



# wwPDB X-ray Structure Validation Summary Report

Apr 23, 2024 – 10:23 pm BST

PDB ID : 8RHK  
Title : Yeast 20S proteasome in complex with a linear oxindole epoxyketone (compound 6)  
Authors : Goetz, M.G.; Godwin, K.; Price, R.; Dorn, R.; Merrill-Steskal, G.; Hansen, H.; Klemmer, W.; Produturi, G.; Rocha, M.; Palmer, M.; Molacek, L.; Strater, Z.; Groll, M.  
Deposited on : 2023-12-15  
Resolution : 2.80 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 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)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36.2  
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) : 2.36.2

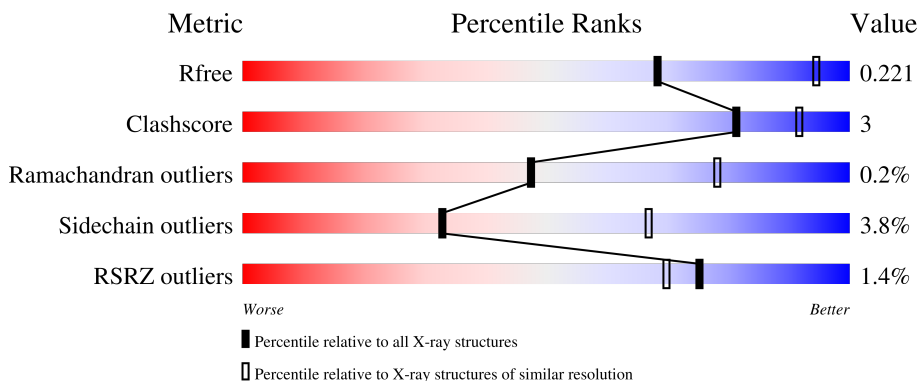
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

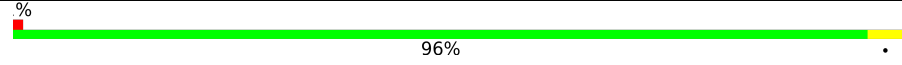
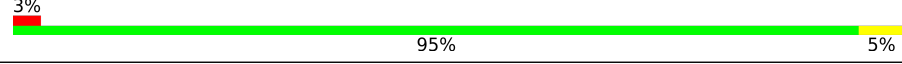
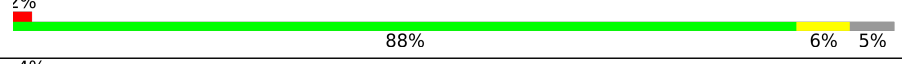
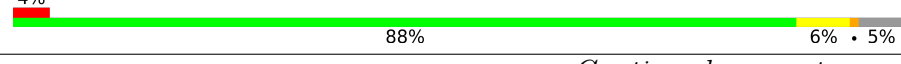
The reported resolution of this entry is 2.80 Å.

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	250	 96% .
1	O	250	 95% 5%
2	B	258	 88% 6% 5%
2	P	258	 88% 6% . 5%



*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	254	3% 87% 6% • 6%
3	Q	254	5% 87% 6% • 6%
4	D	260	% 84% 6% 10%
4	R	260	% 85% 5% 10%
5	E	234	% 92% 6% ..
5	S	234	2% 92% 6% •
6	F	288	% 80% • • 16%
6	T	288	2% 79% 5% • 16%
7	G	252	% 88% 7% •
7	U	252	% 87% 8% •
8	H	232	% 88% 7% • •
8	V	232	% 88% 8% •
9	I	205	91% 8%
9	W	205	91% 8%
10	J	198	% 89% 7% • •
10	X	198	2% 89% 7% • •
11	K	212	86% 11% •
11	Y	212	87% 11% •
12	L	222	91% 9%
12	Z	222	92% 8%
13	M	246	88% 7% 5%
13	a	246	92% • 5%
14	N	196	% 95% 5%
14	b	196	% 98% •
15	g	5	60% 40%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
15	h	5	 60% 40%
16	e	5	 60% 40%
16	f	5	 40% 60%
16	i	5	 60% 40%
16	j	5	 60% 40%

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
16	A1H2H	i	2	-	-	-	X
16	A1H2H	j	2	-	-	-	X

## 2 Entry composition

There are 20 unique types of molecules in this entry. The entry contains 49935 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 Proteasome subunit alpha type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	Total 1915	C 1219	N 315	O 377	S 4	0	0	0
1	O	250	Total 1915	C 1219	N 315	O 377	S 4	0	0	0

- Molecule 2 is a protein called Proteasome subunit alpha type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	244	Total 1904	C 1201	N 321	O 379	S 3	0	0	0
2	P	244	Total 1904	C 1201	N 321	O 379	S 3	0	0	0

- Molecule 3 is a protein called Proteasome subunit alpha type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	240	Total 1881	C 1176	N 329	O 372	S 4	0	0	0
3	Q	240	Total 1881	C 1176	N 329	O 372	S 4	0	0	0

- Molecule 4 is a protein called Proteasome subunit alpha type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	235	Total 1813	C 1136	N 304	O 366	S 7	0	0	0
4	R	235	Total 1813	C 1136	N 304	O 366	S 7	0	0	0

- Molecule 5 is a protein called Proteasome subunit alpha type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	231	Total 1773	C 1114	N 307	O 348	S 4	0	0	0
5	S	231	Total 1773	C 1114	N 307	O 348	S 4	0	0	0

- Molecule 6 is a protein called Probable proteasome subunit alpha type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	243	Total 1892	C 1203	N 329	O 356	S 4	0	0	0
6	T	243	Total 1892	C 1203	N 329	O 356	S 4	0	0	0

- Molecule 7 is a protein called Proteasome subunit alpha type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	241	Total 1907	C 1214	N 320	O 365	S 8	0	0	0
7	U	241	Total 1907	C 1214	N 320	O 365	S 8	0	0	0

- Molecule 8 is a protein called Proteasome subunit beta type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	222	Total 1684	C 1061	N 293	O 323	S 7	0	0	0
8	V	222	Total 1684	C 1061	N 293	O 323	S 7	0	0	0

- Molecule 9 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	204	Total 1581	C 1010	N 258	O 305	S 8	0	0	0
9	W	204	Total 1581	C 1010	N 258	O 305	S 8	0	0	0

- Molecule 10 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	195	Total 1561	C 992	N 264	O 299	S 6	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	195	Total	C	N	O	S	0	0	0
			1561	992	264	299	6			

- Molecule 11 is a protein called Proteasome subunit beta type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			
11	Y	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			

- Molecule 12 is a protein called Proteasome subunit beta type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			
12	Z	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			

- Molecule 13 is a protein called Proteasome subunit beta type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			
13	a	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			

- Molecule 14 is a protein called Proteasome subunit beta type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			
14	b	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			

- Molecule 15 is a protein called Linear oxindole epoxyketone.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	g	5	Total	C	F	N	O	0	0	0
			54	37	1	6	10			
15	h	5	Total	C	F	N	O	0	0	0
			54	37	1	6	10			

- Molecule 16 is a protein called Linear oxindole epoxyketone.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	e	5	Total	C	F	N	O	0	0	0
			54	37	1	6	10			
16	i	5	Total	C	F	N	O	0	0	0
			54	37	1	6	10			
16	f	5	Total	C	F	N	O	0	0	0
			54	37	1	6	10			
16	j	5	Total	C	F	N	O	0	0	0
			54	37	1	6	10			

- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	G	1	Total	Mg	0	0
			1	1		
17	I	1	Total	Mg	0	0
			1	1		
17	K	1	Total	Mg	0	0
			1	1		
17	N	1	Total	Mg	0	0
			1	1		
17	V	1	Total	Mg	0	0
			1	1		
17	X	1	Total	Mg	0	0
			1	1		
17	Y	1	Total	Mg	0	0
			1	1		
17	Z	1	Total	Mg	0	0
			1	1		

- Molecule 18 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	G	1	Total	Cl	0	0
			1	1		
18	U	1	Total	Cl	0	0
			1	1		

- Molecule 19 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
19	K	1	12	6	1	4	1	0	0
19	Y	1	12	6	1	4	1	0	0

- Molecule 20 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
20	A	9	Total	O	0	0
			9	9		
20	B	16	Total	O	0	0
			16	16		
20	C	5	Total	O	0	0
			5	5		
20	D	9	Total	O	0	0
			9	9		
20	E	5	Total	O	0	0
			5	5		
20	F	4	Total	O	0	0
			4	4		
20	G	8	Total	O	0	0
			8	8		
20	H	17	Total	O	0	0
			17	17		
20	I	7	Total	O	0	0
			7	7		
20	J	19	Total	O	0	0
			19	19		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
20	K	12	Total O 12 12	0	0
20	L	21	Total O 21 21	0	0
20	M	14	Total O 14 14	0	0
20	N	13	Total O 13 13	0	0
20	O	2	Total O 2 2	0	0
20	P	10	Total O 10 10	0	0
20	Q	8	Total O 8 8	0	0
20	R	2	Total O 2 2	0	0
20	S	4	Total O 4 4	0	0
20	T	3	Total O 3 3	0	0
20	U	11	Total O 11 11	0	0
20	V	12	Total O 12 12	0	0
20	W	7	Total O 7 7	0	0
20	X	14	Total O 14 14	0	0
20	Y	12	Total O 12 12	0	0
20	Z	11	Total O 11 11	0	0
20	a	12	Total O 12 12	0	0
20	b	7	Total O 7 7	0	0
20	e	2	Total O 2 2	0	0
20	i	2	Total O 2 2	0	0
20	f	1	Total O 1 1	0	0

*Continued on next page...*

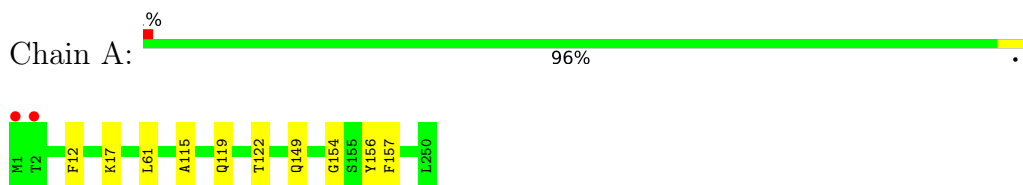
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
20	h	1	Total O 1 1	0	0
20	j	1	Total O 1 1	0	0

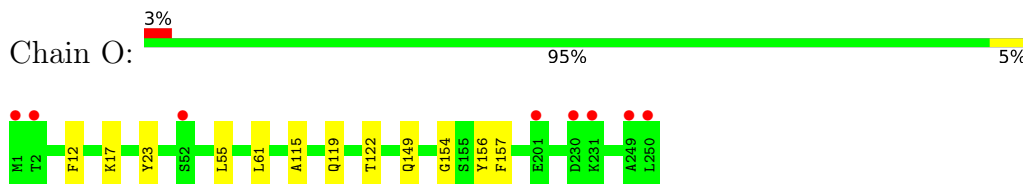
### 3 Residue-property plots [i](#)

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.

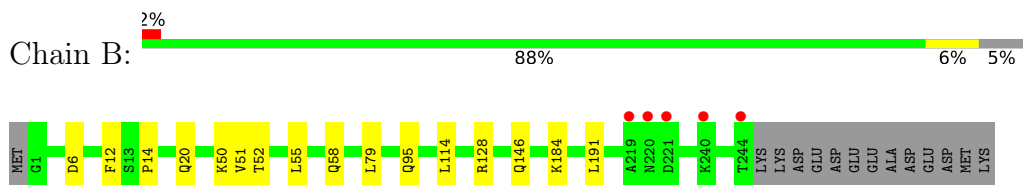
- Molecule 1: Proteasome subunit alpha type-2



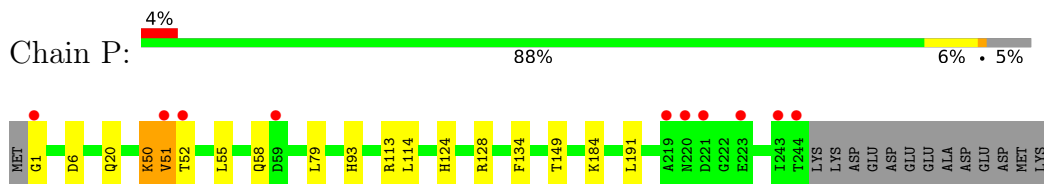
- Molecule 1: Proteasome subunit alpha type-2



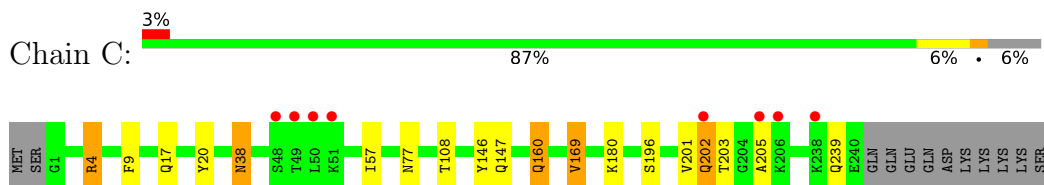
- Molecule 2: Proteasome subunit alpha type-3



- Molecule 2: Proteasome subunit alpha type-3



- Molecule 3: Proteasome subunit alpha type-4



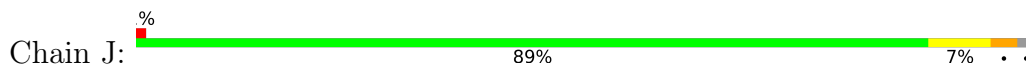
- Molecule 3: Proteasome subunit alpha type-4



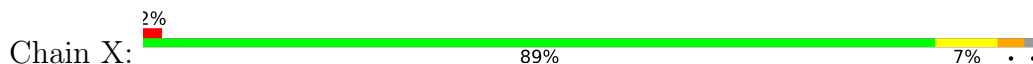




- Molecule 10: Proteasome subunit beta type-4



- Molecule 10: Proteasome subunit beta type-4



- Molecule 11: Proteasome subunit beta type-5



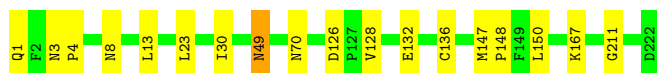
- Molecule 11: Proteasome subunit beta type-5




- Molecule 12: Proteasome subunit beta type-6



- Molecule 12: Proteasome subunit beta type-6



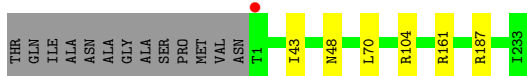
- Molecule 13: Proteasome subunit beta type-7

Chain M:  88% 7% 5%



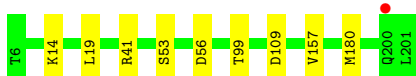
- Molecule 13: Proteasome subunit beta type-7

Chain a:  92% 5%



- Molecule 14: Proteasome subunit beta type-1

Chain N:  95% 5%



- Molecule 14: Proteasome subunit beta type-1

Chain b:  98%



- Molecule 15: Linear oxindole epoxyketone

Chain g:  60% 40%



- Molecule 15: Linear oxindole epoxyketone

Chain h:  60% 40%



- Molecule 16: Linear oxindole epoxyketone

Chain e:  60% 40%



- Molecule 16: Linear oxindole epoxyketone



Chain i:  60% 40%

ACE1
A1H2H2
L3
A1H2I4
6V05

- Molecule 16: Linear oxindole epoxyketone

Chain f:  40% 60%

ACE1
A1H2H2
L3
A1H2I4
6V05

- Molecule 16: Linear oxindole epoxyketone

Chain j:  60% 40%

ACE1
A1H2H2
L3
A1H2I4
6V05

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.10Å 300.68Å 144.70Å 90.00° 113.19° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 29.98 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.2 (30.00-2.80) 96.3 (29.98-2.80)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 2.80Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.182 , 0.218 0.190 , 0.221	Depositor DCC
$R_{free}$ test set	12577 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.1	Xtrriage
Anisotropy	0.541	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.3	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	0.95	EDS
Total number of atoms	49935	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	77.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.28% 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: 6VO, MES, A1H2H, MG, A1H2I, CL, A1H45, ACE

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	0.65	0/1952	0.70	0/2642
1	O	0.65	0/1952	0.70	0/2642
2	B	0.65	0/1934	0.71	0/2618
2	P	0.65	0/1934	0.71	0/2618
3	C	0.65	0/1910	0.72	0/2586
3	Q	0.66	0/1910	0.72	0/2586
4	D	0.66	0/1837	0.70	0/2475
4	R	0.66	0/1837	0.70	0/2475
5	E	0.66	0/1800	0.71	0/2433
5	S	0.66	0/1800	0.71	0/2433
6	F	0.65	0/1932	0.70	0/2609
6	T	0.65	0/1932	0.70	0/2609
7	G	0.64	0/1945	0.69	0/2634
7	U	0.64	0/1945	0.70	0/2634
8	H	0.65	0/1715	0.73	0/2326
8	V	0.65	0/1715	0.74	0/2326
9	I	0.65	0/1611	0.71	0/2174
9	W	0.65	0/1611	0.72	0/2174
10	J	0.63	0/1589	0.70	0/2142
10	X	0.63	0/1589	0.70	0/2142
11	K	0.64	0/1681	0.72	0/2274
11	Y	0.65	0/1681	0.72	0/2274
12	L	0.63	0/1795	0.70	0/2420
12	Z	0.63	0/1795	0.70	0/2420
13	M	0.64	0/1855	0.72	0/2514
13	a	0.64	0/1855	0.73	0/2514
14	N	0.64	0/1541	0.70	0/2087
14	b	0.64	0/1541	0.70	0/2087
15	g	1.17	0/7	1.84	0/8
15	h	1.20	0/7	1.72	0/8
16	e	1.69	0/7	2.52	0/8
16	f	1.11	0/7	1.93	0/8

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
16	i	1.14	0/7	1.92	0/8
16	j	0.93	0/7	1.91	0/8
All	All	0.65	0/50236	0.71	0/67916

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	1915	0	1929	5	0
1	O	1915	0	1929	7	0
2	B	1904	0	1904	8	0
2	P	1904	0	1904	8	0
3	C	1881	0	1895	12	0
3	Q	1881	0	1895	12	0
4	D	1813	0	1797	6	0
4	R	1813	0	1797	5	0
5	E	1773	0	1775	7	0
5	S	1773	0	1775	5	0
6	F	1892	0	1883	4	0
6	T	1892	0	1883	5	0
7	G	1907	0	1901	7	0
7	U	1907	0	1901	10	0
8	H	1684	0	1684	9	0
8	V	1684	0	1684	9	0
9	I	1581	0	1574	13	0
9	W	1581	0	1574	12	0
10	J	1561	0	1569	15	0
10	X	1561	0	1569	14	0
11	K	1644	0	1591	24	0
11	Y	1644	0	1591	21	0
12	L	1757	0	1711	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	Z	1757	0	1711	9	0
13	M	1824	0	1832	7	0
13	a	1824	0	1832	0	0
14	N	1512	0	1477	3	0
14	b	1512	0	1477	0	0
15	g	54	0	14	0	0
15	h	54	0	14	0	0
16	e	54	0	14	0	0
16	f	54	0	14	0	0
16	i	54	0	14	0	0
16	j	54	0	14	0	0
17	G	1	0	0	0	0
17	I	1	0	0	0	0
17	K	1	0	0	0	0
17	N	1	0	0	0	0
17	V	1	0	0	0	0
17	X	1	0	0	0	0
17	Y	1	0	0	0	0
17	Z	1	0	0	0	0
18	G	1	0	0	0	0
18	U	1	0	0	0	0
19	K	12	0	13	1	0
19	Y	12	0	13	1	0
20	A	9	0	0	0	0
20	B	16	0	0	0	0
20	C	5	0	0	0	0
20	D	9	0	0	0	0
20	E	5	0	0	0	0
20	F	4	0	0	0	0
20	G	8	0	0	0	0
20	H	17	0	0	1	0
20	I	7	0	0	0	0
20	J	19	0	0	0	0
20	K	12	0	0	0	0
20	L	21	0	0	0	0
20	M	14	0	0	0	0
20	N	13	0	0	0	0
20	O	2	0	0	0	0
20	P	10	0	0	0	0
20	Q	8	0	0	1	0
20	R	2	0	0	0	0
20	S	4	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
20	T	3	0	0	0	0
20	U	11	0	0	0	0
20	V	12	0	0	0	0
20	W	7	0	0	0	0
20	X	14	0	0	0	0
20	Y	12	0	0	0	0
20	Z	11	0	0	0	0
20	a	12	0	0	0	0
20	b	7	0	0	0	0
20	e	2	0	0	0	0
20	f	1	0	0	0	0
20	h	1	0	0	0	0
20	i	2	0	0	0	0
20	j	1	0	0	0	0
All	All	49935	0	49154	203	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 3.

The worst 5 of 203 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)
10:X:101:ASN:HB3	10:X:133:HIS:CD2	1.80	1.17
10:J:101:ASN:HB3	10:J:133:HIS:CD2	1.82	1.15
10:X:101:ASN:HB3	10:X:133:HIS:HD2	1.35	0.88
3:C:160:GLN:HE21	3:C:160:GLN:HA	1.41	0.84
10:J:101:ASN:HB3	10:J:133:HIS:HD2	1.38	0.82

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	248/250 (99%)	238 (96%)	10 (4%)	0	100	100
1	O	248/250 (99%)	238 (96%)	10 (4%)	0	100	100
2	B	242/258 (94%)	233 (96%)	8 (3%)	1 (0%)	34	66
2	P	242/258 (94%)	233 (96%)	8 (3%)	1 (0%)	34	66
3	C	238/254 (94%)	229 (96%)	6 (2%)	3 (1%)	12	36
3	Q	238/254 (94%)	229 (96%)	6 (2%)	3 (1%)	12	36
4	D	231/260 (89%)	226 (98%)	5 (2%)	0	100	100
4	R	231/260 (89%)	226 (98%)	5 (2%)	0	100	100
5	E	229/234 (98%)	222 (97%)	7 (3%)	0	100	100
5	S	229/234 (98%)	222 (97%)	7 (3%)	0	100	100
6	F	241/288 (84%)	233 (97%)	8 (3%)	0	100	100
6	T	241/288 (84%)	233 (97%)	8 (3%)	0	100	100
7	G	239/252 (95%)	232 (97%)	7 (3%)	0	100	100
7	U	239/252 (95%)	231 (97%)	8 (3%)	0	100	100
8	H	220/232 (95%)	216 (98%)	4 (2%)	0	100	100
8	V	220/232 (95%)	217 (99%)	3 (1%)	0	100	100
9	I	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
9	W	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	J	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	29	61
10	X	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	29	61
11	K	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
11	Y	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
12	L	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
12	Z	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
13	M	231/246 (94%)	224 (97%)	7 (3%)	0	100	100
13	a	231/246 (94%)	224 (97%)	7 (3%)	0	100	100
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
15	g	1/5 (20%)	1 (100%)	0	0	100	100
15	h	1/5 (20%)	1 (100%)	0	0	100	100
16	e	1/5 (20%)	1 (100%)	0	0	100	100
16	f	1/5 (20%)	1 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
16	i	1/5 (20%)	1 (100%)	0	0	100	100
16	j	1/5 (20%)	1 (100%)	0	0	100	100
All	All	6282/6644 (95%)	6102 (97%)	170 (3%)	10 (0%)	47	78

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
3	C	202	GLN
2	P	51	VAL
3	Q	202	GLN
10	J	2	ASP

### 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	209/209 (100%)	206 (99%)	3 (1%)	67	90
1	O	209/209 (100%)	206 (99%)	3 (1%)	67	90
2	B	203/216 (94%)	195 (96%)	8 (4%)	32	66
2	P	203/216 (94%)	195 (96%)	8 (4%)	32	66
3	C	212/226 (94%)	204 (96%)	8 (4%)	33	67
3	Q	212/226 (94%)	204 (96%)	8 (4%)	33	67
4	D	194/215 (90%)	186 (96%)	8 (4%)	30	64
4	R	194/215 (90%)	186 (96%)	8 (4%)	30	64
5	E	190/193 (98%)	182 (96%)	8 (4%)	30	63
5	S	190/193 (98%)	182 (96%)	8 (4%)	30	63
6	F	201/239 (84%)	190 (94%)	11 (6%)	21	52
6	T	201/239 (84%)	190 (94%)	11 (6%)	21	52
7	G	206/210 (98%)	198 (96%)	8 (4%)	32	66
7	U	206/210 (98%)	198 (96%)	8 (4%)	32	66

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
8	H	181/190 (95%)	170 (94%)	11 (6%)	18	48
8	V	181/190 (95%)	170 (94%)	11 (6%)	18	48
9	I	172/173 (99%)	167 (97%)	5 (3%)	42	76
9	W	172/173 (99%)	167 (97%)	5 (3%)	42	76
10	J	173/175 (99%)	166 (96%)	7 (4%)	31	65
10	X	173/175 (99%)	166 (96%)	7 (4%)	31	65
11	K	169/169 (100%)	162 (96%)	7 (4%)	30	64
11	Y	169/169 (100%)	162 (96%)	7 (4%)	30	64
12	L	185/185 (100%)	179 (97%)	6 (3%)	39	73
12	Z	185/185 (100%)	179 (97%)	6 (3%)	39	73
13	M	199/208 (96%)	193 (97%)	6 (3%)	41	75
13	a	199/208 (96%)	193 (97%)	6 (3%)	41	75
14	N	162/162 (100%)	158 (98%)	4 (2%)	47	80
14	b	162/162 (100%)	159 (98%)	3 (2%)	57	85
15	g	1/1 (100%)	1 (100%)	0	100	100
15	h	1/1 (100%)	1 (100%)	0	100	100
16	e	1/1 (100%)	1 (100%)	0	100	100
16	f	1/1 (100%)	0	1 (100%)	0	0
16	i	1/1 (100%)	1 (100%)	0	100	100
16	j	1/1 (100%)	1 (100%)	0	100	100
All	All	5318/5546 (96%)	5118 (96%)	200 (4%)	33	67

5 of 200 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Q	160	GLN
6	T	206	LYS
16	f	3	LEU
4	R	99	ILE
5	S	71	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 127 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	18	ASN
11	Y	138	GLN
3	Q	38	ASN
11	Y	90	ASN
12	Z	165	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

12 non-standard protein/DNA/RNA residues 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
16	A1H2I	e	4	16	14,17,18	4.10	7 (50%)	16,24,26	2.99	2 (12%)
16	A1H2I	j	4	16	14,17,18	4.08	7 (50%)	16,24,26	2.93	4 (25%)
16	A1H2I	i	4	16	14,17,18	3.83	6 (42%)	16,24,26	2.68	3 (18%)
16	A1H2I	f	4	16	14,17,18	3.98	7 (50%)	16,24,26	2.77	4 (25%)
16	A1H2H	e	2	16	13,15,16	3.68	4 (30%)	13,20,22	2.07	2 (15%)
15	A1H2H	g	2	15	13,15,16	3.98	4 (30%)	13,20,22	1.97	4 (30%)
16	A1H2H	j	2	16	13,15,16	4.78	3 (23%)	13,20,22	1.63	2 (15%)
15	A1H45	g	4	15	14,17,18	4.72	7 (50%)	16,24,26	3.63	3 (18%)
15	A1H45	h	4	15	14,17,18	4.28	7 (50%)	16,24,26	2.80	4 (25%)
15	A1H2H	h	2	15	13,15,16	4.08	5 (38%)	13,20,22	2.42	4 (30%)
16	A1H2H	i	2	16	13,15,16	3.29	4 (30%)	13,20,22	1.94	2 (15%)
16	A1H2H	f	2	16	13,15,16	2.92	4 (30%)	13,20,22	2.00	3 (23%)

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
16	A1H2I	e	4	16	-	1/5/18/20	0/2/2/2
16	A1H2I	j	4	16	-	2/5/18/20	0/2/2/2
16	A1H2I	i	4	16	-	1/5/18/20	0/2/2/2
16	A1H2I	f	4	16	-	1/5/18/20	0/2/2/2
16	A1H2H	e	2	16	-	2/7/10/12	0/1/1/1
15	A1H2H	g	2	15	-	3/7/10/12	0/1/1/1
16	A1H2H	j	2	16	-	4/7/10/12	0/1/1/1
15	A1H45	g	4	15	-	1/5/18/20	0/2/2/2
15	A1H45	h	4	15	-	1/5/18/20	0/2/2/2
15	A1H2H	h	2	15	-	2/7/10/12	0/1/1/1
16	A1H2H	i	2	16	-	6/7/10/12	0/1/1/1
16	A1H2H	f	2	16	-	2/7/10/12	0/1/1/1

The worst 5 of 65 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	j	2	A1H2H	O34-N33	14.75	1.47	1.22
15	g	2	A1H2H	F19-CE2	-10.86	1.07	1.35
15	g	4	A1H45	CE2-CD2	-9.46	1.31	1.39
15	g	4	A1H45	CD2-CG	-9.08	1.35	1.51
15	h	4	A1H45	CE2-CD2	-8.98	1.31	1.39

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	e	4	A1H2I	CE2-NE1-CD1	-10.90	105.91	111.85
16	j	4	A1H2I	CE2-NE1-CD1	-10.00	106.40	111.85
15	g	4	A1H45	CD2-CG-CD1	9.79	105.67	102.40
15	g	4	A1H45	CE2-NE1-CD1	-9.72	106.55	111.85
16	f	4	A1H2I	CE2-NE1-CD1	-9.33	106.77	111.85

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	g	2	A1H2H	O-C-CA-CB
15	g	4	A1H45	CA-CB-CG-CD1
15	h	4	A1H45	CA-CB-CG-CD1
16	i	2	A1H2H	C-CA-CB-CG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
16	i	2	A1H2H	N-CA-CB-CG

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 10 are monoatomic - leaving 2 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
19	MES	K	301	-	12,12,12	0.76	0	14,16,16	0.50	0
19	MES	Y	301	-	12,12,12	0.70	0	14,16,16	0.50	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
19	MES	K	301	-	-	2/6/14/14	0/1/1/1
19	MES	Y	301	-	-	3/6/14/14	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
19	K	301	MES	C8-C7-N4-C3
19	K	301	MES	C8-C7-N4-C5
19	Y	301	MES	C7-C8-S-O3S
19	Y	301	MES	C7-C8-S-O2S
19	Y	301	MES	C7-C8-S-O1S

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
19	K	301	MES	1	0
19	Y	301	MES	1	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	250/250 (100%)	-0.30	2 (0%) 86 81	54, 68, 103, 161	0
1	O	250/250 (100%)	-0.18	8 (3%) 47 37	57, 78, 114, 164	0
2	B	244/258 (94%)	-0.18	5 (2%) 65 56	55, 73, 126, 163	0
2	P	244/258 (94%)	-0.13	10 (4%) 37 27	60, 78, 133, 155	0
3	C	240/254 (94%)	-0.19	8 (3%) 46 36	54, 77, 127, 148	0
3	Q	240/254 (94%)	0.07	13 (5%) 25 17	62, 95, 156, 175	0
4	D	235/260 (90%)	-0.37	2 (0%) 84 80	57, 77, 104, 146	0
4	R	235/260 (90%)	-0.24	2 (0%) 84 80	58, 85, 119, 151	0
5	E	231/234 (98%)	-0.22	3 (1%) 77 72	59, 81, 113, 128	0
5	S	231/234 (98%)	-0.08	5 (2%) 62 52	62, 94, 130, 150	0
6	F	243/288 (84%)	-0.29	2 (0%) 86 81	54, 77, 116, 145	0
6	T	243/288 (84%)	-0.13	7 (2%) 51 41	61, 87, 129, 147	0
7	G	241/252 (95%)	-0.37	2 (0%) 86 81	52, 70, 103, 146	0
7	U	241/252 (95%)	-0.40	2 (0%) 86 81	58, 73, 107, 134	0
8	H	222/232 (95%)	-0.39	2 (0%) 84 80	53, 65, 93, 123	0
8	V	222/232 (95%)	-0.34	2 (0%) 84 80	55, 69, 91, 128	0
9	I	204/205 (99%)	-0.46	1 (0%) 91 88	48, 65, 90, 118	0
9	W	204/205 (99%)	-0.45	1 (0%) 91 88	48, 65, 88, 117	0
10	J	195/198 (98%)	-0.35	1 (0%) 91 88	49, 66, 90, 122	0
10	X	195/198 (98%)	-0.44	3 (1%) 73 68	55, 69, 91, 134	0
11	K	212/212 (100%)	-0.35	1 (0%) 91 88	51, 64, 88, 120	0
11	Y	212/212 (100%)	-0.36	1 (0%) 91 88	57, 68, 92, 112	0
12	L	222/222 (100%)	-0.44	0 100 100	50, 67, 90, 105	0
12	Z	222/222 (100%)	-0.43	0 100 100	50, 69, 96, 104	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	M	233/246 (94%)	-0.52	1 (0%) 92 91	50, 66, 87, 103	0
13	a	233/246 (94%)	-0.49	1 (0%) 92 91	50, 66, 85, 94	0
14	N	196/196 (100%)	-0.41	1 (0%) 91 88	51, 62, 91, 112	0
14	b	196/196 (100%)	-0.35	1 (0%) 91 88	52, 64, 92, 115	0
15	g	1/5 (20%)	-0.46	0 100 100	66, 66, 66, 66	0
15	h	1/5 (20%)	-1.08	0 100 100	68, 68, 68, 68	0
16	e	1/5 (20%)	-1.09	0 100 100	60, 60, 60, 60	0
16	f	1/5 (20%)	-0.73	0 100 100	67, 67, 67, 67	0
16	i	1/5 (20%)	-0.86	0 100 100	73, 73, 73, 73	0
16	j	1/5 (20%)	-1.07	0 100 100	71, 71, 71, 71	0
All	All	6342/6644 (95%)	-0.31	87 (1%) 75 70	48, 72, 117, 175	0

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	49	THR	7.7
1	A	1	MET	7.0
3	Q	50	LEU	6.4
3	Q	48	SER	5.3
3	Q	206	LYS	5.3

## 6.2 Non-standard residues in protein, DNA, RNA chains [\(i\)](#)

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
16	A1H2H	j	2	15/16	0.75	0.51	83,111,129,133	0
16	A1H2H	i	2	15/16	0.77	0.50	87,117,142,142	0
16	A1H2I	i	4	16/17	0.83	0.23	59,86,89,94	0
15	A1H2H	h	2	15/16	0.90	0.36	75,93,112,113	0
15	A1H2H	g	2	15/16	0.90	0.40	73,95,105,109	0
16	A1H2H	e	2	15/16	0.91	0.29	64,81,102,111	0
15	A1H45	h	4	16/17	0.91	0.21	59,73,76,82	0
16	A1H2I	j	4	16/17	0.91	0.23	62,84,90,92	0
15	A1H45	g	4	16/17	0.92	0.18	57,72,75,81	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
16	A1H2I	e	4	16/17	0.92	0.22	59,77,81,82	0
16	A1H2I	f	4	16/17	0.94	0.23	61,79,85,90	0
16	A1H2H	f	2	15/16	0.94	0.25	68,80,103,108	0

### 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
17	MG	Z	301	1/1	0.87	0.19	89,89,89,89	0
17	MG	V	301	1/1	0.95	0.26	107,107,107,107	0
17	MG	G	301	1/1	0.97	0.19	81,81,81,81	0
17	MG	N	301	1/1	0.97	0.12	65,65,65,65	0
18	CL	U	301	1/1	0.97	0.06	65,65,65,65	0
19	MES	K	301	12/12	0.97	0.20	48,52,71,82	0
18	CL	G	302	1/1	0.98	0.07	60,60,60,60	0
17	MG	K	302	1/1	0.98	0.13	72,72,72,72	0
17	MG	I	301	1/1	0.98	0.13	76,76,76,76	0
19	MES	Y	301	12/12	0.98	0.17	50,54,67,72	0
17	MG	Y	302	1/1	0.99	0.09	73,73,73,73	0
17	MG	X	201	1/1	0.99	0.32	48,48,48,48	0

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

There are no such residues in this entry.