEU:HD11	1:B:524:ARG:CZ	2.50	0.42
1:B:227:LYS:HD3	1:B:269:HIS:CD2	2.55	0.41
1:A:32:LEU:HD12	1:A:32:LEU:C	2.41	0.41
1:B:243:HIS:CE1	1:B:249:PRO:HD3	2.55	0.41
1:B:62:ILE:CG2	1:B:500:LEU:HD22	2.50	0.40
1:A:327:SER:HB2	1:A:385:VAL:HG13	2.03	0.40
1:A:43:VAL:HG23	1:A:114:LEU:HD21	2.02	0.40
1:A:301:ILE:HB	2:A:601:PG4:H81	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	A	432/564 (77%)	414 (96%)	17 (4%)	1 (0%)	47	73
1	B	439/564 (78%)	422 (96%)	17 (4%)	0	100	100
All	All	871/1128 (77%)	836 (96%)	34 (4%)	1 (0%)	51	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	119	GLU

### 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	383/488 (78%)	381 (100%)	2 (0%)	88	96
1	B	390/488 (80%)	387 (99%)	3 (1%)	81	93
All	All	773/976 (79%)	768 (99%)	5 (1%)	86	95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	125	VAL
1	A	255	HIS
1	B	179	ASN
1	B	255	HIS
1	B	337	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	226	HIS
1	A	242	HIS
1	B	242	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	PG4	A	601	-	12,12,12	0.52	0	11,11,11	0.25	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PG4	A	601	-	-	7/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	PG4	O2-C3-C4-O3
2	A	601	PG4	C3-C4-O3-C5
2	A	601	PG4	O1-C1-C2-O2
2	A	601	PG4	C8-C7-O4-C6
2	A	601	PG4	O4-C7-C8-O5
2	A	601	PG4	C4-C3-O2-C2
2	A	601	PG4	O3-C5-C6-O4

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	PG4	3	0

## 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, 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	A	438/564 (77%)	0.11	19 (4%) 35 33	54, 72, 108, 146	0
1	B	445/564 (78%)	0.28	24 (5%) 25 24	58, 81, 130, 161	1 (0%)
All	All	883/1128 (78%)	0.19	43 (4%) 29 28	54, 77, 122, 161	1 (0%)

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	13	SER	6.2
1	B	398	MET	5.9
1	B	399	ARG	5.5
1	A	539	PRO	5.1
1	A	115	THR	4.8
1	B	156	ARG	4.8
1	B	539	PRO	4.8
1	A	147	ARG	4.3
1	A	233	ARG	4.2
1	B	153	ASN	3.9
1	B	114	LEU	3.8
1	B	516	ARG	3.6
1	A	116	PRO	3.6
1	B	397	VAL	3.4
1	B	479	GLY	3.4
1	B	45	GLY	3.3
1	B	152	GLU	3.2
1	A	156	ARG	3.0
1	B	155	PRO	3.0
1	B	47	SER	2.9
1	B	395	SER	2.9
1	B	15	GLU	2.8
1	A	119	GLU	2.7
1	B	44	GLU	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	45	GLY	2.6
1	A	118	SER	2.6
1	B	48	LEU	2.5
1	A	13	SER	2.5
1	B	119	GLU	2.4
1	A	47	SER	2.4
1	B	120	ASN	2.4
1	B	14	SER	2.4
1	A	52	ARG	2.4
1	B	168	HIS	2.4
1	A	117	GLY	2.3
1	A	198	PHE	2.2
1	A	396	ASN	2.2
1	B	530	SER	2.1
1	A	236	TRP	2.1
1	A	112	ARG	2.1
1	B	19	ILE	2.1
1	A	479	GLY	2.1
1	A	114	LEU	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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	PG4	A	601	13/13	0.89	0.16	75,90,101,101	31

## 6.5 Other polymers [i](#)

There are no such residues in this entry.