wwPDB X-ray Structure Validation Summary Report (i)

Sep 13, 2023 – 10:22 PM EDT

PDB ID	:	4Q1S
Title	:	Yeast 20S proteasome in Complex with Kendomycin
Authors	:	Beck, P.; Heinemeyer, W.; Spaeth, A.; Elnakady, Y.; Mueller, R.; Groll, M.
Deposited on	:	2014-04-04
Resolution	:	2.60 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:

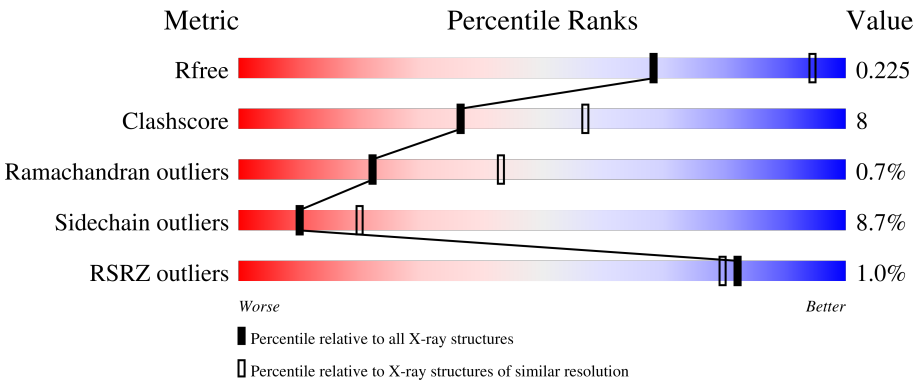
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	84% 14% •
1	O	250	% 84% 13% •
2	B	258	% 69% 22% • 5%
2	P	258	2% 73% 18% • 5%
3	C	254	2% 72% 19% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	4% 65% 25% 5% 5%
4	D	260	2% 74% 17% 7%
4	R	260	2% 71% 20% 7%
5	E	234	% 75% 20%
5	S	234	2% 70% 24%
6	F	288	% 69% 13% 15%
6	T	288	% 62% 19% 15%
7	G	252	% 80% 14%
7	U	252	75% 19%
8	H	232	72% 20%
8	V	232	% 74% 21%
9	I	205	79% 19%
9	W	205	79% 18%
10	J	198	2% 86% 12%
10	X	198	2% 82% 15%
11	K	212	82% 17%
11	Y	212	83% 14%
12	L	222	79% 16%
12	Z	222	72% 25%
13	M	246	78% 15% 5%
13	a	246	88% 7% 5%
14	N	196	85% 13%
14	b	196	96%

2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 50912 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	241	1890	1181	331	374	4	0	0	0
3	Q	241	1890	1181	331	374	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	242	1861	1162	314	378	7	0	0	0
4	R	242	1861	1162	314	378	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	233	1795	1129	312	350	4	0	0	0
5	S	233	1795	1129	312	350	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	244	1896	1205	330	357	4	0	0	0
6	T	244	1896	1205	330	357	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	243	1921	1221	322	370	8	0	0	0
7	U	243	1921	1221	322	370	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	198	1585	1005	269	305	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	198	1585	1005	269	305	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	211	1638	1042	279	310	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	1757	1115	303	335	4	0	0	0
12	Z	222	1757	1115	303	335	4	0	0	0

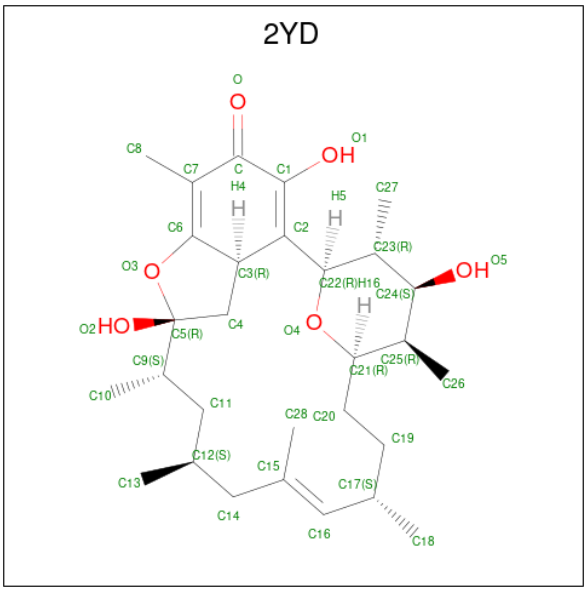
- Molecule 13 is a protein called Proteasome subunit beta type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	233	1824	1154	312	351	7	0	0	0
13	a	233	1824	1154	312	351	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 (5R,6R,7S,8R,9R,12S,13E,16S,18S,19R,20aR)-4,7,19-trihydroxy-2,6,8,12,14,16,18-heptamethyl-6,7,8,9,10,11,12,15,16,17,18,19,20,20a-tetradecahydro-1,19:5,9-diepoxybenzo[18]annulen-3(5H)-one (three-letter code: 2YD) (formula: $C_{29}H_{44}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	H	1	Total C O 35 29 6	0	0
15	V	1	Total C O 35 29 6	0	0

- Molecule 16 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	A	59	Total O 59 59	0	0
16	B	39	Total O 39 39	0	0
16	C	42	Total O 42 42	0	0
16	D	39	Total O 39 39	0	0
16	E	20	Total O 20 20	0	0
16	F	49	Total O 49 49	0	0
16	G	58	Total O 58 58	0	0
16	H	53	Total O 53 53	0	0
16	I	63	Total O 63 63	0	0
16	J	51	Total O 51 51	0	0

Continued on next page...

Continued from previous page...

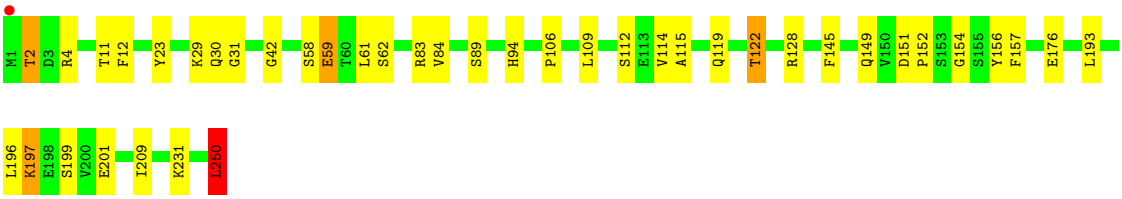
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	K	44	Total O 44 44	0	0
16	L	58	Total O 58 58	0	0
16	M	64	Total O 64 64	0	0
16	N	58	Total O 58 58	0	0
16	O	31	Total O 31 31	0	0
16	P	30	Total O 30 30	0	0
16	Q	25	Total O 25 25	0	0
16	R	32	Total O 32 32	0	0
16	S	21	Total O 21 21	0	0
16	T	38	Total O 38 38	0	0
16	U	66	Total O 66 66	0	0
16	V	44	Total O 44 44	0	0
16	W	62	Total O 62 62	0	0
16	X	49	Total O 49 49	0	0
16	Y	48	Total O 48 48	0	0
16	Z	48	Total O 48 48	0	0
16	a	62	Total O 62 62	0	0
16	b	57	Total O 57 57	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.

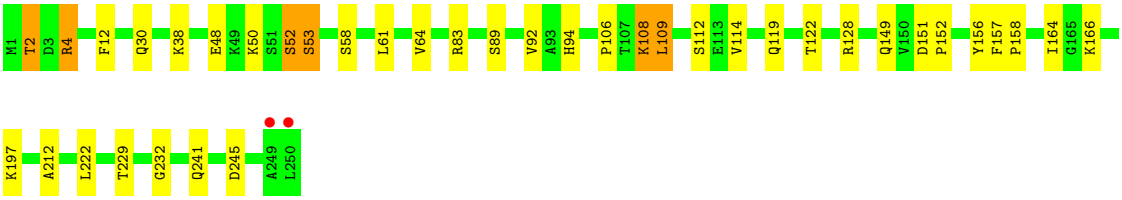
• Molecule 1: Proteasome subunit alpha type-2

Chain A:



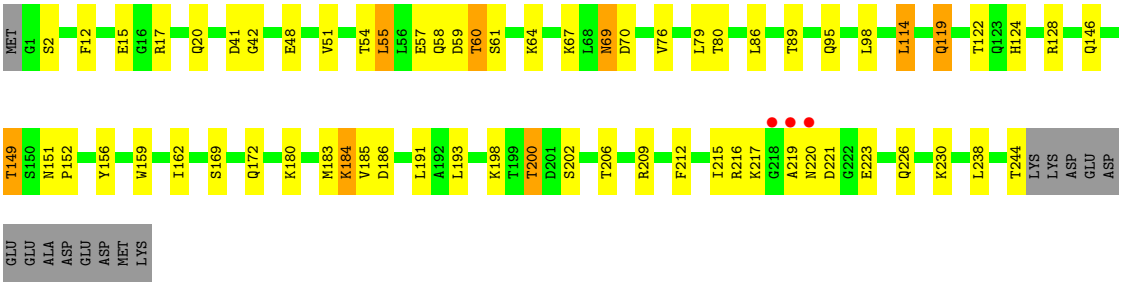
• Molecule 1: Proteasome subunit alpha type-2

Chain O:



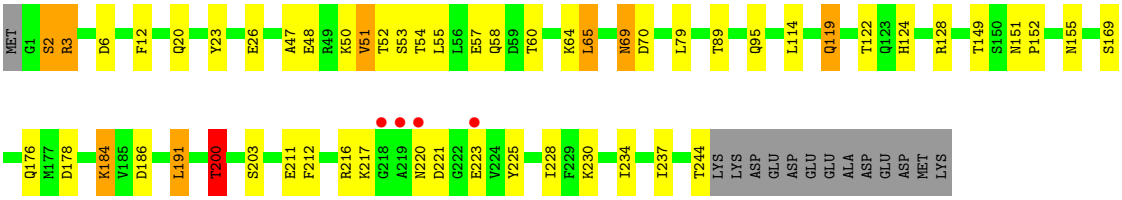
• Molecule 2: Proteasome subunit alpha type-3

Chain B:



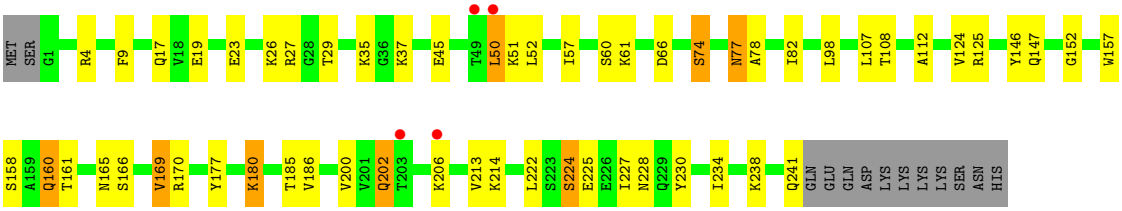
• Molecule 2: Proteasome subunit alpha type-3

Chain P:



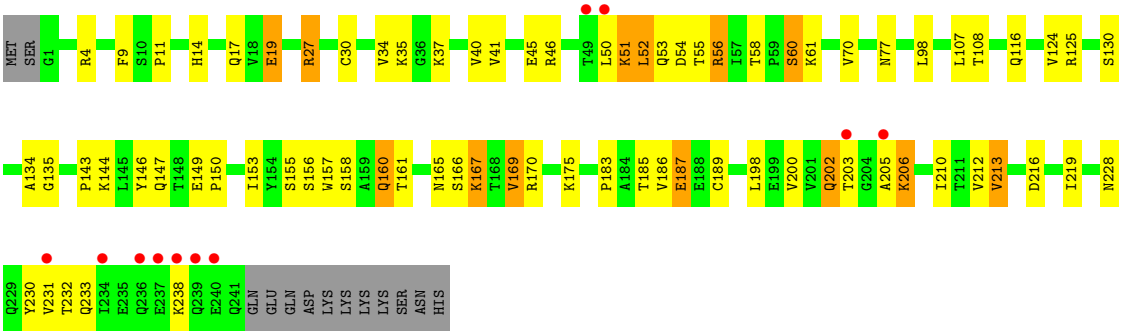
• Molecule 3: Proteasome subunit alpha type-4

Chain C:



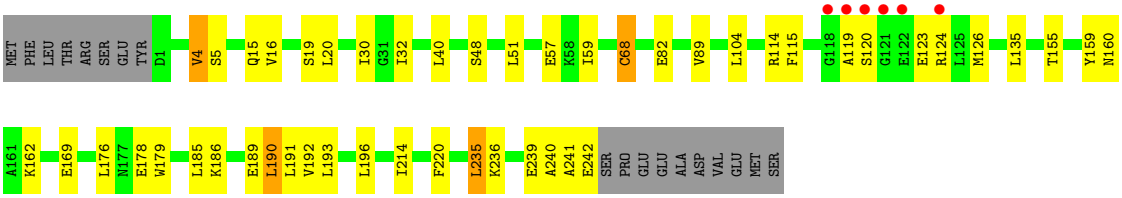
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



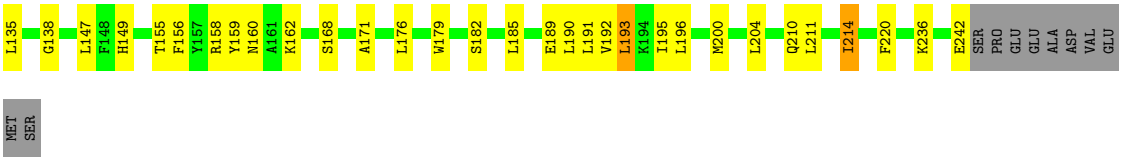
• Molecule 4: Proteasome subunit alpha type-5

Chain D:

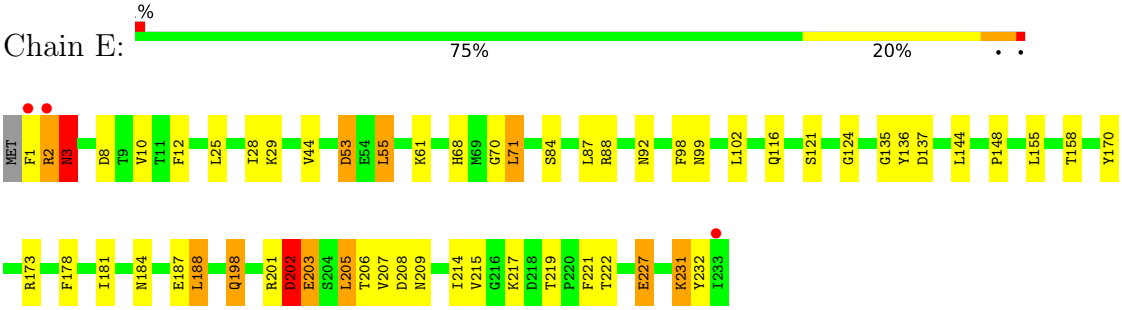


• Molecule 4: Proteasome subunit alpha type-5

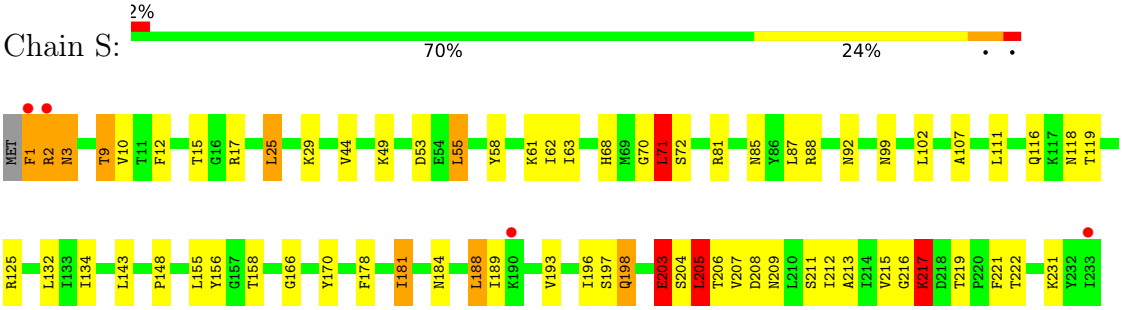
Chain R:



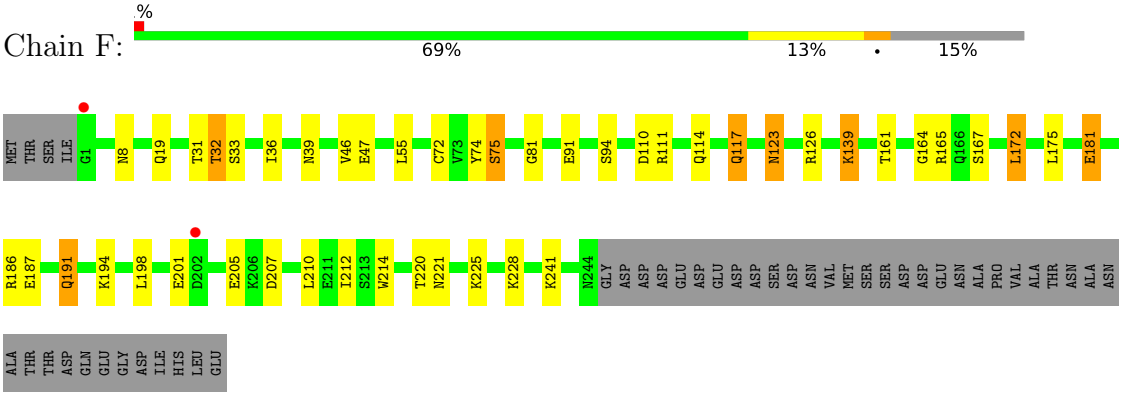
• Molecule 5: Proteasome subunit alpha type-6



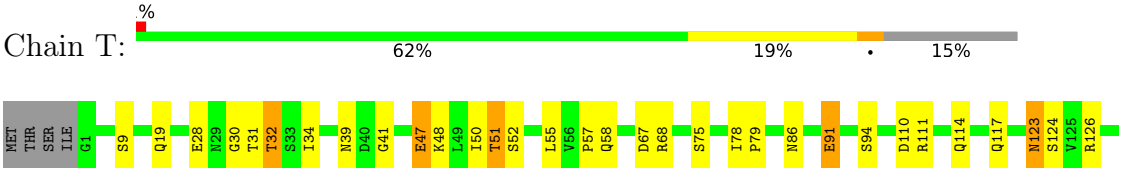
• Molecule 5: Proteasome subunit alpha type-6

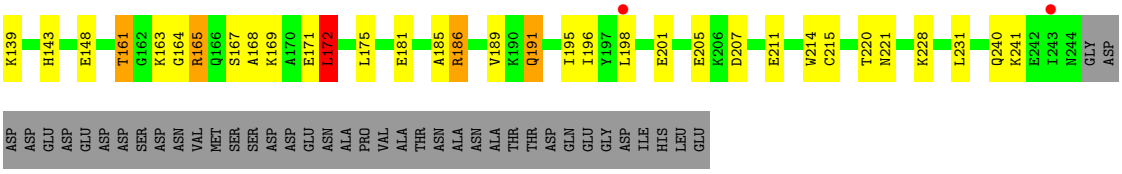


• Molecule 6: Probable proteasome subunit alpha type-7



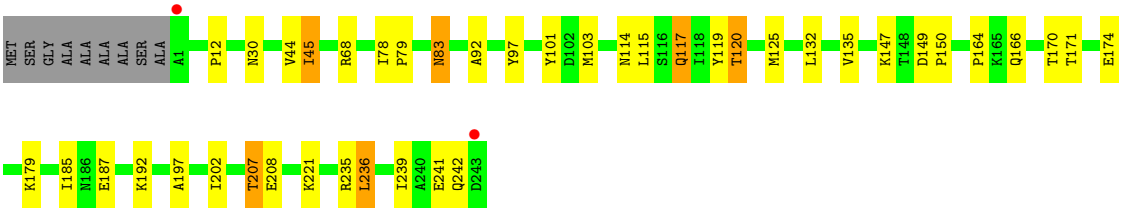
• Molecule 6: Probable proteasome subunit alpha type-7





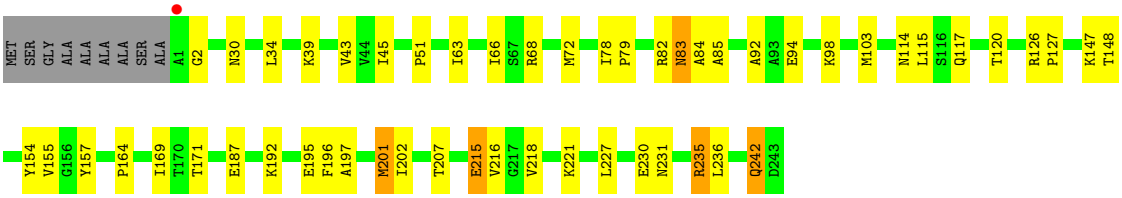
• Molecule 7: Proteasome subunit alpha type-1

Chain G:



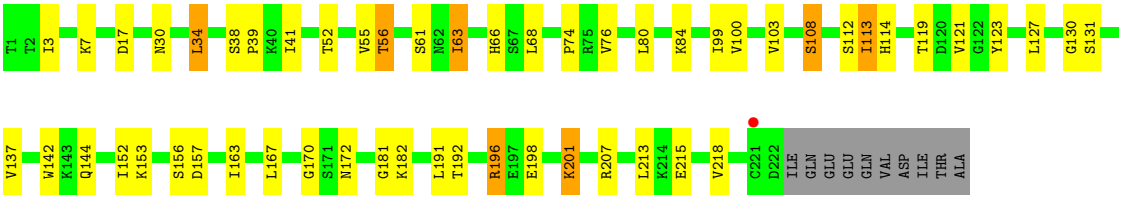
• Molecule 7: Proteasome subunit alpha type-1

Chain U:



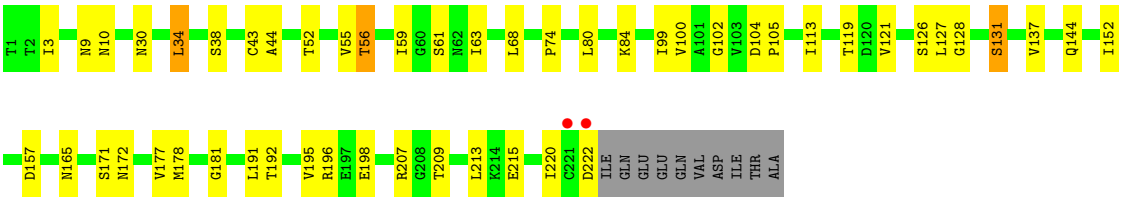
• Molecule 8: Proteasome subunit beta type-2

Chain H:



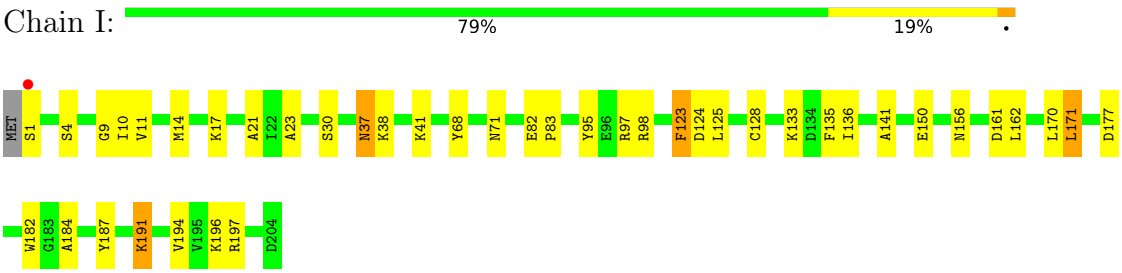
• Molecule 8: Proteasome subunit beta type-2

Chain V:



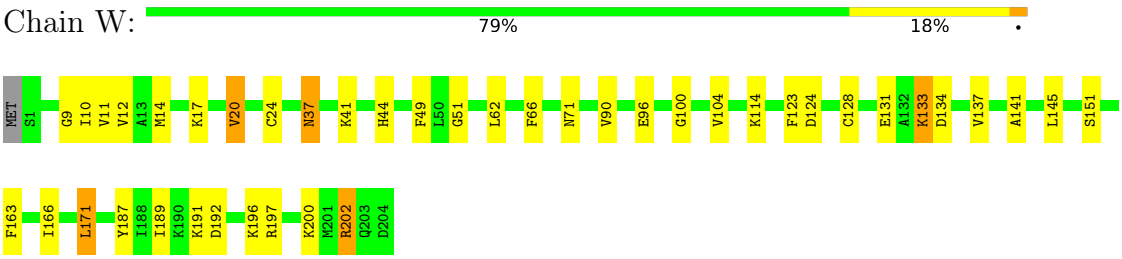
• Molecule 9: Proteasome subunit beta type-3

Chain I:



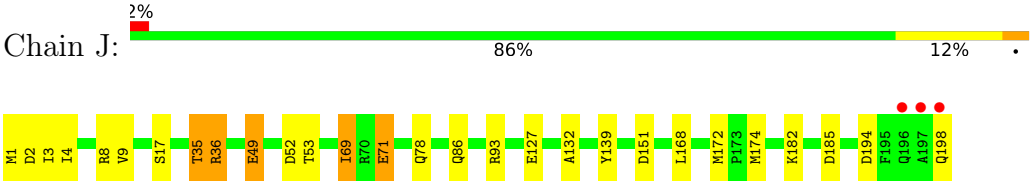
• Molecule 9: Proteasome subunit beta type-3

Chain W:



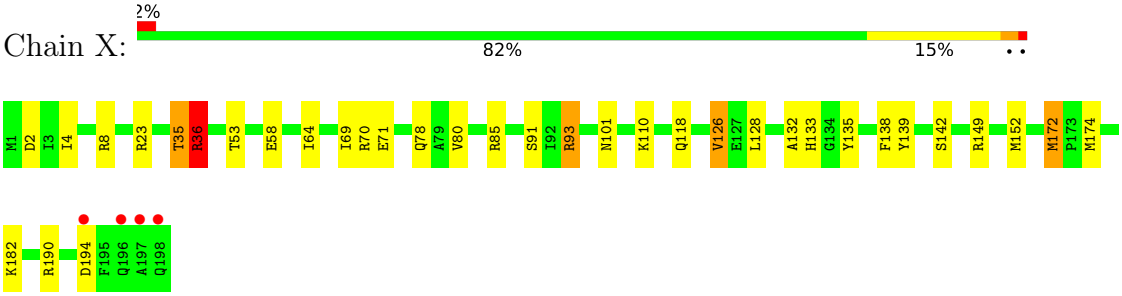
• Molecule 10: Proteasome subunit beta type-4

Chain J:



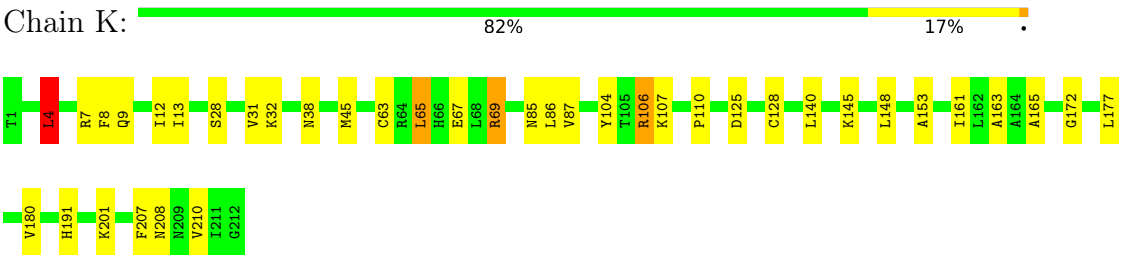
• Molecule 10: Proteasome subunit beta type-4

Chain X:



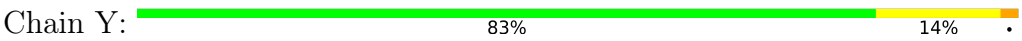
• Molecule 11: Proteasome subunit beta type-5

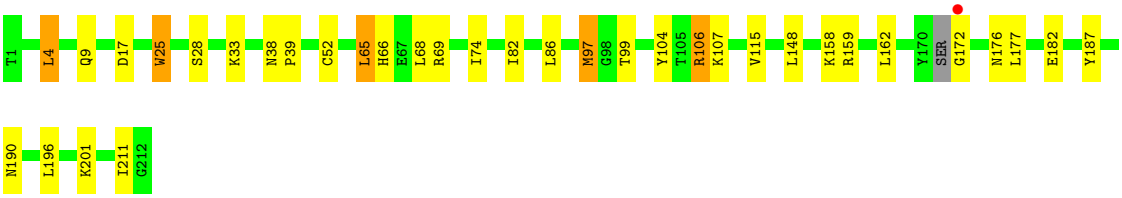
Chain K:



• Molecule 11: Proteasome subunit beta type-5

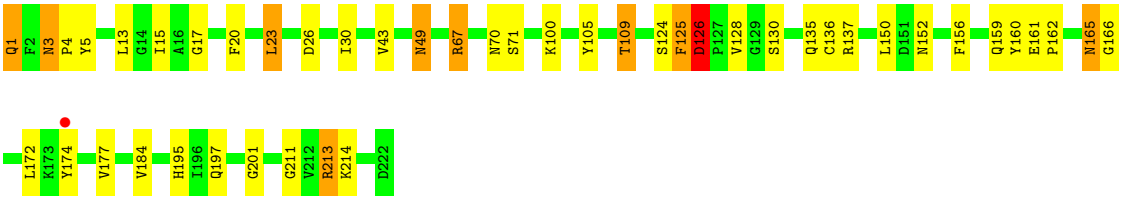
Chain Y:





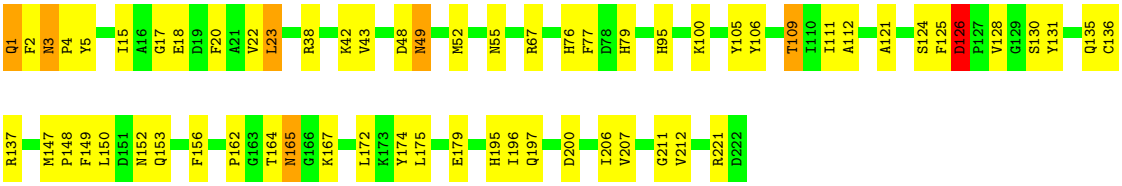
• Molecule 12: Proteasome subunit beta type-6

Chain L:



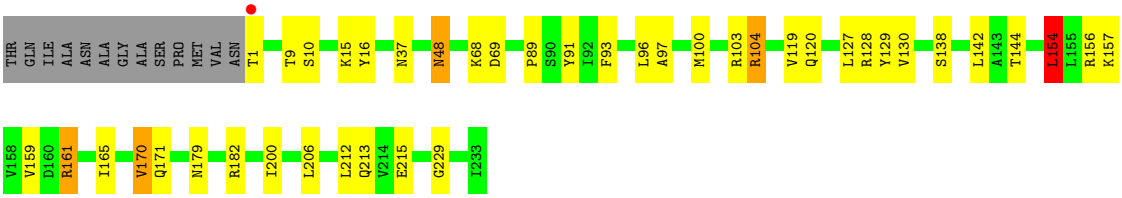
• Molecule 12: Proteasome subunit beta type-6

Chain Z:



• Molecule 13: Proteasome subunit beta type-7

Chain M:

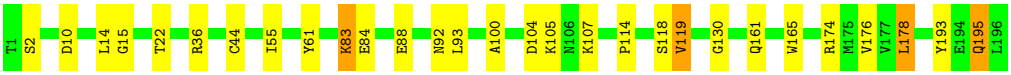


• Molecule 13: Proteasome subunit beta type-7

Chain a:

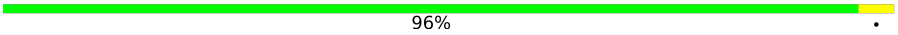
• Molecule 14: Proteasome subunit beta type-1

Chain N:



- Molecule 14: Proteasome subunit beta type-1

Chain b:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	136.44Å 301.12Å 145.54Å 90.00° 113.15° 90.00°	Depositor
Resolution (Å)	48.52 – 2.60 48.53 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.3 (48.52-2.60) 98.4 (48.53-2.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.11 (at 2.61Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.175 , 0.224 0.178 , 0.225	Depositor DCC
R_{free} test set	16212 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	47.5	Xtriage
Anisotropy	0.775	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 45.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	50912	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% 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: 2YD

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	0.61	0/1952	0.79	1/2642 (0.0%)
1	O	0.54	0/1952	0.76	0/2642
2	B	0.63	0/1934	0.79	0/2618
2	P	0.58	0/1934	0.79	2/2618 (0.1%)
3	C	0.56	0/1919	0.77	1/2598 (0.0%)
3	Q	0.52	0/1919	0.74	2/2598 (0.1%)
4	D	0.57	0/1886	0.79	1/2541 (0.0%)
4	R	0.58	0/1886	0.79	0/2541
5	E	0.57	0/1823	0.77	1/2463 (0.0%)
5	S	0.53	0/1823	0.73	2/2463 (0.1%)
6	F	0.58	0/1936	0.78	1/2614 (0.0%)
6	T	0.57	0/1936	0.78	1/2614 (0.0%)
7	G	0.62	0/1959	0.77	0/2652
7	U	0.60	0/1959	0.76	0/2652
8	H	0.61	0/1715	0.82	0/2326
8	V	0.59	0/1715	0.76	0/2326
9	I	0.66	0/1611	0.79	1/2174 (0.0%)
9	W	0.61	0/1611	0.77	0/2174
10	J	0.64	0/1613	0.84	2/2173 (0.1%)
10	X	0.60	0/1613	0.80	2/2173 (0.1%)
11	K	0.63	0/1681	0.80	3/2274 (0.1%)
11	Y	0.68	0/1674	0.87	5/2263 (0.2%)
12	L	0.64	0/1795	0.80	2/2420 (0.1%)
12	Z	0.67	0/1795	0.81	1/2420 (0.0%)
13	M	0.63	0/1855	0.87	4/2514 (0.2%)
13	a	0.65	0/1855	0.88	2/2514 (0.1%)
14	N	0.63	0/1541	0.79	0/2087
14	b	0.59	0/1541	0.75	0/2087
All	All	0.60	0/50433	0.79	34/68181 (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
5	E	0	1
8	H	0	1
8	V	0	1
12	L	0	2
12	Z	0	1
All	All	0	6

There are no bond length outliers.

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	M	104	ARG	NE-CZ-NH1	8.33	124.46	120.30
11	Y	172	GLY	N-CA-C	8.08	133.30	113.10
13	M	154	LEU	CA-CB-CG	7.04	131.49	115.30
12	Z	126	ASP	CB-CA-C	-6.60	97.20	110.40
13	M	104	ARG	NE-CZ-NH2	-6.10	117.25	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	E	201	ARG	Peptide
8	H	181	GLY	Peptide
12	L	125	PHE	Peptide
12	L	174	TYR	Peptide
8	V	181	GLY	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	31	0
1	O	1915	0	1929	30	0
2	B	1904	0	1904	42	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	35	0
3	C	1890	0	1903	35	0
3	Q	1890	0	1903	50	0
4	D	1861	0	1839	26	0
4	R	1861	0	1839	29	0
5	E	1795	0	1800	30	0
5	S	1795	0	1800	44	0
6	F	1896	0	1889	24	0
6	T	1896	0	1889	39	0
7	G	1921	0	1913	34	0
7	U	1921	0	1913	42	0
8	H	1684	0	1688	31	0
8	V	1684	0	1688	26	0
9	I	1581	0	1574	25	0
9	W	1581	0	1574	24	0
10	J	1585	0	1590	21	0
10	X	1585	0	1590	26	0
11	K	1644	0	1595	23	0
11	Y	1638	0	1589	22	0
12	L	1757	0	1711	39	0
12	Z	1757	0	1711	50	0
13	M	1824	0	1832	37	0
13	a	1824	0	1832	0	0
14	N	1512	0	1481	19	0
14	b	1512	0	1481	0	0
15	H	35	0	40	2	0
15	V	35	0	40	2	0
16	A	59	0	0	1	0
16	B	39	0	0	1	0
16	C	42	0	0	1	0
16	D	39	0	0	0	0
16	E	20	0	0	0	0
16	F	49	0	0	0	0
16	G	58	0	0	1	0
16	H	53	0	0	3	0
16	I	63	0	0	2	0
16	J	51	0	0	3	0
16	K	44	0	0	3	0
16	L	58	0	0	1	0
16	M	64	0	0	6	0
16	N	58	0	0	4	0
16	O	31	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	P	30	0	0	0	0
16	Q	25	0	0	2	0
16	R	32	0	0	4	0
16	S	21	0	0	0	0
16	T	38	0	0	2	0
16	U	66	0	0	3	0
16	V	44	0	0	1	0
16	W	62	0	0	2	0
16	X	49	0	0	3	0
16	Y	48	0	0	3	0
16	Z	48	0	0	4	0
16	a	62	0	0	0	0
16	b	57	0	0	0	0
All	All	50912	0	49370	729	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 8.

The worst 5 of 729 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)
7:U:92:ALA:HA	7:U:103:MET:HE2	1.23	1.15
7:G:92:ALA:HA	7:G:103:MET:CE	1.78	1.13
7:G:92:ALA:HA	7:G:103:MET:HE2	1.21	1.11
2:P:122:THR:HG22	3:Q:125:ARG:HH21	1.25	0.98
2:P:69:ASN:HD22	2:P:70:ASP:H	1.07	0.97

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%)	234 (94%)	13 (5%)	1 (0%)	34	57
1	O	248/250 (99%)	232 (94%)	13 (5%)	3 (1%)	13	27
2	B	242/258 (94%)	228 (94%)	10 (4%)	4 (2%)	9	18
2	P	242/258 (94%)	228 (94%)	10 (4%)	4 (2%)	9	18
3	C	239/254 (94%)	226 (95%)	12 (5%)	1 (0%)	34	57
3	Q	239/254 (94%)	223 (93%)	11 (5%)	5 (2%)	7	13
4	D	240/260 (92%)	226 (94%)	11 (5%)	3 (1%)	12	24
4	R	240/260 (92%)	220 (92%)	17 (7%)	3 (1%)	12	24
5	E	231/234 (99%)	216 (94%)	9 (4%)	6 (3%)	5	9
5	S	231/234 (99%)	210 (91%)	16 (7%)	5 (2%)	6	12
6	F	242/288 (84%)	233 (96%)	9 (4%)	0	100	100
6	T	242/288 (84%)	229 (95%)	12 (5%)	1 (0%)	34	57
7	G	241/252 (96%)	233 (97%)	8 (3%)	0	100	100
7	U	241/252 (96%)	230 (95%)	8 (3%)	3 (1%)	13	27
8	H	220/232 (95%)	214 (97%)	6 (3%)	0	100	100
8	V	220/232 (95%)	209 (95%)	10 (4%)	1 (0%)	29	52
9	I	202/205 (98%)	194 (96%)	8 (4%)	0	100	100
9	W	202/205 (98%)	193 (96%)	8 (4%)	1 (0%)	29	52
10	J	196/198 (99%)	188 (96%)	7 (4%)	1 (0%)	29	52
10	X	196/198 (99%)	187 (95%)	9 (5%)	0	100	100
11	K	210/212 (99%)	200 (95%)	10 (5%)	0	100	100
11	Y	207/212 (98%)	200 (97%)	7 (3%)	0	100	100
12	L	220/222 (99%)	208 (94%)	11 (5%)	1 (0%)	29	52
12	Z	220/222 (99%)	207 (94%)	13 (6%)	0	100	100
13	M	231/246 (94%)	225 (97%)	5 (2%)	1 (0%)	34	57
13	a	231/246 (94%)	219 (95%)	11 (5%)	1 (0%)	34	57
14	N	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
All	All	6309/6614 (95%)	5988 (95%)	276 (4%)	45 (1%)	22	43

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
5	E	2	ARG
5	E	3	ASN
2	P	51	VAL
3	Q	52	LEU

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%)	195 (93%)	14 (7%)	16	33
1	O	209/209 (100%)	198 (95%)	11 (5%)	22	45
2	B	203/216 (94%)	176 (87%)	27 (13%)	4	7
2	P	203/216 (94%)	175 (86%)	28 (14%)	3	6
3	C	213/226 (94%)	191 (90%)	22 (10%)	7	13
3	Q	213/226 (94%)	183 (86%)	30 (14%)	3	6
4	D	198/215 (92%)	180 (91%)	18 (9%)	9	18
4	R	198/215 (92%)	180 (91%)	18 (9%)	9	18
5	E	192/193 (100%)	165 (86%)	27 (14%)	3	6
5	S	192/193 (100%)	169 (88%)	23 (12%)	5	9
6	F	201/239 (84%)	178 (89%)	23 (11%)	5	10
6	T	201/239 (84%)	175 (87%)	26 (13%)	4	7
7	G	207/210 (99%)	193 (93%)	14 (7%)	16	32
7	U	207/210 (99%)	196 (95%)	11 (5%)	22	45
8	H	181/190 (95%)	164 (91%)	17 (9%)	8	17
8	V	181/190 (95%)	165 (91%)	16 (9%)	10	19
9	I	172/173 (99%)	161 (94%)	11 (6%)	17	35
9	W	172/173 (99%)	160 (93%)	12 (7%)	15	30
10	J	175/175 (100%)	167 (95%)	8 (5%)	27	51
10	X	175/175 (100%)	164 (94%)	11 (6%)	18	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	K	169/169 (100%)	154 (91%)	15 (9%)	9	19
11	Y	168/169 (99%)	159 (95%)	9 (5%)	22	44
12	L	185/185 (100%)	170 (92%)	15 (8%)	11	23
12	Z	185/185 (100%)	172 (93%)	13 (7%)	15	30
13	M	199/208 (96%)	187 (94%)	12 (6%)	19	39
13	a	199/208 (96%)	184 (92%)	15 (8%)	13	27
14	N	162/162 (100%)	152 (94%)	10 (6%)	18	37
14	b	162/162 (100%)	154 (95%)	8 (5%)	25	48
All	All	5331/5540 (96%)	4867 (91%)	464 (9%)	10	20

5 of 464 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
14	N	178	LEU
13	a	68	LYS
3	Q	130	SER
12	Z	150	LEU
9	W	114	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 154 such sidechains are listed below:

Mol	Chain	Res	Type
7	U	175	ASN
13	a	26	ASN
8	V	66	HIS
11	Y	176	ASN
14	b	38	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

2 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
15	2YD	V	301	8	36,38,38	1.45	6 (16%)	39,59,59	1.71	9 (23%)
15	2YD	H	301	8	36,38,38	1.44	6 (16%)	39,59,59	1.70	7 (17%)

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	2YD	V	301	8	-	3/23/81/81	0/3/4/4
15	2YD	H	301	8	-	3/23/81/81	0/3/4/4

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	V	301	2YD	C4-C5	2.86	1.58	1.52
15	H	301	2YD	C16-C15	2.82	1.37	1.33
15	V	301	2YD	C23-C24	2.72	1.58	1.53
15	H	301	2YD	C25-C24	2.67	1.58	1.53
15	H	301	2YD	O2-C5	2.58	1.45	1.40

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	V	301	2YD	C4-C5-C9	6.07	123.45	114.69

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	H	301	2YD	O3-C5-C4	-5.34	98.19	104.50
15	H	301	2YD	C4-C5-C9	4.60	121.32	114.69
15	V	301	2YD	O3-C5-C4	-3.85	99.95	104.50
15	H	301	2YD	C4-C3-C2	3.10	130.69	117.53

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	H	301	2YD	C1-C2-C22-O4
15	V	301	2YD	C4-C5-C9-C10
15	V	301	2YD	O3-C5-C9-C10
15	H	301	2YD	C12-C14-C15-C28
15	H	301	2YD	C12-C14-C15-C16

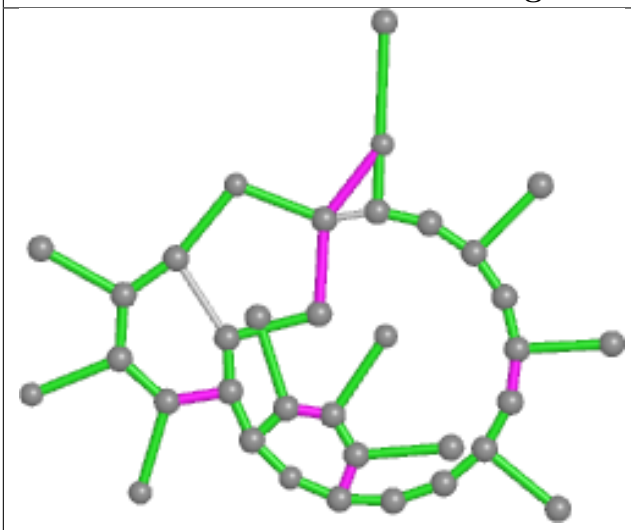
There are no ring outliers.

2 monomers are involved in 4 short contacts:

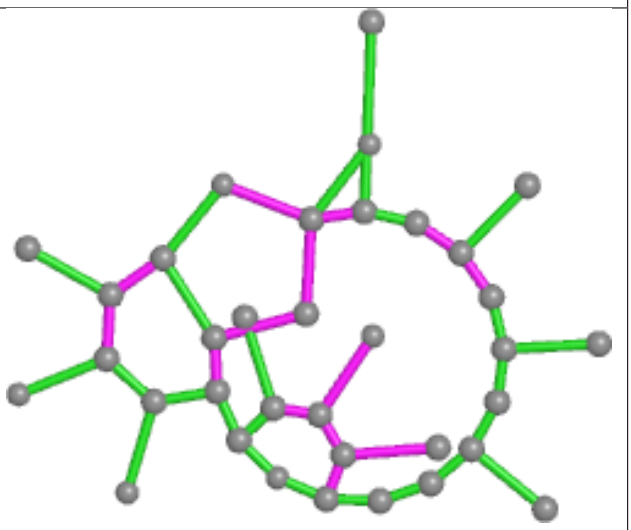
Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	V	301	2YD	2	0
15	H	301	2YD	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

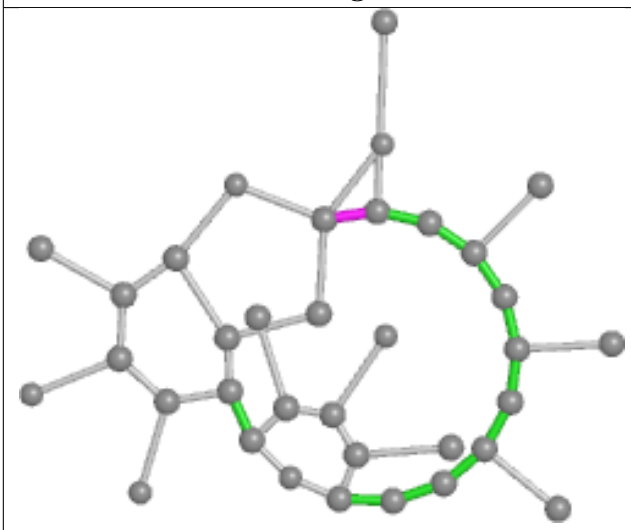
Ligand 2YD V 301



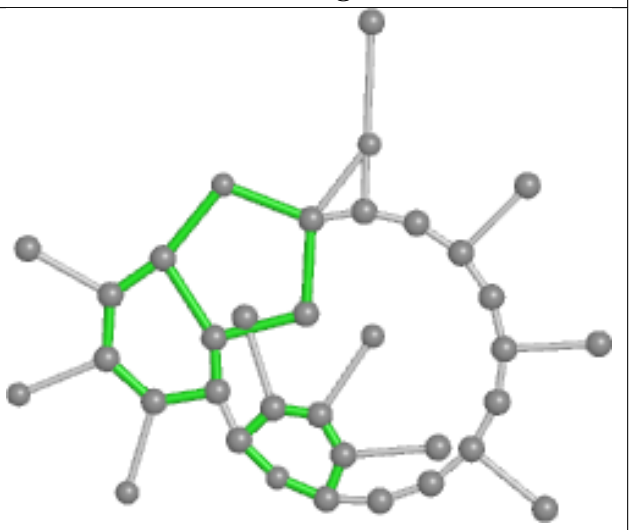
Bond lengths



Bond angles

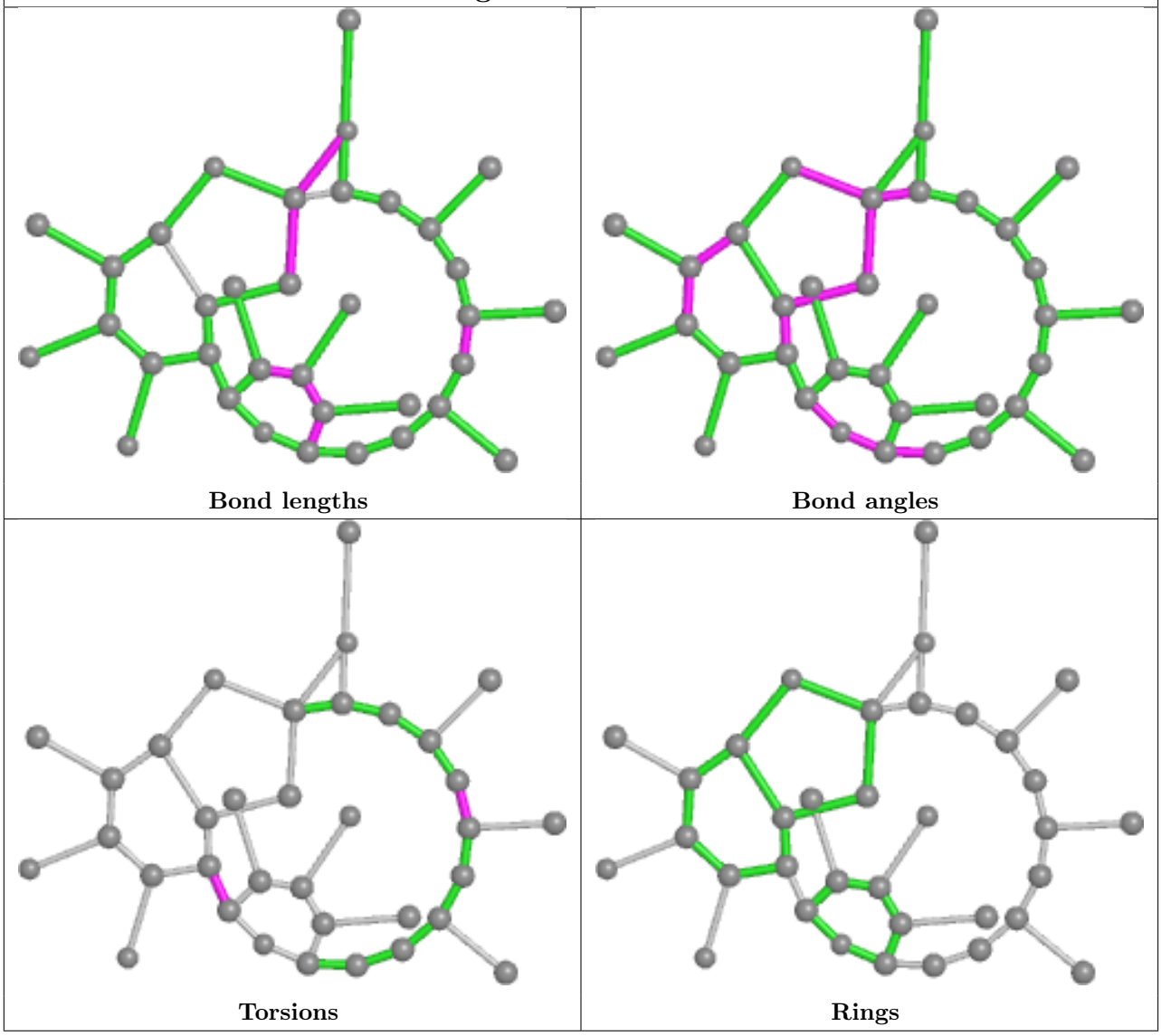


Torsions



Rings

Ligand 2YD H 301



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%)	-0.63	1 (0%)	92	91	31, 45, 76, 130	0
1	O	250/250 (100%)	-0.49	2 (0%)	86	84	35, 51, 87, 116	0
2	B	244/258 (94%)	-0.63	3 (1%)	79	76	29, 49, 90, 147	0
2	P	244/258 (94%)	-0.37	4 (1%)	72	68	37, 53, 102, 151	0
3	C	241/254 (94%)	-0.39	4 (1%)	70	66	35, 55, 114, 135	0
3	Q	241/254 (94%)	-0.12	11 (4%)	32	26	36, 59, 126, 145	0
4	D	242/260 (93%)	-0.35	6 (2%)	57	51	36, 54, 93, 186	0
4	R	242/260 (93%)	-0.40	5 (2%)	63	58	38, 55, 92, 184	0
5	E	233/234 (99%)	-0.46	3 (1%)	77	73	36, 56, 86, 137	0
5	S	233/234 (99%)	-0.30	4 (1%)	70	66	42, 61, 91, 142	0
6	F	244/288 (84%)	-0.58	2 (0%)	86	84	33, 51, 93, 126	0
6	T	244/288 (84%)	-0.36	2 (0%)	86	84	37, 56, 101, 130	0
7	G	243/252 (96%)	-0.67	2 (0%)	86	84	32, 46, 82, 156	0
7	U	243/252 (96%)	-0.62	1 (0%)	92	91	35, 49, 81, 131	0
8	H	222/232 (95%)	-0.58	1 (0%)	91	89	33, 45, 67, 122	0
8	V	222/232 (95%)	-0.64	2 (0%)	84	82	34, 48, 72, 127	0
9	I	204/205 (99%)	-0.62	1 (0%)	91	89	30, 41, 76, 107	0
9	W	204/205 (99%)	-0.49	0	100	100	34, 44, 79, 121	0
10	J	198/198 (100%)	-0.63	3 (1%)	73	70	31, 43, 71, 173	0
10	X	198/198 (100%)	-0.63	4 (2%)	65	60	33, 47, 69, 158	0
11	K	212/212 (100%)	-0.71	0	100	100	31, 43, 62, 77	0
11	Y	211/212 (99%)	-0.70	1 (0%)	91	89	29, 43, 63, 85	0
12	L	222/222 (100%)	-0.65	1 (0%)	91	89	32, 45, 69, 115	0
12	Z	222/222 (100%)	-0.62	0	100	100	32, 43, 66, 104	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%)	-0.68	1 (0%) 92 91	29, 46, 69, 101	0
13	a	233/246 (94%)	-0.66	1 (0%) 92 91	29, 44, 67, 105	0
14	N	196/196 (100%)	-0.72	0 100 100	32, 42, 65, 96	0
14	b	196/196 (100%)	-0.60	0 100 100	31, 44, 68, 90	0
All	All	6367/6614 (96%)	-0.54	65 (1%) 82 80	29, 48, 86, 186	0

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	119	ALA	15.9
4	D	121	GLY	12.0
4	R	121	GLY	11.2
4	D	120	SER	8.6
10	X	197	ALA	7.3

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

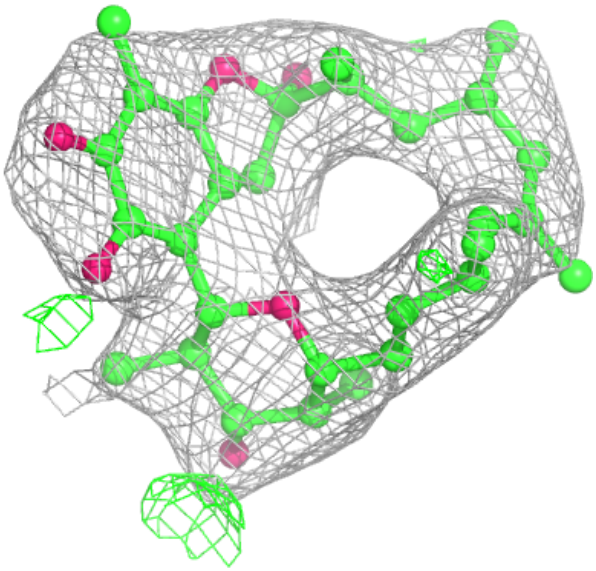
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

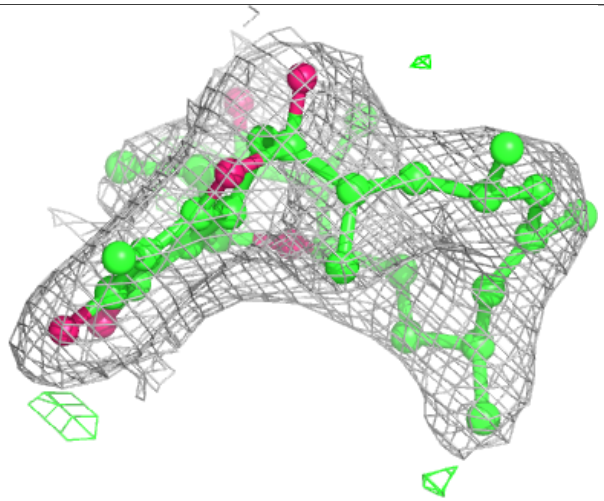
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
15	2YD	V	301	35/35	0.96	0.15	41,56,62,66	0
15	2YD	H	301	35/35	0.97	0.14	45,50,57,66	0

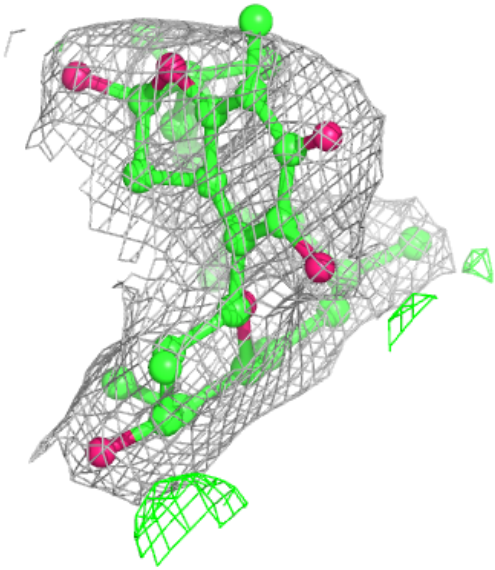
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 2YD V 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

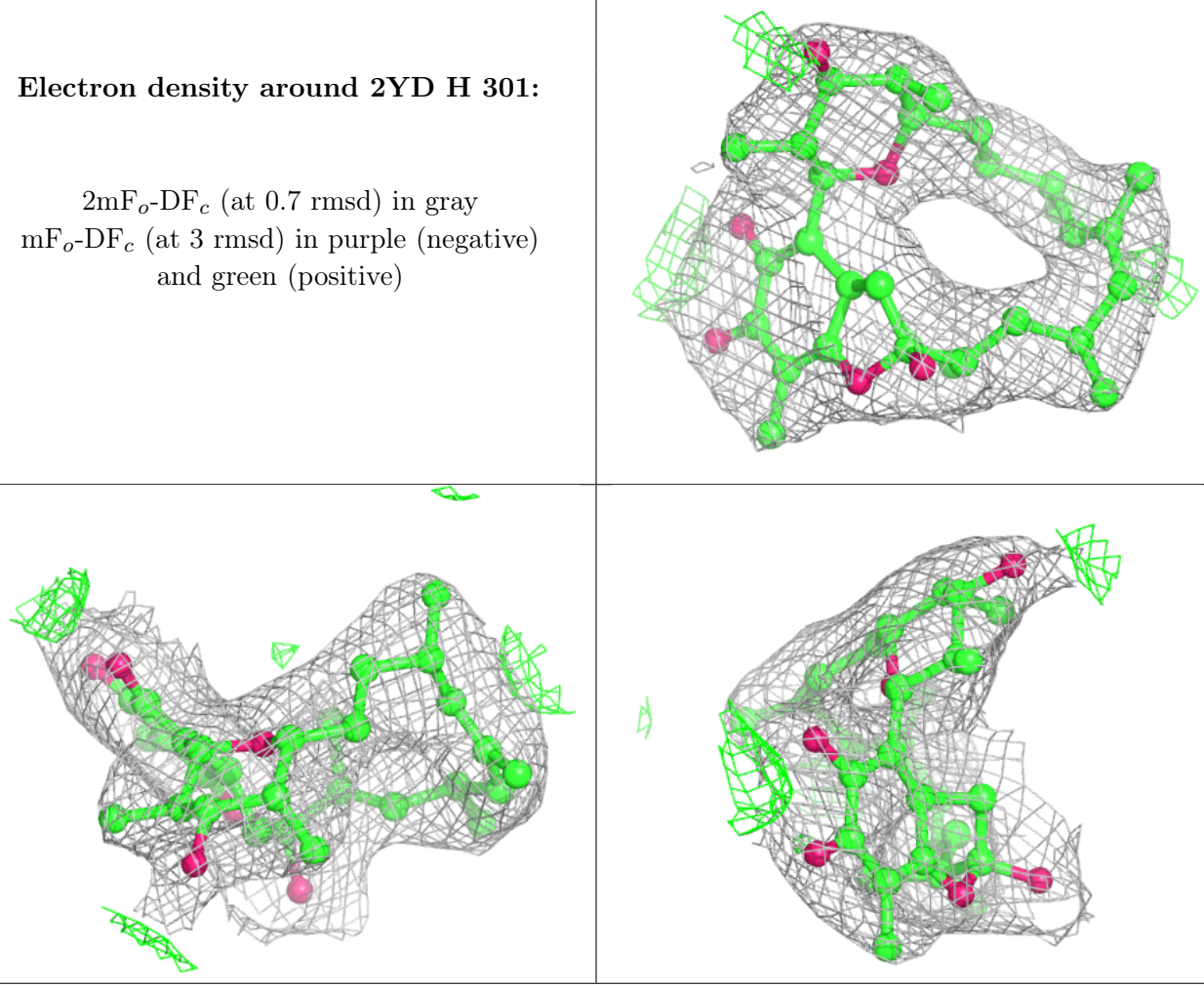






Electron density around 2YD H 301:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.