wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 11:23 pm GMT

PDB ID	:	4Y8T
Title	:	Yeast 20S proteasome beta2-H116D mutant in complex with Ac-PAE-ep
Authors	:	Huber, E.M.; Groll, M.
Deposited on	:	2015-02-16
Resolution	:	2.70 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:

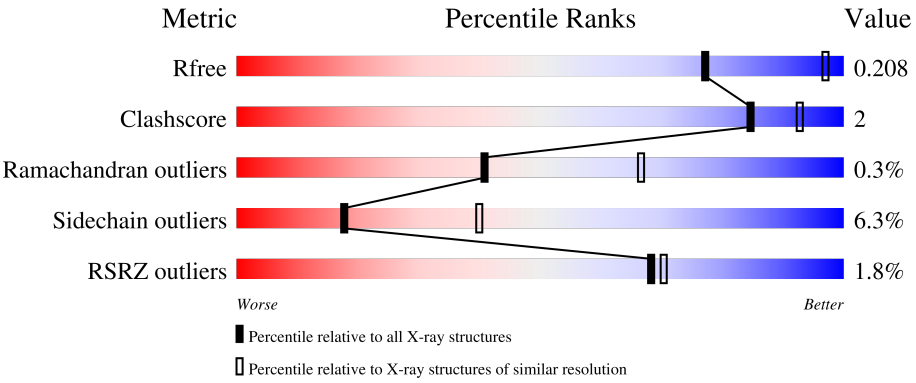
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	94% 5%
1	O	250	3% 94% 5%
2	B	258	3% 80% 13% • 5%
2	P	258	3% 81% 13% • 5%
3	C	254	4% 80% 13% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	6% 81% 12% 6%
4	D	260	77% 12% 10%
4	R	260	77% 12% 10%
5	E	234	83% 15%
5	S	234	84% 14%
6	F	288	76% 8% 16%
6	T	288	76% 8% 16%
7	G	252	85% 10%
7	U	252	86% 9%
8	H	232	89% 7%
8	V	232	89% 7%
9	I	205	92% 7%
9	W	205	92% 7%
10	J	198	88% 10%
10	X	198	88% 9%
11	K	212	90% 8%
11	Y	212	91% 8%
12	L	222	93% 7%
12	Z	222	92% 7%
13	M	246	88% 5% 5%
13	a	246	90% 5%
14	N	196	93% 6%
14	b	196	94% 6%
15	c	5	60% 40%
15	d	5	60% 40%

2 Entry composition (i)

There are 17 unique types of molecules in this entry. The entry contains 50193 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

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

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

- Molecule 2 is a protein called Proteasome subunit alpha type-3.

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

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

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

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

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

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	226	Total	C	N	O	S	0	0	0
			1717	1080	296	334	7			
8	V	226	Total	C	N	O	S	0	0	0
			1717	1080	296	334	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	116	ASP	HIS	engineered mutation	UNP P25043
V	116	ASP	HIS	engineered mutation	UNP P25043

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

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

- Molecule 10 is a protein called Proteasome subunit beta type-4.

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

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

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

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

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
15	c	5	Total C N O 28 18 3 7	0	0	0
15	d	5	Total C N O 28 18 3 7	0	0	0

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

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

- Molecule 17 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	A	30	Total O 30 30	0	0
17	B	26	Total O 26 26	0	0
17	C	34	Total O 34 34	0	0
17	D	11	Total O 11 11	0	0
17	E	15	Total O 15 15	0	0
17	F	21	Total O 21 21	0	0
17	G	36	Total O 36 36	0	0
17	H	57	Total O 57 57	0	0
17	I	29	Total O 29 29	0	0
17	J	32	Total O 32 32	0	0

Continued on next page...

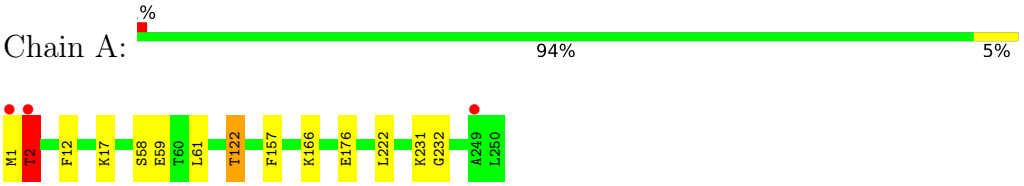
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	K	27	Total O 27 27	0	0
17	L	39	Total O 39 39	0	0
17	M	30	Total O 30 30	0	0
17	N	24	Total O 24 24	0	0
17	O	18	Total O 18 18	0	0
17	P	14	Total O 14 14	0	0
17	Q	15	Total O 15 15	0	0
17	R	19	Total O 19 19	0	0
17	S	10	Total O 10 10	0	0
17	T	24	Total O 24 24	0	0
17	U	35	Total O 35 35	0	0
17	V	36	Total O 36 36	0	0
17	W	23	Total O 23 23	0	0
17	X	24	Total O 24 24	0	0
17	Y	36	Total O 36 36	0	0
17	Z	39	Total O 39 39	0	0
17	a	36	Total O 36 36	0	0
17	b	27	Total O 27 27	0	0
17	c	1	Total O 1 1	0	0

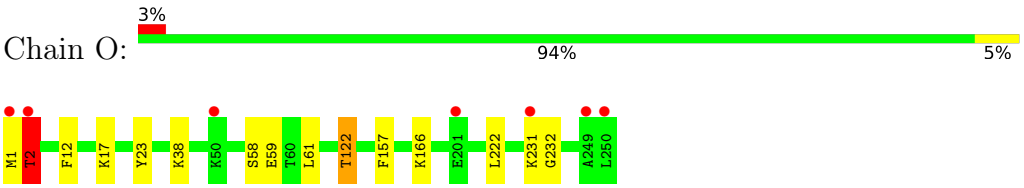
3 Residue-property plots (i)

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.

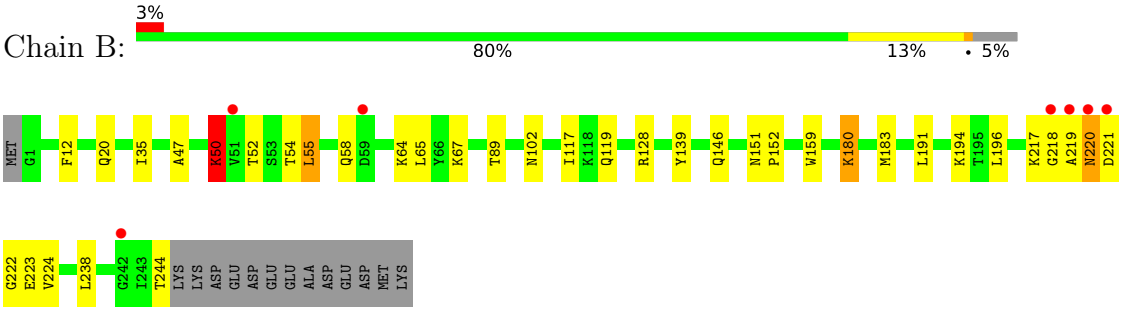
• Molecule 1: Proteasome subunit alpha type-2



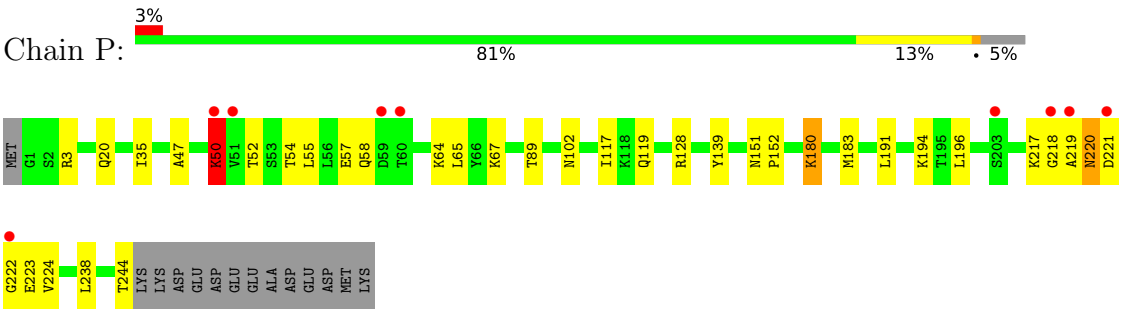
• Molecule 1: Proteasome subunit alpha type-2



• Molecule 2: Proteasome subunit alpha type-3

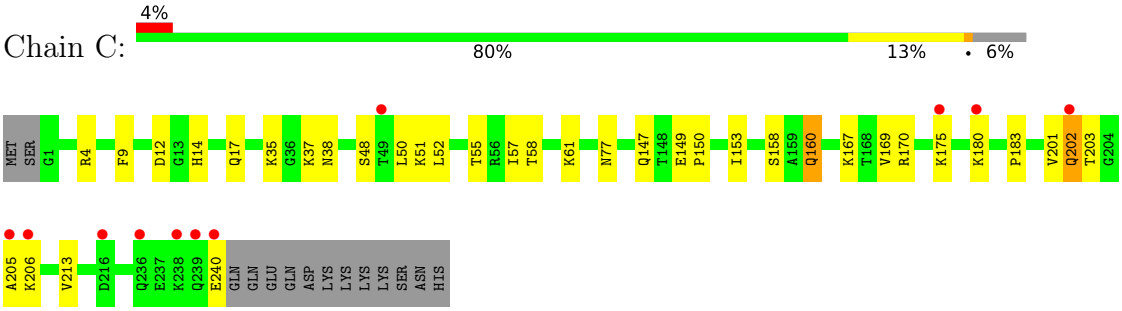


• Molecule 2: Proteasome subunit alpha type-3



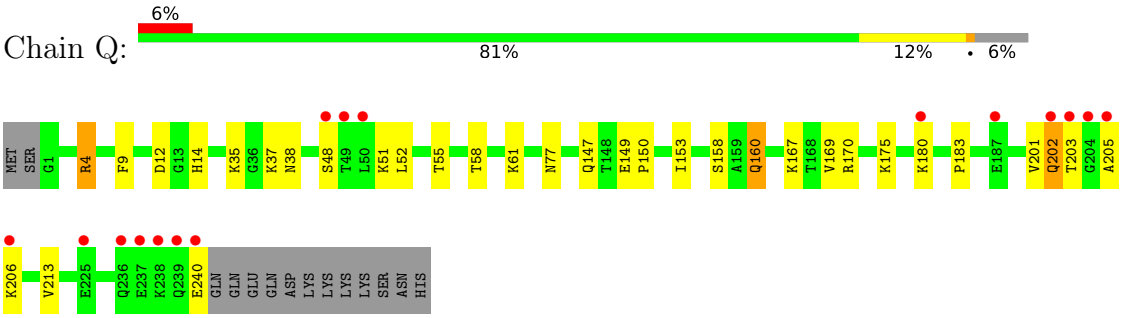
• Molecule 3: Proteasome subunit alpha type-4

Chain C:



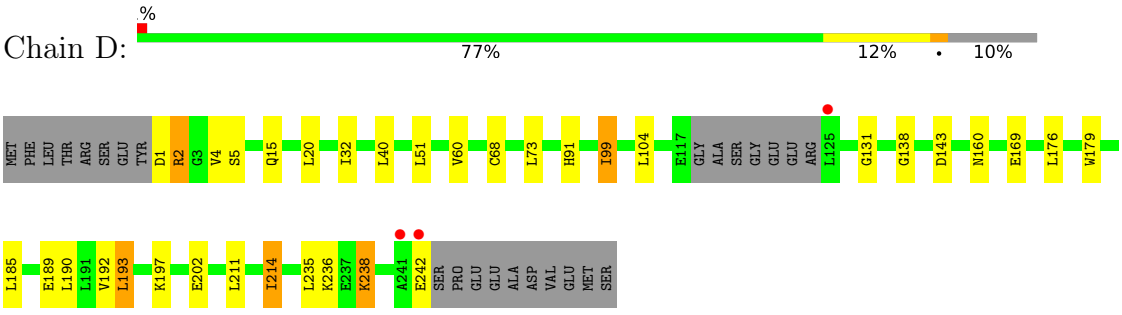
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



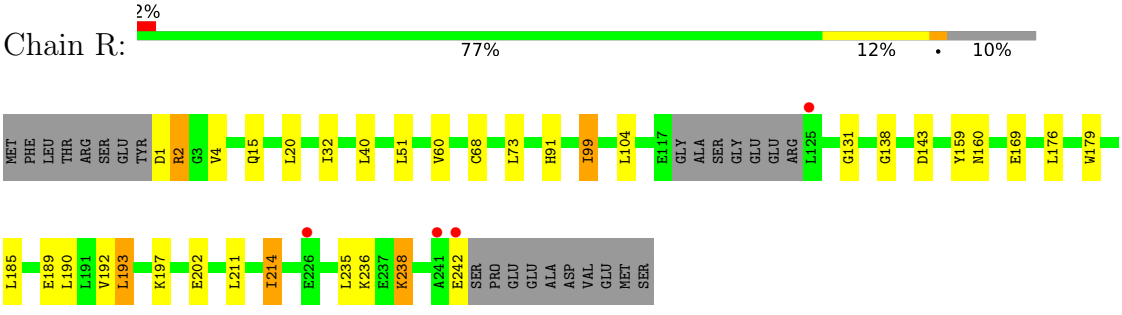
• Molecule 4: Proteasome subunit alpha type-5

Chain D:



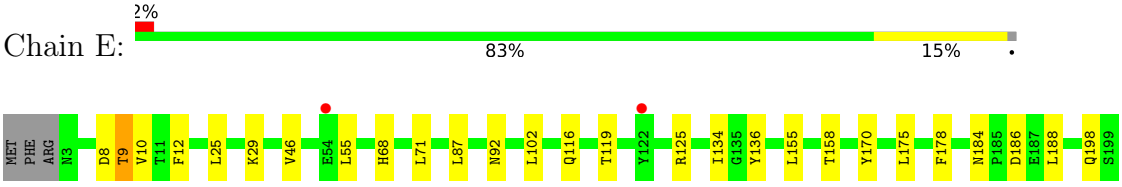
• Molecule 4: Proteasome subunit alpha type-5

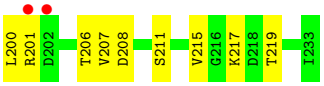
Chain R:



• Molecule 5: Proteasome subunit alpha type-6

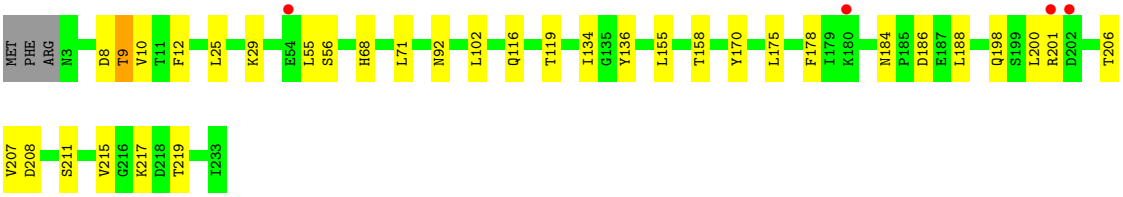
Chain E:





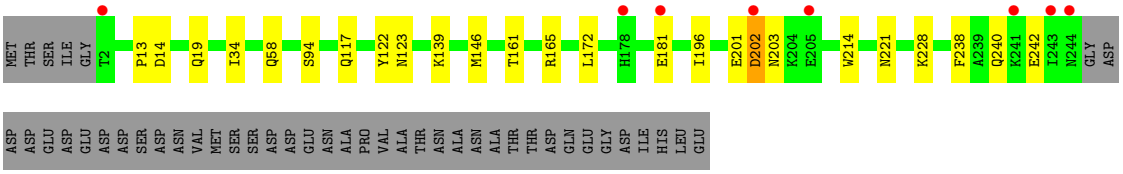
• Molecule 5: Proteasome subunit alpha type-6

Chain S:



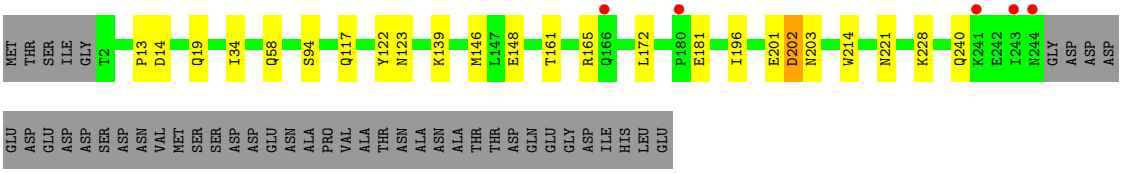
• Molecule 6: Proteasome subunit alpha type-7

Chain F:



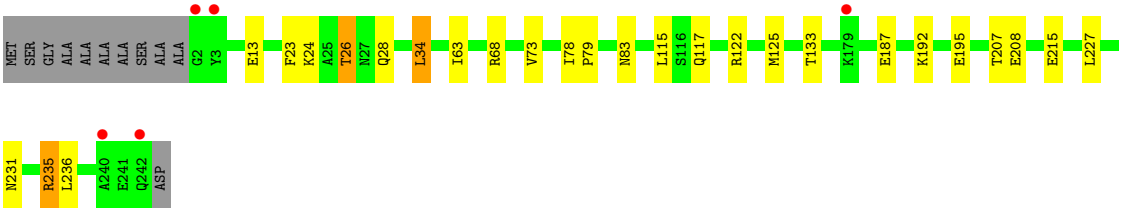
• Molecule 6: Proteasome subunit alpha type-7

Chain T:



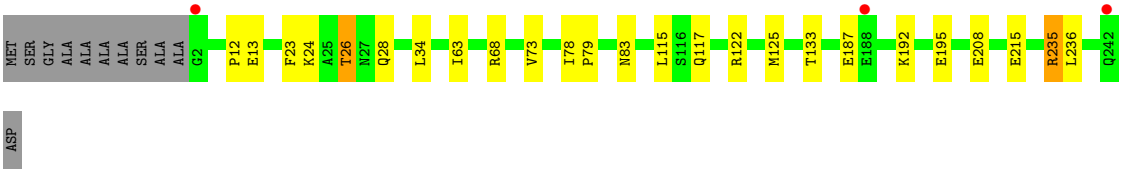
• Molecule 7: Proteasome subunit alpha type-1

Chain G:



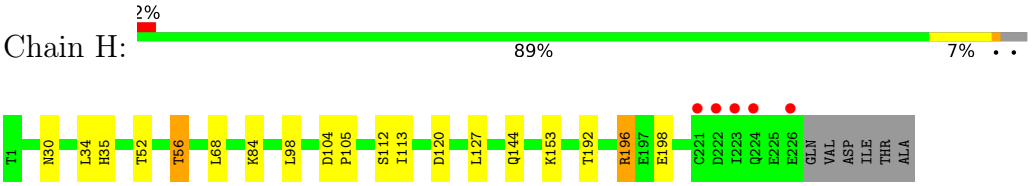
• Molecule 7: Proteasome subunit alpha type-1

Chain U:



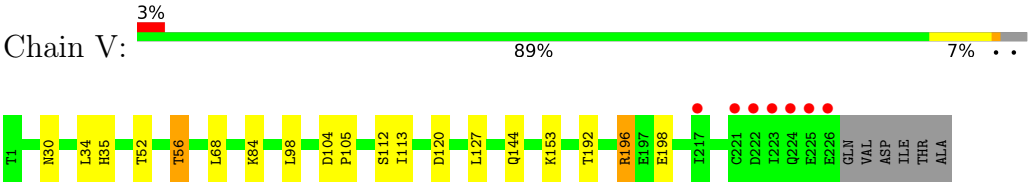
• Molecule 8: Proteasome subunit beta type-2

Chain H:



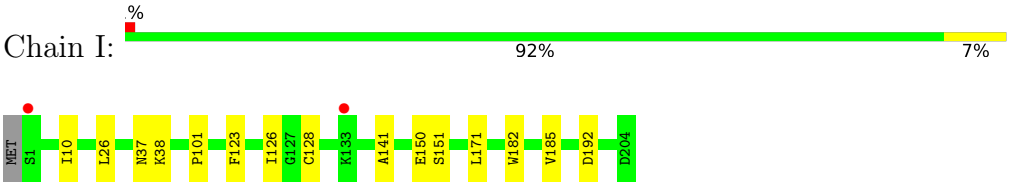
• Molecule 8: Proteasome subunit beta type-2

Chain V:



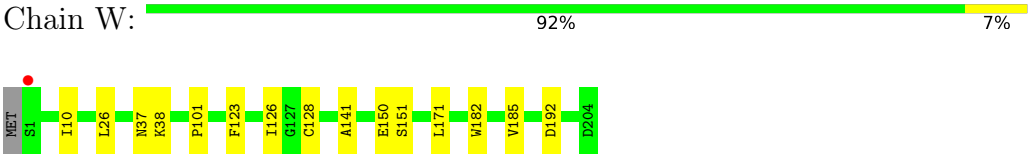
• Molecule 9: Proteasome subunit beta type-3

Chain I:



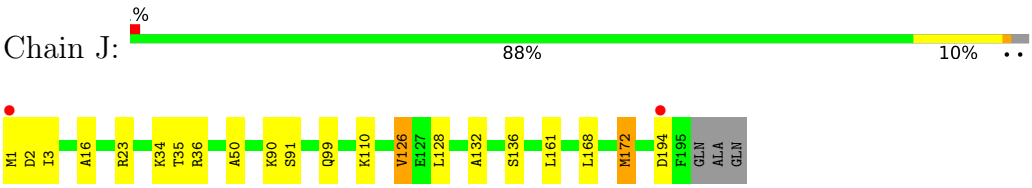
• Molecule 9: Proteasome subunit beta type-3

Chain W:



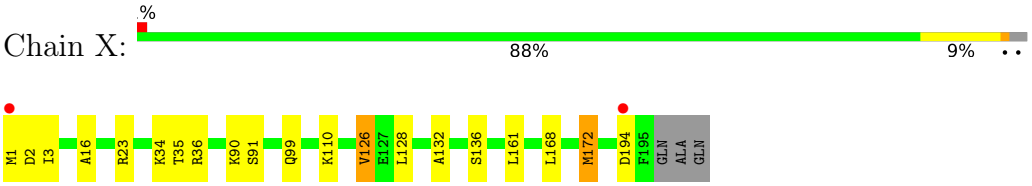
• Molecule 10: Proteasome subunit beta type-4

Chain J:



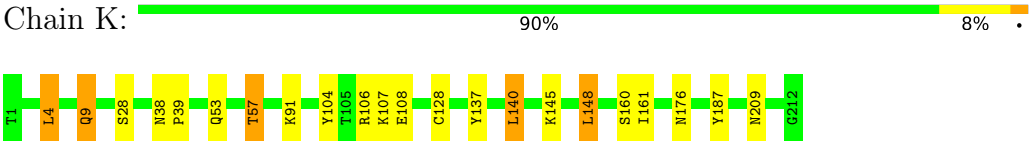
• Molecule 10: Proteasome subunit beta type-4

Chain X:



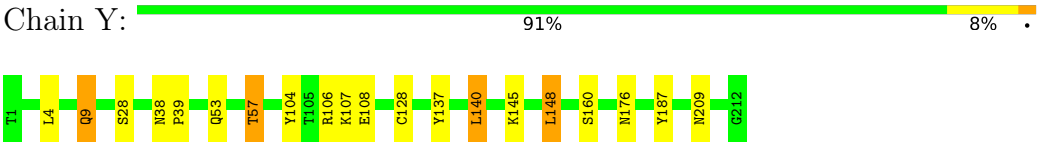
• Molecule 11: Proteasome subunit beta type-5

Chain K:



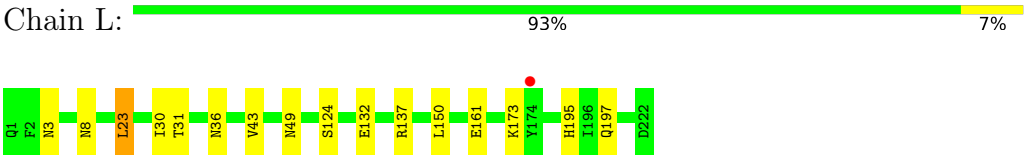
• Molecule 11: Proteasome subunit beta type-5

Chain Y:



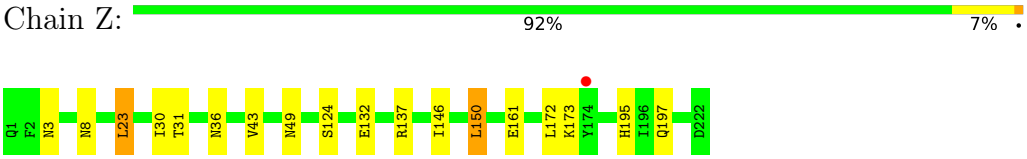
• Molecule 12: Proteasome subunit beta type-6

Chain L:



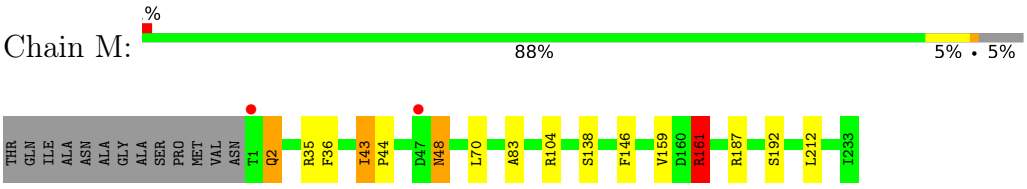
• Molecule 12: Proteasome subunit beta type-6

Chain Z:



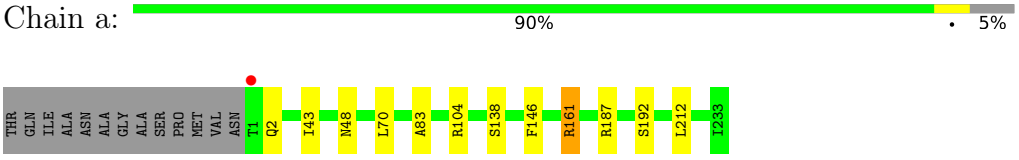
• Molecule 13: Proteasome subunit beta type-7

Chain M:



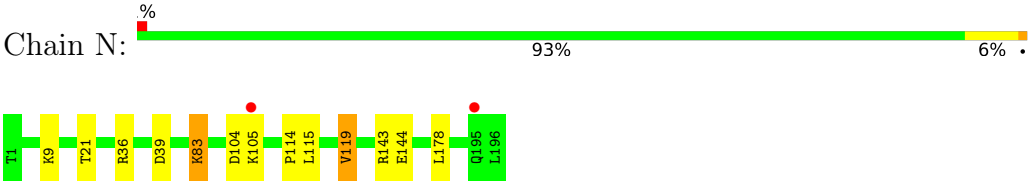
• Molecule 13: Proteasome subunit beta type-7

Chain a:



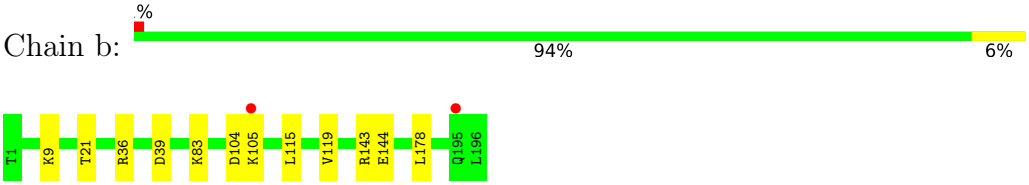
• Molecule 14: Proteasome subunit beta type-1

Chain N:



• Molecule 14: Proteasome subunit beta type-1

Chain b:



• Molecule 15: Ac-PAE-ep

Chain c:

• Molecule 15: Ac-PAE-ep

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	136.40Å 301.59Å 145.63Å 90.00° 113.03° 90.00°	Depositor
Resolution (Å)	15.00 – 2.70 15.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.8 (15.00-2.70) 97.8 (15.00-2.70)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.181 , 0.209 0.183 , 0.208	Depositor DCC
R_{free} test set	14351 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	63.4	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 46.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	50193	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% 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: GAU, POL, ACE, 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	0.27	0/1952	0.51	0/2642
1	O	0.27	0/1952	0.51	0/2642
2	B	0.28	0/1934	0.54	0/2618
2	P	0.27	0/1934	0.54	0/2618
3	C	0.28	0/1910	0.55	0/2586
3	Q	0.28	0/1910	0.55	0/2586
4	D	0.28	0/1837	0.60	2/2475 (0.1%)
4	R	0.28	0/1837	0.59	1/2475 (0.0%)
5	E	0.28	0/1800	0.52	0/2433
5	S	0.27	0/1800	0.52	0/2433
6	F	0.28	0/1932	0.50	0/2609
6	T	0.28	0/1932	0.50	0/2609
7	G	0.28	0/1945	0.53	0/2634
7	U	0.28	0/1945	0.52	0/2634
8	H	0.26	0/1747	0.52	0/2369
8	V	0.26	0/1747	0.52	0/2369
9	I	0.28	0/1611	0.51	0/2174
9	W	0.28	0/1611	0.52	0/2174
10	J	0.27	0/1589	0.51	0/2142
10	X	0.27	0/1589	0.51	0/2142
11	K	0.27	0/1681	0.53	0/2274
11	Y	0.27	0/1681	0.53	0/2274
12	L	0.27	0/1795	0.53	0/2420
12	Z	0.27	0/1795	0.52	0/2420
13	M	0.28	0/1855	0.57	1/2514 (0.0%)
13	a	0.27	0/1855	0.57	1/2514 (0.0%)
14	N	0.26	0/1541	0.50	0/2087
14	b	0.27	0/1541	0.49	0/2087
15	c	1.86	1/13 (7.7%)	0.63	0/18
15	d	1.90	1/13 (7.7%)	0.67	0/18
All	All	0.28	2/50284 (0.0%)	0.53	5/67990 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
2	P	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	d	2	PRO	CA-C	-6.15	1.40	1.52
15	c	2	PRO	CA-C	-6.02	1.40	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	238	LYS	CB-CG-CD	9.55	136.42	111.60
4	R	238	LYS	CB-CG-CD	9.47	136.22	111.60
13	M	161	ARG	NE-CZ-NH1	6.99	123.80	120.30
13	a	161	ARG	NE-CZ-NH1	6.85	123.73	120.30
4	D	238	LYS	CA-CB-CG	6.35	127.37	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	50	LYS	Peptide
2	P	50	LYS	Peptide

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	6	0
1	O	1915	0	1929	8	0
2	B	1904	0	1904	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	P	1904	0	1904	15	0
3	C	1881	0	1895	11	0
3	Q	1881	0	1895	9	0
4	D	1813	0	1797	13	0
4	R	1813	0	1797	13	0
5	E	1773	0	1775	11	0
5	S	1773	0	1775	10	0
6	F	1892	0	1883	8	0
6	T	1892	0	1883	8	0
7	G	1907	0	1901	10	0
7	U	1907	0	1901	9	0
8	H	1717	0	1716	9	0
8	V	1717	0	1716	9	0
9	I	1581	0	1574	6	0
9	W	1581	0	1574	6	0
10	J	1561	0	1569	9	0
10	X	1561	0	1569	8	0
11	K	1644	0	1595	12	0
11	Y	1644	0	1595	10	0
12	L	1757	0	1711	6	0
12	Z	1757	0	1711	7	0
13	M	1824	0	1832	9	0
13	a	1824	0	1832	0	0
14	N	1512	0	1478	4	0
14	b	1512	0	1478	0	0
15	c	28	0	26	0	0
15	d	28	0	26	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	A	30	0	0	0	0
17	B	26	0	0	0	0
17	C	34	0	0	0	0
17	D	11	0	0	0	0
17	E	15	0	0	0	0
17	F	21	0	0	0	0
17	G	36	0	0	0	0
17	H	57	0	0	0	0
17	I	29	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	J	32	0	0	0	0
17	K	27	0	0	0	0
17	L	39	0	0	0	0
17	M	30	0	0	1	0
17	N	24	0	0	0	0
17	O	18	0	0	0	0
17	P	14	0	0	0	0
17	Q	15	0	0	0	0
17	R	19	0	0	0	0
17	S	10	0	0	0	0
17	T	24	0	0	1	0
17	U	35	0	0	0	0
17	V	36	0	0	0	0
17	W	23	0	0	0	0
17	X	24	0	0	0	0
17	Y	36	0	0	0	0
17	Z	39	0	0	0	0
17	a	36	0	0	0	0
17	b	27	0	0	0	0
17	c	1	0	0	0	0
All	All	50193	0	49170	209	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 2.

The worst 5 of 209 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)
13:M:2:GLN:NE2	17:M:301:HOH:O	2.18	0.75
11:K:53:GLN:O	11:K:57:THR:HG23	1.97	0.64
4:R:99:ILE:HD11	4:R:104:LEU:HB2	1.80	0.63
2:P:217:LYS:O	2:P:219:ALA:N	2.32	0.63
4:D:99:ILE:HD11	4:D:104:LEU:HB2	1.81	0.63

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%)	241 (97%)	5 (2%)	2 (1%)	19	43
1	O	248/250 (99%)	240 (97%)	6 (2%)	2 (1%)	19	43
2	B	242/258 (94%)	233 (96%)	7 (3%)	2 (1%)	19	43
2	P	242/258 (94%)	233 (96%)	7 (3%)	2 (1%)	19	43
3	C	238/254 (94%)	232 (98%)	3 (1%)	3 (1%)	12	30
3	Q	238/254 (94%)	232 (98%)	3 (1%)	3 (1%)	12	30
4	D	231/260 (89%)	228 (99%)	2 (1%)	1 (0%)	34	60
4	R	231/260 (89%)	228 (99%)	2 (1%)	1 (0%)	34	60
5	E	229/234 (98%)	219 (96%)	10 (4%)	0	100	100
5	S	229/234 (98%)	220 (96%)	9 (4%)	0	100	100
6	F	241/288 (84%)	232 (96%)	9 (4%)	0	100	100
6	T	241/288 (84%)	232 (96%)	9 (4%)	0	100	100
7	G	239/252 (95%)	236 (99%)	3 (1%)	0	100	100
7	U	239/252 (95%)	236 (99%)	3 (1%)	0	100	100
8	H	224/232 (97%)	218 (97%)	6 (3%)	0	100	100
8	V	224/232 (97%)	218 (97%)	6 (3%)	0	100	100
9	I	202/205 (98%)	197 (98%)	5 (2%)	0	100	100
9	W	202/205 (98%)	197 (98%)	5 (2%)	0	100	100
10	J	193/198 (98%)	186 (96%)	7 (4%)	0	100	100
10	X	193/198 (98%)	186 (96%)	7 (4%)	0	100	100
11	K	210/212 (99%)	204 (97%)	5 (2%)	1 (0%)	29	54
11	Y	210/212 (99%)	204 (97%)	5 (2%)	1 (0%)	29	54
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	231/246 (94%)	223 (96%)	7 (3%)	1 (0%)	34	60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
13	a	231/246 (94%)	223 (96%)	7 (3%)	1 (0%)	34 60
14	N	194/196 (99%)	190 (98%)	4 (2%)	0	100 100
14	b	194/196 (99%)	190 (98%)	4 (2%)	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
All	All	6288/6624 (95%)	6112 (97%)	156 (2%)	20 (0%)	41 66

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	THR
2	B	218	GLY
2	B	222	GLY
3	C	202	GLN
3	C	205	ALA

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%)	201 (96%)	8 (4%)	33 62
1	O	209/209 (100%)	201 (96%)	8 (4%)	33 62
2	B	203/216 (94%)	188 (93%)	15 (7%)	13 32
2	P	203/216 (94%)	188 (93%)	15 (7%)	13 32
3	C	212/226 (94%)	193 (91%)	19 (9%)	9 22
3	Q	212/226 (94%)	193 (91%)	19 (9%)	9 22
4	D	194/215 (90%)	176 (91%)	18 (9%)	9 21
4	R	194/215 (90%)	176 (91%)	18 (9%)	9 21
5	E	190/193 (98%)	172 (90%)	18 (10%)	8 20
5	S	190/193 (98%)	173 (91%)	17 (9%)	9 22
6	F	201/239 (84%)	186 (92%)	15 (8%)	13 31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	T	201/239 (84%)	186 (92%)	15 (8%)	13	31
7	G	206/210 (98%)	192 (93%)	14 (7%)	16	36
7	U	206/210 (98%)	194 (94%)	12 (6%)	20	43
8	H	185/190 (97%)	176 (95%)	9 (5%)	25	52
8	V	185/190 (97%)	176 (95%)	9 (5%)	25	52
9	I	172/173 (99%)	167 (97%)	5 (3%)	42	71
9	W	172/173 (99%)	167 (97%)	5 (3%)	42	71
10	J	173/175 (99%)	163 (94%)	10 (6%)	20	43
10	X	173/175 (99%)	163 (94%)	10 (6%)	20	43
11	K	169/169 (100%)	162 (96%)	7 (4%)	30	59
11	Y	169/169 (100%)	162 (96%)	7 (4%)	30	59
12	L	185/185 (100%)	177 (96%)	8 (4%)	29	57
12	Z	185/185 (100%)	176 (95%)	9 (5%)	25	52
13	M	199/208 (96%)	188 (94%)	11 (6%)	21	46
13	a	199/208 (96%)	188 (94%)	11 (6%)	21	46
14	N	162/162 (100%)	150 (93%)	12 (7%)	13	32
14	b	162/162 (100%)	150 (93%)	12 (7%)	13	32
15	c	1/1 (100%)	1 (100%)	0	100	100
15	d	1/1 (100%)	1 (100%)	0	100	100
All	All	5322/5542 (96%)	4986 (94%)	336 (6%)	18	40

5 of 336 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	S	10	VAL
9	W	192	ASP
5	S	184	ASN
6	T	228	LYS
11	Y	104	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 126 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	48	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
11	Y	85	ASN
3	Q	77	ASN
10	X	147	HIS
12	Z	158	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 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	GAU	c	4	14,15	8,8,8	1.33	1 (12%)	8,9,9	1.88	2 (25%)
15	GAU	d	4	14,15	8,8,8	1.34	1 (12%)	8,9,9	1.85	2 (25%)

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	GAU	c	4	14,15	-	2/7/7/7	-
15	GAU	d	4	14,15	-	2/7/7/7	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	c	4	GAU	C-CA	2.79	1.56	1.52
15	d	4	GAU	C-CA	2.75	1.56	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	c	4	GAU	CG-CB-CA	-4.06	104.22	112.74
15	d	4	GAU	CG-CB-CA	-4.03	104.26	112.74
15	c	4	GAU	OE1-CD-CG	-2.36	115.49	123.08
15	d	4	GAU	OE1-CD-CG	-2.31	115.66	123.08

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	c	4	GAU	O-C-CA-CB
15	d	4	GAU	O-C-CA-CB
15	c	4	GAU	O-C-CA-N
15	d	4	GAU	O-C-CA-N

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 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

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%)	-0.47	3 (1%)	79	80	40, 61, 94, 139	0
1	O	250/250 (100%)	-0.44	7 (2%)	53	54	47, 69, 110, 138	0
2	B	244/258 (94%)	-0.42	7 (2%)	51	52	44, 64, 112, 159	0
2	P	244/258 (94%)	-0.31	9 (3%)	41	41	51, 71, 114, 163	0
3	C	240/254 (94%)	-0.21	11 (4%)	32	31	49, 71, 131, 157	0
3	Q	240/254 (94%)	-0.06	16 (6%)	17	16	54, 81, 148, 164	0
4	D	235/260 (90%)	-0.48	3 (1%)	77	78	50, 71, 105, 134	0
4	R	235/260 (90%)	-0.40	4 (1%)	70	72	50, 72, 107, 140	0
5	E	231/234 (98%)	-0.35	4 (1%)	70	72	53, 75, 109, 153	0
5	S	231/234 (98%)	-0.26	4 (1%)	70	72	54, 81, 115, 158	0
6	F	243/288 (84%)	-0.44	8 (3%)	46	46	47, 69, 113, 141	0
6	T	243/288 (84%)	-0.40	5 (2%)	63	65	52, 75, 123, 151	0
7	G	241/252 (95%)	-0.49	5 (2%)	63	65	47, 62, 95, 146	0
7	U	241/252 (95%)	-0.41	3 (1%)	79	80	50, 67, 100, 129	0
8	H	226/232 (97%)	-0.52	5 (2%)	62	63	45, 58, 90, 144	0
8	V	226/232 (97%)	-0.44	7 (3%)	49	49	47, 63, 95, 163	0
9	I	204/205 (99%)	-0.71	2 (0%)	82	83	43, 58, 84, 111	0
9	W	204/205 (99%)	-0.67	1 (0%)	91	92	47, 61, 92, 124	0
10	J	195/198 (98%)	-0.55	2 (1%)	82	83	47, 61, 89, 144	0
10	X	195/198 (98%)	-0.49	2 (1%)	82	83	47, 63, 88, 150	0
11	K	212/212 (100%)	-0.58	0	100	100	47, 63, 86, 104	0
11	Y	212/212 (100%)	-0.59	0	100	100	47, 62, 85, 107	0
12	L	222/222 (100%)	-0.58	1 (0%)	91	92	47, 62, 94, 139	0
12	Z	222/222 (100%)	-0.60	1 (0%)	91	92	46, 60, 92, 134	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
13	M	233/246 (94%)	-0.63	2 (0%)	84	85	45, 61, 82, 108	0
13	a	233/246 (94%)	-0.61	1 (0%)	92	93	46, 60, 81, 96	0
14	N	196/196 (100%)	-0.69	2 (1%)	82	83	45, 57, 83, 115	0
14	b	196/196 (100%)	-0.64	2 (1%)	82	83	42, 58, 84, 119	0
15	c	2/5 (40%)	-0.84	0	100	100	53, 53, 53, 55	0
15	d	2/5 (40%)	-1.00	0	100	100	52, 52, 52, 54	0
All	All	6348/6624 (95%)	-0.47	117 (1%)	68	70	40, 65, 107, 164	0

The worst 5 of 117 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	49	THR	7.1
3	Q	50	LEU	6.1
3	Q	236	GLN	6.1
8	V	224	GLN	5.8
3	C	206	LYS	5.7

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	GAU	d	4	9/9	0.97	0.11	57,59,62,67	0
15	GAU	c	4	9/9	0.98	0.08	58,62,64,64	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
16	MG	G	301	1/1	0.96	0.11	65,65,65,65	0
16	MG	I	301	1/1	0.96	0.25	63,63,63,63	0
16	MG	Z	301	1/1	0.96	0.17	60,60,60,60	0
16	MG	L	301	1/1	0.97	0.09	74,74,74,74	0
16	MG	K	301	1/1	0.97	0.11	65,65,65,65	0
16	MG	N	201	1/1	0.98	0.06	57,57,57,57	0
16	MG	I	302	1/1	0.98	0.07	56,56,56,56	0

6.5 Other polymers (i)

There are no such residues in this entry.