wwPDB X-ray Structure Validation Summary Report (i)

Jan 14, 2024 – 07:46 am GMT

PDB ID	:	6HUU
Title	:	Yeast 20S proteasome with human beta2c (S171G) in complex with 29
Authors	:	Huber, E.M.; Groll, M.
Deposited on	:	2018-10-09
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:

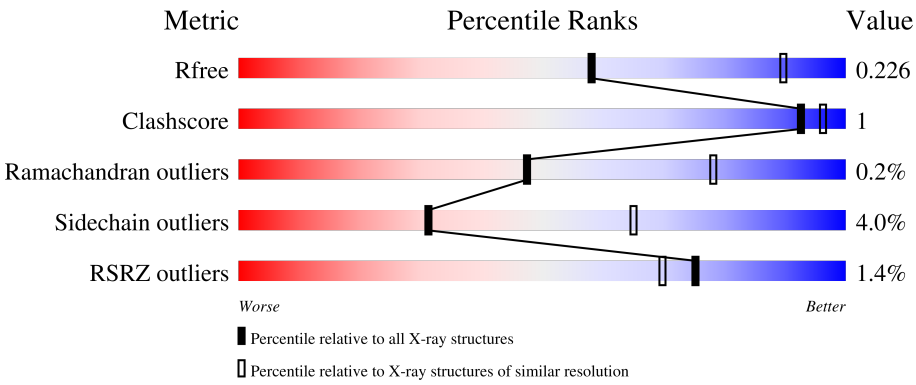
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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	2%, 93%, • •
1	O	250	2%, 93%, • •
2	B	258	3%, 87%, 7% • 5%
2	P	258	2%, 87%, 7% • 5%
3	C	254	2%, 87%, 6% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	6% 87% 6% 6%
4	D	260	85% 5% 10%
4	R	260	85% 5% 10%
5	E	234	2% 92% 6%
5	S	234	5% 93% 6%
6	F	288	% 80% 16%
6	T	288	80% 16%
7	G	252	88% 7%
7	U	252	88% 8%
8	H	234	% 87% 5% 6%
8	V	234	2% 88% 5% 6%
9	I	205	94%
9	W	205	94% 5%
10	J	198	% 89% 9%
10	X	198	% 89% 9%
11	K	212	84% 14%
11	Y	212	87% 11%
12	L	222	94% 5%
12	Z	222	% 92% 7%
13	M	246	2% 87% 9%
13	a	246	2% 88% 9%
14	N	196	% 96%
14	b	196	% 98%

2 Entry composition (i)

There are 19 unique types of molecules in this entry. The entry contains 49563 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	240	1842	1171	305	362	4	0	0	0
1	O	240	1842	1171	305	362	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-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	219	1648	1038	282	316	12	0	0	0
8	V	219	1648	1038	282	316	12	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	171	GLY	SER	engineered mutation	UNP Q99436
V	171	GLY	SER	engineered mutation	UNP Q99436

- 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
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	1644	1045	280	312	7	0	0	0
11	Y	212	1644	1045	280	312	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	224	1753	1108	300	338	7	0	0	0
13	a	224	1753	1108	300	338	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

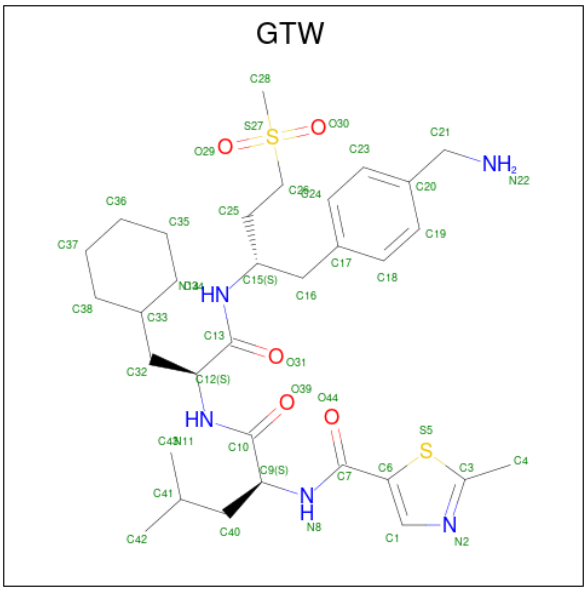
- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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	W	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).

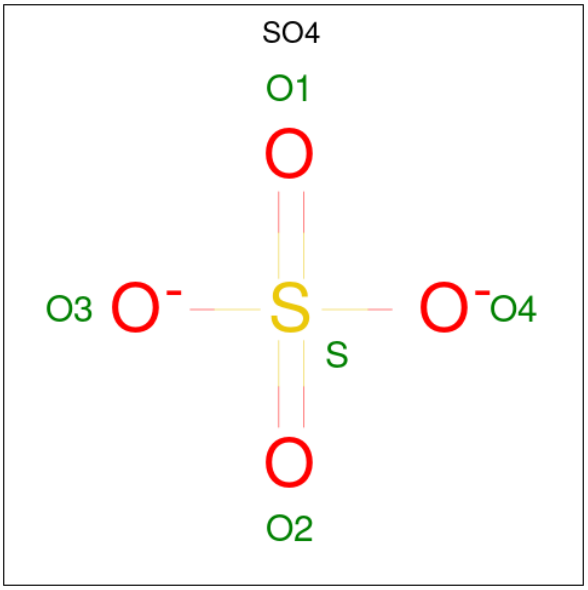
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	G	1	Total Cl 1 1	0	0
16	U	1	Total Cl 1 1	0	0

- Molecule 17 is {N}-[(2 {S})-1-[[(2 {S})-1-[[(2 {S})-1-[4-(aminomethyl)phenyl]-4-methylsulfonyl-butan-2-yl]amino]-3-cyclohexyl-1-oxidanylidene-propan-2-yl]amino]-4-methyl-1-oxidanylidene-pentan-2-yl]-2-methyl-1,3-thiazole-5-carboxamide (three-letter code: GTW) (formula: $C_{32}H_{49}N_5O_5S_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
17	H	1	44	32	5	5	2	0	0
17	K	1	44	32	5	5	2	0	0
17	V	1	44	32	5	5	2	0	0
17	Y	1	44	32	5	5	2	0	0

- Molecule 18 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
18	N	1	5	4	1	0	0
18	b	1	5	4	1	0	0

- Molecule 19 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
19	A	9	9	9	0	0
19	B	15	15	15	0	0
19	C	18	18	18	0	0
19	D	10	10	10	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
19	E	7	7	7	0	0
			Total	O		
19	F	12	12	12	0	0
			Total	O		
19	G	24	24	24	0	0
			Total	O		
19	H	16	16	16	0	0
			Total	O		
19	I	16	16	16	0	0
			Total	O		
19	J	22	22	22	0	0
			Total	O		
19	K	28	28	28	0	0
			Total	O		
19	L	17	17	17	0	0
			Total	O		
19	M	15	15	15	0	0
			Total	O		
19	N	20	20	20	0	0
			Total	O		
19	O	5	5	5	0	0
			Total	O		
19	P	13	13	13	0	0
			Total	O		
19	Q	15	15	15	0	0
			Total	O		
19	R	10	10	10	0	0
			Total	O		
19	S	6	6	6	0	0
			Total	O		
19	T	12	12	12	0	0
			Total	O		
19	U	23	23	23	0	0
			Total	O		
19	V	12	12	12	0	0
			Total	O		
19	W	15	15	15	0	0
			Total	O		
19	X	23	23	23	0	0
			Total	O		
19	Y	22	22	22	0	0

Continued on next page...

Continued from previous page...

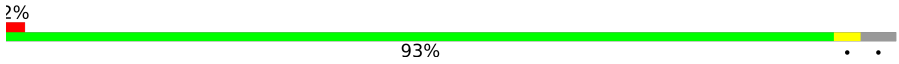
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
19	Z	12	12	12	0	0
			Total	O		
19	a	20	20	20	0	0
			Total	O		
19	b	14	14	14	0	0

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.

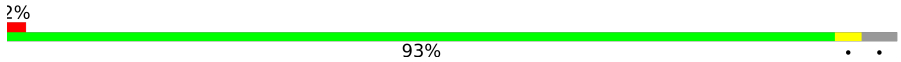
• Molecule 1: Proteasome subunit alpha type-2

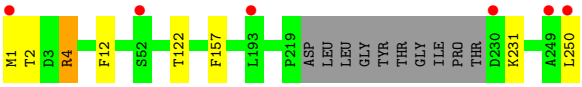
Chain A:



• Molecule 1: Proteasome subunit alpha type-2

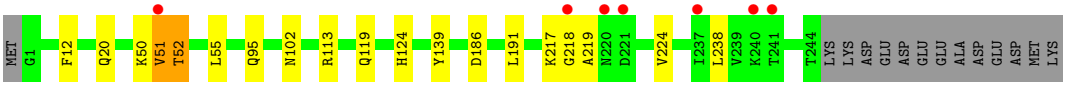
Chain O:





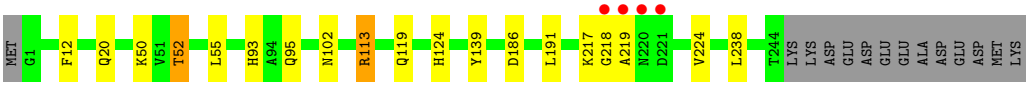
• Molecule 2: Proteasome subunit alpha type-3

Chain B:



• Molecule 2: Proteasome subunit alpha type-3

Chain P:

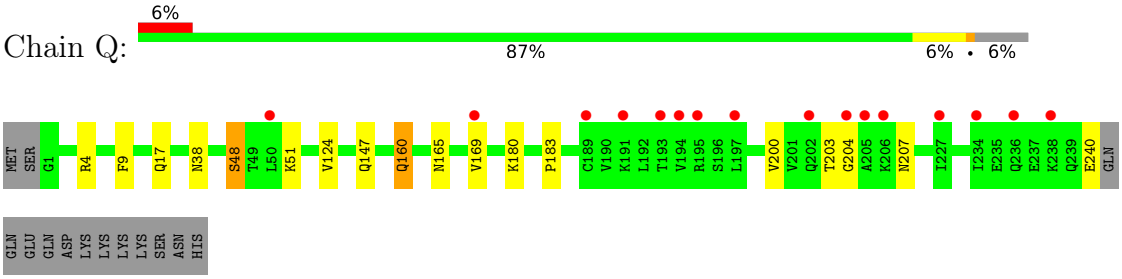


• Molecule 3: Proteasome subunit alpha type-4

Chain C:

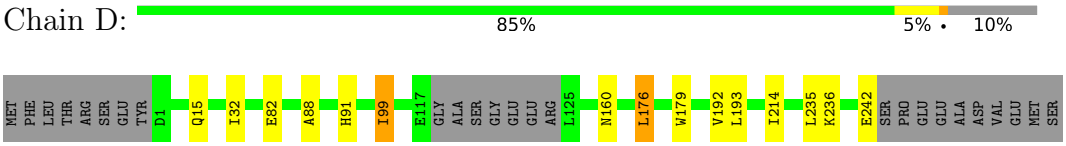
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



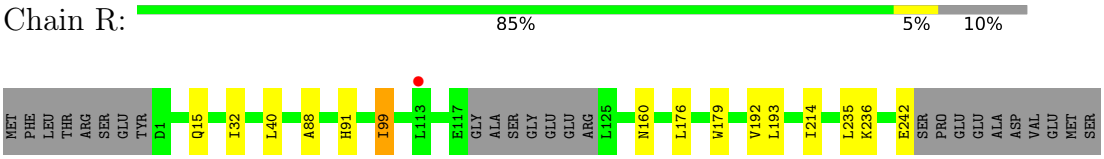
• Molecule 4: Proteasome subunit alpha type-5

Chain D:



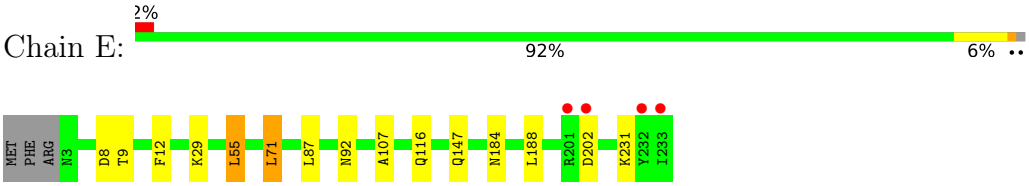
• Molecule 4: Proteasome subunit alpha type-5

Chain R:



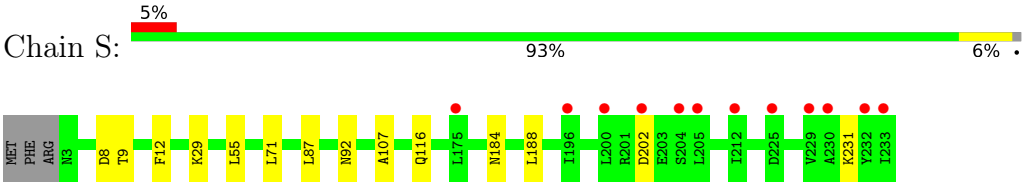
• Molecule 5: Proteasome subunit alpha type-6

Chain E:



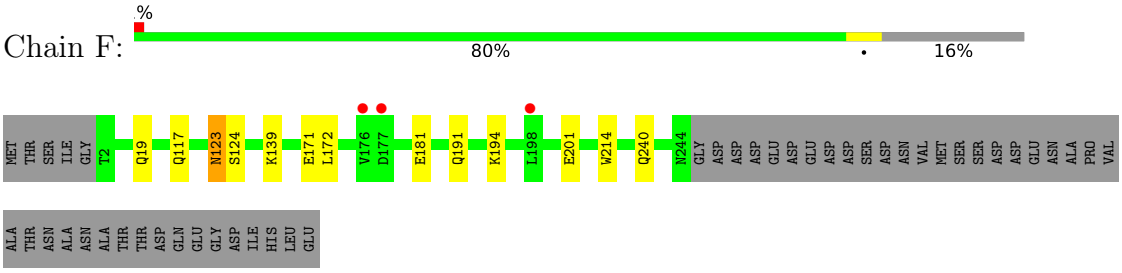
• Molecule 5: Proteasome subunit alpha type-6

Chain S:



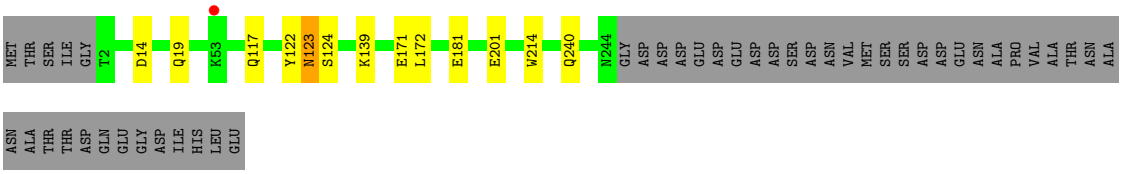
• Molecule 6: Probable proteasome subunit alpha type-7

Chain F:



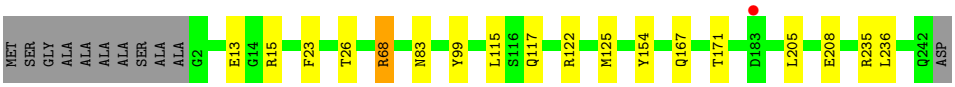
• Molecule 6: Probable proteasome subunit alpha type-7

Chain T:



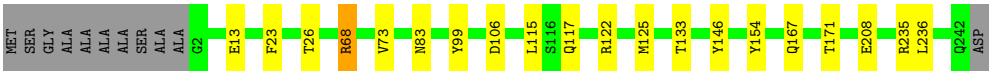
• Molecule 7: Proteasome subunit alpha type-1

Chain G:



• Molecule 7: Proteasome subunit alpha type-1

Chain U:

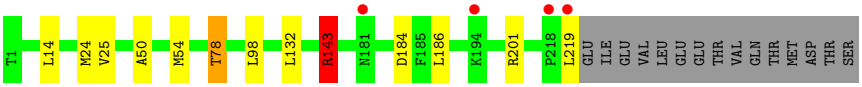


• Molecule 8: Proteasome subunit beta type-7

Chain H:

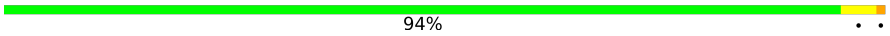
• Molecule 8: Proteasome subunit beta type-7

Chain V:



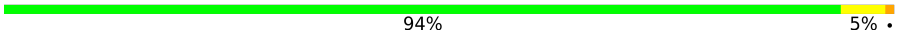
• Molecule 9: Proteasome subunit beta type-3

Chain I:



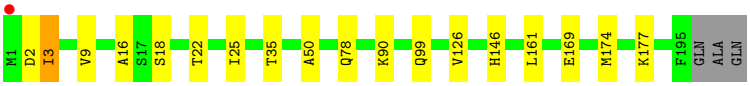
• Molecule 9: Proteasome subunit beta type-3

Chain W:



• Molecule 10: Proteasome subunit beta type-4

Chain J:

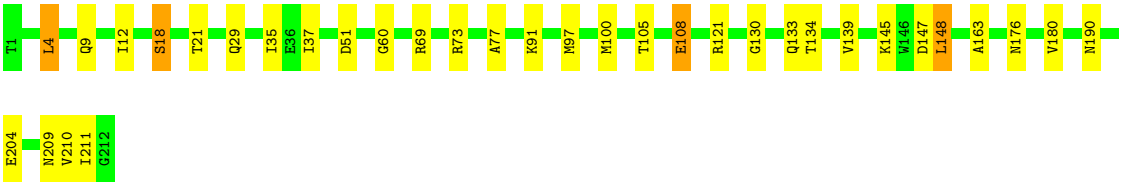


• Molecule 10: Proteasome subunit beta type-4

Chain X:

• Molecule 11: Proteasome subunit beta type-5

Chain K:

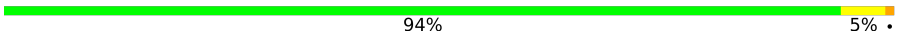


• Molecule 11: Proteasome subunit beta type-5

Chain Y:

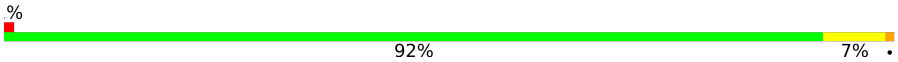
• Molecule 12: Proteasome subunit beta type-6

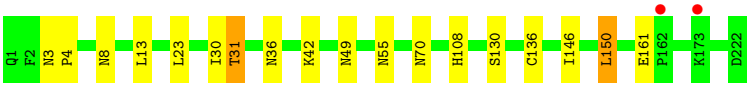
Chain L:



• Molecule 12: Proteasome subunit beta type-6

Chain Z:



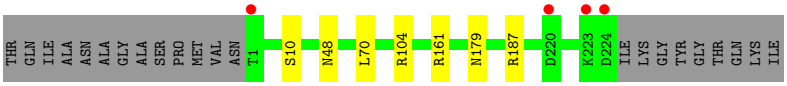


• Molecule 13: Proteasome subunit beta type-7

Chain M:

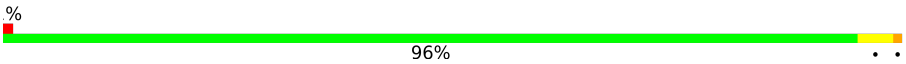
• Molecule 13: Proteasome subunit beta type-7

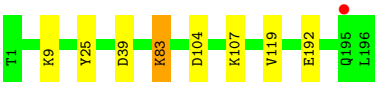
Chain a:



• Molecule 14: Proteasome subunit beta type-1

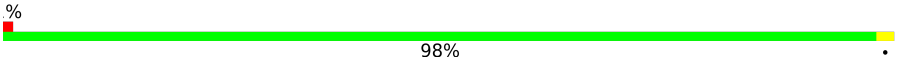
Chain N:

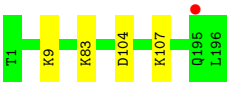




• Molecule 14: Proteasome subunit beta type-1

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	134.67Å 299.61Å 143.62Å 90.00° 112.36° 90.00°	Depositor
Resolution (Å)	15.00 – 2.80 48.93 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.2 (15.00-2.80) 99.2 (48.93-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.194 , 0.221 0.200 , 0.226	Depositor DCC
R_{free} test set	12770 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	62.2	Xtriage
Anisotropy	0.369	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 38.8	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	49563	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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: SO4, GTW, CL, 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	0.37	0/1876	0.81	3/2535 (0.1%)
1	O	0.37	0/1876	0.82	4/2535 (0.2%)
2	B	0.39	0/1934	0.65	1/2618 (0.0%)
2	P	0.38	0/1934	0.64	0/2618
3	C	0.37	0/1910	0.65	0/2586
3	Q	0.37	0/1910	0.64	0/2586
4	D	0.37	0/1837	0.60	0/2475
4	R	0.36	0/1837	0.60	0/2475
5	E	0.37	0/1800	0.59	1/2433 (0.0%)
5	S	0.36	0/1800	0.59	0/2433
6	F	0.38	0/1932	0.57	0/2609
6	T	0.37	0/1932	0.57	0/2609
7	G	0.39	0/1945	0.78	3/2634 (0.1%)
7	U	0.38	0/1945	0.81	3/2634 (0.1%)
8	H	0.36	0/1675	0.88	3/2267 (0.1%)
8	V	0.35	0/1675	0.87	3/2267 (0.1%)
9	I	0.37	0/1611	0.61	0/2174
9	W	0.37	0/1611	0.61	0/2174
10	J	0.37	0/1589	0.63	0/2142
10	X	0.37	0/1589	0.63	0/2142
11	K	0.37	0/1681	0.65	1/2274 (0.0%)
11	Y	0.36	0/1681	0.64	1/2274 (0.0%)
12	L	0.37	0/1795	0.61	0/2420
12	Z	0.37	0/1795	0.61	0/2420
13	M	0.39	0/1783	0.65	0/2420
13	a	0.38	0/1783	0.65	0/2420
14	N	0.36	0/1541	0.60	0/2087
14	b	0.35	0/1541	0.59	0/2087
All	All	0.37	0/49818	0.67	23/67348 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if

the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	O	0	1
7	U	0	1
8	V	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	U	68	ARG	NE-CZ-NH2	-22.10	109.25	120.30
1	A	4	ARG	NE-CZ-NH2	-20.76	109.92	120.30
1	O	4	ARG	NE-CZ-NH1	-20.42	110.09	120.30
8	V	143	ARG	NE-CZ-NH2	-19.98	110.31	120.30
8	H	143	ARG	NE-CZ-NH1	-18.92	110.84	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	4	ARG	Sidechain
1	O	4	ARG	Sidechain
7	U	68	ARG	Sidechain
8	V	143	ARG	Sidechain

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	1842	0	1855	2	0
1	O	1842	0	1855	3	0
2	B	1904	0	1904	6	0
2	P	1904	0	1904	9	0
3	C	1881	0	1895	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	Q	1881	0	1895	8	0
4	D	1813	0	1797	7	0
4	R	1813	0	1797	6	0
5	E	1773	0	1775	7	0
5	S	1773	0	1775	5	0
6	F	1892	0	1883	4	0
6	T	1892	0	1883	4	1
7	G	1907	0	1901	5	1
7	U	1907	0	1901	5	0
8	H	1648	0	1672	7	0
8	V	1648	0	1672	5	0
9	I	1581	0	1574	8	0
9	W	1581	0	1574	8	0
10	J	1561	0	1569	10	0
10	X	1561	0	1569	11	0
11	K	1644	0	1594	18	0
11	Y	1644	0	1594	12	0
12	L	1757	0	1711	7	0
12	Z	1757	0	1711	10	0
13	M	1753	0	1754	1	0
13	a	1753	0	1754	0	0
14	N	1512	0	1481	4	0
14	b	1512	0	1481	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	W	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	44	0	0	1	0
17	K	44	0	0	0	0
17	V	44	0	0	0	0
17	Y	44	0	0	0	0
18	N	5	0	0	0	0
18	b	5	0	0	0	0
19	A	9	0	0	0	0
19	B	15	0	0	0	0
19	C	18	0	0	0	0
19	D	10	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
19	E	7	0	0	1	0
19	F	12	0	0	0	0
19	G	24	0	0	1	0
19	H	16	0	0	0	0
19	I	16	0	0	0	0
19	J	22	0	0	1	0
19	K	28	0	0	0	0
19	L	17	0	0	0	0
19	M	15	0	0	0	0
19	N	20	0	0	1	0
19	O	5	0	0	0	0
19	P	13	0	0	2	0
19	Q	15	0	0	0	0
19	R	10	0	0	0	0
19	S	6	0	0	0	0
19	T	12	0	0	0	0
19	U	23	0	0	0	0
19	V	12	0	0	0	0
19	W	15	0	0	0	0
19	X	23	0	0	1	0
19	Y	22	0	0	0	0
19	Z	12	0	0	0	0
19	a	20	0	0	0	0
19	b	14	0	0	0	0
All	All	49563	0	48730	134	1

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 134 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)
5:E:92:ASN:HD21	12:L:70:ASN:HD21	1.17	0.89
5:S:92:ASN:HD21	12:Z:70:ASN:HD21	1.20	0.88
3:C:160:GLN:HE21	3:C:160:GLN:HA	1.55	0.72
3:Q:160:GLN:HE21	3:Q:160:GLN:HA	1.56	0.69
11:K:77:ALA:O	11:K:121:ARG:NH1	2.26	0.69

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:G:205:LEU:O	6:T:14:ASP:OD1[2_656]	1.85	0.35

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	236/250 (94%)	227 (96%)	8 (3%)	1 (0%)	34	66
1	O	236/250 (94%)	226 (96%)	9 (4%)	1 (0%)	34	66
2	B	242/258 (94%)	236 (98%)	4 (2%)	2 (1%)	19	49
2	P	242/258 (94%)	235 (97%)	5 (2%)	2 (1%)	19	49
3	C	238/254 (94%)	235 (99%)	2 (1%)	1 (0%)	34	66
3	Q	238/254 (94%)	235 (99%)	2 (1%)	1 (0%)	34	66
4	D	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
4	R	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
5	E	229/234 (98%)	219 (96%)	10 (4%)	0	100	100
5	S	229/234 (98%)	220 (96%)	9 (4%)	0	100	100
6	F	241/288 (84%)	239 (99%)	2 (1%)	0	100	100
6	T	241/288 (84%)	240 (100%)	1 (0%)	0	100	100
7	G	239/252 (95%)	237 (99%)	2 (1%)	0	100	100
7	U	239/252 (95%)	236 (99%)	3 (1%)	0	100	100
8	H	217/234 (93%)	215 (99%)	2 (1%)	0	100	100
8	V	217/234 (93%)	215 (99%)	2 (1%)	0	100	100
9	I	202/205 (98%)	194 (96%)	8 (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%)	201 (96%)	7 (3%)	2 (1%)	15	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
11	Y	210/212 (99%)	202 (96%)	6 (3%)	2 (1%)	15	44
12	L	220/222 (99%)	214 (97%)	6 (3%)	0	100	100
12	Z	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
13	M	222/246 (90%)	213 (96%)	9 (4%)	0	100	100
13	a	222/246 (90%)	213 (96%)	9 (4%)	0	100	100
14	N	194/196 (99%)	190 (98%)	4 (2%)	0	100	100
14	b	194/196 (99%)	190 (98%)	4 (2%)	0	100	100
All	All	6228/6618 (94%)	6075 (98%)	139 (2%)	14 (0%)	47	78

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	231	LYS
2	B	52	THR
2	B	218	GLY
1	O	231	LYS
2	P	52	THR

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	201/209 (96%)	197 (98%)	4 (2%)	55	84
1	O	201/209 (96%)	197 (98%)	4 (2%)	55	84
2	B	203/216 (94%)	193 (95%)	10 (5%)	25	57
2	P	203/216 (94%)	194 (96%)	9 (4%)	28	61
3	C	212/226 (94%)	203 (96%)	9 (4%)	30	63
3	Q	212/226 (94%)	203 (96%)	9 (4%)	30	63
4	D	194/215 (90%)	187 (96%)	7 (4%)	35	69
4	R	194/215 (90%)	186 (96%)	8 (4%)	30	64
5	E	190/193 (98%)	180 (95%)	10 (5%)	22	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	S	190/193 (98%)	180 (95%)	10 (5%)	22	54
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%)	196 (95%)	10 (5%)	25	57
7	U	206/210 (98%)	196 (95%)	10 (5%)	25	57
8	H	179/194 (92%)	169 (94%)	10 (6%)	21	51
8	V	179/194 (92%)	171 (96%)	8 (4%)	27	60
9	I	172/173 (99%)	167 (97%)	5 (3%)	42	76
9	W	172/173 (99%)	166 (96%)	6 (4%)	36	70
10	J	173/175 (99%)	167 (96%)	6 (4%)	36	70
10	X	173/175 (99%)	167 (96%)	6 (4%)	36	70
11	K	169/169 (100%)	161 (95%)	8 (5%)	26	59
11	Y	169/169 (100%)	161 (95%)	8 (5%)	26	59
12	L	185/185 (100%)	177 (96%)	8 (4%)	29	62
12	Z	185/185 (100%)	177 (96%)	8 (4%)	29	62
13	M	192/208 (92%)	185 (96%)	7 (4%)	35	69
13	a	192/208 (92%)	185 (96%)	7 (4%)	35	69
14	N	162/162 (100%)	158 (98%)	4 (2%)	47	80
14	b	162/162 (100%)	158 (98%)	4 (2%)	47	80
All	All	5278/5548 (95%)	5065 (96%)	213 (4%)	31	65

5 of 213 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	P	113	ARG
5	S	184	ASN
12	Z	136	CYS
2	P	238	LEU
4	R	99	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 126 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	18	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
11	Y	85	ASN
2	P	176	GLN
11	Y	9	GLN
12	Z	80	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 16 ligands modelled in this entry, 10 are monoatomic - leaving 6 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	$\#\|Z\| > 2$
17	GTW	K	301	11	43,46,46	1.44	6 (13%)	55,63,63	1.59	11 (20%)
17	GTW	H	301	8	43,46,46	1.61	8 (18%)	55,63,63	1.53	14 (25%)
18	SO4	b	201	-	4,4,4	0.29	0	6,6,6	0.11	0
17	GTW	Y	301	11	43,46,46	1.47	6 (13%)	55,63,63	1.66	12 (21%)
18	SO4	N	202	-	4,4,4	0.29	0	6,6,6	0.13	0
17	GTW	V	301	8	43,46,46	1.49	6 (13%)	55,63,63	1.59	12 (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
17	GTW	V	301	8	-	13/40/52/52	0/3/3/3
17	GTW	K	301	11	-	6/40/52/52	0/3/3/3
17	GTW	Y	301	11	-	6/40/52/52	0/3/3/3
17	GTW	H	301	8	-	13/40/52/52	0/3/3/3

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	Y	301	GTW	C16-C17	-4.65	1.40	1.51
17	K	301	GTW	C16-C17	-4.64	1.40	1.51
17	H	301	GTW	C3-S5	-4.38	1.60	1.73
17	Y	301	GTW	O29-S27	4.20	1.53	1.44
17	V	301	GTW	O29-S27	4.16	1.53	1.44

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
17	V	301	GTW	O30-S27-C26	-5.88	104.23	108.34
17	Y	301	GTW	O30-S27-C26	-4.73	105.03	108.34
17	H	301	GTW	O30-S27-C26	-4.68	105.07	108.34
17	K	301	GTW	O29-S27-O30	-4.13	108.27	117.09
17	V	301	GTW	C28-S27-C26	3.73	119.60	105.21

There are no chirality outliers.

5 of 38 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	H	301	GTW	C25-C26-S27-C28
17	H	301	GTW	C25-C26-S27-O30
17	H	301	GTW	C25-C26-S27-O29
17	K	301	GTW	C25-C26-S27-C28
17	K	301	GTW	C25-C26-S27-O30

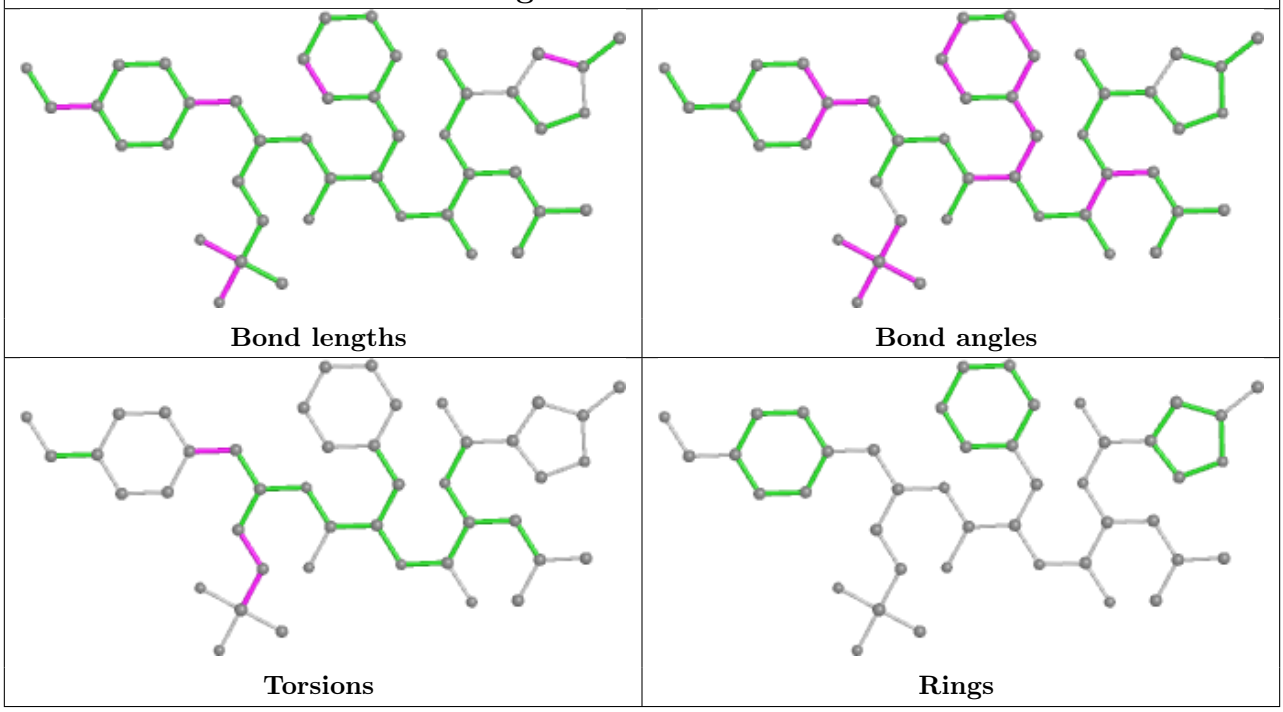
There are no ring outliers.

1 monomer is involved in 1 short contact:

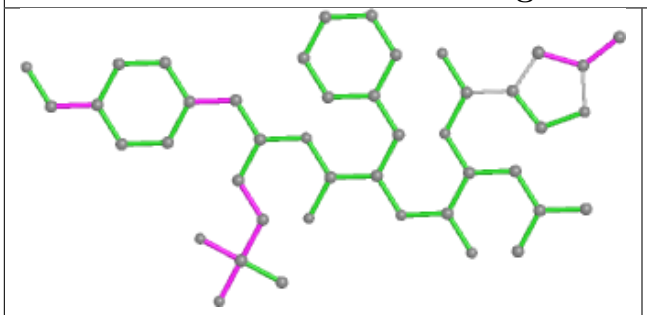
Mol	Chain	Res	Type	Clashes	Symm-Clashes
17	H	301	GTW	1	0

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.

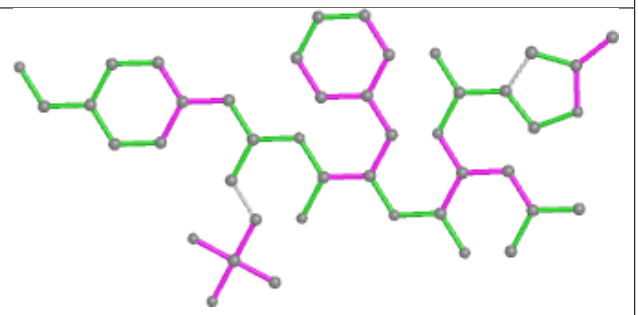
Ligand GTW K 301



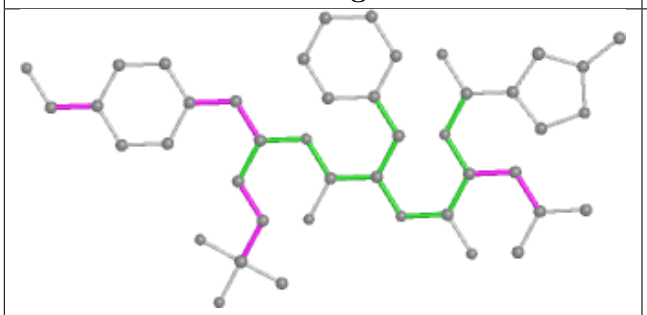
Ligand GTW H 301



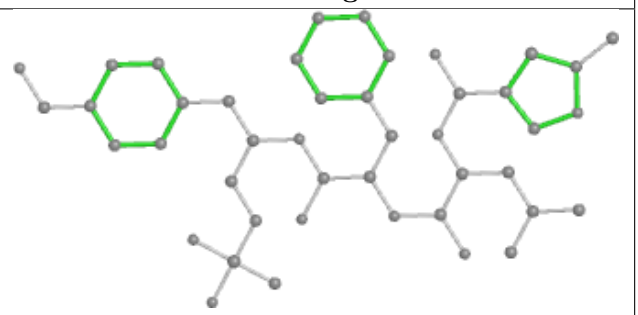
Bond lengths



Bond angles

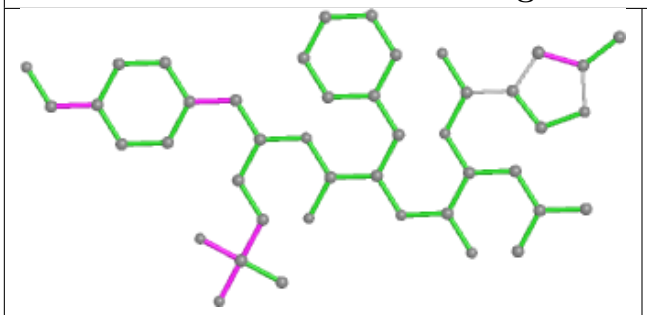


Torsions

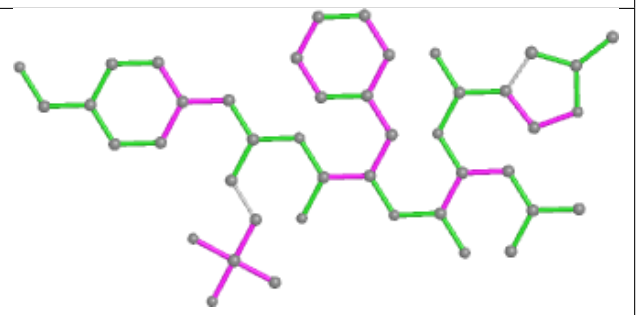


Rings

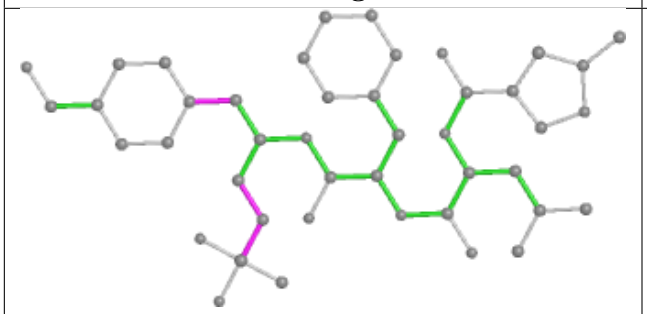
Ligand GTW Y 301



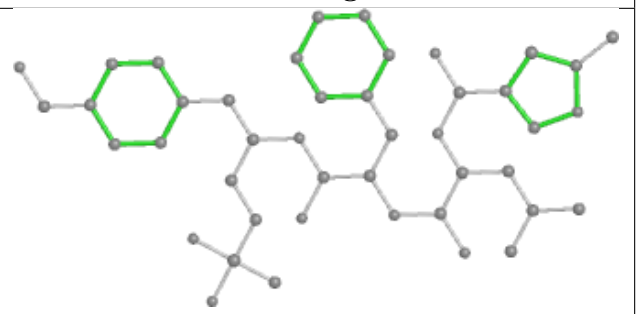
Bond lengths



Bond angles

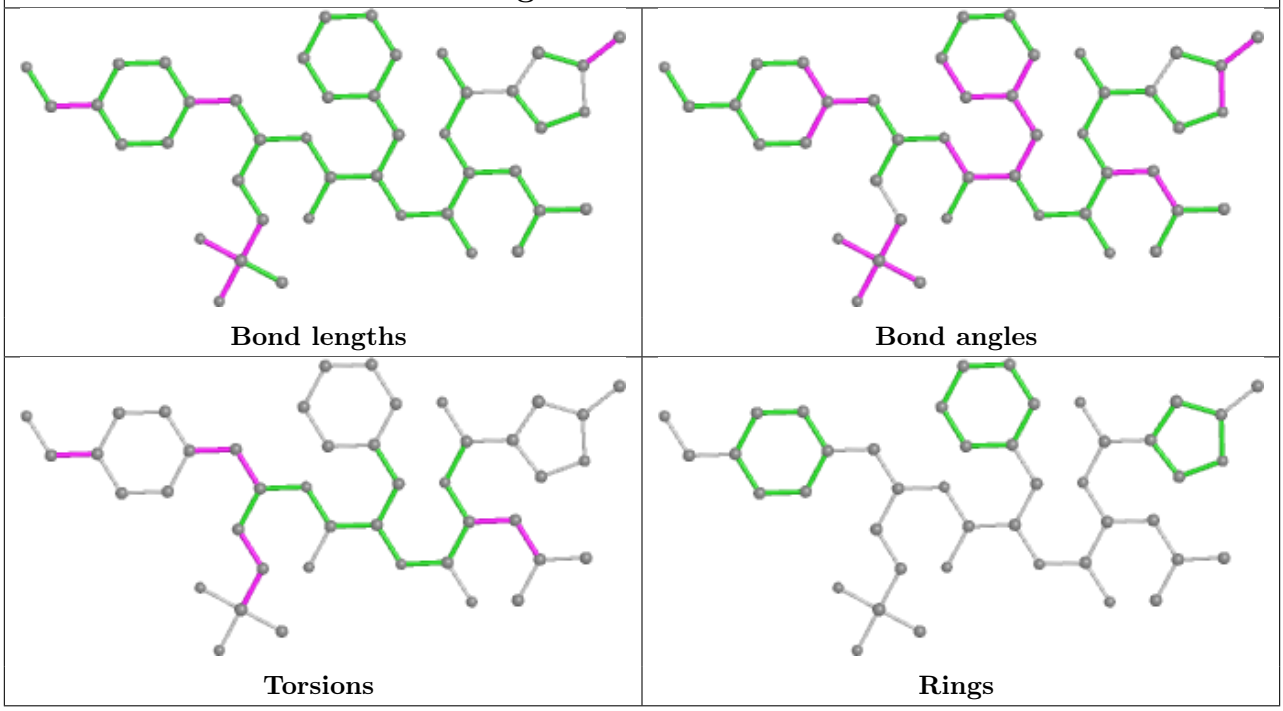


Torsions



Rings

Ligand GTW V 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	240/250 (96%)	-0.24	4 (1%)	70	63	51, 75, 124, 160	0
1	O	240/250 (96%)	-0.23	6 (2%)	57	47	54, 77, 125, 160	0
2	B	244/258 (94%)	-0.24	7 (2%)	51	41	50, 74, 129, 180	0
2	P	244/258 (94%)	-0.30	4 (1%)	72	66	55, 75, 122, 175	0
3	C	240/254 (94%)	-0.26	5 (2%)	63	54	50, 74, 135, 168	0
3	Q	240/254 (94%)	0.04	16 (6%)	17	10	62, 91, 170, 199	0
4	D	235/260 (90%)	-0.43	0	100	100	55, 76, 110, 151	0
4	R	235/260 (90%)	-0.30	1 (0%)	92	91	61, 84, 126, 158	0
5	E	231/234 (98%)	-0.18	4 (1%)	70	63	61, 84, 122, 162	0
5	S	231/234 (98%)	-0.04	12 (5%)	27	18	59, 88, 134, 170	0
6	F	243/288 (84%)	-0.32	3 (1%)	79	73	54, 78, 132, 159	0
6	T	243/288 (84%)	-0.30	1 (0%)	92	91	50, 81, 133, 163	0
7	G	241/252 (95%)	-0.31	1 (0%)	92	91	53, 76, 118, 178	0
7	U	241/252 (95%)	-0.43	0	100	100	54, 72, 112, 160	0
8	H	219/234 (93%)	-0.29	2 (0%)	84	80	49, 70, 126, 141	0
8	V	219/234 (93%)	-0.23	4 (1%)	68	61	51, 70, 124, 161	0
9	I	204/205 (99%)	-0.46	0	100	100	43, 62, 93, 122	0
9	W	204/205 (99%)	-0.38	1 (0%)	91	88	44, 61, 90, 113	0
10	J	195/198 (98%)	-0.45	1 (0%)	91	88	46, 65, 94, 136	0
10	X	195/198 (98%)	-0.49	2 (1%)	82	77	47, 66, 92, 155	0
11	K	212/212 (100%)	-0.36	0	100	100	46, 64, 95, 122	0
11	Y	212/212 (100%)	-0.36	0	100	100	46, 69, 102, 125	0
12	L	222/222 (100%)	-0.44	1 (0%)	91	88	46, 66, 112, 140	0
12	Z	222/222 (100%)	-0.36	2 (0%)	84	80	45, 66, 114, 148	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
13	M	224/246 (91%)	-0.25	4 (1%)	68	61	49, 71, 102, 143	0
13	a	224/246 (91%)	-0.37	4 (1%)	68	61	47, 69, 98, 141	0
14	N	196/196 (100%)	-0.45	1 (0%)	91	88	48, 65, 103, 129	0
14	b	196/196 (100%)	-0.47	1 (0%)	91	88	47, 64, 98, 130	0
All	All	6292/6618 (95%)	-0.31	87 (1%)	75	70	43, 72, 123, 199	0

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
13	M	1	THR	6.7
3	Q	50	LEU	6.2
1	A	1	MET	6.0
13	a	1	THR	5.2
1	O	249	ALA	5.1

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	I	302	1/1	0.81	0.38	95,95,95,95	0
16	CL	U	301	1/1	0.86	0.19	88,88,88,88	0
15	MG	Z	301	1/1	0.90	0.27	76,76,76,76	0
16	CL	G	302	1/1	0.90	0.21	81,81,81,81	0
15	MG	L	301	1/1	0.90	0.12	85,85,85,85	0
15	MG	W	301	1/1	0.94	0.39	87,87,87,87	0
17	GTW	H	301	44/44	0.94	0.18	48,66,74,78	0

Continued on next page...

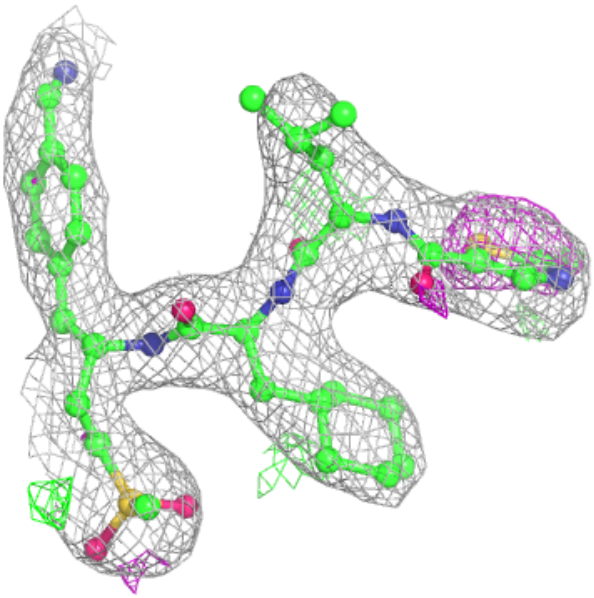
Continued from previous page...

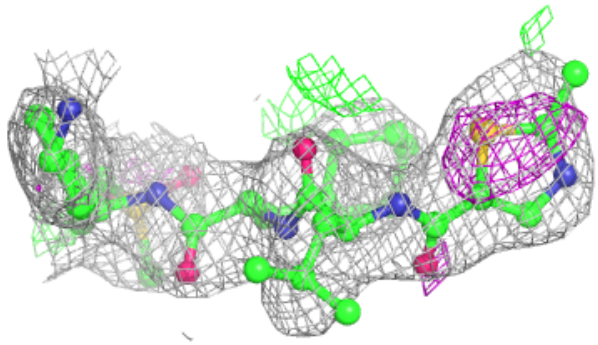
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
17	GTW	Y	301	44/44	0.94	0.16	37,50,54,58	0
18	SO4	N	202	5/5	0.94	0.29	49,57,67,69	5
18	SO4	b	201	5/5	0.94	0.37	51,52,59,67	5
17	GTW	V	301	44/44	0.95	0.19	43,60,65,69	0
17	GTW	K	301	44/44	0.95	0.17	34,48,53,60	0
15	MG	N	201	1/1	0.96	0.11	57,57,57,57	0
15	MG	I	301	1/1	0.97	0.28	81,81,81,81	0
15	MG	G	301	1/1	0.99	0.04	66,66,66,66	0
15	MG	K	302	1/1	0.99	0.15	68,68,68,68	0

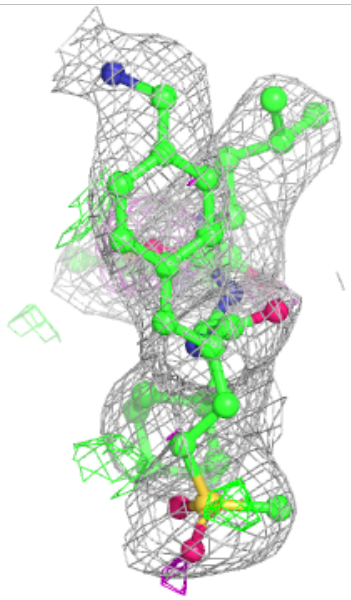
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around GTW 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)

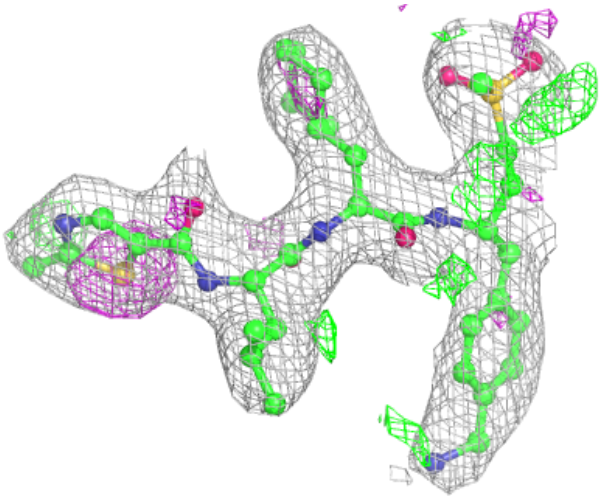


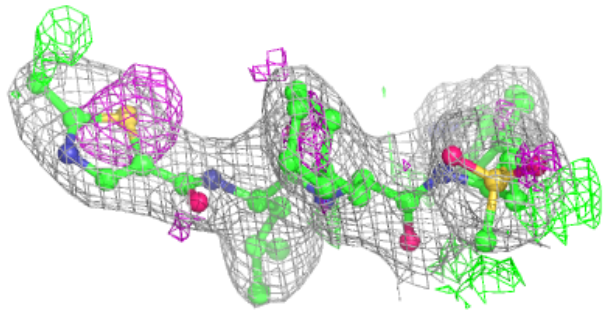


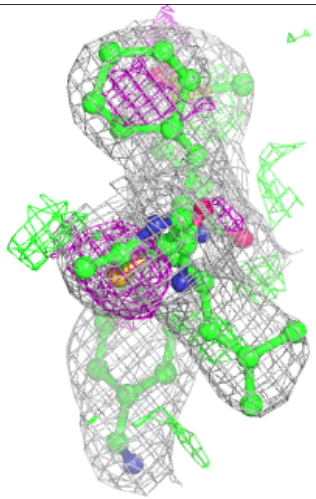


Electron density around GTW Y 301:

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

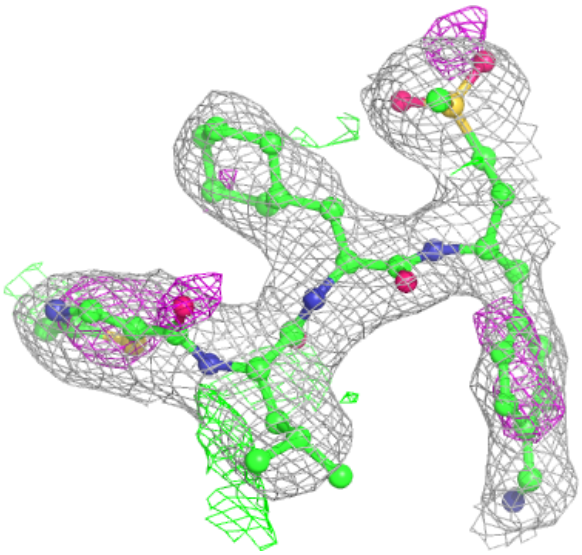


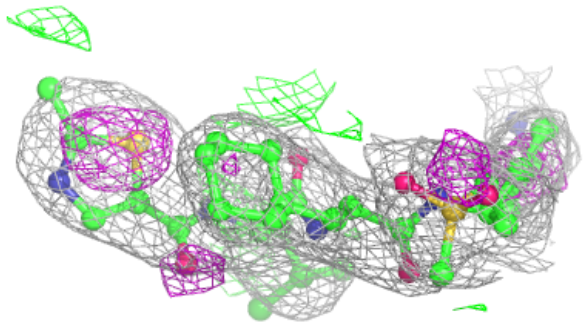


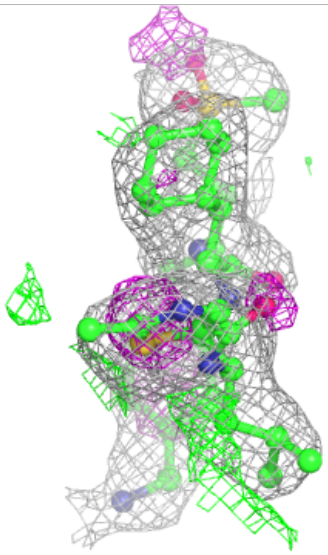


Electron density around GTW V 301:

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

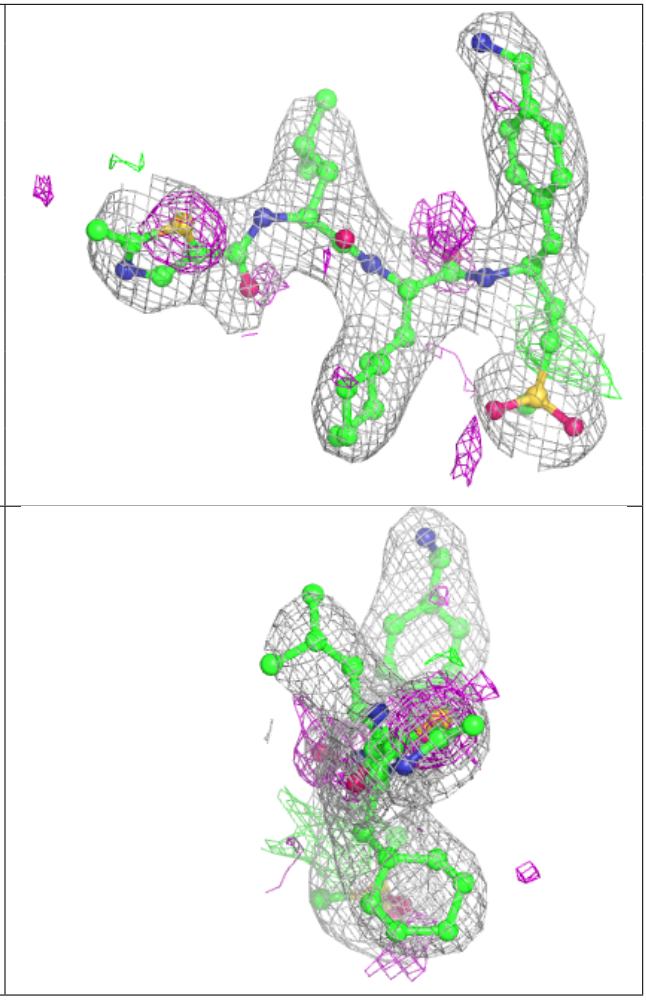






Electron density around GTW K 301:

$2mF_o$-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.