



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 10:32 PM BST

PDB ID : 4KQ0  
Title : Crystal structure of double-helical CGG-repetitive RNA 19mer complexed with RSS p19  
Authors : Cabo, A.; Katorcha, E.; Tamjar, J.; Popov, A.N.; Malinina, L.  
Deposited on : 2013-05-14  
Resolution : 2.10 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

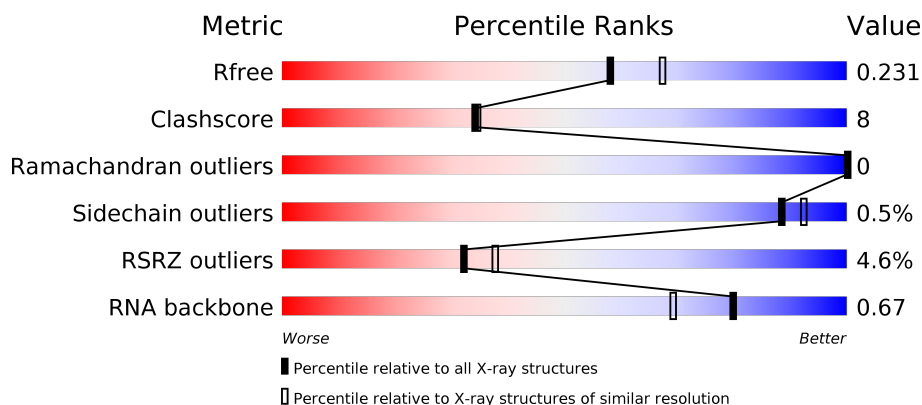
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)
RNA backbone	3102	1000 (2.54-1.66)

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	135	<div> <div>5%</div> <div> <div></div> <div>76%</div> <div>15%</div> <div>10%</div> </div> </div>
1	D	135	<div> <div>4%</div> <div> <div></div> <div>75%</div> <div>16%</div> <div>10%</div> </div> </div>
2	B	19	<div> <div>68%</div> <div>26%</div> <div>5%</div> </div>
2	E	19	<div> <div>68%</div> <div>26%</div> <div>5%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	D	303	-	-	-	X



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 RNA silencing suppressor p19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	122	Total 985	C 623	N 175	O 181	S 6	0	0	0
1	D	122	Total 985	C 623	N 175	O 181	S 6	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP P69517
A	2	SER	-	EXPRESSION TAG	UNP P69517
A	3	HIS	-	EXPRESSION TAG	UNP P69517
A	4	MET	-	EXPRESSION TAG	UNP P69517
A	122	MET	LEU	ENGINEERED MUTATION	UNP P69517
A	125	MET	LEU	ENGINEERED MUTATION	UNP P69517
D	1	GLY	-	EXPRESSION TAG	UNP P69517
D	2	SER	-	EXPRESSION TAG	UNP P69517
D	3	HIS	-	EXPRESSION TAG	UNP P69517
D	4	MET	-	EXPRESSION TAG	UNP P69517
D	122	MET	LEU	ENGINEERED MUTATION	UNP P69517
D	125	MET	LEU	ENGINEERED MUTATION	UNP P69517

- Molecule 2 is a RNA chain called 5'-R(P\*GP\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*GP\*GP\*CP\*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	19	Total 417	C 183	N 81	O 134	P 19	0	0	0
2	E	19	Total 417	C 183	N 81	O 134	P 19	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	115	Total	O	0	0
			115	115		
4	B	19	Total	O	0	0
			19	19		
4	D	115	Total	O	0	0
			115	115		
4	E	18	Total	O	0	0
			18	18		



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.93Å 90.93Å 147.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 2.10 33.41 – 2.10	Depositor EDS
% Data completeness (in resolution range)	94.9 (15.00-2.10) 49.7 (33.41-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.19 (at 2.10Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.179 , 0.227 0.190 , 0.231	Depositor DCC
$R_{free}$ test set	656 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtriage
Anisotropy	0.614	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.499 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3101	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.91% 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: 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	1.00	0/1011	0.92	1/1363 (0.1%)
1	D	1.00	0/1011	0.92	1/1363 (0.1%)
2	B	0.74	1/466 (0.2%)	0.98	1/726 (0.1%)
2	E	0.74	1/466 (0.2%)	0.95	1/726 (0.1%)
All	All	0.93	2/2954 (0.1%)	0.94	4/4178 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	G	OP3-P	-10.20	1.49	1.61
2	E	1	G	OP3-P	-10.20	1.49	1.61

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	95	ARG	NE-CZ-NH2	-7.28	116.66	120.30
1	D	95	ARG	NE-CZ-NH2	-6.92	116.84	120.30
2	B	15	C	O5'-P-OP1	5.30	117.06	110.70
2	E	15	C	O5'-P-OP1	5.23	116.97	110.70

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	985	0	940	20	0
1	D	985	0	940	20	0
2	B	417	0	210	4	0
2	E	417	0	210	4	0
3	A	15	0	0	0	0
3	D	15	0	0	0	0
4	A	115	0	0	7	0
4	B	19	0	0	1	0
4	D	115	0	0	7	0
4	E	18	0	0	1	0
All	All	3101	0	2300	41	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 8.

All (41) 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:92:ILE:HG23	1:A:122:MET:HE1	1.21	1.16
1:D:92:ILE:HG23	1:D:122:MET:HE1	1.23	1.07
1:D:92:ILE:HG23	1:D:122:MET:CE	1.85	1.06
1:A:92:ILE:HG23	1:A:122:MET:CE	1.86	1.04
1:D:92:ILE:CG2	1:D:122:MET:CE	2.65	0.75
1:A:92:ILE:CG2	1:A:122:MET:CE	2.65	0.73
1:A:92:ILE:CG2	1:A:122:MET:HE1	2.11	0.71
1:D:107:THR:HG23	4:D:483:HOH:O	1.93	0.69
1:A:107:THR:HG23	4:A:477:HOH:O	1.93	0.68
2:B:4:G:N3	2:B:4:G:H5"	2.11	0.66
2:E:4:G:H5"	2:E:4:G:N3	2.12	0.65
2:B:19:C:O2	4:B:105:HOH:O	2.15	0.59
2:E:19:C:O2	4:E:107:HOH:O	2.13	0.59
1:A:127:PRO:HD3	4:A:457:HOH:O	2.03	0.58
1:A:27:THR:C	4:A:483:HOH:O	2.41	0.58
1:D:127:PRO:HD3	4:D:463:HOH:O	2.02	0.57
1:D:27:THR:C	4:D:489:HOH:O	2.42	0.57
1:A:116:ILE:O	1:A:120:GLN:HG3	2.07	0.55
1:D:23:HIS:O	1:D:27:THR:HG23	2.08	0.53
1:A:23:HIS:O	1:A:27:THR:HG23	2.08	0.53
1:D:116:ILE:O	1:D:120:GLN:HG3	2.08	0.53
1:A:56:ARG:HD2	4:A:451:HOH:O	2.09	0.52
1:A:114:MET:HE3	1:D:125:MET:HB2	1.91	0.52
1:D:56:ARG:HD2	4:D:457:HOH:O	2.10	0.50

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:92:ILE:CG2	1:D:122:MET:HE3	2.41	0.48
1:A:125:MET:HB2	1:D:114:MET:HE3	1.95	0.46
1:D:23:HIS:NE2	1:D:27:THR:HG21	2.31	0.45
1:A:117:ARG:HB3	4:A:513:HOH:O	2.16	0.45
1:D:82:GLY:N	4:D:479:HOH:O	2.49	0.45
1:A:23:HIS:NE2	1:A:27:THR:HG21	2.31	0.45
2:B:10:G:C8	2:E:11:G:C2	3.04	0.45
1:D:85:GLN:OE1	1:D:106:ARG:HB2	2.16	0.45
1:A:85:GLN:OE1	1:A:106:ARG:HB2	2.16	0.45
1:D:117:ARG:HB3	4:D:404:HOH:O	2.16	0.45
1:A:82:GLY:N	4:A:473:HOH:O	2.49	0.45
1:A:97:VAL:HG13	1:D:111:LEU:HD23	2.00	0.43
1:A:111:LEU:HD23	1:D:97:VAL:HG13	2.00	0.43
1:D:126:ALA:HA	4:D:463:HOH:O	2.19	0.42
1:A:45:LYS:HE2	2:B:8:G:O2'	2.20	0.41
1:A:126:ALA:HA	4:A:457:HOH:O	2.20	0.41
1:D:45:LYS:HE2	2:E:8:G:O2'	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	118/135 (87%)	118 (100%)	0	0	100	100
1	D	118/135 (87%)	118 (100%)	0	0	100	100
All	All	236/270 (87%)	236 (100%)	0	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	A	106/119 (89%)	106 (100%)	0	100	100
1	D	106/119 (89%)	105 (99%)	1 (1%)	78	84
All	All	212/238 (89%)	211 (100%)	1 (0%)	88	92

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

Mol	Chain	Res	Type
1	D	98	SER

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	23	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	18/19 (94%)	1 (5%)	0
2	E	18/19 (94%)	1 (5%)	0
All	All	36/38 (94%)	2 (5%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	4	G
2	E	4	G

There are no RNA pucker outliers to report.

## 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

6 ligands are 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	$\# Z  > 2$
3	SO4	D	303	-	4,4,4	0.59	0	6,6,6	1.03	1 (16%)
3	SO4	D	301	-	4,4,4	0.39	0	6,6,6	1.15	0
3	SO4	A	301	-	4,4,4	0.69	0	6,6,6	1.21	1 (16%)
3	SO4	D	302	-	4,4,4	0.30	0	6,6,6	0.63	0
3	SO4	A	303	-	4,4,4	0.90	0	6,6,6	0.67	0
3	SO4	A	302	-	4,4,4	0.58	0	6,6,6	0.45	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	303	SO4	O4-S-O3	2.25	118.66	109.06
3	A	301	SO4	O4-S-O3	2.01	117.66	109.06

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

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	A	122/135 (90%)	0.56	7 (5%) 23 29	28, 38, 70, 103	0
1	D	122/135 (90%)	0.48	6 (4%) 29 35	28, 38, 70, 102	0
2	B	19/19 (100%)	-0.66	0 100 100	44, 62, 73, 77	0
2	E	19/19 (100%)	-0.75	0 100 100	44, 62, 73, 77	0
All	All	282/308 (91%)	0.35	13 (4%) 32 38	28, 40, 73, 103	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	126	ALA	6.1
1	A	82	GLY	4.6
1	D	125	MET	4.5
1	A	127	PRO	4.1
1	A	81	PHE	4.0
1	D	83	PHE	3.8
1	D	126	ALA	3.6
1	D	127	PRO	3.6
1	A	83	PHE	3.3
1	A	125	MET	2.4
1	A	80	PHE	2.2
1	D	32	ASP	2.2
1	D	110	HIS	2.1

### 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
3	SO4	D	303	5/5	0.72	0.41	69,81,90,107	0
3	SO4	A	303	5/5	0.72	0.26	70,83,90,108	0
3	SO4	D	302	5/5	0.95	0.13	53,54,60,68	0
3	SO4	D	301	5/5	0.98	0.12	38,45,50,50	0
3	SO4	A	302	5/5	0.98	0.07	52,54,60,68	0
3	SO4	A	301	5/5	0.99	0.07	38,45,50,50	0

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

There are no such residues in this entry.