



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 08:11 pm BST

PDB ID : 1IRU  
Title : Crystal Structure of the mammalian 20S proteasome at 2.75 Å resolution  
Authors : Unno, M.; Mizushima, T.; Morimoto, Y.; Tomisugi, Y.; Tanaka, K.; Yasuoka, N.; Tsukihara, T.  
Deposited on : 2001-10-24  
Resolution : 2.75 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
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

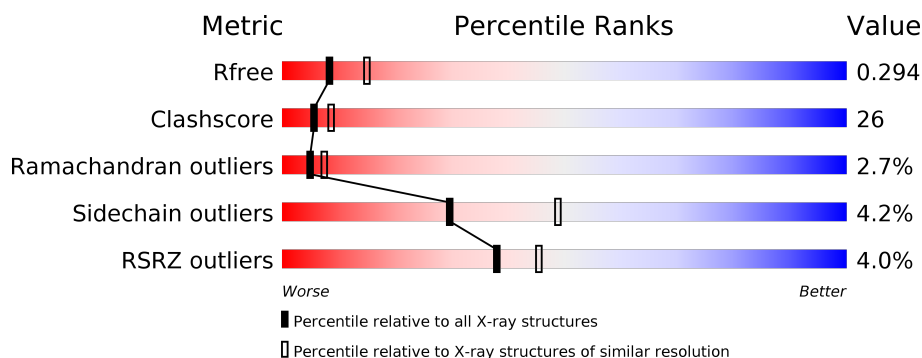
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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

Mol	Chain	Length	Quality of chain
1	A	246	<div> <div>10%</div> <div>52%</div> <div>45%</div> <div>..</div> </div>
1	O	246	<div> <div>51%</div> <div>46%</div> <div>..</div> </div>
2	B	233	<div> <div>13%</div> <div>51%</div> <div>44%</div> <div>6%</div> </div>
2	P	233	<div> <div>3%</div> <div>53%</div> <div>43%</div> <div>.</div> </div>
3	C	261	<div> <div>17%</div> <div>47%</div> <div>44%</div> <div>5%</div> <div>.</div> </div>
3	Q	261	<div> <div>5%</div> <div>47%</div> <div>43%</div> <div>5%</div> <div>.</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	D	248	
4	R	248	
5	E	241	
5	S	241	
6	F	263	
6	T	263	
7	G	254	
7	U	254	
8	H	205	
8	V	205	
9	I	234	
9	W	234	
10	J	205	
10	X	205	
11	K	201	
11	Y	201	
12	L	204	
12	Z	204	
13	1	213	
13	M	213	
14	2	219	
14	N	219	

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
15	MG	G	319	-	-	-	X

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 47757 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 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	244	Total	C	N	O	S	0	0	0
			1842	1170	309	350	13			
1	O	244	Total	C	N	O	S	0	0	0
			1842	1170	309	350	13			

- Molecule 2 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	233	Total	C	N	O	S	0	0	0
			1707	1081	287	334	5			
2	P	233	Total	C	N	O	S	0	0	0
			1707	1081	287	334	5			

- Molecule 3 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	250	Total	C	N	O	S	0	0	0
			1902	1195	329	370	8			
3	Q	250	Total	C	N	O	S	0	0	0
			1902	1195	329	370	8			

- Molecule 4 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	243	Total	C	N	O	S	0	0	0
			1665	1032	307	322	4			
4	R	243	Total	C	N	O	S	0	0	0
			1665	1032	307	322	4			

- Molecule 5 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	234	Total	C	N	O	S	0	0	0
			1763	1104	290	358	11			
5	S	234	Total	C	N	O	S	0	0	0
			1763	1104	290	358	11			

- Molecule 6 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	238	Total	C	N	O	S	0	0	0
			1850	1159	334	346	11			
6	T	238	Total	C	N	O	S	0	0	0
			1850	1159	334	346	11			

- Molecule 7 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	245	Total	C	N	O	S	0	0	0
			1885	1195	319	360	11			
7	U	245	Total	C	N	O	S	0	0	0
			1885	1195	319	360	11			

- Molecule 8 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	202	Total	C	N	O	S	0	0	0
			1509	945	258	294	12			
8	V	202	Total	C	N	O	S	0	0	0
			1509	945	258	294	12			

- Molecule 9 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	220	Total	C	N	O	S	0	0	0
			1645	1034	282	317	12			
9	W	220	Total	C	N	O	S	0	0	0
			1645	1034	282	317	12			

- Molecule 10 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	204	Total	C	N	O	S	0	0	0
			1585	1011	262	294	18			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	204	Total	C	N	O	S	0	0	0
			1585	1011	262	294	18			

- Molecule 11 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	199	Total	C	N	O	S	0	0	0
			1570	1006	265	290	9			
11	Y	199	Total	C	N	O	S	0	0	0
			1570	1006	265	290	9			

- Molecule 12 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	201	Total	C	N	O	S	0	0	0
			1548	974	273	292	9			
12	Z	201	Total	C	N	O	S	0	0	0
			1548	974	273	292	9			

- Molecule 13 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	213	Total	C	N	O	S	0	0	0
			1639	1034	282	313	10			
13	1	213	Total	C	N	O	S	0	0	0
			1639	1034	282	313	10			

- Molecule 14 is a protein called 20S proteasome.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	217	Total	C	N	O	S	0	0	0
			1671	1053	287	319	12			
14	2	217	Total	C	N	O	S	0	0	0
			1671	1053	287	319	12			

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	P	1	Total	Mg	0	0
			1	1		
15	G	3	Total	Mg	0	0
			3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	J	4	Total 4	Mg 4	0	0
15	Q	1	Total 1	Mg 1	0	0
15	D	1	Total 1	Mg 1	0	0
15	K	1	Total 1	Mg 1	0	0
15	B	1	Total 1	Mg 1	0	0
15	I	1	Total 1	Mg 1	0	0
15	C	1	Total 1	Mg 1	0	0
15	1	2	Total 2	Mg 2	0	0
15	W	1	Total 1	Mg 1	0	0
15	A	1	Total 1	Mg 1	0	0
15	U	3	Total 3	Mg 3	0	0
15	X	4	Total 4	Mg 4	0	0
15	O	1	Total 1	Mg 1	0	0
15	R	1	Total 1	Mg 1	0	0
15	Y	1	Total 1	Mg 1	0	0
15	M	2	Total 2	Mg 2	0	0

- Molecule 16 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	2	Total 2	O 2	0	0
16	B	3	Total 3	O 3	0	0
16	C	3	Total 3	O 3	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	F	6	Total O 6 6	0	0
16	G	4	Total O 4 4	0	0
16	H	7	Total O 7 7	0	0
16	I	6	Total O 6 6	0	0
16	J	1	Total O 1 1	0	0
16	K	6	Total O 6 6	0	0
16	L	8	Total O 8 8	0	0
16	M	10	Total O 10 10	0	0
16	N	9	Total O 9 9	0	0
16	O	7	Total O 7 7	0	0
16	P	6	Total O 6 6	0	0
16	Q	3	Total O 3 3	0	0
16	R	3	Total O 3 3	0	0
16	S	1	Total O 1 1	0	0
16	T	7	Total O 7 7	0	0
16	U	9	Total O 9 9	0	0
16	V	15	Total O 15 15	0	0
16	W	9	Total O 9 9	0	0
16	X	17	Total O 17 17	0	0
16	Y	3	Total O 3 3	0	0
16	Z	8	Total O 8 8	0	0

*Continued on next page...*

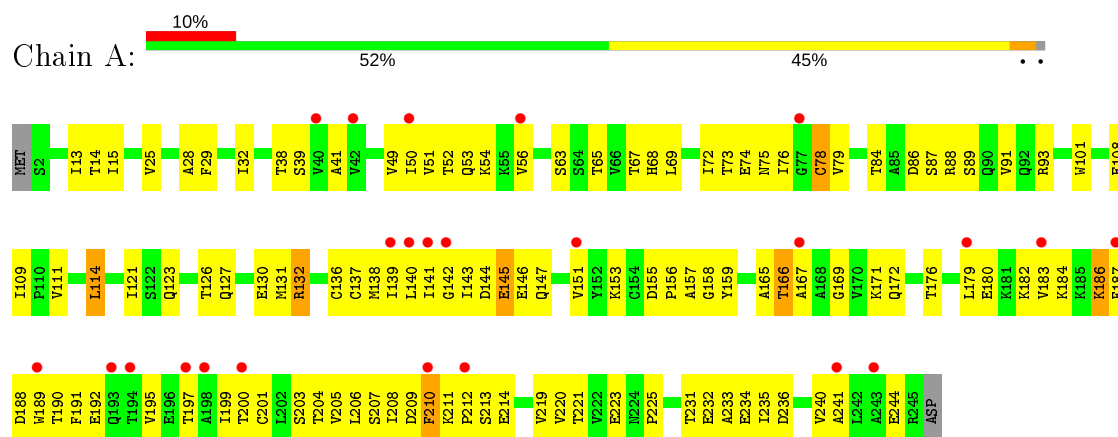
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	1	4	Total	O	0	0
			4	4		
16	2	8	Total	O	0	0
			8	8		

### 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.

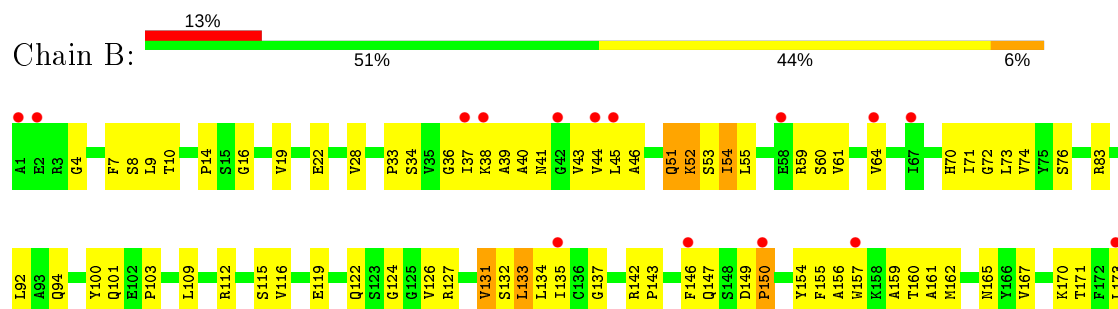
- Molecule 1: 20S proteasome



- Molecule 1: 20S proteasome

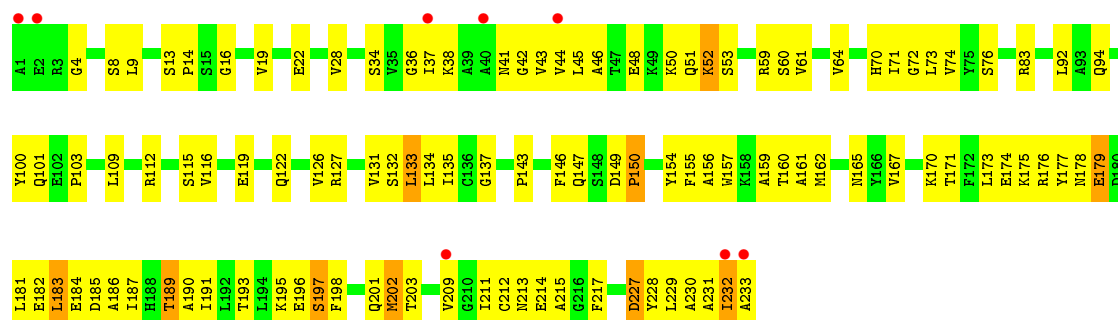


- Molecule 2: 20S proteasome

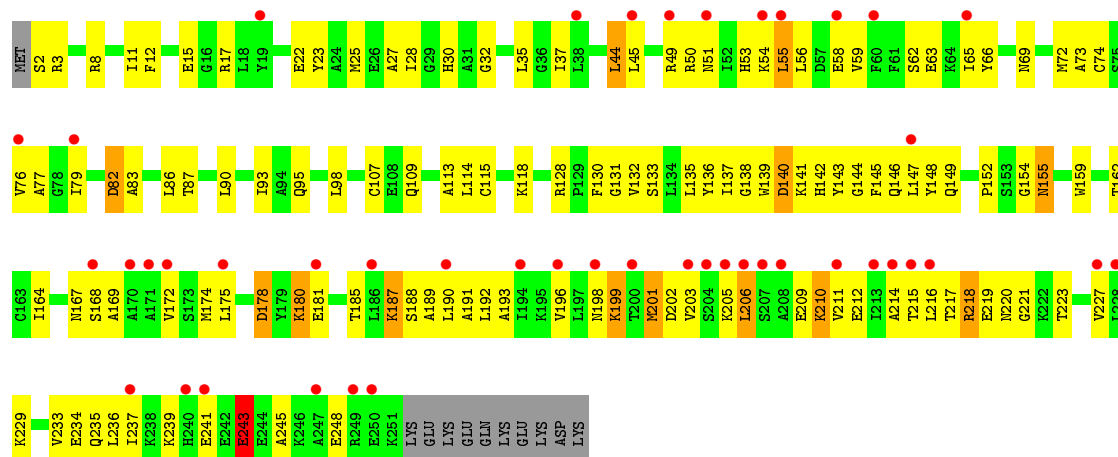




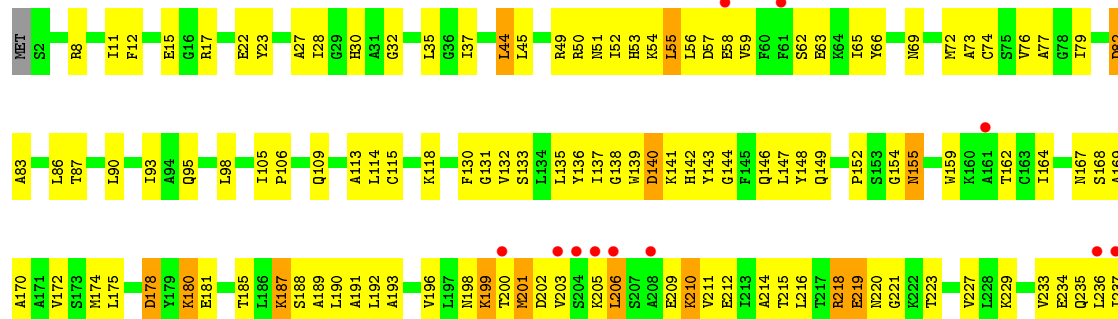
• Molecule 2: 20S proteasome

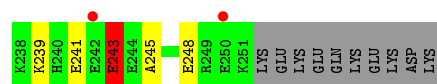


• Molecule 3: 20S proteasome

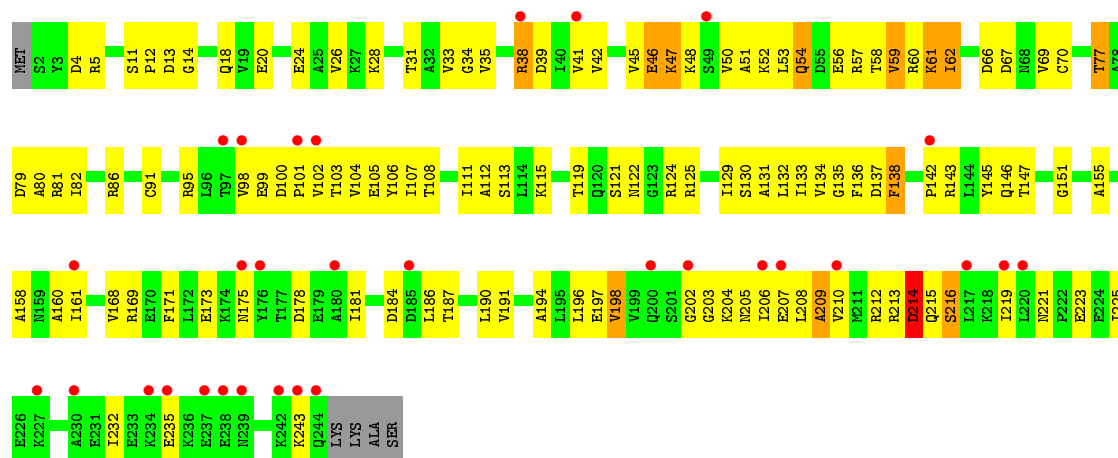


• Molecule 3: 20S proteasome

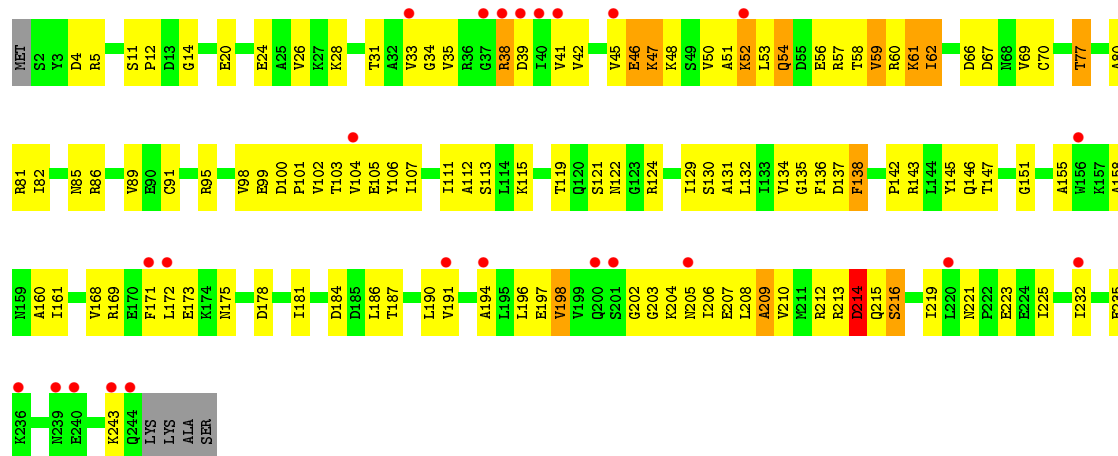




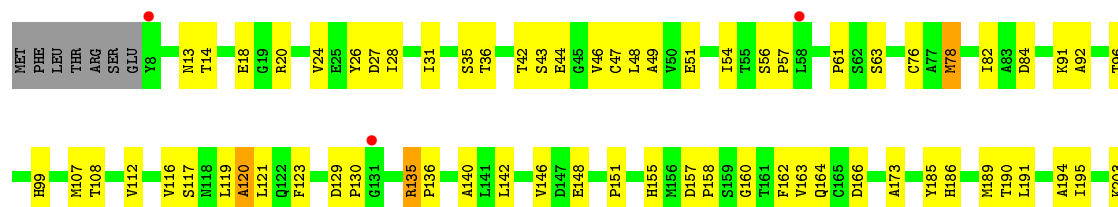
• Molecule 4: 20S proteasome



• Molecule 4: 20S proteasome

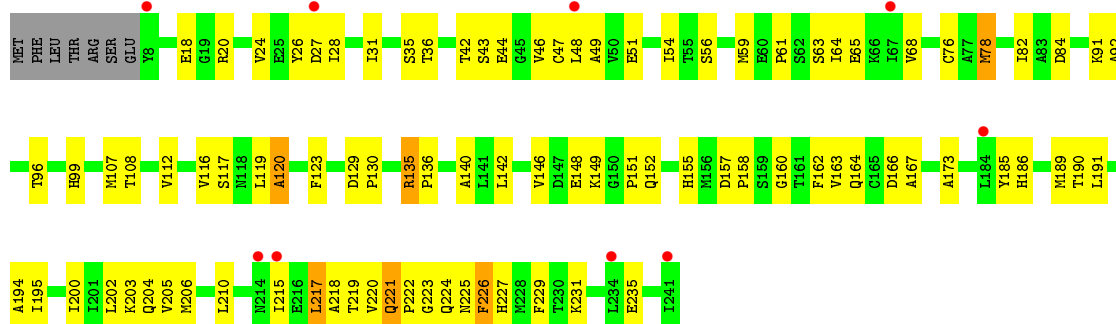


• Molecule 5: 20S proteasome

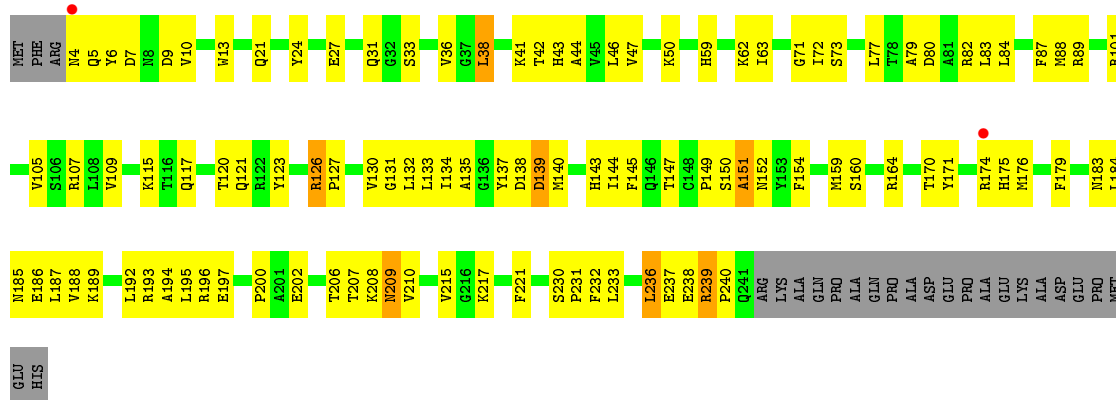




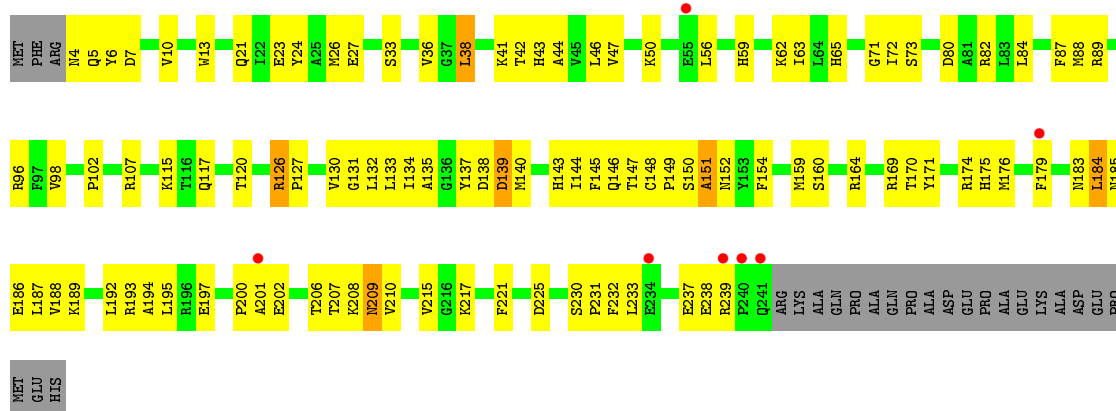
• Molecule 5: 20S proteasome



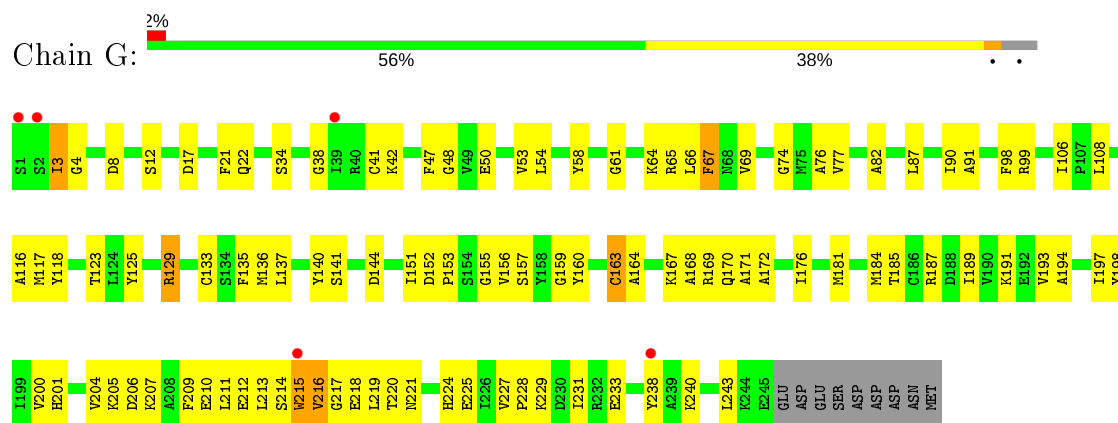
• Molecule 6: 20S proteasome



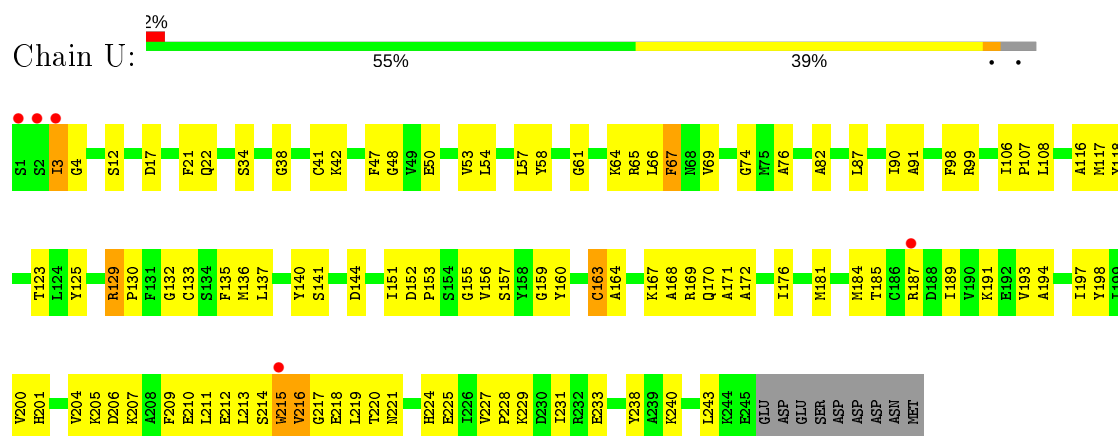
• Molecule 6: 20S proteasome



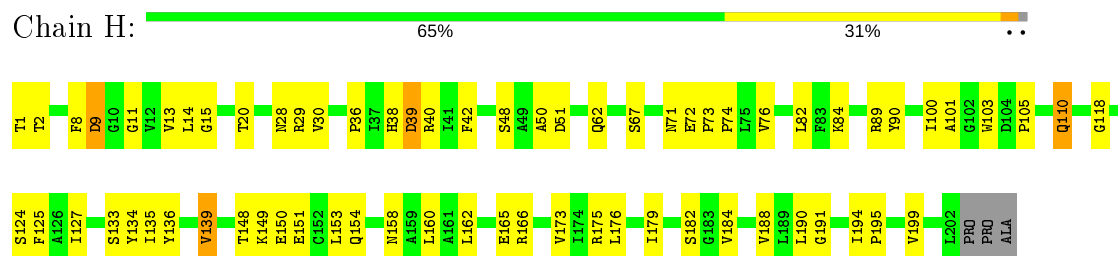
- Molecule 7: 20S proteasome



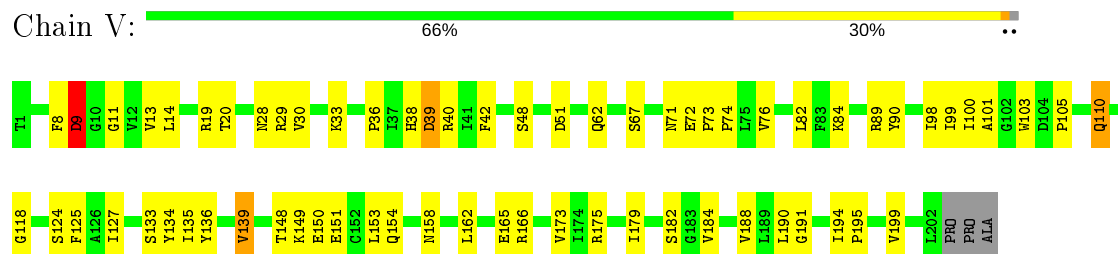
- Molecule 7: 20S proteasome



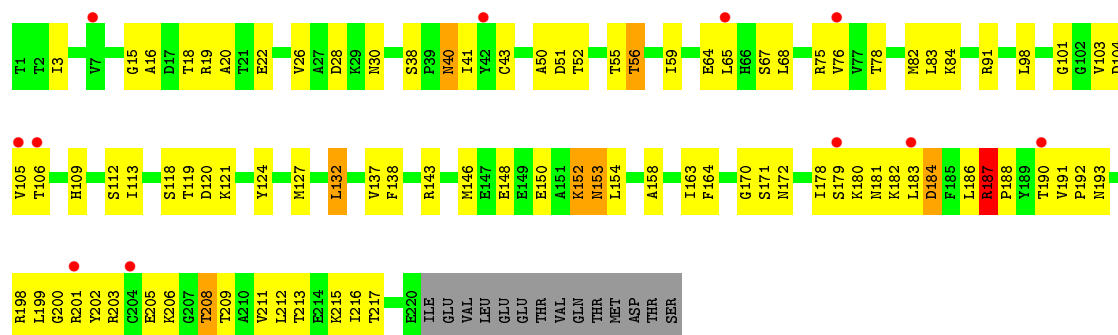
- Molecule 8: 20S proteasome



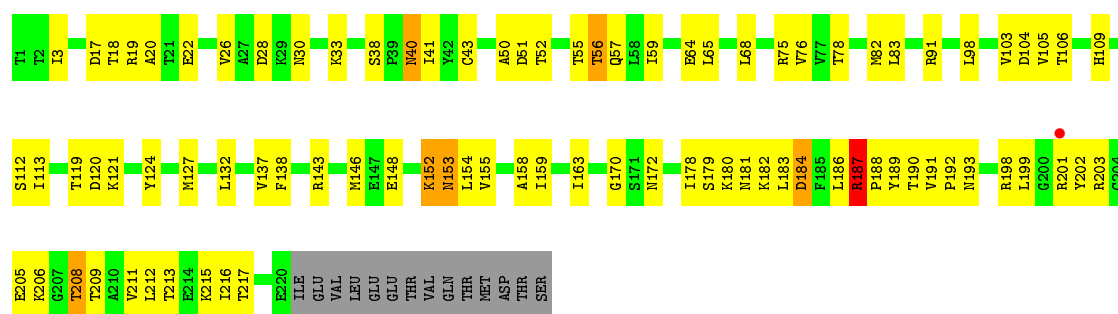
- Molecule 8: 20S proteasome



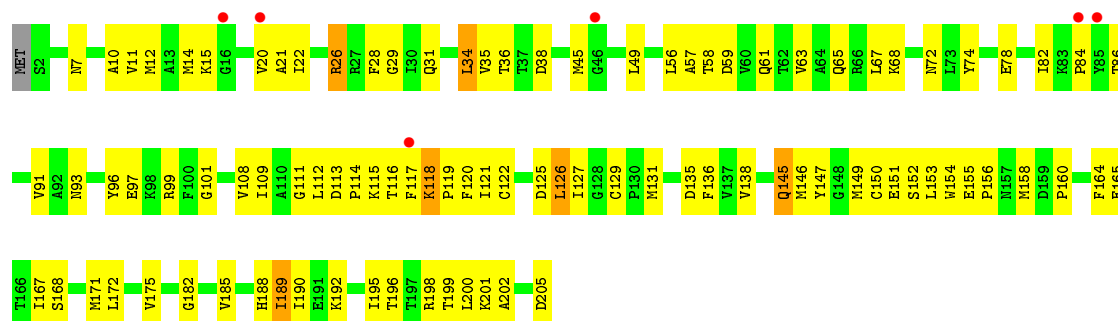
- Molecule 9: 20S proteasome



• Molecule 9: 20S proteasome



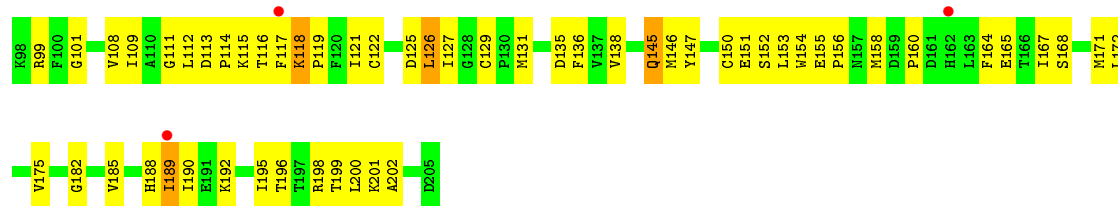
• Molecule 10: 20S proteasome



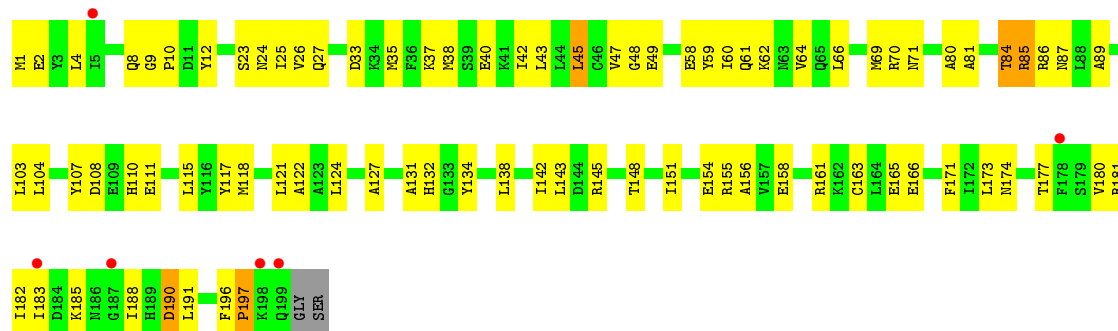
• Molecule 10: 20S proteasome







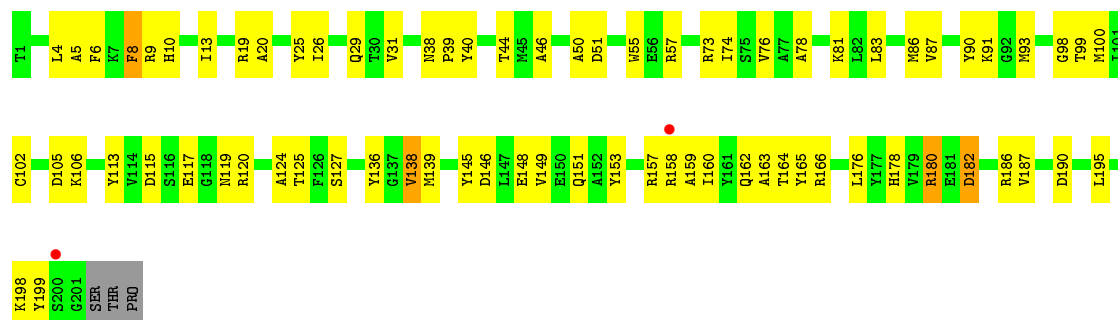
• Molecule 11: 20S proteasome



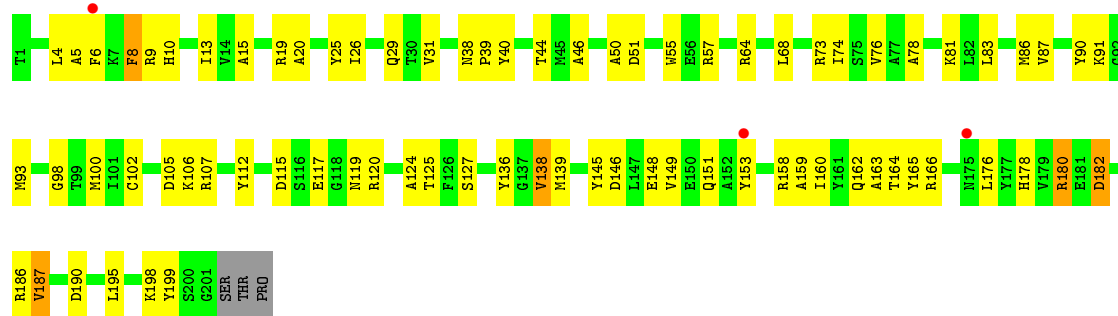
• Molecule 11: 20S proteasome



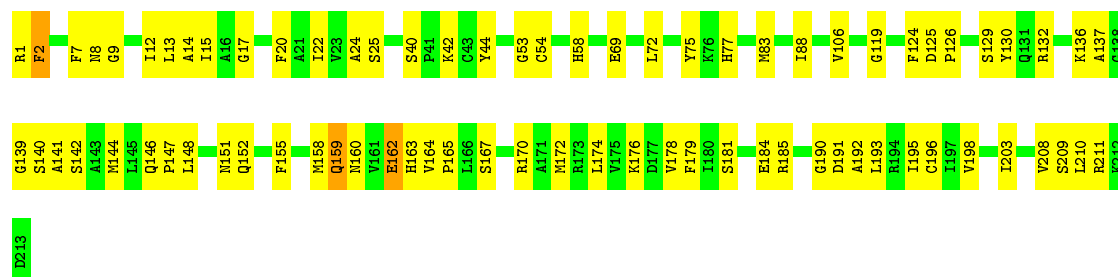
• Molecule 12: 20S proteasome



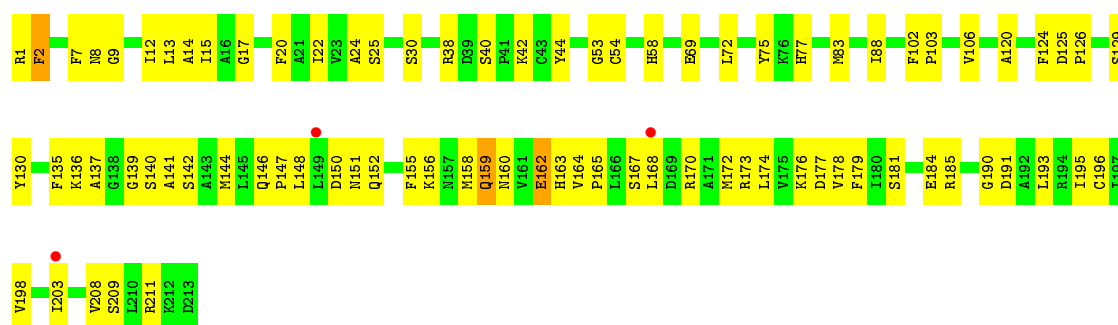
• Molecule 12: 20S proteasome



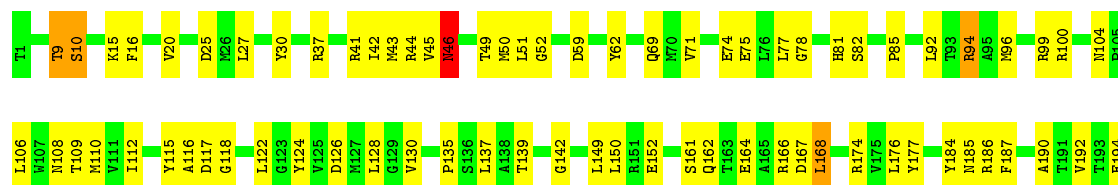
• Molecule 13: 20S proteasome

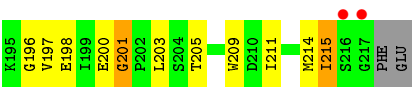


• Molecule 13: 20S proteasome

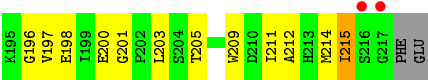
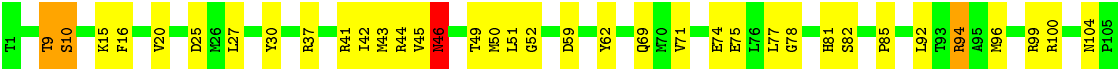


• Molecule 14: 20S proteasome





● Molecule 14: 20S proteasome



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	316.70Å 205.90Å 116.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	65.00 – 2.75 67.08 – 2.75	Depositor EDS
% Data completeness (in resolution range)	(Not available) (65.00-2.75) 96.1 (67.08-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.09 (at 2.73Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.250 , 0.294 0.249 , 0.294	Depositor DCC
$R_{free}$ test set	9420 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.6	Xtriage
Anisotropy	0.397	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 56.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	47757	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.96% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.36	0/1875	0.66	0/2545
1	O	0.43	0/1875	0.67	0/2545
2	B	0.33	0/1742	0.63	0/2372
2	P	0.38	0/1742	0.64	0/2372
3	C	0.34	0/1931	0.61	0/2613
3	Q	0.39	0/1931	0.61	0/2613
4	D	0.33	0/1688	0.60	0/2310
4	R	0.35	0/1688	0.62	0/2310
5	E	0.42	0/1790	0.64	0/2424
5	S	0.40	0/1790	0.63	0/2424
6	F	0.40	0/1885	0.71	0/2552
6	T	0.40	0/1885	0.71	1/2552 (0.0%)
7	G	0.41	0/1920	0.64	0/2591
7	U	0.44	0/1920	0.65	0/2591
8	H	0.43	0/1535	0.70	0/2078
8	V	0.46	0/1535	0.70	0/2078
9	I	0.39	0/1672	0.70	2/2267 (0.1%)
9	W	0.43	0/1672	0.71	2/2267 (0.1%)
10	J	0.40	0/1614	0.70	0/2178
10	X	0.46	0/1614	0.72	0/2178
11	K	0.39	0/1603	0.68	0/2174
11	Y	0.42	0/1603	0.69	0/2174
12	L	0.43	0/1579	0.68	0/2134
12	Z	0.42	0/1579	0.67	0/2134
13	1	0.41	0/1669	0.66	0/2250
13	M	0.45	0/1669	0.67	0/2250
14	2	0.43	0/1704	0.73	2/2311 (0.1%)
14	N	0.45	0/1704	0.73	2/2311 (0.1%)
All	All	0.41	0/48414	0.67	9/65598 (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
14	2	0	1
14	N	0	1
All	All	0	2

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	I	187	ARG	C-N-CD	5.89	140.76	128.40
9	W	187	ARG	C-N-CD	5.80	140.59	128.40
14	N	109	THR	N-CA-C	-5.38	96.49	111.00
14	N	106	LEU	N-CA-C	-5.24	96.84	111.00
14	2	109	THR	N-CA-C	-5.19	96.99	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
14	2	30	TYR	Sidechain
14	N	30	TYR	Sidechain

## 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	1842	0	1803	101	0
1	O	1842	0	1803	107	0
2	B	1707	0	1591	96	0
2	P	1707	0	1591	101	0
3	C	1902	0	1835	134	0
3	Q	1902	0	1835	136	0
4	D	1665	0	1433	100	0
4	R	1665	0	1433	100	0
5	E	1763	0	1708	83	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	S	1763	0	1708	82	0
6	F	1850	0	1822	102	0
6	T	1850	0	1822	95	0
7	G	1885	0	1843	88	0
7	U	1885	0	1843	84	0
8	H	1509	0	1473	58	0
8	V	1509	0	1473	55	0
9	I	1645	0	1649	105	0
9	W	1645	0	1649	94	0
10	J	1585	0	1600	109	0
10	X	1585	0	1600	108	0
11	K	1570	0	1546	103	0
11	Y	1570	0	1546	102	0
12	L	1548	0	1499	78	0
12	Z	1548	0	1499	82	0
13	1	1639	0	1609	81	0
13	M	1639	0	1609	69	0
14	2	1671	0	1625	77	0
14	N	1671	0	1625	80	0
15	1	2	0	0	0	0
15	A	1	0	0	0	0
15	B	1	0	0	0	0
15	C	1	0	0	0	0
15	D	1	0	0	0	0
15	G	3	0	0	0	0
15	I	1	0	0	0	0
15	J	4	0	0	0	0
15	K	1	0	0	0	0
15	M	2	0	0	0	0
15	O	1	0	0	0	0
15	P	1	0	0	0	0
15	Q	1	0	0	0	0
15	R	1	0	0	0	0
15	U	3	0	0	0	0
15	W	1	0	0	0	0
15	X	4	0	0	0	0
15	Y	1	0	0	0	0
16	1	4	0	0	1	0
16	2	8	0	0	1	0
16	A	2	0	0	0	0
16	B	3	0	0	0	0
16	C	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	F	6	0	0	1	0
16	G	4	0	0	1	0
16	H	7	0	0	0	0
16	I	6	0	0	2	0
16	J	1	0	0	0	0
16	K	6	0	0	0	0
16	L	8	0	0	0	0
16	M	10	0	0	0	0
16	N	9	0	0	4	0
16	O	7	0	0	1	0
16	P	6	0	0	1	0
16	Q	3	0	0	1	0
16	R	3	0	0	0	0
16	S	1	0	0	0	0
16	T	7	0	0	1	0
16	U	9	0	0	0	0
16	V	15	0	0	0	0
16	W	9	0	0	0	0
16	X	17	0	0	0	0
16	Y	3	0	0	1	0
16	Z	8	0	0	0	0
All	All	47757	0	46072	2398	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 26.

The worst 5 of 2398 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)
11:K:145:ARG:HE	12:Z:158:ARG:HD2	1.01	1.14
3:Q:180:LYS:H	3:Q:180:LYS:HD3	1.10	1.11
12:L:158:ARG:HD2	11:Y:145:ARG:HE	1.10	1.08
3:C:180:LYS:H	3:C:180:LYS:HD3	1.10	1.07
4:D:57:ARG:HA	4:D:60:ARG:HE	1.18	1.06

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	242/246 (98%)	220 (91%)	19 (8%)	3 (1%)	13	23
1	O	242/246 (98%)	220 (91%)	19 (8%)	3 (1%)	13	23
2	B	231/233 (99%)	185 (80%)	30 (13%)	16 (7%)	1	1
2	P	231/233 (99%)	188 (81%)	28 (12%)	15 (6%)	1	1
3	C	248/261 (95%)	201 (81%)	34 (14%)	13 (5%)	2	2
3	Q	248/261 (95%)	202 (82%)	33 (13%)	13 (5%)	2	2
4	D	241/248 (97%)	192 (80%)	33 (14%)	16 (7%)	1	1
4	R	241/248 (97%)	191 (79%)	34 (14%)	16 (7%)	1	1
5	E	232/241 (96%)	205 (88%)	24 (10%)	3 (1%)	12	21
5	S	232/241 (96%)	205 (88%)	24 (10%)	3 (1%)	12	21
6	F	236/263 (90%)	214 (91%)	16 (7%)	6 (2%)	5	9
6	T	236/263 (90%)	214 (91%)	17 (7%)	5 (2%)	7	12
7	G	243/254 (96%)	223 (92%)	14 (6%)	6 (2%)	5	9
7	U	243/254 (96%)	221 (91%)	16 (7%)	6 (2%)	5	9
8	H	200/205 (98%)	188 (94%)	9 (4%)	3 (2%)	10	18
8	V	200/205 (98%)	188 (94%)	9 (4%)	3 (2%)	10	18
9	I	218/234 (93%)	193 (88%)	21 (10%)	4 (2%)	8	15
9	W	218/234 (93%)	193 (88%)	21 (10%)	4 (2%)	8	15
10	J	202/205 (98%)	180 (89%)	18 (9%)	4 (2%)	7	13
10	X	202/205 (98%)	180 (89%)	18 (9%)	4 (2%)	7	13
11	K	197/201 (98%)	174 (88%)	20 (10%)	3 (2%)	10	18
11	Y	197/201 (98%)	174 (88%)	20 (10%)	3 (2%)	10	18
12	L	199/204 (98%)	180 (90%)	17 (8%)	2 (1%)	15	27
12	Z	199/204 (98%)	181 (91%)	16 (8%)	2 (1%)	15	27
13	1	211/213 (99%)	191 (90%)	19 (9%)	1 (0%)	29	47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	M	211/213 (99%)	192 (91%)	18 (8%)	1 (0%)	29	47
14	2	215/219 (98%)	195 (91%)	16 (7%)	4 (2%)	8	14
14	N	215/219 (98%)	196 (91%)	15 (7%)	4 (2%)	8	14
All	All	6230/6454 (96%)	5486 (88%)	578 (9%)	166 (3%)	5	7

5 of 166 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	40	ALA
2	B	52	LYS
2	B	54	ILE
2	B	60	SER
2	B	179	GLU

### 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	192/210 (91%)	184 (96%)	8 (4%)	30	49
1	O	192/210 (91%)	184 (96%)	8 (4%)	30	49
2	B	163/190 (86%)	153 (94%)	10 (6%)	18	33
2	P	163/190 (86%)	153 (94%)	10 (6%)	18	33
3	C	191/221 (86%)	180 (94%)	11 (6%)	20	35
3	Q	191/221 (86%)	179 (94%)	12 (6%)	18	31
4	D	136/211 (64%)	128 (94%)	8 (6%)	19	34
4	R	136/211 (64%)	128 (94%)	8 (6%)	19	34
5	E	190/204 (93%)	184 (97%)	6 (3%)	39	59
5	S	190/204 (93%)	184 (97%)	6 (3%)	39	59
6	F	198/224 (88%)	193 (98%)	5 (2%)	47	67
6	T	198/224 (88%)	193 (98%)	5 (2%)	47	67
7	G	195/211 (92%)	186 (95%)	9 (5%)	27	46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	U	195/211 (92%)	186 (95%)	9 (5%)	27	46
8	H	155/159 (98%)	150 (97%)	5 (3%)	39	59
8	V	155/159 (98%)	150 (97%)	5 (3%)	39	59
9	I	177/195 (91%)	171 (97%)	6 (3%)	37	58
9	W	177/195 (91%)	171 (97%)	6 (3%)	37	58
10	J	172/174 (99%)	165 (96%)	7 (4%)	30	50
10	X	172/174 (99%)	164 (95%)	8 (5%)	26	45
11	K	164/171 (96%)	157 (96%)	7 (4%)	29	48
11	Y	164/171 (96%)	157 (96%)	7 (4%)	29	48
12	L	153/159 (96%)	146 (95%)	7 (5%)	27	46
12	Z	153/159 (96%)	146 (95%)	7 (5%)	27	46
13	1	173/178 (97%)	167 (96%)	6 (4%)	36	56
13	M	173/178 (97%)	167 (96%)	6 (4%)	36	56
14	2	174/181 (96%)	169 (97%)	5 (3%)	42	62
14	N	174/181 (96%)	168 (97%)	6 (3%)	37	58
All	All	4866/5376 (90%)	4663 (96%)	203 (4%)	30	49

5 of 203 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	162	GLU
2	P	133	LEU
12	Z	102	CYS
13	M	170	ARG
1	O	114	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 163 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	151	ASN
2	P	168	ASN
13	1	77	HIS
13	M	159	GLN
1	O	90	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](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 30 ligands modelled in this entry, 30 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	244/246 (99%)	0.73	24 (9%) 7 8	56, 89, 120, 135	0
1	O	244/246 (99%)	0.00	1 (0%) 92 95	38, 56, 81, 100	0
2	B	233/233 (100%)	0.82	31 (13%) 3 4	62, 97, 133, 142	0
2	P	233/233 (100%)	0.30	8 (3%) 45 53	36, 67, 105, 129	0
3	C	250/261 (95%)	1.03	44 (17%) 1 1	59, 103, 142, 148	0
3	Q	250/261 (95%)	0.44	13 (5%) 27 33	37, 70, 123, 141	0
4	D	243/248 (97%)	0.73	31 (12%) 3 4	54, 93, 146, 163	0
4	R	243/248 (97%)	0.47	24 (9%) 7 8	40, 82, 134, 157	0
5	E	234/241 (97%)	0.27	3 (1%) 77 84	39, 64, 94, 110	0
5	S	234/241 (97%)	0.41	9 (3%) 40 48	39, 74, 102, 132	0
6	F	238/263 (90%)	0.19	2 (0%) 86 90	42, 63, 91, 124	0
6	T	238/263 (90%)	0.23	7 (2%) 51 61	37, 58, 95, 135	0
7	G	245/254 (96%)	0.37	5 (2%) 65 73	49, 74, 104, 118	0
7	U	245/254 (96%)	0.17	5 (2%) 65 73	35, 53, 89, 126	0
8	H	202/205 (98%)	0.11	0 100 100	35, 60, 78, 94	0
8	V	202/205 (98%)	0.02	0 100 100	33, 49, 70, 102	0
9	I	220/234 (94%)	0.46	11 (5%) 28 35	49, 71, 100, 115	0
9	W	220/234 (94%)	0.07	1 (0%) 91 94	30, 53, 81, 99	0
10	J	204/205 (99%)	0.56	6 (2%) 51 61	49, 79, 101, 111	0
10	X	204/205 (99%)	0.20	4 (1%) 65 73	36, 52, 78, 97	0
11	K	199/201 (99%)	0.39	6 (3%) 50 59	49, 69, 93, 110	0
11	Y	199/201 (99%)	0.16	3 (1%) 73 81	36, 56, 80, 113	0
12	L	201/204 (98%)	0.10	2 (0%) 82 87	35, 53, 75, 108	0
12	Z	201/204 (98%)	0.25	3 (1%) 73 81	45, 65, 85, 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	1	213/213 (100%)	0.23	3 (1%) 75 82	40, 70, 97, 112	0
13	M	213/213 (100%)	0.04	0 100 100	31, 50, 73, 97	0
14	2	217/219 (99%)	0.12	2 (0%) 84 89	33, 55, 81, 109	0
14	N	217/219 (99%)	0.05	2 (0%) 84 89	33, 51, 70, 109	0
All	All	6286/6454 (97%)	0.33	250 (3%) 38 45	30, 66, 115, 163	0

The worst 5 of 250 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	232	ILE	12.8
2	P	233	ALA	10.0
11	K	199	GLN	8.5
4	R	244	GLN	7.9
7	U	1	SER	7.8

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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
15	MG	G	319	1/1	0.59	0.43	66,66,66,66	0
15	MG	G	311	1/1	0.66	0.25	97,97,97,97	0
15	MG	1	418	1/1	0.69	0.25	90,90,90,90	0
15	MG	O	401	1/1	0.76	0.22	57,57,57,57	0
15	MG	G	317	1/1	0.81	0.18	85,85,85,85	0
15	MG	J	407	1/1	0.82	0.12	70,70,70,70	0
15	MG	J	320	1/1	0.84	0.45	112,112,112,112	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
15	MG	K	313	1/1	0.84	0.24	77,77,77,77	0
15	MG	D	309	1/1	0.84	0.13	87,87,87,87	0
15	MG	Q	414	1/1	0.86	0.20	90,90,90,90	0
15	MG	A	301	1/1	0.86	0.23	83,83,83,83	0
15	MG	B	316	1/1	0.87	0.34	84,84,84,84	0
15	MG	W	406	1/1	0.87	0.16	59,59,59,59	0
15	MG	Y	413	1/1	0.88	0.31	51,51,51,51	0
15	MG	X	404	1/1	0.89	0.37	70,70,70,70	0
15	MG	U	419	1/1	0.90	0.69	70,70,70,70	0
15	MG	U	417	1/1	0.90	0.45	64,64,64,64	0
15	MG	X	420	1/1	0.90	0.41	77,77,77,77	0
15	MG	P	416	1/1	0.90	0.20	72,72,72,72	0
15	MG	X	307	1/1	0.91	0.14	39,39,39,39	0
15	MG	M	318	1/1	0.92	0.30	81,81,81,81	0
15	MG	I	306	1/1	0.92	0.23	87,87,87,87	0
15	MG	J	304	1/1	0.93	0.10	56,56,56,56	0
15	MG	U	411	1/1	0.93	0.22	76,76,76,76	0
15	MG	X	415	1/1	0.93	0.21	59,59,59,59	0
15	MG	J	315	1/1	0.93	0.41	69,69,69,69	0
15	MG	C	314	1/1	0.94	0.35	76,76,76,76	0
15	MG	1	410	1/1	0.96	0.21	66,66,66,66	0
15	MG	M	310	1/1	0.96	0.09	67,67,67,67	0
15	MG	R	409	1/1	0.96	0.07	59,59,59,59	0

## 6.5 Other polymers

There are no such residues in this entry.