wwPDB X-ray Structure Validation Summary Report (i)

Apr 24, 2024 – 05:35 pm BST

PDB ID	:	4Y82
Title	:	Yeast 20S proteasome in complex with Ac-LAY-ep
Authors	:	Huber, E.M.; Groll, M.
Deposited on	:	2015-02-16
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 (i)) were used in the production of this report:

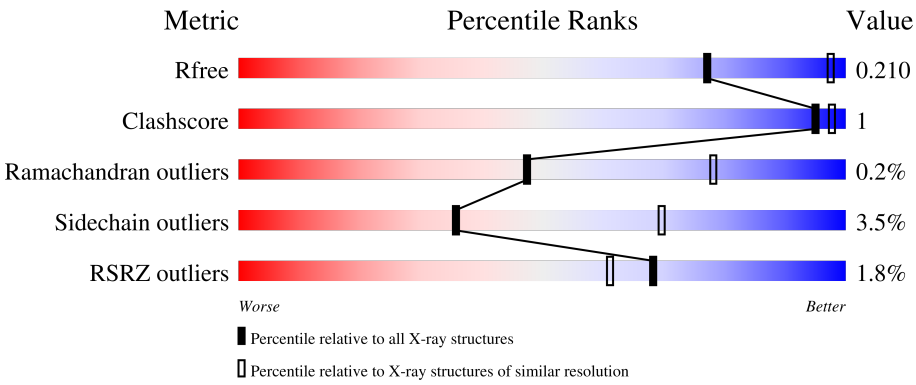
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	2%, 97%, •
1	O	250	2%, 98%, •
2	B	258	4%, 88%, 6%, 5%
2	P	258	4%, 89%, 5%, 5%
3	C	254	5%, 88%, 6%, • , 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	6% 88% 6% 6%
4	D	260	86% 10%
4	R	260	85% 5% 10%
5	E	234	92% 6%
5	S	234	92% 6%
6	F	288	80% 16%
6	T	288	80% 16%
7	G	252	91% 5%
7	U	252	90% 6%
8	H	232	89% 6%
8	V	232	90% 5%
9	I	205	92% 8%
9	W	205	93% 6%
10	J	198	91% 6%
10	X	198	92% 5%
11	K	212	96%
11	Y	212	96%
12	L	222	94% 5%
12	Z	222	95% 5%
13	M	246	88% 5% 7%
13	a	246	91% 6%
14	N	196	96%
14	b	196	99%
15	c	5	80% 20%
15	d	5	80% 20%

Continued on next page...

Continued from previous page...

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

2 Entry composition (i)

There are 19 unique types of molecules in this entry. The entry contains 49969 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Proteasome subunit alpha type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	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 subunit alpha type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	244	1904	1201	321	379	3	0	0	0
2	P	244	1904	1201	321	379	3	0	0	0

- Molecule 3 is a protein called Proteasome subunit alpha type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	240	1881	1176	329	372	4	0	0	0
3	Q	240	1881	1176	329	372	4	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	235	1813	1136	304	366	7	0	0	0
4	R	235	1813	1136	304	366	7	0	0	0

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	222	1684	1061	293	323	7	0	0	0
8	V	222	1684	1061	293	323	7	0	0	0

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

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 subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	195	1561	992	264	299	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	195	1561	992	264	299	6	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	212	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 subunit beta type-6.

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	230	1797	1137	307	346	7	0	0	0
13	a	232	1815	1148	311	349	7	0	0	0

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

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

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
15	c	5	32	23	3	6	0	0	0
15	d	5	32	23	3	6	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
15	e	5	32	23	3	6	0	0	0
15	f	5	32	23	3	6	0	0	0
15	g	5	32	23	3	6	0	0	0
15	h	5	32	23	3	6	0	0	0

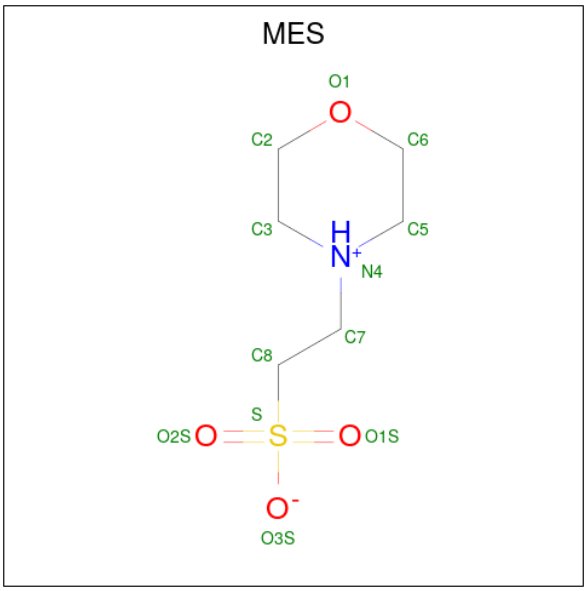
- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

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

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

- Molecule 18 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		
18	K	1	12	6	1	4	1	0	0
			Total	C	N	O	S		
18	Y	1	12	6	1	4	1	0	0

- Molecule 19 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
19	A	20	20	20	0	0
			Total	O		
19	B	17	17	17	0	0
			Total	O		
19	C	20	20	20	0	0
			Total	O		
19	D	5	5	5	0	0
			Total	O		
19	E	8	8	8	0	0
			Total	O		
19	F	19	19	19	0	0
			Total	O		
19	G	18	18	18	0	0
			Total	O		
19	H	17	17	17	0	0
			Total	O		
19	I	20	20	20	0	0
			Total	O		
19	J	21	21	21	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
19	K	22	Total O 22 22	0	0
19	L	24	Total O 24 24	0	0
19	M	21	Total O 21 21	0	0
19	N	10	Total O 10 10	0	0
19	O	17	Total O 17 17	0	0
19	P	11	Total O 11 11	0	0
19	Q	14	Total O 14 14	0	0
19	R	10	Total O 10 10	0	0
19	S	7	Total O 7 7	0	0
19	T	14	Total O 14 14	0	0
19	U	11	Total O 11 11	0	0
19	V	14	Total O 14 14	0	0
19	W	14	Total O 14 14	0	0
19	X	22	Total O 22 22	0	0
19	Y	21	Total O 21 21	0	0
19	Z	25	Total O 25 25	0	0
19	a	28	Total O 28 28	0	0
19	b	19	Total O 19 19	0	0
19	c	1	Total O 1 1	0	0
19	d	1	Total O 1 1	0	0
19	e	2	Total O 2 2	0	0

Continued on next page...

Continued from previous page...

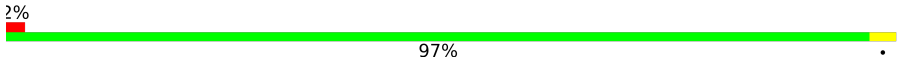
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
19	f	1	1	1	0	0
			Total	O		
19	g	1	1	1	0	0
			Total	O		
19	h	2	2	2	0	0

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.

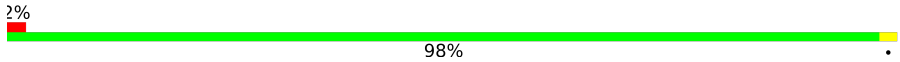
• Molecule 1: Proteasome subunit alpha type-2

Chain A:



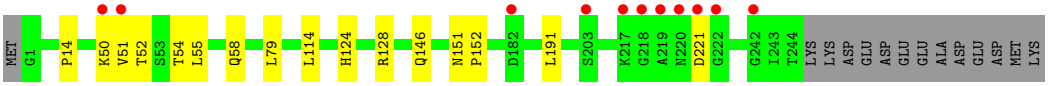
• Molecule 1: Proteasome subunit alpha type-2

Chain O:



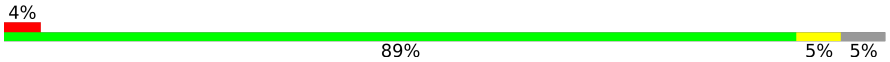
• Molecule 2: Proteasome subunit alpha type-3

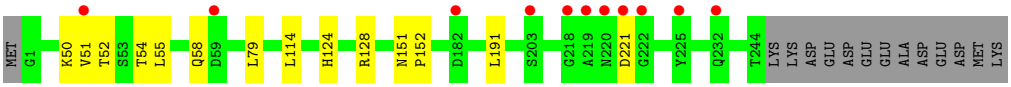
Chain B:



• Molecule 2: Proteasome subunit alpha type-3

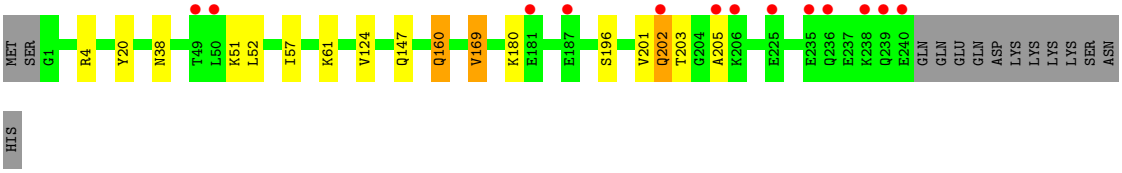
Chain P:





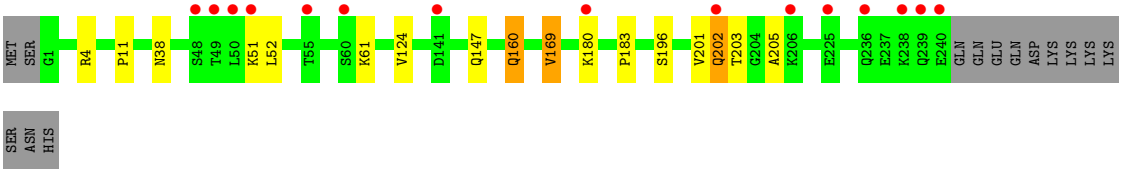
• Molecule 3: Proteasome subunit alpha type-4

Chain C:



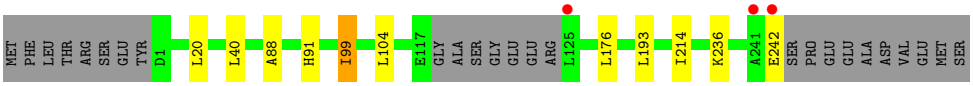
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



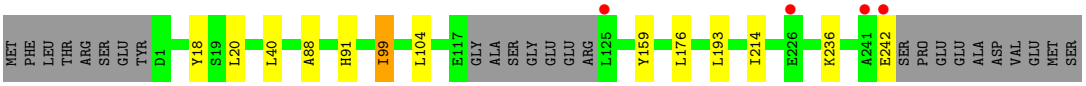
• Molecule 4: Proteasome subunit alpha type-5

Chain D:



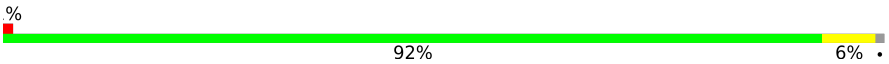
• Molecule 4: Proteasome subunit alpha type-5

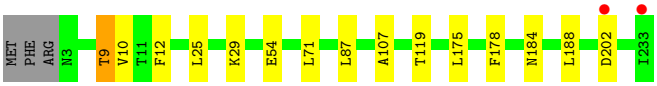
Chain R:



• Molecule 5: Proteasome subunit alpha type-6

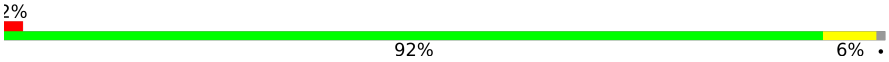
Chain E:

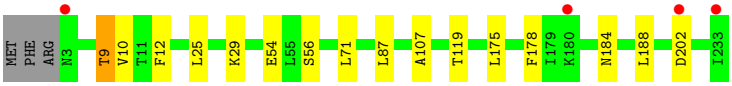




• Molecule 5: Proteasome subunit alpha type-6

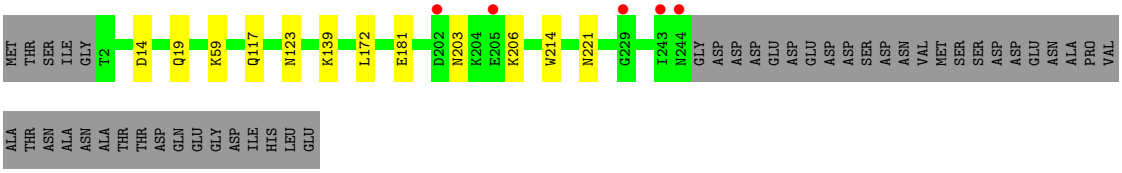
Chain S:





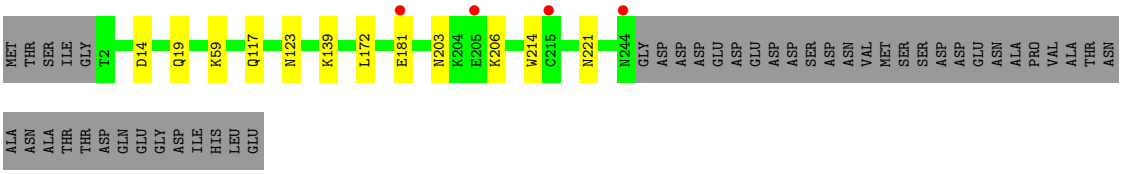
• Molecule 6: Probable proteasome subunit alpha type-7

Chain F:



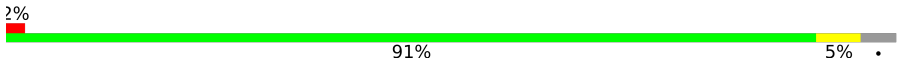
• Molecule 6: Probable proteasome subunit alpha type-7

Chain T:



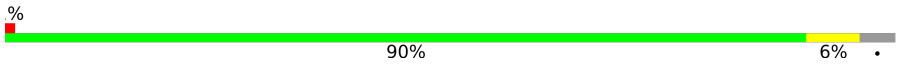
• Molecule 7: Proteasome subunit alpha type-1

Chain G:



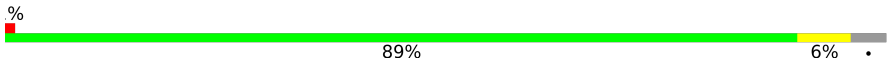
• Molecule 7: Proteasome subunit alpha type-1

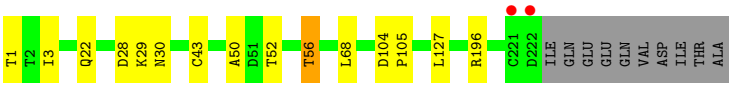
Chain U:



• Molecule 8: Proteasome subunit beta type-2

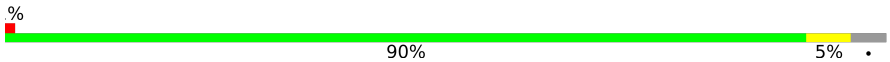
Chain H:

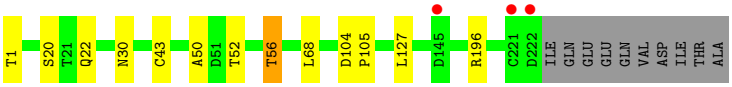




• Molecule 8: Proteasome subunit beta type-2

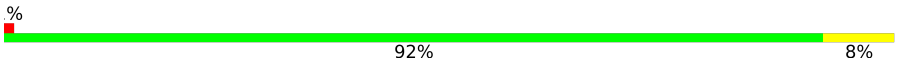
Chain V:

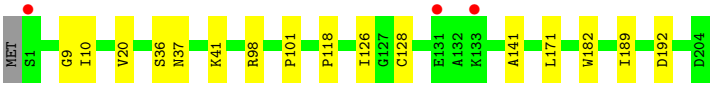




• Molecule 9: Proteasome subunit beta type-3

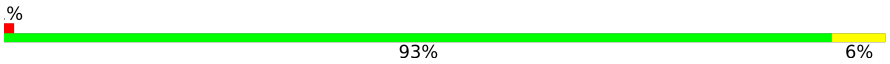
Chain I:

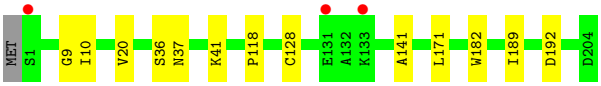




• Molecule 9: Proteasome subunit beta type-3

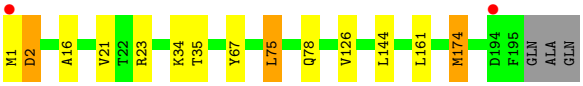
Chain W:





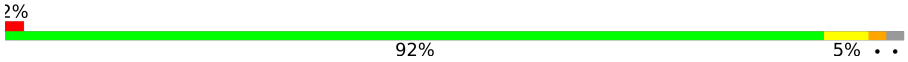
• Molecule 10: Proteasome subunit beta type-4

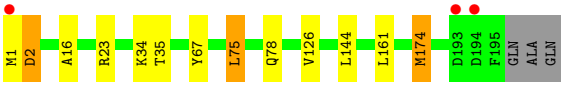
Chain J:



• Molecule 10: Proteasome subunit beta type-4

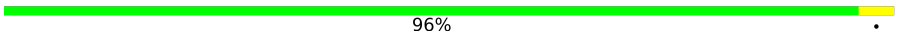
Chain X:

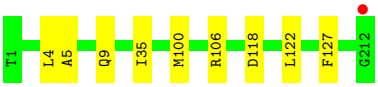




• Molecule 11: Proteasome subunit beta type-5

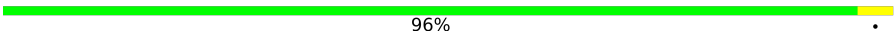
Chain K:





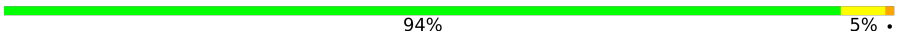
• Molecule 11: Proteasome subunit beta type-5

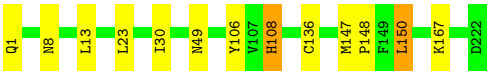
Chain Y:



• Molecule 12: Proteasome subunit beta type-6

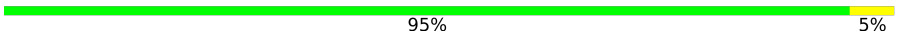
Chain L:





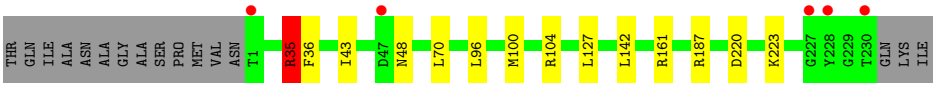
• Molecule 12: Proteasome subunit beta type-6

Chain Z:

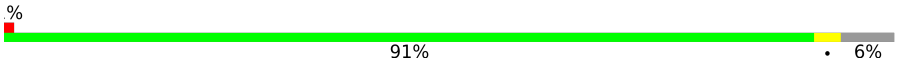


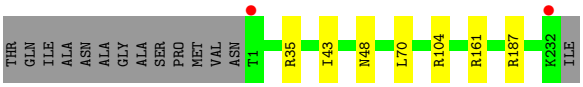
• Molecule 13: Proteasome subunit beta type-7

Chain M:

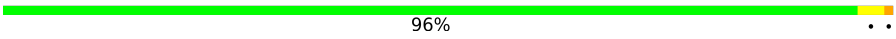


• Molecule 13: Proteasome subunit beta type-7

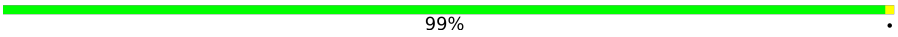
Chain a: 



• Molecule 14: Proteasome subunit beta type-1

Chain N: 

• Molecule 14: Proteasome subunit beta type-1

Chain b: 

• Molecule 15: Ac-LAY-ep

Chain c:

• Molecule 15: Ac-LAY-ep

Chain d:

• Molecule 15: Ac-LAY-ep

Chain e:

• Molecule 15: Ac-LAY-ep

Chain f:

- Molecule 15: Ac-LAY-ep

Chain g:

- Molecule 15: Ac-LAY-ep

Chain h:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	136.24Å 300.91Å 145.47Å 90.00° 113.09° 90.00°	Depositor
Resolution (Å)	15.00 – 2.80 15.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.2 (15.00-2.80) 98.3 (15.00-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.86 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.189 , 0.207 0.193 , 0.210	Depositor DCC
R_{free} test set	12861 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	58.5	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 42.9	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	0.95	EDS
Total number of atoms	49969	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% 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: TYE, POL, ACE, MES, CL, 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	0.28	0/1952	0.48	0/2642
1	O	0.28	0/1952	0.47	0/2642
2	B	0.29	0/1934	0.49	0/2618
2	P	0.28	0/1934	0.49	0/2618
3	C	0.29	0/1910	0.51	0/2586
3	Q	0.29	0/1910	0.51	0/2586
4	D	0.28	0/1837	0.48	0/2475
4	R	0.28	0/1837	0.48	0/2475
5	E	0.28	0/1800	0.48	0/2433
5	S	0.28	0/1800	0.48	0/2433
6	F	0.29	0/1932	0.46	0/2609
6	T	0.28	0/1932	0.46	0/2609
7	G	0.29	0/1945	0.47	0/2634
7	U	0.29	0/1945	0.47	0/2634
8	H	0.34	0/1715	0.50	0/2326
8	V	0.31	0/1715	0.50	1/2326 (0.0%)
9	I	0.29	0/1611	0.48	0/2174
9	W	0.29	0/1611	0.49	0/2174
10	J	0.28	0/1589	0.49	0/2142
10	X	0.27	0/1589	0.49	0/2142
11	K	0.31	0/1681	0.50	0/2274
11	Y	0.31	0/1681	0.51	0/2274
12	L	0.37	2/1806 (0.1%)	0.57	4/2435 (0.2%)
12	Z	0.32	0/1795	0.48	0/2420
13	M	0.31	0/1828	0.56	2/2480 (0.1%)
13	a	0.31	0/1846	0.56	2/2503 (0.1%)
14	N	0.36	0/1541	0.49	0/2087
14	b	0.37	0/1541	0.49	0/2087
15	c	1.04	0/13	1.38	0/17
15	d	1.46	0/13	1.70	0/17
15	e	1.24	0/13	1.41	0/17
15	f	0.65	0/13	1.05	0/17

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
15	g	1.35	0/13	1.50	0/17
15	h	1.04	0/13	1.41	0/17
All	All	0.30	2/50247 (0.0%)	0.50	9/67940 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	L	108[A]	HIS	CA-C	6.47	1.69	1.52
12	L	108[B]	HIS	CA-C	6.47	1.69	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	a	35	ARG	NE-CZ-NH2	9.30	124.95	120.30
13	M	35	ARG	NE-CZ-NH2	8.84	124.72	120.30
12	L	108[A]	HIS	CA-C-O	6.80	134.39	120.10
12	L	108[B]	HIS	CA-C-O	6.80	134.39	120.10
13	M	35	ARG	NH1-CZ-NH2	-5.79	113.03	119.40

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	3	0
1	O	1915	0	1929	3	0
2	B	1904	0	1904	6	0
2	P	1904	0	1904	4	0
3	C	1881	0	1895	9	0
3	Q	1881	0	1895	8	0
4	D	1813	0	1797	3	0
4	R	1813	0	1797	5	0
5	E	1773	0	1775	4	0
5	S	1773	0	1775	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	F	1892	0	1883	1	0
6	T	1892	0	1883	1	0
7	G	1907	0	1901	3	0
7	U	1907	0	1901	4	0
8	H	1684	0	1685	5	0
8	V	1684	0	1685	3	0
9	I	1581	0	1574	8	0
9	W	1581	0	1574	6	0
10	J	1561	0	1569	8	0
10	X	1561	0	1569	6	0
11	K	1644	0	1592	4	0
11	Y	1644	0	1592	3	0
12	L	1764	0	1718	6	0
12	Z	1757	0	1711	4	0
13	M	1797	0	1800	4	0
13	a	1815	0	1821	0	0
14	N	1512	0	1478	3	0
14	b	1512	0	1478	0	0
15	c	32	0	32	0	0
15	d	32	0	32	0	0
15	e	32	0	32	0	0
15	f	32	0	33	0	0
15	g	32	0	32	0	0
15	h	32	0	32	0	0
16	G	1	0	0	0	0
16	I	2	0	0	0	0
16	K	1	0	0	0	0
16	L	1	0	0	0	0
16	N	1	0	0	0	0
16	Z	1	0	0	0	0
17	G	1	0	0	0	0
17	U	1	0	0	0	0
18	K	12	0	13	0	0
18	Y	12	0	13	0	0
19	A	20	0	0	0	0
19	B	17	0	0	0	0
19	C	20	0	0	0	0
19	D	5	0	0	0	0
19	E	8	0	0	0	0
19	F	19	0	0	0	0
19	G	18	0	0	0	0
19	H	17	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
19	I	20	0	0	0	0
19	J	21	0	0	0	0
19	K	22	0	0	0	0
19	L	24	0	0	3	0
19	M	21	0	0	0	0
19	N	10	0	0	0	0
19	O	17	0	0	0	0
19	P	11	0	0	0	0
19	Q	14	0	0	0	0
19	R	10	0	0	0	0
19	S	7	0	0	0	0
19	T	14	0	0	0	0
19	U	11	0	0	0	0
19	V	14	0	0	0	0
19	W	14	0	0	0	0
19	X	22	0	0	0	0
19	Y	21	0	0	0	0
19	Z	25	0	0	1	0
19	a	28	0	0	0	0
19	b	19	0	0	0	0
19	c	1	0	0	0	0
19	d	1	0	0	0	0
19	e	2	0	0	0	0
19	f	1	0	0	0	0
19	g	1	0	0	0	0
19	h	2	0	0	0	0
All	All	49969	0	49233	100	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

The worst 5 of 100 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)
12:L:108[A]:HIS:ND1	19:L:401:HOH:O	1.64	1.22
12:L:108[A]:HIS:CE1	19:L:401:HOH:O	2.07	1.01
12:Z:106:TYR:OH	19:Z:401:HOH:O	1.97	0.82
10:J:1:MET:O	10:J:2:ASP:HB2	1.84	0.76
10:X:1:MET:O	10:X:2:ASP:HB2	1.85	0.75

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	248/250 (99%)	242 (98%)	6 (2%)	0	100	100
1	O	248/250 (99%)	242 (98%)	6 (2%)	0	100	100
2	B	242/258 (94%)	234 (97%)	6 (2%)	2 (1%)	19	49
2	P	242/258 (94%)	234 (97%)	6 (2%)	2 (1%)	19	49
3	C	238/254 (94%)	231 (97%)	5 (2%)	2 (1%)	19	49
3	Q	238/254 (94%)	231 (97%)	4 (2%)	3 (1%)	12	36
4	D	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
4	R	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
5	E	229/234 (98%)	224 (98%)	5 (2%)	0	100	100
5	S	229/234 (98%)	224 (98%)	5 (2%)	0	100	100
6	F	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
6	T	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
7	G	239/252 (95%)	236 (99%)	3 (1%)	0	100	100
7	U	239/252 (95%)	235 (98%)	4 (2%)	0	100	100
8	H	220/232 (95%)	213 (97%)	7 (3%)	0	100	100
8	V	220/232 (95%)	213 (97%)	7 (3%)	0	100	100
9	I	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
9	W	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	J	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	29	61
10	X	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	29	61
11	K	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
11	Y	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
12	L	221/222 (100%)	217 (98%)	4 (2%)	0	100	100
12	Z	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
13	M	228/246 (93%)	221 (97%)	7 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	a	230/246 (94%)	223 (97%)	7 (3%)	0	100	100
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
15	c	2/5 (40%)	2 (100%)	0	0	100	100
15	d	2/5 (40%)	2 (100%)	0	0	100	100
15	e	2/5 (40%)	2 (100%)	0	0	100	100
15	f	2/5 (40%)	2 (100%)	0	0	100	100
15	g	2/5 (40%)	2 (100%)	0	0	100	100
15	h	2/5 (40%)	2 (100%)	0	0	100	100
All	All	6285/6644 (95%)	6130 (98%)	144 (2%)	11 (0%)	47	78

5 of 11 Ramachandran outliers are listed below:

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

5.3.2 Protein sidechains (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%)	206 (99%)	3 (1%)	67	90
1	O	209/209 (100%)	206 (99%)	3 (1%)	67	90
2	B	203/216 (94%)	195 (96%)	8 (4%)	32	66
2	P	203/216 (94%)	195 (96%)	8 (4%)	32	66
3	C	212/226 (94%)	204 (96%)	8 (4%)	33	67
3	Q	212/226 (94%)	204 (96%)	8 (4%)	33	67
4	D	194/215 (90%)	186 (96%)	8 (4%)	30	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	R	194/215 (90%)	186 (96%)	8 (4%)	30	64
5	E	190/193 (98%)	181 (95%)	9 (5%)	26	59
5	S	190/193 (98%)	181 (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	206/210 (98%)	199 (97%)	7 (3%)	37	71
7	U	206/210 (98%)	198 (96%)	8 (4%)	32	66
8	H	181/190 (95%)	174 (96%)	7 (4%)	32	66
8	V	181/190 (95%)	173 (96%)	8 (4%)	28	61
9	I	172/173 (99%)	168 (98%)	4 (2%)	50	82
9	W	172/173 (99%)	168 (98%)	4 (2%)	50	82
10	J	173/175 (99%)	167 (96%)	6 (4%)	36	70
10	X	173/175 (99%)	167 (96%)	6 (4%)	36	70
11	K	169/169 (100%)	164 (97%)	5 (3%)	41	75
11	Y	169/169 (100%)	164 (97%)	5 (3%)	41	75
12	L	186/185 (100%)	179 (96%)	7 (4%)	33	67
12	Z	185/185 (100%)	179 (97%)	6 (3%)	39	73
13	M	196/208 (94%)	189 (96%)	7 (4%)	35	69
13	a	198/208 (95%)	192 (97%)	6 (3%)	41	75
14	N	162/162 (100%)	159 (98%)	3 (2%)	57	85
14	b	162/162 (100%)	160 (99%)	2 (1%)	71	92
15	c	1/1 (100%)	1 (100%)	0	100	100
15	d	1/1 (100%)	1 (100%)	0	100	100
15	e	1/1 (100%)	1 (100%)	0	100	100
15	f	1/1 (100%)	1 (100%)	0	100	100
15	g	1/1 (100%)	1 (100%)	0	100	100
15	h	1/1 (100%)	1 (100%)	0	100	100
All	All	5315/5546 (96%)	5130 (96%)	185 (4%)	36	70

5 of 185 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	R	176	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
7	U	154	TYR
4	R	242	GLU
6	T	59	LYS
8	V	30	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 92 such sidechains are listed below:

Mol	Chain	Res	Type
4	R	15	GLN
6	T	240	GLN
4	R	225	ASN
5	S	184	ASN
7	U	121	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	TYE	g	4	15,11	12,12,12	1.87	1 (8%)	13,15,15	1.32	2 (15%)
15	TYE	d	4	15,11	12,12,12	1.77	1 (8%)	13,15,15	1.48	1 (7%)
15	TYE	h	4	15,14	12,12,12	1.97	1 (8%)	13,15,15	1.44	3 (23%)
15	TYE	e	4	15,14	12,12,12	1.86	1 (8%)	13,15,15	1.33	1 (7%)
15	TYE	c	4	15,8	12,12,12	1.96	1 (8%)	13,15,15	1.70	4 (30%)
15	TYE	f	4	15,8	12,12,12	1.28	1 (8%)	13,15,15	1.33	3 (23%)

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	TYE	g	4	15,11	-	2/6/6/6	0/1/1/1
15	TYE	d	4	15,11	-	2/6/6/6	0/1/1/1
15	TYE	h	4	15,14	-	2/6/6/6	0/1/1/1
15	TYE	e	4	15,14	-	2/6/6/6	0/1/1/1
15	TYE	c	4	15,8	-	3/6/6/6	0/1/1/1
15	TYE	f	4	15,8	-	3/6/6/6	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)
15	h	4	TYE	CB-CG	-5.86	1.37	1.51
15	c	4	TYE	CB-CG	-5.81	1.37	1.51
15	g	4	TYE	CB-CG	-5.67	1.37	1.51
15	e	4	TYE	CB-CG	-5.56	1.37	1.51
15	d	4	TYE	CB-CG	-5.12	1.39	1.51

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	f	4	TYE	CB-CA-C	-3.25	106.34	112.21
15	e	4	TYE	CB-CA-C	-3.23	106.36	112.21
15	d	4	TYE	CG-CB-CA	3.23	119.40	113.24
15	c	4	TYE	CB-CA-C	-3.04	106.71	112.21
15	h	4	TYE	CB-CA-C	-2.93	106.92	112.21

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	g	4	TYE	CA-CB-CG-CD2
15	g	4	TYE	CA-CB-CG-CD1
15	c	4	TYE	CA-CB-CG-CD1
15	d	4	TYE	CA-CB-CG-CD2
15	c	4	TYE	CA-CB-CG-CD2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 9 are monoatomic - leaving 2 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
18	MES	Y	301	-	12,12,12	2.26	1 (8%)	14,16,16	1.26	2 (14%)
18	MES	K	302	-	12,12,12	2.26	1 (8%)	14,16,16	1.23	2 (14%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
18	MES	Y	301	-	-	0/6/14/14	0/1/1/1
18	MES	K	302	-	-	0/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	K	302	MES	C8-S	-7.57	1.66	1.77
18	Y	301	MES	C8-S	-7.56	1.66	1.77

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	K	302	MES	O3S-S-C8	2.81	110.31	105.77
18	Y	301	MES	O2S-S-C8	2.73	110.20	106.92
18	Y	301	MES	O3S-S-C8	2.68	110.11	105.77
18	K	302	MES	O2S-S-C8	2.17	109.52	106.92

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (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%)	-0.41	5 (2%)	65	56	39, 53, 91, 132	0
1	O	250/250 (100%)	-0.33	6 (2%)	59	49	43, 61, 105, 135	0
2	B	244/258 (94%)	-0.29	11 (4%)	33	23	39, 58, 110, 168	0
2	P	244/258 (94%)	-0.19	11 (4%)	33	23	44, 62, 113, 169	0
3	C	240/254 (94%)	-0.19	13 (5%)	25	17	38, 62, 127, 163	0
3	Q	240/254 (94%)	0.09	15 (6%)	20	12	49, 79, 162, 213	0
4	D	235/260 (90%)	-0.35	3 (1%)	77	72	43, 65, 98, 147	0
4	R	235/260 (90%)	-0.24	4 (1%)	70	63	52, 70, 107, 155	0
5	E	231/234 (98%)	-0.26	2 (0%)	84	80	47, 67, 102, 145	0
5	S	231/234 (98%)	-0.17	4 (1%)	70	63	48, 75, 119, 157	0
6	F	243/288 (84%)	-0.39	5 (2%)	63	54	40, 61, 111, 142	0
6	T	243/288 (84%)	-0.20	4 (1%)	72	66	39, 73, 128, 166	0
7	G	241/252 (95%)	-0.42	4 (1%)	70	63	39, 55, 98, 150	0
7	U	241/252 (95%)	-0.42	3 (1%)	79	73	42, 58, 92, 140	0
8	H	222/232 (95%)	-0.33	2 (0%)	84	80	39, 53, 83, 127	0
8	V	222/232 (95%)	-0.35	3 (1%)	75	70	42, 56, 83, 138	0
9	I	204/205 (99%)	-0.64	3 (1%)	73	68	35, 48, 80, 99	0
9	W	204/205 (99%)	-0.58	3 (1%)	73	68	36, 51, 81, 105	0
10	J	195/198 (98%)	-0.59	2 (1%)	82	77	36, 49, 74, 122	0
10	X	195/198 (98%)	-0.57	3 (1%)	73	68	37, 51, 78, 129	0
11	K	212/212 (100%)	-0.60	1 (0%)	91	88	36, 49, 75, 97	0
11	Y	212/212 (100%)	-0.56	1 (0%)	91	88	37, 50, 78, 103	0
12	L	222/222 (100%)	-0.55	0	100	100	36, 53, 81, 107	0
12	Z	222/222 (100%)	-0.49	1 (0%)	91	88	37, 52, 83, 110	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
13	M	230/246 (93%)	-0.52	5 (2%)	62	52	32, 52, 76, 101	0
13	a	232/246 (94%)	-0.54	2 (0%)	84	80	34, 51, 77, 101	0
14	N	196/196 (100%)	-0.59	0	100	100	36, 48, 78, 105	0
14	b	196/196 (100%)	-0.54	0	100	100	36, 49, 80, 101	0
15	c	2/5 (40%)	-1.04	0	100	100	56, 56, 56, 63	0
15	d	2/5 (40%)	-1.29	0	100	100	43, 43, 43, 50	0
15	e	2/5 (40%)	-0.91	0	100	100	42, 42, 42, 52	0
15	f	2/5 (40%)	0.03	0	100	100	60, 60, 60, 70	0
15	g	2/5 (40%)	-0.91	0	100	100	42, 42, 42, 53	0
15	h	2/5 (40%)	-1.01	0	100	100	42, 42, 42, 57	0
All	All	6344/6644 (95%)	-0.39	116 (1%)	68	61	32, 57, 105, 213	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	220	ASN	7.9
3	Q	49	THR	7.1
2	P	219	ALA	7.1
8	V	222	ASP	6.8
3	C	206	LYS	6.2

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	TYE	c	4	12/12	0.95	0.13	61,65,69,70	0
15	TYE	e	4	12/12	0.96	0.12	42,50,52,53	0
15	TYE	f	4	12/12	0.97	0.12	63,67,69,70	0
15	TYE	h	4	12/12	0.97	0.10	45,48,51,52	0
15	TYE	g	4	12/12	0.98	0.08	45,49,51,51	0
15	TYE	d	4	12/12	0.98	0.10	42,44,44,45	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
18	MES	K	302	12/12	0.92	0.21	51,76,84,86	0
16	MG	N	201	1/1	0.94	0.13	50,50,50,50	0
16	MG	G	301	1/1	0.94	0.07	47,47,47,47	0
18	MES	Y	301	12/12	0.94	0.21	51,74,82,83	0
16	MG	I	301	1/1	0.97	0.19	61,61,61,61	0
17	CL	G	302	1/1	0.97	0.23	30,30,30,30	0
16	MG	K	301	1/1	0.98	0.06	51,51,51,51	0
17	CL	U	301	1/1	0.98	0.22	30,30,30,30	0
16	MG	I	302	1/1	0.99	0.05	43,43,43,43	0
16	MG	Z	301	1/1	0.99	0.31	59,59,59,59	0
16	MG	L	301	1/1	0.99	0.06	51,51,51,51	0

6.5 Other polymers (i)

There are no such residues in this entry.