wwPDB X-ray Structure Validation Summary Report (i)

Nov 6, 2023 – 12:12 PM EST

PDB ID	:	4JSQ
Title	:	Yeast 20S proteasome in complex with the dimerized linear mimetic of TMC-95A - yCP:4e
Authors	:	Desvergne, A.; Genin, E.; Marechal, X.; Gallastegui, N.; Dufau, L.; Richy, N.; Groll, M.; Vidal, J.; Reboud-Ravaux, M.
Deposited on	:	2013-03-22
Resolution	:	2.80 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
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 (1)) were used in the production of this report:

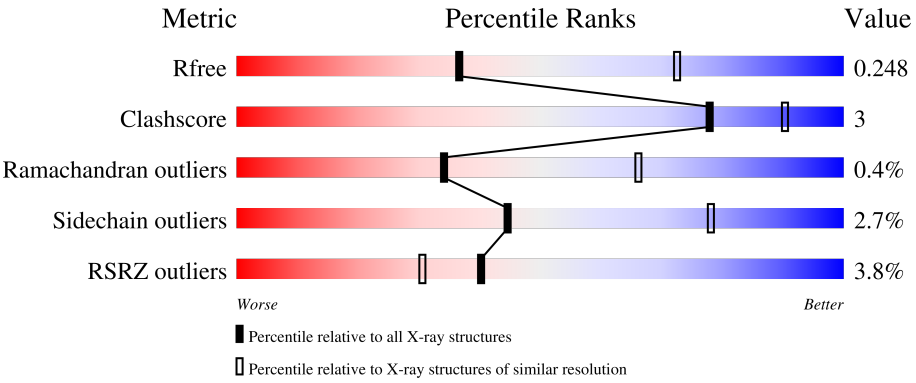
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

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



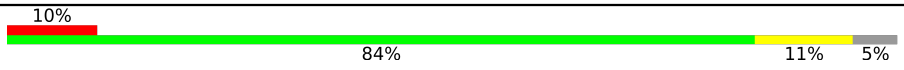
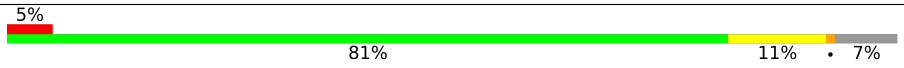
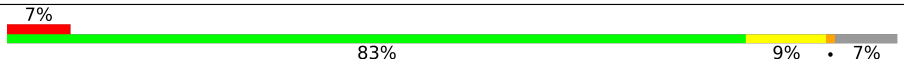
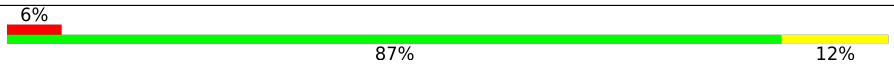
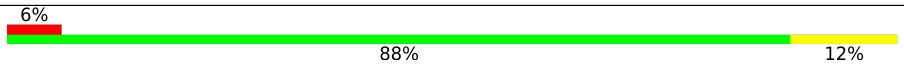
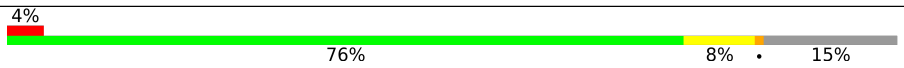
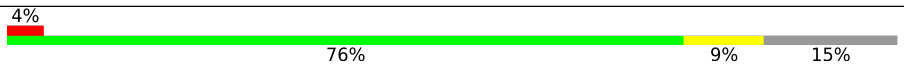
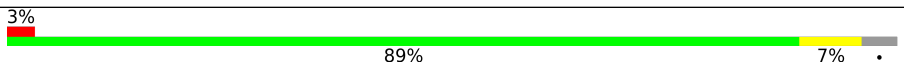
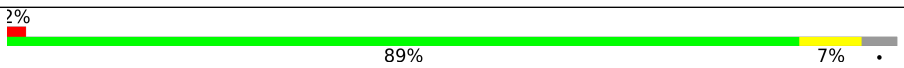
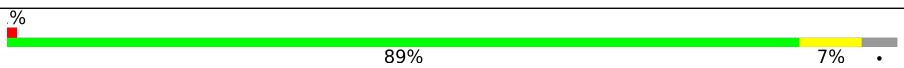
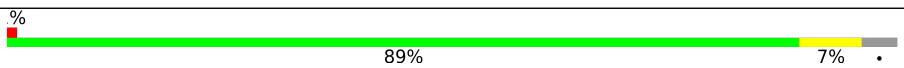
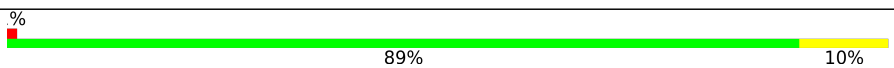
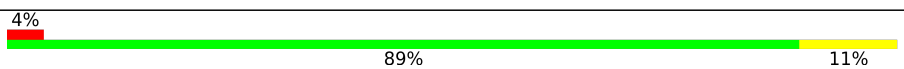
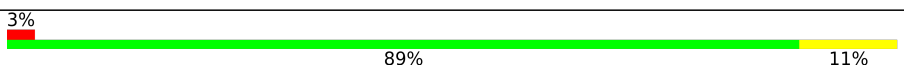
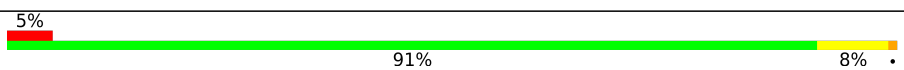
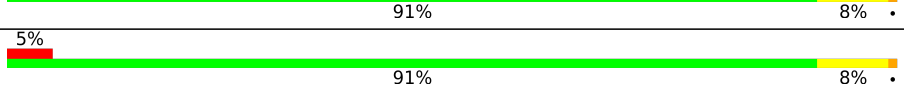
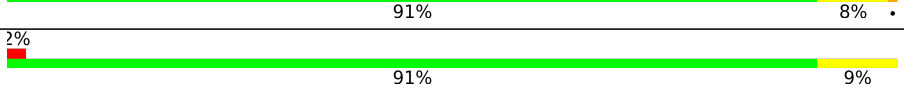
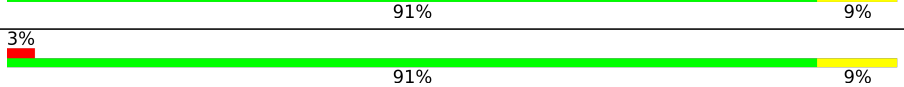
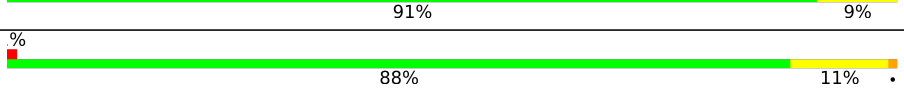
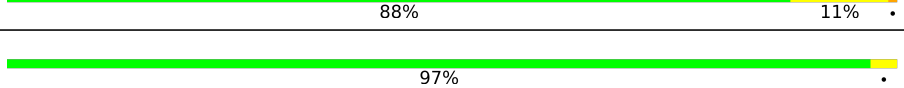
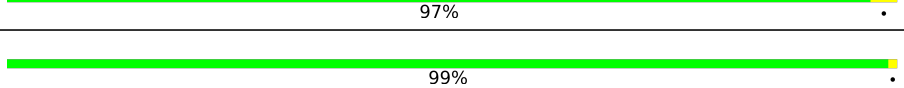
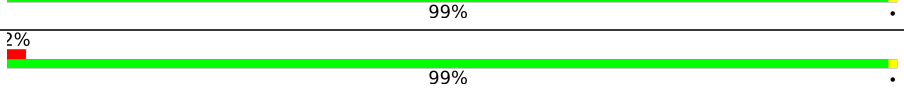
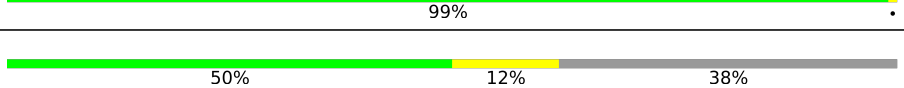
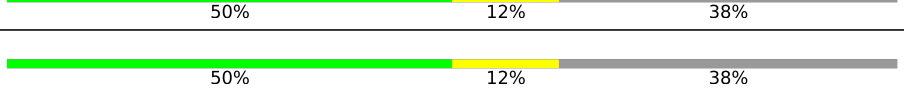
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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% 93% 7%
1	O	250	4% 92% 8%
2	B	258	7% 81% 14% 5%
2	P	258	7% 81% 14% 5%
3	C	254	5% 81% 13% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	
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	233	
13	a	233	
14	N	196	
14	b	196	
15	c	8	
15	d	8	

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	ACA	c	4	-	-	-	X
15	ACA	d	4	-	-	-	X

2 Entry composition (i)

There are 17 unique types of molecules in this entry. The entry contains 51011 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	241	Total	C	N	O	S	0	0	0
			1890	1181	331	374	4			
3	Q	241	Total	C	N	O	S	0	0	0
			1890	1181	331	374	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	242	Total	C	N	O	S	0	0	0
			1861	1162	314	378	7			
4	R	242	Total	C	N	O	S	0	0	0
			1861	1162	314	378	7			

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	244	Total	C	N	O	S	0	0	0
			1896	1205	330	357	4			
6	T	244	Total	C	N	O	S	0	0	0
			1896	1205	330	357	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			
7	U	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	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	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			

Continued on next page...

Continued from previous page...

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

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

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

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

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

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

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

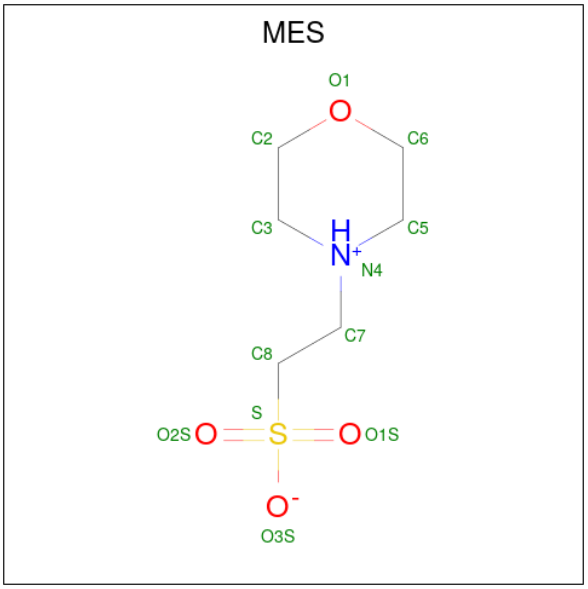
- Molecule 14 is a protein called Proteasome subunit beta type-1.

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

- Molecule 15 is a protein called TMC-95A mimic ligand yCP:4e.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	c	5	Total	C	N	O	0	0	0
			56	43	6	7			
15	d	5	Total	C	N	O	0	0	0
			56	43	6	7			

- Molecule 16 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



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

- Molecule 17 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	59	Total	O	0	0
			59	59		
17	B	39	Total	O	0	0
			39	39		
17	C	43	Total	O	0	0
			43	43		
17	D	36	Total	O	0	0
			36	36		
17	E	21	Total	O	0	0
			21	21		
17	F	47	Total	O	0	0
			47	47		
17	G	60	Total	O	0	0
			60	60		
17	H	52	Total	O	0	0
			52	52		

Continued on next page...

Continued from previous page...

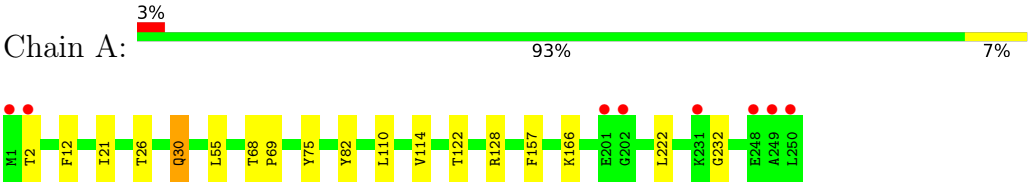
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	I	64	Total O 64 64	0	0
17	J	53	Total O 53 53	0	0
17	K	48	Total O 48 48	0	0
17	L	54	Total O 54 54	0	0
17	M	81	Total O 81 81	0	0
17	N	58	Total O 58 58	0	0
17	O	34	Total O 34 34	0	0
17	P	30	Total O 30 30	0	0
17	Q	29	Total O 29 29	0	0
17	R	27	Total O 27 27	0	0
17	S	18	Total O 18 18	0	0
17	T	43	Total O 43 43	0	0
17	U	55	Total O 55 55	0	0
17	V	50	Total O 50 50	0	0
17	W	62	Total O 62 62	0	0
17	X	41	Total O 41 41	0	0
17	Y	50	Total O 50 50	0	0
17	Z	49	Total O 49 49	0	0
17	a	78	Total O 78 78	0	0
17	b	56	Total O 56 56	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

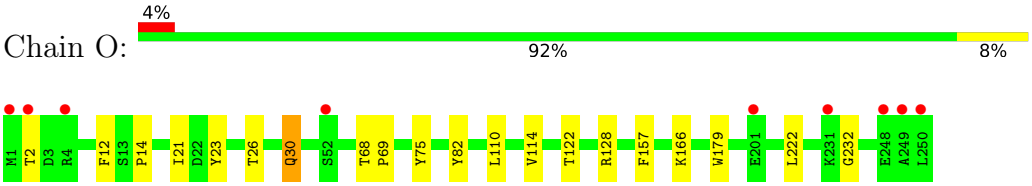
• Molecule 1: Proteasome subunit alpha type-2

Chain A:



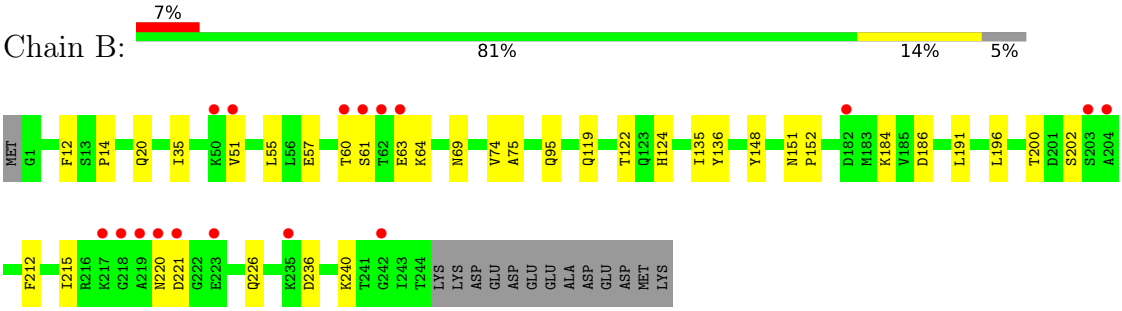
• Molecule 1: Proteasome subunit alpha type-2

Chain O:



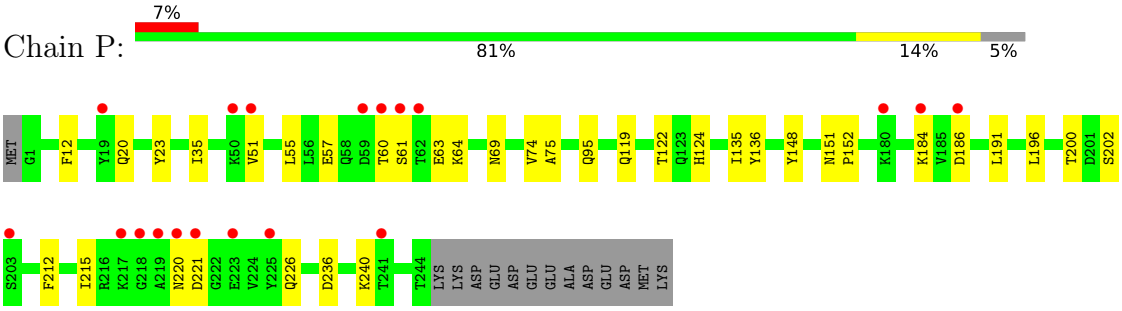
• Molecule 2: Proteasome subunit alpha type-3

Chain B:



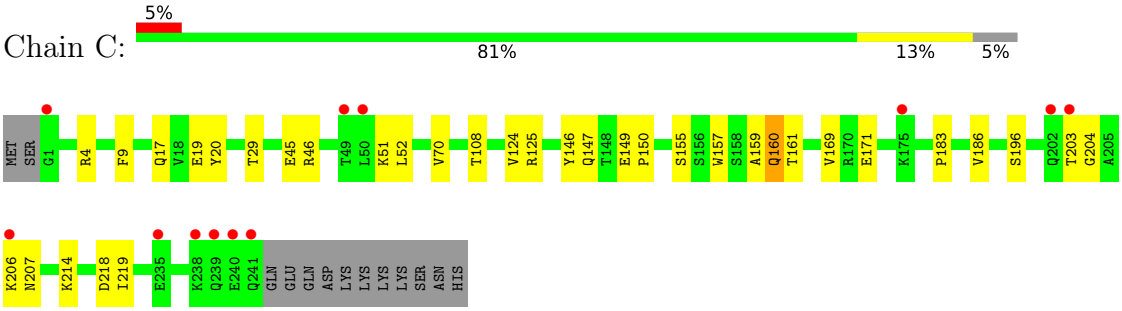
• Molecule 2: Proteasome subunit alpha type-3

Chain P:



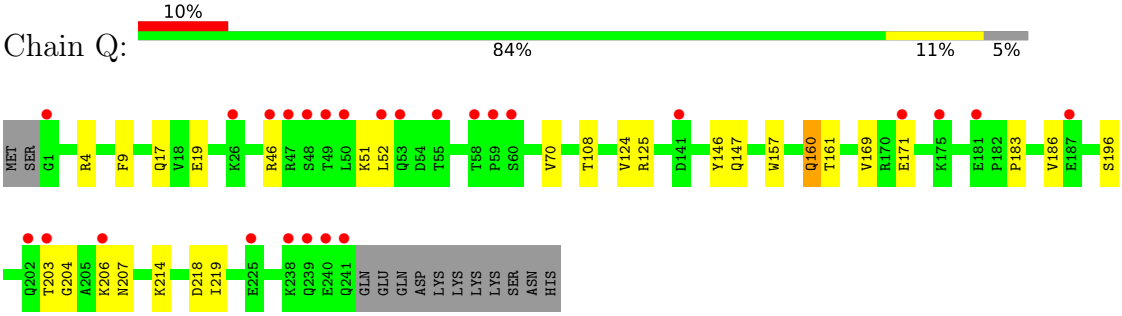
• Molecule 3: Proteasome subunit alpha type-4

Chain C:



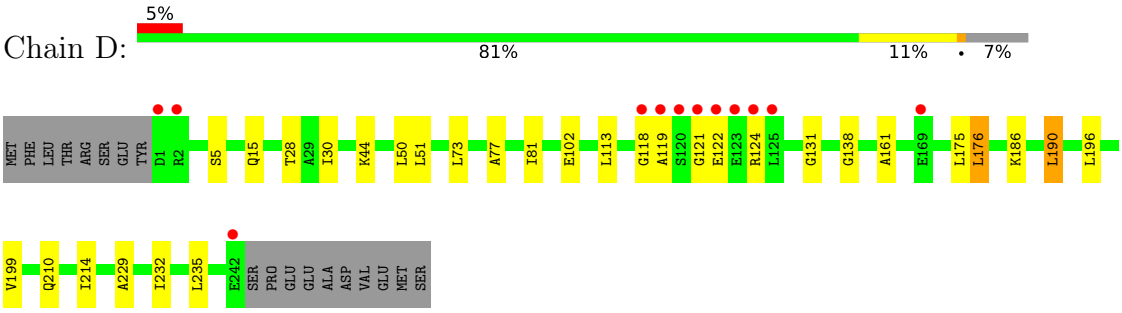
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



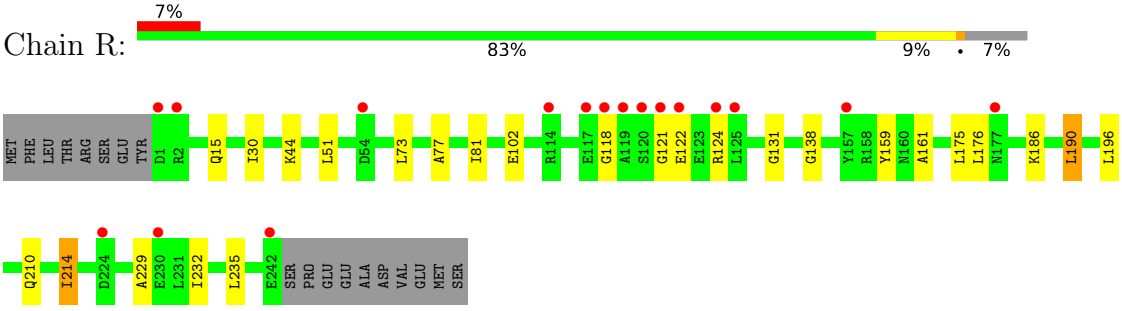
• Molecule 4: Proteasome subunit alpha type-5

Chain D:



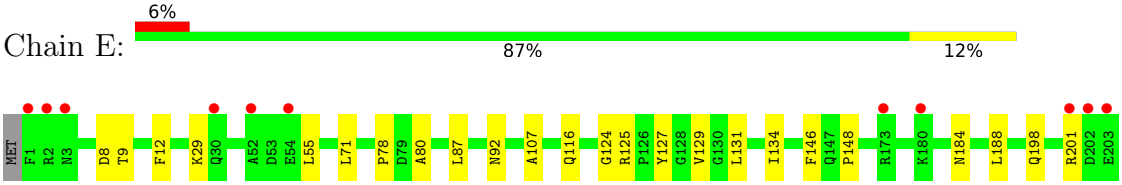
• Molecule 4: Proteasome subunit alpha type-5

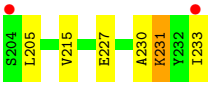
Chain R:



• Molecule 5: Proteasome subunit alpha type-6

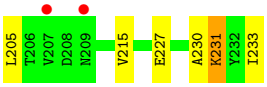
Chain E:





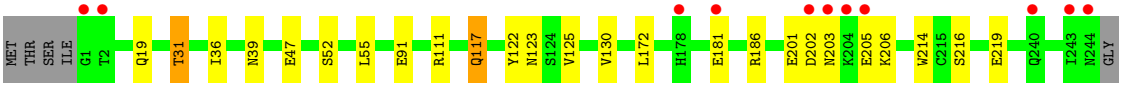
• Molecule 5: Proteasome subunit alpha type-6

Chain S:



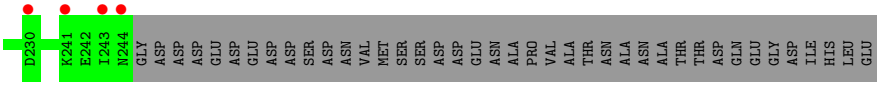
• Molecule 6: Probable proteasome subunit alpha type-7

Chain F:



• Molecule 6: Probable proteasome subunit alpha type-7

Chain T:

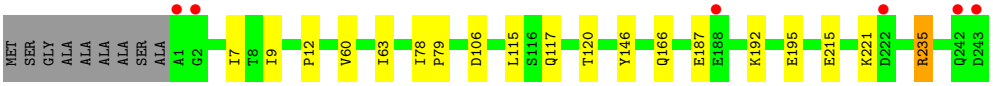


• Molecule 7: Proteasome subunit alpha type-1

Chain G:

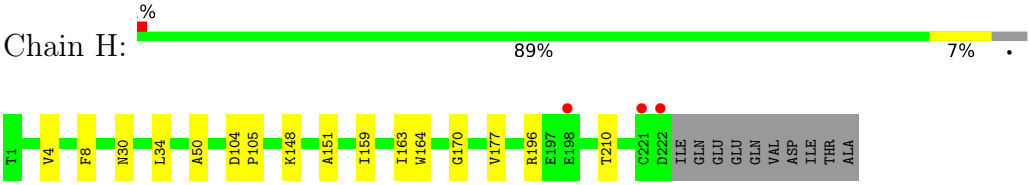
• Molecule 7: Proteasome subunit alpha type-1

Chain U:



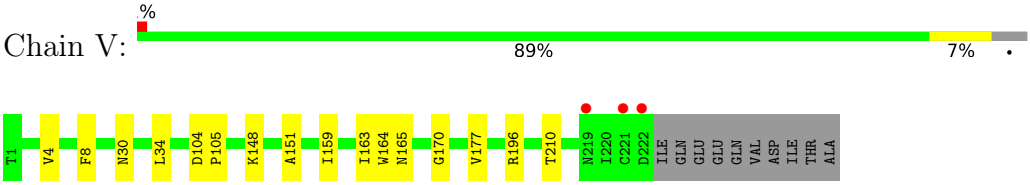
• Molecule 8: Proteasome subunit beta type-2

Chain H:



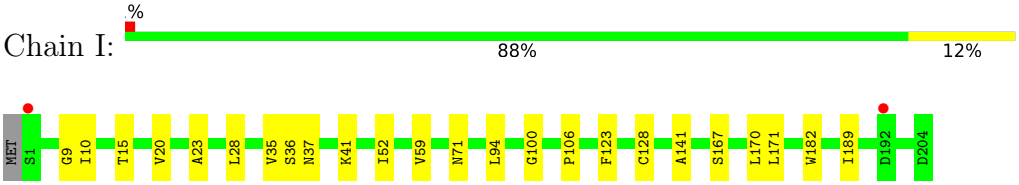
• Molecule 8: Proteasome subunit beta type-2

Chain V:



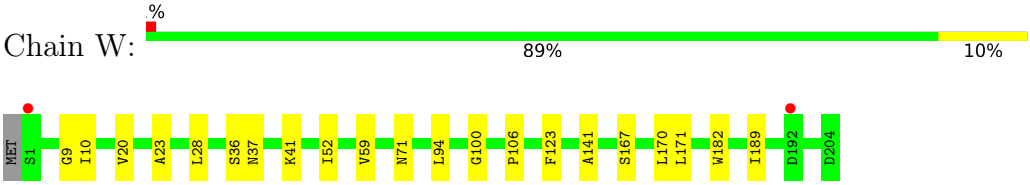
• Molecule 9: Proteasome subunit beta type-3

Chain I:



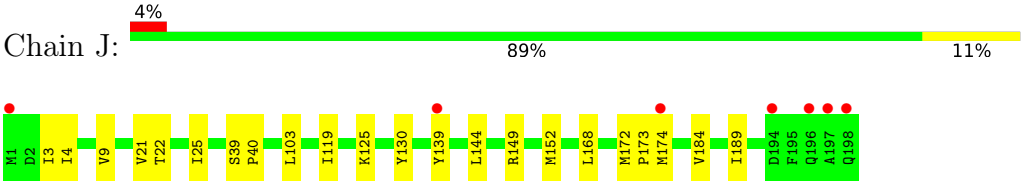
• Molecule 9: Proteasome subunit beta type-3

Chain W:



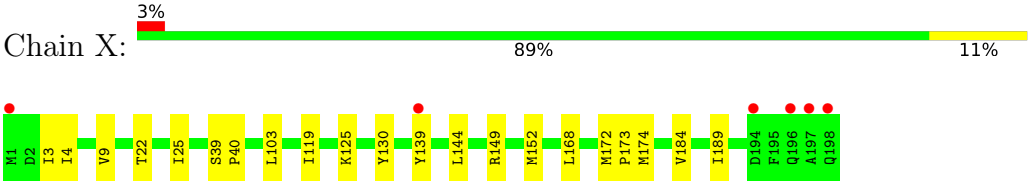
• Molecule 10: Proteasome subunit beta type-4

Chain J:



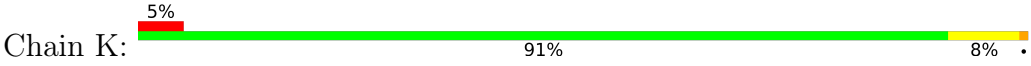
• Molecule 10: Proteasome subunit beta type-4

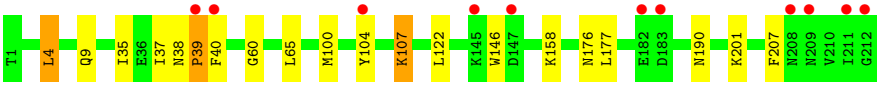
Chain X:



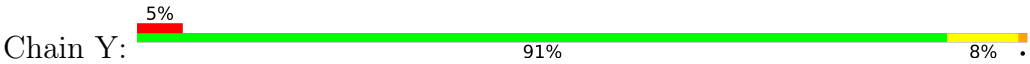
• Molecule 11: Proteasome subunit beta type-5

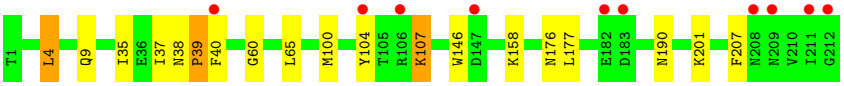
Chain K:



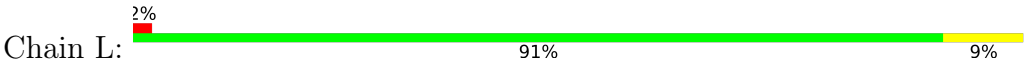


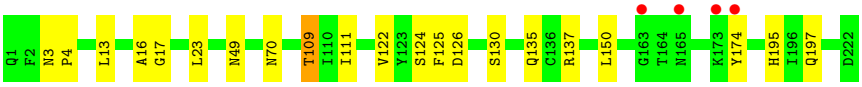
• Molecule 11: Proteasome subunit beta type-5



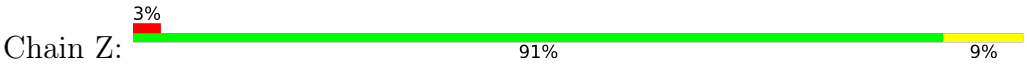


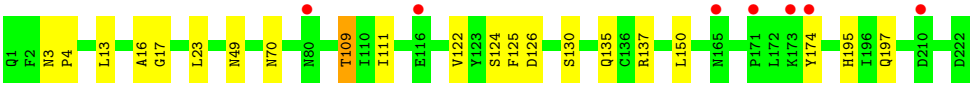
• Molecule 12: Proteasome subunit beta type-6



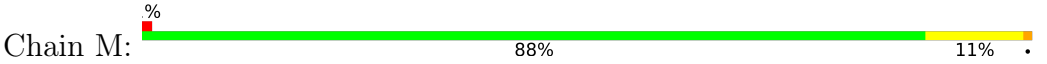


• Molecule 12: Proteasome subunit beta type-6

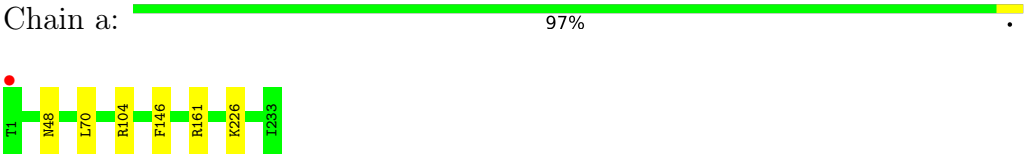




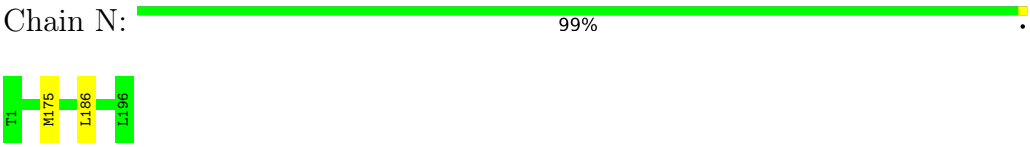
• Molecule 13: Proteasome subunit beta type-7



• Molecule 13: Proteasome subunit beta type-7

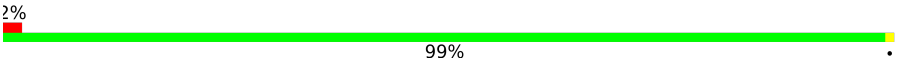


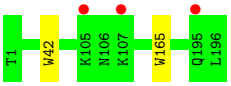
• Molecule 14: Proteasome subunit beta type-1



• Molecule 14: Proteasome subunit beta type-1

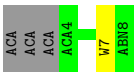
Chain b:





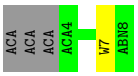
• Molecule 15: TMC-95A mimic ligand yCP:4e

Chain c:



• Molecule 15: TMC-95A mimic ligand yCP:4e

Chain d:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	135.77Å 300.22Å 144.26Å 90.00° 112.86° 90.00°	Depositor
Resolution (Å)	15.00 – 2.80 15.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.1 (15.00-2.80) 99.3 (15.00-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.37 (at 2.81Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.235 , 0.241 0.241 , 0.248	Depositor DCC
R_{free} test set	12835 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	57.1	Xtriage
Anisotropy	0.218	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 45.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	51011	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.64% 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: RE0, ACA, ABN, TY5, 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	0.37	0/1952	0.47	0/2642
1	O	0.37	1/1952 (0.1%)	0.47	0/2642
2	B	0.33	0/1934	0.46	0/2618
2	P	0.34	0/1934	0.47	0/2618
3	C	0.34	0/1919	0.48	0/2598
3	Q	0.34	0/1919	0.48	0/2598
4	D	0.36	0/1886	0.49	0/2541
4	R	0.36	0/1886	0.49	0/2541
5	E	0.31	0/1823	0.46	0/2463
5	S	0.31	0/1823	0.46	0/2463
6	F	0.41	0/1936	0.45	0/2614
6	T	0.41	0/1936	0.45	0/2614
7	G	0.35	0/1959	0.46	0/2652
7	U	0.34	0/1959	0.46	0/2652
8	H	0.44	1/1715 (0.1%)	0.47	0/2326
8	V	0.44	1/1715 (0.1%)	0.47	0/2326
9	I	0.34	0/1611	0.47	0/2174
9	W	0.34	0/1611	0.47	0/2174
10	J	0.31	0/1613	0.46	0/2173
10	X	0.31	0/1613	0.46	0/2173
11	K	0.50	1/1681 (0.1%)	0.50	1/2274 (0.0%)
11	Y	0.50	1/1681 (0.1%)	0.50	1/2274 (0.0%)
12	L	0.36	0/1795	0.46	0/2420
12	Z	0.36	0/1795	0.46	0/2420
13	M	0.36	0/1855	0.48	0/2514
13	a	0.36	0/1855	0.48	0/2514
14	N	0.39	0/1541	0.45	0/2087
14	b	0.39	2/1541 (0.1%)	0.45	0/2087
15	c	0.79	0/4	0.50	0/4
15	d	0.80	0/4	0.46	0/4
All	All	0.37	7/50448 (0.0%)	0.47	2/68200 (0.0%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	Y	146	TRP	CD2-CE2	5.07	1.47	1.41
11	K	146	TRP	CD2-CE2	5.06	1.47	1.41
1	O	179	TRP	CD2-CE2	5.03	1.47	1.41
8	H	164	TRP	CD2-CE2	5.03	1.47	1.41
14	b	42	TRP	CD2-CE2	5.01	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
11	Y	4	LEU	CA-CB-CG	5.34	127.59	115.30
11	K	4	LEU	CA-CB-CG	5.33	127.56	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	1929	10	0
1	O	1915	0	1929	10	0
2	B	1904	0	1904	16	0
2	P	1904	0	1904	16	0
3	C	1890	0	1903	19	0
3	Q	1890	0	1903	14	0
4	D	1861	0	1839	17	0
4	R	1861	0	1839	12	0
5	E	1795	0	1800	17	0
5	S	1795	0	1800	15	0
6	F	1896	0	1889	11	0
6	T	1896	0	1889	11	0
7	G	1921	0	1913	8	0
7	U	1921	0	1913	9	0
8	H	1684	0	1688	7	0
8	V	1684	0	1688	7	0
9	I	1581	0	1574	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	W	1581	0	1574	10	0
10	J	1585	0	1590	16	0
10	X	1585	0	1590	14	0
11	K	1644	0	1595	8	0
11	Y	1644	0	1595	7	0
12	L	1757	0	1711	13	0
12	Z	1757	0	1711	14	0
13	M	1824	0	1832	17	0
13	a	1824	0	1832	0	0
14	N	1512	0	1481	1	0
14	b	1512	0	1481	0	0
15	c	56	0	42	0	0
15	d	56	0	42	0	0
16	K	12	0	13	0	0
16	Y	12	0	13	0	0
17	A	59	0	0	0	0
17	B	39	0	0	0	0
17	C	43	0	0	0	0
17	D	36	0	0	0	0
17	E	21	0	0	0	0
17	F	47	0	0	0	0
17	G	60	0	0	0	0
17	H	52	0	0	0	0
17	I	64	0	0	0	0
17	J	53	0	0	2	0
17	K	48	0	0	0	0
17	L	54	0	0	0	0
17	M	81	0	0	0	0
17	N	58	0	0	0	0
17	O	34	0	0	0	0
17	P	30	0	0	0	0
17	Q	29	0	0	0	0
17	R	27	0	0	0	0
17	S	18	0	0	0	0
17	T	43	0	0	0	0
17	U	55	0	0	0	0
17	V	50	0	0	0	0
17	W	62	0	0	0	0
17	X	41	0	0	0	0
17	Y	50	0	0	0	0
17	Z	49	0	0	0	0
17	a	78	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	b	56	0	0	0	0
All	All	51011	0	49406	264	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 264 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)
3:C:9:PHE:H	4:D:15:GLN:HE22	1.33	0.77
6:T:91:GLU:HG2	6:T:111:ARG:HB3	1.67	0.76
3:Q:160:GLN:HA	3:Q:160:GLN:HE21	1.51	0.76
3:C:160:GLN:HE21	3:C:160:GLN:HA	1.51	0.75
6:F:91:GLU:HG2	6:F:111:ARG:HB3	1.67	0.74

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	248/250 (99%)	240 (97%)	6 (2%)	2 (1%)	19	49
1	O	248/250 (99%)	240 (97%)	6 (2%)	2 (1%)	19	49
2	B	242/258 (94%)	235 (97%)	5 (2%)	2 (1%)	19	49
2	P	242/258 (94%)	235 (97%)	5 (2%)	2 (1%)	19	49
3	C	239/254 (94%)	233 (98%)	3 (1%)	3 (1%)	12	36
3	Q	239/254 (94%)	233 (98%)	3 (1%)	3 (1%)	12	36
4	D	240/260 (92%)	235 (98%)	2 (1%)	3 (1%)	12	36
4	R	240/260 (92%)	235 (98%)	2 (1%)	3 (1%)	12	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	E	231/234 (99%)	224 (97%)	6 (3%)	1 (0%)	34	66
5	S	231/234 (99%)	224 (97%)	6 (3%)	1 (0%)	34	66
6	F	242/288 (84%)	233 (96%)	9 (4%)	0	100	100
6	T	242/288 (84%)	233 (96%)	9 (4%)	0	100	100
7	G	241/252 (96%)	236 (98%)	5 (2%)	0	100	100
7	U	241/252 (96%)	236 (98%)	5 (2%)	0	100	100
8	H	220/232 (95%)	214 (97%)	6 (3%)	0	100	100
8	V	220/232 (95%)	214 (97%)	6 (3%)	0	100	100
9	I	202/205 (98%)	194 (96%)	7 (4%)	1 (0%)	29	61
9	W	202/205 (98%)	194 (96%)	7 (4%)	1 (0%)	29	61
10	J	196/198 (99%)	189 (96%)	6 (3%)	1 (0%)	29	61
10	X	196/198 (99%)	189 (96%)	6 (3%)	1 (0%)	29	61
11	K	210/212 (99%)	202 (96%)	7 (3%)	1 (0%)	29	61
11	Y	210/212 (99%)	202 (96%)	7 (3%)	1 (0%)	29	61
12	L	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
12	Z	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
13	M	231/233 (99%)	221 (96%)	10 (4%)	0	100	100
13	a	231/233 (99%)	221 (96%)	10 (4%)	0	100	100
14	N	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
14	b	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
15	c	1/8 (12%)	1 (100%)	0	0	100	100
15	d	1/8 (12%)	1 (100%)	0	0	100	100
All	All	6314/6604 (96%)	6120 (97%)	166 (3%)	28 (0%)	34	66

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	52	LEU
3	Q	52	LEU
1	A	2	THR
1	A	166	LYS
4	D	122	GLU

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%)	207 (99%)	2 (1%)	76	93
1	O	209/209 (100%)	207 (99%)	2 (1%)	76	93
2	B	203/216 (94%)	193 (95%)	10 (5%)	25	57
2	P	203/216 (94%)	193 (95%)	10 (5%)	25	57
3	C	213/226 (94%)	206 (97%)	7 (3%)	38	72
3	Q	213/226 (94%)	206 (97%)	7 (3%)	38	72
4	D	198/215 (92%)	193 (98%)	5 (2%)	47	80
4	R	198/215 (92%)	192 (97%)	6 (3%)	41	75
5	E	192/193 (100%)	183 (95%)	9 (5%)	26	59
5	S	192/193 (100%)	183 (95%)	9 (5%)	26	59
6	F	201/239 (84%)	190 (94%)	11 (6%)	21	52
6	T	201/239 (84%)	190 (94%)	11 (6%)	21	52
7	G	207/210 (99%)	202 (98%)	5 (2%)	49	81
7	U	207/210 (99%)	202 (98%)	5 (2%)	49	81
8	H	181/190 (95%)	178 (98%)	3 (2%)	60	87
8	V	181/190 (95%)	178 (98%)	3 (2%)	60	87
9	I	172/173 (99%)	169 (98%)	3 (2%)	60	87
9	W	172/173 (99%)	169 (98%)	3 (2%)	60	87
10	J	175/175 (100%)	175 (100%)	0	100	100
10	X	175/175 (100%)	175 (100%)	0	100	100
11	K	169/169 (100%)	162 (96%)	7 (4%)	30	64
11	Y	169/169 (100%)	162 (96%)	7 (4%)	30	64
12	L	185/185 (100%)	182 (98%)	3 (2%)	62	88
12	Z	185/185 (100%)	182 (98%)	3 (2%)	62	88
13	M	199/199 (100%)	193 (97%)	6 (3%)	41	75
13	a	199/199 (100%)	193 (97%)	6 (3%)	41	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
14	N	162/162 (100%)	162 (100%)	0	100 100
14	b	162/162 (100%)	162 (100%)	0	100 100
All	All	5332/5522 (97%)	5189 (97%)	143 (3%)	44 78

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	T	214	TRP
7	U	221	LYS
11	Y	65	LEU
7	G	235	ARG
7	G	221	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 148 such sidechains are listed below:

Mol	Chain	Res	Type
7	U	167	GLN
13	a	108	ASN
8	V	30	ASN
11	Y	9	GLN
8	H	144	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

6 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
15	ACA	d	4	15	7,7,8	0.70	0	6,6,8	0.59	0
15	TY5	d	5	15	19,20,21	1.01	0	22,25,27	0.49	0
15	RE0	d	7	15	15,17,18	1.19	1 (6%)	19,25,27	2.02	5 (26%)
15	RE0	c	7	15	15,17,18	1.25	1 (6%)	19,25,27	2.03	5 (26%)
15	ACA	c	4	15	7,7,8	0.72	0	6,6,8	0.53	0
15	TY5	c	5	15	19,20,21	1.02	0	22,25,27	0.55	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
15	ACA	d	4	15	-	3/4/5/6	-
15	TY5	d	5	15	-	5/10/11/13	0/2/2/2
15	RE0	d	7	15	-	0/6/23/25	0/2/2/2
15	RE0	c	7	15	-	0/6/23/25	0/2/2/2
15	ACA	c	4	15	-	3/4/5/6	-
15	TY5	c	5	15	-	4/10/11/13	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	c	7	RE0	CG-CD2	3.28	1.54	1.51
15	d	7	RE0	CG-CD2	3.05	1.54	1.51

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	c	7	RE0	CG-CD2-CE2	-4.37	107.11	108.86
15	d	7	RE0	CG-CD2-CE2	-4.23	107.16	108.86
15	d	7	RE0	CE2-NE1-CD1	-3.60	109.70	111.86
15	c	7	RE0	CE2-NE1-CD1	-3.52	109.75	111.86
15	c	7	RE0	CD2-CE2-NE1	3.42	111.89	109.59

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	c	4	ACA	C-C2-C3-C4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
15	c	5	TY5	CE2-CZ-OH-C49
15	d	5	TY5	CE1-CZ-OH-C49
15	c	5	TY5	CE1-CZ-OH-C49
15	d	5	TY5	CE2-CZ-OH-C49

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
16	MES	Y	301	-	12,12,12	1.37	1 (8%)	14,16,16	2.01	5 (35%)
16	MES	K	301	-	12,12,12	1.29	1 (8%)	14,16,16	1.65	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	MES	Y	301	-	-	4/6/14/14	0/1/1/1
16	MES	K	301	-	-	5/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	Y	301	MES	C8-S	3.00	1.81	1.77
16	K	301	MES	C8-S	2.89	1.81	1.77

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
16	Y	301	MES	O2S-S-C8	-4.00	102.10	106.92
16	K	301	MES	O2S-S-C8	-3.58	102.60	106.92
16	K	301	MES	O3S-S-O1S	3.00	118.62	111.27
16	Y	301	MES	O3S-S-O1S	3.00	118.60	111.27
16	Y	301	MES	C6-C5-N4	2.54	113.96	110.10

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

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

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^{th} 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%)	-0.07	8 (3%)	47	37	58, 71, 92, 109	0
1	O	250/250 (100%)	-0.03	9 (3%)	42	32	61, 77, 101, 118	0
2	B	244/258 (94%)	0.21	17 (6%)	16	9	58, 75, 111, 120	0
2	P	244/258 (94%)	0.25	19 (7%)	13	7	64, 79, 109, 124	0
3	C	241/254 (94%)	0.13	12 (4%)	28	19	56, 75, 112, 143	0
3	Q	241/254 (94%)	0.40	26 (10%)	5	3	69, 92, 139, 169	0
4	D	242/260 (93%)	0.12	12 (4%)	28	19	61, 76, 103, 120	0
4	R	242/260 (93%)	0.21	17 (7%)	16	9	65, 84, 114, 129	0
5	E	233/234 (99%)	0.06	13 (5%)	24	16	64, 79, 99, 112	0
5	S	233/234 (99%)	0.17	15 (6%)	19	12	65, 86, 112, 124	0
6	F	244/288 (84%)	-0.02	11 (4%)	33	23	59, 74, 102, 125	0
6	T	244/288 (84%)	0.05	11 (4%)	33	23	61, 78, 112, 134	0
7	G	243/252 (96%)	-0.02	8 (3%)	46	36	56, 73, 96, 130	0
7	U	243/252 (96%)	0.01	6 (2%)	57	47	60, 72, 91, 116	0
8	H	222/232 (95%)	-0.14	3 (1%)	75	70	56, 67, 81, 98	0
8	V	222/232 (95%)	-0.20	3 (1%)	75	70	55, 66, 81, 105	0
9	I	204/205 (99%)	-0.37	2 (0%)	82	77	53, 63, 79, 83	0
9	W	204/205 (99%)	-0.25	2 (0%)	82	77	59, 66, 82, 91	0
10	J	198/198 (100%)	-0.08	7 (3%)	44	34	54, 66, 83, 116	0
10	X	198/198 (100%)	-0.07	6 (3%)	50	40	59, 68, 83, 116	0
11	K	212/212 (100%)	-0.05	11 (5%)	27	18	52, 66, 85, 91	0
11	Y	212/212 (100%)	-0.07	10 (4%)	31	22	58, 68, 88, 96	0
12	L	222/222 (100%)	-0.19	4 (1%)	68	61	55, 65, 89, 96	0
12	Z	222/222 (100%)	-0.17	7 (3%)	47	37	56, 66, 88, 95	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
13	M	233/233 (100%)	-0.29	2 (0%) 84 80	53, 64, 77, 80	0
13	a	233/233 (100%)	-0.29	1 (0%) 92 91	53, 65, 76, 80	0
14	N	196/196 (100%)	-0.32	0 100 100	55, 62, 78, 86	0
14	b	196/196 (100%)	-0.30	3 (1%) 73 68	54, 62, 77, 85	0
15	c	1/8 (12%)	-0.23	0 100 100	58, 58, 58, 58	0
15	d	1/8 (12%)	-0.02	0 100 100	56, 56, 56, 56	0
All	All	6370/6604 (96%)	-0.04	245 (3%) 40 30	52, 71, 104, 169	0

The worst 5 of 245 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	R	119	ALA	10.6
4	D	119	ALA	10.0
2	P	219	ALA	9.3
4	D	120	SER	9.3
2	B	220	ASN	8.9

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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	ACA	d	4	8/9	0.61	0.45	61,63,64,64	0
15	ACA	c	4	8/9	0.78	0.47	63,64,65,65	0
15	TY5	c	5	19/20	0.87	0.29	60,62,63,63	0
15	TY5	d	5	19/20	0.88	0.26	58,60,61,61	0
15	RE0	c	7	16/17	0.89	0.20	56,57,58,58	0
15	RE0	d	7	16/17	0.91	0.20	56,57,57,58	0

6.3 Carbohydrates (i)

There are no monosaccharides 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, 95^{th} 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
16	MES	Y	301	12/12	0.87	0.33	60,64,66,66	0
16	MES	K	301	12/12	0.91	0.27	61,62,65,65	0

6.5 Other polymers

There are no such residues in this entry.