wwPDB X-ray Structure Validation Summary Report (i)

Aug 29, 2023 – 03:24 PM EDT

PDB ID	:	3MG7
Title	:	Structure of yeast 20S open-gate proteasome with Compound 8
Authors	:	Sintchak, M.D.
Deposited on	:	2010-04-05
Resolution	:	2.78 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 (i) 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 (i)) were used in the production of this report:

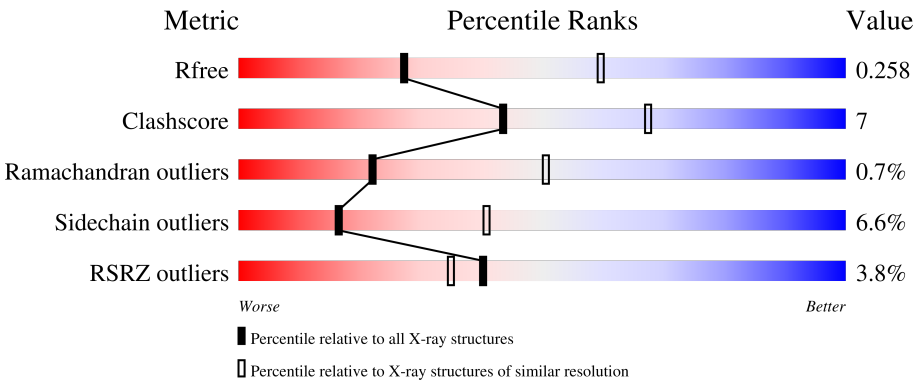
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.78 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	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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 >=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 <=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	3% 86% 12% •
1	O	250	6% 80% 18% •
2	B	245	3% 75% 18% • •
2	P	245	3% 76% 18% • •
3	C	243	11% 83% 15% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	243	20% 77% 20%
4	D	250	7% 80% 15%
4	R	250	6% 78% 17%
5	E	234	8% 76% 22%
5	S	234	10% 77% 20%
6	F	248	2% 78% 16%
6	T	248	2% 77% 17%
7	G	252	2% 79% 16%
7	U	252	% 82% 12%
8	H	222	80% 18%
8	V	222	78% 21%
9	I	205	% 82% 16%
9	W	205	% 82% 16%
10	J	198	4% 80% 17%
10	X	198	3% 85% 14%
11	K	212	2% 82% 16%
11	Y	212	83% 14%
12	L	241	78% 12% 8%
12	Z	241	2% 75% 15% 8%
13	1	266	% 70% 15% 12%
13	M	266	76% 11% 12%
14	2	196	% 82% 17%
14	N	196	% 84% 16%

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	F	242	-	-	-	X
15	MG	F	243	-	-	-	X

2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 49424 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 component Y7.

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

- Molecule 2 is a protein called Proteasome component Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	235	1829	1158	303	365	3	0	0	0
2	P	235	1829	1158	303	365	3	0	0	0

- Molecule 3 is a protein called Proteasome component PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	241	1891	1181	331	375	4	0	0	0
3	Q	241	1891	1181	331	375	4	0	0	0

- Molecule 4 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	242	1862	1162	314	379	7	0	0	0
4	R	242	1862	1162	314	379	7	0	0	0

- Molecule 5 is a protein called Proteasome component PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	233	1795	1129	312	350	4	0	0	0
5	S	233	1795	1129	312	350	4	0	0	0

- Molecule 6 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	237	1848	1175	323	346	4	0	0	0
6	T	237	1848	1175	323	346	4	0	0	0

- Molecule 7 is a protein called Proteasome component C7-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	243	1921	1221	322	370	8	0	0	0
7	U	243	1921	1221	322	370	8	0	0	0

- Molecule 8 is a protein called Proteasome component PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	222	1685	1061	293	324	7	0	0	0
8	V	222	1685	1061	293	324	7	0	0	0

- Molecule 9 is a protein called Proteasome component PUP3.

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

- Molecule 10 is a protein called Proteasome component C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	198	1585	1005	269	305	6	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	X	198	1585	1005	269	305	6	0	0	0

- Molecule 11 is a protein called Proteasome component PRE2.

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

- Molecule 12 is a protein called Proteasome component C5.

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

- Molecule 13 is a protein called Proteasome component PRE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	233	1824	1154	312	351	7	0	0	0
13	1	233	1824	1154	312	351	7	0	0	0

- Molecule 14 is a protein called Proteasome component PRE3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
14	N	196	1512	955	250	300	7	0	0	0
14	2	196	1512	955	250	300	7	0	0	0

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

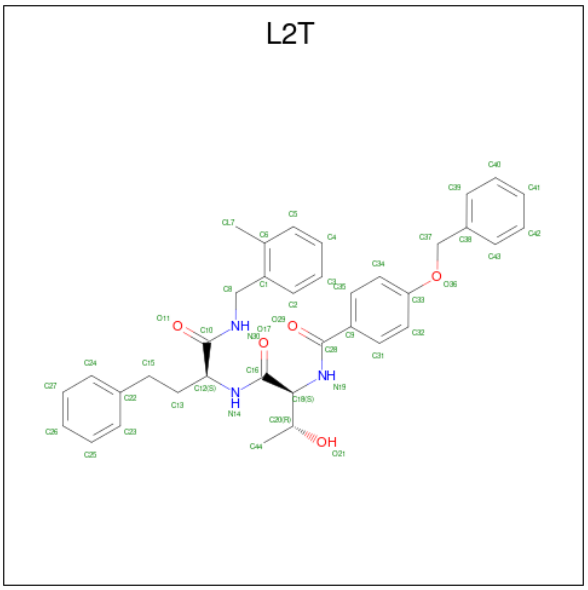
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
15	F	2	2	2	0	0
15	G	1	1	1	0	0

Continued on next page...

Continued from previous page...

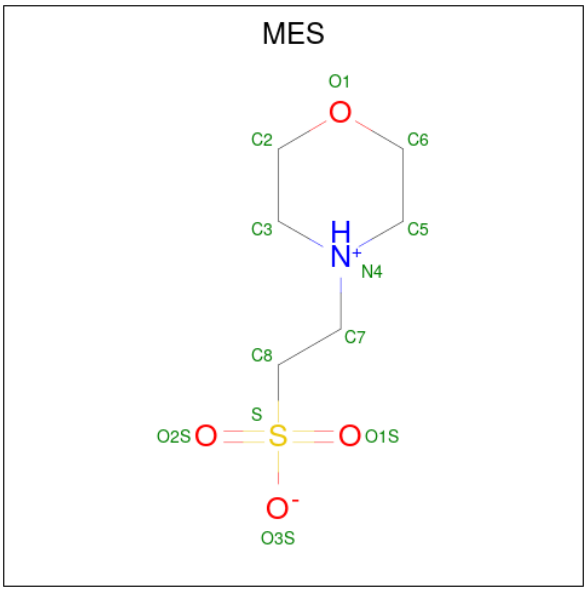
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
15	H	1	1	1	0	0
			Total	Mg		
15	I	2	2	2	0	0
			Total	Mg		
15	K	1	1	1	0	0
			Total	Mg		
15	L	2	2	2	0	0
			Total	Mg		
15	N	1	1	1	0	0

- Molecule 16 is 4-(benzyloxy)-N-[(1S,2R)-2-hydroxy-1-({(1S)-1-[(2-methylbenzyl)carbamoyl]-3-phenylpropyl}carbamoyl)propyl]benzamide (three-letter code: L2T) (formula: $C_{36}H_{39}N_3O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
16	K	1	44	36	3	5	0	0
			Total	C	N	O		
16	Y	1	44	36	3	5	0	0

- Molecule 17 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



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

- Molecule 18 is water.

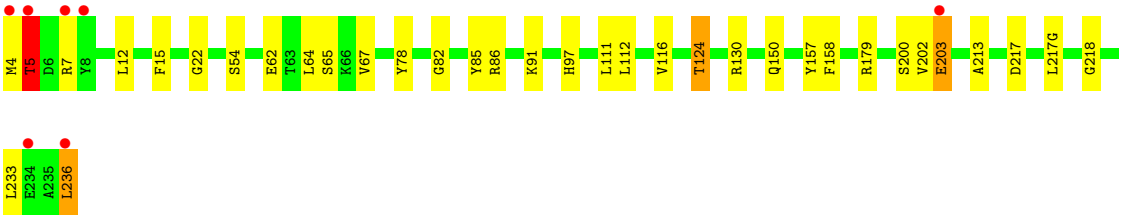
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
18	L	2	2	2	0	0
			Total	O		
18	Z	2	2	2	0	0

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.

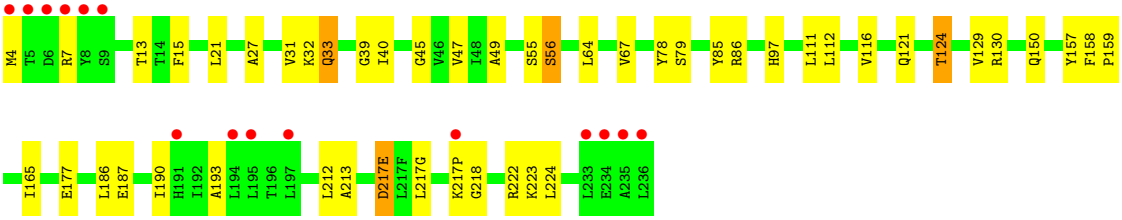
• Molecule 1: Proteasome component Y7

Chain A:



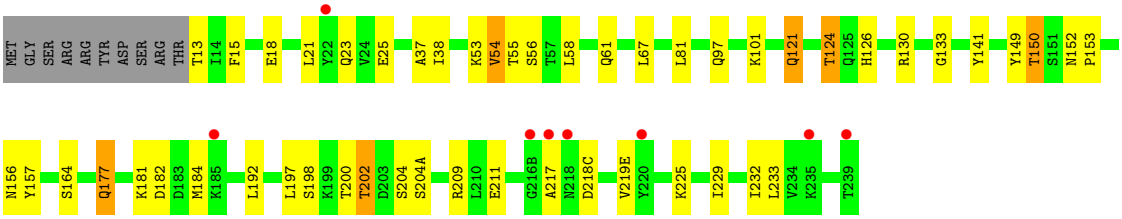
• Molecule 1: Proteasome component Y7

Chain O:



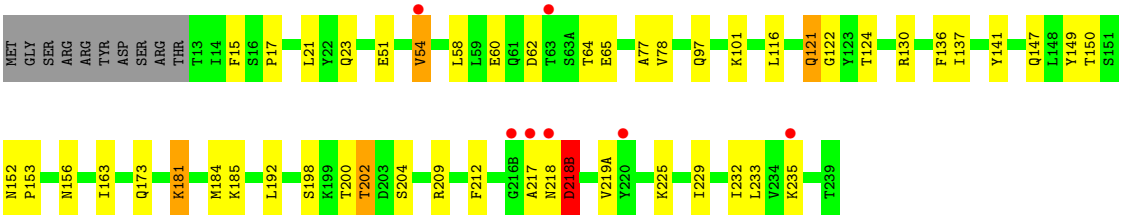
• Molecule 2: Proteasome component Y13

Chain B:



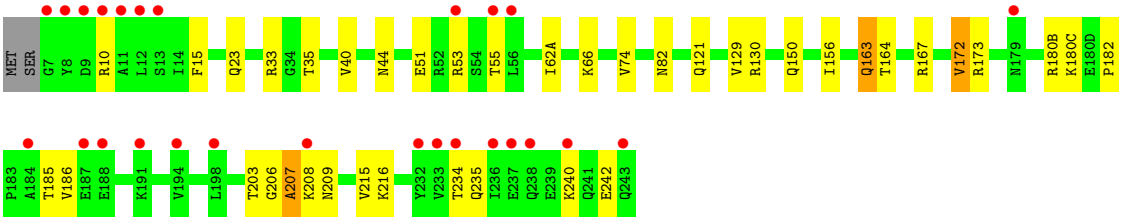
• Molecule 2: Proteasome component Y13

Chain P:



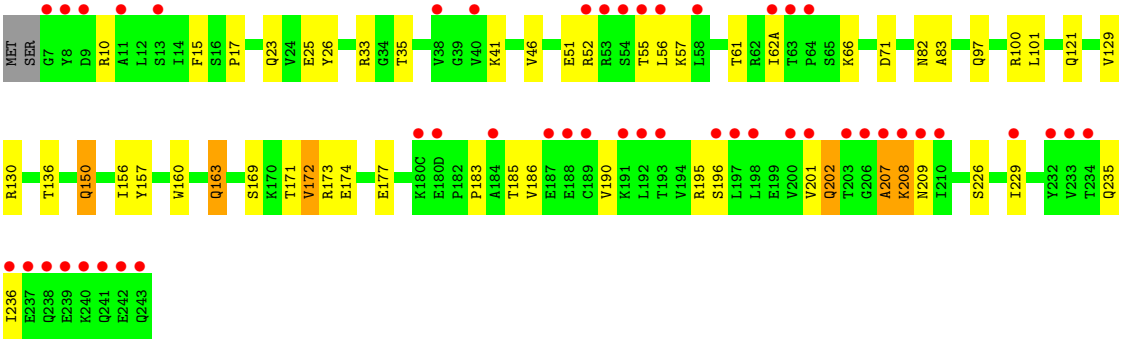
- Molecule 3: Proteasome component PRE6

Chain C:



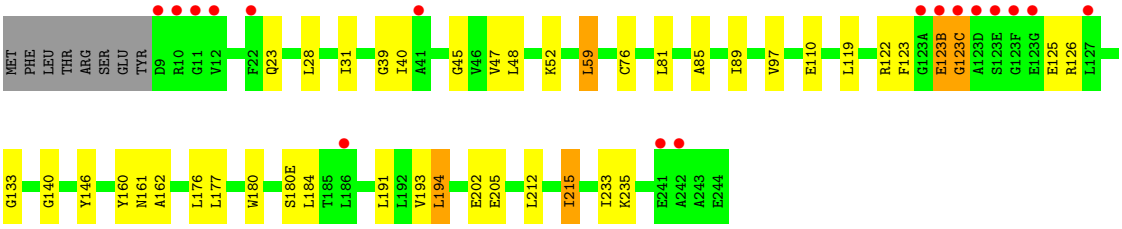
- Molecule 3: Proteasome component PRE6

Chain Q:



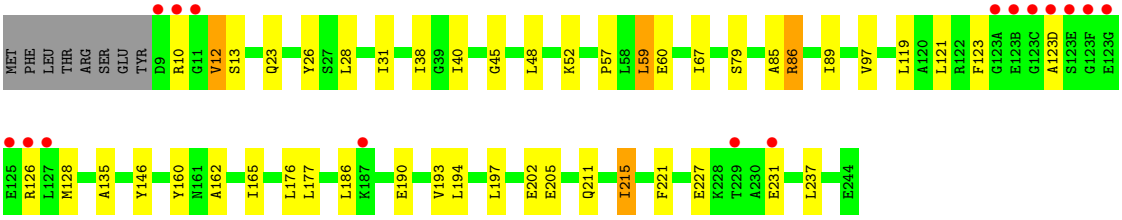
- Molecule 4: Proteasome component PUP2

Chain D:

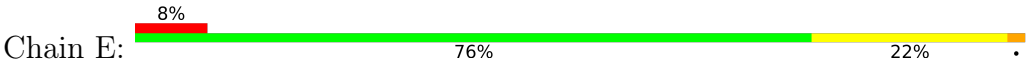


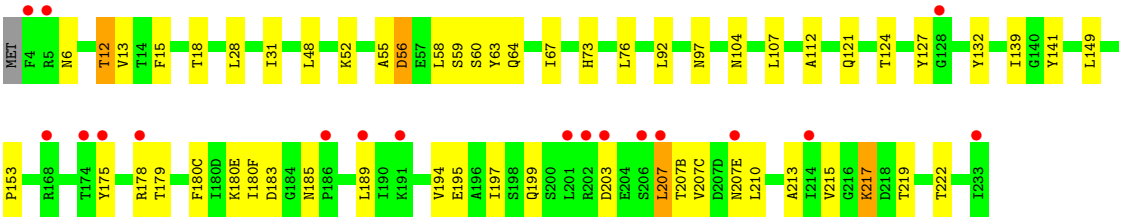
- Molecule 4: Proteasome component PUP2

Chain R:

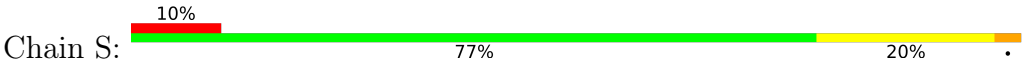


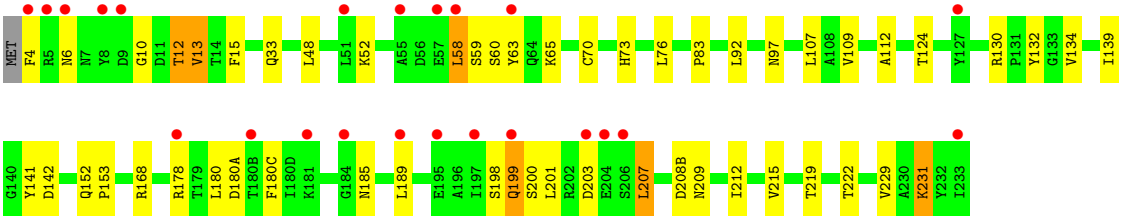
• Molecule 5: Proteasome component PRE5



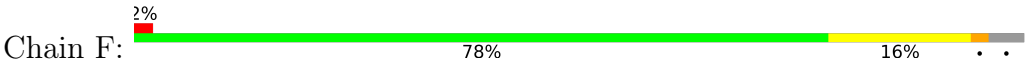


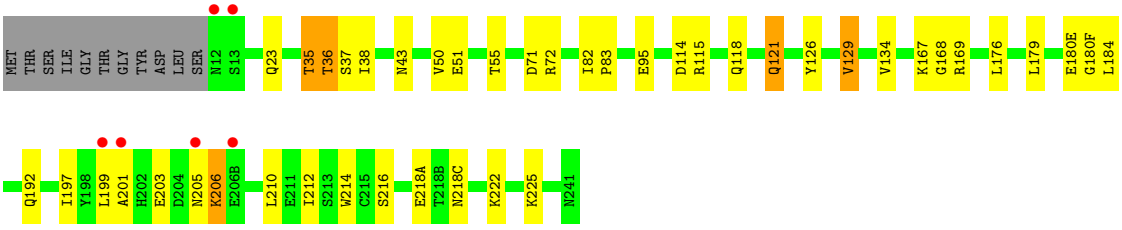
• Molecule 5: Proteasome component PRE5



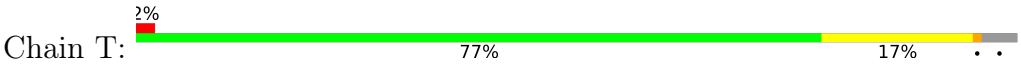


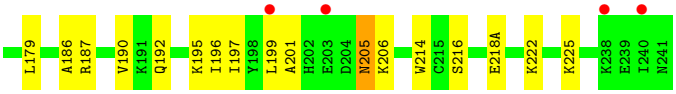
• Molecule 6: Proteasome component C1





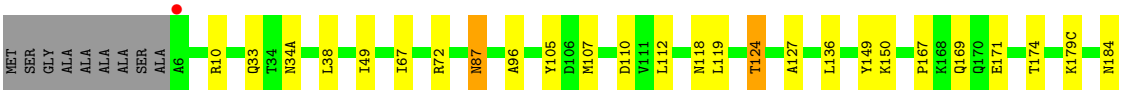
• Molecule 6: Proteasome component C1

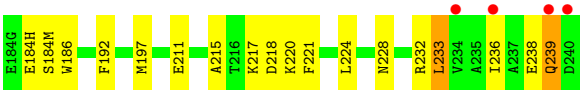




• Molecule 7: Proteasome component C7-alpha

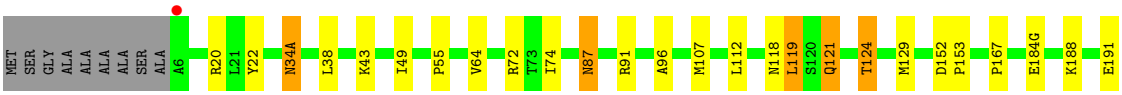
Chain G:

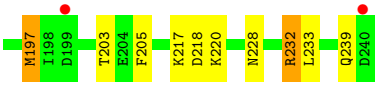




• Molecule 7: Proteasome component C7-alpha

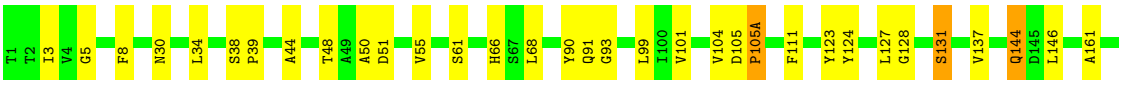
Chain U:

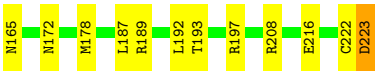




• Molecule 8: Proteasome component PUP1

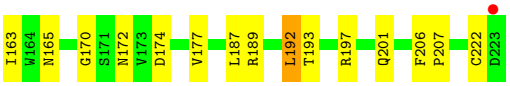
Chain H:





• Molecule 8: Proteasome component PUP1

Chain V:



• Molecule 9: Proteasome component PUP3

Chain I: %
82% 16%

MET S-8 G-1 G1 I2 V3 V4 K9 D10 L20 S28 H36 Y38 E57 N64 P99 I104 K105B K107 I110 F113 C118 A122 I125 A130 E139 N145 L160 D166 S169 G170 W171 V174 I178 K179 K181

D182 R187 D194

• Molecule 9: Proteasome component PUP3

Chain W: %
82% 16%

MET S-8 D-7 S-4 G-1 G1 I2 V12 R19 E32 K33 H36 Y38 G44 Y61 K69 L87 P99 I104 K107 P108 F109 I110 F113 C118 E121 I125 A130 L160 L168 W171 V174 I178 K179 K181 K186

R187 Y188 D194

• Molecule 10: Proteasome component C11

Chain J: 4%
80% 17%

M-1 D1 I2 I6 R7 V8 Q9 V12 A15 S16 S17 I24 T34 R35 E48 A49 I68 R69 E70 Q77 Q85 I90A R92 R93 P94 Y95 Q96 P108 I113 L116 G117 T118 K119 Y129 L138 H141 Y142 R143 M146 L154 L155

L162 M168 D169 F170 D179 K180 D181 G183 Q191 A192 Q193

• Molecule 10: Proteasome component C11

Chain X: 3%
85% 14%

M-1 D1 I2 I6 R7 V8 Q9 D10 I24 D32 K33 T34 T52 A61 L65 I68 Q77 G103 K105B I113 T118 K119 S130 L157 E161 M168 V178 D179 K180 D181 G183 V187 D188 D189 F190 Q191 A192 Q193

• Molecule 11: Proteasome component PRE2

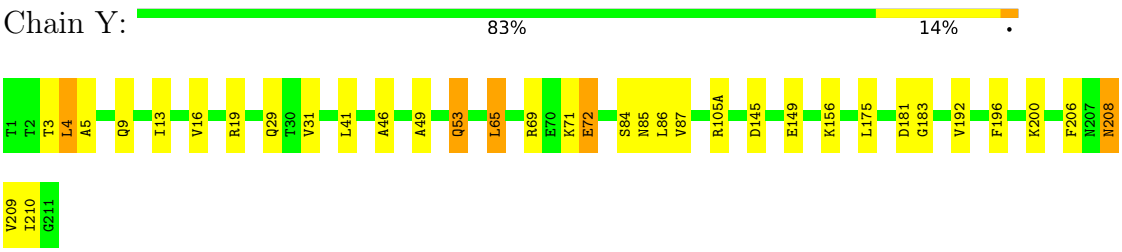
Chain K: 2%
82% 16%

T1 L4 A5 F6 Q9 I13 V31 I35 E36 I37 N38 P39 F40 L41 A46 G60 L65 R69 R73 L86 V87 T99 Y104 T105 R105A K105B T109 V113 D123 V137 D145 V148 A151 I159 L160 A161 N174

E180 D181 G183 N189 V192 F196 K200 E201 F206 N207 N208 V209 I210 G211

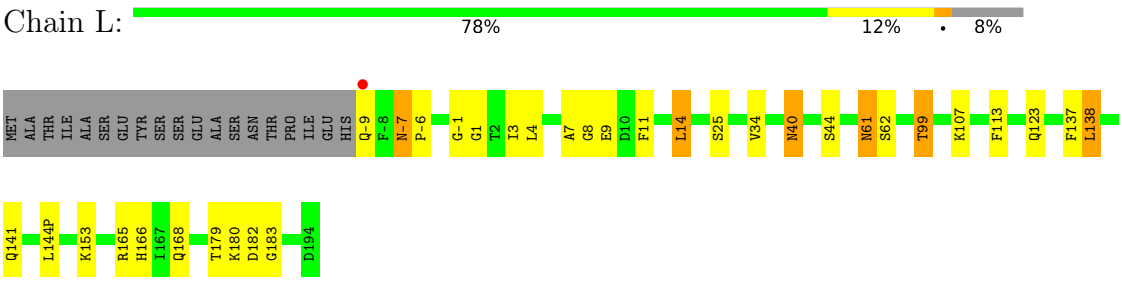
• Molecule 11: Proteasome component PRE2

Chain Y:



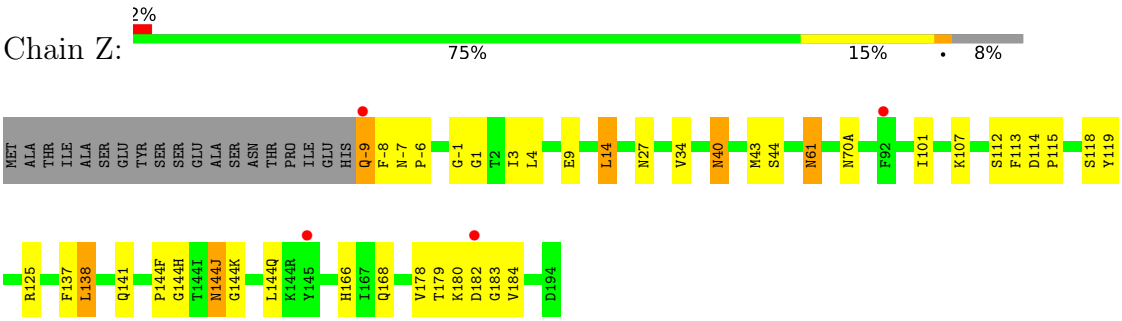
• Molecule 12: Proteasome component C5

Chain L:



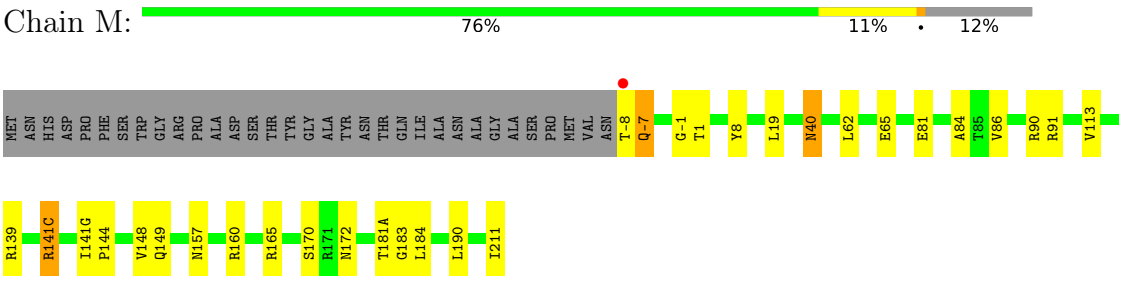
• Molecule 12: Proteasome component C5

Chain Z:



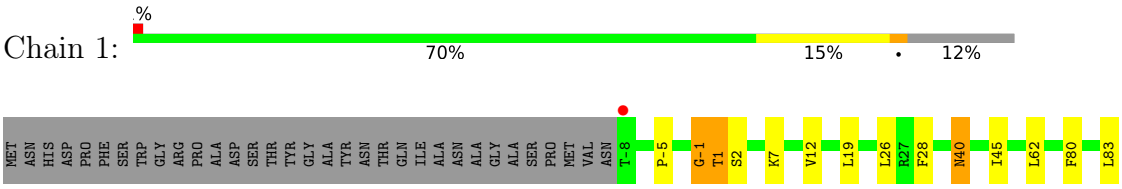
• Molecule 13: Proteasome component PRE4

Chain M:



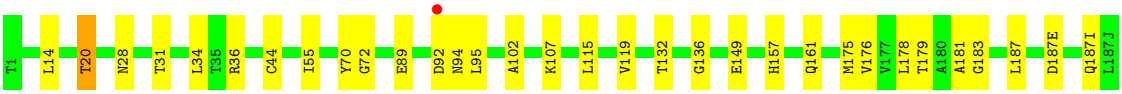
• Molecule 13: Proteasome component PRE4

Chain 1:



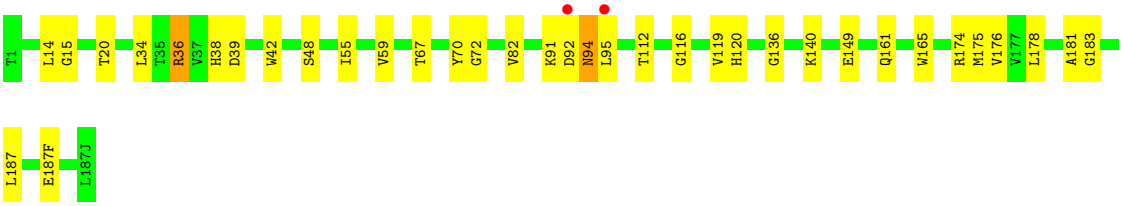
• Molecule 14: Proteasome component PRE3

Chain N:



• Molecule 14: Proteasome component PRE3

Chain 2:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	136.02Å 300.38Å 145.81Å 90.00° 113.73° 90.00°	Depositor
Resolution (Å)	48.68 – 2.78 48.68 – 2.78	Depositor EDS
% Data completeness (in resolution range)	95.8 (48.68-2.78) 95.8 (48.68-2.78)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.96 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.217 , 0.259 0.217 , 0.258	Depositor DCC
R_{free} test set	5191 reflections (2.02%)	wwPDB-VP
Wilson B-factor (Å^2)	58.1	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 25.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	49424	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage'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.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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, MES, L2T

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	0.41	0/1951	0.56	0/2639
1	O	0.42	0/1951	0.55	0/2639
2	B	0.42	0/1857	0.57	0/2513
2	P	0.40	0/1857	0.55	0/2513
3	C	0.41	0/1918	0.56	0/2591
3	Q	0.41	0/1918	0.57	0/2591
4	D	0.42	0/1883	0.61	1/2529 (0.0%)
4	R	0.40	0/1884	0.58	1/2532 (0.0%)
5	E	0.40	0/1819	0.57	0/2451
5	S	0.39	0/1821	0.57	0/2457
6	F	0.43	0/1885	0.55	0/2540
6	T	0.43	0/1886	0.57	0/2543
7	G	0.45	0/1956	0.57	0/2643
7	U	0.45	0/1958	0.57	0/2649
8	H	0.40	0/1714	0.56	0/2320
8	V	0.40	0/1714	0.56	0/2320
9	I	0.46	0/1608	0.60	0/2165
9	W	1.03	1/1609 (0.1%)	0.69	1/2168 (0.0%)
10	J	0.43	0/1611	0.58	0/2167
10	X	0.43	0/1611	0.59	0/2167
11	K	0.42	0/1680	0.58	1/2271 (0.0%)
11	Y	0.42	0/1680	0.59	1/2271 (0.0%)
12	L	0.43	0/1793	0.56	0/2414
12	Z	0.43	0/1793	0.57	0/2414
13	1	0.69	1/1853 (0.1%)	0.81	2/2507 (0.1%)
13	M	0.42	0/1852	0.63	0/2504
14	2	0.42	0/1538	0.56	0/2078
14	N	0.45	0/1538	0.54	0/2078
All	All	0.46	2/50138 (0.0%)	0.59	7/67674 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	W	-1	GLY	C-N	37.20	2.00	1.33
13	1	-1	GLY	C-N	23.09	1.87	1.34

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	1	-1	GLY	O-C-N	-25.36	82.12	122.70
9	W	-1	GLY	O-C-N	-16.35	95.40	123.20
4	D	59	LEU	CA-CB-CG	6.11	129.36	115.30
4	R	59	LEU	CA-CB-CG	5.96	129.02	115.30
11	Y	4	LEU	CA-CB-CG	5.63	128.26	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1915	0	1925	24	0
1	O	1915	0	1925	34	0
2	B	1829	0	1828	30	0
2	P	1829	0	1828	30	0
3	C	1891	0	1899	24	0
3	Q	1891	0	1899	36	0
4	D	1862	0	1832	25	0
4	R	1862	0	1833	27	0
5	E	1795	0	1793	29	0
5	S	1795	0	1795	29	0
6	F	1848	0	1842	24	0
6	T	1848	0	1843	30	0
7	G	1921	0	1907	28	0
7	U	1921	0	1909	24	0
8	H	1685	0	1686	29	0
8	V	1685	0	1686	36	0
9	I	1581	0	1574	29	0
9	W	1581	0	1574	26	0
10	J	1585	0	1591	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	X	1585	0	1591	15	0
11	K	1644	0	1594	23	0
11	Y	1644	0	1594	18	0
12	L	1757	0	1712	23	0
12	Z	1757	0	1712	29	0
13	1	1824	0	1833	31	0
13	M	1824	0	1833	23	0
14	2	1512	0	1478	25	0
14	N	1512	0	1478	22	0
15	F	2	0	0	0	0
15	G	1	0	0	0	0
15	H	1	0	0	0	0
15	I	2	0	0	0	0
15	K	1	0	0	0	0
15	L	2	0	0	0	0
15	N	1	0	0	0	0
16	K	44	0	39	5	0
16	Y	44	0	39	2	0
17	K	12	0	13	1	0
17	Y	12	0	13	1	0
18	L	2	0	0	0	0
18	Z	2	0	0	0	0
All	All	49424	0	49098	660	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 7.

The worst 5 of 660 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)
16:K:213:L2T:C6	16:K:213:L2T:CL7	1.75	1.60
13:1:-1:GLY:C	13:1:1:THR:N	1.87	1.28
11:K:181:ASP:C	11:K:183:GLY:HA2	1.64	1.16
9:W:-1:GLY:C	9:W:1:GLY:N	2.00	1.15
13:1:-1:GLY:C	13:1:1:THR:H1	1.45	1.12

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	246/250 (98%)	234 (95%)	9 (4%)	3 (1%)	13	36
1	O	246/250 (98%)	235 (96%)	10 (4%)	1 (0%)	34	64
2	B	231/245 (94%)	216 (94%)	10 (4%)	5 (2%)	6	20
2	P	231/245 (94%)	218 (94%)	8 (4%)	5 (2%)	6	20
3	C	235/243 (97%)	229 (97%)	5 (2%)	1 (0%)	34	64
3	Q	235/243 (97%)	228 (97%)	3 (1%)	4 (2%)	9	27
4	D	232/250 (93%)	219 (94%)	11 (5%)	2 (1%)	17	44
4	R	234/250 (94%)	216 (92%)	17 (7%)	1 (0%)	34	64
5	E	223/234 (95%)	211 (95%)	8 (4%)	4 (2%)	8	25
5	S	227/234 (97%)	211 (93%)	12 (5%)	4 (2%)	8	25
6	F	231/248 (93%)	217 (94%)	14 (6%)	0	100	100
6	T	233/248 (94%)	222 (95%)	10 (4%)	1 (0%)	34	64
7	G	235/252 (93%)	224 (95%)	10 (4%)	1 (0%)	34	64
7	U	239/252 (95%)	234 (98%)	3 (1%)	2 (1%)	19	47
8	H	216/222 (97%)	207 (96%)	8 (4%)	1 (0%)	29	58
8	V	216/222 (97%)	205 (95%)	11 (5%)	0	100	100
9	I	196/205 (96%)	188 (96%)	8 (4%)	0	100	100
9	W	198/205 (97%)	190 (96%)	8 (4%)	0	100	100
10	J	193/198 (98%)	183 (95%)	8 (4%)	2 (1%)	15	41
10	X	193/198 (98%)	184 (95%)	6 (3%)	3 (2%)	9	28
11	K	208/212 (98%)	202 (97%)	5 (2%)	1 (0%)	29	58
11	Y	208/212 (98%)	198 (95%)	9 (4%)	1 (0%)	29	58
12	L	216/241 (90%)	212 (98%)	4 (2%)	0	100	100
12	Z	216/241 (90%)	208 (96%)	7 (3%)	1 (0%)	29	58
13	1	227/266 (85%)	217 (96%)	9 (4%)	1 (0%)	34	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	M	225/266 (85%)	215 (96%)	9 (4%)	1 (0%)	34	64
14	2	188/196 (96%)	181 (96%)	7 (4%)	0	100	100
14	N	188/196 (96%)	181 (96%)	7 (4%)	0	100	100
All	All	6166/6524 (94%)	5885 (95%)	236 (4%)	45 (1%)	22	50

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	54	VAL
5	E	6	ASN
5	E	203	ASP
5	E	217	LYS
7	G	239	GLN

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%)	195 (93%)	14 (7%)	16	40
1	O	209/209 (100%)	198 (95%)	11 (5%)	22	51
2	B	195/204 (96%)	181 (93%)	14 (7%)	14	36
2	P	195/204 (96%)	180 (92%)	15 (8%)	13	32
3	C	213/215 (99%)	193 (91%)	20 (9%)	8	23
3	Q	213/215 (99%)	200 (94%)	13 (6%)	18	45
4	D	198/206 (96%)	186 (94%)	12 (6%)	18	45
4	R	198/206 (96%)	187 (94%)	11 (6%)	21	48
5	E	192/193 (100%)	174 (91%)	18 (9%)	8	23
5	S	192/193 (100%)	175 (91%)	17 (9%)	9	26
6	F	196/205 (96%)	181 (92%)	15 (8%)	13	32
6	T	196/205 (96%)	181 (92%)	15 (8%)	13	32
7	G	207/210 (99%)	189 (91%)	18 (9%)	10	27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	U	207/210 (99%)	192 (93%)	15 (7%)	14	36
8	H	181/181 (100%)	173 (96%)	8 (4%)	28	58
8	V	181/181 (100%)	172 (95%)	9 (5%)	24	53
9	I	172/173 (99%)	164 (95%)	8 (5%)	26	56
9	W	172/173 (99%)	163 (95%)	9 (5%)	23	52
10	J	175/175 (100%)	164 (94%)	11 (6%)	18	43
10	X	175/175 (100%)	165 (94%)	10 (6%)	20	48
11	K	169/169 (100%)	154 (91%)	15 (9%)	9	26
11	Y	169/169 (100%)	155 (92%)	14 (8%)	11	29
12	L	185/201 (92%)	175 (95%)	10 (5%)	22	50
12	Z	185/201 (92%)	177 (96%)	8 (4%)	29	59
13	1	199/224 (89%)	184 (92%)	15 (8%)	13	34
13	M	199/224 (89%)	190 (96%)	9 (4%)	27	58
14	2	162/162 (100%)	155 (96%)	7 (4%)	29	59
14	N	162/162 (100%)	152 (94%)	10 (6%)	18	44
All	All	5306/5454 (97%)	4955 (93%)	351 (7%)	16	41

5 of 351 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	R	237	LEU
8	V	197	ARG
5	S	76	LEU
6	T	205	ASN
10	X	52	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 151 such sidechains are listed below:

Mol	Chain	Res	Type
8	V	66	HIS
13	1	157	ASN
10	X	54	GLN
12	Z	40	ASN
14	2	161	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 10 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
16	L2T	K	213	-	47,47,47	1.95	2 (4%)	61,62,62	0.81	1 (1%)
16	L2T	Y	212	-	47,47,47	1.86	2 (4%)	61,62,62	0.81	2 (3%)
17	MES	K	214	-	12,12,12	2.02	1 (8%)	14,16,16	1.45	1 (7%)
17	MES	Y	213	-	12,12,12	2.19	1 (8%)	14,16,16	1.48	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
16	L2T	K	213	-	-	2/39/39/39	0/4/4/4
16	L2T	Y	212	-	-	6/39/39/39	0/4/4/4
17	MES	K	214	-	-	5/6/14/14	0/1/1/1
17	MES	Y	213	-	-	1/6/14/14	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	K	213	L2T	CL7-C6	12.01	1.75	1.51
16	Y	212	L2T	CL7-C6	11.35	1.73	1.51
17	Y	213	MES	C8-S	-7.21	1.67	1.77
17	K	214	MES	C8-S	-6.57	1.68	1.77
16	K	213	L2T	C37-C38	-4.71	1.39	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
17	K	214	MES	O3S-S-C8	3.75	111.83	105.77
17	Y	213	MES	O3S-S-C8	3.29	111.09	105.77
16	Y	212	L2T	O36-C37-C38	3.27	118.96	109.16
17	Y	213	MES	O2S-S-C8	2.85	110.34	106.92
16	K	213	L2T	O36-C37-C38	2.26	115.95	109.16

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	K	214	MES	C8-C7-N4-C3
17	K	214	MES	C8-C7-N4-C5
17	K	214	MES	C7-C8-S-O1S
17	K	214	MES	C7-C8-S-O2S
17	K	214	MES	C7-C8-S-O3S

There are no ring outliers.

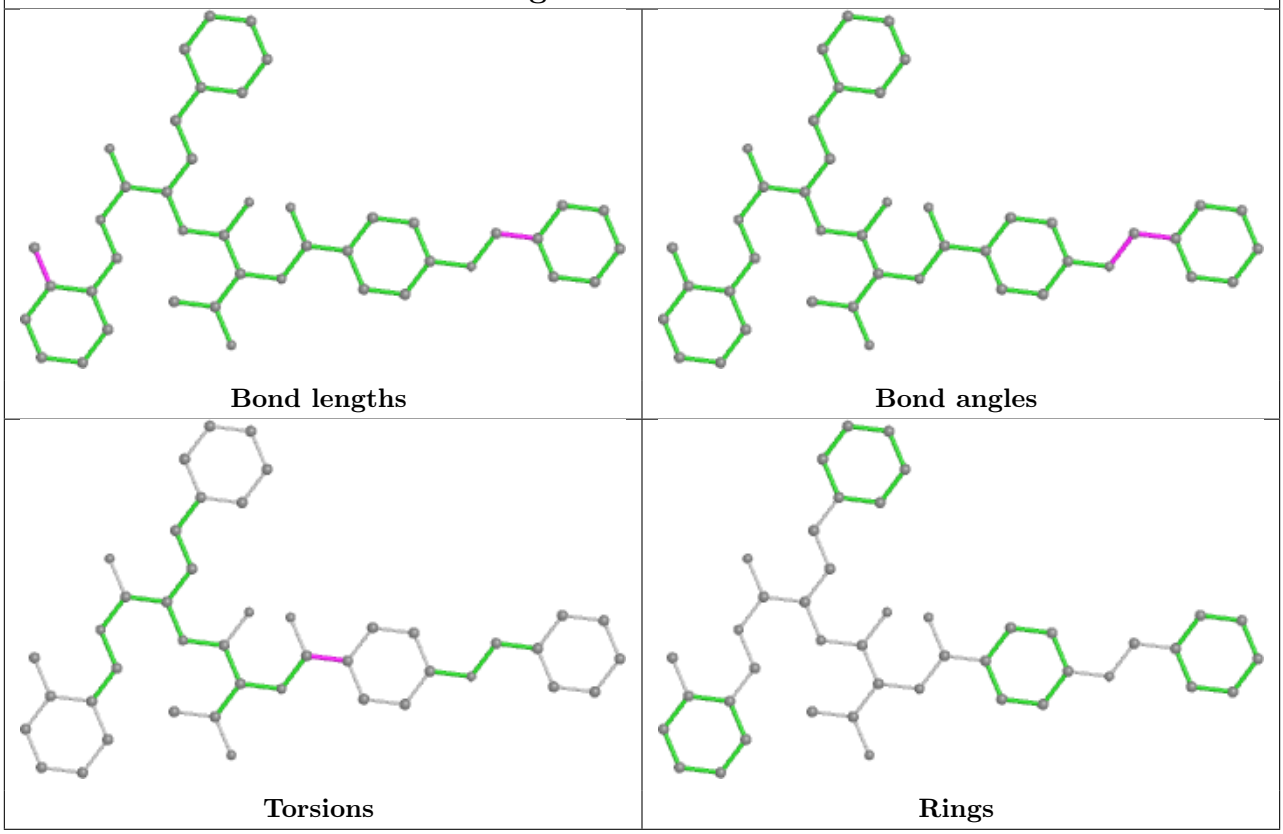
4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	K	213	L2T	5	0
16	Y	212	L2T	2	0
17	K	214	MES	1	0
17	Y	213	MES	1	0

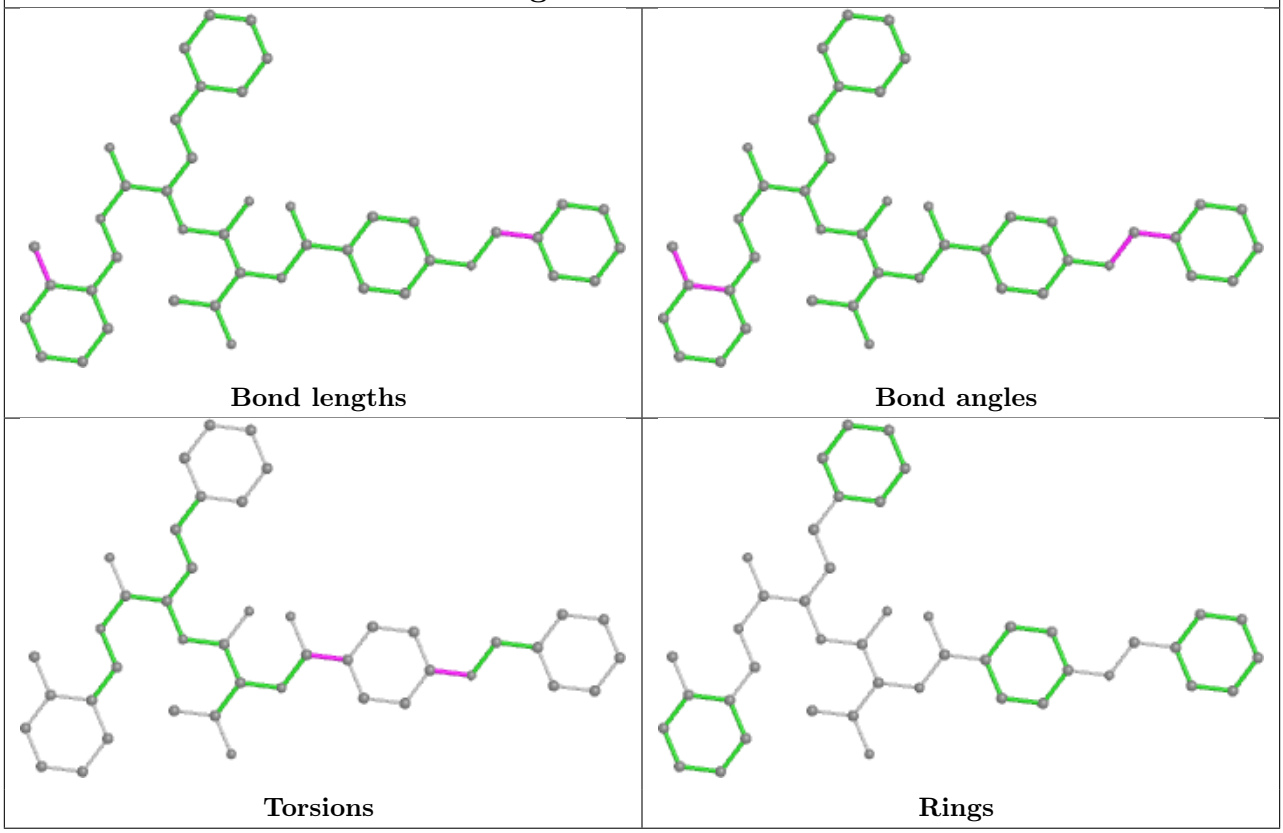
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring

in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand L2T K 213



Ligand L2T Y 212



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	D	4
5	E	4
4	R	3
13	M	3
14	N	3
7	G	3
14	2	3
13	1	3
9	I	3
9	W	3
3	C	2
10	J	2

Continued on next page...

Continued from previous page...

Mol	Chain	Number of breaks
10	X	2
3	Q	2
5	S	2
6	F	2
12	Z	2
12	L	2
8	V	2
8	H	2
11	K	1
1	A	1
1	O	1
2	P	1
11	Y	1
6	T	1
7	U	1
2	B	1

The worst 5 of 60 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	123(G):GLU	C	125:GLU	N	4.65
1	C	203:THR	C	206:GLY	N	4.56
1	R	123(G):GLU	C	125:GLU	N	4.48
1	J	-1:MET	C	1:ASP	N	4.02
1	X	-1:MET	C	1:ASP	N	3.71

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, 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%)	0.13	7 (2%)	53	48	37, 54, 80, 110	0
1	O	250/250 (100%)	0.27	15 (6%)	21	16	41, 62, 88, 115	0
2	B	235/245 (95%)	0.21	8 (3%)	45	39	41, 60, 92, 106	0
2	P	235/245 (95%)	0.29	7 (2%)	50	45	39, 63, 97, 108	0
3	C	241/243 (99%)	0.54	26 (10%)	5	4	43, 69, 110, 125	0
3	Q	241/243 (99%)	0.90	48 (19%)	1	0	45, 74, 121, 140	0
4	D	242/250 (96%)	0.42	17 (7%)	16	11	40, 65, 101, 122	0
4	R	242/250 (96%)	0.43	16 (6%)	18	13	41, 65, 101, 129	0
5	E	233/234 (99%)	0.34	18 (7%)	13	9	48, 63, 91, 111	0
5	S	233/234 (99%)	0.48	23 (9%)	7	5	46, 64, 97, 125	0
6	F	237/248 (95%)	0.06	6 (2%)	57	52	39, 57, 95, 107	0
6	T	237/248 (95%)	0.12	6 (2%)	57	52	38, 58, 85, 109	0
7	G	243/252 (96%)	0.08	5 (2%)	63	59	36, 51, 80, 111	0
7	U	243/252 (96%)	-0.05	3 (1%)	79	76	35, 51, 73, 101	0
8	H	222/222 (100%)	-0.07	0	100	100	34, 48, 65, 99	0
8	V	222/222 (100%)	-0.06	1 (0%)	91	90	39, 52, 72, 107	0
9	I	204/205 (99%)	-0.03	2 (0%)	82	80	36, 48, 66, 82	0
9	W	204/205 (99%)	0.06	2 (0%)	82	80	35, 51, 69, 86	0
10	J	198/198 (100%)	0.09	7 (3%)	44	38	38, 51, 68, 127	0
10	X	198/198 (100%)	0.10	6 (3%)	50	45	39, 53, 68, 125	0
11	K	212/212 (100%)	-0.04	4 (1%)	66	63	36, 50, 69, 88	0
11	Y	212/212 (100%)	-0.13	0	100	100	36, 51, 70, 77	0
12	L	222/241 (92%)	-0.05	1 (0%)	91	90	36, 51, 71, 82	0
12	Z	222/241 (92%)	0.02	4 (1%)	68	65	34, 50, 72, 83	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
13	1	233/266 (87%)	-0.10	2 (0%)	84	82	34, 48, 63, 67	0
13	M	233/266 (87%)	-0.06	1 (0%)	92	92	36, 51, 65, 71	0
14	2	196/196 (100%)	-0.12	2 (1%)	82	80	34, 46, 64, 79	0
14	N	196/196 (100%)	-0.13	1 (0%)	91	90	37, 45, 60, 76	0
All	All	6336/6524 (97%)	0.14	238 (3%)	40	35	34, 54, 91, 140	0

The worst 5 of 238 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	123(C)	GLY	12.0
4	R	123(E)	SER	11.5
4	R	123(F)	GLY	11.1
3	C	7	GLY	10.9
4	R	123(C)	GLY	10.8

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 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, 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	MG	F	242	1/1	0.69	0.41	74,74,74,74	0
15	MG	I	196	1/1	0.69	0.31	68,68,68,68	0
15	MG	F	243	1/1	0.80	0.81	118,118,118,118	0
15	MG	L	195	1/1	0.81	0.24	66,66,66,66	0
15	MG	I	195	1/1	0.87	0.58	72,72,72,72	0
15	MG	L	196	1/1	0.88	0.20	52,52,52,52	0
16	L2T	K	213	44/44	0.92	0.21	45,47,55,55	1

Continued on next page...

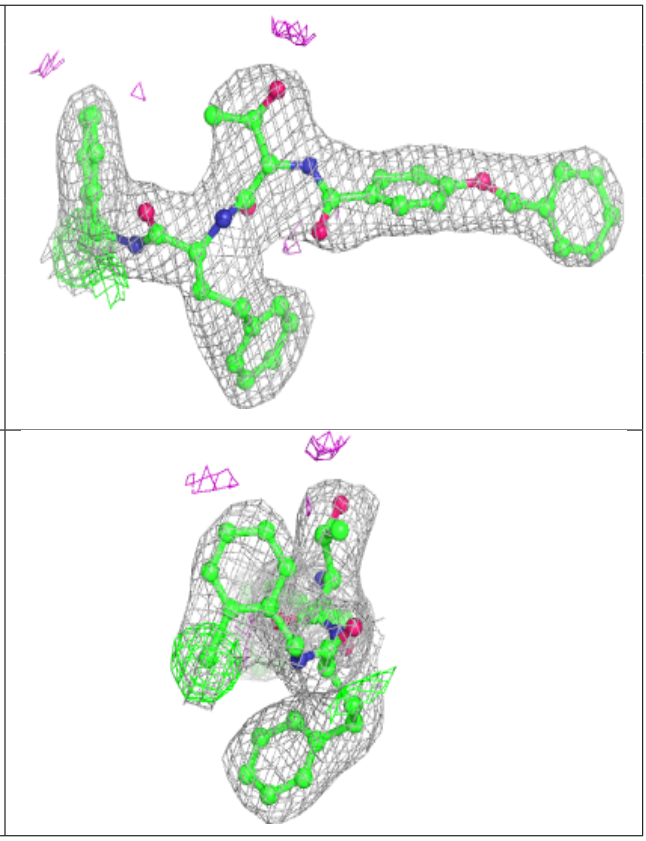
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
17	MES	K	214	12/12	0.92	0.20	90,90,90,90	0
15	MG	H	224	1/1	0.93	0.15	57,57,57,57	0
15	MG	N	188	1/1	0.93	0.11	50,50,50,50	0
15	MG	K	212	1/1	0.93	0.28	55,55,55,55	0
16	L2T	Y	212	44/44	0.93	0.22	49,51,56,57	1
15	MG	G	241	1/1	0.93	0.10	54,54,54,54	0
17	MES	Y	213	12/12	0.95	0.20	97,98,98,98	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

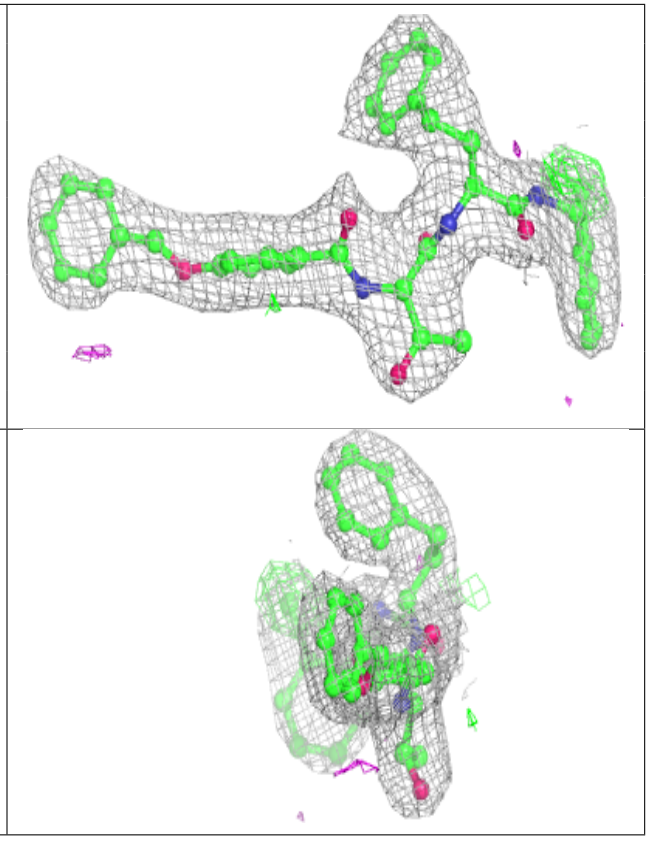
Electron density around L2T K 213:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around L2T Y 212:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.