wwPDB X-ray Structure Validation Summary Report (i)

Jan 2, 2024 – 02:34 pm GMT

PDB ID	:	5CGI
Title	:	Yeast 20S proteasome beta5-G48C mutant in complex with ONX 0914
Authors	:	Dubiella, C.; Groll, M.
Deposited on	:	2015-07-09
Resolution	:	2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

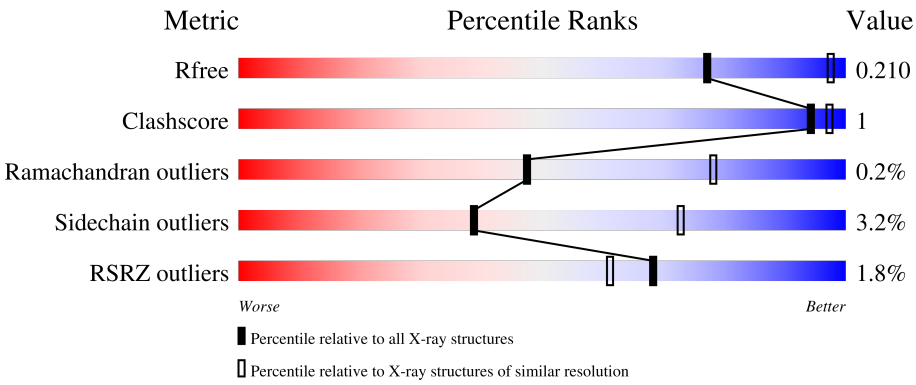
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	% 98% •
1	O	250	2% 98% •
2	B	258	3% 90% • • 5%
2	P	258	3% 89% 5% • 5%
3	C	254	4% 89% • • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	7% 89% 6%
4	D	260	.% 87% 10%
4	R	260	2% 87% 10%
5	E	234	2% 94% 5%
5	S	234	3% 94% 5%
6	F	288	.% 81% 16%
6	T	288	2% 81% 16%
7	G	252	2% 90% 6%
7	U	252	.% 90% 6%
8	H	232	.% 91%
8	V	232	.% 91% 5%
9	I	205	95% 5%
9	W	205	95% 5%
10	J	198	.% 92%
10	X	198	.% 93%
11	K	212	91% 8%
11	Y	212	92% 7%
12	L	222	93% 7%
12	Z	222	93% 7%
13	M	246	3% 89% 5% 5%
13	a	246	2% 92% 5%
14	N	196	.% 93% 7%
14	b	196	98%

2 Entry composition (i)

There are 20 unique types of molecules in this entry. The entry contains 49878 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	240	1881	1176	329	372	4	0	0	0
3	Q	240	1881	1176	329	372	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	235	1813	1136	304	366	7	0	0	0
4	R	235	1813	1136	304	366	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	231	1773	1114	307	348	4	0	0	0
5	S	231	1773	1114	307	348	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	243	1892	1203	329	356	4	0	0	0
6	T	243	1892	1203	329	356	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	241	1907	1214	320	365	8	0	0	0
7	U	241	1907	1214	320	365	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	195	1561	992	264	299	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	195	1561	992	264	299	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	1646	1046	280	312	8	0	0	0
			Total	C	N	O	S			
11	Y	212	1646	1046	280	312	8	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	48	CYS	GLY	conflict	UNP P30656
Y	48	CYS	GLY	conflict	UNP P30656

- 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
			Total	C	N	O	S			
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
			Total	C	N	O	S			
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
			Total	C	N	O	S			
14	b	196	1512	955	250	300	7	0	0	0

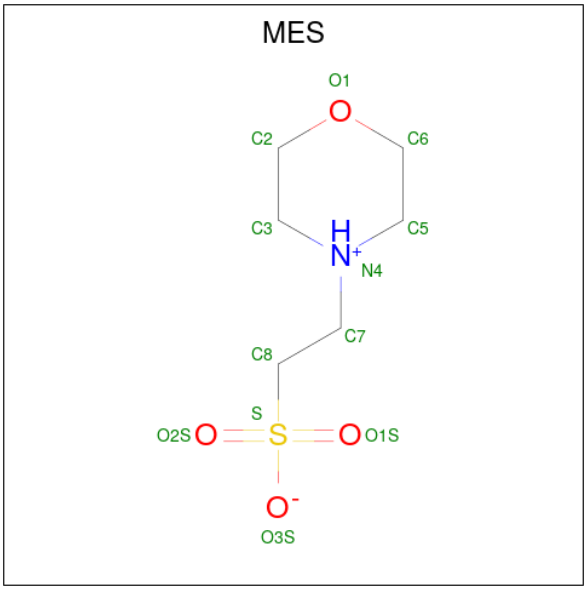
- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	G	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	1	Total Mg 1 1	0	0
15	N	1	Total Mg 1 1	0	0
15	X	1	Total Mg 1 1	0	0
15	Z	1	Total Mg 1 1	0	0

- Molecule 16 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

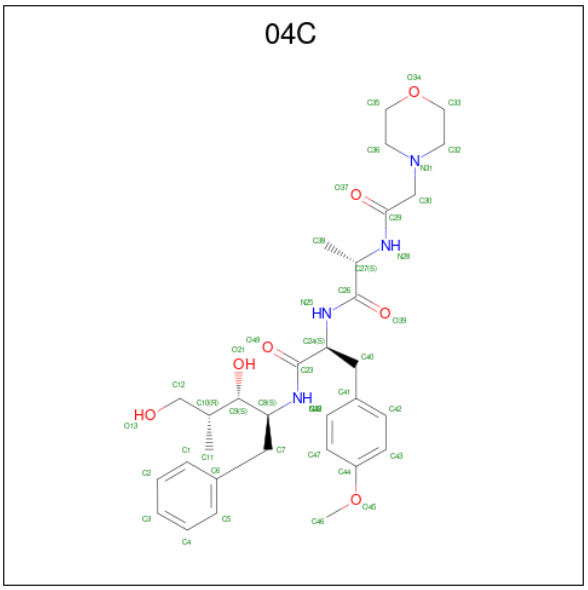
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	G	1	Total Cl 1 1	0	0
16	U	1	Total Cl 1 1	0	0

- 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	H	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
17	K	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
17	V	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
17	Y	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 18 is 1,2,4-trideoxy-4-methyl-2-{[N-(morpholin-4-ylacetyl)-L-alanyl-O-methyl-L-tyrosyl]amino}-1-phenyl-D-xylitol (three-letter code: 04C) (formula: $C_{31}H_{44}N_4O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
18	H	1	Total	C	N	O	0	0
			42	31	4	7		
18	K	1	Total	C	N	O	0	0
			42	31	4	7		
18	N	1	Total	C	N	O	0	0
			42	31	4	7		
18	V	1	Total	C	N	O	0	0
			42	31	4	7		
18	Y	1	Total	C	N	O	0	0
			42	31	4	7		
18	b	1	Total	C	N	O	0	0
			42	31	4	7		

- Molecule 19 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
19	K	1	1	1	0	0
			Total	Na		
19	Y	1	1	1	0	0

- Molecule 20 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
20	A	10	10	10	0	0
			Total	O		
20	B	11	11	11	0	0
			Total	O		
20	C	9	9	9	0	0
			Total	O		
20	D	5	5	5	0	0
			Total	O		
20	E	5	5	5	0	0
			Total	O		
20	F	3	3	3	0	0
			Total	O		
20	G	11	11	11	0	0
			Total	O		
20	H	13	13	13	0	0
			Total	O		
20	I	9	9	9	0	0
			Total	O		
20	J	17	17	17	0	0
			Total	O		
20	K	11	11	11	0	0
			Total	O		
20	L	14	14	14	0	0
			Total	O		
20	M	10	10	10	0	0
			Total	O		
20	N	7	7	7	0	0
			Total	O		
20	O	8	8	8	0	0
			Total	O		
20	P	12	12	12	0	0
			Total	O		
20	Q	5	5	5	0	0
			Total	O		
20	R	2	2	2	0	0

Continued on next page...

Continued from previous page...

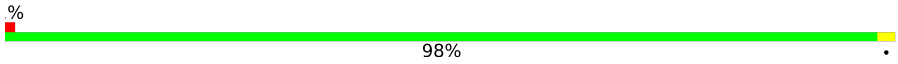
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
20	S	6	6	6	0	0
			Total	O		
20	T	8	8	8	0	0
			Total	O		
20	U	9	9	9	0	0
			Total	O		
20	V	11	11	11	0	0
			Total	O		
20	W	7	7	7	0	0
			Total	O		
20	X	13	13	13	0	0
			Total	O		
20	Y	13	13	13	0	0
			Total	O		
20	Z	12	12	12	0	0
			Total	O		
20	a	16	16	16	0	0
			Total	O		
20	b	9	9	9	0	0

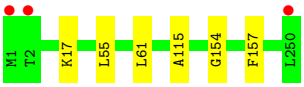
3 Residue-property plots (i)

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

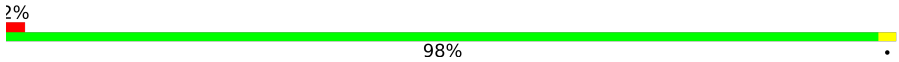
Chain A:





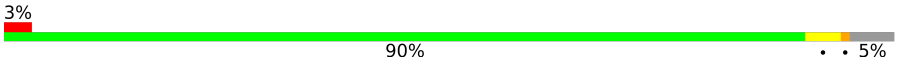
• Molecule 1: Proteasome subunit alpha type-2

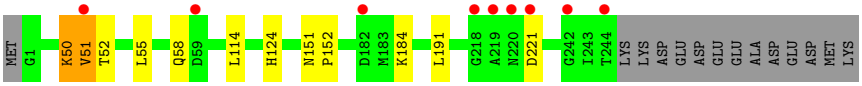
Chain O:



• Molecule 2: Proteasome subunit alpha type-3

Chain B:



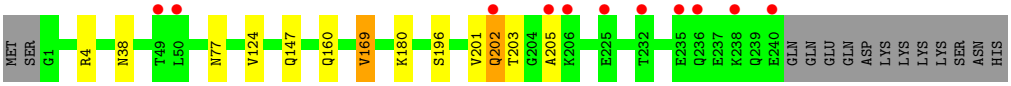


• Molecule 2: Proteasome subunit alpha type-3

Chain P:

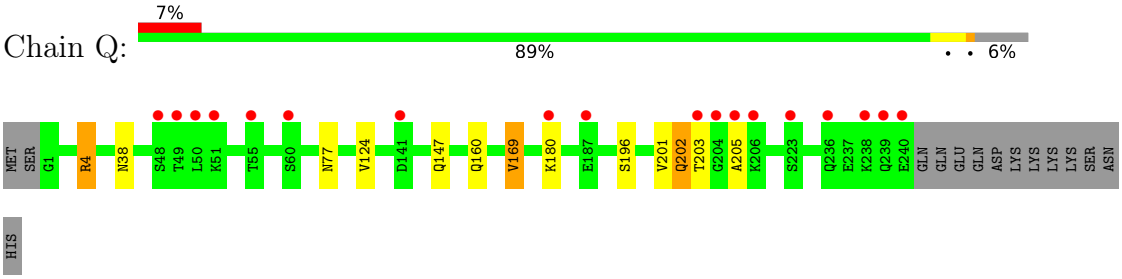
• Molecule 3: Proteasome subunit alpha type-4

Chain C:



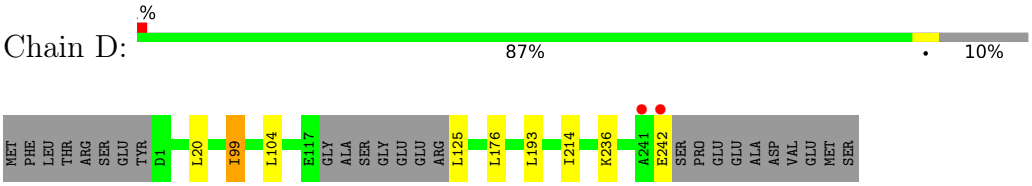
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



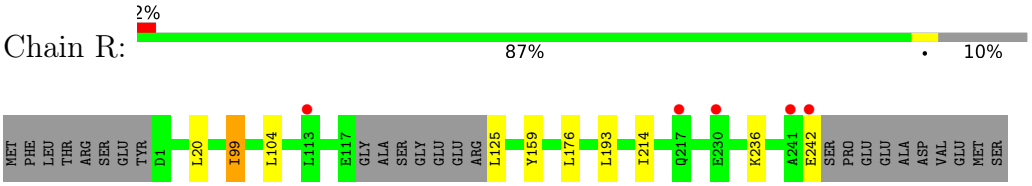
• Molecule 4: Proteasome subunit alpha type-5

Chain D:



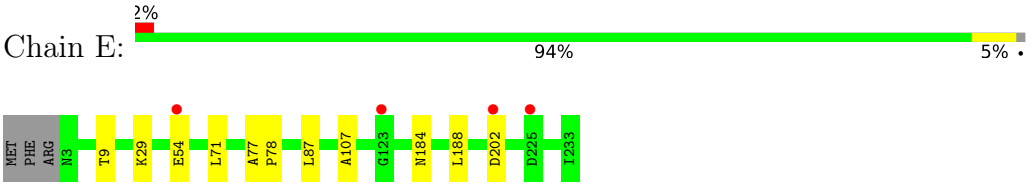
• Molecule 4: Proteasome subunit alpha type-5

Chain R:



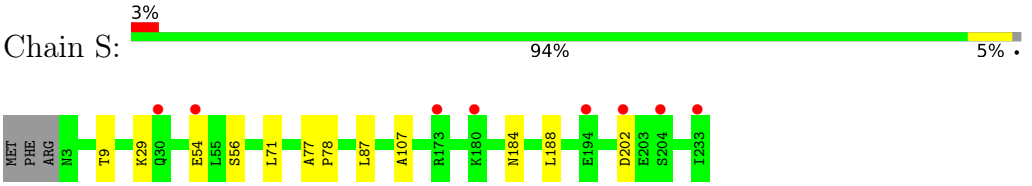
• Molecule 5: Proteasome subunit alpha type-6

Chain E:



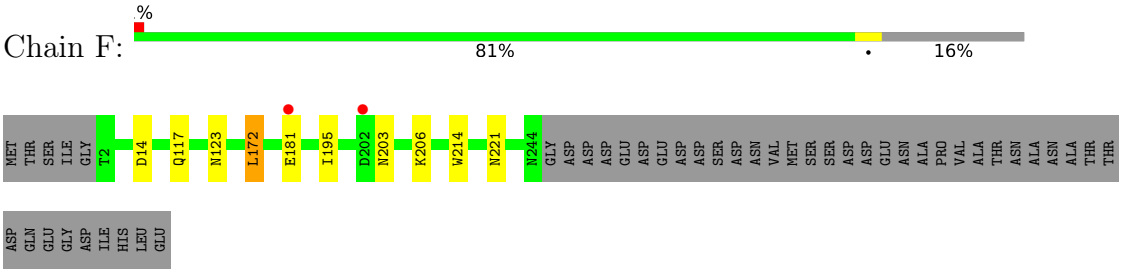
• Molecule 5: Proteasome subunit alpha type-6

Chain S:



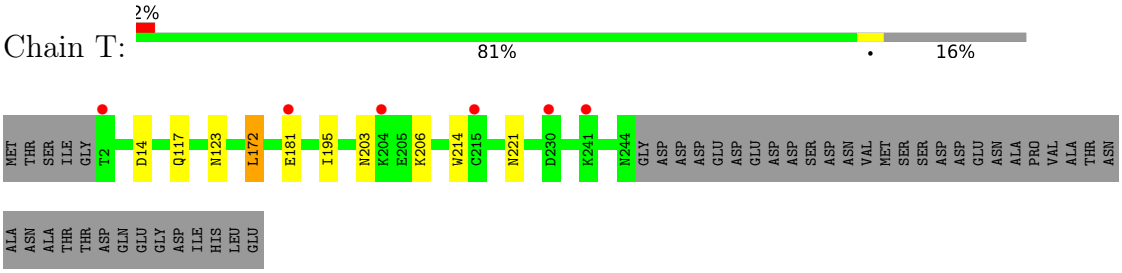
• Molecule 6: Probable proteasome subunit alpha type-7

Chain F:



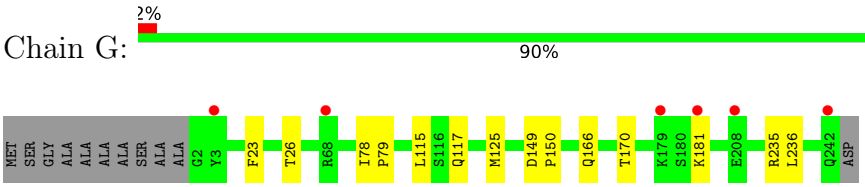
• Molecule 6: Probable proteasome subunit alpha type-7

Chain T:



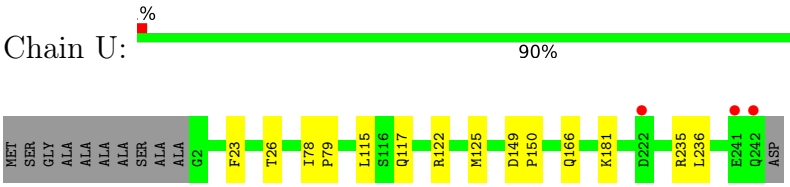
- Molecule 7: Proteasome subunit alpha type-1

Chain G:



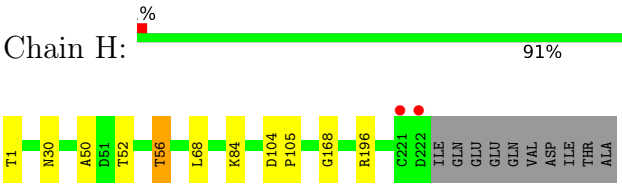
- Molecule 7: Proteasome subunit alpha type-1

Chain U:



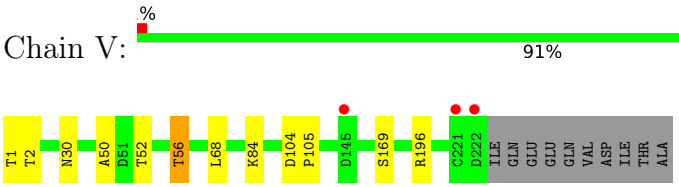
- Molecule 8: Proteasome subunit beta type-2

Chain H:



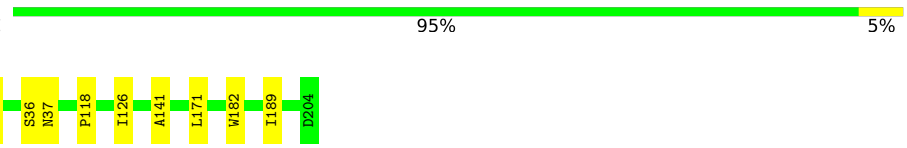
- Molecule 8: Proteasome subunit beta type-2

Chain V:



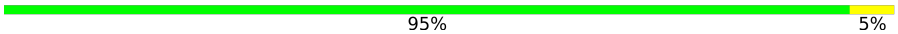
- Molecule 9: Proteasome subunit beta type-3

Chain I:



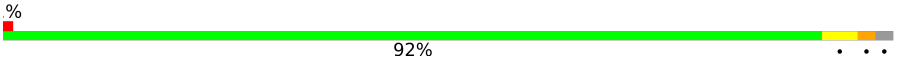
- Molecule 9: Proteasome subunit beta type-3

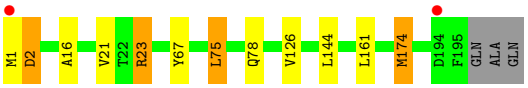
Chain W:



• Molecule 10: Proteasome subunit beta type-4

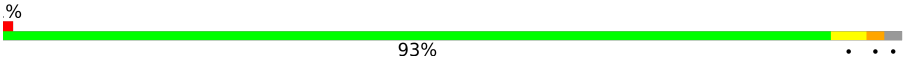
Chain J:





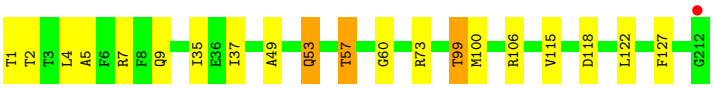
• Molecule 10: Proteasome subunit beta type-4

Chain X:



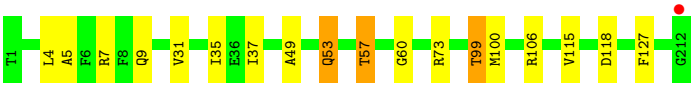
• Molecule 11: Proteasome subunit beta type-5

Chain K:



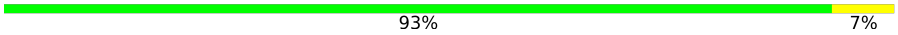
• Molecule 11: Proteasome subunit beta type-5

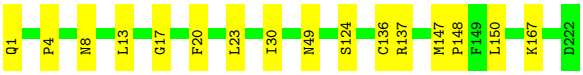
Chain Y:



• Molecule 12: Proteasome subunit beta type-6

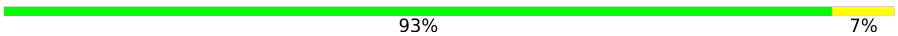
Chain L:





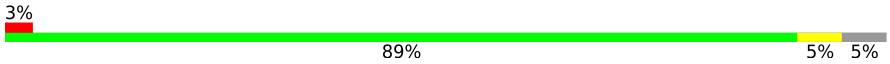
• Molecule 12: Proteasome subunit beta type-6

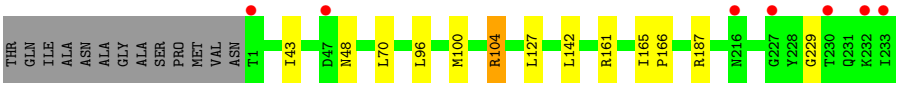
Chain Z:



• Molecule 13: Proteasome subunit beta type-7

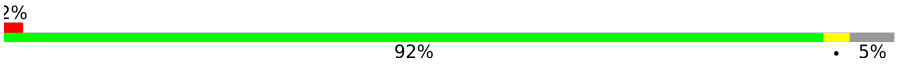
Chain M:





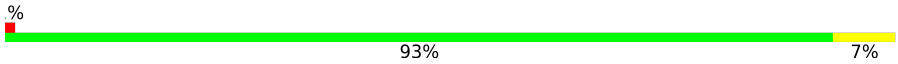
• Molecule 13: Proteasome subunit beta type-7

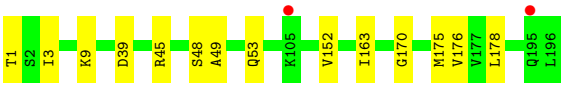
Chain a:



• Molecule 14: Proteasome subunit beta type-1

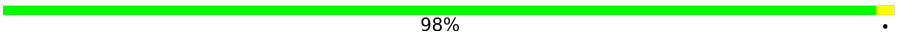
Chain N:





• Molecule 14: Proteasome subunit beta type-1

Chain b:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	136.19Å 300.03Å 145.04Å 90.00° 112.96° 90.00°	Depositor
Resolution (Å)	15.00 – 2.80 15.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.8 (15.00-2.80) 98.9 (15.00-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.40 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.186 , 0.206 0.192 , 0.210	Depositor DCC
R_{free} test set	12874 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	56.9	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 43.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	49878	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.36% 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: MG, CL, MES, NA, 04C

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	0.27	0/1952	0.47	0/2642
1	O	0.27	0/1952	0.46	0/2642
2	B	0.27	0/1934	0.48	0/2618
2	P	0.27	0/1934	0.48	0/2618
3	C	0.27	0/1910	0.50	0/2586
3	Q	0.27	0/1910	0.50	0/2586
4	D	0.26	0/1837	0.47	0/2475
4	R	0.26	0/1837	0.47	0/2475
5	E	0.26	0/1800	0.46	0/2433
5	S	0.26	0/1800	0.46	0/2433
6	F	0.27	0/1932	0.45	0/2609
6	T	0.27	0/1932	0.44	0/2609
7	G	0.27	0/1945	0.46	0/2634
7	U	0.27	0/1945	0.46	0/2634
8	H	0.26	0/1715	0.52	1/2326 (0.0%)
8	V	0.28	0/1715	0.50	0/2326
9	I	0.28	0/1611	0.53	0/2174
9	W	0.28	0/1611	0.53	0/2174
10	J	0.26	0/1589	0.48	0/2142
10	X	0.26	0/1589	0.48	0/2142
11	K	0.26	0/1683	0.53	0/2277
11	Y	0.26	0/1683	0.53	0/2277
12	L	0.32	0/1795	0.50	0/2420
12	Z	0.27	0/1795	0.50	0/2420
13	M	0.26	0/1855	0.53	0/2514
13	a	0.26	0/1855	0.53	0/2514
14	N	0.25	0/1541	0.49	0/2087
14	b	0.27	0/1541	0.52	1/2087 (0.0%)
All	All	0.27	0/50198	0.49	2/67874 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	H	1	THR	N-CA-C	6.22	127.80	111.00
14	b	1	THR	N-CA-C	5.63	126.20	111.00

There are no chirality outliers.

There are no planarity outliers.

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	2	0
1	O	1915	0	1929	1	0
2	B	1904	0	1904	4	0
2	P	1904	0	1904	6	0
3	C	1881	0	1895	5	0
3	Q	1881	0	1895	6	0
4	D	1813	0	1797	1	0
4	R	1813	0	1797	2	0
5	E	1773	0	1775	2	0
5	S	1773	0	1775	3	0
6	F	1892	0	1883	1	0
6	T	1892	0	1883	1	0
7	G	1907	0	1901	4	0
7	U	1907	0	1901	4	0
8	H	1684	0	1685	4	0
8	V	1684	0	1685	8	0
9	I	1581	0	1574	5	0
9	W	1581	0	1574	5	0
10	J	1561	0	1569	8	0
10	X	1561	0	1569	7	0
11	K	1646	0	1594	11	0
11	Y	1646	0	1594	10	0
12	L	1757	0	1711	7	0
12	Z	1757	0	1711	6	0
13	M	1824	0	1832	4	0
13	a	1824	0	1832	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	N	1512	0	1478	7	0
14	b	1512	0	1478	0	0
15	G	1	0	0	0	0
15	I	2	0	0	0	0
15	K	1	0	0	0	0
15	L	1	0	0	0	0
15	N	1	0	0	0	0
15	X	1	0	0	0	0
15	Z	1	0	0	0	0
16	G	1	0	0	0	0
16	U	1	0	0	0	0
17	H	12	0	13	0	0
17	K	12	0	13	1	0
17	V	12	0	13	0	0
17	Y	12	0	13	2	0
18	H	42	0	41	2	0
18	K	42	0	41	1	0
18	N	42	0	42	2	0
18	V	42	0	41	2	0
18	Y	42	0	41	5	0
18	b	42	0	42	0	0
19	K	1	0	0	0	0
19	Y	1	0	0	0	0
20	A	10	0	0	0	0
20	B	11	0	0	0	0
20	C	9	0	0	0	0
20	D	5	0	0	0	0
20	E	5	0	0	0	0
20	F	3	0	0	0	0
20	G	11	0	0	0	0
20	H	13	0	0	0	0
20	I	9	0	0	0	0
20	J	17	0	0	0	0
20	K	11	0	0	0	0
20	L	14	0	0	0	0
20	M	10	0	0	0	0
20	N	7	0	0	0	0
20	O	8	0	0	0	0
20	P	12	0	0	0	0
20	Q	5	0	0	0	0
20	R	2	0	0	0	0
20	S	6	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
20	T	8	0	0	0	0
20	U	9	0	0	1	0
20	V	11	0	0	0	0
20	W	7	0	0	0	0
20	X	13	0	0	0	0
20	Y	13	0	0	0	0
20	Z	12	0	0	0	0
20	a	16	0	0	0	0
20	b	9	0	0	0	0
All	All	49878	0	49354	119	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 1.

The worst 5 of 119 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)
10:J:1:MET:O	10:J:2:ASP:HB2	1.88	0.74
10:X:1:MET:O	10:X:2:ASP:HB2	1.88	0.72
18:H:302:04C:H9	18:H:302:04C:H25	1.74	0.70
11:Y:5:ALA:HB3	11:Y:100:MET:HE2	1.77	0.67
11:Y:53:GLN:O	11:Y:57:THR:OG1	2.12	0.67

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%)	242 (98%)	6 (2%)	0	100	100
1	O	248/250 (99%)	242 (98%)	6 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	242/258 (94%)	233 (96%)	7 (3%)	2 (1%)	19	49
2	P	242/258 (94%)	232 (96%)	8 (3%)	2 (1%)	19	49
3	C	238/254 (94%)	230 (97%)	6 (2%)	2 (1%)	19	49
3	Q	238/254 (94%)	230 (97%)	6 (2%)	2 (1%)	19	49
4	D	231/260 (89%)	225 (97%)	6 (3%)	0	100	100
4	R	231/260 (89%)	225 (97%)	6 (3%)	0	100	100
5	E	229/234 (98%)	222 (97%)	7 (3%)	0	100	100
5	S	229/234 (98%)	222 (97%)	7 (3%)	0	100	100
6	F	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
6	T	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
7	G	239/252 (95%)	235 (98%)	4 (2%)	0	100	100
7	U	239/252 (95%)	235 (98%)	4 (2%)	0	100	100
8	H	220/232 (95%)	215 (98%)	5 (2%)	0	100	100
8	V	220/232 (95%)	215 (98%)	5 (2%)	0	100	100
9	I	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
9	W	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	J	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	29	61
10	X	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	29	61
11	K	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
11	Y	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
12	L	220/222 (99%)	217 (99%)	3 (1%)	0	100	100
12	Z	220/222 (99%)	217 (99%)	3 (1%)	0	100	100
13	M	231/246 (94%)	220 (95%)	10 (4%)	1 (0%)	34	66
13	a	231/246 (94%)	220 (95%)	10 (4%)	1 (0%)	34	66
14	N	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
14	b	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
All	All	6276/6614 (95%)	6099 (97%)	165 (3%)	12 (0%)	47	78

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
3	C	202	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
10	J	2	ASP
2	P	51	VAL
3	Q	202	GLN

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%)	206 (99%)	3 (1%)	67	90
1	O	209/209 (100%)	206 (99%)	3 (1%)	67	90
2	B	203/216 (94%)	196 (97%)	7 (3%)	37	71
2	P	203/216 (94%)	196 (97%)	7 (3%)	37	71
3	C	212/226 (94%)	204 (96%)	8 (4%)	33	67
3	Q	212/226 (94%)	204 (96%)	8 (4%)	33	67
4	D	194/215 (90%)	186 (96%)	8 (4%)	30	64
4	R	194/215 (90%)	186 (96%)	8 (4%)	30	64
5	E	190/193 (98%)	183 (96%)	7 (4%)	34	68
5	S	190/193 (98%)	183 (96%)	7 (4%)	34	68
6	F	201/239 (84%)	192 (96%)	9 (4%)	27	60
6	T	201/239 (84%)	192 (96%)	9 (4%)	27	60
7	G	206/210 (98%)	199 (97%)	7 (3%)	37	71
7	U	206/210 (98%)	199 (97%)	7 (3%)	37	71
8	H	181/190 (95%)	176 (97%)	5 (3%)	43	77
8	V	181/190 (95%)	176 (97%)	5 (3%)	43	77
9	I	172/173 (99%)	169 (98%)	3 (2%)	60	87
9	W	172/173 (99%)	169 (98%)	3 (2%)	60	87
10	J	173/175 (99%)	168 (97%)	5 (3%)	42	76
10	X	173/175 (99%)	168 (97%)	5 (3%)	42	76
11	K	170/170 (100%)	161 (95%)	9 (5%)	22	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	Y	170/170 (100%)	161 (95%)	9 (5%)	22	54
12	L	185/185 (100%)	180 (97%)	5 (3%)	44	78
12	Z	185/185 (100%)	180 (97%)	5 (3%)	44	78
13	M	199/208 (96%)	193 (97%)	6 (3%)	41	75
13	a	199/208 (96%)	193 (97%)	6 (3%)	41	75
14	N	162/162 (100%)	160 (99%)	2 (1%)	71	92
14	b	162/162 (100%)	160 (99%)	2 (1%)	71	92
All	All	5314/5542 (96%)	5146 (97%)	168 (3%)	39	73

5 of 168 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	S	188	LEU
10	X	78	GLN
6	T	123	ASN
7	U	181	LYS
11	Y	53	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

Mol	Chain	Res	Type
6	T	86	ASN
12	Z	49	ASN
6	T	123	ASN
7	U	121	GLN
13	a	102	GLN

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 22 ligands modelled in this entry, 12 are monoatomic - leaving 10 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	$\#\|Z\| > 2$
18	04C	Y	303	11	44,44,44	1.72	5 (11%)	56,58,58	1.35	6 (10%)
17	MES	Y	301	-	12,12,12	2.36	1 (8%)	14,16,16	1.34	2 (14%)
17	MES	K	303	-	12,12,12	2.33	1 (8%)	14,16,16	1.38	2 (14%)
18	04C	N	202	14	44,44,44	1.45	4 (9%)	56,58,58	1.41	8 (14%)
18	04C	b	201	14	44,44,44	1.43	3 (6%)	56,58,58	1.41	9 (16%)
18	04C	H	302	8	44,44,44	1.61	3 (6%)	56,58,58	1.26	4 (7%)
18	04C	K	304	11	44,44,44	1.74	5 (11%)	56,58,58	1.40	7 (12%)
17	MES	H	301	-	12,12,12	2.19	1 (8%)	14,16,16	1.39	2 (14%)
17	MES	V	301	-	12,12,12	2.24	1 (8%)	14,16,16	1.34	3 (21%)
18	04C	V	302	8	44,44,44	1.55	4 (9%)	56,58,58	1.11	5 (8%)

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
18	04C	Y	303	11	-	12/44/52/52	0/3/3/3
17	MES	Y	301	-	-	1/6/14/14	0/1/1/1
17	MES	K	303	-	-	0/6/14/14	0/1/1/1
18	04C	N	202	14	-	16/44/52/52	0/3/3/3
18	04C	b	201	14	-	18/44/52/52	0/3/3/3
18	04C	H	302	8	-	13/44/52/52	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
18	04C	K	304	11	-	12/44/52/52	0/3/3/3
17	MES	H	301	-	-	3/6/14/14	0/1/1/1
17	MES	V	301	-	-	0/6/14/14	0/1/1/1
18	04C	V	302	8	-	16/44/52/52	0/3/3/3

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	Y	301	MES	C8-S	-7.83	1.66	1.77
17	K	303	MES	C8-S	-7.75	1.66	1.77
17	V	301	MES	C8-S	-7.49	1.66	1.77
17	H	301	MES	C8-S	-7.32	1.67	1.77
18	H	302	04C	C40-C41	-6.08	1.36	1.51

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	N	202	04C	C11-C10-C12	-5.68	102.41	109.88
18	Y	303	04C	C7-C8-N22	-4.68	103.25	110.07
18	K	304	04C	C7-C8-N22	-4.44	103.59	110.07
18	N	202	04C	C33-O34-C35	3.70	122.25	109.89
18	K	304	04C	C11-C10-C12	-3.59	105.17	109.88

There are no chirality outliers.

5 of 91 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	H	301	MES	C7-C8-S-O2S
18	K	304	04C	C11-C10-C9-C8
18	N	202	04C	C11-C10-C9-C8
18	N	202	04C	C12-C10-C9-C8
18	V	302	04C	C11-C10-C9-O21

There are no ring outliers.

7 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
18	Y	303	04C	5	0
17	Y	301	MES	2	0
17	K	303	MES	1	0

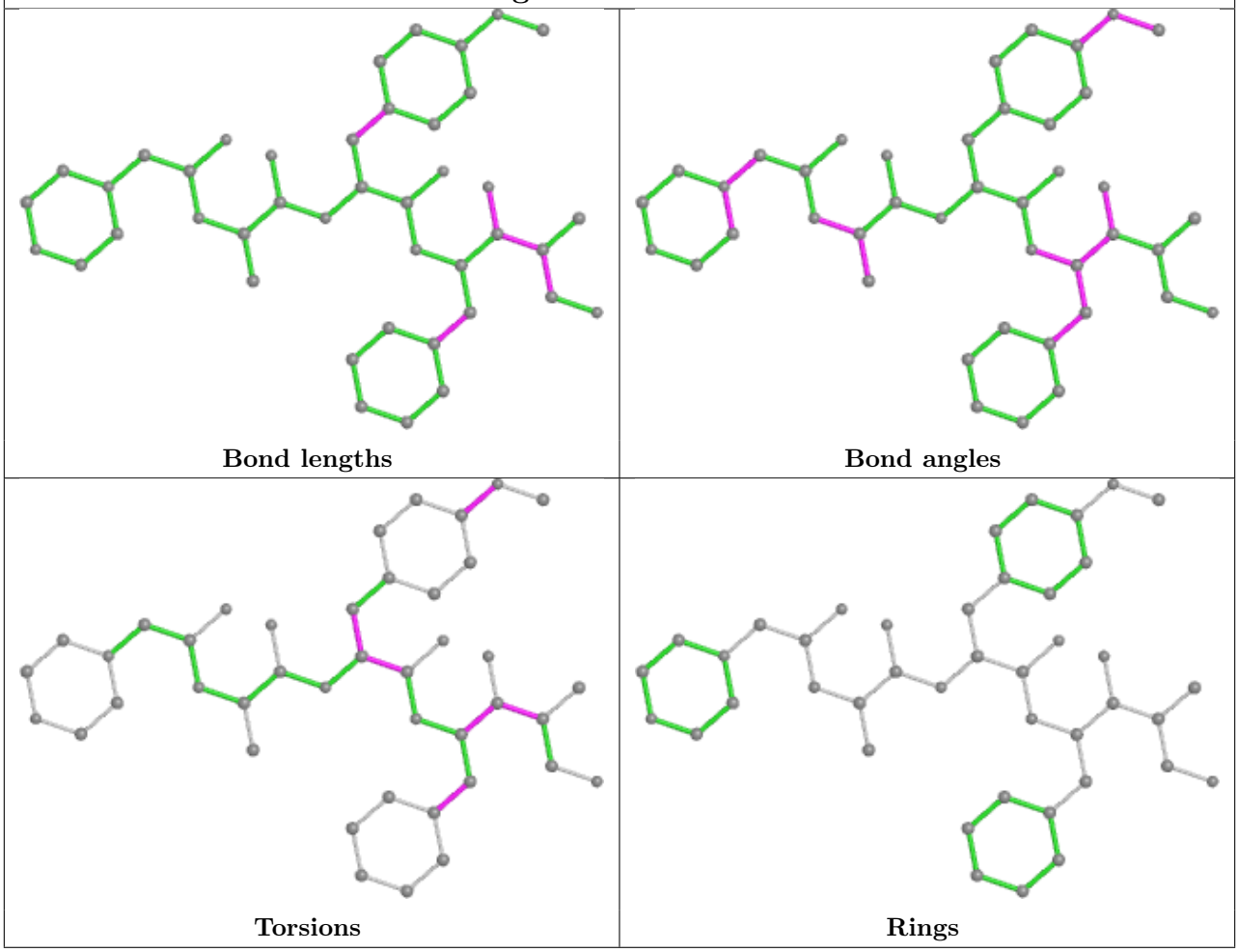
Continued on next page...

Continued from previous page...

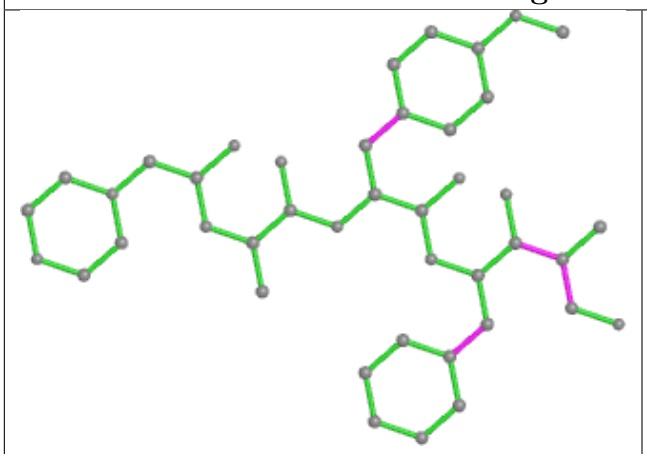
Mol	Chain	Res	Type	Clashes	Symm-Clashes
18	N	202	04C	2	0
18	H	302	04C	2	0
18	K	304	04C	1	0
18	V	302	04C	2	0

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.

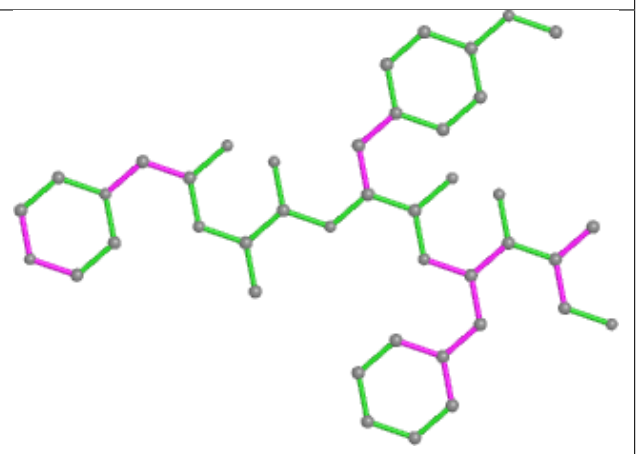
Ligand 04C Y 303



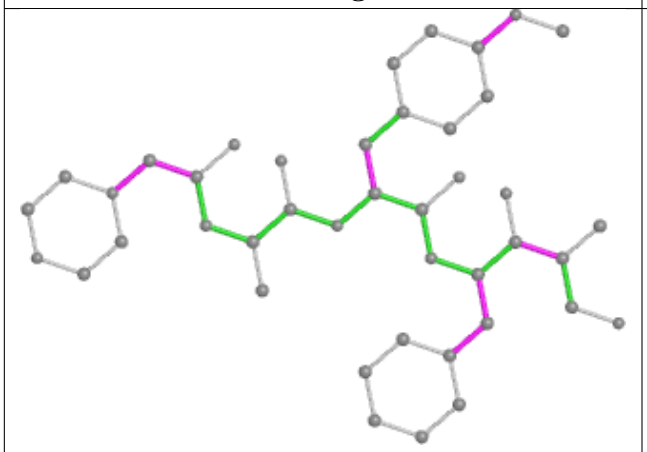
Ligand 04C N 202



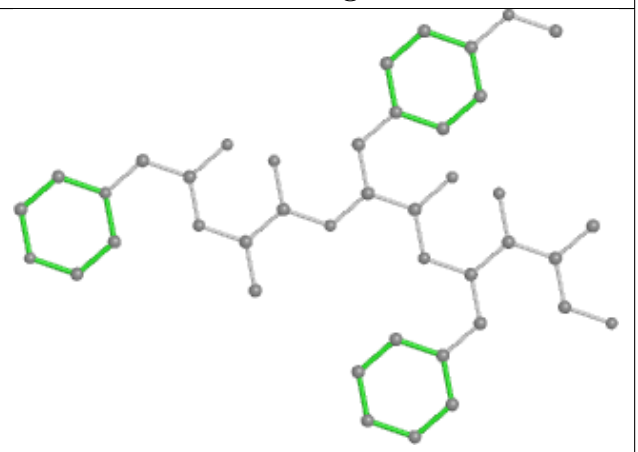
Bond lengths



Bond angles

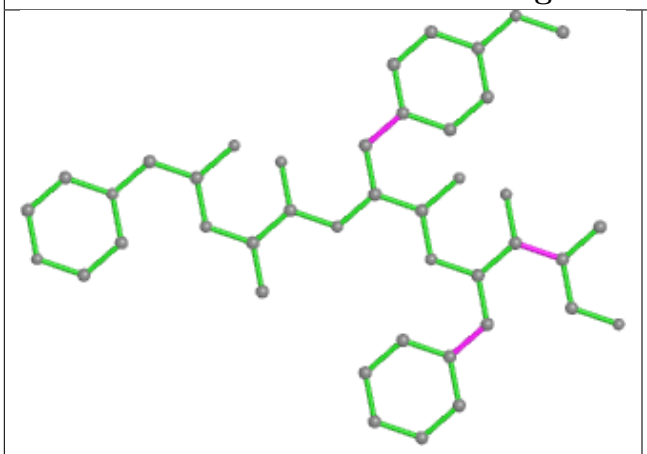


Torsions

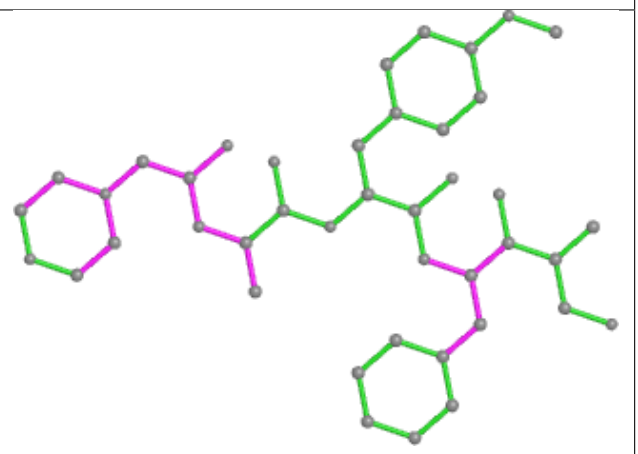


Rings

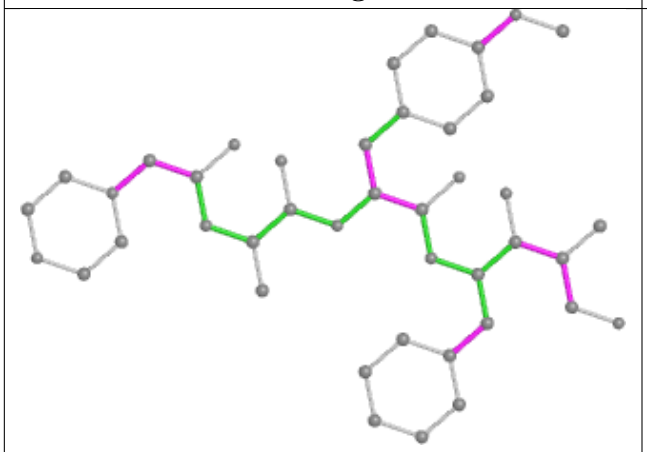
Ligand 04C b 201



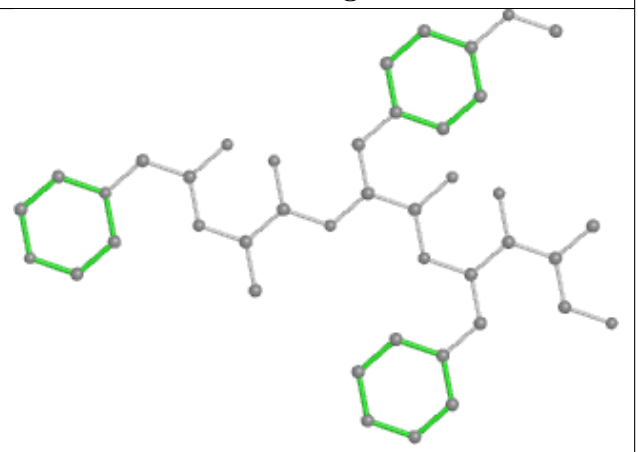
Bond lengths



Bond angles

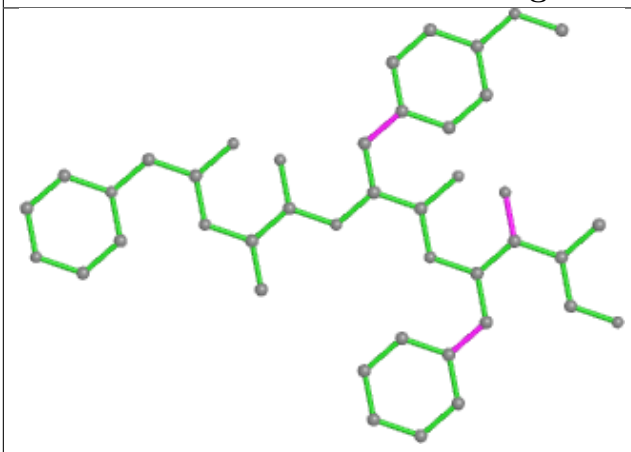


Torsions

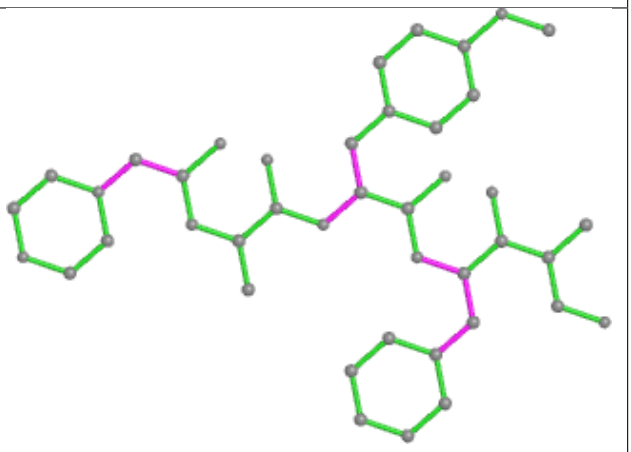


Rings

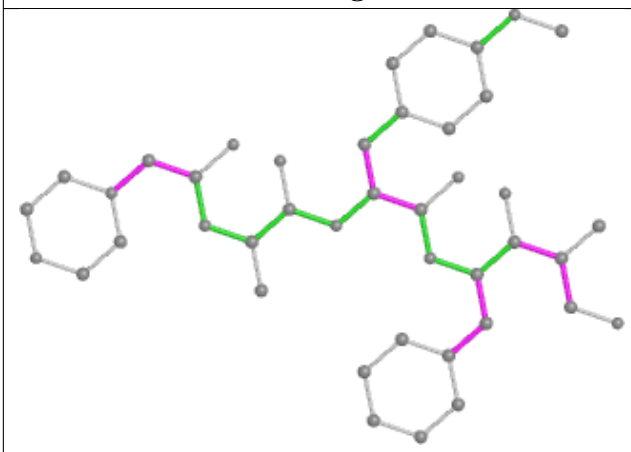
Ligand 04C H 302



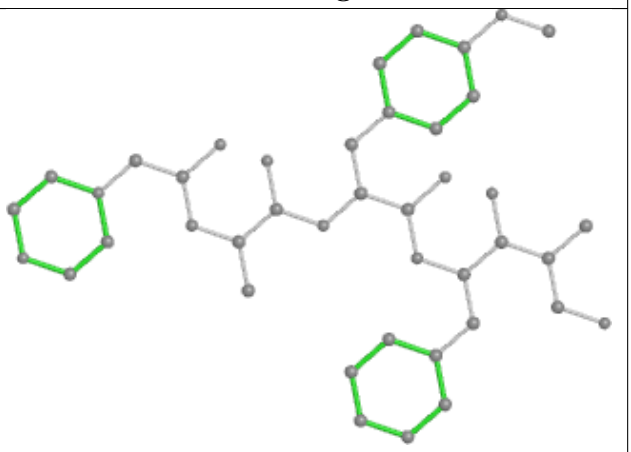
Bond lengths



Bond angles

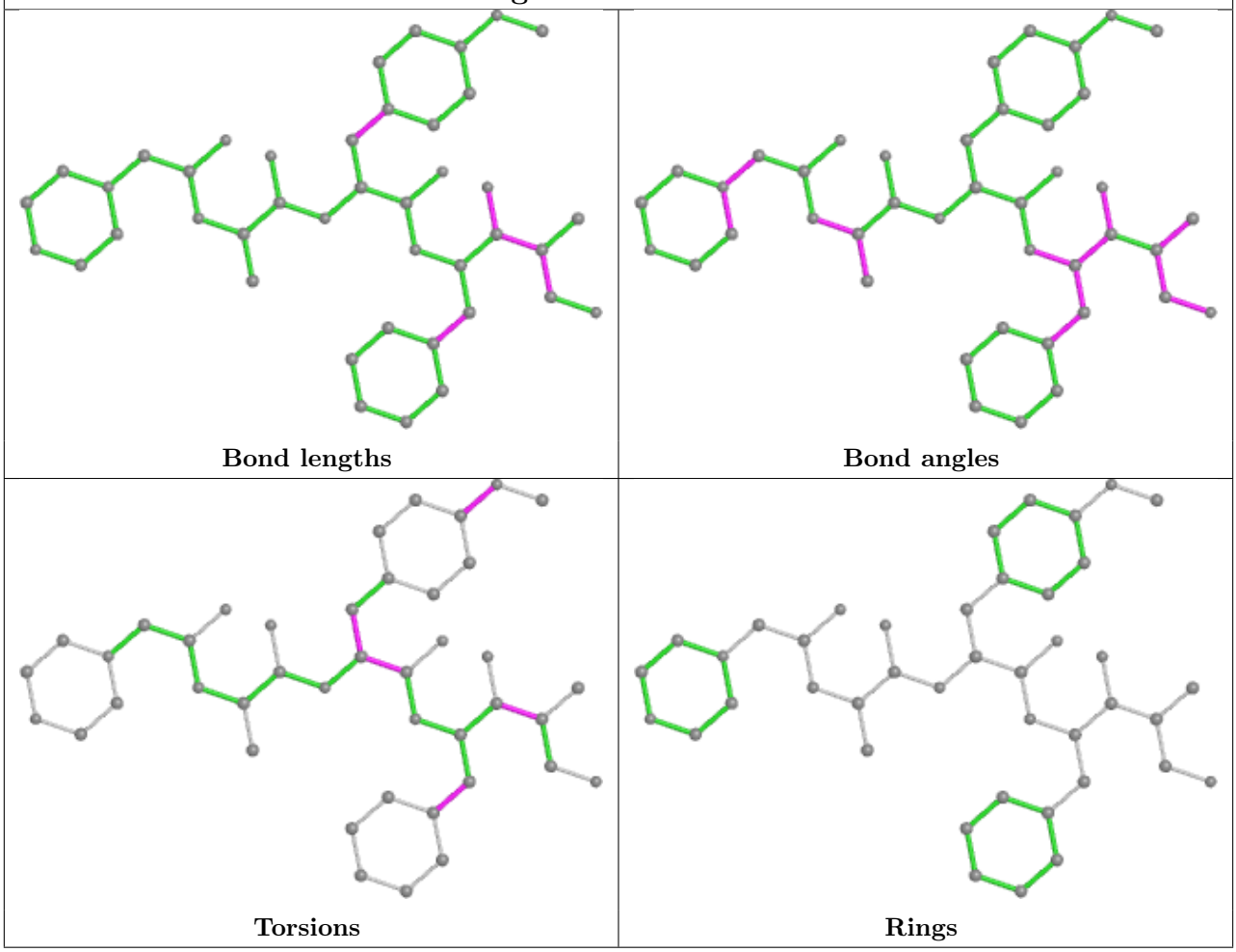


Torsions

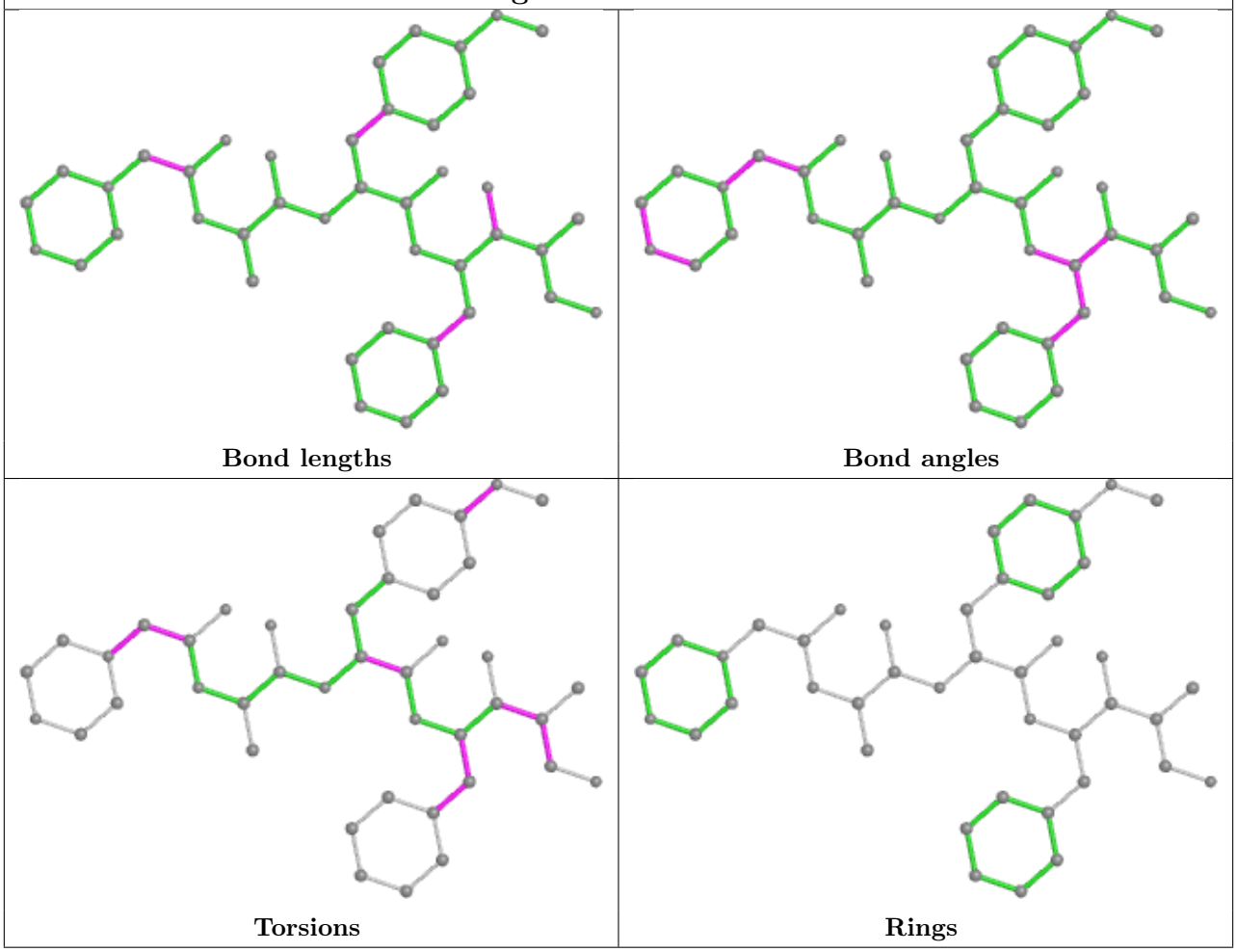


Rings

Ligand 04C K 304



Ligand 04C V 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 (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%)	-0.44	3 (1%)	79	73	40, 55, 96, 142	0
1	O	250/250 (100%)	-0.39	5 (2%)	65	56	42, 62, 103, 147	0
2	B	244/258 (94%)	-0.33	9 (3%)	41	31	38, 59, 108, 161	0
2	P	244/258 (94%)	-0.31	7 (2%)	51	41	41, 64, 110, 151	0
3	C	240/254 (94%)	-0.25	11 (4%)	32	22	38, 64, 126, 156	0
3	Q	240/254 (94%)	0.10	18 (7%)	14	8	50, 83, 165, 200	0
4	D	235/260 (90%)	-0.35	2 (0%)	84	80	45, 68, 100, 141	0
4	R	235/260 (90%)	-0.26	5 (2%)	63	54	50, 71, 112, 162	0
5	E	231/234 (98%)	-0.26	4 (1%)	70	63	47, 73, 111, 157	0
5	S	231/234 (98%)	-0.07	8 (3%)	44	34	51, 84, 139, 178	0
6	F	243/288 (84%)	-0.41	2 (0%)	86	81	42, 67, 116, 149	0
6	T	243/288 (84%)	-0.22	6 (2%)	57	47	47, 77, 138, 173	0
7	G	241/252 (95%)	-0.43	6 (2%)	57	47	38, 61, 102, 146	0
7	U	241/252 (95%)	-0.46	3 (1%)	79	73	41, 60, 97, 133	0
8	H	222/232 (95%)	-0.59	2 (0%)	84	80	38, 56, 84, 108	0
8	V	222/232 (95%)	-0.53	3 (1%)	75	70	38, 57, 85, 112	0
9	I	204/205 (99%)	-0.76	0	100	100	34, 49, 80, 108	0
9	W	204/205 (99%)	-0.71	0	100	100	35, 52, 81, 112	0
10	J	195/198 (98%)	-0.61	2 (1%)	82	77	32, 49, 80, 120	0
10	X	195/198 (98%)	-0.61	2 (1%)	82	77	37, 52, 85, 135	0
11	K	212/212 (100%)	-0.65	1 (0%)	91	88	31, 52, 77, 101	0
11	Y	212/212 (100%)	-0.61	1 (0%)	91	88	25, 53, 82, 110	0
12	L	222/222 (100%)	-0.63	0	100	100	36, 56, 84, 103	0
12	Z	222/222 (100%)	-0.64	1 (0%)	91	88	35, 55, 84, 102	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%)	-0.52	7 (3%)	50	40	33, 57, 89, 115	0
13	a	233/246 (94%)	-0.55	4 (1%)	70	63	30, 56, 89, 124	0
14	N	196/196 (100%)	-0.63	2 (1%)	82	77	33, 53, 84, 110	0
14	b	196/196 (100%)	-0.61	0	100	100	35, 53, 87, 109	0
All	All	6336/6614 (95%)	-0.44	114 (1%)	68	61	25, 60, 111, 200	0

The worst 5 of 114 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	49	THR	8.5
3	Q	50	LEU	6.8
5	S	202	ASP	6.2
3	Q	206	LYS	6.2
1	A	1	MET	5.6

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	301	1/1	0.86	0.14	64,64,64,64	0
15	MG	G	301	1/1	0.89	0.18	57,57,57,57	0
18	04C	b	201	42/42	0.89	0.30	38,55,76,83	0
18	04C	N	202	42/42	0.90	0.29	35,52,68,69	0
18	04C	H	302	42/42	0.94	0.16	34,41,64,66	0
18	04C	K	304	42/42	0.95	0.15	29,41,54,55	0
17	MES	V	301	12/12	0.95	0.26	57,60,72,79	0

Continued on next page...

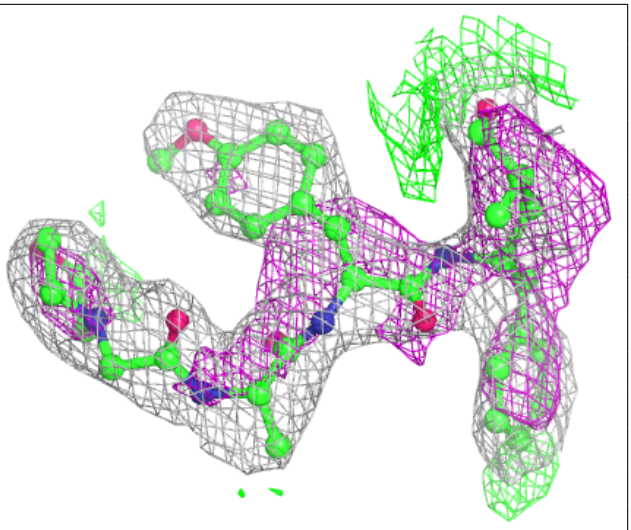
Continued from previous page...

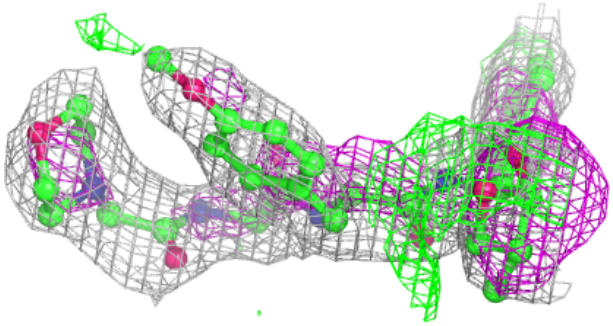
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
18	04C	V	302	42/42	0.95	0.14	35,39,72,73	0
15	MG	Z	301	1/1	0.95	0.10	63,63,63,63	0
18	04C	Y	303	42/42	0.96	0.15	25,40,59,59	0
17	MES	H	301	12/12	0.97	0.23	58,60,63,66	0
15	MG	N	201	1/1	0.97	0.10	46,46,46,46	0
15	MG	X	201	1/1	0.97	0.24	29,29,29,29	0
15	MG	I	302	1/1	0.97	0.08	53,53,53,53	0
15	MG	K	301	1/1	0.98	0.09	55,55,55,55	0
17	MES	Y	301	12/12	0.98	0.20	22,35,40,41	0
15	MG	I	301	1/1	0.98	0.08	51,51,51,51	0
17	MES	K	303	12/12	0.98	0.20	33,36,43,52	0
19	NA	K	302	1/1	0.98	0.11	26,26,26,26	0
16	CL	U	301	1/1	0.99	0.13	40,40,40,40	0
16	CL	G	302	1/1	0.99	0.15	43,43,43,43	0
19	NA	Y	302	1/1	0.99	0.11	12,12,12,12	0

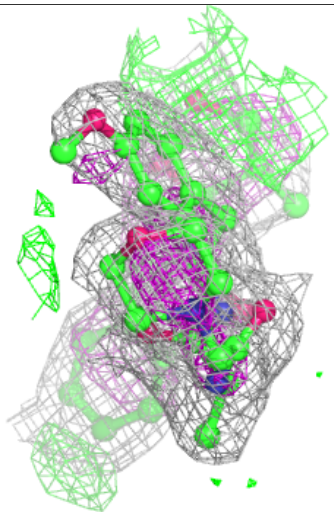
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 04C b 201:

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

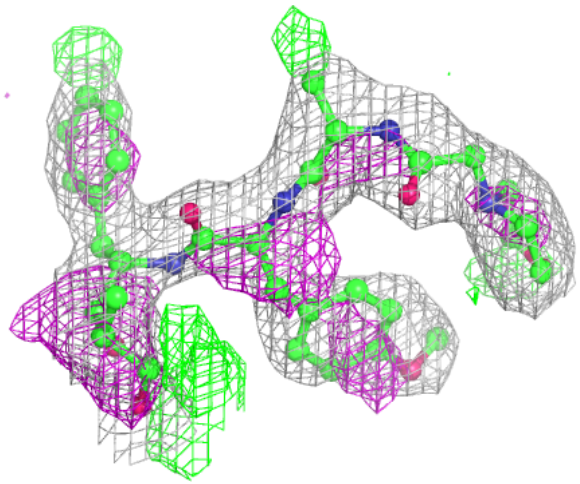


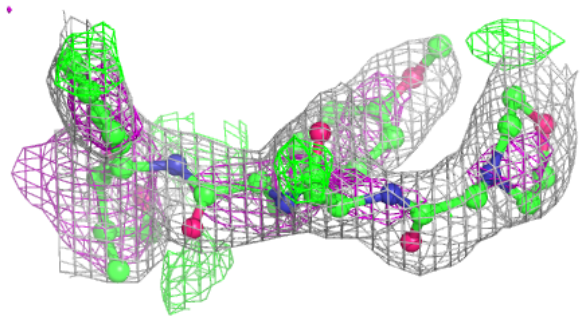


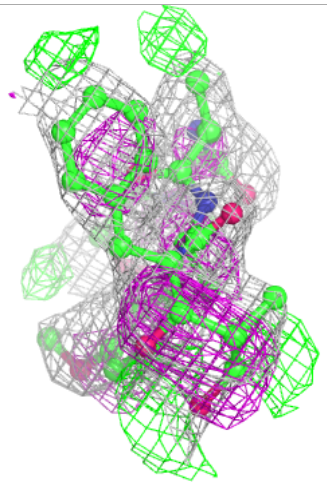


Electron density around 04C N 202:

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

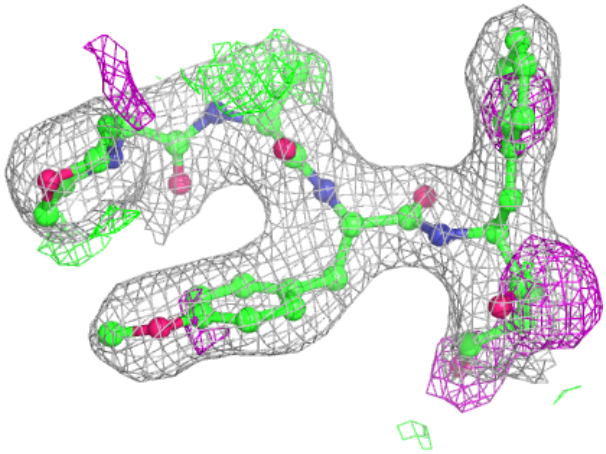


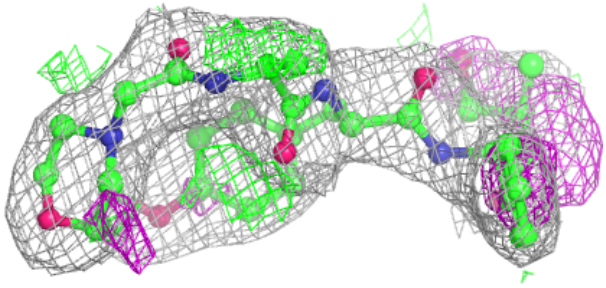


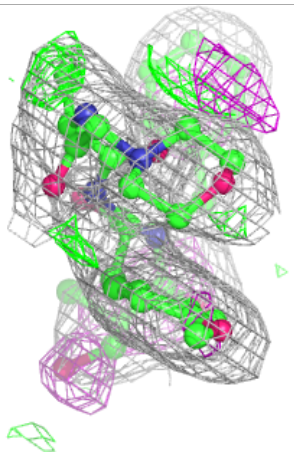


Electron density around 04C H 302:

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

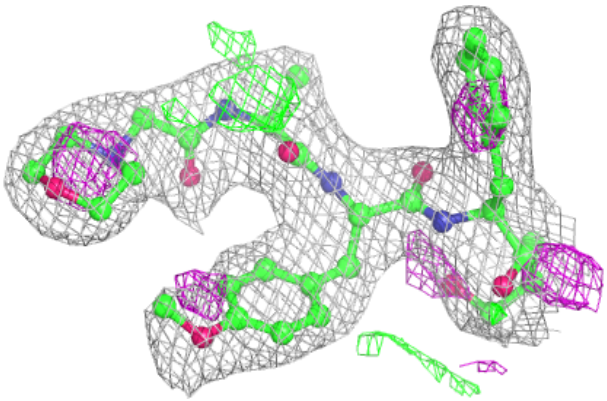


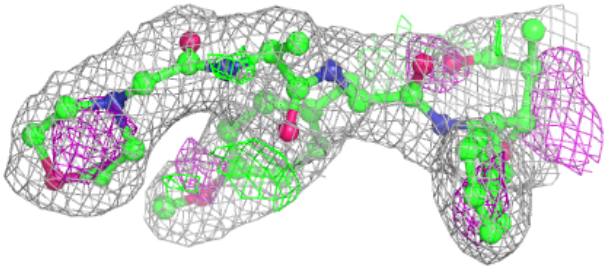


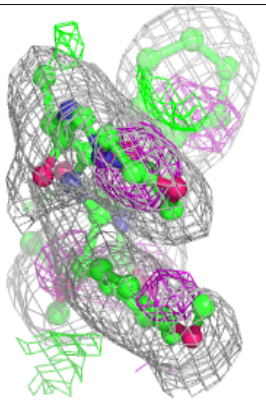


Electron density around 04C K 304:

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

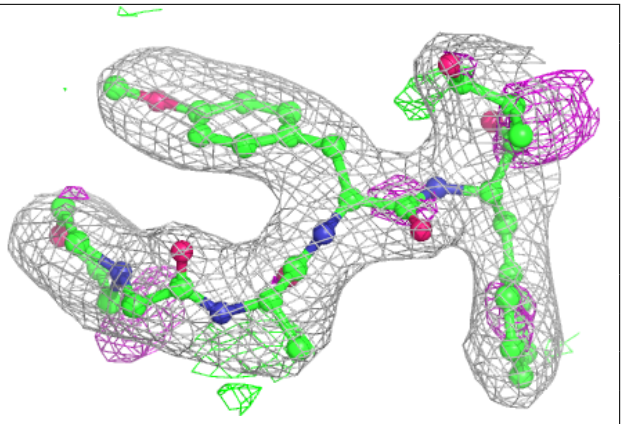


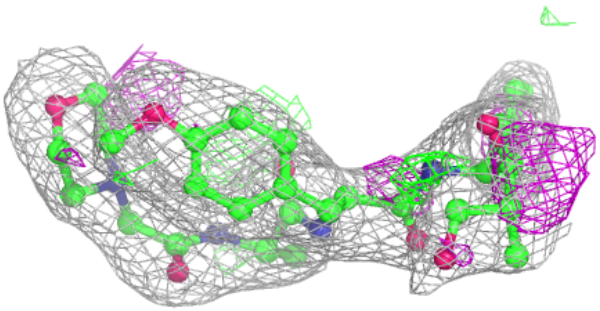


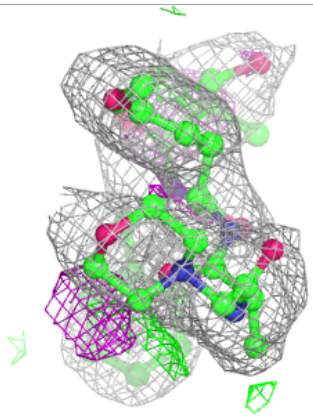


Electron density around 04C V 302:

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

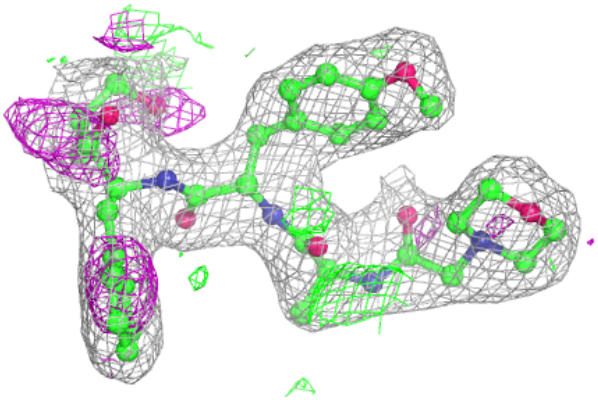


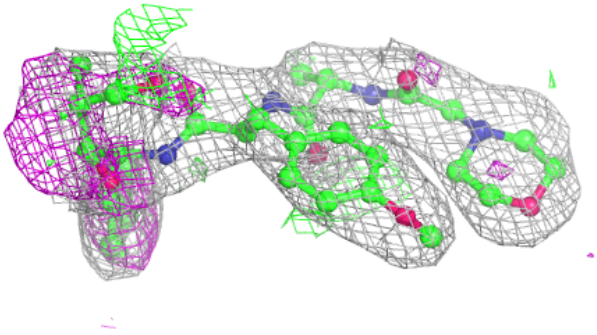


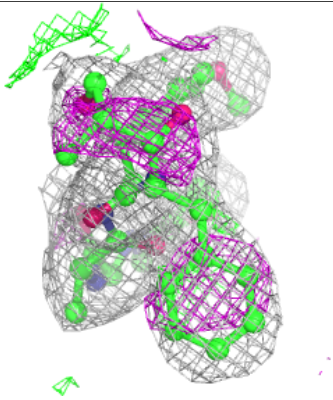


Electron density around 04C Y 303:

$2mF_o$-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.