



# wwPDB X-ray Structure Validation Summary Report

Nov 11, 2023 – 08:39 am GMT

PDB ID : 4Y8S  
Title : Yeast 20S proteasome beta2-H116D mutant in complex with Ac-LAE-ep  
Authors : Huber, E.M.; Groll, M.  
Deposited on : 2015-02-16  
Resolution : 2.70 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)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

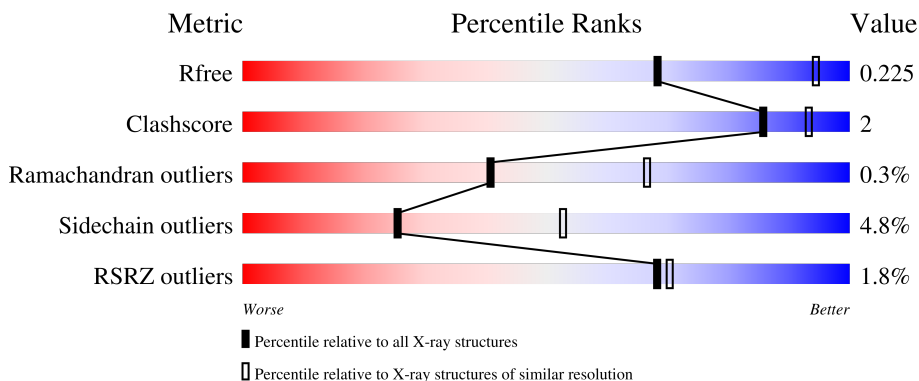
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	250	 2% 97%
1	O	250	 2% 97%
2	B	258	 4% 83% 10% 5%
2	P	258	 3% 84% 10% 5%
3	C	254	 4% 85% 7% 6%





*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	6% 85% 7% • 6%
4	D	260	80% 9% • 10%
4	R	260	2% 80% 10% • 10%
5	E	234	2% 90% 8% •
5	S	234	% 90% 9% •
6	F	288	% 79% 5% • 16%
6	T	288	4% 79% 5% • 16%
7	G	252	2% 88% 6% • •
7	U	252	2% 88% 7% • •
8	H	232	% 89% 6% •
8	V	232	% 90% • • •
9	I	205	91% 9%
9	W	205	% 91% 8%
10	J	198	% 87% 10% • •
10	X	198	% 88% 9% • •
11	K	212	93% 6%
11	Y	212	% 92% 7%
12	L	222	90% 8% •
12	Z	222	89% 10% •
13	M	246	% 84% 9% • 5%
13	a	246	88% 7% 5%
14	N	196	% 95% • •
14	b	196	97% •
15	c	5	80% 20%
15	d	5	80% 20%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
15	e	5	 80% 20%
15	f	5	 80% 20%
15	g	5	 80% 20%
15	h	5	 80% 20%

## 2 Entry composition

There are 18 unique types of molecules in this entry. The entry contains 49825 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	C	N	O	S	0	0	0
			1915	1219	315	377	4			
1	O	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			

- 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	C	N	O	S	0	0	0
			1904	1201	321	379	3			
2	P	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			

- 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	C	N	O	S	0	0	0
			1881	1176	329	372	4			
3	Q	240	Total	C	N	O	S	0	0	0
			1881	1176	329	372	4			

- 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	C	N	O	S	0	0	0
			1813	1136	304	366	7			
4	R	235	Total	C	N	O	S	0	0	0
			1813	1136	304	366	7			

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	222	Total	C	N	O	S	0	0	0
			1682	1059	291	325	7			
8	V	222	Total	C	N	O	S	0	1	0
			1689	1064	293	325	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	116	ASP	HIS	engineered mutation	UNP P25043
V	116	ASP	HIS	engineered mutation	UNP P25043

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

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

- 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
10	X	195	Total 1561	C 992	N 264	O 299	S 6	0	0	0

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	222	Total 1757	C 1115	N 303	O 335	S 4	0	0	0
12	Z	222	Total 1764	C 1120	N 305	O 335	S 4	0	1	0

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

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

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

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

- Molecule 15 is a protein called Ac-LAE-ep.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	c	5	Total	C	N	O	0	0	0
			29	19	3	7			
15	d	5	Total	C	N	O	0	0	0
			29	19	3	7			
15	e	5	Total	C	N	O	0	0	0
			29	19	3	7			
15	f	5	Total	C	N	O	0	0	0
			29	19	3	7			
15	g	5	Total	C	N	O	0	0	0
			29	19	3	7			
15	h	5	Total	C	N	O	0	0	0
			29	19	3	7			

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	G	1	Total	Mg	0	0
			1	1		
16	H	1	Total	Mg	0	0
			1	1		
16	I	2	Total	Mg	0	0
			2	2		
16	K	1	Total	Mg	0	0
			1	1		
16	L	1	Total	Mg	0	0
			1	1		
16	N	1	Total	Mg	0	0
			1	1		

- Molecule 17 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		
17	K	1	12	6	1	4	1	0	0
17	Y	1	12	6	1	4	1	0	0

- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	6	Total	O	0	0
			6	6		
18	B	13	Total	O	0	0
			13	13		
18	C	10	Total	O	0	0
			10	10		
18	D	2	Total	O	0	0
			2	2		
18	E	5	Total	O	0	0
			5	5		
18	F	13	Total	O	0	0
			13	13		
18	G	15	Total	O	0	0
			15	15		
18	H	23	Total	O	0	0
			23	23		
18	I	10	Total	O	0	0
			10	10		
18	J	11	Total	O	0	0
			11	11		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	K	14	Total O 14 14	0	0
18	L	14	Total O 14 14	0	0
18	M	13	Total O 13 13	0	0
18	N	6	Total O 6 6	0	0
18	O	3	Total O 3 3	0	0
18	P	7	Total O 7 7	0	0
18	Q	8	Total O 8 8	0	0
18	R	3	Total O 3 3	0	0
18	S	4	Total O 4 4	0	0
18	T	9	Total O 9 9	0	0
18	U	21	Total O 21 21	0	0
18	V	13	Total O 13 13	0	0
18	W	8	Total O 8 8	0	0
18	X	14	Total O 14 14	0	0
18	Y	16	Total O 16 16	0	0
18	Z	20	Total O 20 20	0	0
18	a	17	Total O 17 17	0	0
18	b	11	Total O 11 11	0	0
18	c	1	Total O 1 1	0	0
18	d	1	Total O 1 1	0	0
18	e	1	Total O 1 1	0	0

*Continued on next page...*

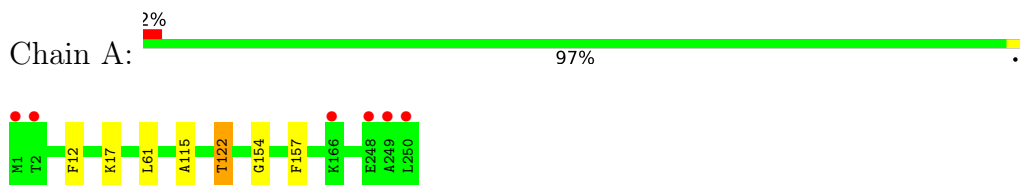
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
18	f	1	Total O 1 1	0	0
18	g	1	Total O 1 1	0	0

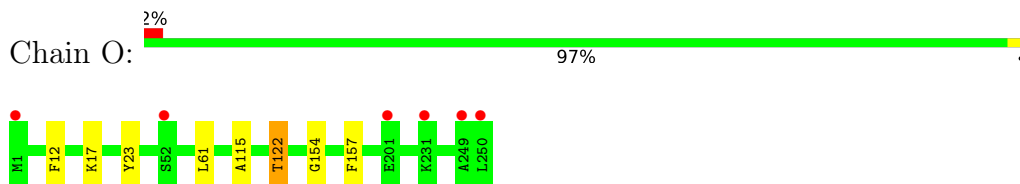
### 3 Residue-property plots [i](#)

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.

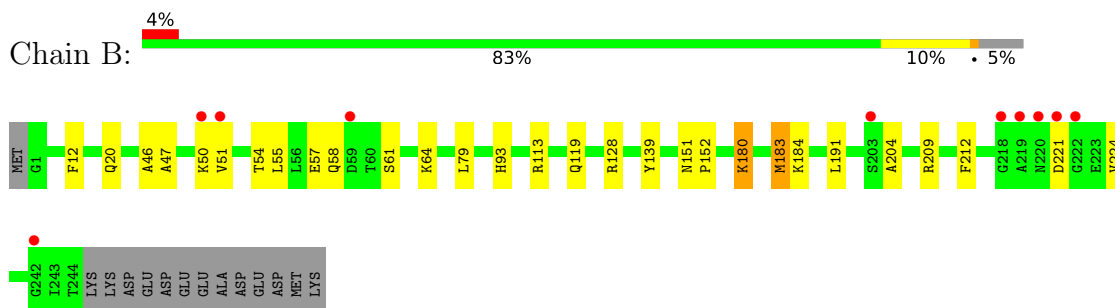
- Molecule 1: Proteasome subunit alpha type-2



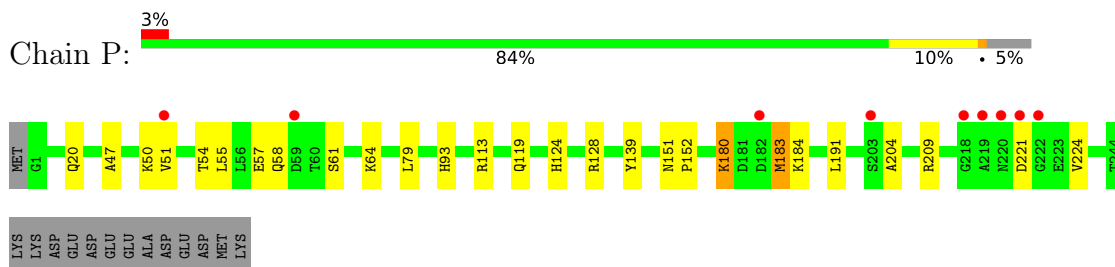
- Molecule 1: Proteasome subunit alpha type-2



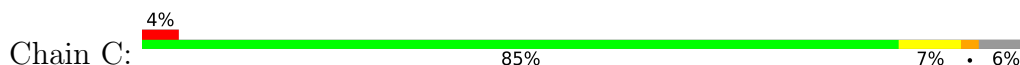
- Molecule 2: Proteasome subunit alpha type-3

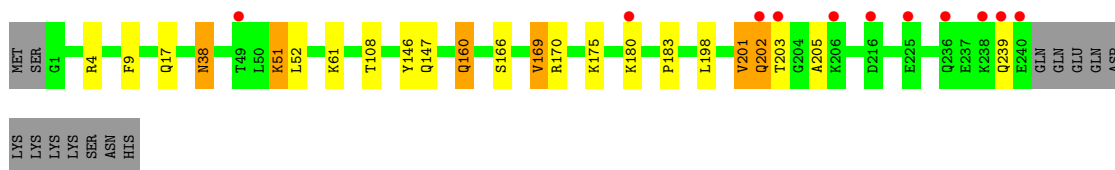


- Molecule 2: Proteasome subunit alpha type-3

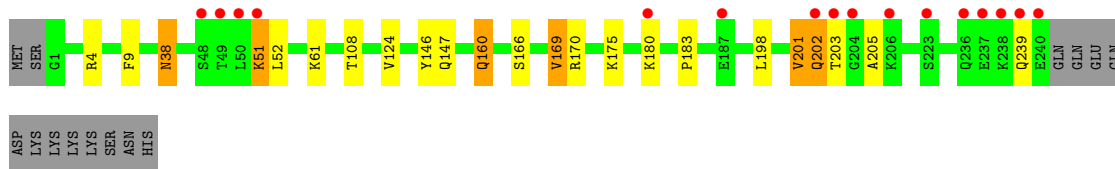
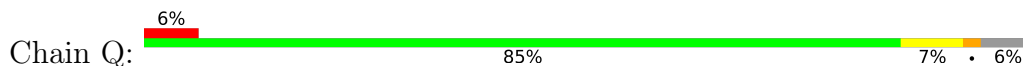


- Molecule 3: Proteasome subunit alpha type-4

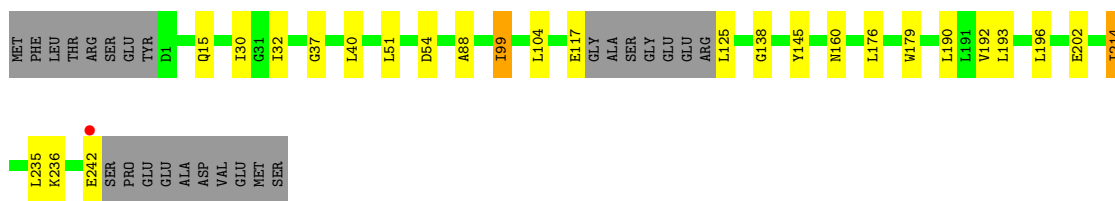
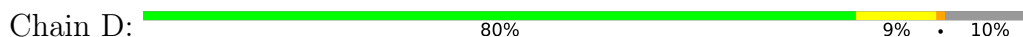




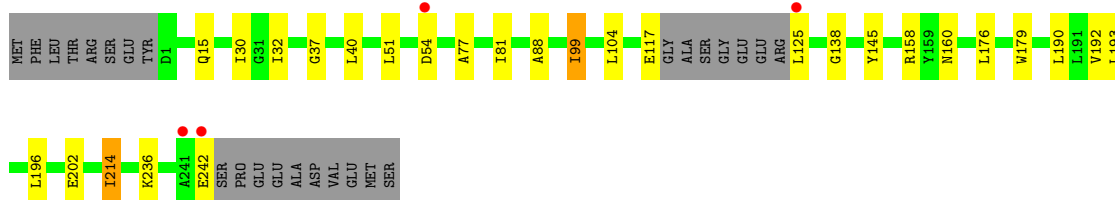
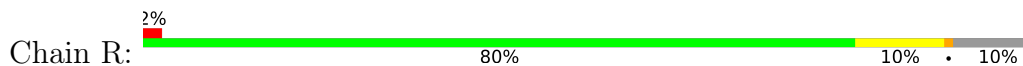
• Molecule 3: Proteasome subunit alpha type-4



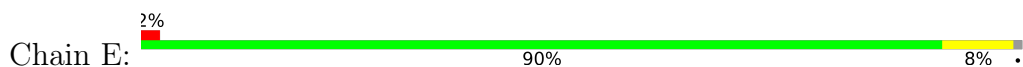
• Molecule 4: Proteasome subunit alpha type-5



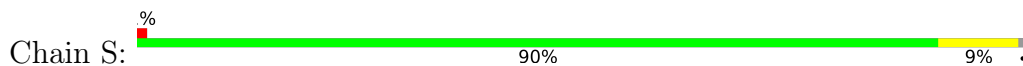
• Molecule 4: Proteasome subunit alpha type-5

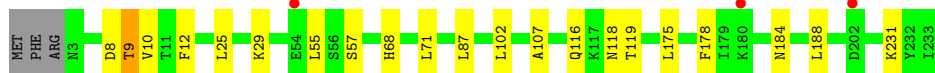


• Molecule 5: Proteasome subunit alpha type-6

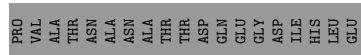
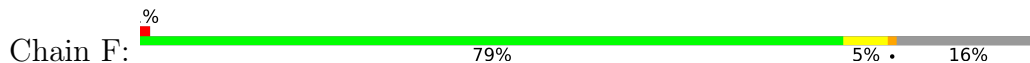


• Molecule 5: Proteasome subunit alpha type-6

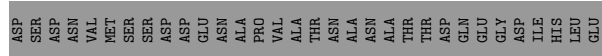
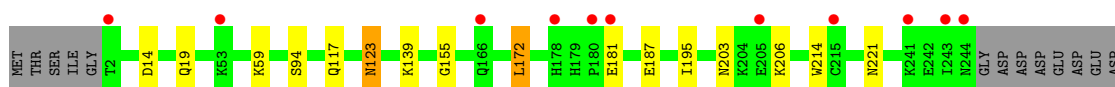
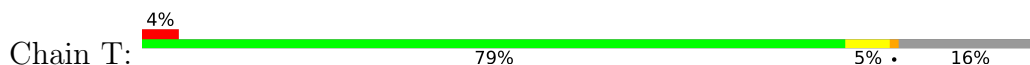




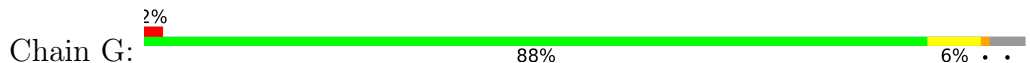
● Molecule 6: Probable proteasome subunit alpha type-7



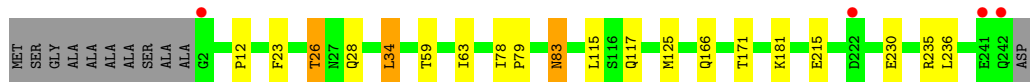
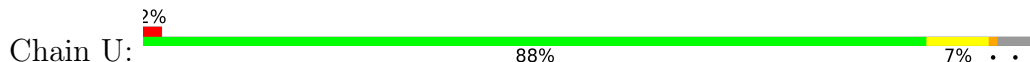
● Molecule 6: Probable proteasome subunit alpha type-7



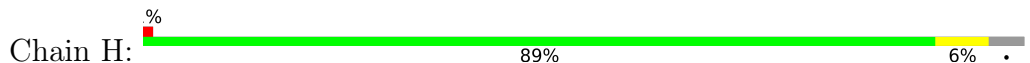
● Molecule 7: Proteasome subunit alpha type-1



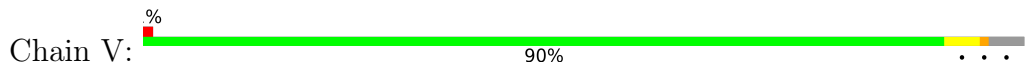
● Molecule 7: Proteasome subunit alpha type-1



● Molecule 8: Proteasome subunit beta type-2



● Molecule 8: Proteasome subunit beta type-2





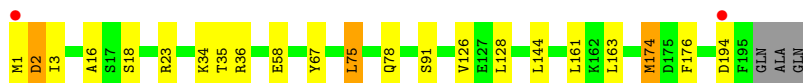
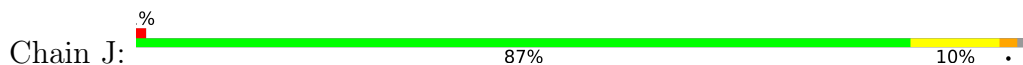
- Molecule 9: Proteasome subunit beta type-3



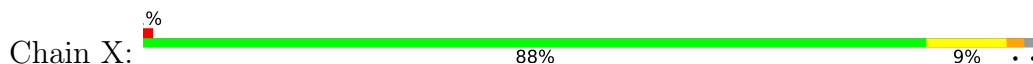
- Molecule 9: Proteasome subunit beta type-3



- 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

Chain L:  90% 8%




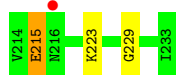
- Molecule 12: Proteasome subunit beta type-6

Chain Z:  89% 10%




- Molecule 13: Proteasome subunit beta type-7

Chain M:  84% 9% 5%



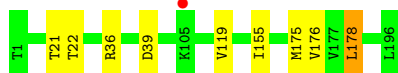
- Molecule 13: Proteasome subunit beta type-7

Chain a:  88% 7% 5%



- Molecule 14: Proteasome subunit beta type-1

Chain N:  95%




- Molecule 14: Proteasome subunit beta type-1

Chain b:  97%



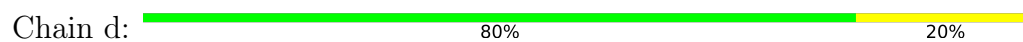
- Molecule 15: Ac-LAE-ep

Chain c:  80% 20%

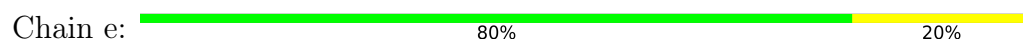




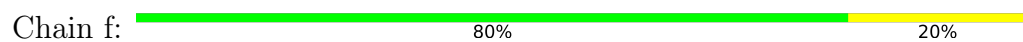
- Molecule 15: Ac-LAE-ep



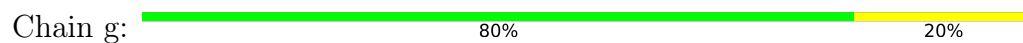
- Molecule 15: Ac-LAE-ep



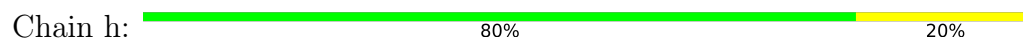
- Molecule 15: Ac-LAE-ep



- Molecule 15: Ac-LAE-ep



- Molecule 15: Ac-LAE-ep



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.35Å 300.08Å 144.78Å 90.00° 112.99° 90.00°	Depositor
Resolution (Å)	15.00 – 2.70 15.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.2 (15.00-2.70) 97.3 (15.00-2.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.63 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.188 , 0.221 0.193 , 0.225	Depositor DCC
$R_{free}$ test set	14123 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.6	Xtrriage
Anisotropy	0.035	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 43.8	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	49825	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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: ACE, POL, GAU, MG, MES

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	0.27	0/1952	0.51	0/2642
1	O	0.27	0/1952	0.51	0/2642
2	B	0.27	0/1934	0.53	0/2618
2	P	0.27	0/1934	0.53	0/2618
3	C	0.28	0/1910	0.55	0/2586
3	Q	0.28	0/1910	0.55	0/2586
4	D	0.27	0/1837	0.53	0/2475
4	R	0.27	0/1837	0.53	0/2475
5	E	0.27	0/1800	0.51	0/2433
5	S	0.27	0/1800	0.51	0/2433
6	F	0.28	0/1932	0.49	0/2609
6	T	0.28	0/1932	0.49	0/2609
7	G	0.28	0/1945	0.52	0/2634
7	U	0.28	0/1945	0.52	0/2634
8	H	0.30	0/1712	0.52	0/2322
8	V	0.27	0/1723	0.56	1/2337 (0.0%)
9	I	0.27	0/1611	0.53	0/2174
9	W	0.27	0/1611	0.52	0/2174
10	J	0.27	0/1589	0.52	0/2142
10	X	0.27	0/1589	0.52	0/2142
11	K	0.29	0/1681	0.53	0/2274
11	Y	0.28	0/1681	0.53	0/2274
12	L	0.27	0/1795	0.51	0/2420
12	Z	0.29	0/1806	0.51	0/2435
13	M	0.28	0/1855	0.56	0/2514
13	a	0.28	0/1855	0.56	0/2514
14	N	0.34	0/1541	0.50	0/2087
14	b	0.32	0/1541	0.50	0/2087
15	c	1.07	0/13	1.18	0/17
15	d	1.03	0/13	1.30	0/17
15	e	0.99	0/13	0.99	0/17
15	f	0.51	0/13	1.26	0/17

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
15	g	0.72	0/13	1.18	0/17
15	h	1.09	0/13	1.23	0/17
All	All	0.28	0/50288	0.53	1/67992 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
8	V	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	V	115	ALA	C-N-CA	5.98	136.66	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
8	V	115	ALA	Peptide

## 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	4	0
1	O	1915	0	1929	5	0
2	B	1904	0	1904	13	0
2	P	1904	0	1904	12	0
3	C	1881	0	1895	12	0
3	Q	1881	0	1895	13	0
4	D	1813	0	1797	8	0
4	R	1813	0	1797	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	1773	0	1775	5	0
5	S	1773	0	1775	6	0
6	F	1892	0	1883	4	0
6	T	1892	0	1883	5	0
7	G	1907	0	1901	5	0
7	U	1907	0	1901	7	0
8	H	1682	0	1682	6	0
8	V	1689	0	1689	4	0
9	I	1581	0	1574	10	0
9	W	1581	0	1574	8	0
10	J	1561	0	1569	11	0
10	X	1561	0	1569	10	0
11	K	1644	0	1592	8	0
11	Y	1644	0	1592	10	0
12	L	1757	0	1711	11	0
12	Z	1764	0	1718	13	0
13	M	1824	0	1832	10	0
13	a	1824	0	1832	0	0
14	N	1512	0	1478	3	0
14	b	1512	0	1478	0	0
15	c	29	0	30	0	0
15	d	29	0	30	0	0
15	e	29	0	30	0	0
15	f	29	0	30	0	0
15	g	29	0	30	0	0
15	h	29	0	30	0	0
16	G	1	0	0	0	0
16	H	1	0	0	0	0
16	I	2	0	0	0	0
16	K	1	0	0	0	0
16	L	1	0	0	0	0
16	N	1	0	0	0	0
17	K	12	0	13	0	0
17	Y	12	0	13	0	0
18	A	6	0	0	0	0
18	B	13	0	0	0	0
18	C	10	0	0	0	0
18	D	2	0	0	0	0
18	E	5	0	0	0	0
18	F	13	0	0	0	0
18	G	15	0	0	0	0
18	H	23	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
18	I	10	0	0	0	0
18	J	11	0	0	0	0
18	K	14	0	0	0	0
18	L	14	0	0	0	0
18	M	13	0	0	0	0
18	N	6	0	0	0	0
18	O	3	0	0	0	0
18	P	7	0	0	0	0
18	Q	8	0	0	0	0
18	R	3	0	0	0	0
18	S	4	0	0	0	0
18	T	9	0	0	0	0
18	U	21	0	0	0	0
18	V	13	0	0	0	0
18	W	8	0	0	0	0
18	X	14	0	0	0	0
18	Y	16	0	0	0	0
18	Z	20	0	0	0	0
18	a	17	0	0	0	0
18	b	11	0	0	0	0
18	c	1	0	0	0	0
18	d	1	0	0	0	0
18	e	1	0	0	0	0
18	f	1	0	0	0	0
18	g	1	0	0	0	0
All	All	49825	0	49264	192	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 2.

The worst 5 of 192 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)
4:R:99:ILE:HD11	4:R:104:LEU:HB2	1.57	0.86
4:D:99:ILE:HD11	4:D:104:LEU:HB2	1.57	0.85
11:Y:5:ALA:HB3	11:Y:100:MET:HE2	1.65	0.79
12:Z:108[B]:HIS:HD2	12:Z:139:GLY:HA3	1.48	0.79
11:K:53:GLN:O	11:K:57:THR:HG23	1.86	0.75

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%)	237 (96%)	11 (4%)	0	100	100
1	O	248/250 (99%)	237 (96%)	11 (4%)	0	100	100
2	B	242/258 (94%)	231 (96%)	9 (4%)	2 (1%)	19	43
2	P	242/258 (94%)	231 (96%)	9 (4%)	2 (1%)	19	43
3	C	238/254 (94%)	229 (96%)	4 (2%)	5 (2%)	7	18
3	Q	238/254 (94%)	229 (96%)	4 (2%)	5 (2%)	7	18
4	D	231/260 (89%)	225 (97%)	6 (3%)	0	100	100
4	R	231/260 (89%)	225 (97%)	6 (3%)	0	100	100
5	E	229/234 (98%)	218 (95%)	11 (5%)	0	100	100
5	S	229/234 (98%)	218 (95%)	11 (5%)	0	100	100
6	F	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
6	T	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
7	G	239/252 (95%)	234 (98%)	5 (2%)	0	100	100
7	U	239/252 (95%)	234 (98%)	5 (2%)	0	100	100
8	H	220/232 (95%)	214 (97%)	6 (3%)	0	100	100
8	V	221/232 (95%)	216 (98%)	5 (2%)	0	100	100
9	I	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
9	W	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
10	J	193/198 (98%)	188 (97%)	4 (2%)	1 (0%)	29	54
10	X	193/198 (98%)	188 (97%)	4 (2%)	1 (0%)	29	54
11	K	210/212 (99%)	203 (97%)	7 (3%)	0	100	100
11	Y	210/212 (99%)	203 (97%)	7 (3%)	0	100	100
12	L	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
12	Z	221/222 (100%)	217 (98%)	4 (2%)	0	100	100
13	M	231/246 (94%)	223 (96%)	6 (3%)	2 (1%)	17	40

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	a	231/246 (94%)	223 (96%)	6 (3%)	2 (1%)	17	40
14	N	194/196 (99%)	186 (96%)	8 (4%)	0	100	100
14	b	194/196 (99%)	186 (96%)	8 (4%)	0	100	100
15	c	2/5 (40%)	2 (100%)	0	0	100	100
15	d	2/5 (40%)	2 (100%)	0	0	100	100
15	e	2/5 (40%)	2 (100%)	0	0	100	100
15	f	2/5 (40%)	2 (100%)	0	0	100	100
15	g	2/5 (40%)	2 (100%)	0	0	100	100
15	h	2/5 (40%)	2 (100%)	0	0	100	100
All	All	6290/6644 (95%)	6083 (97%)	187 (3%)	20 (0%)	41	66

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	202	GLN
3	Q	202	GLN
2	B	51	VAL
3	C	239	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%)	205 (98%)	4 (2%)	57	82
1	O	209/209 (100%)	205 (98%)	4 (2%)	57	82
2	B	203/216 (94%)	193 (95%)	10 (5%)	25	52
2	P	203/216 (94%)	193 (95%)	10 (5%)	25	52
3	C	212/226 (94%)	202 (95%)	10 (5%)	26	54
3	Q	212/226 (94%)	202 (95%)	10 (5%)	26	54
4	D	194/215 (90%)	180 (93%)	14 (7%)	14	34

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	R	194/215 (90%)	181 (93%)	13 (7%)	16	37
5	E	190/193 (98%)	178 (94%)	12 (6%)	18	40
5	S	190/193 (98%)	178 (94%)	12 (6%)	18	40
6	F	201/239 (84%)	188 (94%)	13 (6%)	17	38
6	T	201/239 (84%)	188 (94%)	13 (6%)	17	38
7	G	206/210 (98%)	193 (94%)	13 (6%)	18	40
7	U	206/210 (98%)	193 (94%)	13 (6%)	18	40
8	H	181/190 (95%)	175 (97%)	6 (3%)	38	67
8	V	182/190 (96%)	175 (96%)	7 (4%)	33	62
9	I	172/173 (99%)	168 (98%)	4 (2%)	50	78
9	W	172/173 (99%)	168 (98%)	4 (2%)	50	78
10	J	173/175 (99%)	163 (94%)	10 (6%)	20	43
10	X	173/175 (99%)	163 (94%)	10 (6%)	20	43
11	K	169/169 (100%)	164 (97%)	5 (3%)	41	70
11	Y	169/169 (100%)	164 (97%)	5 (3%)	41	70
12	L	185/185 (100%)	177 (96%)	8 (4%)	29	57
12	Z	186/185 (100%)	179 (96%)	7 (4%)	33	62
13	M	199/208 (96%)	185 (93%)	14 (7%)	15	35
13	a	199/208 (96%)	185 (93%)	14 (7%)	15	35
14	N	162/162 (100%)	157 (97%)	5 (3%)	40	69
14	b	162/162 (100%)	157 (97%)	5 (3%)	40	69
15	c	1/1 (100%)	1 (100%)	0	100	100
15	d	1/1 (100%)	1 (100%)	0	100	100
15	e	1/1 (100%)	1 (100%)	0	100	100
15	f	1/1 (100%)	1 (100%)	0	100	100
15	g	1/1 (100%)	1 (100%)	0	100	100
15	h	1/1 (100%)	1 (100%)	0	100	100
All	All	5320/5546 (96%)	5065 (95%)	255 (5%)	25	53

5 of 255 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	161	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
10	X	163	LEU
3	Q	38	ASN
10	X	91	SER
13	a	10	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 121 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	179	ASN
12	Z	76	HIS
3	Q	147	GLN
12	Z	70	ASN
13	a	213	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

6 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
15	GAU	g	4	15,11	8,8,8	1.25	1 (12%)	8,9,9	1.09	0
15	GAU	c	4	8,15	8,8,8	1.14	0	8,9,9	1.33	1 (12%)
15	GAU	e	4	15,14	8,8,8	1.21	1 (12%)	8,9,9	1.63	2 (25%)
15	GAU	h	4	15,14	8,8,8	1.09	1 (12%)	8,9,9	1.03	0
15	GAU	d	4	15,11	8,8,8	1.37	2 (25%)	8,9,9	1.16	0
15	GAU	f	4	8,15	8,8,8	1.21	1 (12%)	8,9,9	1.80	3 (37%)

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
15	GAU	g	4	15,11	-	0/7/7/7	-
15	GAU	c	4	8,15	-	0/7/7/7	-
15	GAU	e	4	15,14	-	0/7/7/7	-
15	GAU	h	4	15,14	-	0/7/7/7	-
15	GAU	d	4	15,11	-	0/7/7/7	-
15	GAU	f	4	8,15	-	2/7/7/7	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	e	4	GAU	OE2-CD	-2.61	1.22	1.30
15	h	4	GAU	OE2-CD	-2.47	1.22	1.30
15	d	4	GAU	OE2-CD	-2.36	1.22	1.30
15	g	4	GAU	CG-CD	2.17	1.55	1.50
15	f	4	GAU	CG-CD	2.03	1.55	1.50

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	f	4	GAU	CG-CB-CA	-3.49	105.40	112.74
15	e	4	GAU	CB-CA-N	-2.75	101.05	109.03
15	e	4	GAU	CG-CB-CA	-2.72	107.03	112.74
15	c	4	GAU	CB-CA-N	-2.35	102.20	109.03
15	f	4	GAU	OE1-CD-CG	-2.13	116.24	123.08

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	f	4	GAU	C-CA-CB-CG
15	f	4	GAU	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 9 ligands modelled in this entry, 7 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
17	MES	Y	301	-	12,12,12	2.19	1 (8%)	14,16,16	1.31	2 (14%)
17	MES	K	302	-	12,12,12	2.19	1 (8%)	14,16,16	1.35	3 (21%)

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
17	MES	Y	301	-	-	0/6/14/14	0/1/1/1
17	MES	K	302	-	-	0/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	K	302	MES	C8-S	-7.28	1.67	1.77
17	Y	301	MES	C8-S	-7.27	1.67	1.77

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	Y	301	MES	O3S-S-C8	2.75	110.22	105.77
17	K	302	MES	O3S-S-C8	2.67	110.09	105.77
17	Y	301	MES	O2S-S-C8	2.52	109.95	106.92
17	K	302	MES	O2S-S-C8	2.43	109.84	106.92
17	K	302	MES	O1S-S-C8	2.04	109.38	106.92

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 [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%)	-0.38	6 (2%) 59 60	43, 59, 89, 130	0
1	O	250/250 (100%)	-0.32	6 (2%) 59 60	46, 67, 102, 141	0
2	B	244/258 (94%)	-0.27	10 (4%) 37 36	45, 65, 107, 150	0
2	P	244/258 (94%)	-0.21	9 (3%) 41 41	49, 69, 114, 160	0
3	C	240/254 (94%)	-0.11	11 (4%) 32 31	46, 70, 129, 157	0
3	Q	240/254 (94%)	0.09	16 (6%) 17 16	53, 81, 151, 165	0
4	D	235/260 (90%)	-0.35	1 (0%) 92 93	49, 71, 102, 135	0
4	R	235/260 (90%)	-0.27	4 (1%) 70 72	51, 72, 105, 136	0
5	E	231/234 (98%)	-0.25	5 (2%) 62 63	51, 74, 104, 154	0
5	S	231/234 (98%)	-0.19	3 (1%) 77 78	53, 78, 112, 140	0
6	F	243/288 (84%)	-0.42	3 (1%) 79 80	47, 65, 110, 134	0
6	T	243/288 (84%)	-0.31	11 (4%) 33 31	50, 71, 113, 138	0
7	G	241/252 (95%)	-0.39	4 (1%) 70 72	43, 61, 95, 138	0
7	U	241/252 (95%)	-0.41	4 (1%) 70 72	45, 62, 93, 119	0
8	H	222/232 (95%)	-0.50	2 (0%) 84 85	42, 57, 82, 107	0
8	V	222/232 (95%)	-0.46	2 (0%) 84 85	43, 59, 85, 116	0
9	I	204/205 (99%)	-0.64	1 (0%) 91 92	42, 56, 82, 106	0
9	W	204/205 (99%)	-0.58	3 (1%) 73 76	44, 59, 84, 118	0
10	J	195/198 (98%)	-0.49	2 (1%) 82 83	41, 59, 85, 121	0
10	X	195/198 (98%)	-0.50	2 (1%) 82 83	45, 61, 86, 147	0
11	K	212/212 (100%)	-0.49	0 100 100	45, 60, 89, 106	0
11	Y	212/212 (100%)	-0.43	3 (1%) 75 77	47, 63, 98, 123	0
12	L	222/222 (100%)	-0.55	1 (0%) 91 92	45, 61, 88, 109	0
12	Z	222/222 (100%)	-0.54	0 100 100	43, 60, 88, 106	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%)	-0.55	2 (0%) 84 85	41, 59, 77, 91	0
13	a	233/246 (94%)	-0.56	1 (0%) 92 93	43, 57, 76, 90	0
14	N	196/196 (100%)	-0.59	1 (0%) 91 92	42, 55, 80, 108	0
14	b	196/196 (100%)	-0.55	0 100 100	41, 56, 82, 104	0
15	c	2/5 (40%)	-1.35	0 100 100	48, 48, 48, 53	0
15	d	2/5 (40%)	-0.67	0 100 100	60, 60, 60, 61	0
15	e	2/5 (40%)	-0.66	0 100 100	45, 45, 45, 54	0
15	f	2/5 (40%)	-0.88	0 100 100	55, 55, 55, 62	0
15	g	2/5 (40%)	-0.09	0 100 100	72, 72, 72, 75	0
15	h	2/5 (40%)	-0.90	0 100 100	53, 53, 53, 57	0
All	All	6348/6644 (95%)	-0.39	113 (1%) 68 70	41, 63, 103, 165	0

The worst 5 of 113 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	206	LYS	7.6
3	Q	50	LEU	6.4
2	P	51	VAL	6.0
3	Q	206	LYS	5.9
2	P	219	ALA	5.8

## 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
15	GAU	f	4	9/9	0.94	0.14	55,58,65,71	0
15	GAU	c	4	9/9	0.95	0.11	54,58,65,68	0
15	GAU	d	4	9/9	0.96	0.11	41,57,60,61	0
15	GAU	g	4	9/9	0.96	0.13	64,77,81,81	0
15	GAU	h	4	9/9	0.96	0.11	51,52,60,67	0
15	GAU	e	4	9/9	0.97	0.10	49,55,61,64	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(Å <sup>2</sup> )	Q<0.9
17	MES	K	302	12/12	0.91	0.30	91,95,99,100	0
17	MES	Y	301	12/12	0.91	0.28	84,89,98,99	0
16	MG	G	301	1/1	0.96	0.05	57,57,57,57	0
16	MG	I	301	1/1	0.96	0.14	69,69,69,69	0
16	MG	H	301	1/1	0.97	0.17	57,57,57,57	0
16	MG	I	302	1/1	0.98	0.03	55,55,55,55	0
16	MG	N	201	1/1	0.98	0.08	54,54,54,54	0
16	MG	L	301	1/1	0.99	0.03	54,54,54,54	0
16	MG	K	301	1/1	0.99	0.04	61,61,61,61	0

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

There are no such residues in this entry.