wwPDB X-ray Structure Validation Summary Report (i)

Sep 3, 2023 – 02:11 PM EDT

PDB ID	:	3SDI
Title	:	Structure of yeast 20S open-gate proteasome with Compound 20
Authors	:	Sintchak, M.D.
Deposited on	:	2011-06-09
Resolution	:	2.65 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:

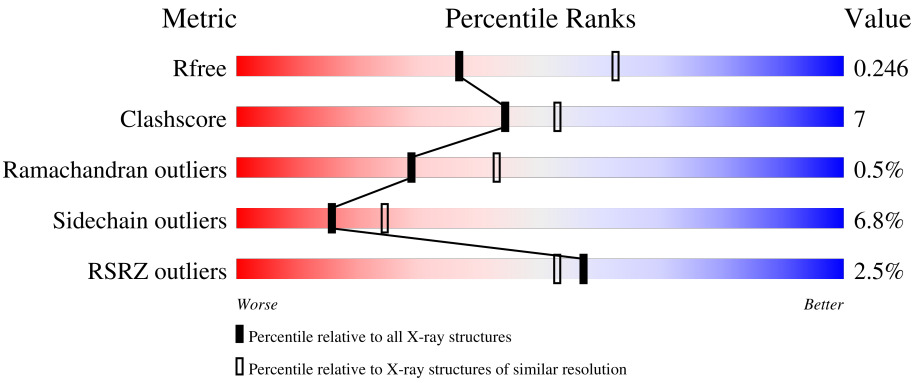
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	250	2% 86% 12% ·
1	O	250	2% 84% 14% · ·
2	B	235	6% 76% 21% ·
2	P	235	4% 75% 23% ·
3	C	241	7% 80% 16% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	241	10% 82% 14%
4	D	260	% 72% 14% 12%
4	R	260	2% 73% 13% 12%
5	E	233	3% 77% 19%
5	S	233	6% 74% 21%
6	F	242	4% 78% 18%
6	T	242	5% 76% 18% 6%
7	G	243	% 78% 19%
7	U	243	% 77% 20%
8	H	222	% 89% 10%
8	V	222	% 86% 13%
9	I	204	88% 10%
9	W	204	% 86% 13%
10	J	198	3% 82% 18%
10	X	198	3% 78% 19%
11	K	212	81% 16%
11	Y	212	% 83% 14%
12	L	222	81% 17%
12	Z	222	85% 14%
13	1	233	% 83% 15%
13	M	233	% 87% 11%
14	2	196	% 83% 15%
14	N	196	86% 13%

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	MG	F	302	-	-	-	X

2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 49012 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 component Y7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	246	Total	C	N	O	S	0	0	0
			1881	1200	308	370	3			
1	O	246	Total	C	N	O	S	0	0	0
			1881	1200	308	370	3			

- Molecule 2 is a protein called Proteasome component Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	235	Total	C	N	O	S	0	0	0
			1827	1157	303	364	3			
2	P	235	Total	C	N	O	S	0	0	0
			1827	1157	303	364	3			

- Molecule 3 is a protein called Proteasome component PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	238	Total	C	N	O	S	0	0	0
			1861	1163	325	369	4			
3	Q	238	Total	C	N	O	S	0	0	0
			1861	1163	325	369	4			

- Molecule 4 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	228	Total	C	N	O	S	0	0	0
			1747	1094	291	355	7			
4	R	229	Total	C	N	O	S	0	0	0
			1752	1097	292	356	7			

- Molecule 5 is a protein called Proteasome component PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	230	Total	C	N	O	S	0	0	0
			1755	1103	304	344	4			
5	S	230	Total	C	N	O	S	0	0	0
			1755	1103	304	344	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	127	ALA	TYR	conflict	UNP P40302
S	127	ALA	TYR	conflict	UNP P40302

- Molecule 6 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	242	Total	C	N	O	S	0	0	0
			1886	1199	328	355	4			
6	T	242	Total	C	N	O	S	0	0	0
			1886	1199	328	355	4			

- Molecule 7 is a protein called Proteasome component C7-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	240	Total	C	N	O	S	0	0	0
			1897	1206	319	364	8			
7	U	240	Total	C	N	O	S	0	0	0
			1897	1206	319	364	8			

- Molecule 8 is a protein called Proteasome component PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	222	Total	C	N	O	S	0	0	0
			1685	1061	293	324	7			
8	V	222	Total	C	N	O	S	0	0	0
			1685	1061	293	324	7			

- Molecule 9 is a protein called Proteasome component PUP3.

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 component C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	198	Total	C	N	O	S	0	0	0
			1582	1003	269	305	5			
10	X	198	Total	C	N	O	S	0	0	0
			1582	1003	269	305	5			

- Molecule 11 is a protein called Proteasome component PRE2.

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 component C5.

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 component PRE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			
13	1	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			

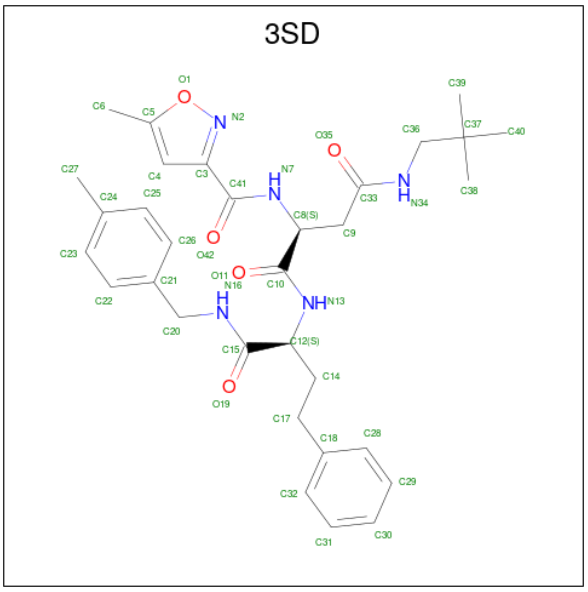
- Molecule 14 is a protein called Proteasome component PRE3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			
14	2	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

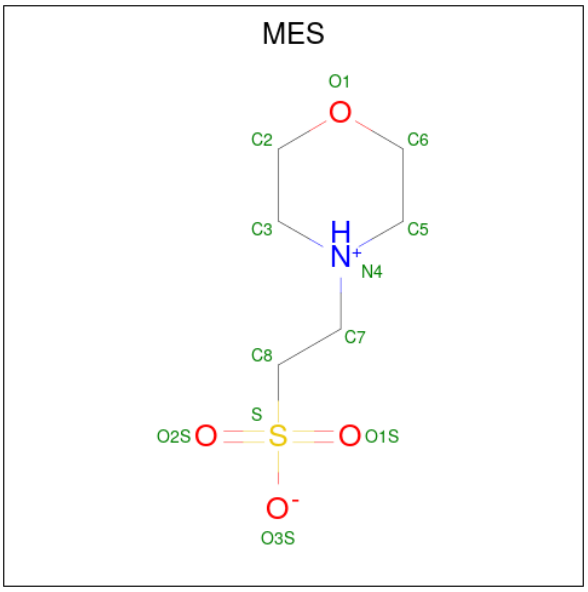
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	F	2	Total Mg 2 2	0	0
15	G	1	Total Mg 1 1	0	0
15	H	1	Total Mg 1 1	0	0
15	I	2	Total Mg 2 2	0	0
15	K	1	Total Mg 1 1	0	0
15	L	2	Total Mg 2 2	0	0
15	N	1	Total Mg 1 1	0	0
15	T	2	Total Mg 2 2	0	0
15	U	1	Total Mg 1 1	0	0
15	V	1	Total Mg 1 1	0	0
15	W	2	Total Mg 2 2	0	0
15	Y	1	Total Mg 1 1	0	0
15	Z	2	Total Mg 2 2	0	0
15	2	1	Total Mg 1 1	0	0

- Molecule 16 is N 4 -(2,2-dimethylpropyl)-N 1 -{(2S)-1-[(4-methylbenzyl)amino]-1-oxo-4-phenylbutan-2-yl}-N 2 -[(5-methyl-1,2-oxazol-3-yl)carbonyl]-L-aspartamide (three-letter code: 3SD) (formula: $C_{32}H_{41}N_5O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
16	K	1	Total	C	N	O	0	0
			42	32	5	5		
16	Y	1	Total	C	N	O	0	0
			42	32	5	5		

- Molecule 17 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
17	K	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
17	Y	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

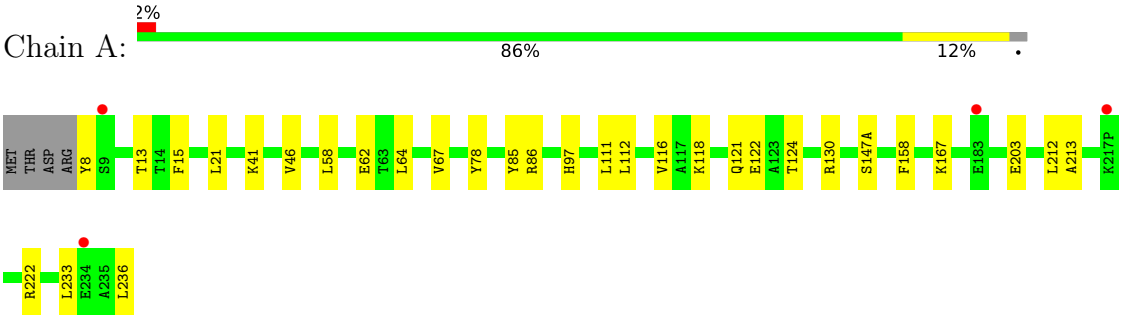
- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	L	1	Total	O	0	0
			1	1		

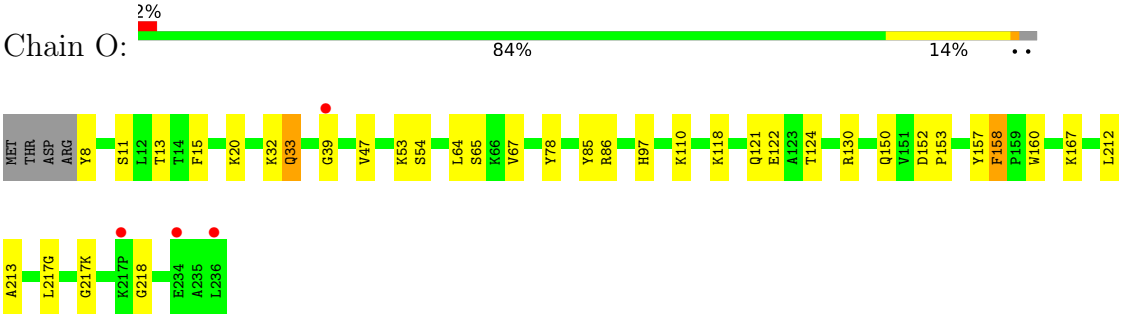
3 Residue-property plots (i)

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.

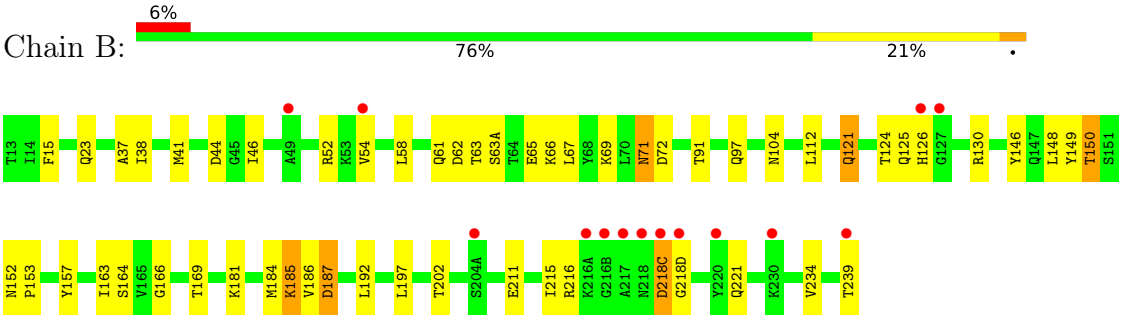
• Molecule 1: Proteasome component Y7



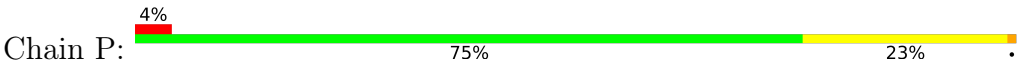
• Molecule 1: Proteasome component Y7

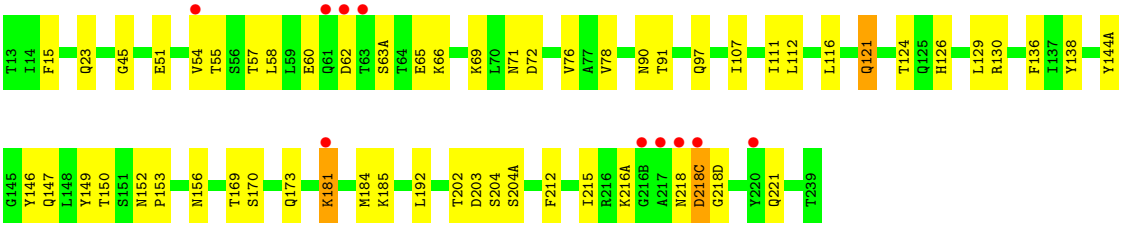


• Molecule 2: Proteasome component Y13

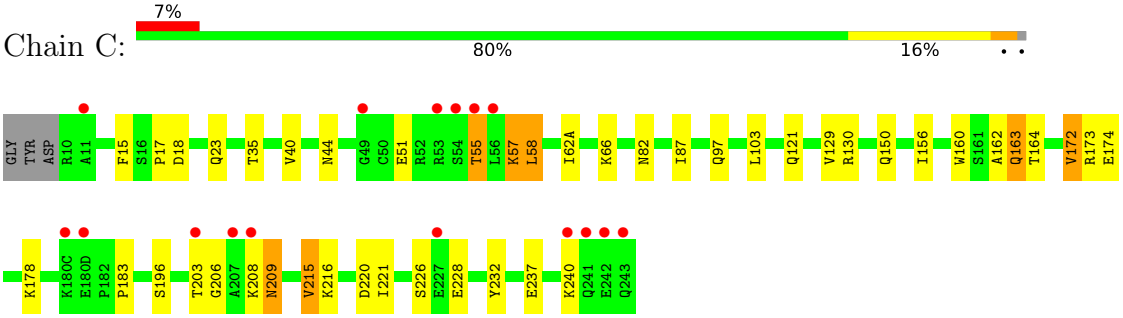


• Molecule 2: Proteasome component Y13

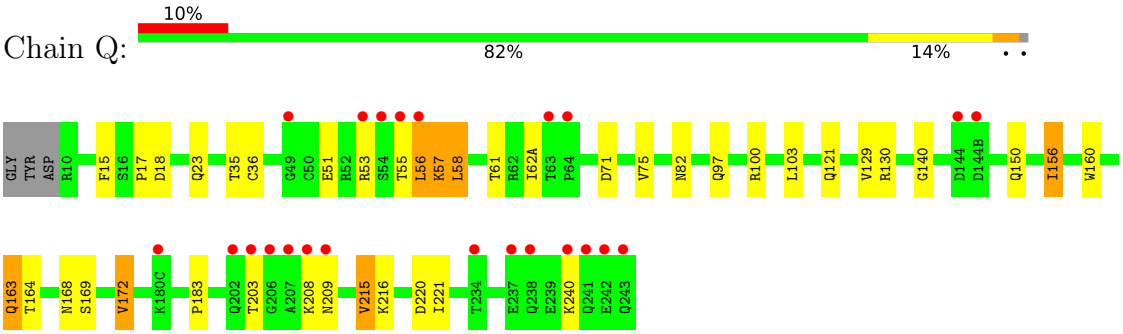




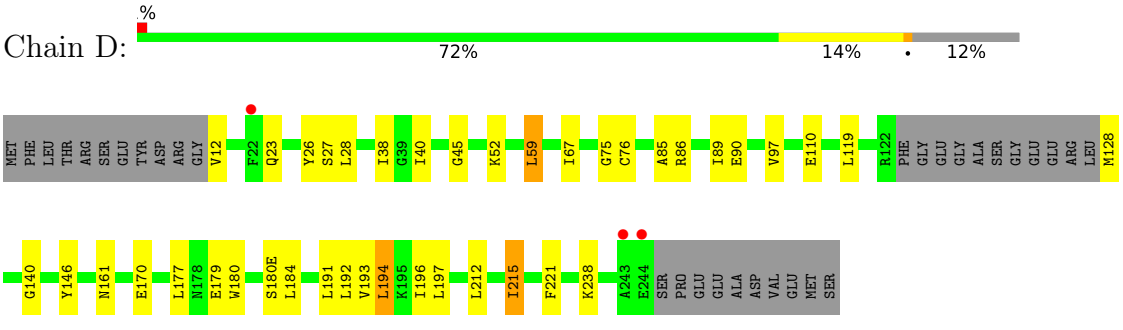
• Molecule 3: Proteasome component PRE6



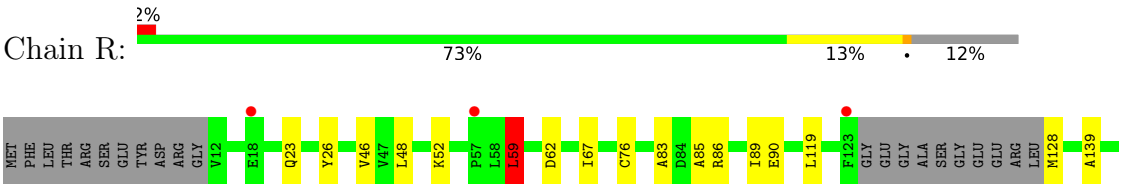
• Molecule 3: Proteasome component PRE6

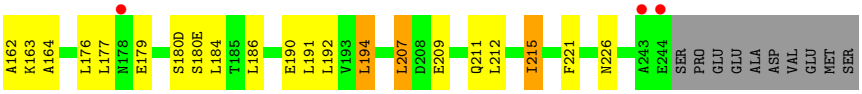


• Molecule 4: Proteasome component PUP2

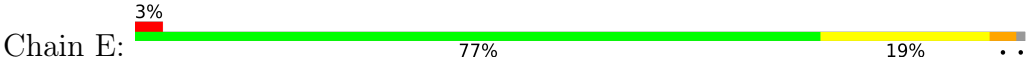


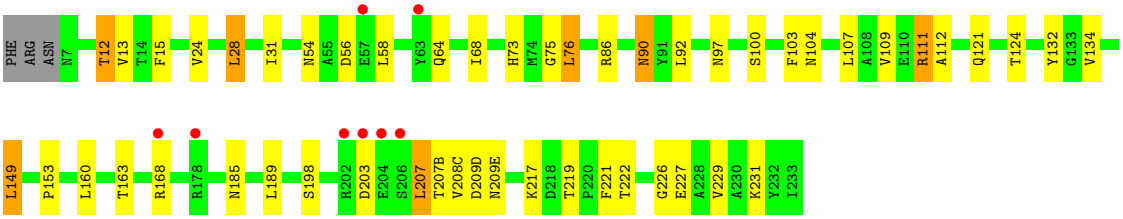
• Molecule 4: Proteasome component PUP2



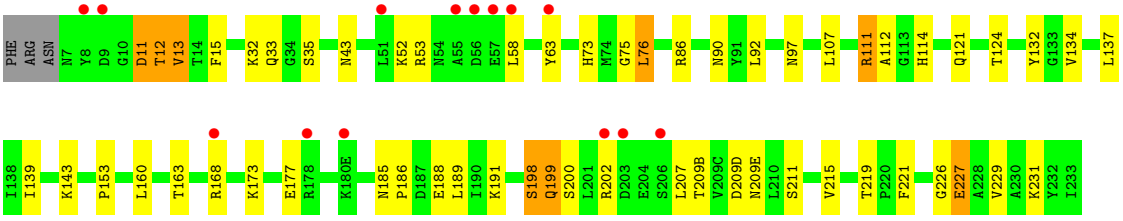


• Molecule 5: Proteasome component PRE5

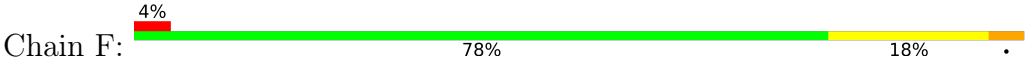


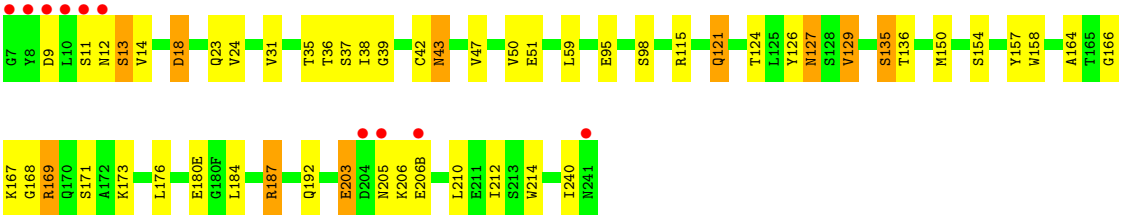


• Molecule 5: Proteasome component PRE5

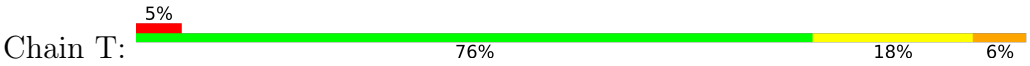


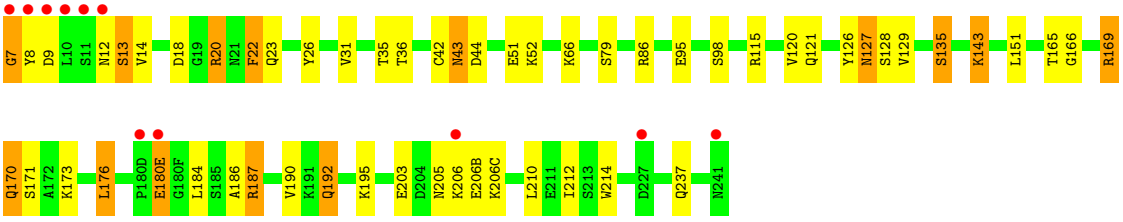
• Molecule 6: Proteasome component C1





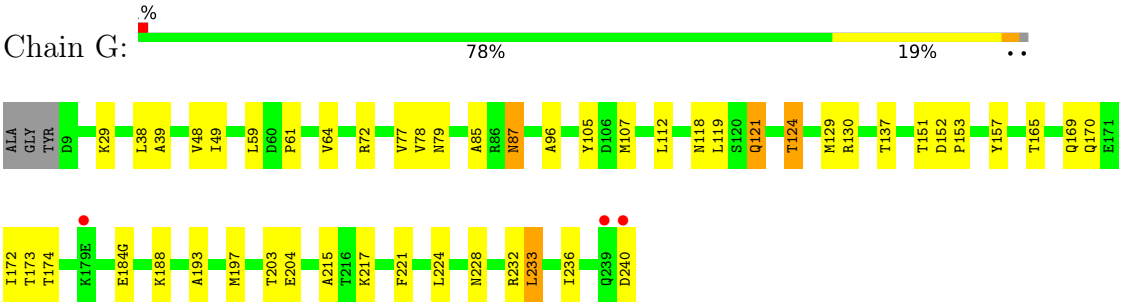
• Molecule 6: Proteasome component C1





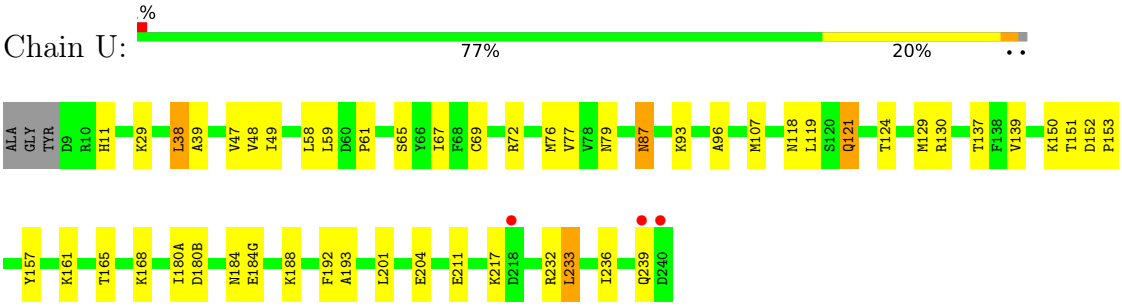
• Molecule 7: Proteasome component C7-alpha

Chain G:



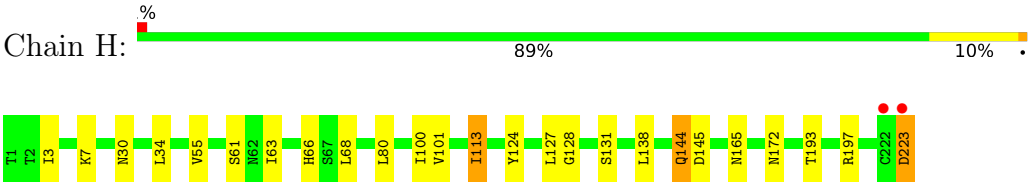
• Molecule 7: Proteasome component C7-alpha

Chain U:



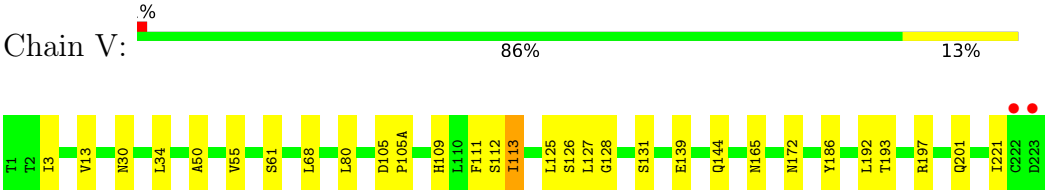
• Molecule 8: Proteasome component PUP1

Chain H:



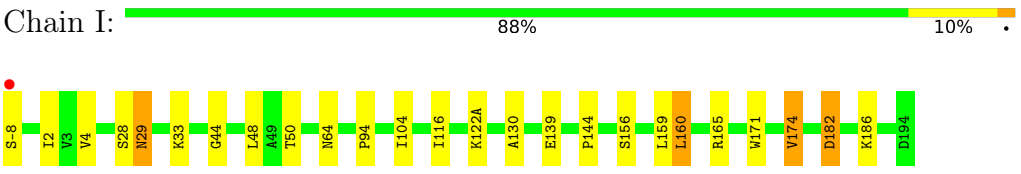
• Molecule 8: Proteasome component PUP1

Chain V:



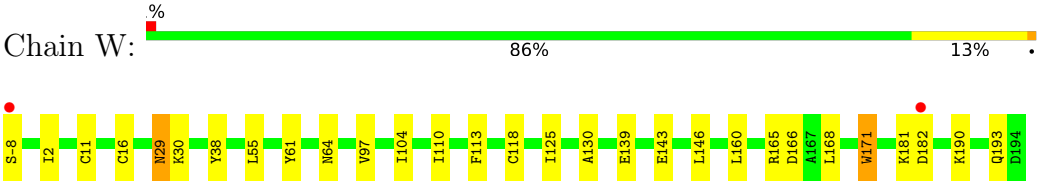
• Molecule 9: Proteasome component PUP3

Chain I:



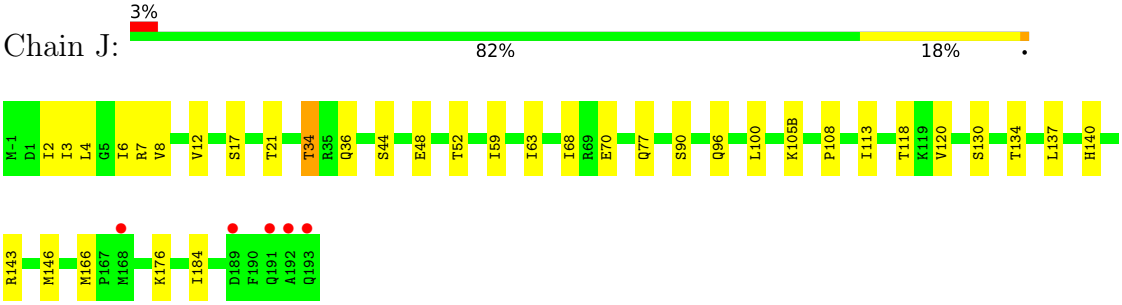
• Molecule 9: Proteasome component PUP3

Chain W:



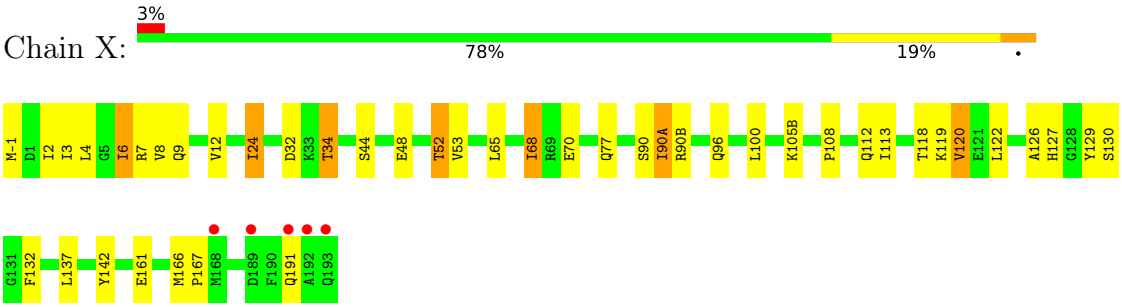
• Molecule 10: Proteasome component C11

Chain J:



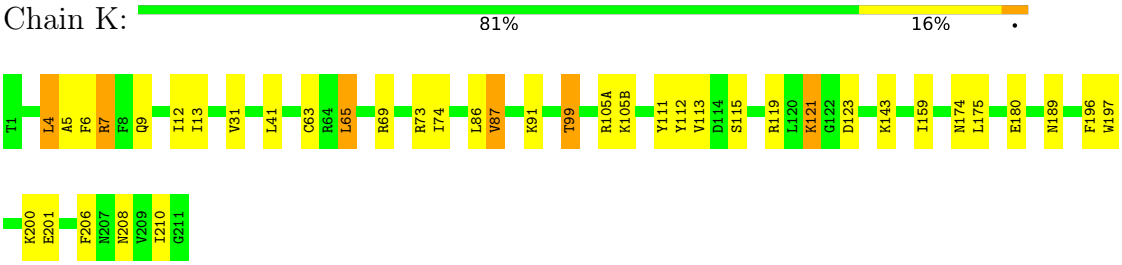
• Molecule 10: Proteasome component C11

Chain X:



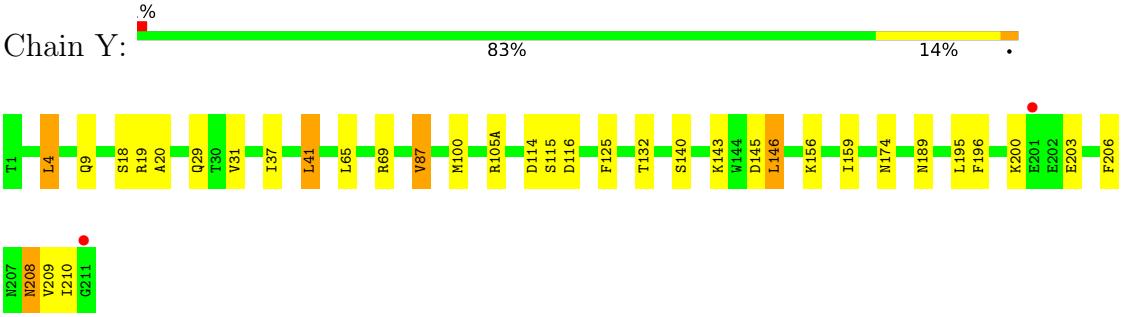
• Molecule 11: Proteasome component PRE2

Chain K:



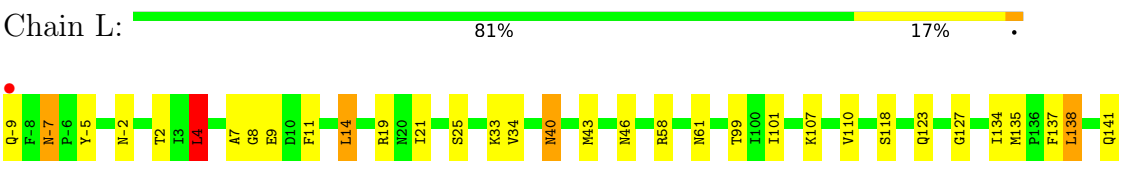
• Molecule 11: Proteasome component PRE2

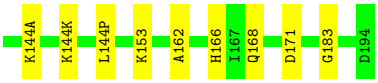
Chain Y:



• Molecule 12: Proteasome component C5

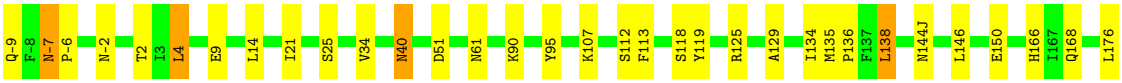
Chain L:





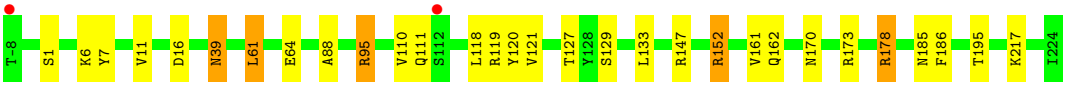
• Molecule 12: Proteasome component C5

Chain Z:



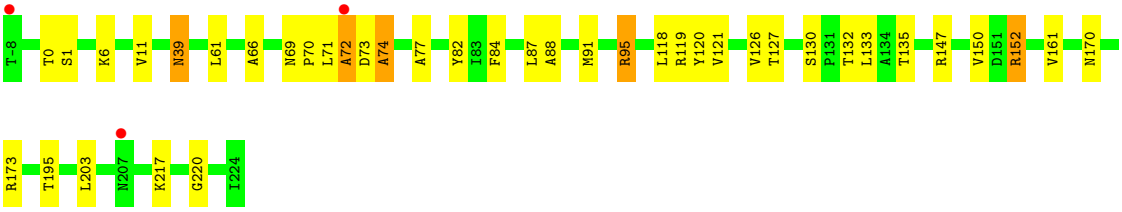
• Molecule 13: Proteasome component PRE4

Chain M:



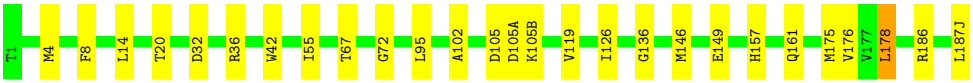
• Molecule 13: Proteasome component PRE4

Chain 1:



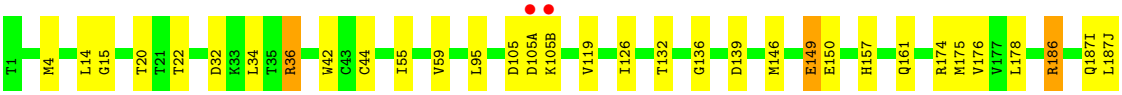
• Molecule 14: Proteasome component PRE3

Chain N:



• Molecule 14: Proteasome component PRE3

Chain 2:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	136.60Å 299.45Å 145.48Å 90.00° 113.17° 90.00°	Depositor
Resolution (Å)	30.00 – 2.65 29.95 – 2.65	Depositor EDS
% Data completeness (in resolution range)	95.4 (30.00-2.65) 95.5 (29.95-2.65)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.44 (at 2.64Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.219 , 0.253 0.215 , 0.246	Depositor DCC
R_{free} test set	5996 reflections (2.02%)	wwPDB-VP
Wilson B-factor (Å^2)	49.8	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 28.9	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	0.93	EDS
Total number of atoms	49012	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.25% 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: 3SD, MES, 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	0.50	0/1918	0.60	0/2597
1	O	0.49	0/1918	0.59	0/2597
2	B	0.46	0/1856	0.61	0/2513
2	P	0.48	0/1856	0.61	0/2513
3	C	0.46	0/1889	0.59	1/2557 (0.0%)
3	Q	0.47	0/1889	0.61	2/2557 (0.1%)
4	D	0.51	0/1770	0.66	1/2387 (0.0%)
4	R	0.47	0/1775	0.63	2/2394 (0.1%)
5	E	0.49	0/1781	0.61	0/2407
5	S	0.47	0/1781	0.60	0/2407
6	F	0.50	0/1926	0.59	0/2599
6	T	0.54	1/1926 (0.1%)	0.63	0/2599
7	G	0.52	0/1934	0.60	0/2618
7	U	0.52	0/1934	0.62	0/2618
8	H	0.51	0/1716	0.64	0/2326
8	V	0.49	0/1716	0.61	0/2326
9	I	0.58	0/1611	0.63	0/2174
9	W	0.64	1/1611 (0.1%)	0.63	0/2174
10	J	0.52	0/1610	0.66	0/2170
10	X	0.54	0/1610	0.68	2/2170 (0.1%)
11	K	0.51	0/1681	0.65	0/2274
11	Y	0.49	0/1681	0.63	0/2274
12	L	0.56	0/1795	0.63	1/2420 (0.0%)
12	Z	0.56	0/1795	0.62	1/2420 (0.0%)
13	1	0.60	0/1855	0.76	1/2514 (0.0%)
13	M	0.63	0/1855	0.70	1/2514 (0.0%)
14	2	0.59	0/1541	0.63	0/2087
14	N	0.62	0/1541	0.61	0/2087
All	All	0.53	2/49771 (0.0%)	0.63	12/67293 (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
13	1	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	W	38	TYR	CD1-CE1	-5.17	1.31	1.39
6	T	7	GLY	N-CA	5.01	1.53	1.46

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	1	74	ALA	O-C-N	-16.84	95.75	122.70
4	D	128	MET	N-CA-C	-7.50	90.76	111.00
10	X	-1	MET	O-C-N	-6.45	112.38	122.70
10	X	-1	MET	C-N-CA	6.29	137.43	121.70
4	R	59	LEU	CA-CB-CG	6.21	129.59	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
13	1	74	ALA	Mainchain

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	1881	0	1893	24	0
1	O	1881	0	1893	26	0
2	B	1827	0	1824	36	0
2	P	1827	0	1824	39	0
3	C	1861	0	1873	31	0
3	Q	1861	0	1873	25	0
4	D	1747	0	1718	22	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	R	1752	0	1720	18	0
5	E	1755	0	1761	27	0
5	S	1755	0	1761	34	0
6	F	1886	0	1876	36	0
6	T	1886	0	1875	50	0
7	G	1897	0	1891	44	0
7	U	1897	0	1891	42	0
8	H	1685	0	1688	15	0
8	V	1685	0	1688	22	0
9	I	1581	0	1574	15	0
9	W	1581	0	1574	24	0
10	J	1582	0	1583	17	0
10	X	1582	0	1582	24	0
11	K	1644	0	1595	24	0
11	Y	1644	0	1595	21	0
12	L	1757	0	1711	23	0
12	Z	1757	0	1711	24	0
13	1	1824	0	1832	35	0
13	M	1824	0	1832	20	0
14	2	1512	0	1481	21	0
14	N	1512	0	1481	17	0
15	2	1	0	0	0	0
15	F	2	0	0	0	0
15	G	1	0	0	0	0
15	H	1	0	0	0	0
15	I	2	0	0	0	0
15	K	1	0	0	0	0
15	L	2	0	0	0	0
15	N	1	0	0	0	0
15	T	2	0	0	0	0
15	U	1	0	0	0	0
15	V	1	0	0	0	0
15	W	2	0	0	0	0
15	Y	1	0	0	0	0
15	Z	2	0	0	0	0
16	K	42	0	41	1	0
16	Y	42	0	41	2	0
17	K	12	0	13	1	0
17	Y	12	0	13	0	0
18	L	1	0	0	0	0
All	All	49012	0	48708	641	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 7.

The worst 5 of 641 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)
14:N:161:GLN:HE21	14:2:136:GLY:HA2	1.17	1.09
7:G:96:ALA:HA	7:G:107:MET:HE2	1.31	1.06
7:U:96:ALA:HA	7:U:107:MET:HE2	1.36	1.04
3:C:163:GLN:NE2	3:C:164:THR:H	1.59	1.00
1:A:130:ARG:HH21	7:G:124:THR:CG2	1.76	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	244/250 (98%)	239 (98%)	4 (2%)	1 (0%)	34	48
1	O	244/250 (98%)	232 (95%)	10 (4%)	2 (1%)	19	29
2	B	233/235 (99%)	218 (94%)	12 (5%)	3 (1%)	12	18
2	P	233/235 (99%)	217 (93%)	12 (5%)	4 (2%)	9	13
3	C	236/241 (98%)	227 (96%)	5 (2%)	4 (2%)	9	13
3	Q	236/241 (98%)	226 (96%)	7 (3%)	3 (1%)	12	18
4	D	224/260 (86%)	215 (96%)	8 (4%)	1 (0%)	34	48
4	R	225/260 (86%)	216 (96%)	8 (4%)	1 (0%)	34	48
5	E	228/233 (98%)	214 (94%)	12 (5%)	2 (1%)	17	26
5	S	228/233 (98%)	211 (92%)	14 (6%)	3 (1%)	12	18
6	F	240/242 (99%)	228 (95%)	11 (5%)	1 (0%)	34	48
6	T	240/242 (99%)	225 (94%)	12 (5%)	3 (1%)	12	18
7	G	238/243 (98%)	225 (94%)	13 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	U	238/243 (98%)	229 (96%)	9 (4%)	0	100	100
8	H	220/222 (99%)	209 (95%)	11 (5%)	0	100	100
8	V	220/222 (99%)	210 (96%)	10 (4%)	0	100	100
9	I	202/204 (99%)	196 (97%)	6 (3%)	0	100	100
9	W	202/204 (99%)	192 (95%)	10 (5%)	0	100	100
10	J	196/198 (99%)	188 (96%)	7 (4%)	1 (0%)	29	43
10	X	196/198 (99%)	188 (96%)	7 (4%)	1 (0%)	29	43
11	K	210/212 (99%)	204 (97%)	5 (2%)	1 (0%)	29	43
11	Y	210/212 (99%)	201 (96%)	9 (4%)	0	100	100
12	L	220/222 (99%)	214 (97%)	6 (3%)	0	100	100
12	Z	220/222 (99%)	217 (99%)	3 (1%)	0	100	100
13	1	231/233 (99%)	221 (96%)	8 (4%)	2 (1%)	17	26
13	M	231/233 (99%)	221 (96%)	9 (4%)	1 (0%)	34	48
14	2	194/196 (99%)	186 (96%)	8 (4%)	0	100	100
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
All	All	6233/6382 (98%)	5958 (96%)	241 (4%)	34 (0%)	29	43

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	58	LEU
4	D	180(E)	SER
11	K	180	GLU
2	P	218(C)	ASP
3	Q	58	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	205/209 (98%)	199 (97%)	6 (3%)	42	60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	O	205/209 (98%)	197 (96%)	8 (4%)	32 48
2	B	194/195 (100%)	179 (92%)	15 (8%)	13 20
2	P	194/195 (100%)	179 (92%)	15 (8%)	13 20
3	C	210/213 (99%)	192 (91%)	18 (9%)	10 15
3	Q	210/213 (99%)	195 (93%)	15 (7%)	14 22
4	D	186/215 (86%)	171 (92%)	15 (8%)	11 17
4	R	186/215 (86%)	173 (93%)	13 (7%)	15 23
5	E	187/191 (98%)	167 (89%)	20 (11%)	6 9
5	S	187/191 (98%)	167 (89%)	20 (11%)	6 9
6	F	200/200 (100%)	179 (90%)	21 (10%)	7 10
6	T	200/200 (100%)	179 (90%)	21 (10%)	7 10
7	G	205/207 (99%)	190 (93%)	15 (7%)	14 21
7	U	205/207 (99%)	193 (94%)	12 (6%)	19 30
8	H	181/181 (100%)	171 (94%)	10 (6%)	21 33
8	V	181/181 (100%)	173 (96%)	8 (4%)	28 43
9	I	172/172 (100%)	161 (94%)	11 (6%)	17 27
9	W	172/172 (100%)	165 (96%)	7 (4%)	30 46
10	J	174/175 (99%)	164 (94%)	10 (6%)	20 31
10	X	174/175 (99%)	158 (91%)	16 (9%)	9 13
11	K	169/169 (100%)	153 (90%)	16 (10%)	8 12
11	Y	169/169 (100%)	159 (94%)	10 (6%)	19 30
12	L	185/185 (100%)	172 (93%)	13 (7%)	15 23
12	Z	185/185 (100%)	176 (95%)	9 (5%)	25 38
13	1	199/199 (100%)	191 (96%)	8 (4%)	31 47
13	M	199/199 (100%)	189 (95%)	10 (5%)	24 38
14	2	162/162 (100%)	151 (93%)	11 (7%)	16 24
14	N	162/162 (100%)	155 (96%)	7 (4%)	29 44
All	All	5258/5346 (98%)	4898 (93%)	360 (7%)	16 24

5 of 360 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	R	177	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
7	U	204	GLU
5	S	11	ASP
6	T	36	THR
9	W	16	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 166 such sidechains are listed below:

Mol	Chain	Res	Type
7	U	11	HIS
11	Y	174	ASN
7	U	121	GLN
8	V	165	ASN
12	Z	70(A)	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)

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)

Of 24 ligands modelled in this entry, 20 are monoatomic - leaving 4 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
16	3SD	K	302	-	42,44,44	1.07	3 (7%)	54,60,60	1.04	3 (5%)
17	MES	Y	303	-	12,12,12	2.32	1 (8%)	14,16,16	1.39	2 (14%)
16	3SD	Y	302	-	42,44,44	1.17	3 (7%)	54,60,60	0.86	1 (1%)
17	MES	K	303	-	12,12,12	2.04	1 (8%)	14,16,16	1.74	3 (21%)

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
16	3SD	K	302	-	-	0/37/40/40	0/3/3/3
17	MES	Y	303	-	-	0/6/14/14	0/1/1/1
16	3SD	Y	302	-	-	1/37/40/40	0/3/3/3
17	MES	K	303	-	-	0/6/14/14	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	Y	303	MES	C8-S	-7.76	1.66	1.77
17	K	303	MES	C8-S	-6.75	1.67	1.77
16	Y	302	3SD	C6-C5	4.20	1.53	1.48
16	Y	302	3SD	C4-C5	-4.16	1.33	1.39
16	K	302	3SD	C6-C5	3.98	1.53	1.48

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
17	K	303	MES	O3S-S-C8	4.52	113.07	105.77
17	Y	303	MES	O3S-S-C8	3.46	111.36	105.77
17	K	303	MES	O1S-S-C8	3.06	110.60	106.92
16	K	302	3SD	C9-C8-C10	-2.85	103.70	110.42
16	K	302	3SD	C9-C33-N34	-2.73	112.20	115.97

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	Y	302	3SD	C4-C3-C41-O42

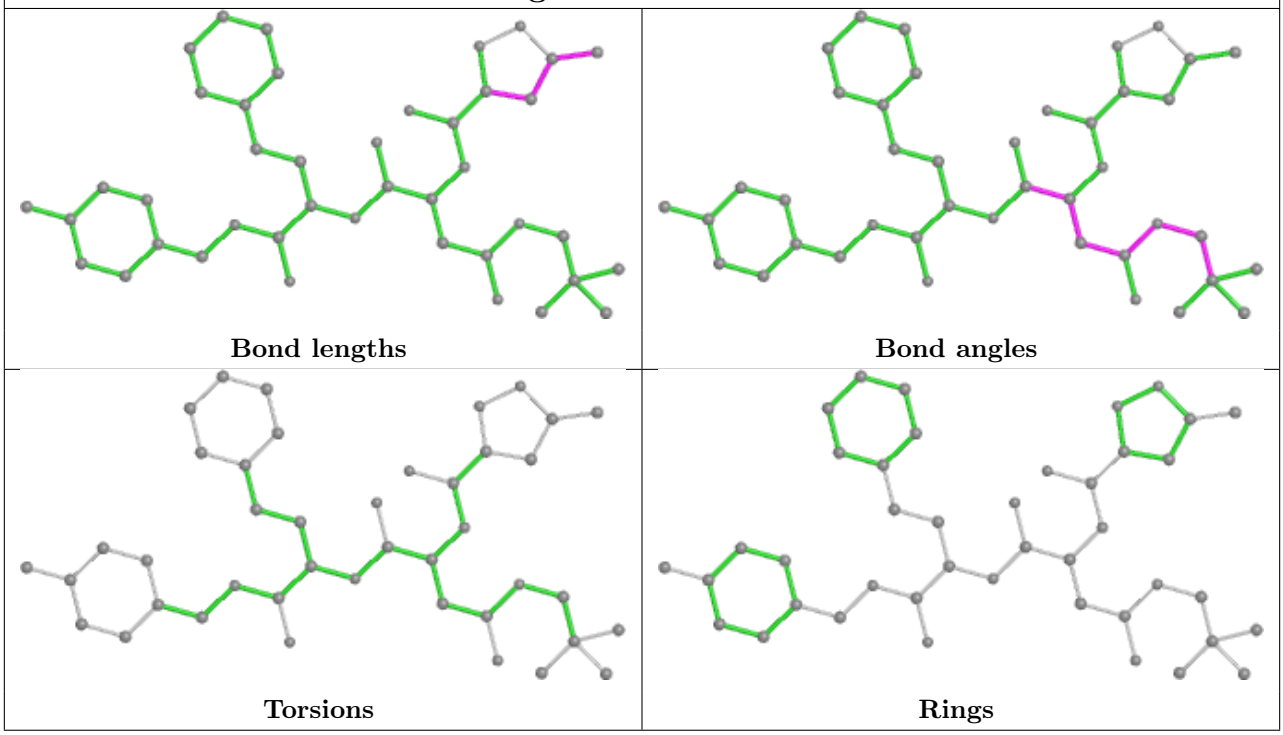
There are no ring outliers.

3 monomers are involved in 3 short contacts:

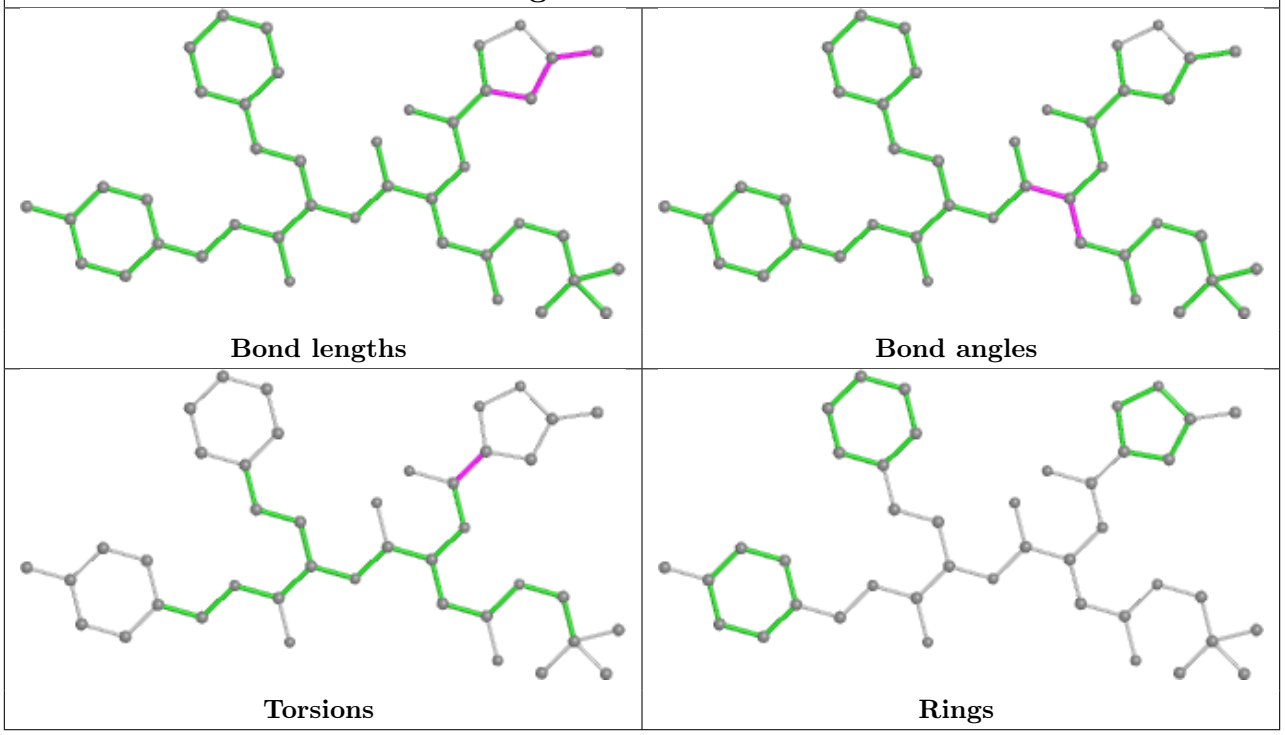
Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	K	302	3SD	1	0
16	Y	302	3SD	2	0
17	K	303	MES	1	0

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 3SD K 302



Ligand 3SD Y 302



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

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	246/250 (98%)	-0.29	4 (1%) 72 69	31, 43, 64, 89	0
1	O	246/250 (98%)	-0.08	4 (1%) 72 69	32, 50, 74, 94	0
2	B	235/235 (100%)	0.13	14 (5%) 21 18	26, 51, 82, 90	0
2	P	235/235 (100%)	0.14	10 (4%) 35 31	27, 51, 85, 94	0
3	C	238/241 (98%)	0.21	16 (6%) 17 14	31, 59, 103, 118	0
3	Q	238/241 (98%)	0.53	23 (9%) 7 6	34, 63, 111, 127	0
4	D	228/260 (87%)	-0.20	3 (1%) 77 75	28, 51, 73, 95	0
4	R	229/260 (88%)	-0.04	6 (2%) 56 52	29, 53, 77, 102	0
5	E	230/233 (98%)	-0.06	8 (3%) 44 40	37, 52, 78, 92	0
5	S	230/233 (98%)	0.18	14 (6%) 21 18	34, 54, 89, 99	0
6	F	242/242 (100%)	0.05	10 (4%) 37 33	29, 49, 85, 109	0
6	T	242/242 (100%)	0.20	11 (4%) 33 30	29, 48, 76, 109	0
7	G	240/243 (98%)	-0.26	3 (1%) 77 75	28, 42, 68, 99	0
7	U	240/243 (98%)	-0.38	3 (1%) 77 75	26, 42, 63, 85	0
8	H	222/222 (100%)	-0.33	2 (0%) 84 83	27, 37, 56, 85	0
8	V	222/222 (100%)	-0.30	2 (0%) 84 83	29, 41, 59, 87	0
9	I	204/204 (100%)	-0.42	1 (0%) 91 91	25, 37, 53, 61	0
9	W	204/204 (100%)	-0.37	2 (0%) 82 81	28, 38, 56, 62	0
10	J	198/198 (100%)	-0.14	5 (2%) 57 53	28, 41, 57, 114	0
10	X	198/198 (100%)	-0.16	5 (2%) 57 53	30, 42, 55, 115	0
11	K	212/212 (100%)	-0.33	0 100 100	26, 37, 54, 63	0
11	Y	212/212 (100%)	-0.30	2 (0%) 84 83	29, 39, 57, 61	0
12	L	222/222 (100%)	-0.40	1 (0%) 91 91	27, 39, 58, 65	0
12	Z	222/222 (100%)	-0.37	0 100 100	24, 38, 57, 65	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
13	1	233/233 (100%)	-0.36	3 (1%) 77 75	23, 35, 53, 62	0
13	M	233/233 (100%)	-0.40	2 (0%) 84 83	25, 39, 55, 61	0
14	2	196/196 (100%)	-0.34	2 (1%) 82 81	27, 35, 52, 69	0
14	N	196/196 (100%)	-0.39	0 100 100	26, 36, 52, 64	0
All	All	6293/6382 (98%)	-0.15	156 (2%) 57 53	23, 44, 80, 127	0

The worst 5 of 156 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	T	10	LEU	17.3
6	T	8	TYR	16.6
6	T	9	ASP	15.1
6	T	11	SER	13.2
6	F	9	ASP	12.6

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	MG	L	201	1/1	0.59	0.11	52,52,52,52	0
15	MG	F	302	1/1	0.69	0.60	120,120,120,120	0
15	MG	T	301	1/1	0.77	0.32	69,69,69,69	0
15	MG	W	201	1/1	0.77	0.14	53,53,53,53	0
15	MG	Z	201	1/1	0.77	0.11	65,65,65,65	0
15	MG	F	301	1/1	0.84	0.79	74,74,74,74	0
15	MG	I	202	1/1	0.84	0.27	50,50,50,50	0

Continued on next page...

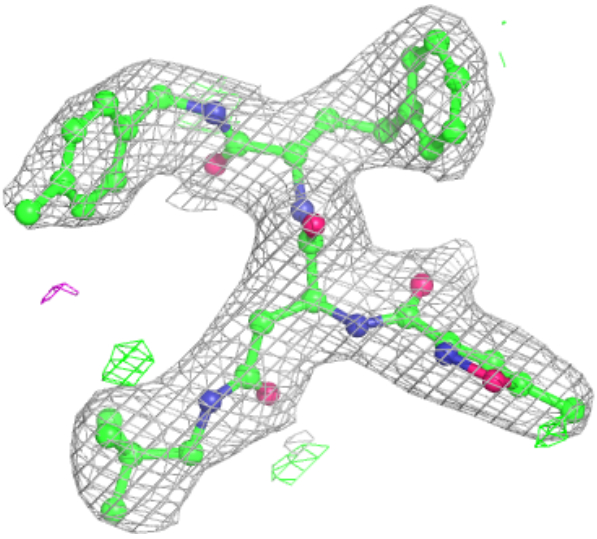
Continued from previous page...

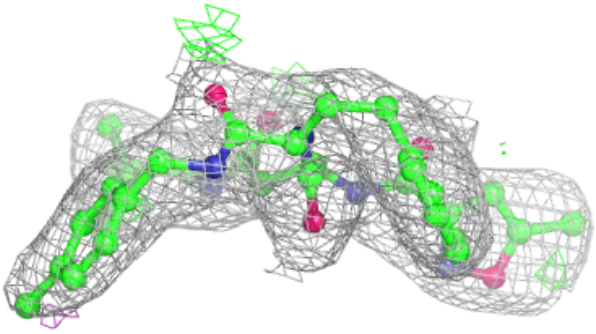
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
15	MG	I	201	1/1	0.86	0.10	46,46,46,46	0
15	MG	T	302	1/1	0.92	0.50	110,110,110,110	0
15	MG	L	202	1/1	0.93	0.10	46,46,46,46	0
15	MG	U	301	1/1	0.93	0.07	38,38,38,38	0
15	MG	Z	202	1/1	0.93	0.13	49,49,49,49	0
15	MG	W	202	1/1	0.94	0.14	50,50,50,50	0
16	3SD	Y	302	42/42	0.94	0.16	37,39,48,50	0
15	MG	N	201	1/1	0.95	0.14	41,41,41,41	0
15	MG	H	301	1/1	0.96	0.07	47,47,47,47	0
16	3SD	K	302	42/42	0.96	0.15	28,35,49,50	0
15	MG	G	9001	1/1	0.96	0.10	45,45,45,45	0
17	MES	K	303	12/12	0.96	0.18	55,57,58,58	0
15	MG	K	301	1/1	0.97	0.13	50,50,50,50	0
15	MG	V	301	1/1	0.97	0.08	42,42,42,42	0
15	MG	2	201	1/1	0.97	0.14	30,30,30,30	0
17	MES	Y	303	12/12	0.97	0.14	57,61,64,64	0
15	MG	Y	301	1/1	0.98	0.13	53,53,53,53	0

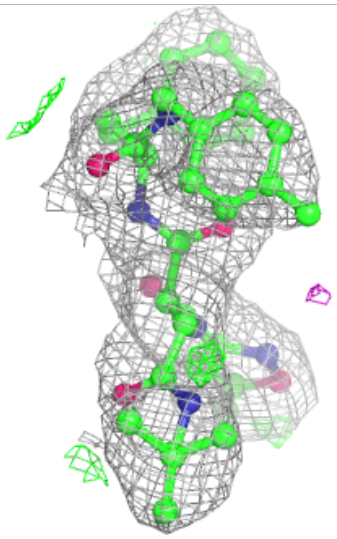
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 3SD Y 302:

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

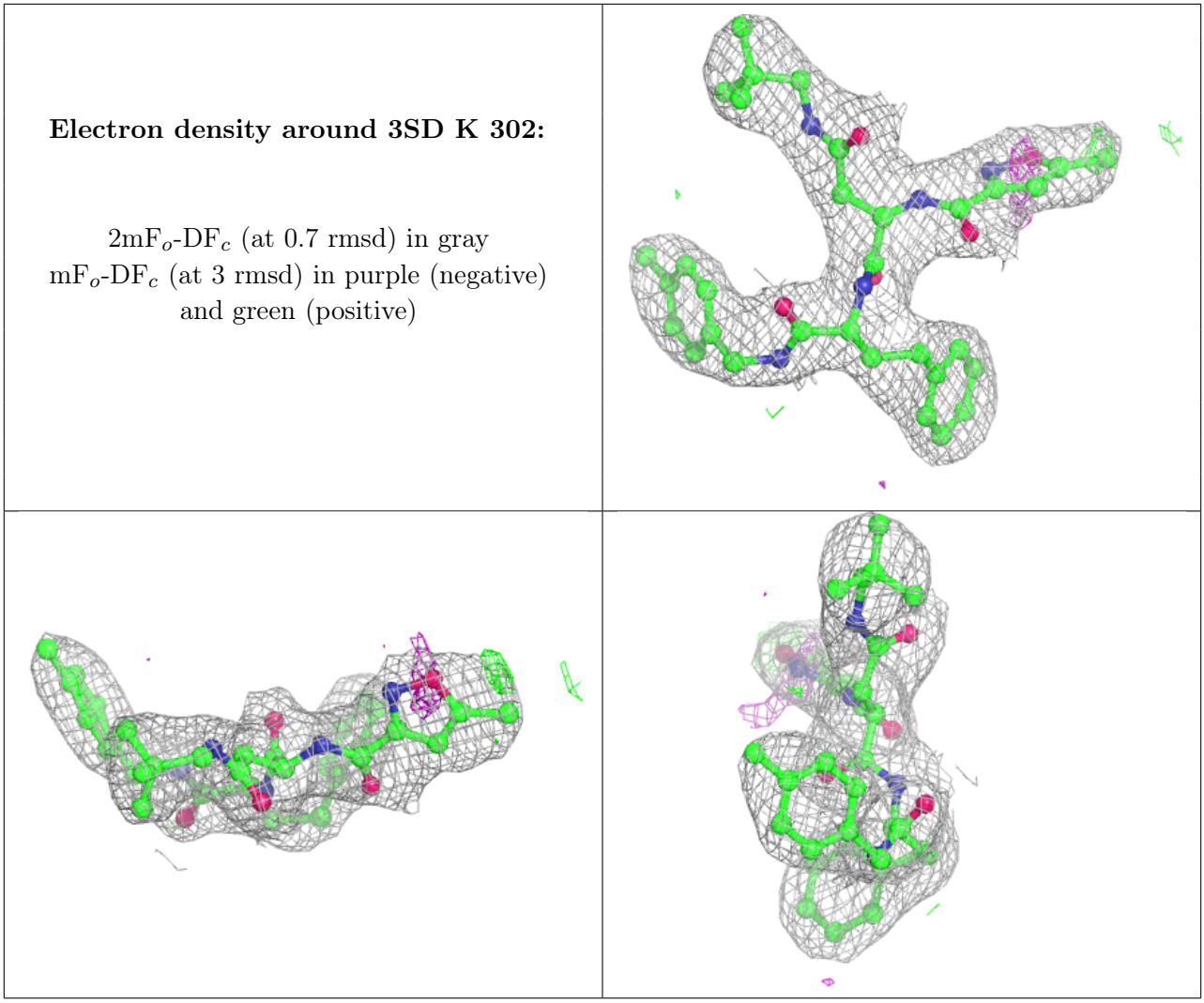






Electron density around 3SD K 302:

$2mF_o$-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.