



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 11:43 AM BST

PDB ID : 6LSF  
Title : Crystal structure of the enterovirus 71 polymerase elongation complex (C2S6RA/C2S6RB form)  
Authors : Wang, M.; Shu, B.; Jing, X.; Gong, P.  
Deposited on : 2020-01-17  
Resolution : 2.15 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

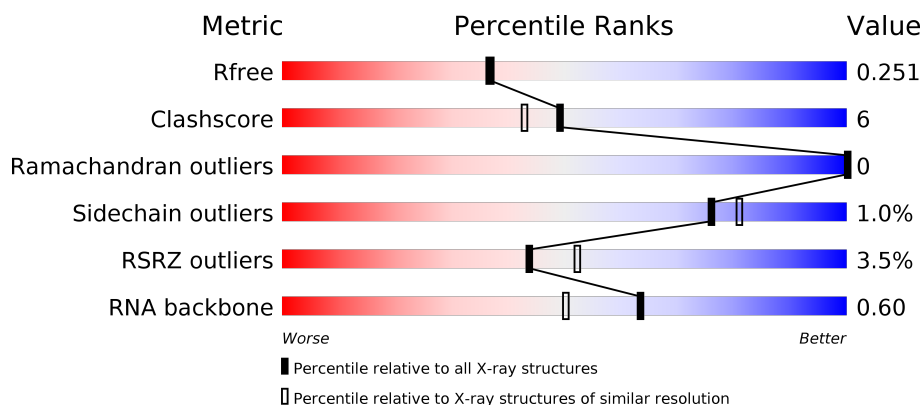
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)
RNA backbone	3102	1060 (2.60-1.72)

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 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	468	<div> <div>3%</div> <div> <div></div> <div>90%</div> <div>9%</div> </div> </div>
2	B	35	<div> <div>6%</div> <div>23%</div> <div>71%</div> </div>
3	C	18	<div> <div>11%</div> <div>50%</div> <div>50%</div> </div>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 4329 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	463	Total	C	N	O	S	0	1	0
			3629	2326	602	683	18			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	MET	CYS	engineered mutation	UNP E5RPG3
A	463	HIS	-	expression tag	UNP E5RPG3
A	464	HIS	-	expression tag	UNP E5RPG3
A	465	HIS	-	expression tag	UNP E5RPG3
A	466	HIS	-	expression tag	UNP E5RPG3
A	467	HIS	-	expression tag	UNP E5RPG3
A	468	HIS	-	expression tag	UNP E5RPG3

- Molecule 2 is a RNA chain called RNA (35-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	10	Total	C	N	O	P	0	0	0
			208	93	33	72	10			

- Molecule 3 is a RNA chain called RNA (5'-R(\*UP\*GP\*UP\*UP\*CP\*GP\*AP\*CP\*GP\*AP\*GP\*AP\*GP\*AP\*GP\*AP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	P	0	9	0
			394	176	82	118	18			

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Zn	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	85	Total	O	0	0
			85	85		
6	B	6	Total	O	0	0
			6	6		
6	C	1	Total	O	0	0
			1	1		



- Molecule 1: Genome polyprotein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.35Å 76.45Å 155.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.73 – 2.15 47.85 – 2.15	Depositor EDS
% Data completeness (in resolution range)	89.5 (28.73-2.15) 89.5 (47.85-2.15)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.82 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.14	Depositor
R, $R_{free}$	0.215 , 0.251 0.215 , 0.251	Depositor DCC
$R_{free}$ test set	1847 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.9	Xtriage
Anisotropy	0.773	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4329	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 3.80% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SO4

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	0.43	0/3727	0.57	0/5070
2	B	0.70	0/230	1.20	0/355
3	C	0.39	0/442	0.85	0/686
All	All	0.45	0/4399	0.66	0/6111

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	3629	0	3479	25	0
2	B	208	0	108	12	0
3	C	394	0	200	28	0
4	A	5	0	0	0	0
5	A	1	0	0	0	0
6	A	85	0	0	2	0
6	B	6	0	0	0	0
6	C	1	0	0	0	0
All	All	4329	0	3787	46	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 (46) 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:328:GLY:HA3	3:C:701[B]:C:H4'	1.46	0.96
2:B:604:C:H42	3:C:697[B]:A:H61	1.28	0.79
2:B:605:U:H3	3:C:696[B]:G:H1	1.35	0.72
1:A:17:ILE:HG21	1:A:155:SER:HB3	1.73	0.70
1:A:166:ASP:OD2	1:A:170:LYS:NZ	2.28	0.66
1:A:277:ARG:NH1	6:A:601:HOH:O	2.28	0.66
3:C:700[B]:C:O2'	3:C:701[B]:C:H5'	1.97	0.64
2:B:606:C:H42	3:C:695[B]:A:H61	1.46	0.61
2:B:603:U:H3	3:C:698[B]:G:H1	1.49	0.60
1:A:91:ASP:OD1	1:A:93:SER:HB3	2.01	0.60
1:A:114:THR:HG21	2:B:601:G:OP2	2.05	0.57
1:A:427:LYS:NZ	1:A:452:GLU:OE1	2.36	0.57
2:B:604:C:H42	3:C:697[A]:A:H61	1.52	0.56
2:B:608:C:N4	3:C:693[A]:A:H61	2.03	0.56
1:A:376:LYS:HD2	3:C:700[A]:C:H5''	1.87	0.56
1:A:328:GLY:HA3	3:C:701[B]:C:C4'	2.29	0.55
1:A:100:ALA:O	1:A:188:ARG:NH2	2.37	0.55
2:B:608:C:H41	3:C:693[A]:A:H61	1.55	0.54
2:B:602:G:H22	3:C:699[B]:A:H2	1.56	0.54
1:A:421:LEU:HD11	3:C:698[B]:G:O2'	2.10	0.51
3:C:698[A]:G:H2'	3:C:699[A]:A:C8	2.46	0.51
1:A:438[B]:ARG:NH1	6:A:602:HOH:O	2.34	0.49
1:A:259:SER:O	1:A:263:VAL:HG23	2.13	0.48
1:A:330:ASP:HB2	3:C:701[B]:C:OP1	2.12	0.48
1:A:42:VAL:HG23	1:A:168:ILE:HG23	1.96	0.48
2:B:602:G:N2	3:C:699[B]:A:C2	2.81	0.47
1:A:123:LEU:HB2	1:A:125:ILE:HD12	1.96	0.46
1:A:274:HIS:HB2	1:A:281:TYR:CZ	2.50	0.46
1:A:12:THR:HG21	1:A:280:THR:HB	1.98	0.46
2:B:607:U:N3	3:C:695[A]:A:N1	2.66	0.44
2:B:604:C:C2	3:C:698[B]:G:N2	2.85	0.44
1:A:222:PRO:HA	1:A:369:TRP:CZ2	2.53	0.43
3:C:696[A]:G:H2'	3:C:697[A]:A:O4'	2.18	0.43
3:C:698[B]:G:H2'	3:C:699[B]:A:C8	2.54	0.43
3:C:695[B]:A:C2	3:C:696[B]:G:C4	3.06	0.43
1:A:58:LEU:HD13	1:A:175:LEU:HD21	2.01	0.42
1:A:186:TYR:OH	1:A:261:GLU:OE1	2.31	0.42
3:C:697[B]:A:H2'	3:C:698[B]:G:C8	2.54	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:41:ALA:HB2	1:A:163:ARG:HG3	2.00	0.42
1:A:159:LYS:HZ1	3:C:701[B]:C:H42	1.68	0.41
3:C:698[B]:G:C6	3:C:699[B]:A:C6	3.08	0.41
1:A:450:ASN:ND2	1:A:452:GLU:OE2	2.53	0.41
3:C:694[A]:G:H3'	3:C:695[A]:A:H5''	2.03	0.40
1:A:238:ASP:OD2	3:C:701[B]:C:H2'	2.21	0.40
3:C:699[B]:A:O5'	3:C:699[B]:A:H8	2.04	0.40
3:C:700[B]:C:C2'	3:C:701[B]:C:H5'	2.51	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	462/468 (99%)	452 (98%)	10 (2%)	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	386/412 (94%)	382 (99%)	4 (1%)	76	81

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	93	SER
1	A	147	LYS
1	A	218	TRP
1	A	327	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	457	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	9/35 (25%)	0	0
3	C	0/18	-	-
All	All	9/53 (16%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

## 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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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	$\# Z  > 2$
4	SO4	A	501	-	4,4,4	0.29	0	6,6,6	0.30	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	463/468 (98%)	0.13	15 (3%) 47 56	36, 50, 74, 89	0
2	B	10/35 (28%)	-0.39	0 100 100	45, 52, 83, 102	0
3	C	9/18 (50%)	0.72	2 (22%) 0 0	49, 54, 96, 103	0
All	All	482/521 (92%)	0.13	17 (3%) 44 52	36, 50, 78, 103	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	134	THR	4.0
1	A	143	PHE	3.5
1	A	20	PRO	3.4
1	A	122	ALA	3.0
1	A	15	LEU	3.0
1	A	123	LEU	2.8
1	A	132	ASP	2.7
1	A	135	THR	2.4
3	C	693[A]	A	2.3
3	C	694[A]	G	2.2
1	A	209	ALA	2.2
1	A	277	ARG	2.2
1	A	113	HIS	2.1
1	A	327	TYR	2.1
1	A	295	SER	2.1
1	A	279	LYS	2.0
1	A	147	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 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( $\text{\AA}^2$ )	Q<0.9
5	ZN	A	502	1/1	0.98	0.17	44,44,44,44	1
4	SO4	A	501	5/5	0.99	0.18	40,44,51,57	5

### 6.5 Other polymers [i](#)

There are no such residues in this entry.