



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:29 PM BST

PDB ID : 5A0Q  
EMDB ID: : EMD-2981  
Title : Cryo-EM reveals the conformation of a substrate analogue in the human 20S proteasome core  
Authors : daFonseca, P.C.A.; Morris, E.P.  
Deposited on : 2015-04-22  
Resolution : 3.50 Å(reported)  
Based on PDB ID : 3UNE

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.  
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

---

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk27241

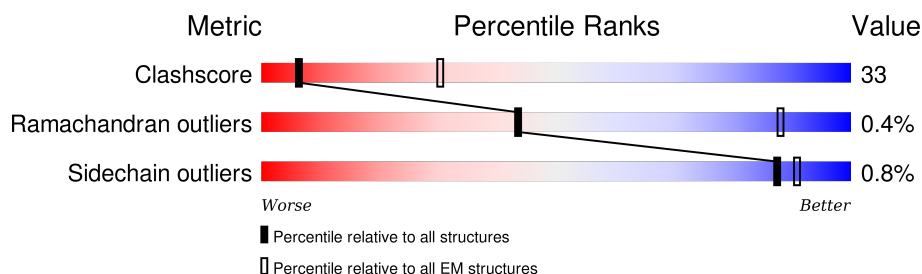
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.50 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.











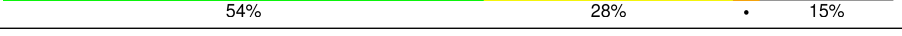


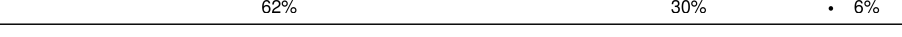

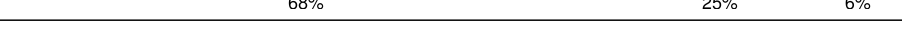


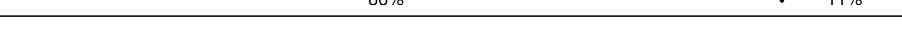

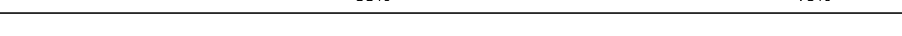
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	246	66% 13% 21%
1	O	246	67% 12% 21%
2	B	234	59% 29% • 9%
2	P	234	60% 27% • 9%
3	C	261	67% 23% • 9%
3	Q	261	66% 23% • 9%
4	D	248	68% 17% • 14%
4	R	248	68% 17% • 14%
5	E	241	67% 20% • 11%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	S	241	
6	F	263	
6	T	263	
7	G	255	
7	U	255	
8	H	205	
8	V	205	
9	I	234	
9	W	234	
10	J	204	
10	X	204	
11	K	201	
11	Y	201	
12	L	204	
12	Z	204	
13	M	213	
13	a	213	
14	N	219	
14	b	219	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
15	KNM	H	300	-	-	X	-
15	KNM	I	300	-	-	X	-
15	KNM	L	300	-	-	X	-
15	KNM	V	300	-	-	X	-
15	KNM	W	300	-	-	X	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
15	KNM	Z	300	-	-	X	-

## 2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 43448 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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-6.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	195	Total	C	N	O	S	0	0
			1514	963	254	284	13		
1	O	195	Total	C	N	O	S	0	0
			1514	963	254	284	13		

- Molecule 2 is a protein called PROTEASOME SUBUNIT ALPHA TYPE-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	214	Total	C	N	O	S	0	0
			1671	1072	285	309	5		
2	P	214	Total	C	N	O	S	0	0
			1671	1072	285	309	5		

- Molecule 3 is a protein called PROTEASOME SUBUNIT ALPHA TYPE-4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	237	Total	C	N	O	S	0	0
			1860	1175	321	354	10		
3	Q	237	Total	C	N	O	S	0	0
			1860	1175	321	354	10		

- Molecule 4 is a protein called PROTEASOME SUBUNIT ALPHA TYPE-7.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	214	Total	C	N	O	S	0	0
			1674	1056	300	313	5		
4	R	214	Total	C	N	O	S	0	0
			1674	1056	300	313	5		

- Molecule 5 is a protein called PROTEASOME SUBUNIT ALPHA TYPE-5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	215	Total	C	N	O	S	0	0
			1643	1036	272	325	10		
5	S	215	Total	C	N	O	S	0	0
			1643	1036	272	325	10		

- Molecule 6 is a protein called PROTEASOME SUBUNIT ALPHA TYPE-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	195	Total	C	N	O	S	0	0
			1535	969	278	278	10		
6	T	195	Total	C	N	O	S	0	0
			1535	969	278	278	10		

- Molecule 7 is a protein called PROTEASOME SUBUNIT ALPHA TYPE-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	209	Total	C	N	O	S	0	0
			1626	1032	279	304	11		
7	U	209	Total	C	N	O	S	0	0
			1626	1032	279	304	11		

- Molecule 8 is a protein called PROTEASOME SUBUNIT BETA TYPE-6.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	183	Total	C	N	O	S	0	0
			1372	858	236	266	12		
8	V	183	Total	C	N	O	S	0	0
			1372	858	236	266	12		

- Molecule 9 is a protein called PROTEASOME SUBUNIT BETA TYPE-7.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	198	Total	C	N	O	S	0	0
			1490	939	251	288	12		
9	W	198	Total	C	N	O	S	0	0
			1490	939	251	288	12		

- Molecule 10 is a protein called PROTEASOME SUBUNIT BETA TYPE-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	176	Total	C	N	O	S	0	0
			1374	882	227	251	14		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
10	X	176	Total	C	N	O	S	0	0
			1374	882	227	251	14		

- Molecule 11 is a protein called PROTEASOME SUBUNIT BETA TYPE-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	189	Total	C	N	O	S	0	0
			1512	970	259	275	8		
11	Y	189	Total	C	N	O	S	0	0
			1512	970	259	275	8		

- Molecule 12 is a protein called PROTEASOME SUBUNIT BETA TYPE-5.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	192	Total	C	N	O	S	0	0
			1480	933	258	280	9		
12	Z	192	Total	C	N	O	S	0	0
			1480	933	258	280	9		

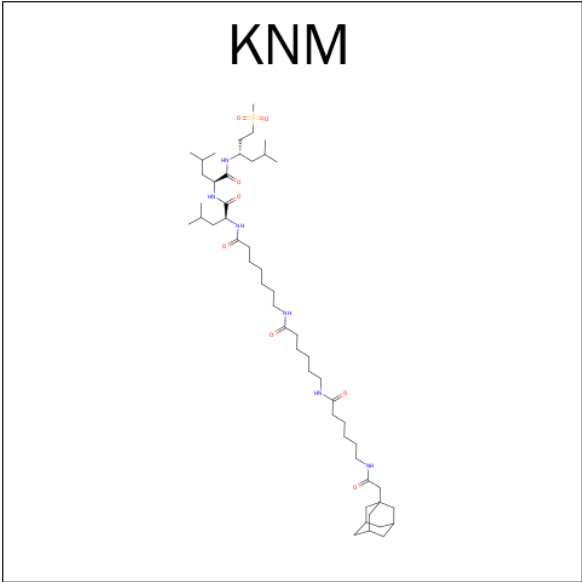
- Molecule 13 is a protein called PROTEASOME SUBUNIT BETA TYPE-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	190	Total	C	N	O	S	0	0
			1453	919	250	275	9		
13	a	190	Total	C	N	O	S	0	0
			1453	919	250	275	9		

- Molecule 14 is a protein called PROTEASOME SUBUNIT BETA TYPE-4.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	184	Total	C	N	O	S	0	0
			1428	905	245	267	11		
14	b	184	Total	C	N	O	S	0	0
			1428	905	245	267	11		

- Molecule 15 is ADA-(AHX)3-(LEU)3-VINYLSULFONE (three-letter code: KNM) (formula: C<sub>51</sub>H<sub>92</sub>N<sub>6</sub>O<sub>8</sub>S).



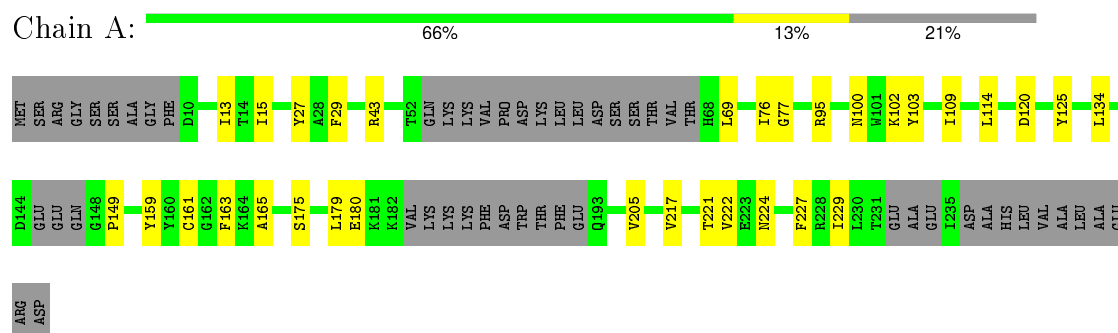
Mol	Chain	Residues	Atoms					AltConf
15	H	1	Total	C	N	O	S	0
			31	22	3	5	1	
15	I	1	Total	C	N	O	S	0
			30	22	3	4	1	
15	L	1	Total	C	N	O	S	0
			31	22	3	5	1	
15	V	1	Total	C	N	O	S	0
			31	22	3	5	1	
15	W	1	Total	C	N	O	S	0
			30	22	3	4	1	
15	Z	1	Total	C	N	O	S	0
			31	22	3	5	1	



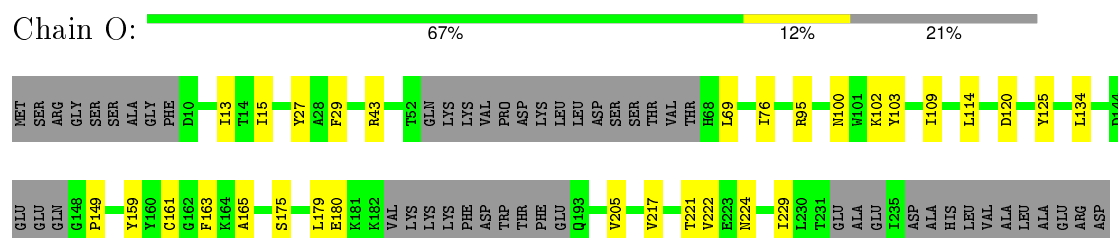
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

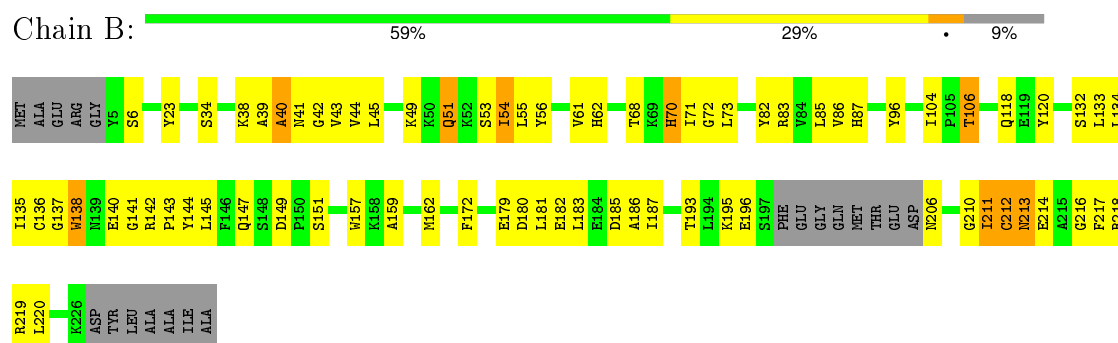
#### • Molecule 1: PROTEASOME SUBUNIT ALPHA TYPE-6



#### • Molecule 1: PROTEASOME SUBUNIT ALPHA TYPE-6



#### • Molecule 2: PROTEASOME SUBUNIT ALPHA TYPE-2



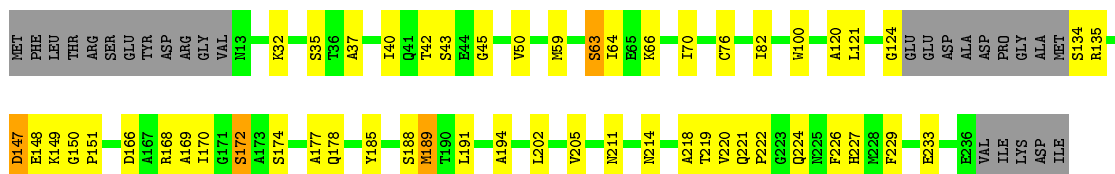
#### • Molecule 2: PROTEASOME SUBUNIT ALPHA TYPE-2





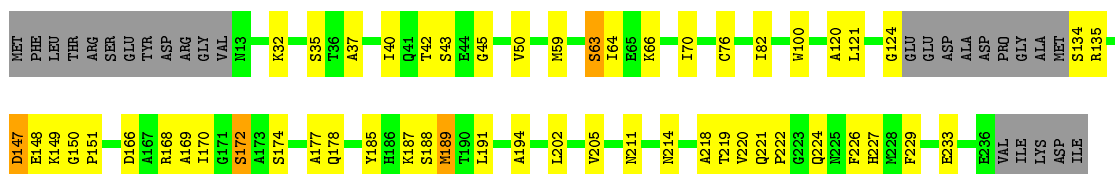
- Molecule 5: PROTEASOME SUBUNIT ALPHA TYPE-5

Chain E:  67% 20% • 11%



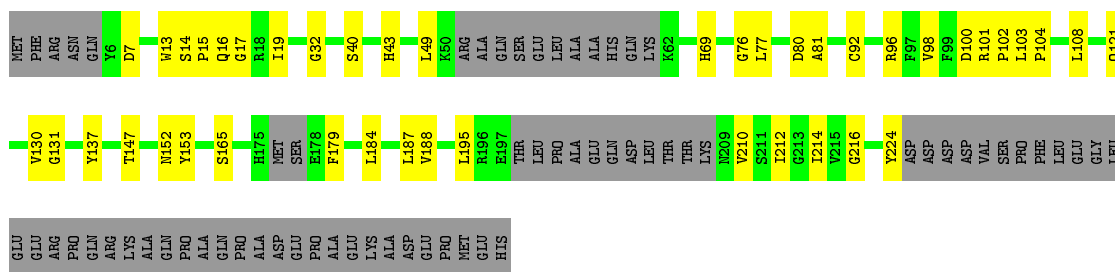
- Molecule 5: PROTEASOME SUBUNIT ALPHA TYPE-5

Chain S:  67% 21% • 11%



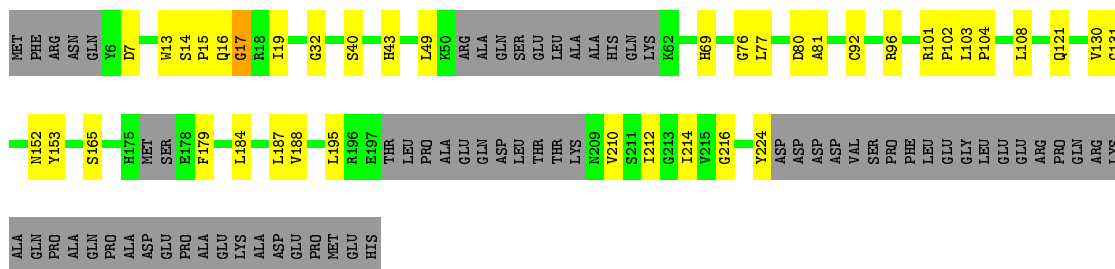
- Molecule 6: PROTEASOME SUBUNIT ALPHA TYPE-1

Chain F:  58% 16% 26%

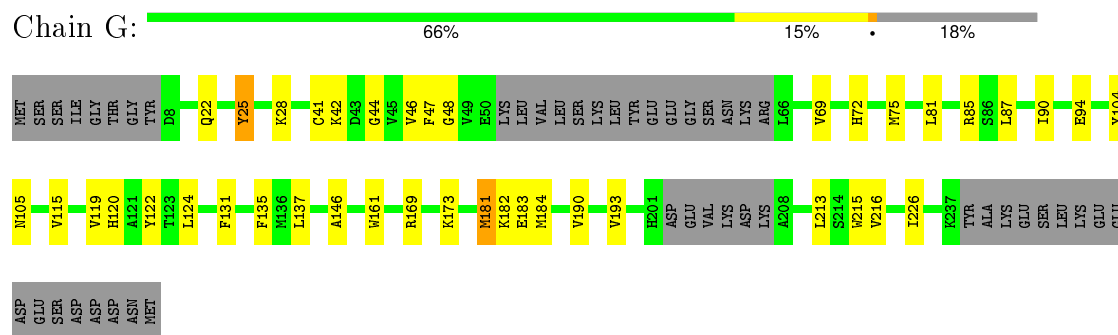


- Molecule 6: PROTEASOME SUBUNIT ALPHA TYPE-1

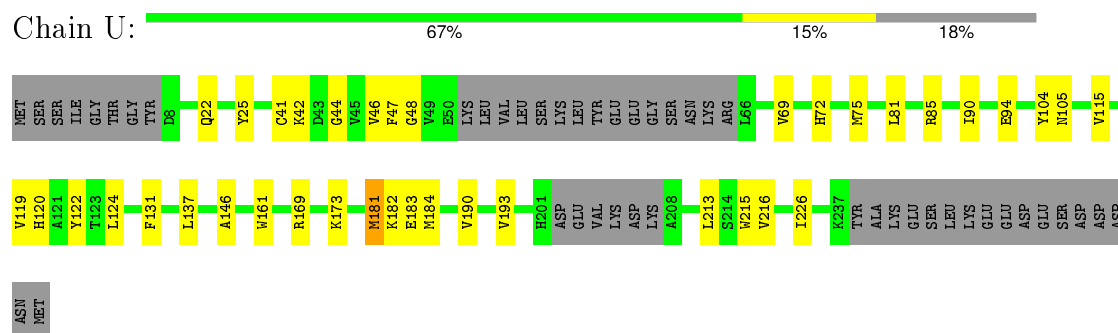
Chain T:  59% 14% 26%



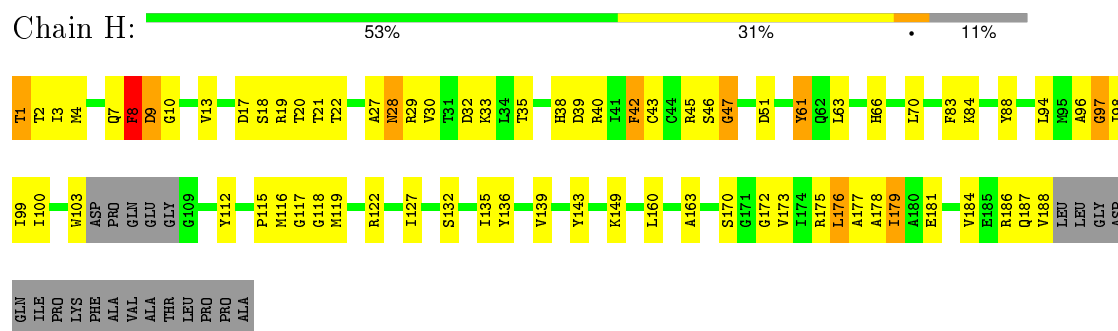
- Molecule 7: PROTEASOME SUBUNIT ALPHA TYPE-3



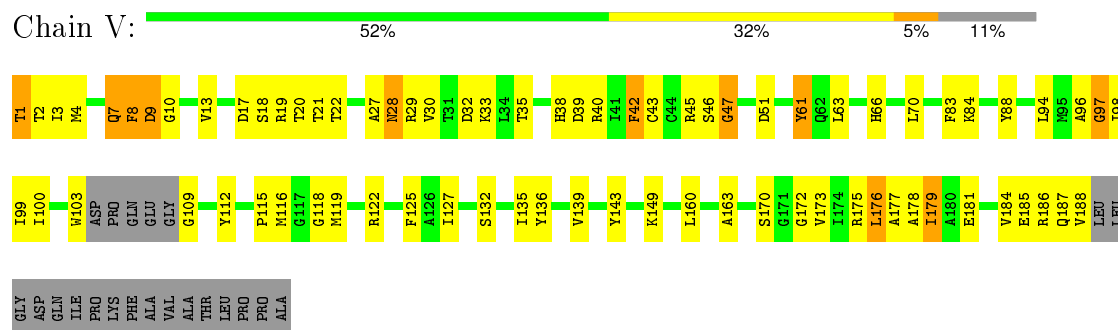
- Molecule 7: PROTEASOME SUBUNIT ALPHA TYPE-3



- Molecule 8: PROTEASOME SUBUNIT BETA TYPE-6



- Molecule 8: PROTEASOME SUBUNIT BETA TYPE-6



- Molecule 9: PROTEASOME SUBUNIT BETA TYPE-7



- Molecule 9: PROTEASOME SUBUNIT BETA TYPE-7

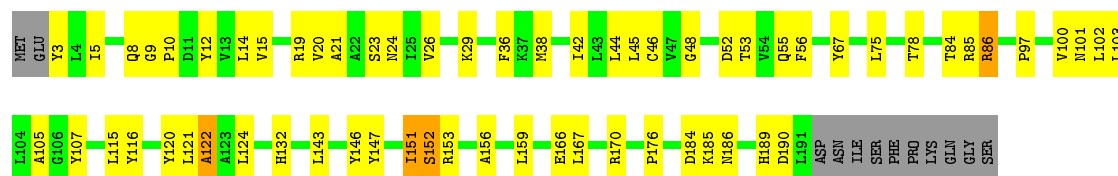
- Molecule 10: PROTEASOME SUBUNIT BETA TYPE-3

- Molecule 10: PROTEASOME SUBUNIT BETA TYPE-3

- Molecule 11: PROTEASOME SUBUNIT BETA TYPE-2

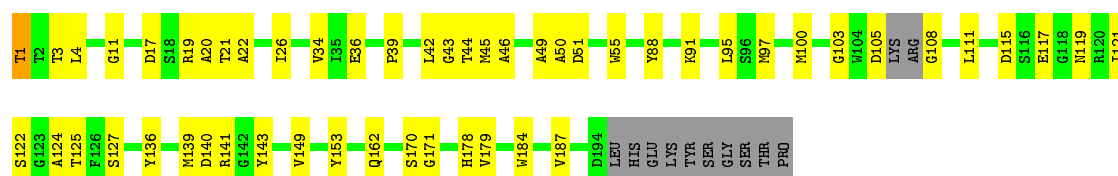
- Molecule 11: PROTEASOME SUBUNIT BETA TYPE-2

Chain Y: 



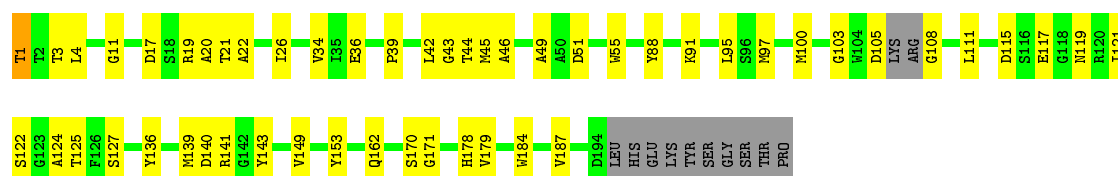
• Molecule 12: PROTEASOME SUBUNIT BETA TYPE-5

Chain L: 



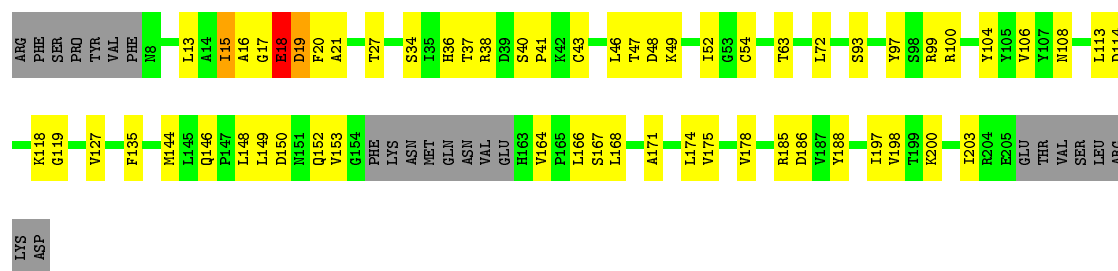
• Molecule 12: PROTEASOME SUBUNIT BETA TYPE-5

Chain Z: 




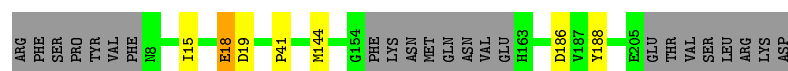
• Molecule 13: PROTEASOME SUBUNIT BETA TYPE-1

Chain M: 



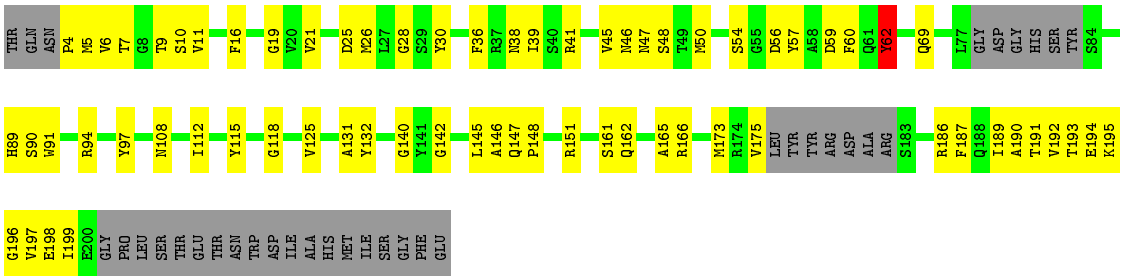
• Molecule 13: PROTEASOME SUBUNIT BETA TYPE-1

Chain a: 



• Molecule 14: PROTEASOME SUBUNIT BETA TYPE-4

Chain N: 



● Molecule 14: PROTEASOME SUBUNIT BETA TYPE-4

Chain b:  83%  16%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	FULL RECORDED IMAGE	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	4.8	Depositor
Minimum defocus (nm)	1700	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	75000	Depositor
Image detector	FEI FALCON II (4K X 4K)	Depositor



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.95	0/1537	1.10	5/2073 (0.2%)
1	O	0.95	0/1537	1.10	5/2073 (0.2%)
10	J	0.92	1/1398 (0.1%)	1.07	3/1884 (0.2%)
10	X	0.92	1/1398 (0.1%)	1.07	3/1884 (0.2%)
11	K	0.87	0/1543	1.06	4/2088 (0.2%)
11	Y	0.87	0/1543	1.06	4/2088 (0.2%)
12	L	1.06	2/1508 (0.1%)	1.13	3/2038 (0.1%)
12	Z	1.06	2/1508 (0.1%)	1.13	3/2038 (0.1%)
13	M	0.91	0/1477	1.11	3/1990 (0.2%)
13	a	0.91	0/1477	1.11	3/1990 (0.2%)
14	N	0.91	1/1451 (0.1%)	1.08	2/1957 (0.1%)
14	b	0.91	1/1451 (0.1%)	1.08	2/1957 (0.1%)
2	B	0.90	0/1707	1.05	5/2312 (0.2%)
2	P	0.90	0/1707	1.05	5/2312 (0.2%)
3	C	0.94	3/1887 (0.2%)	1.07	7/2542 (0.3%)
3	Q	0.94	3/1887 (0.2%)	1.07	7/2542 (0.3%)
4	D	1.00	1/1695 (0.1%)	1.09	2/2283 (0.1%)
4	R	1.00	1/1695 (0.1%)	1.09	2/2283 (0.1%)
5	E	0.82	0/1668	1.02	4/2252 (0.2%)
5	S	0.82	0/1668	1.02	4/2252 (0.2%)
6	F	0.99	0/1562	1.12	6/2105 (0.3%)
6	T	1.00	0/1562	1.19	7/2105 (0.3%)
7	G	0.97	0/1656	1.12	2/2232 (0.1%)
7	U	0.97	0/1656	1.12	2/2232 (0.1%)
8	H	1.05	1/1394 (0.1%)	1.16	7/1884 (0.4%)
8	V	1.04	1/1394 (0.1%)	1.16	7/1884 (0.4%)
9	I	0.96	2/1515 (0.1%)	1.14	4/2050 (0.2%)
9	W	0.96	2/1515 (0.1%)	1.14	4/2050 (0.2%)
All	All	0.95	22/43996 (0.1%)	1.10	115/59380 (0.2%)

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
1	A	0	1
1	O	0	1
10	J	0	1
10	X	0	1
11	K	0	1
11	Y	0	1
13	M	0	2
13	a	0	2
5	E	0	1
5	S	0	1
7	G	0	1
7	U	0	1
9	I	0	2
9	W	0	2
All	All	0	18

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	H	1	THR	C-N	19.41	1.78	1.34
8	V	1	THR	C-N	19.41	1.78	1.34
9	I	123	PRO	N-CD	-7.63	1.37	1.47
9	W	123	PRO	N-CD	-7.63	1.37	1.47
12	L	1	THR	C-N	6.47	1.49	1.34

The worst 5 of 115 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	T	17	GLY	N-CA-C	15.39	151.57	113.10
8	H	47	GLY	N-CA-C	-9.29	89.87	113.10
8	V	47	GLY	N-CA-C	-9.29	89.87	113.10
11	Y	38	MET	CG-SD-CE	-8.17	87.13	100.20
11	K	38	MET	CG-SD-CE	-8.15	87.16	100.20

There are no chirality outliers.

5 of 18 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	95	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
5	E	147	ASP	Peptide
7	G	181	MET	Peptide
9	I	1	THR	Mainchain
9	I	111	TYR	Sidechain

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1514	0	1527	73	0
1	O	1514	0	1527	74	0
2	B	1671	0	1681	153	0
2	P	1671	0	1681	151	0
3	C	1860	0	1886	84	0
3	Q	1860	0	1886	86	0
4	D	1674	0	1712	67	0
4	R	1674	0	1712	68	0
5	E	1643	0	1636	93	0
5	S	1643	0	1636	94	0
6	F	1535	0	1540	73	0
6	T	1535	0	1540	68	0
7	G	1626	0	1602	96	0
7	U	1626	0	1602	93	0
8	H	1372	0	1338	141	0
8	V	1372	0	1338	118	0
9	I	1490	0	1490	142	0
9	W	1490	0	1490	164	0
10	J	1374	0	1383	74	0
10	X	1374	0	1383	86	0
11	K	1512	0	1520	121	0
11	Y	1512	0	1520	117	0
12	L	1480	0	1441	106	0
12	Z	1480	0	1441	96	0
13	M	1453	0	1455	163	0
13	a	1453	0	1455	0	0
14	N	1428	0	1432	188	0
14	b	1428	0	1432	0	0
15	H	31	0	39	33	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
15	I	30	0	39	45	0
15	L	31	0	39	30	0
15	V	31	0	39	31	0
15	W	30	0	39	43	0
15	Z	31	0	39	29	0
All	All	43448	0	43520	2537	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 33.

The worst 5 of 2537 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:152:GLN:HG3	9:W:202:TYR:CD2	1.31	1.63
1:A:109:ILE:CG2	1:A:114:LEU:HD21	1.26	1.62
14:N:26:MET:CE	14:N:186:ARG:HH12	1.10	1.60
1:O:109:ILE:CG2	1:O:114:LEU:HD21	1.26	1.57
14:N:16:PHE:CE1	14:N:166:ARG:N	1.74	1.56

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 PDB entries followed by that with respect to all EM entries.

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	186/246 (76%)	178 (96%)	8 (4%)	0	100	100
1	O	186/246 (76%)	178 (96%)	8 (4%)	0	100	100
2	B	210/234 (90%)	198 (94%)	10 (5%)	2 (1%)	19	66
2	P	210/234 (90%)	198 (94%)	10 (5%)	2 (1%)	19	66
3	C	233/261 (89%)	225 (97%)	8 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	Q	233/261 (89%)	225 (97%)	8 (3%)	0	100	100
4	D	206/248 (83%)	201 (98%)	5 (2%)	0	100	100
4	R	206/248 (83%)	201 (98%)	5 (2%)	0	100	100
5	E	211/241 (88%)	201 (95%)	9 (4%)	1 (0%)	34	78
5	S	211/241 (88%)	201 (95%)	9 (4%)	1 (0%)	34	78
6	F	187/263 (71%)	184 (98%)	3 (2%)	0	100	100
6	T	187/263 (71%)	185 (99%)	2 (1%)	0	100	100
7	G	203/255 (80%)	196 (97%)	7 (3%)	0	100	100
7	U	203/255 (80%)	196 (97%)	7 (3%)	0	100	100
8	H	179/205 (87%)	164 (92%)	13 (7%)	2 (1%)	17	63
8	V	179/205 (87%)	165 (92%)	12 (7%)	2 (1%)	17	63
9	I	194/234 (83%)	187 (96%)	6 (3%)	1 (0%)	34	78
9	W	194/234 (83%)	187 (96%)	6 (3%)	1 (0%)	34	78
10	J	168/204 (82%)	158 (94%)	10 (6%)	0	100	100
10	X	168/204 (82%)	158 (94%)	10 (6%)	0	100	100
11	K	187/201 (93%)	173 (92%)	10 (5%)	4 (2%)	9	50
11	Y	187/201 (93%)	173 (92%)	10 (5%)	4 (2%)	9	50
12	L	188/204 (92%)	182 (97%)	6 (3%)	0	100	100
12	Z	188/204 (92%)	182 (97%)	6 (3%)	0	100	100
13	M	186/213 (87%)	175 (94%)	10 (5%)	1 (0%)	34	78
13	a	186/213 (87%)	175 (94%)	10 (5%)	1 (0%)	34	78
14	N	178/219 (81%)	161 (90%)	17 (10%)	0	100	100
14	b	178/219 (81%)	161 (90%)	17 (10%)	0	100	100
All	All	5432/6456 (84%)	5168 (95%)	242 (4%)	22 (0%)	43	81

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	40	ALA
2	B	54	ILE
5	E	189	MET
8	H	8	PHE
11	K	122	ALA

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	166/210 (79%)	166 (100%)	0	100	100
1	O	166/210 (79%)	166 (100%)	0	100	100
2	B	177/191 (93%)	170 (96%)	7 (4%)	38	75
2	P	177/191 (93%)	170 (96%)	7 (4%)	38	75
3	C	199/221 (90%)	197 (99%)	2 (1%)	82	93
3	Q	199/221 (90%)	197 (99%)	2 (1%)	82	93
4	D	179/211 (85%)	177 (99%)	2 (1%)	80	92
4	R	179/211 (85%)	177 (99%)	2 (1%)	80	92
5	E	181/203 (89%)	180 (99%)	1 (1%)	90	97
5	S	181/203 (89%)	180 (99%)	1 (1%)	90	97
6	F	166/224 (74%)	166 (100%)	0	100	100
6	T	166/224 (74%)	166 (100%)	0	100	100
7	G	170/212 (80%)	170 (100%)	0	100	100
7	U	170/212 (80%)	170 (100%)	0	100	100
8	H	142/159 (89%)	140 (99%)	2 (1%)	74	91
8	V	142/159 (89%)	140 (99%)	2 (1%)	74	91
9	I	162/195 (83%)	162 (100%)	0	100	100
9	W	162/195 (83%)	162 (100%)	0	100	100
10	J	149/173 (86%)	149 (100%)	0	100	100
10	X	149/173 (86%)	149 (100%)	0	100	100
11	K	160/171 (94%)	159 (99%)	1 (1%)	90	97
11	Y	160/171 (94%)	159 (99%)	1 (1%)	90	97
12	L	148/159 (93%)	148 (100%)	0	100	100
12	Z	148/159 (93%)	148 (100%)	0	100	100
13	M	155/178 (87%)	153 (99%)	2 (1%)	76	91
13	a	155/178 (87%)	153 (99%)	2 (1%)	76	91

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
14	N	152/181 (84%)	151 (99%)	1 (1%)	88	96
14	b	152/181 (84%)	151 (99%)	1 (1%)	88	96
All	All	4612/5376 (86%)	4576 (99%)	36 (1%)	87	95

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	18	GLU
2	P	106	THR
13	a	15	ILE
2	P	51	GLN
2	P	138	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
2	P	51	GLN
2	P	213	ASN
11	Y	189	HIS
2	P	62	HIS
2	P	87	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

6 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	$\# Z  > 2$
15	KNM	H	300	8	30,30,68	1.80	1 (3%)	40,41,91	0.92	3 (7%)
15	KNM	I	300	9	29,29,68	1.79	1 (3%)	36,39,91	1.24	3 (8%)
15	KNM	L	300	12	30,30,68	1.79	2 (6%)	40,41,91	0.87	3 (7%)
15	KNM	V	300	8	30,30,68	1.80	1 (3%)	40,41,91	0.92	3 (7%)
15	KNM	W	300	9	29,29,68	1.78	1 (3%)	36,39,91	1.24	3 (8%)
15	KNM	Z	300	12	30,30,68	1.79	2 (6%)	40,41,91	0.86	3 (7%)

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	KNM	H	300	8	-	0/38/38/98	0/0/0/3
15	KNM	I	300	9	-	0/37/37/98	0/0/0/3
15	KNM	L	300	12	-	0/38/38/98	0/0/0/3
15	KNM	V	300	8	-	0/38/38/98	0/0/0/3
15	KNM	W	300	9	-	0/37/37/98	0/0/0/3
15	KNM	Z	300	12	-	0/38/38/98	0/0/0/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	V	300	KNM	C16-S1	-9.55	1.66	1.78
15	H	300	KNM	C16-S1	-9.55	1.66	1.78
15	L	300	KNM	C16-S1	-9.32	1.66	1.78
15	Z	300	KNM	C16-S1	-9.30	1.66	1.78
15	I	300	KNM	C16-S1	-9.00	1.66	1.78

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	W	300	KNM	O4-S1-O5	-3.11	108.90	116.90

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	I	300	KNM	O4-S1-O5	-3.11	108.90	116.90
15	V	300	KNM	O4-S1-O5	-2.92	109.40	116.90
15	H	300	KNM	O4-S1-O5	-2.92	109.40	116.90
15	L	300	KNM	O4-S1-O5	-2.91	109.42	116.90

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 211 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	H	300	KNM	33	0
15	I	300	KNM	45	0
15	L	300	KNM	30	0
15	V	300	KNM	31	0
15	W	300	KNM	43	0
15	Z	300	KNM	29	0

## 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.