



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 6, 2021 – 02:23 PM EST

PDB ID : 7M50  
Title : Crystallographic structure of a cubic crystal form of STMV grown from ammonium sulfate  
Authors : McPherson, A.  
Deposited on : 2021-03-22  
Resolution : 2.31 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	2.24
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)	:	2.24

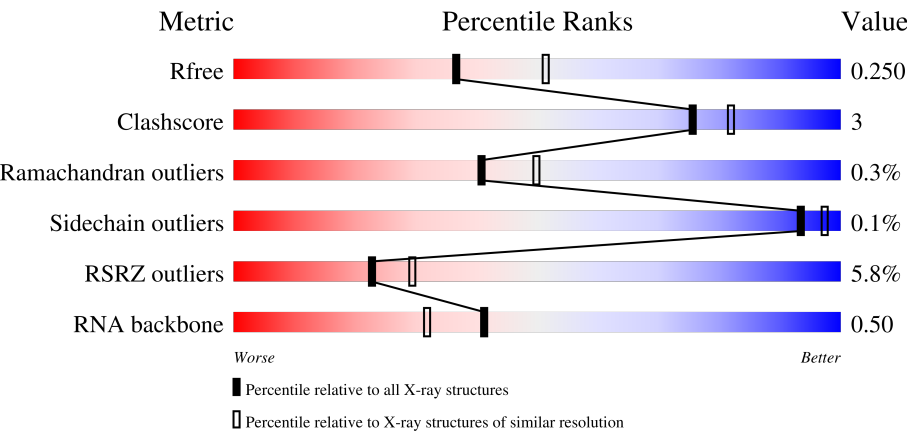
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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 <sub>free</sub>	130704	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)
RNA backbone	3102	1031 (2.70-1.94)

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	159	<div><div>3%</div><div><div></div><div>86%</div><div>6%</div><div>8%</div></div></div>
1	B	159	<div><div>%</div><div><div></div><div>84%</div><div>6%</div><div>10%</div></div></div>
1	C	159	<div><div>%</div><div><div></div><div>85%</div><div>5%</div><div>10%</div></div></div>
1	D	159	<div><div>%</div><div><div></div><div>84%</div><div>6%</div><div>9%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	159	% 82% 9% 9%
1	F	159	% 87% 5% 8%
1	G	159	% 82% 9% 8%
1	GG	159	2% 86% 5% 9%
1	H	159	3% 84% 8% 8%
1	HH	159	87% 9%
1	I	159	% 87% 9%
1	II	159	% 81% 8% 10%
1	J	159	3% 85% 8% 8%
1	JJ	159	3% 87% 6% 8%
1	K	159	4% 84% 8% 8%
1	KK	159	% 88% 9%
1	L	159	10% 87% 10% ..
1	M	159	3% 89% 9%
1	N	159	4% 85% 5% 9%
1	O	159	4% 89% 5% 6%
2	P	8	88% 62% 25% 12%
2	TT	8	88% 25% 12% 50% 12%
2	V	8	75% 62% 25% 12%
3	S	6	67% 67% 17% 17%
3	T	6	33% 17% 67% 17%
4	UU	10	80% 60% 10% 10% 20%
4	X	10	60% 60% 20% 20%
4	bb	10	80% 60% 30% 10%
5	Y	7	86% 86% 14%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	a	7	<div>86%</div> <div>57% 29% 14%</div>
6	WW	12	<div>75%</div> <div>17% 42% 25% 17%</div>
6	e	12	<div>83%</div> <div>42% 42% 17%</div>
7	h	10	<div>60%</div> <div>50% 30% 20%</div>
8	i	9	<div>22%</div> <div>33% 22% 44%</div>
8	ll	9	<div>22%</div> <div>44% 11% 44%</div>
8	qq	9	<div>33%</div> <div>56% 44%</div>
9	kk	8	<div>50%</div> <div>62% 12% 25%</div>
9	m	8	<div>38%</div> <div>75% 12% 12%</div>
9	n	8	<div>50%</div> <div>88% 12%</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
11	MG	B	702	-	-	-	X

## 2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 29786 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 Coat protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	146	Total	C	N	O	S	0	8	0
			1130	709	201	214	6			
1	B	143	Total	C	N	O	S	0	10	0
			1110	699	197	208	6			
1	C	143	Total	C	N	O	S	0	14	0
			1110	699	197	208	6			
1	D	144	Total	C	N	O	S	0	12	0
			1118	703	199	210	6			
1	E	145	Total	C	N	O	S	0	12	0
			1126	707	200	213	6			
1	F	147	Total	C	N	O	S	0	10	0
			1137	713	202	216	6			
1	G	146	Total	C	N	O	S	0	8	0
			1130	709	201	214	6			
1	H	147	Total	C	N	O	S	0	10	0
			1137	713	202	216	6			
1	I	144	Total	C	N	O	S	0	8	0
			1118	703	199	210	6			
1	J	147	Total	C	N	O	S	0	9	0
			1137	713	202	216	6			
1	K	146	Total	C	N	O	S	0	8	0
			1130	709	201	214	6			
1	L	157	Total	C	N	O	S	0	6	0
			1215	761	221	227	6			
1	M	144	Total	C	N	O	S	0	7	0
			1118	703	199	210	6			
1	N	144	Total	C	N	O	S	0	3	0
			1118	703	199	210	6			
1	O	150	Total	C	N	O	S	0	17	0
			1163	728	209	220	6			
1	GG	144	Total	C	N	O	S	0	12	0
			1118	703	199	210	6			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	HH	144	Total	C	N	O	S	0	9	0
			1118	703	199	210	6			
1	II	143	Total	C	N	O	S	0	9	0
			1110	699	197	208	6			
1	JJ	147	Total	C	N	O	S	0	13	0
			1137	713	202	216	6			
1	KK	145	Total	C	N	O	S	0	9	0
			1126	707	200	213	6			

- Molecule 2 is a RNA chain called RNA (5'-R(P\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	P	0	7	0
			154	70	35	42	7			
2	V	7	Total	C	N	O	P	0	6	0
			154	70	35	42	7			
2	TT	7	Total	C	N	O	P	0	6	0
			154	70	35	42	7			

- Molecule 3 is a RNA chain called RNA (5'-R(P\*AP\*AP\*AP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	S	5	Total	C	N	O	P	0	5	0
			110	50	25	30	5			
3	T	5	Total	C	N	O	P	0	5	0
			110	50	25	30	5			

- Molecule 4 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	X	8	Total	C	N	O	P	0	8	0
			176	80	40	48	8			
4	UU	8	Total	C	N	O	P	0	6	0
			176	80	40	48	8			
4	bb	9	Total	C	N	O	P	0	8	0
			198	90	45	54	9			

- Molecule 5 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	Y	6	Total	C	N	O	P	0	6	0
			132	60	30	36	6			
5	a	6	Total	C	N	O	P	0	6	0
			132	60	30	36	6			

- Molecule 6 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	e	12	Total	C	N	O	P	0	10	0
			240	108	24	96	12			
6	WW	10	Total	C	N	O	P	0	2	0
			200	90	20	80	10			

- Molecule 7 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	h	8	Total	C	N	O	P	0	7	0
			160	72	16	64	8			

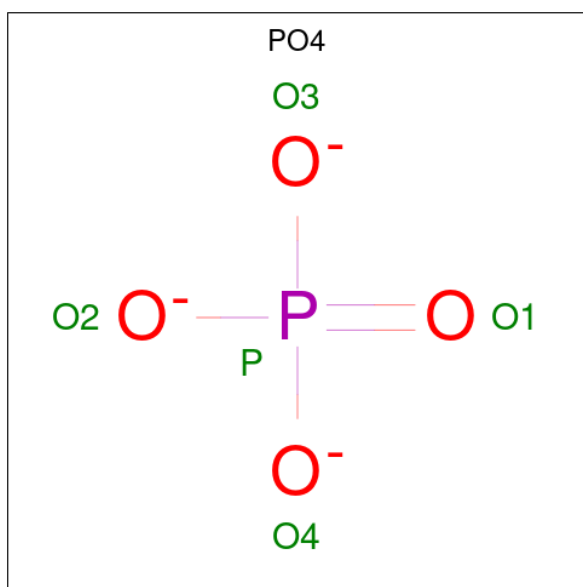
- Molecule 8 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	i	5	Total	C	N	O	P	0	5	0
			100	45	10	40	5			
8	ll	5	Total	C	N	O	P	0	5	0
			100	45	10	40	5			
8	qq	5	Total	C	N	O	P	0	4	0
			100	45	10	40	5			

- Molecule 9 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	m	7	Total	C	N	O	P	0	7	0
			140	63	14	56	7			
9	n	7	Total	C	N	O	P	0	7	0
			140	63	14	56	7			
9	kk	6	Total	C	N	O	P	0	6	0
			120	54	12	48	6			

- Molecule 10 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	B	1	Total	O	P	0	0
			5	4	1		
10	H	1	Total	O	P	0	0
			5	4	1		
10	L	1	Total	O	P	0	0
			5	4	1		
10	HH	1	Total	O	P	0	0
			5	4	1		
10	II	1	Total	O	P	0	0
			5	4	1		
10	R	1	Total	O	P	0	0
			5	4	1		

- Molecule 11 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

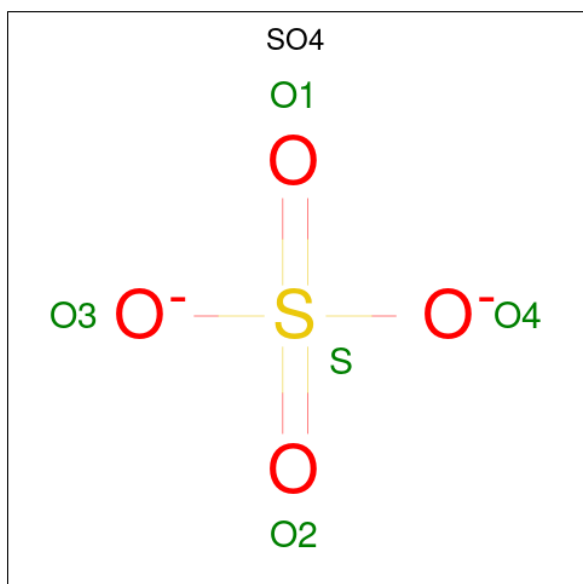
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	B	1	Total	Mg	0	0
			1	1		
11	G	1	Total	Mg	0	0
			1	1		
11	H	1	Total	Mg	0	0
			1	1		
11	JJ	2	Total	Mg	0	0
			2	2		

- Molecule 12 is CHLORIDE ION (three-letter code: CL) (formula: Cl).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	C	1	Total Cl 1 1	0	0

- Molecule 13 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	F	1	Total O S 5 4 1	0	0
13	M	1	Total O S 5 4 1	0	0
13	M	1	Total O S 5 4 1	0	0
13	O	1	Total O S 5 4 1	0	0
13	JJ	1	Total O S 5 4 1	0	0
13	KK	1	Total O S 5 4 1	0	0

- Molecule 14 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	A	128	Total O 128 128	0	0
14	B	149	Total O 149 149	0	1
14	C	118	Total O 118 118	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	D	177	Total 177	O 177	0	1
14	E	110	Total 110	O 110	0	0
14	F	127	Total 127	O 127	0	0
14	G	153	Total 153	O 153	0	0
14	H	149	Total 149	O 149	0	0
14	I	105	Total 105	O 105	0	0
14	J	155	Total 155	O 155	0	0
14	K	198	Total 198	O 198	0	0
14	L	138	Total 138	O 138	0	0
14	M	150	Total 150	O 150	0	1
14	N	96	Total 96	O 96	0	0
14	O	252	Total 252	O 252	0	0
14	P	39	Total 39	O 39	0	0
14	S	29	Total 29	O 29	0	0
14	T	22	Total 22	O 22	0	0
14	V	61	Total 61	O 61	0	0
14	X	36	Total 36	O 36	0	0
14	Y	55	Total 55	O 55	0	0
14	a	67	Total 67	O 67	0	0
14	e	193	Total 193	O 193	0	0
14	h	64	Total 64	O 64	0	0

*Continued on next page...*

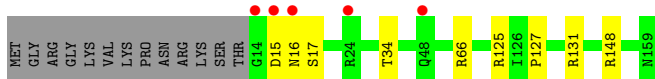
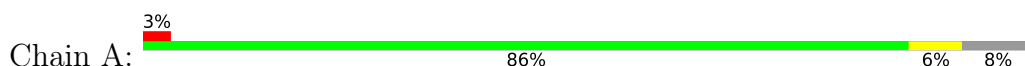
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	i	17	Total 17	O 17	0	0
14	m	39	Total 39	O 39	0	0
14	n	48	Total 48	O 48	0	0
14	GG	127	Total 127	O 127	0	1
14	HH	188	Total 188	O 188	0	0
14	II	271	Total 271	O 271	0	0
14	JJ	125	Total 125	O 125	0	0
14	KK	200	Total 200	O 200	0	0
14	TT	46	Total 46	O 46	0	0
14	UU	96	Total 96	O 96	0	0
14	WW	76	Total 76	O 76	0	0
14	bb	212	Total 212	O 212	0	0
14	kk	65	Total 65	O 65	0	0
14	ll	24	Total 24	O 24	0	0
14	qq	13	Total 13	O 13	0	0

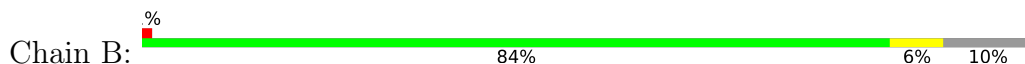
### 3 Residue-property plots [i](#)

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.

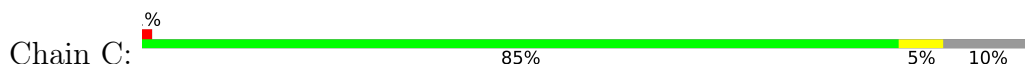
- Molecule 1: Coat protein



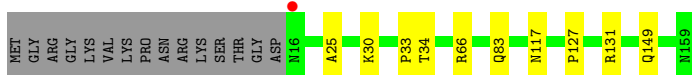
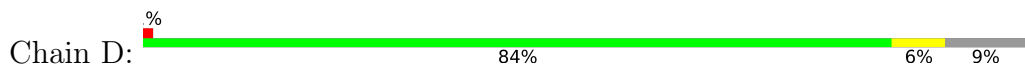
- Molecule 1: Coat protein



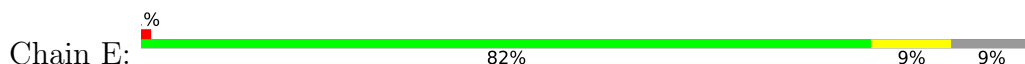
- Molecule 1: Coat protein



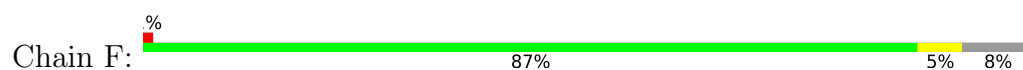
- Molecule 1: Coat protein



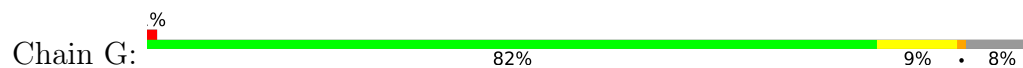
- Molecule 1: Coat protein



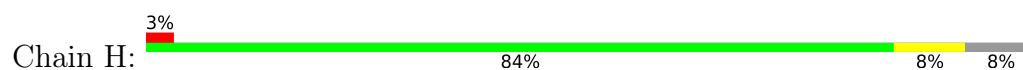
- Molecule 1: Coat protein



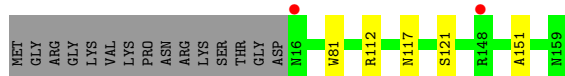
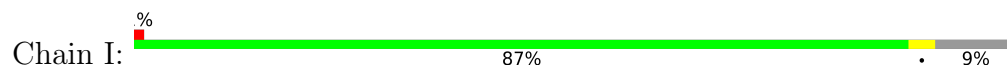
- Molecule 1: Coat protein



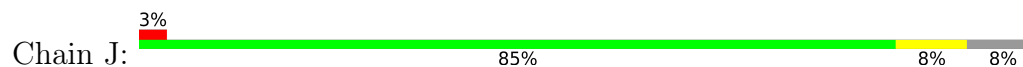
- Molecule 1: Coat protein



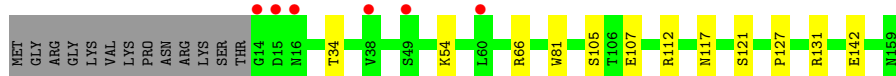
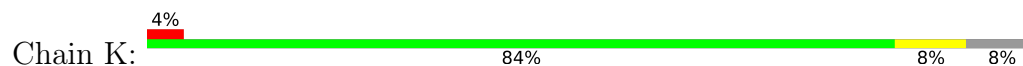
- Molecule 1: Coat protein



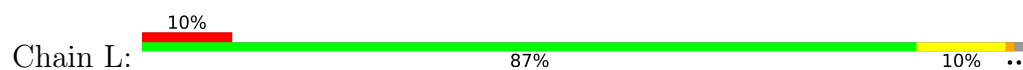
- Molecule 1: Coat protein



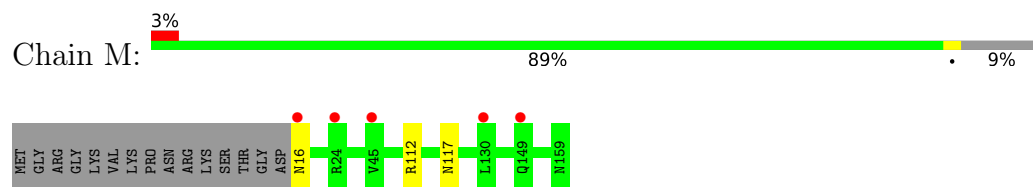
- Molecule 1: Coat protein



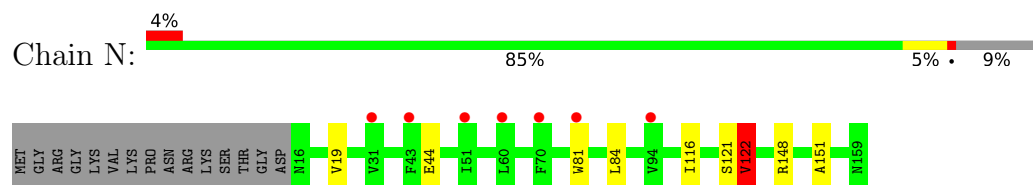
- Molecule 1: Coat protein



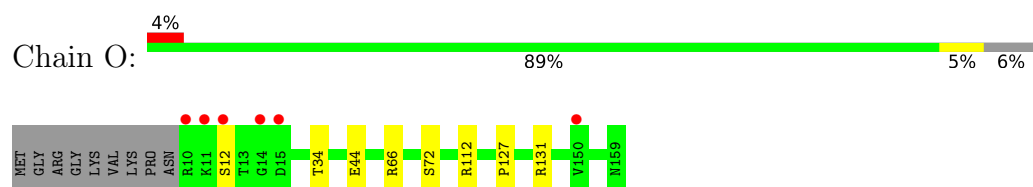
- Molecule 1: Coat protein



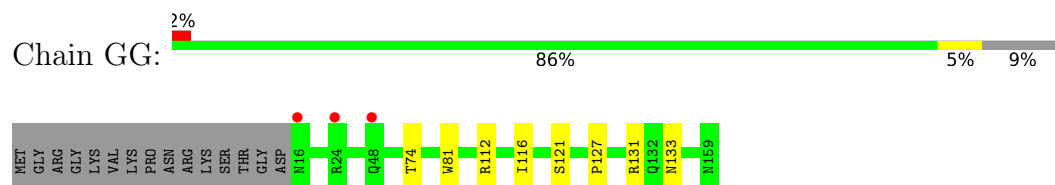
- Molecule 1: Coat protein



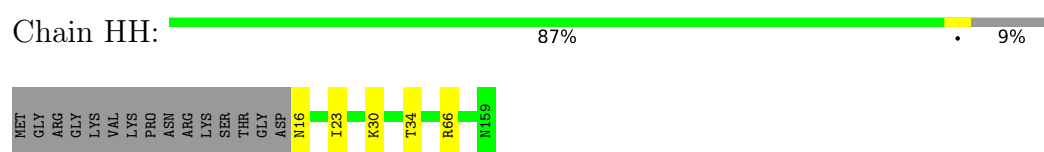
- Molecule 1: Coat protein



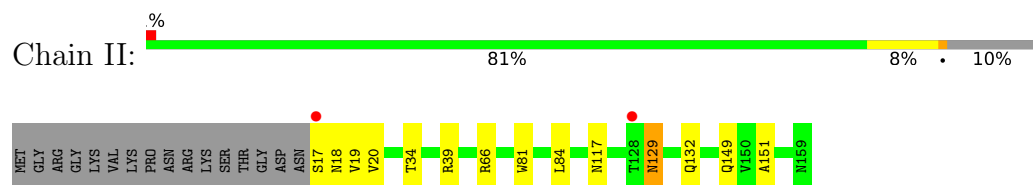
- Molecule 1: Coat protein



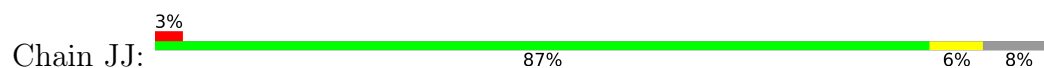
- Molecule 1: Coat protein

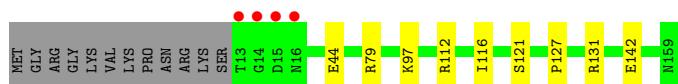


- Molecule 1: Coat protein

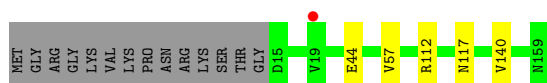
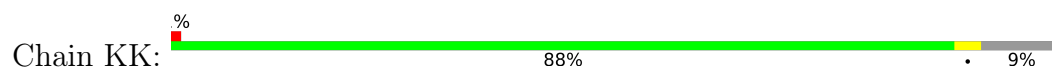


- Molecule 1: Coat protein

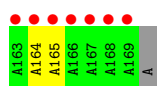




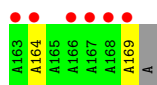
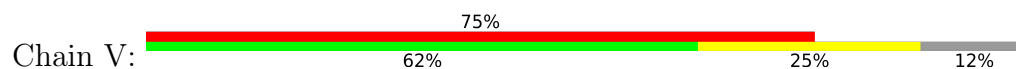
- Molecule 1: Coat protein



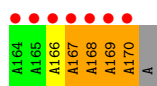
- Molecule 2: RNA (5'-R(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')



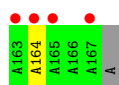
- Molecule 2: RNA (5'-R(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')



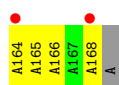
- Molecule 2: RNA (5'-R(P\*AP\*AP\*AP\*AP\*AP\*AP\*AP\*A)-3')



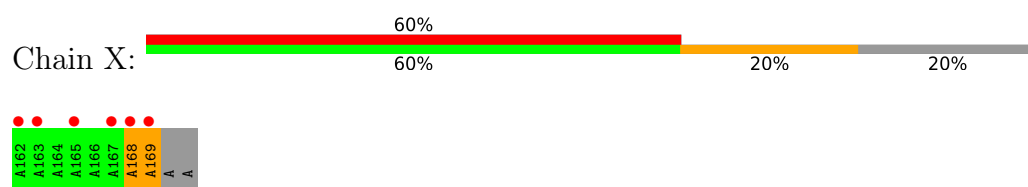
- Molecule 3: RNA (5'-R(P\*AP\*AP\*AP\*AP\*AP\*A)-3')



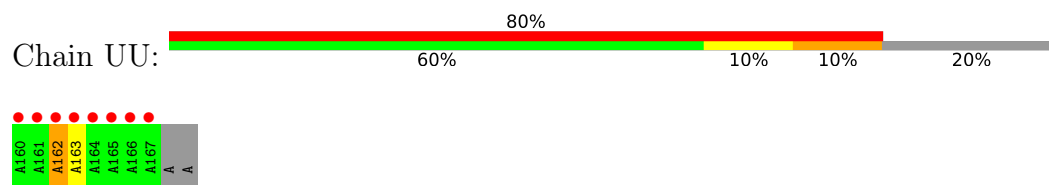
- Molecule 3: RNA (5'-R(P\*AP\*AP\*AP\*AP\*AP\*A)-3')



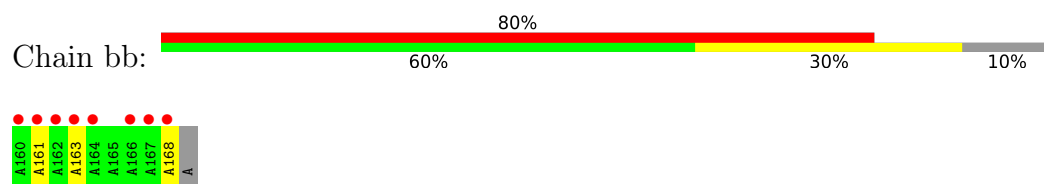
- Molecule 4: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



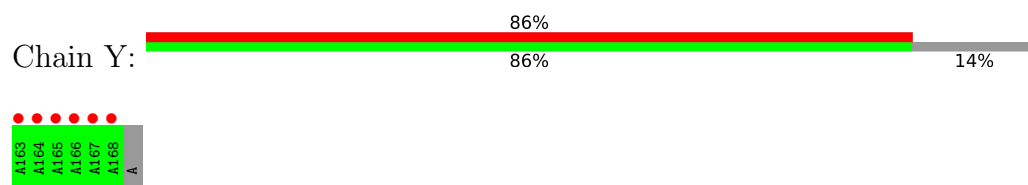
- Molecule 4: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



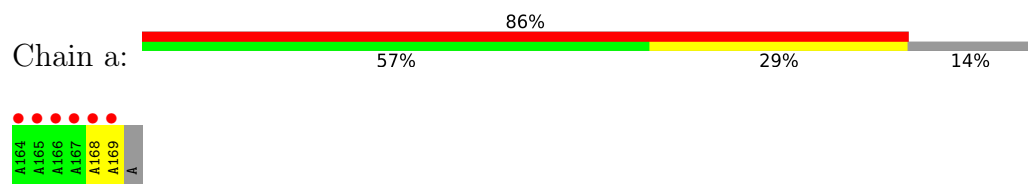
- Molecule 4: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



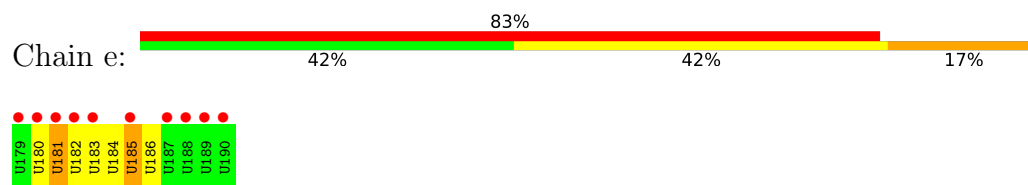
- Molecule 5: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



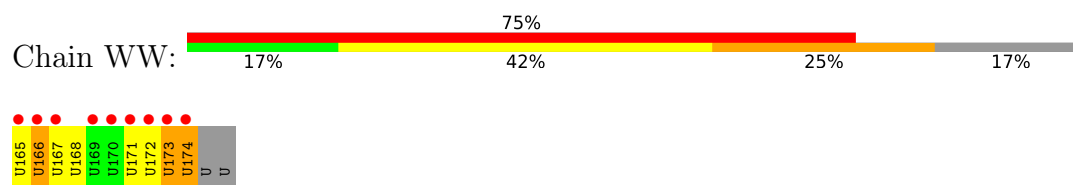
- Molecule 5: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 6: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U))-3')

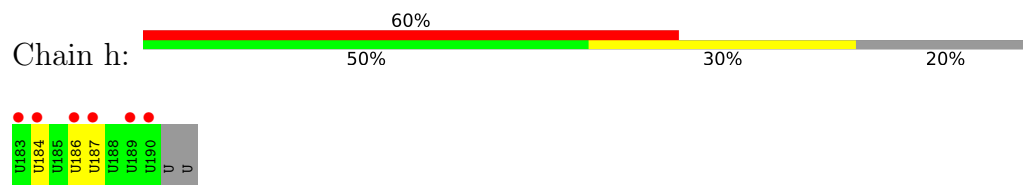


- Molecule 6: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U))-3')

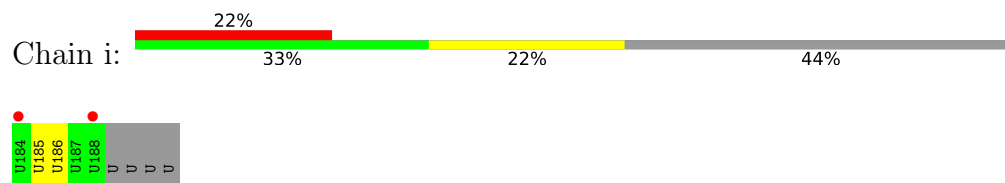




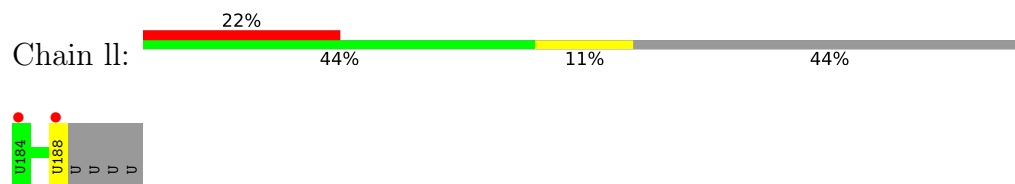
- Molecule 7: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



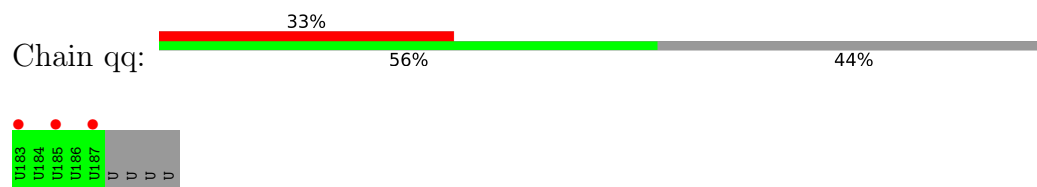
- Molecule 8: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



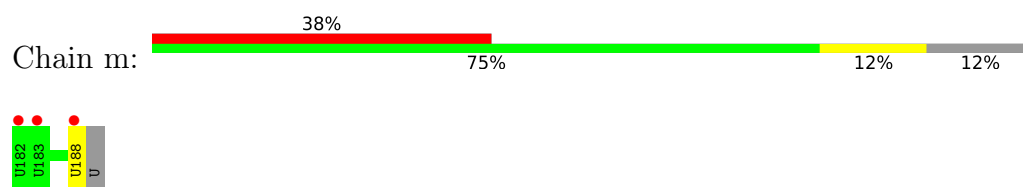
- Molecule 8: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



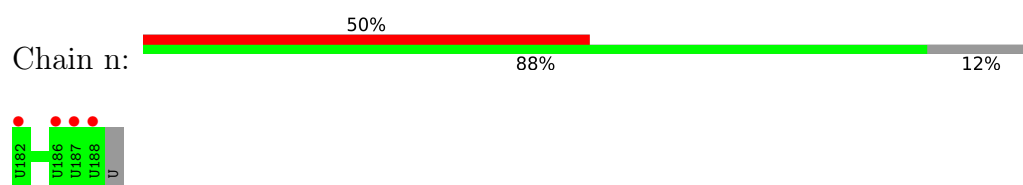
- Molecule 8: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 9: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 9: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 9: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



U183	U184	U185	U186	U187	U188	U	U
------	------	------	------	------	------	---	---

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	234.05Å 234.05Å 234.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	74.01 – 2.31 74.01 – 2.31	Depositor EDS
% Data completeness (in resolution range)	86.0 (74.01-2.31) 82.7 (74.01-2.31)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	0.18	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.71 (at 2.32Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, $R_{free}$	0.246 , 0.271 0.232 , 0.250	Depositor DCC
$R_{free}$ test set	1918 reflections (1.25%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	3.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , -5.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.28$ , $\langle L^2 \rangle = 0.13$	Xtriage
Estimated twinning fraction	0.147 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.72	EDS
Total number of atoms	29786	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	9.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.96% 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: PO4, SO4, CL, MG

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	0.24	0/1154	0.51	0/1572
1	B	0.24	0/1134	0.51	0/1545
1	C	0.24	0/1134	0.52	0/1545
1	D	0.24	0/1142	0.51	0/1556
1	E	0.24	0/1150	0.51	0/1567
1	F	0.24	0/1161	0.52	0/1582
1	G	0.25	0/1154	0.54	0/1572
1	GG	0.24	0/1142	0.51	0/1556
1	H	0.24	0/1161	0.51	0/1582
1	HH	0.24	0/1142	0.51	0/1556
1	I	0.24	0/1142	0.51	0/1556
1	II	0.24	0/1134	0.52	0/1545
1	J	0.24	0/1161	0.51	0/1582
1	JJ	0.24	0/1161	0.52	0/1582
1	K	0.24	0/1154	0.51	0/1572
1	KK	0.24	0/1150	0.51	0/1567
1	L	0.24	0/1240	0.51	0/1685
1	M	0.24	0/1142	0.51	0/1556
1	N	0.25	0/1142	0.52	0/1556
1	O	0.24	0/1187	0.52	0/1615
2	P	0.19	0/174	0.72	0/269
2	TT	0.19	0/174	1.23	3/269 (1.1%)
2	V	0.19	0/174	0.70	0/269
3	S	0.19	0/124	0.73	0/191
3	T	0.17	0/124	0.67	0/191
4	UU	0.27	0/199	0.84	0/308
4	X	0.17	0/199	0.74	0/308
4	bb	0.14	0/224	0.64	0/347
5	Y	0.14	0/149	0.66	0/230
5	a	0.32	0/149	0.88	0/230
6	WW	0.25	0/219	1.17	3/336 (0.9%)
6	e	0.23	0/263	1.16	4/404 (1.0%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
7	h	0.12	0/175	0.74	0/268
8	i	0.20	0/109	0.76	0/166
8	ll	0.13	0/109	0.73	0/166
8	qq	0.18	0/109	1.00	0/166
9	kk	0.13	0/131	0.72	0/200
9	m	0.10	0/153	0.71	0/234
9	n	0.10	0/153	0.71	0/234
All	All	0.24	0/26198	0.57	10/36235 (0.0%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	TT	169[A]	A	OP1-P-O3'	-10.66	81.74	105.20
2	TT	169[A]	A	OP2-P-O3'	-9.45	84.41	105.20
2	TT	170[A]	A	OP1-P-OP2	8.29	132.04	119.60
6	WW	166[A]	U	C2-N1-C1'	7.63	126.86	117.70
6	WW	166[A]	U	N3-C2-O2	-7.08	117.24	122.20

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	1130	0	1117	5	0
1	B	1110	0	1103	6	0
1	C	1110	0	1099	4	1
1	D	1118	0	1103	8	0
1	E	1126	0	1108	11	0
1	F	1137	0	1120	6	0
1	G	1130	0	1115	12	0
1	GG	1118	0	1103	5	0
1	H	1137	0	1118	9	0
1	HH	1118	0	1109	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	1118	0	1106	5	0
1	II	1110	0	1105	11	0
1	J	1137	0	1122	9	0
1	JJ	1137	0	1118	7	0
1	K	1130	0	1115	8	0
1	KK	1126	0	1110	4	0
1	L	1215	0	1216	12	0
1	M	1118	0	1109	3	0
1	N	1118	0	1112	7	0
1	O	1163	0	1146	6	0
2	P	154	0	78	2	0
2	TT	154	0	78	2	0
2	V	154	0	78	1	0
3	S	110	0	56	0	0
3	T	110	0	56	2	0
4	UU	176	0	89	2	0
4	X	176	0	89	1	0
4	bb	198	0	100	0	0
5	Y	132	0	67	0	0
5	a	132	0	67	0	0
6	WW	200	0	95	1	0
6	e	240	0	121	0	0
7	h	160	0	81	0	0
8	i	100	0	51	0	0
8	ll	100	0	51	0	0
8	qq	100	0	51	0	0
9	kk	120	0	61	0	0
9	m	140	0	71	0	0
9	n	140	0	71	0	0
10	B	5	0	0	0	0
10	H	5	0	0	0	0
10	HH	5	0	0	0	0
10	II	5	0	0	0	0
10	L	5	0	0	0	0
10	R	5	0	0	0	0
11	B	1	0	0	0	0
11	G	1	0	0	0	0
11	H	1	0	0	0	0
11	JJ	2	0	0	0	0
12	C	1	0	0	0	0
13	F	5	0	0	0	0
13	JJ	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	KK	5	0	0	0	0
13	M	10	0	0	0	0
13	O	5	0	0	0	0
14	A	128	0	0	2	0
14	B	149	0	0	0	0
14	C	118	0	0	0	0
14	D	177	0	0	1	0
14	E	110	0	0	1	0
14	F	127	0	0	0	0
14	G	153	0	0	2	0
14	GG	127	0	0	0	0
14	H	149	0	0	2	0
14	HH	188	0	0	0	0
14	I	105	0	0	1	0
14	II	271	0	0	1	0
14	J	155	0	0	1	0
14	JJ	125	0	0	2	0
14	K	198	0	0	0	0
14	KK	200	0	0	0	0
14	L	138	0	0	0	0
14	M	150	0	0	1	0
14	N	96	0	0	0	0
14	O	252	0	0	1	0
14	P	39	0	0	0	0
14	S	29	0	0	0	0
14	T	22	0	0	0	0
14	TT	46	0	0	1	0
14	UU	96	0	0	1	0
14	V	61	0	0	1	0
14	WW	76	0	0	1	0
14	X	36	0	0	0	0
14	Y	55	0	0	0	0
14	a	67	0	0	0	0
14	bb	212	0	0	0	0
14	e	193	0	0	0	0
14	h	64	0	0	0	0
14	i	17	0	0	0	0
14	kk	65	0	0	0	0
14	ll	24	0	0	0	0
14	m	39	0	0	0	0
14	n	48	0	0	0	0
14	qq	13	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	29786	0	23765	118	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 3.

The worst 5 of 118 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:L:12:SER:OG	1:L:16:ASN:O	2.15	0.65
1:JJ:79[A]:ARG:NH2	14:JJ:304:HOH:O	2.33	0.61
4:UU:162[A]:A:N6	14:UU:202:HOH:O	2.33	0.60
1:N:81:TRP:O	1:N:151:ALA:N	2.35	0.59
1:G:38:VAL:O	14:WW:201:HOH:O	2.17	0.58

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)
1:C:49[A]:SER:OG	1:C:49[A]:SER:OG[2_675]	2.14	0.06

## 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	144/159 (91%)	135 (94%)	6 (4%)	3 (2%)	7	5
1	B	141/159 (89%)	134 (95%)	6 (4%)	1 (1%)	22	26
1	C	141/159 (89%)	134 (95%)	7 (5%)	0	100	100
1	D	142/159 (89%)	132 (93%)	10 (7%)	0	100	100
1	E	143/159 (90%)	136 (95%)	7 (5%)	0	100	100

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	145/159 (91%)	133 (92%)	12 (8%)	0	100	100
1	G	144/159 (91%)	135 (94%)	8 (6%)	1 (1%)	22	26
1	GG	142/159 (89%)	133 (94%)	9 (6%)	0	100	100
1	H	145/159 (91%)	137 (94%)	8 (6%)	0	100	100
1	HH	142/159 (89%)	135 (95%)	7 (5%)	0	100	100
1	I	142/159 (89%)	134 (94%)	8 (6%)	0	100	100
1	II	141/159 (89%)	129 (92%)	11 (8%)	1 (1%)	22	26
1	J	145/159 (91%)	135 (93%)	10 (7%)	0	100	100
1	JJ	145/159 (91%)	137 (94%)	8 (6%)	0	100	100
1	K	144/159 (91%)	133 (92%)	11 (8%)	0	100	100
1	KK	143/159 (90%)	137 (96%)	6 (4%)	0	100	100
1	L	155/159 (98%)	137 (88%)	16 (10%)	2 (1%)	12	12
1	M	142/159 (89%)	132 (93%)	10 (7%)	0	100	100
1	N	142/159 (89%)	131 (92%)	10 (7%)	1 (1%)	22	26
1	O	148/159 (93%)	138 (93%)	9 (6%)	1 (1%)	22	26
All	All	2876/3180 (90%)	2687 (93%)	179 (6%)	10 (0%)	41	50

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	N	122	VAL
1	II	129	ASN
1	L	12	SER
1	L	16	ASN
1	A	17	SER

### 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	129/140 (92%)	129 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	127/140 (91%)	127 (100%)	0	100	100
1	C	127/140 (91%)	127 (100%)	0	100	100
1	D	128/140 (91%)	128 (100%)	0	100	100
1	E	129/140 (92%)	128 (99%)	1 (1%)	81	90
1	F	130/140 (93%)	130 (100%)	0	100	100
1	G	129/140 (92%)	129 (100%)	0	100	100
1	GG	128/140 (91%)	128 (100%)	0	100	100
1	H	130/140 (93%)	130 (100%)	0	100	100
1	HH	128/140 (91%)	128 (100%)	0	100	100
1	I	128/140 (91%)	128 (100%)	0	100	100
1	II	127/140 (91%)	127 (100%)	0	100	100
1	J	130/140 (93%)	130 (100%)	0	100	100
1	JJ	130/140 (93%)	130 (100%)	0	100	100
1	K	129/140 (92%)	129 (100%)	0	100	100
1	KK	129/140 (92%)	129 (100%)	0	100	100
1	L	138/140 (99%)	138 (100%)	0	100	100
1	M	128/140 (91%)	128 (100%)	0	100	100
1	N	128/140 (91%)	127 (99%)	1 (1%)	81	90
1	O	133/140 (95%)	133 (100%)	0	100	100
All	All	2585/2800 (92%)	2583 (100%)	2 (0%)	93	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	159	ASN
1	N	122	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	48	GLN
1	G	48	GLN
1	J	117	ASN
1	L	9	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	P	6/8 (75%)	0	0
2	TT	6/8 (75%)	5 (83%)	0
2	V	6/8 (75%)	1 (16%)	0
3	S	4/6 (66%)	1 (25%)	0
3	T	4/6 (66%)	1 (25%)	0
4	UU	7/10 (70%)	1 (14%)	0
4	X	7/10 (70%)	1 (14%)	1 (14%)
4	bb	8/10 (80%)	3 (37%)	0
5	Y	5/7 (71%)	0	0
5	a	5/7 (71%)	2 (40%)	0
6	WW	10/12 (83%)	7 (70%)	2 (20%)
6	e	11/12 (91%)	7 (63%)	0
7	h	7/10 (70%)	3 (42%)	0
8	i	4/9 (44%)	2 (50%)	0
8	ll	4/9 (44%)	1 (25%)	0
8	qq	4/9 (44%)	0	0
9	kk	5/8 (62%)	1 (20%)	0
9	m	6/8 (75%)	1 (16%)	0
9	n	6/8 (75%)	0	0
All	All	115/165 (69%)	37 (32%)	3 (2%)

5 of 37 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	S	164[A]	A
3	T	168[A]	A
2	V	169[A]	A
4	X	169[A]	A
5	a	168[A]	A

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	X	168[A]	A
6	WW	165	U
6	WW	173	U

## 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 [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 6 are monoatomic - leaving 12 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	$\# Z  > 2$
10	PO4	R	1	-	4,4,4	0.91	0	6,6,6	0.44	0
13	SO4	JJ	203	-	4,4,4	0.14	0	6,6,6	0.05	0
13	SO4	O	201	-	4,4,4	0.14	0	6,6,6	0.06	0
13	SO4	M	201	-	4,4,4	0.14	0	6,6,6	0.07	0
13	SO4	KK	201	-	4,4,4	0.14	0	6,6,6	1.59	1 (16%)
13	SO4	M	202	-	4,4,4	0.14	0	6,6,6	0.12	0
10	PO4	L	901	-	4,4,4	0.92	0	6,6,6	0.43	0
13	SO4	F	201	-	4,4,4	0.14	0	6,6,6	1.32	1 (16%)
10	PO4	II	201	-	4,4,4	0.90	0	6,6,6	1.09	0
10	PO4	HH	401	-	4,4,4	0.92	0	6,6,6	0.43	0
10	PO4	B	701	-	4,4,4	0.92	0	6,6,6	0.43	0
10	PO4	H	501	-	4,4,4	0.92	0	6,6,6	0.42	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	KK	201	SO4	O3-S-O1	3.27	126.37	109.31
13	F	201	SO4	O4-S-O3	2.14	118.21	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 [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	146/159 (91%)	0.33	5 (3%) 45 52	4, 7, 11, 19	0
1	B	143/159 (89%)	0.32	1 (0%) 87 91	4, 6, 9, 11	0
1	C	143/159 (89%)	0.27	1 (0%) 87 91	3, 4, 7, 19	0
1	D	144/159 (90%)	0.16	1 (0%) 87 91	2, 4, 7, 14	0
1	E	145/159 (91%)	0.26	1 (0%) 87 91	4, 5, 8, 15	0
1	F	147/159 (92%)	0.42	2 (1%) 75 80	4, 8, 10, 17	0
1	G	146/159 (91%)	0.32	1 (0%) 87 91	3, 6, 10, 12	0
1	GG	144/159 (90%)	0.28	3 (2%) 63 70	2, 3, 7, 14	0
1	H	147/159 (92%)	0.35	5 (3%) 45 52	5, 7, 10, 18	0
1	HH	144/159 (90%)	0.22	0 100 100	3, 4, 7, 15	0
1	I	144/159 (90%)	0.38	2 (1%) 75 80	7, 10, 12, 19	0
1	II	143/159 (89%)	0.22	2 (1%) 75 80	2, 4, 7, 12	0
1	J	147/159 (92%)	0.43	4 (2%) 54 62	8, 10, 13, 22	0
1	JJ	147/159 (92%)	0.36	4 (2%) 54 62	2, 4, 8, 18	0
1	K	146/159 (91%)	0.50	6 (4%) 37 44	7, 9, 13, 23	0
1	KK	145/159 (91%)	0.17	1 (0%) 87 91	2, 3, 7, 16	0
1	L	157/159 (98%)	0.77	16 (10%) 6 10	9, 11, 17, 22	0
1	M	144/159 (90%)	0.57	5 (3%) 44 51	10, 11, 15, 21	0
1	N	144/159 (90%)	0.53	7 (4%) 29 37	6, 10, 14, 16	0
1	O	150/159 (94%)	0.45	6 (4%) 38 45	5, 8, 13, 26	0
2	P	7/8 (87%)	3.46	7 (100%) 0 0	16, 18, 23, 28	0
2	TT	7/8 (87%)	5.29	7 (100%) 0 0	17, 20, 24, 29	0
2	V	7/8 (87%)	3.16	6 (85%) 0 0	14, 17, 22, 23	0
3	S	5/6 (83%)	3.06	4 (80%) 0 0	15, 16, 19, 21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
3	T	5/6 (83%)	1.78	2 (40%) 0 0	16, 17, 18, 19	0
4	UU	8/10 (80%)	4.72	8 (100%) 0 0	17, 18, 26, 26	0
4	X	8/10 (80%)	3.28	6 (75%) 0 0	21, 22, 26, 26	0
4	bb	9/10 (90%)	5.30	8 (88%) 0 0	14, 19, 25, 27	0
5	Y	6/7 (85%)	3.41	6 (100%) 0 0	20, 22, 23, 23	0
5	a	6/7 (85%)	5.67	6 (100%) 0 0	20, 23, 29, 31	0
6	WW	10/12 (83%)	3.90	9 (90%) 0 0	16, 24, 29, 30	0
6	e	12/12 (100%)	3.75	10 (83%) 0 0	19, 23, 28, 28	0
7	h	8/10 (80%)	2.95	6 (75%) 0 0	14, 17, 24, 26	0
8	i	5/9 (55%)	1.47	2 (40%) 0 0	15, 15, 19, 23	0
8	ll	5/9 (55%)	2.17	2 (40%) 0 0	16, 16, 22, 23	0
8	qq	5/9 (55%)	2.12	3 (60%) 0 0	14, 15, 22, 22	0
9	kk	6/8 (75%)	2.66	4 (66%) 0 0	17, 19, 23, 28	0
9	m	7/8 (87%)	2.80	3 (42%) 0 0	17, 20, 28, 28	0
9	n	7/8 (87%)	2.63	4 (57%) 0 0	19, 20, 24, 25	0
All	All	3049/3345 (91%)	0.50	176 (5%) 23 29	2, 7, 16, 31	0

The worst 5 of 176 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	6	VAL	10.0
1	J	13	THR	9.7
4	bb	161[A]	A	8.9
2	TT	168[A]	A	8.7
2	TT	169[A]	A	8.5

## 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 ⓘ

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
10	PO4	II	201	5/5	0.76	0.37	23,26,26,33	5
11	MG	B	702	1/1	0.80	0.57	4,4,4,4	0
13	SO4	JJ	203	5/5	0.83	0.48	9,9,12,18	5
13	SO4	M	202	5/5	0.89	0.18	31,31,34,39	5
13	SO4	O	201	5/5	0.90	0.49	30,30,31,34	5
13	SO4	M	201	5/5	0.90	0.51	35,40,44,45	5
13	SO4	F	201	5/5	0.93	0.21	11,12,13,14	5
12	CL	C	201	1/1	0.94	0.09	13,13,13,13	0
11	MG	G	201	1/1	0.94	0.66	14,14,14,14	0
11	MG	JJ	202	1/1	0.94	0.33	4,4,4,4	0
10	PO4	HH	401	5/5	0.95	0.17	2,2,2,2	0
13	SO4	KK	201	5/5	0.95	0.55	24,28,29,30	5
10	PO4	L	901	5/5	0.96	0.11	9,10,11,15	0
11	MG	JJ	201	1/1	0.96	0.24	6,6,6,6	0
10	PO4	H	501	5/5	0.97	0.11	7,7,7,7	0
10	PO4	B	701	5/5	0.97	0.15	3,4,4,4	0
11	MG	H	502	1/1	0.97	0.10	7,7,7,7	0
10	PO4	R	1	5/5	0.99	0.35	42,49,49,49	5

## 6.5 Other polymers [i](#)

There are no such residues in this entry.