wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 31, 2023 – 12:49 PM JST

PDB ID	:	5DSV
Title	:	Crystal structure of human proteasome alpha7 tetradecamer
Authors	:	Satoh, T.; Thammaporn, R.; Seetaha, S.; Kato, K.
Deposited on	:	2015-09-17
Resolution	:	3.75 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

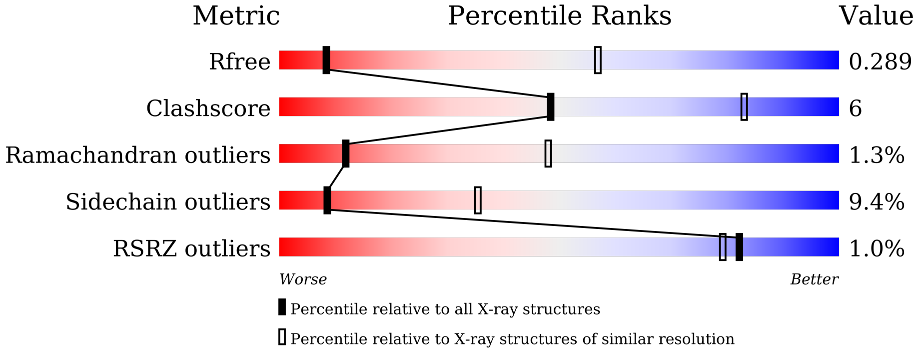
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

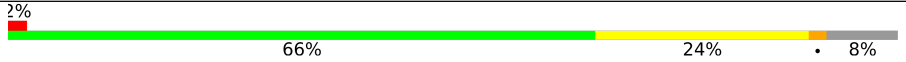
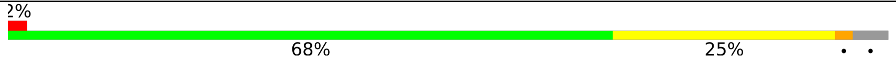
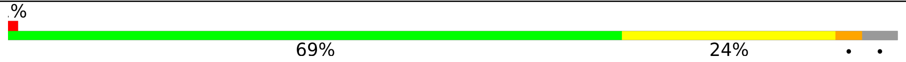
The reported resolution of this entry is 3.75 Å.

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.



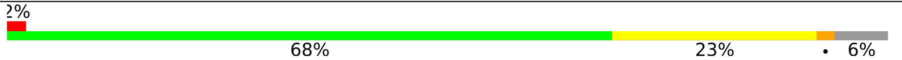
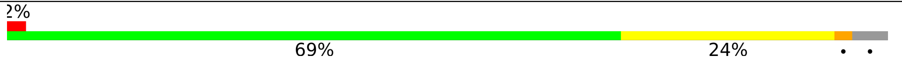
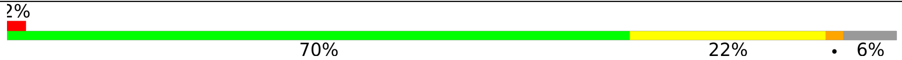
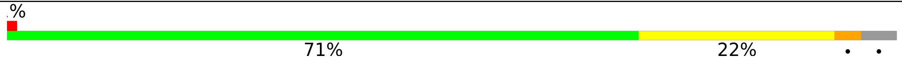
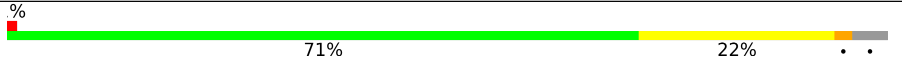
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1039 (3.94-3.58)
Clashscore	141614	1051 (3.92-3.60)
Ramachandran outliers	138981	1015 (3.92-3.60)
Sidechain outliers	138945	1011 (3.92-3.60)
RSRZ outliers	127900	1050 (3.96-3.56)

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	255	
1	B	255	
1	C	255	
1	D	255	
1	E	255	
1	F	255	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	255	
1	H	255	
1	I	255	
1	J	255	
1	K	255	
1	L	255	
1	M	255	
1	N	255	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 26493 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.

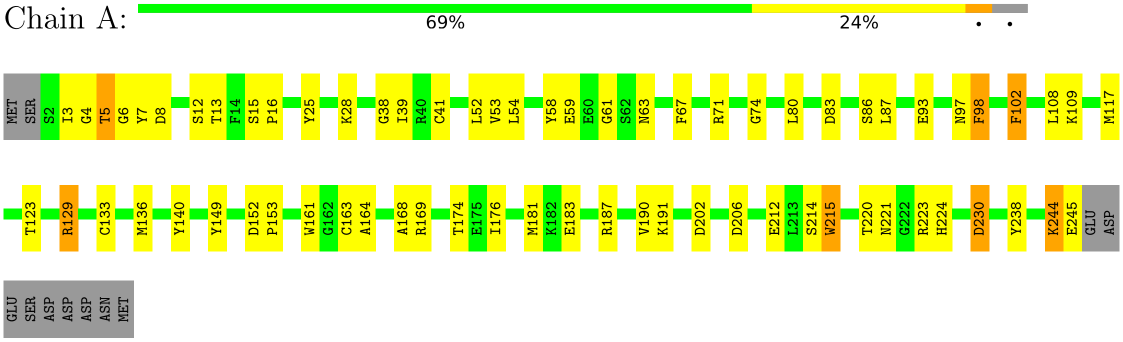
- Molecule 1 is a protein called Proteasome subunit alpha type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	244	Total 1908	C 1209	N 325	O 363	S 11	0	0	0
1	B	245	Total 1914	C 1212	N 326	O 365	S 11	0	0	0
1	C	234	Total 1825	C 1155	N 312	O 348	S 10	0	0	0
1	D	244	Total 1908	C 1209	N 325	O 363	S 11	0	0	0
1	E	245	Total 1914	C 1212	N 326	O 365	S 11	0	0	0
1	F	244	Total 1908	C 1209	N 325	O 363	S 11	0	0	0
1	G	242	Total 1894	C 1200	N 323	O 360	S 11	0	0	0
1	H	239	Total 1863	C 1179	N 318	O 355	S 11	0	0	0
1	I	244	Total 1905	C 1207	N 325	O 362	S 11	0	0	0
1	J	239	Total 1863	C 1179	N 318	O 355	S 11	0	0	0
1	K	244	Total 1905	C 1207	N 325	O 362	S 11	0	0	0
1	L	245	Total 1914	C 1212	N 326	O 365	S 11	0	0	0
1	M	239	Total 1864	C 1178	N 318	O 357	S 11	0	0	0
1	N	244	Total 1908	C 1209	N 325	O 363	S 11	0	0	0

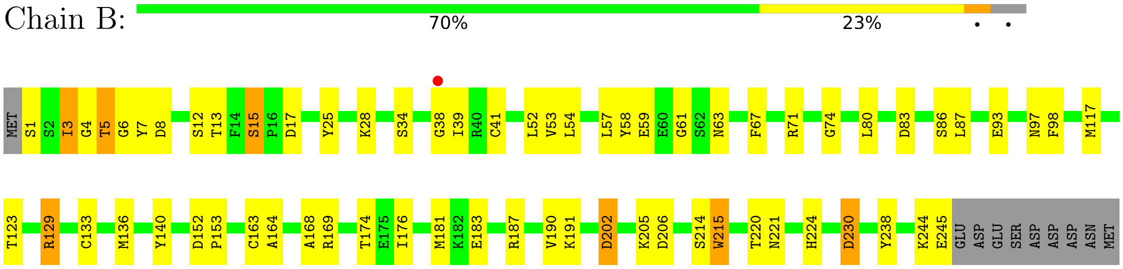
3 Residue-property plots

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.

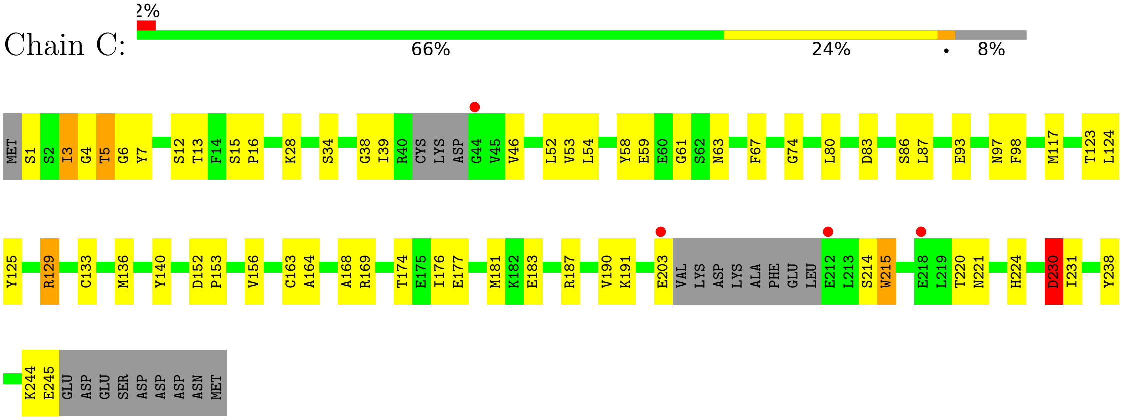
- Molecule 1: Proteasome subunit alpha type-3



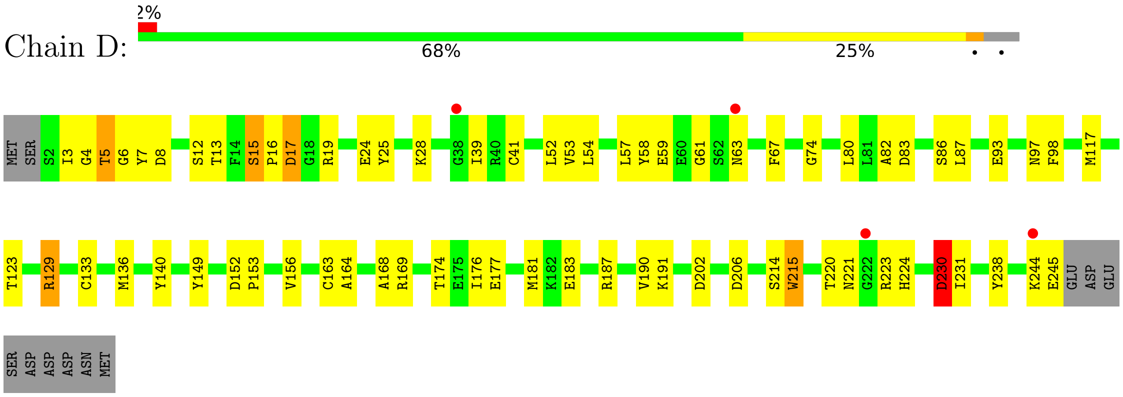
- Molecule 1: Proteasome subunit alpha type-3



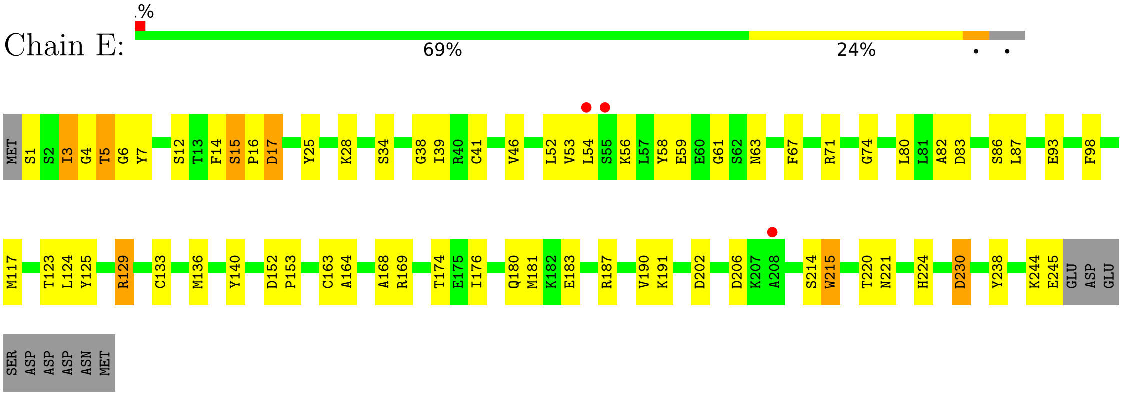
- Molecule 1: Proteasome subunit alpha type-3



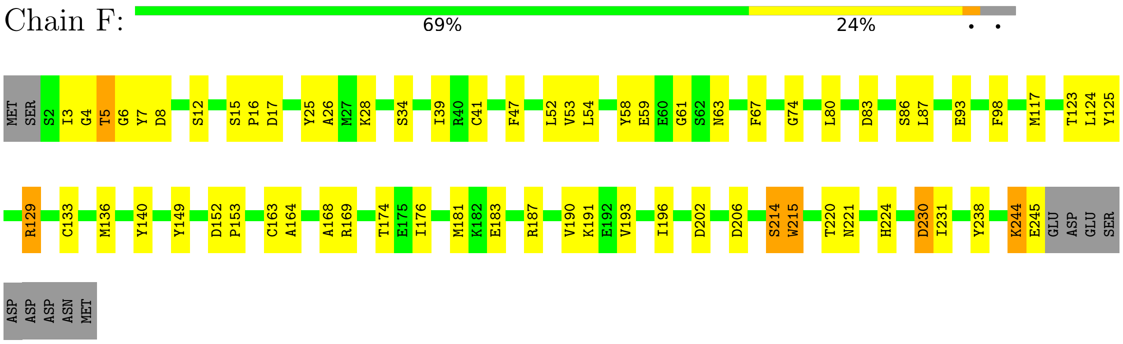
- Molecule 1: Proteasome subunit alpha type-3



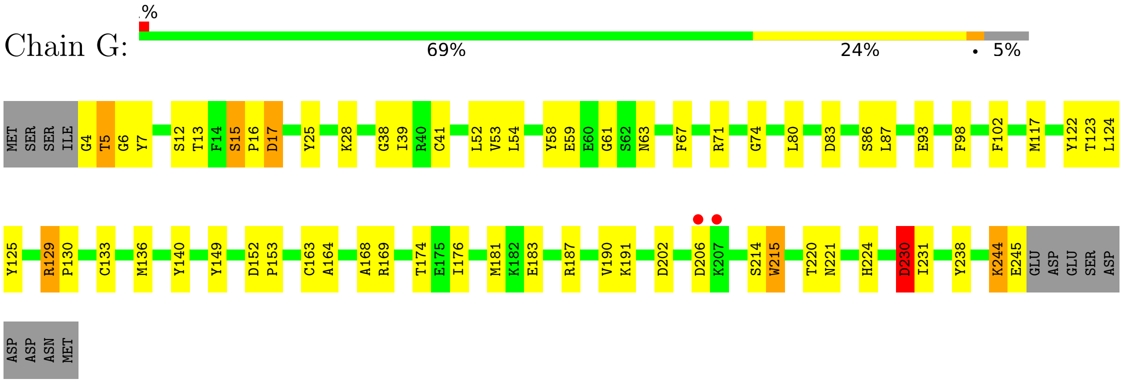
- Molecule 1: Proteasome subunit alpha type-3



- Molecule 1: Proteasome subunit alpha type-3

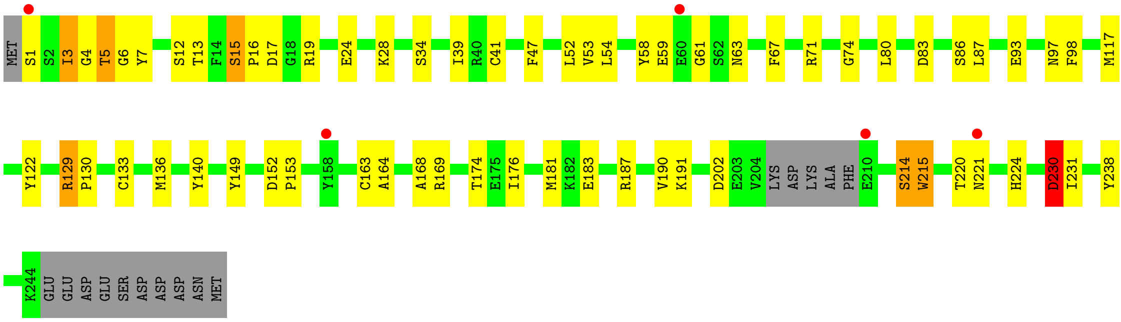


- Molecule 1: Proteasome subunit alpha type-3



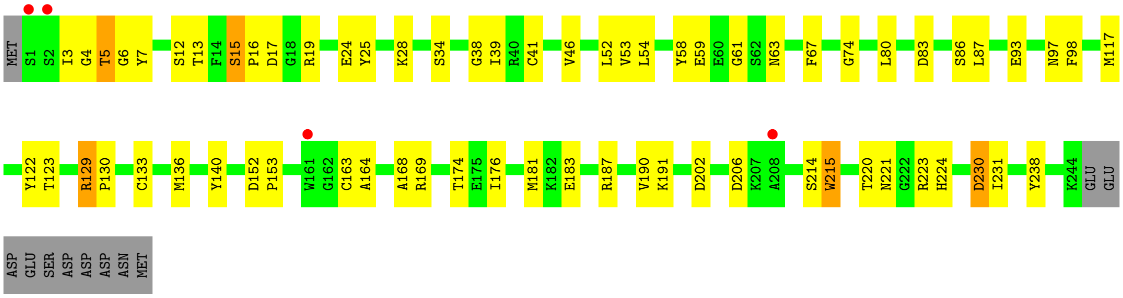
- Molecule 1: Proteasome subunit alpha type-3

Chain H:



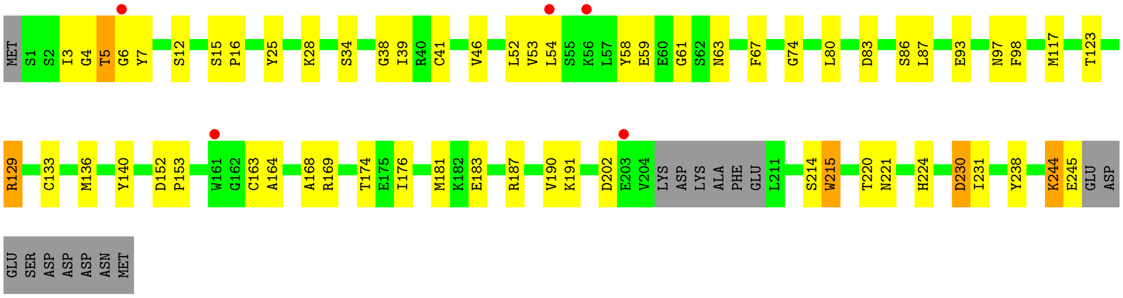
- Molecule 1: Proteasome subunit alpha type-3

Chain I:



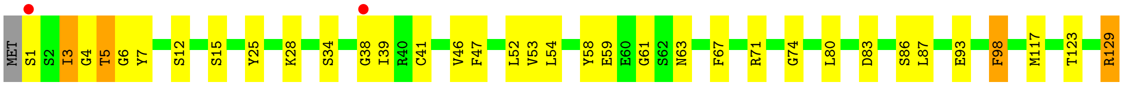
- Molecule 1: Proteasome subunit alpha type-3

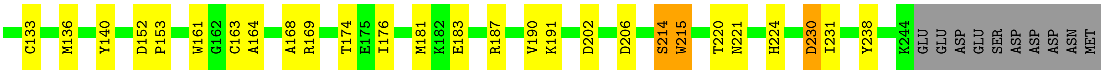
Chain J:



- Molecule 1: Proteasome subunit alpha type-3

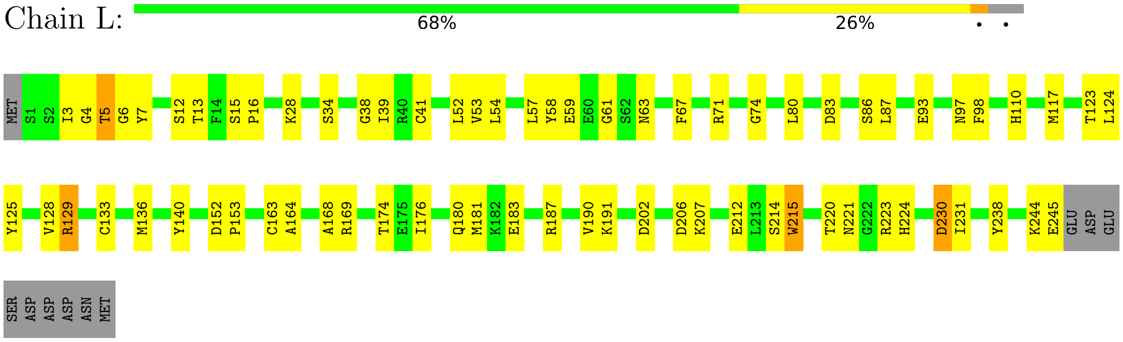
Chain K:





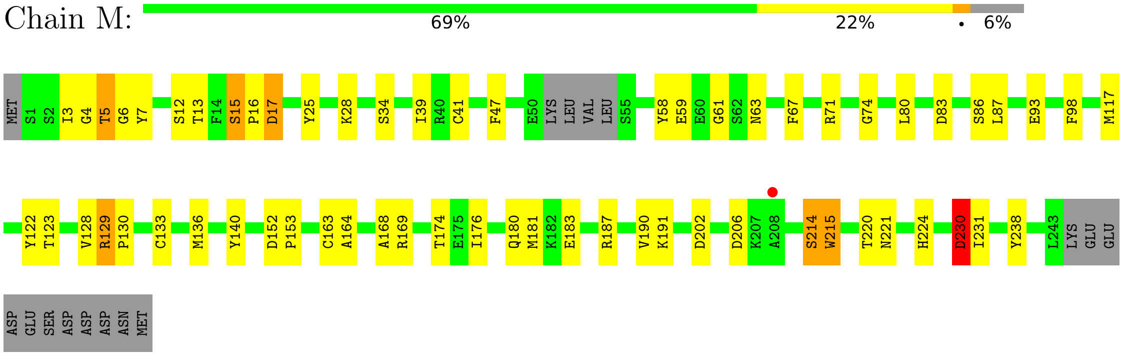
- Molecule 1: Proteasome subunit alpha type-3

Chain L:



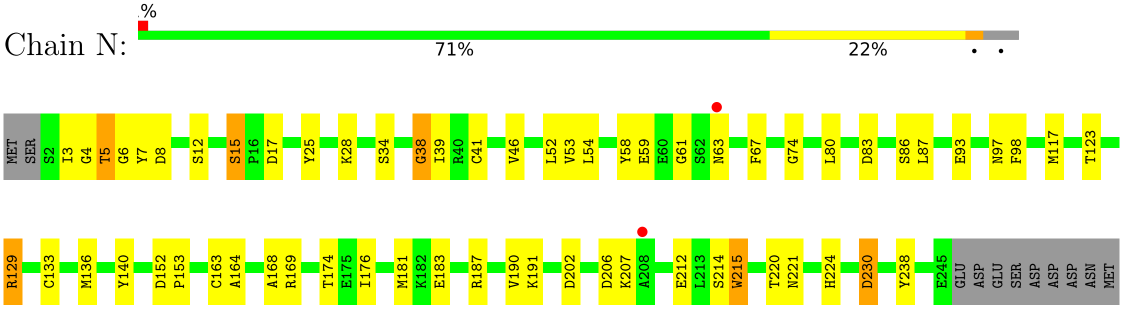
- Molecule 1: Proteasome subunit alpha type-3

Chain M:



- Molecule 1: Proteasome subunit alpha type-3

Chain N:



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	132.48Å 132.48Å 444.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.75 19.97 – 3.75	Depositor EDS
% Data completeness (in resolution range)	96.5 (20.00-3.75) 97.4 (19.97-3.75)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.05 (at 3.71Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.209 , 0.291 0.211 , 0.289	Depositor DCC
R_{free} test set	2039 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	126.1	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 74.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	26493	wwPDB-VP
Average B, all atoms (Å^2)	153.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.55	0/1943	0.69	0/2616
1	B	0.54	0/1949	0.68	0/2624
1	C	0.52	0/1857	0.67	1/2499 (0.0%)
1	D	0.51	0/1943	0.67	0/2616
1	E	0.57	0/1949	0.69	0/2624
1	F	0.57	0/1943	0.68	0/2616
1	G	0.56	0/1929	0.69	0/2597
1	H	0.48	0/1896	0.66	0/2553
1	I	0.47	0/1940	0.65	0/2612
1	J	0.47	0/1896	0.66	0/2553
1	K	0.51	0/1940	0.67	0/2612
1	L	0.53	0/1949	0.68	1/2624 (0.0%)
1	M	0.51	0/1898	0.68	1/2555 (0.0%)
1	N	0.49	0/1943	0.66	0/2616
All	All	0.52	0/26975	0.67	3/36317 (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
1	A	0	16
1	B	0	15
1	C	0	14
1	D	0	15
1	E	0	14
1	F	0	13
1	G	0	15
1	H	0	14
1	I	0	13
1	J	0	13
1	K	0	11

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	L	0	15
1	M	0	15
1	N	0	14
All	All	0	197

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	M	128	VAL	CB-CA-C	-5.01	101.88	111.40
1	C	203	GLU	OE1-CD-OE2	-5.01	117.29	123.30
1	L	128	VAL	CB-CA-C	-5.00	101.89	111.40

There are no chirality outliers.

5 of 197 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	13	THR	Peptide
1	A	3	ILE	Peptide
1	A	38	GLY	Peptide
1	A	4	GLY	Peptide
1	A	58	TYR	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	1908	0	1891	37	0
1	B	1914	0	1899	26	0
1	C	1825	0	1807	26	0
1	D	1908	0	1891	30	0
1	E	1914	0	1899	31	0
1	F	1908	0	1891	26	0
1	G	1894	0	1875	28	0
1	H	1863	0	1848	26	0
1	I	1905	0	1893	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	1863	0	1848	21	0
1	K	1905	0	1893	28	0
1	L	1914	0	1899	28	0
1	M	1864	0	1835	25	0
1	N	1908	0	1891	21	0
All	All	26493	0	26260	318	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 6.

The worst 5 of 318 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)
1:A:102:PHE:HZ	1:K:98:PHE:CE2	1.84	0.94
1:A:102:PHE:HZ	1:K:98:PHE:HE2	1.22	0.86
1:A:102:PHE:CZ	1:K:98:PHE:HE2	1.94	0.85
1:A:71:ARG:HD3	1:L:71:ARG:HD3	1.64	0.79
1:A:71:ARG:HB3	1:L:71:ARG:NH1	1.99	0.76

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	242/255 (95%)	213 (88%)	26 (11%)	3 (1%)	13	49
1	B	243/255 (95%)	215 (88%)	25 (10%)	3 (1%)	13	49
1	C	228/255 (89%)	201 (88%)	24 (10%)	3 (1%)	12	48
1	D	242/255 (95%)	213 (88%)	26 (11%)	3 (1%)	13	49
1	E	243/255 (95%)	213 (88%)	27 (11%)	3 (1%)	13	49

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	242/255 (95%)	213 (88%)	26 (11%)	3 (1%)	13	49
1	G	240/255 (94%)	212 (88%)	25 (10%)	3 (1%)	12	48
1	H	235/255 (92%)	209 (89%)	23 (10%)	3 (1%)	12	48
1	I	242/255 (95%)	214 (88%)	25 (10%)	3 (1%)	13	49
1	J	235/255 (92%)	207 (88%)	25 (11%)	3 (1%)	12	48
1	K	242/255 (95%)	213 (88%)	26 (11%)	3 (1%)	13	49
1	L	243/255 (95%)	215 (88%)	24 (10%)	4 (2%)	9	44
1	M	235/255 (92%)	209 (89%)	24 (10%)	2 (1%)	17	54
1	N	242/255 (95%)	213 (88%)	25 (10%)	4 (2%)	9	43
All	All	3354/3570 (94%)	2960 (88%)	351 (10%)	43 (1%)	12	48

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	53	VAL
1	C	53	VAL
1	D	53	VAL
1	E	53	VAL
1	F	53	VAL

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 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	201/212 (95%)	181 (90%)	20 (10%)	7	32
1	B	202/212 (95%)	183 (91%)	19 (9%)	8	35
1	C	192/212 (91%)	176 (92%)	16 (8%)	11	41
1	D	201/212 (95%)	183 (91%)	18 (9%)	9	37
1	E	202/212 (95%)	183 (91%)	19 (9%)	8	35
1	F	201/212 (95%)	182 (90%)	19 (10%)	8	35
1	G	199/212 (94%)	181 (91%)	18 (9%)	9	37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	H	197/212 (93%)	179 (91%)	18 (9%)	9	37
1	I	201/212 (95%)	182 (90%)	19 (10%)	8	35
1	J	197/212 (93%)	179 (91%)	18 (9%)	9	37
1	K	201/212 (95%)	182 (90%)	19 (10%)	8	35
1	L	202/212 (95%)	182 (90%)	20 (10%)	8	33
1	M	196/212 (92%)	177 (90%)	19 (10%)	8	34
1	N	201/212 (95%)	181 (90%)	20 (10%)	7	32
All	All	2793/2968 (94%)	2531 (91%)	262 (9%)	8	35

5 of 262 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	86	SER
1	M	202	ASP
1	N	214	SER
1	F	41	CYS
1	F	15	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	63	ASN
1	N	63	ASN
1	H	110	HIS
1	M	63	ASN
1	H	63	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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, 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	244/255 (95%)	-0.52	0	100	100	78, 123, 190, 244	0
1	B	245/255 (96%)	-0.49	1 (0%)	92	91	84, 132, 192, 207	0
1	C	234/255 (91%)	-0.31	4 (1%)	70	65	91, 159, 220, 253	0
1	D	244/255 (95%)	-0.29	4 (1%)	72	67	84, 169, 226, 265	0
1	E	245/255 (96%)	-0.53	3 (1%)	79	75	81, 127, 199, 260	0
1	F	244/255 (95%)	-0.55	0	100	100	75, 119, 193, 213	0
1	G	242/255 (94%)	-0.51	2 (0%)	86	84	72, 124, 190, 232	0
1	H	239/255 (93%)	-0.34	5 (2%)	63	58	112, 166, 220, 239	0
1	I	244/255 (95%)	-0.21	4 (1%)	72	67	110, 180, 241, 272	0
1	J	239/255 (93%)	-0.24	5 (2%)	63	58	101, 172, 235, 268	0
1	K	244/255 (95%)	-0.44	2 (0%)	86	84	84, 144, 197, 247	0
1	L	245/255 (96%)	-0.44	0	100	100	93, 144, 209, 249	0
1	M	239/255 (93%)	-0.33	1 (0%)	92	91	94, 152, 212, 258	0
1	N	244/255 (95%)	-0.27	2 (0%)	86	84	98, 161, 235, 364	0
All	All	3392/3570 (95%)	-0.39	33 (0%)	82	79	72, 148, 220, 364	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	1	SER	4.3
1	C	44	GLY	4.3
1	D	38	GLY	3.2
1	D	63	ASN	3.1
1	I	161	TRP	2.8

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.