



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 05:31 am BST

PDB ID : 4Y8I
Title : Yeast 20S proteasome in complex with Ac-PLL-ep
Authors : Huber, E.M.; Groll, M.
Deposited on : 2015-02-16
Resolution : 2.60 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

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 : 2.11
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) : 2.11

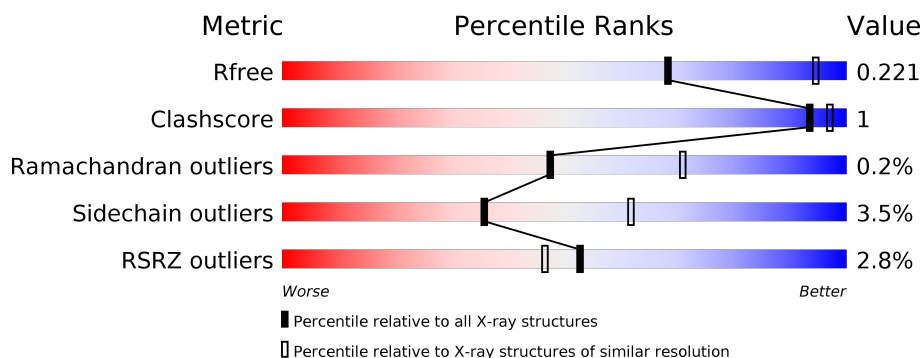
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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.



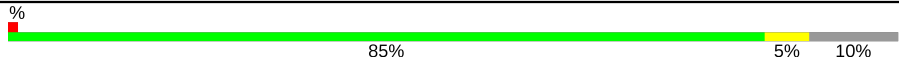
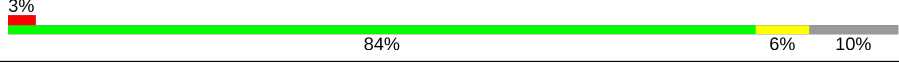
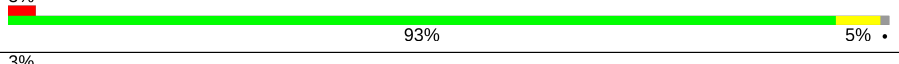
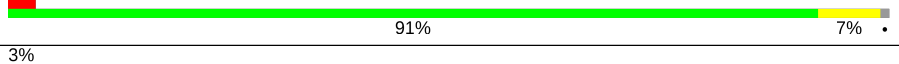

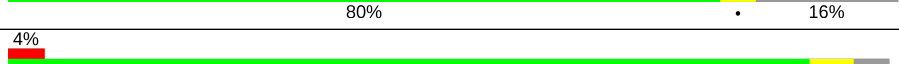
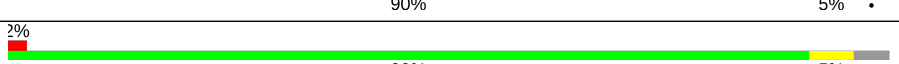
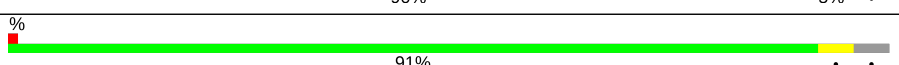
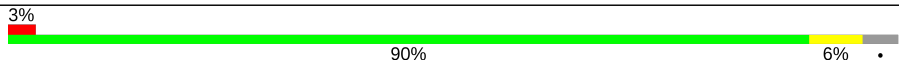
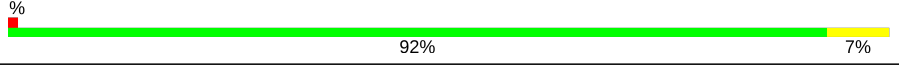
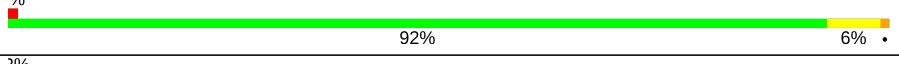
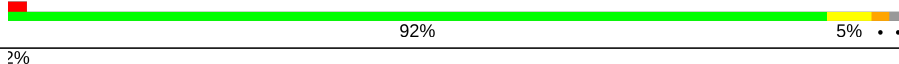
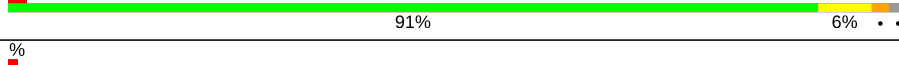
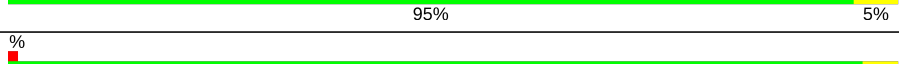
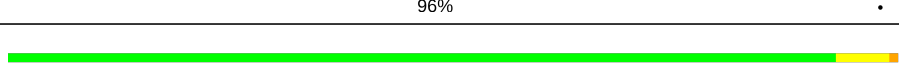
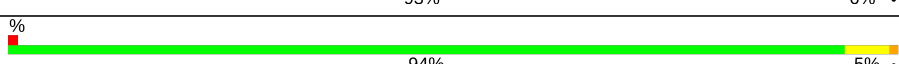

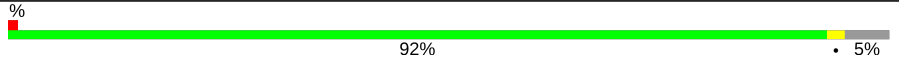
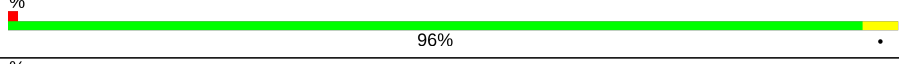
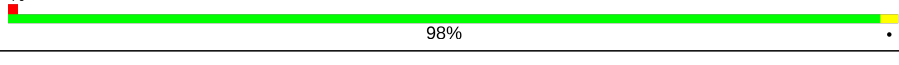





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	<div> <div>2%</div> <div> <div></div> <div>96%</div> <div></div> </div> <div></div> </div>
1	O	250	<div> <div>3%</div> <div> <div></div> <div>98%</div> <div></div> </div> <div></div> </div>
2	B	258	<div> <div>5%</div> <div> <div></div> <div>89%</div> <div>5% 5%</div> </div> <div></div> </div>
2	P	258	<div> <div>6%</div> <div> <div></div> <div>90%</div> <div>• • 5%</div> </div> <div></div> </div>
3	C	254	<div> <div>6%</div> <div> <div></div> <div>88%</div> <div>6% • 6%</div> </div> <div></div> </div>
3	Q	254	<div> <div>9%</div> <div> <div></div> <div>89%</div> <div>5% • 6%</div> </div> <div></div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	260	
4	R	260	
5	E	234	
5	S	234	
6	F	288	
6	T	288	
7	G	252	
7	U	252	
8	H	232	
8	V	232	
9	I	205	
9	W	205	
10	J	198	
10	X	198	
11	K	212	
11	Y	212	
12	L	222	
12	Z	222	
13	M	246	
13	a	246	
14	N	196	
14	b	196	
15	c	5	
15	d	5	
15	e	5	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
15	f	5	<div><div></div><div>40%</div><div>80%</div><div>20%</div></div>
15	g	5	<div><div></div><div>80%</div><div>20%</div></div>
15	h	5	<div><div></div><div>80%</div><div>20%</div></div>

2 Entry composition

There are 19 unique types of molecules in this entry. The entry contains 50137 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
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
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
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
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
			1684	1061	293	323	7			
8	V	222	Total	C	N	O	S	0	0	0
			1684	1061	293	323	7			

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

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	1	0
			1764	1120	305	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 Ac-PLL-ep.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	c	5	Total	C	N	O	0	0	0
			30	22	3	5			
15	d	5	Total	C	N	O	0	0	0
			30	22	3	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	e	5	Total	C	N	O	0	0	0
			30	22	3	5			
15	f	5	Total	C	N	O	0	0	0
			30	22	3	5			
15	g	5	Total	C	N	O	0	0	0
			30	22	3	5			
15	h	5	Total	C	N	O	0	0	0
			30	22	3	5			

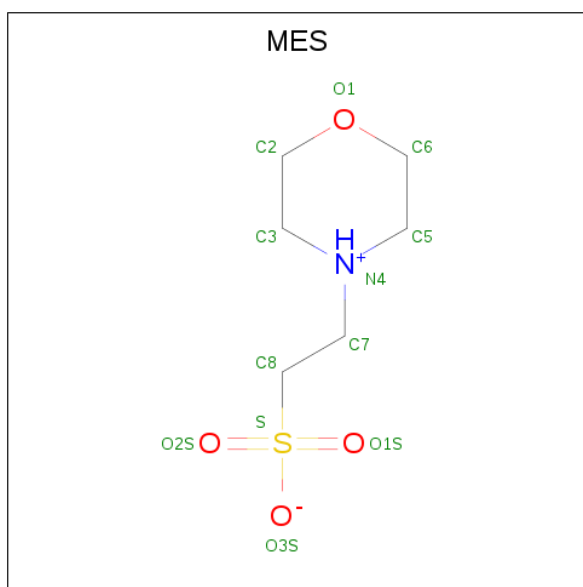
- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 17 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	G	1	Total	Cl	0	0
			1	1		
17	b	1	Total	Cl	0	0
			1	1		
17	N	1	Total	Cl	0	0
			1	1		
17	U	1	Total	Cl	0	0
			1	1		

- Molecule 18 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C₆H₁₃NO₄S).



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

- Molecule 19 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
19	A	29	Total	O	0	0
			29	29		
19	B	21	Total	O	0	0
			21	21		
19	C	14	Total	O	0	0
			14	14		
19	D	16	Total	O	0	0
			16	16		
19	E	14	Total	O	0	0
			14	14		
19	F	17	Total	O	0	0
			17	17		
19	G	29	Total	O	0	0
			29	29		
19	H	24	Total	O	0	0
			24	24		
19	I	34	Total	O	0	0
			34	34		
19	J	22	Total	O	0	0
			22	22		

Continued on next page...

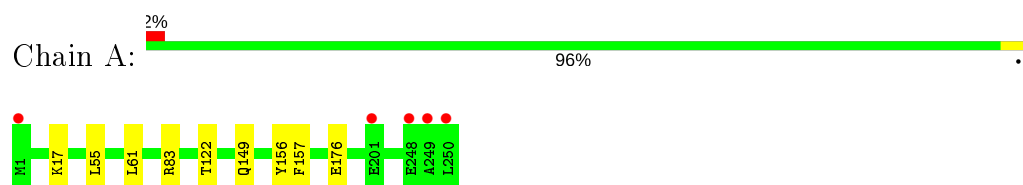
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
19	K	21	Total 21	O 21	0	0
19	L	30	Total 30	O 30	0	0
19	M	28	Total 28	O 28	0	0
19	N	25	Total 25	O 25	0	0
19	O	19	Total 19	O 19	0	0
19	P	11	Total 11	O 11	0	0
19	Q	11	Total 11	O 11	0	0
19	R	13	Total 13	O 13	0	0
19	S	10	Total 10	O 10	0	0
19	T	14	Total 14	O 14	0	0
19	U	36	Total 36	O 36	0	0
19	V	18	Total 18	O 18	0	0
19	W	19	Total 19	O 19	0	0
19	X	28	Total 28	O 28	0	0
19	Y	25	Total 25	O 25	0	0
19	Z	28	Total 28	O 28	0	0
19	a	34	Total 34	O 34	0	0
19	b	25	Total 25	O 25	0	0
19	c	1	Total 1	O 1	0	0
19	e	2	Total 2	O 2	0	0
19	h	1	Total 1	O 1	0	0

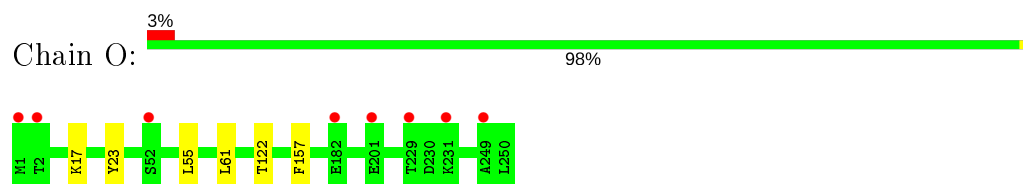
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

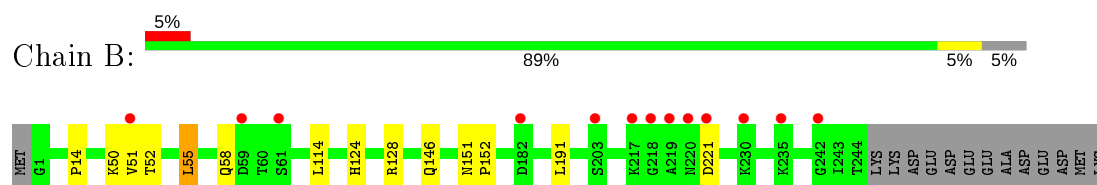
- Molecule 1: Proteasome subunit alpha type-2



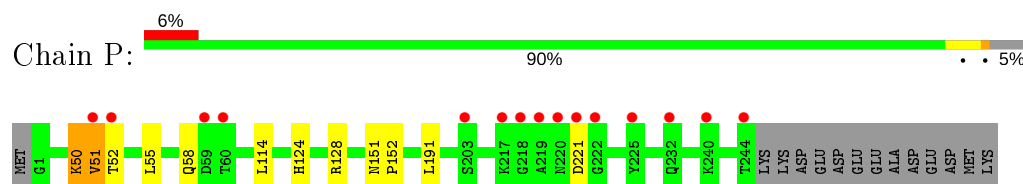
- Molecule 1: Proteasome subunit alpha type-2



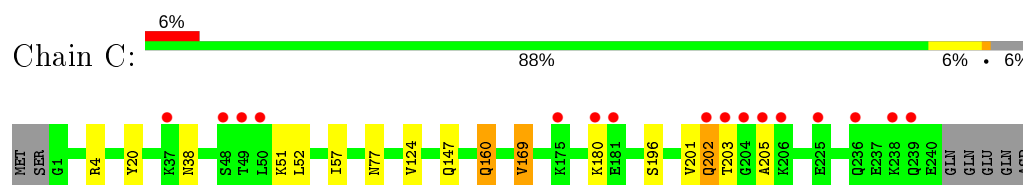
- Molecule 2: Proteasome subunit alpha type-3



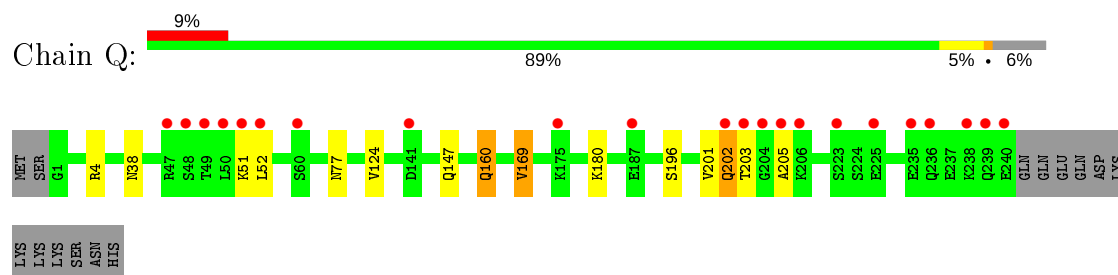
- Molecule 2: Proteasome subunit alpha type-3



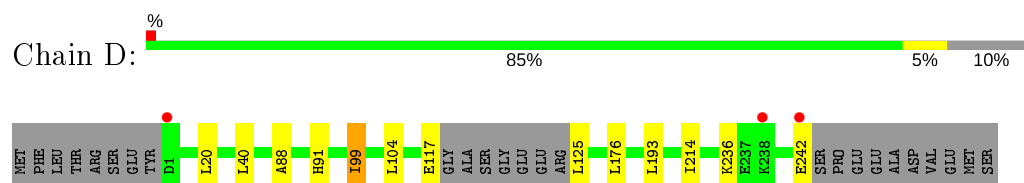
- Molecule 3: Proteasome subunit alpha type-4



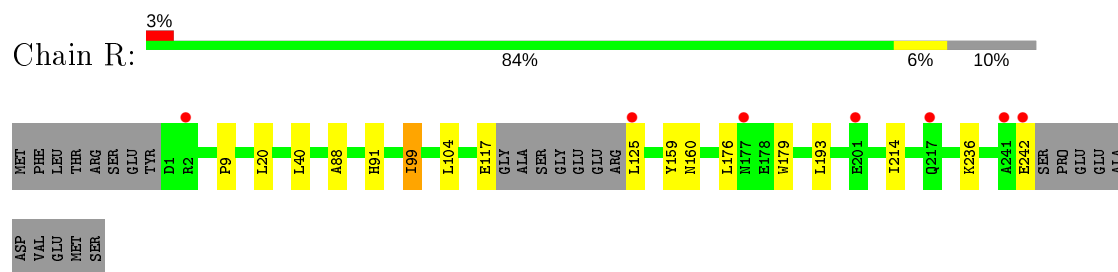
- Molecule 3: Proteasome subunit alpha type-4



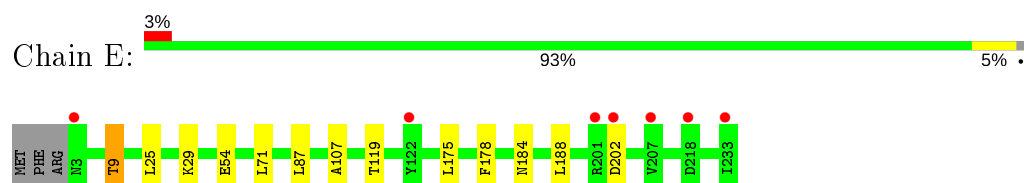
- Molecule 4: Proteasome subunit alpha type-5



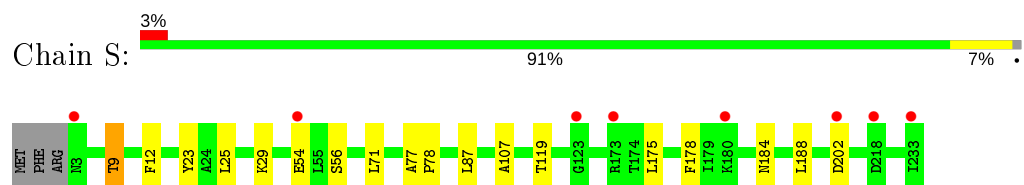
- Molecule 4: Proteasome subunit alpha type-5



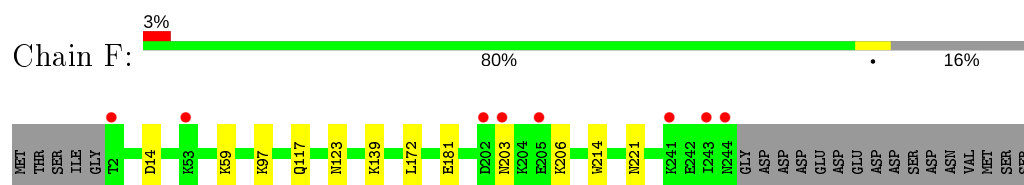
- Molecule 5: Proteasome subunit alpha type-6



- Molecule 5: Proteasome subunit alpha type-6

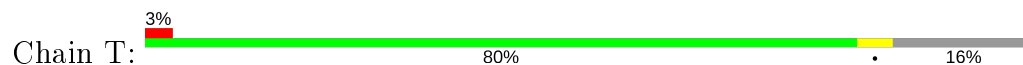


- Molecule 6: Probable proteasome subunit alpha type-7



PRO VAL ALA THR THR ASN ASN ALA ALA THR THR ASP GLN GLU GLY ASP ILE HIS LEU GLU

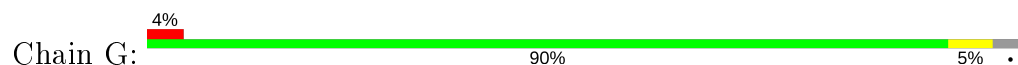
- Molecule 6: Probable proteasome subunit alpha type-7



MET THR SER ILE GLY T2 D14 Q19 K69 Q117 N123 K139 L172 P180 E181 N203 K204 E205 K206 W214 N221 D230 D237 E241 E242 L243 N244 GLY ASP ASP ASP GLU GLU GLU ASP ASP SER ASP ASP ASN VAL MET SER SER ASP ASP GLU

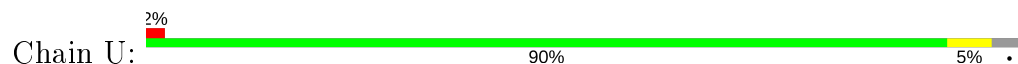
ASN ALA PRO VAL THR ASN ALA ALA THR THR ASP GLN GLU GLY ASP ILE HIS LEU GLU

- Molecule 7: Proteasome subunit alpha type-1



MET SER GLY ALA ALA ALA SER ALA ALA G2 Y3 F23 T26 I78 P79 N83 L115 M125 G152 Q166 T170 K181 N186 E187 E188 D222 E230 R235 L236 A240 E241 Q242 ASP

- Molecule 7: Proteasome subunit alpha type-1



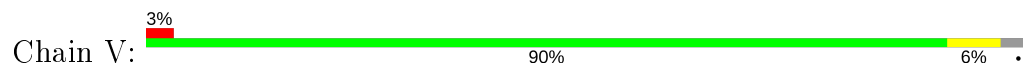
MET SER GLY ALA ALA ALA SER ALA ALA G2 P12 F23 T26 I78 P79 N83 L115 M125 A159 Q166 K181 D203 D222 P235 L236 E241 Q242 ASP

- Molecule 8: Proteasome subunit beta type-2



T1 Q22 N30 C43 T56 L68 D104 P105 L127 R196 E215 V218 C221 D222 ILE GLN GLU GLU GLN VAL ASP THR ALA

- Molecule 8: Proteasome subunit beta type-2



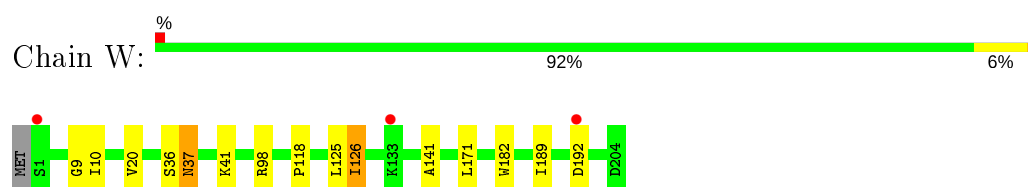
T1 T2 Q22 G23 A27 N30 C43 T56 L68 D104 P105 L127 D145 R196 R207 E215 S216 T217 C221 D222 ILE GLN GLU GLU GLN VAL ASP ILE THR ALA

- Molecule 9: Proteasome subunit beta type-3

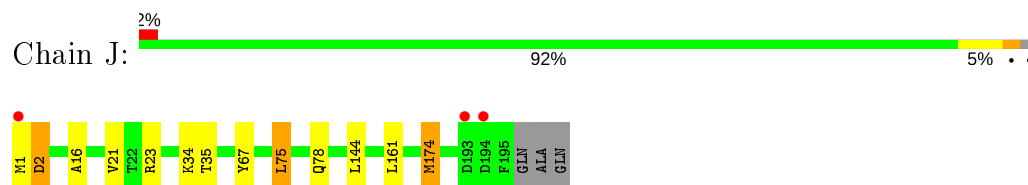


MET S1 G9 I10 V20 N37 K41 R93 P118 L125 I126 E131 A132 K133 A141 L171 W182 S216 I189 D192 K196 P204

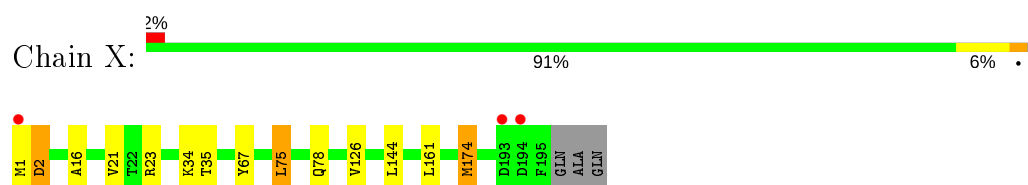
- Molecule 9: Proteasome subunit beta type-3



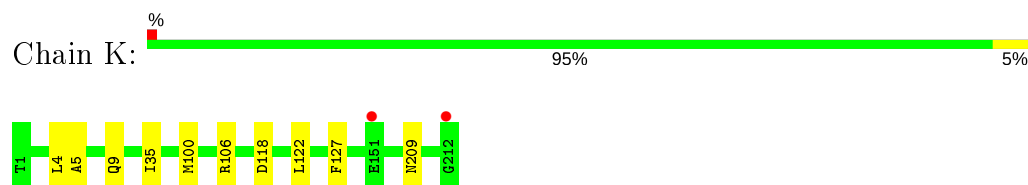
- Molecule 10: Proteasome subunit beta type-4



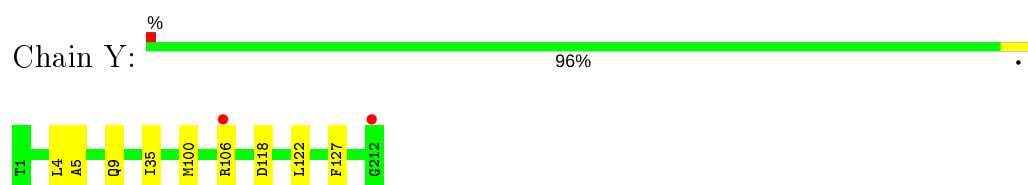
- Molecule 10: Proteasome subunit beta type-4



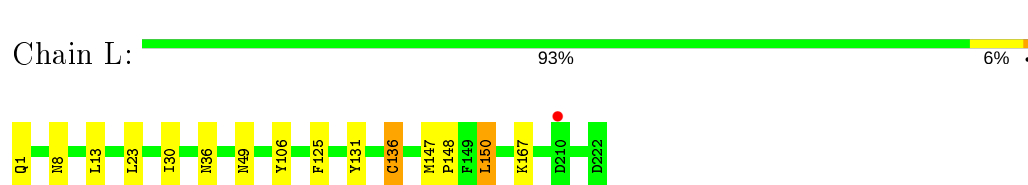
- Molecule 11: Proteasome subunit beta type-5



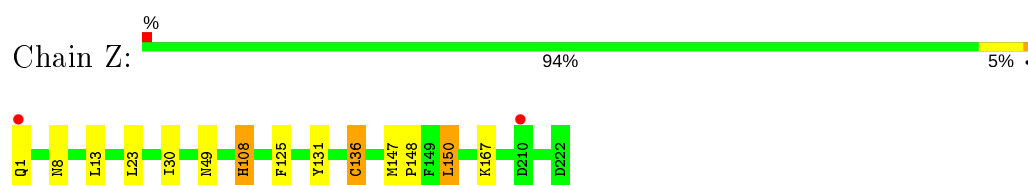
- Molecule 11: Proteasome subunit beta type-5



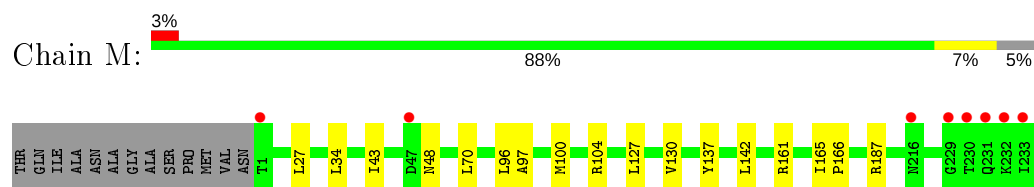
- Molecule 12: Proteasome subunit beta type-6



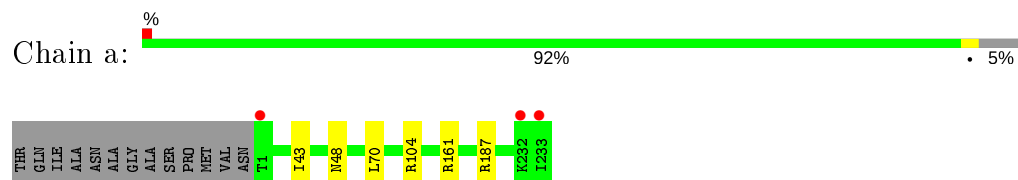
- Molecule 12: Proteasome subunit beta type-6



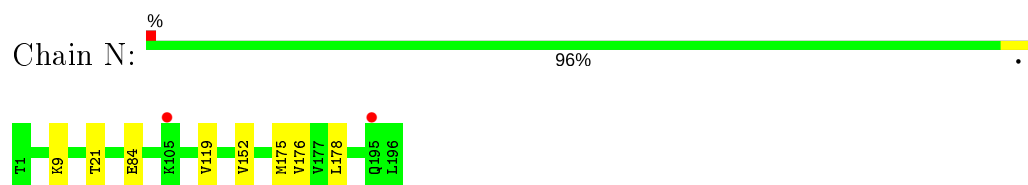
- Molecule 13: Proteasome subunit beta type-7



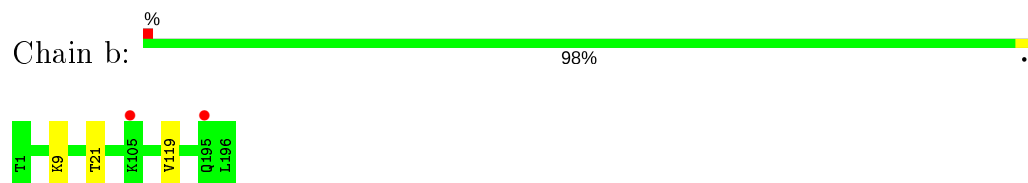
- Molecule 13: Proteasome subunit beta type-7



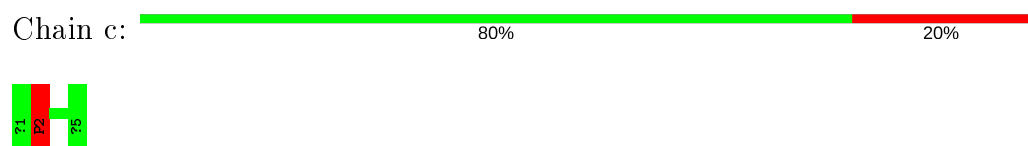
- Molecule 14: Proteasome subunit beta type-1



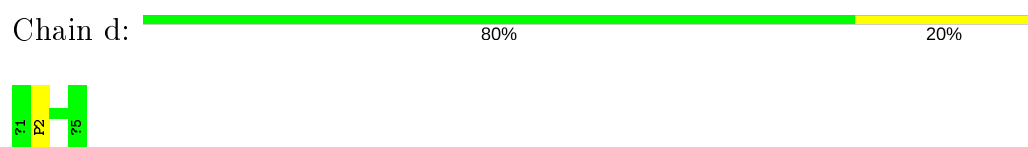
- Molecule 14: Proteasome subunit beta type-1



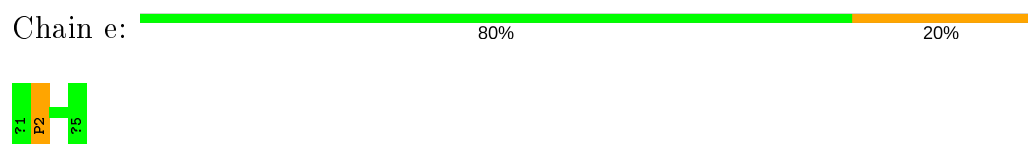
- Molecule 15: Ac-PLL-ep



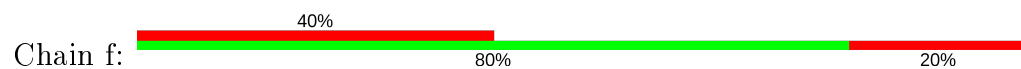
- Molecule 15: Ac-PLL-ep



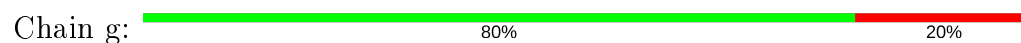
- Molecule 15: Ac-PLL-ep



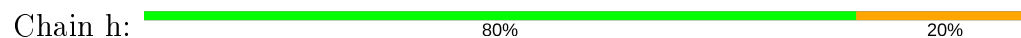
● Molecule 15: Ac-PLL-ep



● Molecule 15: Ac-PLL-ep



● Molecule 15: Ac-PLL-ep



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	136.82Å 298.32Å 145.62Å 90.00° 113.06° 90.00°	Depositor
Resolution (Å)	15.00 – 2.60 14.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	97.9 (15.00-2.60) 98.0 (14.99-2.60)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.68 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.199 , 0.220 0.202 , 0.221	Depositor DCC
R_{free} test set	15979 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	50.5	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 35.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	50137	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MG, ACE, CL, DCL, POL, 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	0.48	0/2642
1	O	0.27	0/1952	0.48	0/2642
2	B	0.27	0/1934	0.49	0/2618
2	P	0.27	0/1934	0.49	0/2618
3	C	0.27	0/1910	0.51	0/2586
3	Q	0.27	0/1910	0.51	0/2586
4	D	0.27	0/1837	0.48	0/2475
4	R	0.27	0/1837	0.48	0/2475
5	E	0.27	0/1800	0.47	0/2433
5	S	0.27	0/1800	0.48	0/2433
6	F	0.27	0/1932	0.45	0/2609
6	T	0.27	0/1932	0.45	0/2609
7	G	0.27	0/1945	0.48	0/2634
7	U	0.27	0/1945	0.47	0/2634
8	H	0.28	0/1715	0.48	0/2326
8	V	0.29	0/1715	0.49	0/2326
9	I	0.27	0/1611	0.50	0/2174
9	W	0.27	0/1611	0.49	0/2174
10	J	0.26	0/1589	0.49	0/2142
10	X	0.26	0/1589	0.49	0/2142
11	K	0.27	0/1681	0.49	0/2274
11	Y	0.28	0/1681	0.50	0/2274
12	L	0.27	0/1795	0.48	0/2420
12	Z	0.38	2/1806 (0.1%)	0.55	4/2435 (0.2%)
13	M	0.28	0/1855	0.51	0/2514
13	a	0.28	0/1855	0.51	0/2514
14	N	0.34	0/1541	0.49	0/2087
14	b	0.32	0/1541	0.49	0/2087
15	c	1.77	1/16 (6.2%)	2.08	1/22 (4.5%)
15	d	2.95	2/16 (12.5%)	1.46	0/22
15	e	2.47	1/16 (6.2%)	1.32	0/22
15	f	1.78	1/16 (6.2%)	1.95	1/22 (4.5%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
15	g	2.57	1/16 (6.2%)	1.83	0/22
15	h	2.58	1/16 (6.2%)	1.45	0/22
All	All	0.30	9/50301 (0.0%)	0.49	6/68015 (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
12	Z	0	2

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	h	2	PRO	CA-C	-8.07	1.36	1.52
15	e	2	PRO	CA-C	-7.79	1.37	1.52
15	g	2	PRO	CA-C	-7.70	1.37	1.52
15	d	2	PRO	CA-C	-7.28	1.38	1.52
15	d	2	PRO	N-CD	-6.86	1.38	1.47

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	c	2	PRO	CA-N-CD	-6.87	101.88	111.50
15	f	2	PRO	CA-N-CD	-6.21	102.81	111.50
12	Z	108[A]	HIS	CA-C-N	-5.55	105.00	117.20
12	Z	108[B]	HIS	CA-C-N	-5.55	105.00	117.20
12	Z	108[A]	HIS	CA-C-O	5.53	131.71	120.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
12	Z	108[A]	HIS	Mainchain
12	Z	108[B]	HIS	Mainchain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1915	0	1929	5	0
1	O	1915	0	1929	3	0
2	B	1904	0	1904	8	0
2	P	1904	0	1904	5	0
3	C	1881	0	1895	10	0
3	Q	1881	0	1895	7	0
4	D	1813	0	1797	3	0
4	R	1813	0	1797	6	0
5	E	1773	0	1775	3	0
5	S	1773	0	1775	7	0
6	F	1892	0	1883	1	0
6	T	1892	0	1883	1	0
7	G	1907	0	1901	4	0
7	U	1907	0	1901	4	0
8	H	1684	0	1685	4	0
8	V	1684	0	1685	4	0
9	I	1581	0	1574	10	0
9	W	1581	0	1574	10	0
10	J	1561	0	1569	7	0
10	X	1561	0	1569	8	0
11	K	1644	0	1592	6	0
11	Y	1644	0	1592	4	0
12	L	1757	0	1711	6	0
12	Z	1764	0	1718	5	0
13	M	1824	0	1832	6	0
13	a	1824	0	1832	0	0
14	N	1512	0	1478	3	0
14	b	1512	0	1478	0	0
15	c	30	0	38	0	0
15	d	30	0	37	0	0
15	e	30	0	37	0	0
15	f	30	0	39	0	0
15	g	30	0	37	0	0
15	h	30	0	37	0	0
16	G	1	0	0	0	0
16	I	2	0	0	0	0
16	K	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	L	1	0	0	0	0
16	N	1	0	0	0	0
16	Z	1	0	0	0	0
17	G	1	0	0	0	0
17	N	1	0	0	0	0
17	U	1	0	0	0	0
17	b	1	0	0	0	0
18	K	12	0	13	0	0
18	Y	12	0	13	0	0
19	A	29	0	0	0	0
19	B	21	0	0	0	0
19	C	14	0	0	0	0
19	D	16	0	0	0	0
19	E	14	0	0	0	0
19	F	17	0	0	0	0
19	G	29	0	0	0	0
19	H	24	0	0	0	0
19	I	34	0	0	0	0
19	J	22	0	0	0	0
19	K	21	0	0	0	0
19	L	30	0	0	0	0
19	M	28	0	0	0	0
19	N	25	0	0	0	0
19	O	19	0	0	0	0
19	P	11	0	0	0	0
19	Q	11	0	0	0	0
19	R	13	0	0	0	0
19	S	10	0	0	0	0
19	T	14	0	0	0	0
19	U	36	0	0	0	0
19	V	18	0	0	0	0
19	W	19	0	0	0	0
19	X	28	0	0	0	0
19	Y	25	0	0	0	0
19	Z	28	0	0	0	0
19	a	34	0	0	0	0
19	b	25	0	0	0	0
19	c	1	0	0	0	0
19	e	2	0	0	0	0
19	h	1	0	0	0	0
All	All	50137	0	49308	116	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 1.

The worst 5 of 116 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:J:1:MET:O	10:J:2:ASP:HB2	1.86	0.73
10:X:1:MET:O	10:X:2:ASP:HB2	1.87	0.73
9:I:98:ARG:HD2	9:I:126:ILE:HD12	1.71	0.72
9:W:98:ARG:HD2	9:W:126:ILE:HD12	1.71	0.71
12:L:13:LEU:HD13	12:L:150:LEU:HD21	1.75	0.69

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	248/250 (99%)	242 (98%)	6 (2%)	0	100	100
1	O	248/250 (99%)	242 (98%)	6 (2%)	0	100	100
2	B	242/258 (94%)	234 (97%)	6 (2%)	2 (1%)	19	39
2	P	242/258 (94%)	234 (97%)	6 (2%)	2 (1%)	19	39
3	C	238/254 (94%)	230 (97%)	6 (2%)	2 (1%)	19	39
3	Q	238/254 (94%)	230 (97%)	6 (2%)	2 (1%)	19	39
4	D	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
4	R	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
5	E	229/234 (98%)	225 (98%)	4 (2%)	0	100	100
5	S	229/234 (98%)	225 (98%)	4 (2%)	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

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	U	239/252 (95%)	235 (98%)	4 (2%)	0	100	100
8	H	220/232 (95%)	215 (98%)	5 (2%)	0	100	100
8	V	220/232 (95%)	214 (97%)	6 (3%)	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%)	190 (98%)	2 (1%)	1 (0%)	29	52
10	X	193/198 (98%)	190 (98%)	2 (1%)	1 (0%)	29	52
11	K	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
11	Y	210/212 (99%)	207 (99%)	3 (1%)	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%)	225 (97%)	6 (3%)	0	100	100
13	a	231/246 (94%)	225 (97%)	6 (3%)	0	100	100
14	N	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
15	c	2/5 (40%)	1 (50%)	0	1 (50%)	0	0
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%)	1 (50%)	0	1 (50%)	0	0
15	g	2/5 (40%)	1 (50%)	0	1 (50%)	0	0
15	h	2/5 (40%)	2 (100%)	0	0	100	100
All	All	6289/6644 (95%)	6135 (98%)	141 (2%)	13 (0%)	47	71

5 of 13 Ramachandran outliers are listed below:

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

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	209/209 (100%)	206 (99%)	3 (1%)	67	85
1	O	209/209 (100%)	206 (99%)	3 (1%)	67	85
2	B	203/216 (94%)	197 (97%)	6 (3%)	41	67
2	P	203/216 (94%)	197 (97%)	6 (3%)	41	67
3	C	212/226 (94%)	204 (96%)	8 (4%)	33	59
3	Q	212/226 (94%)	204 (96%)	8 (4%)	33	59
4	D	194/215 (90%)	184 (95%)	10 (5%)	23	46
4	R	194/215 (90%)	184 (95%)	10 (5%)	23	46
5	E	190/193 (98%)	182 (96%)	8 (4%)	30	55
5	S	190/193 (98%)	182 (96%)	8 (4%)	30	55
6	F	201/239 (84%)	190 (94%)	11 (6%)	21	43
6	T	201/239 (84%)	190 (94%)	11 (6%)	21	43
7	G	206/210 (98%)	199 (97%)	7 (3%)	37	63
7	U	206/210 (98%)	199 (97%)	7 (3%)	37	63
8	H	181/190 (95%)	174 (96%)	7 (4%)	32	58
8	V	181/190 (95%)	175 (97%)	6 (3%)	38	64
9	I	172/173 (99%)	167 (97%)	5 (3%)	42	68
9	W	172/173 (99%)	167 (97%)	5 (3%)	42	68
10	J	173/175 (99%)	167 (96%)	6 (4%)	36	62
10	X	173/175 (99%)	167 (96%)	6 (4%)	36	62
11	K	169/169 (100%)	164 (97%)	5 (3%)	41	67
11	Y	169/169 (100%)	164 (97%)	5 (3%)	41	67
12	L	185/185 (100%)	178 (96%)	7 (4%)	33	59
12	Z	186/185 (100%)	180 (97%)	6 (3%)	39	65
13	M	199/208 (96%)	193 (97%)	6 (3%)	41	67
13	a	199/208 (96%)	193 (97%)	6 (3%)	41	67

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
14	N	162/162 (100%)	159 (98%)	3 (2%)	57	79
14	b	162/162 (100%)	159 (98%)	3 (2%)	57	79
15	c	2/2 (100%)	1 (50%)	1 (50%)	0	0
15	d	2/2 (100%)	2 (100%)	0	100	100
15	e	2/2 (100%)	1 (50%)	1 (50%)	0	0
15	f	2/2 (100%)	1 (50%)	1 (50%)	0	0
15	g	2/2 (100%)	1 (50%)	1 (50%)	0	0
15	h	2/2 (100%)	1 (50%)	1 (50%)	0	0
All	All	5325/5552 (96%)	5138 (96%)	187 (4%)	36	62

5 of 187 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	48	ASN
3	Q	147	GLN
12	Z	167	LYS
13	M	161	ARG
1	O	157	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	102	GLN
3	Q	120	GLN
12	Z	79	HIS
13	M	194	ASN
2	P	123	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	DCL	h	4	15,14	7,7,7	1.03	0	6,8,8	1.78	1 (16%)
15	DCL	d	4	11,15	7,7,7	0.74	0	6,8,8	1.33	1 (16%)
15	DCL	g	4	11,15	7,7,7	0.94	1 (14%)	6,8,8	1.38	1 (16%)
15	DCL	e	4	15,14	7,7,7	1.21	1 (14%)	6,8,8	1.81	1 (16%)
15	DCL	c	4	8,15	7,7,7	0.53	0	6,8,8	1.32	1 (16%)
15	DCL	f	4	8,15	7,7,7	0.34	0	6,8,8	1.48	1 (16%)

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	DCL	h	4	15,14	-	4/6/6/6	-
15	DCL	d	4	11,15	-	2/6/6/6	-
15	DCL	g	4	11,15	-	3/6/6/6	-
15	DCL	e	4	15,14	-	4/6/6/6	-
15	DCL	c	4	8,15	-	3/6/6/6	-
15	DCL	f	4	8,15	-	4/6/6/6	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	e	4	DCL	OXT-C	-2.39	1.32	1.42
15	g	4	DCL	OXT-C	-2.23	1.33	1.42

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	h	4	DCL	CB-CA-C	-3.89	105.41	111.79
15	e	4	DCL	CB-CA-C	-3.71	105.71	111.79
15	f	4	DCL	CB-CA-C	-2.94	106.97	111.79
15	c	4	DCL	CB-CA-C	-2.65	107.45	111.79
15	g	4	DCL	CB-CA-C	-2.52	107.66	111.79

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	h	4	DCL	N-CA-CB-CG
15	h	4	DCL	C-CA-CB-CG
15	d	4	DCL	N-CA-CB-CG
15	d	4	DCL	C-CA-CB-CG
15	g	4	DCL	N-CA-CB-CG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 11 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
18	MES	K	302	-	12,12,12	2.22	1 (8%)	14,16,16	1.26	2 (14%)
18	MES	Y	301	-	12,12,12	2.21	1 (8%)	14,16,16	1.30	2 (14%)

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
18	MES	K	302	-	-	0/6/14/14	0/1/1/1
18	MES	Y	301	-	-	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)
18	K	302	MES	C8-S	-7.41	1.67	1.77
18	Y	301	MES	C8-S	-7.37	1.67	1.77

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
18	K	302	MES	O3S-S-C8	2.68	110.10	105.77
18	Y	301	MES	O2S-S-C8	2.67	110.13	106.92
18	Y	301	MES	O3S-S-C8	2.60	109.98	105.77
18	K	302	MES	O2S-S-C8	2.00	109.33	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 [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, 95th 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 ²)	Q<0.9
1	A	250/250 (100%)	-0.28	5 (2%) 65 60	34, 47, 83, 127	0
1	O	250/250 (100%)	-0.20	8 (3%) 47 40	37, 53, 95, 127	0
2	B	244/258 (94%)	-0.09	13 (5%) 26 20	34, 52, 101, 154	0
2	P	244/258 (94%)	-0.07	15 (6%) 21 16	37, 55, 104, 156	0
3	C	240/254 (94%)	-0.00	16 (6%) 17 13	35, 56, 120, 158	0
3	Q	240/254 (94%)	0.27	22 (9%) 9 6	39, 69, 150, 192	0
4	D	235/260 (90%)	-0.30	3 (1%) 77 73	37, 57, 90, 132	0
4	R	235/260 (90%)	-0.18	7 (2%) 50 43	41, 61, 96, 145	0
5	E	231/234 (98%)	-0.17	7 (3%) 50 43	42, 60, 94, 135	0
5	S	231/234 (98%)	-0.05	8 (3%) 44 36	39, 62, 100, 137	0
6	F	243/288 (84%)	-0.27	8 (3%) 46 39	34, 53, 103, 129	0
6	T	243/288 (84%)	-0.15	9 (3%) 41 34	29, 57, 104, 140	0
7	G	241/252 (95%)	-0.23	9 (3%) 41 34	34, 50, 90, 147	0
7	U	241/252 (95%)	-0.36	5 (2%) 63 58	35, 48, 82, 125	0
8	H	222/232 (95%)	-0.24	3 (1%) 75 71	36, 48, 76, 112	0
8	V	222/232 (95%)	-0.06	6 (2%) 54 48	36, 50, 82, 120	0
9	I	204/205 (99%)	-0.52	3 (1%) 73 70	29, 42, 71, 93	0
9	W	204/205 (99%)	-0.50	3 (1%) 73 70	30, 45, 73, 98	0
10	J	195/198 (98%)	-0.49	3 (1%) 73 70	31, 43, 68, 115	0
10	X	195/198 (98%)	-0.47	3 (1%) 73 70	33, 46, 72, 123	0
11	K	212/212 (100%)	-0.44	2 (0%) 84 82	33, 46, 71, 90	0
11	Y	212/212 (100%)	-0.53	2 (0%) 84 82	33, 44, 71, 96	0
12	L	222/222 (100%)	-0.45	1 (0%) 91 89	32, 47, 70, 97	0
12	Z	222/222 (100%)	-0.42	2 (0%) 84 82	30, 46, 73, 100	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
13	M	233/246 (94%)	-0.33	8 (3%) 45 38	28, 46, 70, 101	0
13	a	233/246 (94%)	-0.43	3 (1%) 77 73	30, 44, 67, 85	0
14	N	196/196 (100%)	-0.58	2 (1%) 82 80	28, 40, 68, 95	0
14	b	196/196 (100%)	-0.48	2 (1%) 82 80	29, 41, 69, 92	0
15	c	2/5 (40%)	1.32	0 100 100	81, 81, 81, 82	0
15	d	2/5 (40%)	-0.34	0 100 100	67, 67, 67, 72	0
15	e	2/5 (40%)	-0.69	0 100 100	51, 51, 51, 55	0
15	f	2/5 (40%)	3.75	2 (100%) 0 0	81, 81, 81, 88	0
15	g	2/5 (40%)	-0.48	0 100 100	62, 62, 62, 67	0
15	h	2/5 (40%)	-0.70	0 100 100	54, 54, 54, 60	0
All	All	6348/6644 (95%)	-0.27	180 (2%) 53 46	28, 50, 92, 192	0

The worst 5 of 180 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	7.4
2	B	220	ASN	7.2
2	P	219	ALA	6.9
2	P	222	GLY	6.4
3	Q	206	LYS	6.4

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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, 95th 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 ²)	Q<0.9
15	DCL	c	4	8/8	0.76	0.37	81,82,83,86	0
15	DCL	f	4	8/8	0.82	0.35	84,88,89,90	0
15	DCL	h	4	8/8	0.93	0.23	52,62,65,66	0
15	DCL	g	4	8/8	0.94	0.19	58,68,74,77	0
15	DCL	e	4	8/8	0.94	0.20	50,58,60,61	0
15	DCL	d	4	8/8	0.95	0.23	59,72,74,75	0

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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, 95th 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 ²)	Q<0.9
18	MES	Y	301	12/12	0.87	0.27	63,81,84,86	0
18	MES	K	302	12/12	0.88	0.28	66,83,86,87	0
17	CL	b	201	1/1	0.94	0.11	69,69,69,69	0
16	MG	Z	301	1/1	0.95	0.18	52,52,52,52	0
16	MG	I	302	1/1	0.95	0.10	44,44,44,44	0
17	CL	N	202	1/1	0.96	0.07	47,47,47,47	0
16	MG	G	301	1/1	0.97	0.07	55,55,55,55	0
16	MG	K	301	1/1	0.97	0.08	53,53,53,53	0
16	MG	I	301	1/1	0.98	0.21	54,54,54,54	0
16	MG	L	301	1/1	0.98	0.06	56,56,56,56	0
16	MG	N	201	1/1	0.98	0.06	41,41,41,41	0
17	CL	U	301	1/1	0.99	0.17	30,30,30,30	0
17	CL	G	302	1/1	0.99	0.14	30,30,30,30	0

6.5 Other polymers ⓘ

There are no such residues in this entry.