wwPDB X-ray Structure Validation Summary Report (i)

Jan 13, 2024 – 12:57 pm GMT

PDB ID	:	6HVY
Title	:	Yeast 20S proteasome in complex with 5 (7- and 6-membered ring)
Authors	:	Huber, E.M.; Groll, M.
Deposited on	:	2018-10-11
Resolution	:	2.70 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:

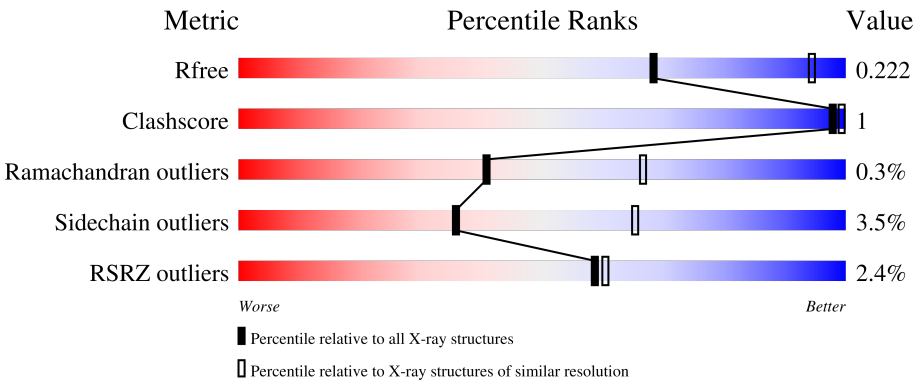
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	3% 98% •
1	O	250	2% 98% •
2	B	258	2% 89% 6% 5%
2	P	258	4% 89% 5% 5%
3	C	254	6% 88% 6% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	8% 89% 5% 6%
4	D	260	2% 86% 10%
4	R	260	2% 85% 5% 10%
5	E	234	3% 94% 5%
5	S	234	3% 92% 6%
6	F	288	2% 80% 16%
6	T	288	3% 80% 16%
7	G	252	2% 91% 5%
7	U	252	2% 91% 5%
8	H	232	3% 94%
8	V	232	2% 93%
9	I	205	94% 5%
9	W	205	% 94% 5%
10	J	198	2% 94%
10	X	198	2% 94%
11	K	212	94%
11	Y	212	94%
12	L	222	96%
12	Z	222	% 96%
13	M	246	% 92% 5%
13	a	246	92% 5%
14	N	196	% 94% 6%
14	b	196	% 97%

2 Entry composition (i)

There are 20 unique types of molecules in this entry. The entry contains 50142 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	226	1726	1086	299	334	7	0	1	0
8	V	226	1726	1086	299	334	7	0	1	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	1644	1045	280	312	7	0	0	0
			Total	C	N	O	S			
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
			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
			Total	Mg		
15	G	1	1	1	0	0
			Total	Mg		
15	I	2	2	2	0	0

Continued on next page...

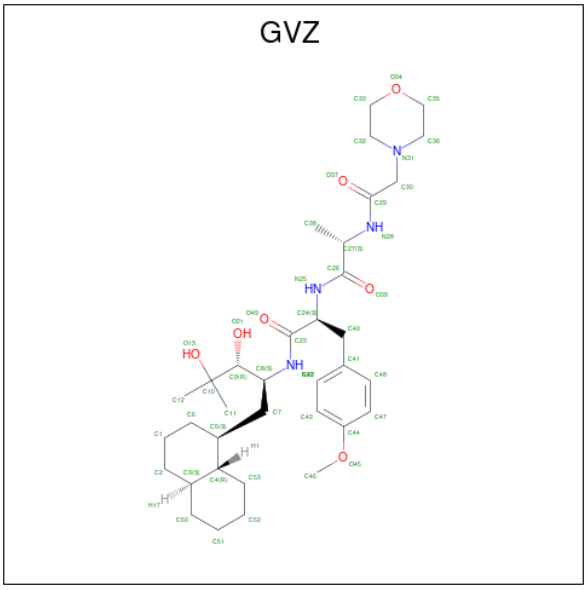
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
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 (2 {S})- {N}-[(2 {S},3 {R})-1-[(1 {S},4 {a} {S},8 {a} {R})-1,2,3,4,4 {a} ,5,6,7,8,8 {a}-decahydronaphthalen-1-yl]-4-methyl-3,4-bis(oxidanyl)pentan-2-yl]-3-(4-met hoxyphenyl)-2-[[(2 {S})-2-(2-morpholin-4-ylethanoylamino)propanoyl]amino]propanamide (three-letter code: GVZ) (formula: $C_{35}H_{56}N_4O_7$).



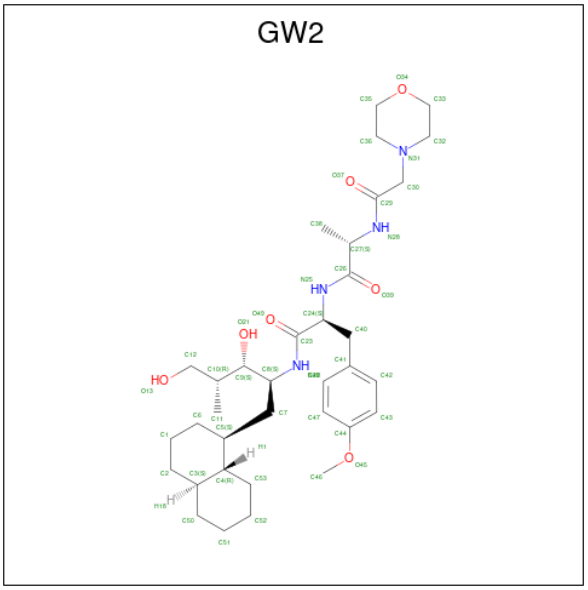
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	H	1	Total C N O 46 35 4 7	0	1

Continued on next page...

Continued from previous page...

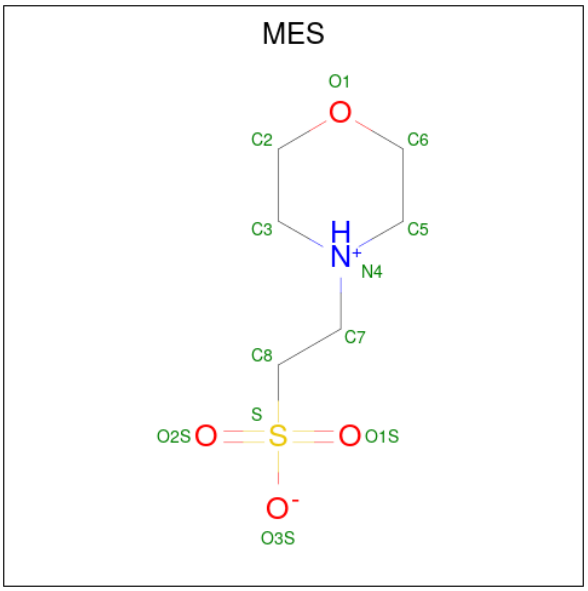
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
17	V	1	46	35	4	7	0	1

- Molecule 18 is (2 {S})- {N}-[(2 {S},3 {S},4 {R})-1-[(1 {S},4 {a} {S},8 {a} {R})-1,2,3,4,4 {a},5,6,7,8,8 {a}-decahydronaphthalen-1-yl]-4-methyl-3,5-bis(oxidanyl)pentan-2-yl]-3-(4-methoxyphenyl)-2-[[(2 {S})-2-(2-morpholin-4-ylethanoylamino)propanoyl]amino]propanamide (three-letter code: GW2) (formula: $C_{35}H_{56}N_4O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
18	H	1	46	35	4	7	0	1
18	V	1	46	35	4	7	0	1

- Molecule 19 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
19	H	1	12	6	1	4	1	0	0
			Total	C	N	O	S		
19	V	1	12	6	1	4	1	0	0

- Molecule 20 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
20	A	26	26	26	0	0
			Total	O		
20	B	16	16	16	0	0
			Total	O		
20	C	20	20	20	0	0
			Total	O		
20	D	4	4	4	0	0
			Total	O		
20	E	9	9	9	0	0
			Total	O		
20	F	16	16	16	0	0
			Total	O		
20	G	23	23	23	0	0
			Total	O		
20	H	28	28	28	0	0
			Total	O		
20	I	30	30	30	0	0
			Total	O		
20	J	19	19	19	0	0

Continued on next page...

Continued from previous page...

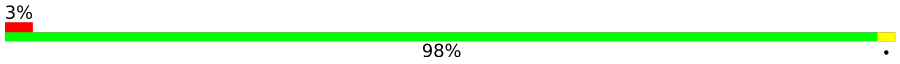
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
20	K	21	Total O 21 21	0	0
20	L	22	Total O 22 22	0	0
20	M	22	Total O 22 22	0	0
20	N	19	Total O 19 19	0	0
20	O	13	Total O 13 13	0	0
20	P	25	Total O 25 25	0	0
20	Q	10	Total O 10 10	0	0
20	R	20	Total O 20 20	0	0
20	S	13	Total O 13 13	0	0
20	T	12	Total O 12 12	0	0
20	U	22	Total O 22 22	0	0
20	V	18	Total O 18 18	0	0
20	W	17	Total O 17 17	0	0
20	X	24	Total O 24 24	0	0
20	Y	33	Total O 33 33	0	0
20	Z	21	Total O 21 21	0	0
20	a	25	Total O 25 25	0	0
20	b	16	Total O 16 16	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.

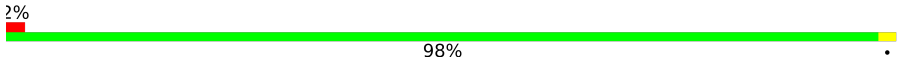
• Molecule 1: Proteasome subunit alpha type-2

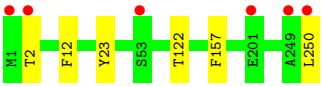
Chain A:



• Molecule 1: Proteasome subunit alpha type-2

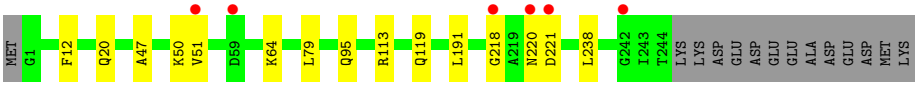
Chain O:





• Molecule 2: Proteasome subunit alpha type-3

Chain B:

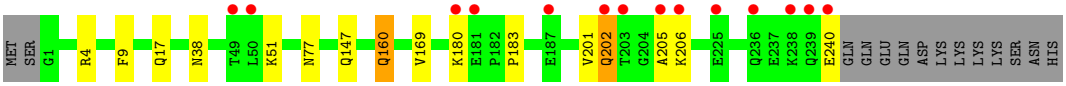


• Molecule 2: Proteasome subunit alpha type-3

Chain P:

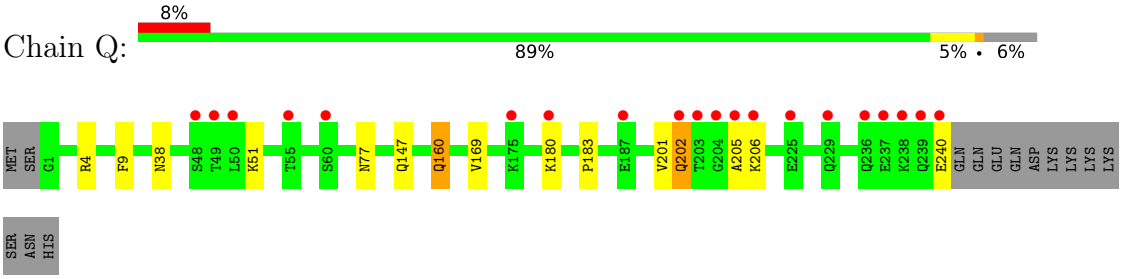
• Molecule 3: Proteasome subunit alpha type-4

Chain C:



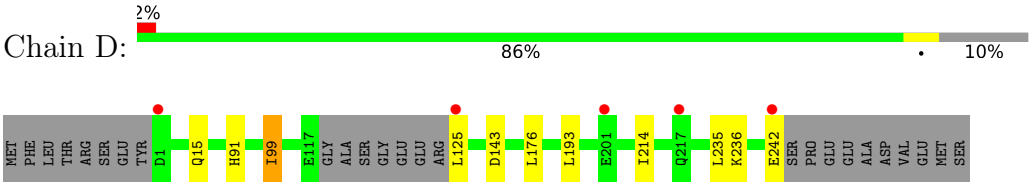
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



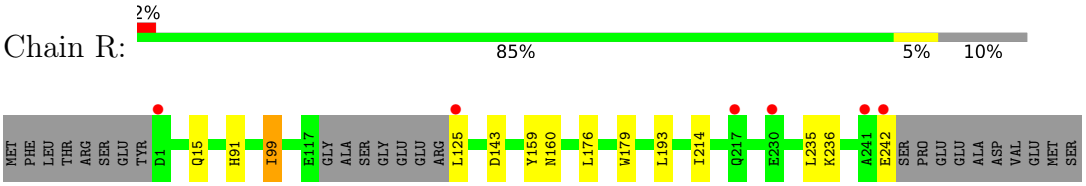
- Molecule 4: Proteasome subunit alpha type-5

Chain D:



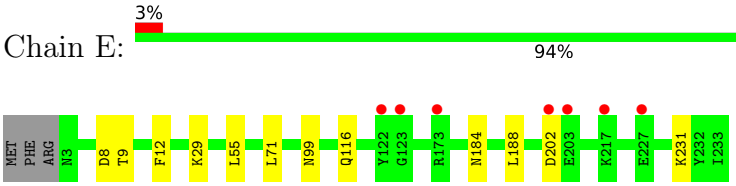
- Molecule 4: Proteasome subunit alpha type-5

Chain R:



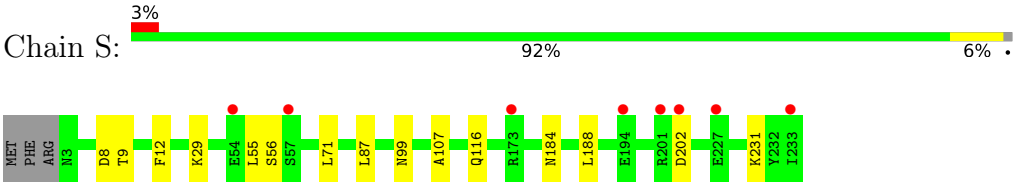
- Molecule 5: Proteasome subunit alpha type-6

Chain E:



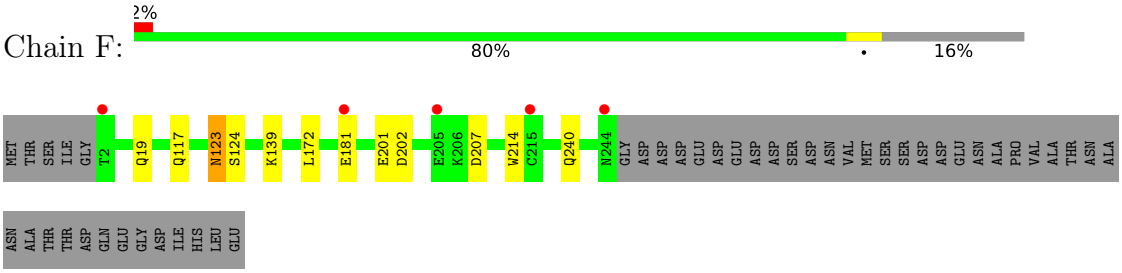
- Molecule 5: Proteasome subunit alpha type-6

Chain S:



- Molecule 6: Probable proteasome subunit alpha type-7

Chain F:



- Molecule 6: Probable proteasome subunit alpha type-7

Chain T: 3% 80% 16%

MET THR SER ILE GLY T2 Q19 K53 Q117 N123 S124 K139 L172 H178 E181 E201 D202 N203 K204 E205 K206 D207 W214 D230 Q240 K241 E242 I243 N244 GLY ASP ASP ASP GLU GLU ASP ASP SER ASP ASN VAL MET SER SER ASP ASP GLU

ASN ALA PRO VAL ALA THR ASN ALA ASN ALA THR THR ASP GLN GLU GLY ASP ILE HIS LEU GLU

- Molecule 7: Proteasome subunit alpha type-1

Chain G: 2% 91% 5%

MET SER GLY ALA ALA ALA ALA SER ALA ALA G2 F23 T26 R68 N75 N83 L115 S116 Q117 R122 M125 K179 E208 D222 R235 L236 A240 E241 Q242 ASP

- Molecule 7: Proteasome subunit alpha type-1

Chain U: 2% 91% 5%

MET SER GLY ALA ALA ALA ALA SER ALA ALA G2 P12 F23 T26 N75 N83 L115 S116 Q117 R122 M125 G206 T207 E208 D222 R235 L236 E241 Q242 ASP

- Molecule 8: Proteasome subunit beta type-2

Chain H: 3% 94%

T1 T2 I3 A16 Q22 G45 T52 L68 D145 R196 E197 E198 C221 D222 I223 Q224 E225 E226 GLN VAL ASP ILE THR ALA

- Molecule 8: Proteasome subunit beta type-2

Chain V: 2% 93%

T1 T2 I3 A16 Q22 N30 K33 G45 T52 L68 D145 R196 C221 D222 I223 Q224 E225 E226 GLN VAL ASP ILE THR ALA

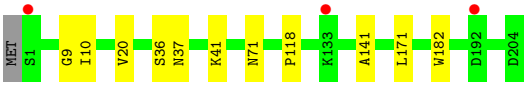
- Molecule 9: Proteasome subunit beta type-3

Chain I: 94% 5%

MET S1 G9 I10 V20 S36 N37 K41 N71 P118 A141 L171 W182 D204

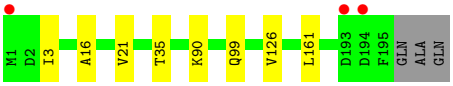
- Molecule 9: Proteasome subunit beta type-3

Chain W: % 94% 5%



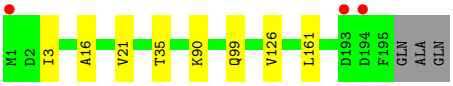
• Molecule 10: Proteasome subunit beta type-4

Chain J: 2% 94%



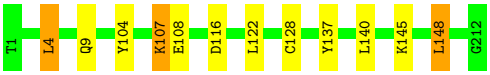
• Molecule 10: Proteasome subunit beta type-4

Chain X: 2% 94%



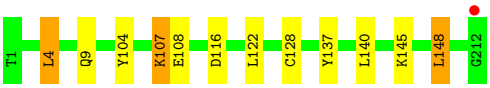
• Molecule 11: Proteasome subunit beta type-5

Chain K: 94%



• Molecule 11: Proteasome subunit beta type-5

Chain Y: 94%



• Molecule 12: Proteasome subunit beta type-6

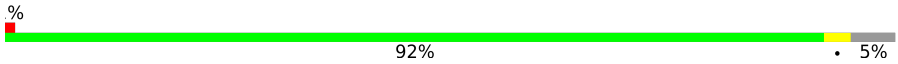
Chain L: 96%

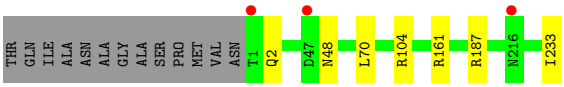
• Molecule 12: Proteasome subunit beta type-6

Chain Z: % 96%

• Molecule 13: Proteasome subunit beta type-7

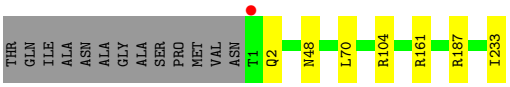
Chain M:





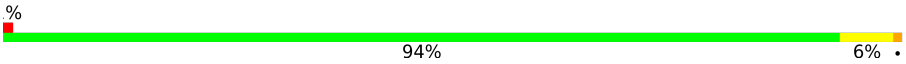
- Molecule 13: Proteasome subunit beta type-7

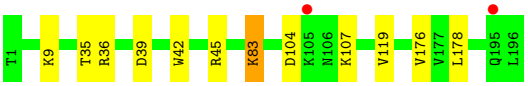
Chain a:



- Molecule 14: Proteasome subunit beta type-1

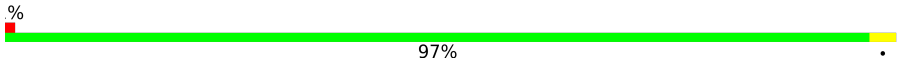
Chain N:

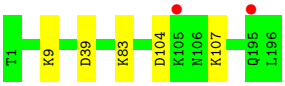




- Molecule 14: Proteasome subunit beta type-1

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	136.63Å 301.80Å 145.57Å 90.00° 113.06° 90.00°	Depositor
Resolution (Å)	15.00 – 2.70 15.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.5 (15.00-2.70) 98.1 (15.00-2.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.52 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.181 , 0.218 0.187 , 0.222	Depositor DCC
R_{free} test set	14418 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	55.4	Xtriage
Anisotropy	0.193	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 39.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	50142	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.51% 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: MES, GW2, GVZ, MG, CL

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	0.35	0/1952	0.55	0/2642
1	O	0.35	0/1952	0.55	0/2642
2	B	0.36	0/1934	0.59	0/2618
2	P	0.36	0/1934	0.59	0/2618
3	C	0.35	0/1910	0.61	0/2586
3	Q	0.35	0/1910	0.60	0/2586
4	D	0.34	0/1837	0.57	0/2475
4	R	0.34	0/1837	0.56	0/2475
5	E	0.35	0/1800	0.56	0/2433
5	S	0.35	0/1800	0.56	0/2433
6	F	0.35	0/1932	0.53	0/2609
6	T	0.35	0/1932	0.53	0/2609
7	G	0.35	0/1945	0.55	0/2634
7	U	0.35	0/1945	0.55	0/2634
8	H	0.32	0/1757	0.57	1/2383 (0.0%)
8	V	0.32	0/1757	0.57	0/2383
9	I	0.34	0/1611	0.58	0/2174
9	W	0.34	0/1611	0.58	0/2174
10	J	0.35	0/1589	0.58	0/2142
10	X	0.34	0/1589	0.58	0/2142
11	K	0.33	0/1681	0.58	1/2274 (0.0%)
11	Y	0.33	0/1681	0.58	1/2274 (0.0%)
12	L	0.35	0/1795	0.58	0/2420
12	Z	0.35	0/1795	0.58	0/2420
13	M	0.35	0/1855	0.61	0/2514
13	a	0.35	0/1855	0.61	0/2514
14	N	0.33	0/1541	0.57	0/2087
14	b	0.33	0/1541	0.57	0/2087
All	All	0.34	0/50278	0.57	3/67982 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
11	Y	4	LEU	CA-CB-CG	5.46	127.86	115.30
11	K	4	LEU	CA-CB-CG	5.44	127.80	115.30
8	H	196	ARG	NE-CZ-NH1	5.22	122.91	120.30

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	1	0
1	O	1915	0	1929	2	0
2	B	1904	0	1904	4	0
2	P	1904	0	1904	5	0
3	C	1881	0	1895	4	0
3	Q	1881	0	1895	3	0
4	D	1813	0	1797	3	0
4	R	1813	0	1797	5	0
5	E	1773	0	1775	1	0
5	S	1773	0	1775	3	0
6	F	1892	0	1883	3	0
6	T	1892	0	1883	3	0
7	G	1907	0	1901	2	0
7	U	1907	0	1901	2	0
8	H	1726	0	1722	2	0
8	V	1726	0	1722	3	0
9	I	1581	0	1574	5	0
9	W	1581	0	1574	5	0
10	J	1561	0	1569	3	0
10	X	1561	0	1569	3	0
11	K	1644	0	1595	5	0
11	Y	1644	0	1595	5	0
12	L	1757	0	1711	2	0
12	Z	1757	0	1711	2	0
13	M	1824	0	1832	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	a	1824	0	1832	0	0
14	N	1512	0	1481	5	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	46	0	0	0	0
17	V	46	0	0	0	0
18	H	46	0	0	0	0
18	V	46	0	0	0	0
19	H	12	0	13	0	0
19	V	12	0	13	0	0
20	A	26	0	0	0	0
20	B	16	0	0	0	0
20	C	20	0	0	0	0
20	D	4	0	0	0	0
20	E	9	0	0	0	0
20	F	16	0	0	0	0
20	G	23	0	0	0	0
20	H	28	0	0	0	0
20	I	30	0	0	0	0
20	J	19	0	0	0	0
20	K	21	0	0	0	0
20	L	22	0	0	0	0
20	M	22	0	0	0	0
20	N	19	0	0	0	0
20	O	13	0	0	0	0
20	P	25	0	0	1	0
20	Q	10	0	0	0	0
20	R	20	0	0	0	0
20	S	13	0	0	0	0
20	T	12	0	0	0	0
20	U	22	0	0	0	0
20	V	18	0	0	0	0
20	W	17	0	0	0	0
20	X	24	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
20	Y	33	0	0	0	0
20	Z	21	0	0	0	0
20	a	25	0	0	0	0
20	b	16	0	0	0	0
All	All	50142	0	49162	65	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 65 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)
3:C:160:GLN:HE21	3:C:160:GLN:HA	1.65	0.62
3:Q:160:GLN:HE21	3:Q:160:GLN:HA	1.65	0.60
5:S:12:PHE:H	6:T:19:GLN:HE22	1.55	0.55
2:B:12:PHE:H	3:C:17:GLN:HE22	1.53	0.55
8:H:45:GLY:HA3	8:H:52:THR:HG21	1.89	0.54

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%)	239 (96%)	8 (3%)	1 (0%)	34 60
1	O	248/250 (99%)	239 (96%)	8 (3%)	1 (0%)	34 60
2	B	242/258 (94%)	235 (97%)	3 (1%)	4 (2%)	9 23
2	P	242/258 (94%)	235 (97%)	3 (1%)	4 (2%)	9 23
3	C	238/254 (94%)	233 (98%)	2 (1%)	3 (1%)	12 30
3	Q	238/254 (94%)	233 (98%)	2 (1%)	3 (1%)	12 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	231/260 (89%)	228 (99%)	3 (1%)	0	100	100
4	R	231/260 (89%)	228 (99%)	3 (1%)	0	100	100
5	E	229/234 (98%)	223 (97%)	6 (3%)	0	100	100
5	S	229/234 (98%)	223 (97%)	6 (3%)	0	100	100
6	F	241/288 (84%)	237 (98%)	4 (2%)	0	100	100
6	T	241/288 (84%)	237 (98%)	4 (2%)	0	100	100
7	G	239/252 (95%)	238 (100%)	1 (0%)	0	100	100
7	U	239/252 (95%)	237 (99%)	2 (1%)	0	100	100
8	H	224/232 (97%)	219 (98%)	5 (2%)	0	100	100
8	V	224/232 (97%)	219 (98%)	5 (2%)	0	100	100
9	I	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
9	W	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
10	J	193/198 (98%)	188 (97%)	5 (3%)	0	100	100
10	X	193/198 (98%)	188 (97%)	5 (3%)	0	100	100
11	K	210/212 (99%)	205 (98%)	5 (2%)	0	100	100
11	Y	210/212 (99%)	205 (98%)	5 (2%)	0	100	100
12	L	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
12	Z	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
13	M	231/246 (94%)	224 (97%)	7 (3%)	0	100	100
13	a	231/246 (94%)	224 (97%)	7 (3%)	0	100	100
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
14	b	194/196 (99%)	190 (98%)	4 (2%)	0	100	100
All	All	6284/6614 (95%)	6140 (98%)	128 (2%)	16 (0%)	41	66

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
2	B	221	ASP
3	C	202	GLN
2	P	51	VAL
2	P	221	ASP

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	86
1	O	209/209 (100%)	206 (99%)	3 (1%)	67	86
2	B	203/216 (94%)	197 (97%)	6 (3%)	41	70
2	P	203/216 (94%)	197 (97%)	6 (3%)	41	70
3	C	212/226 (94%)	202 (95%)	10 (5%)	26	54
3	Q	212/226 (94%)	202 (95%)	10 (5%)	26	54
4	D	194/215 (90%)	185 (95%)	9 (5%)	27	54
4	R	194/215 (90%)	185 (95%)	9 (5%)	27	54
5	E	190/193 (98%)	179 (94%)	11 (6%)	20	43
5	S	190/193 (98%)	179 (94%)	11 (6%)	20	43
6	F	201/239 (84%)	191 (95%)	10 (5%)	24	51
6	T	201/239 (84%)	191 (95%)	10 (5%)	24	51
7	G	206/210 (98%)	197 (96%)	9 (4%)	28	56
7	U	206/210 (98%)	197 (96%)	9 (4%)	28	56
8	H	186/190 (98%)	183 (98%)	3 (2%)	62	85
8	V	186/190 (98%)	182 (98%)	4 (2%)	52	79
9	I	172/173 (99%)	169 (98%)	3 (2%)	60	84
9	W	172/173 (99%)	169 (98%)	3 (2%)	60	84
10	J	173/175 (99%)	169 (98%)	4 (2%)	50	78
10	X	173/175 (99%)	169 (98%)	4 (2%)	50	78
11	K	169/169 (100%)	162 (96%)	7 (4%)	30	59
11	Y	169/169 (100%)	162 (96%)	7 (4%)	30	59
12	L	185/185 (100%)	179 (97%)	6 (3%)	39	68
12	Z	185/185 (100%)	179 (97%)	6 (3%)	39	68
13	M	199/208 (96%)	192 (96%)	7 (4%)	36	65
13	a	199/208 (96%)	192 (96%)	7 (4%)	36	65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
14	N	162/162 (100%)	157 (97%)	5 (3%)	40	69
14	b	162/162 (100%)	157 (97%)	5 (3%)	40	69
All	All	5322/5540 (96%)	5135 (96%)	187 (4%)	36	65

5 of 187 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	R	176	LEU
7	U	83	ASN
4	R	236	LYS
5	S	202	ASP
7	U	236	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 125 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	48	ASN
11	Y	85	ASN
3	Q	38	ASN
10	X	55	GLN
13	a	18	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	$\#\|Z\| > 2$
18	GW2	H	302[B]	8	49,49,49	1.63	11 (22%)	60,66,66	1.39	7 (11%)
19	MES	H	303	-	12,12,12	2.21	1 (8%)	14,16,16	1.45	2 (14%)
18	GW2	V	302[B]	8	49,49,49	1.36	7 (14%)	60,66,66	1.49	5 (8%)
17	GVZ	V	301[A]	8	49,49,49	1.31	5 (10%)	63,68,68	1.53	9 (14%)
19	MES	V	303	-	12,12,12	2.19	1 (8%)	14,16,16	1.50	2 (14%)
17	GVZ	H	301[A]	8	49,49,49	1.36	4 (8%)	63,68,68	1.54	8 (12%)

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	GW2	H	302[B]	8	-	2/44/73/73	0/4/4/4
19	MES	H	303	-	-	5/6/14/14	0/1/1/1
18	GW2	V	302[B]	8	-	3/44/73/73	0/4/4/4
17	GVZ	V	301[A]	8	-	9/44/73/73	0/4/4/4
19	MES	V	303	-	-	2/6/14/14	0/1/1/1
17	GVZ	H	301[A]	8	-	9/44/73/73	0/4/4/4

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
19	H	303	MES	C8-S	-7.37	1.67	1.77
19	V	303	MES	C8-S	-7.25	1.67	1.77
18	H	302[B]	GW2	C40-C41	-6.20	1.36	1.51
17	H	301[A]	GVZ	C40-C41	-4.92	1.39	1.51
18	V	302[B]	GW2	C40-C41	-4.91	1.39	1.51

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	V	302[B]	GW2	C11-C10-C12	-6.68	101.11	109.88
17	H	301[A]	GVZ	C12-C10-C11	-6.57	101.17	110.56
17	V	301[A]	GVZ	C12-C10-C11	-6.47	101.31	110.56
18	H	302[B]	GW2	C11-C10-C12	-5.76	102.31	109.88
19	V	303	MES	O1S-S-C8	4.15	111.91	106.92

There are no chirality outliers.

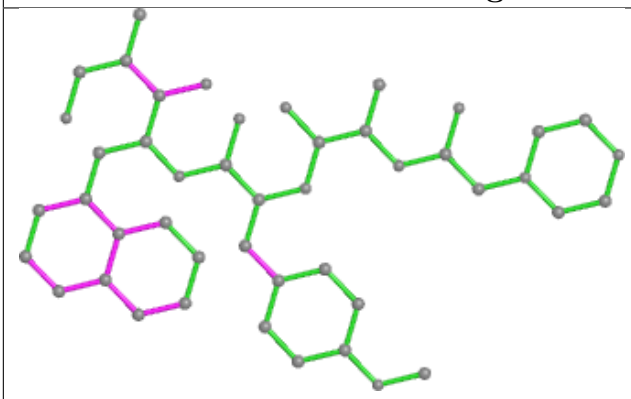
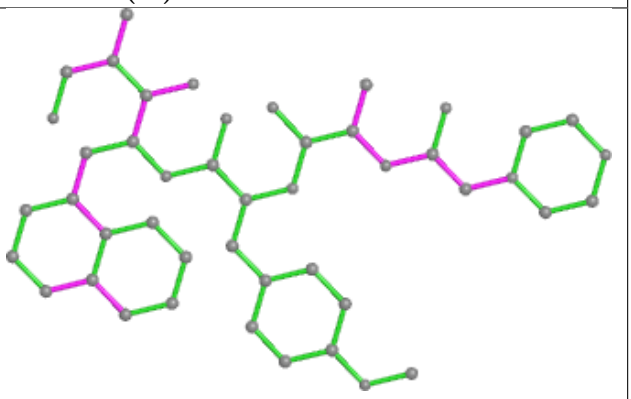
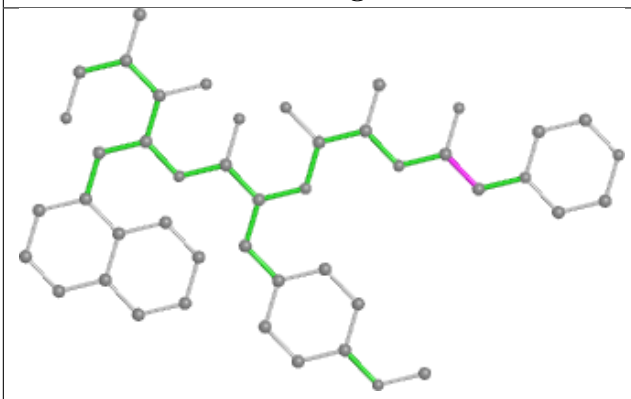
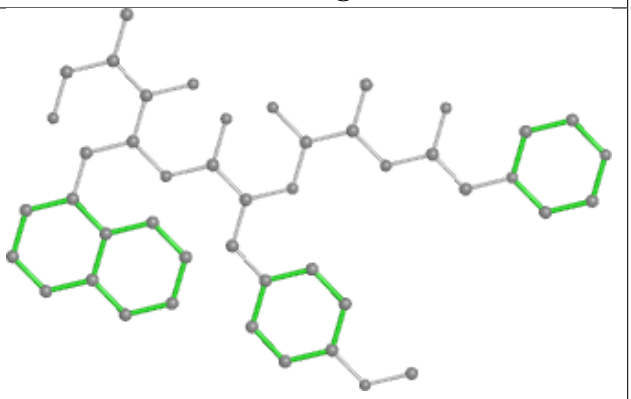
5 of 30 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	H	301[A]	GVZ	C11-C10-C9-C8
17	H	301[A]	GVZ	C11-C10-C9-O21
17	H	301[A]	GVZ	C12-C10-C9-C8
17	H	301[A]	GVZ	C12-C10-C9-O21
17	H	301[A]	GVZ	O13-C10-C9-C8

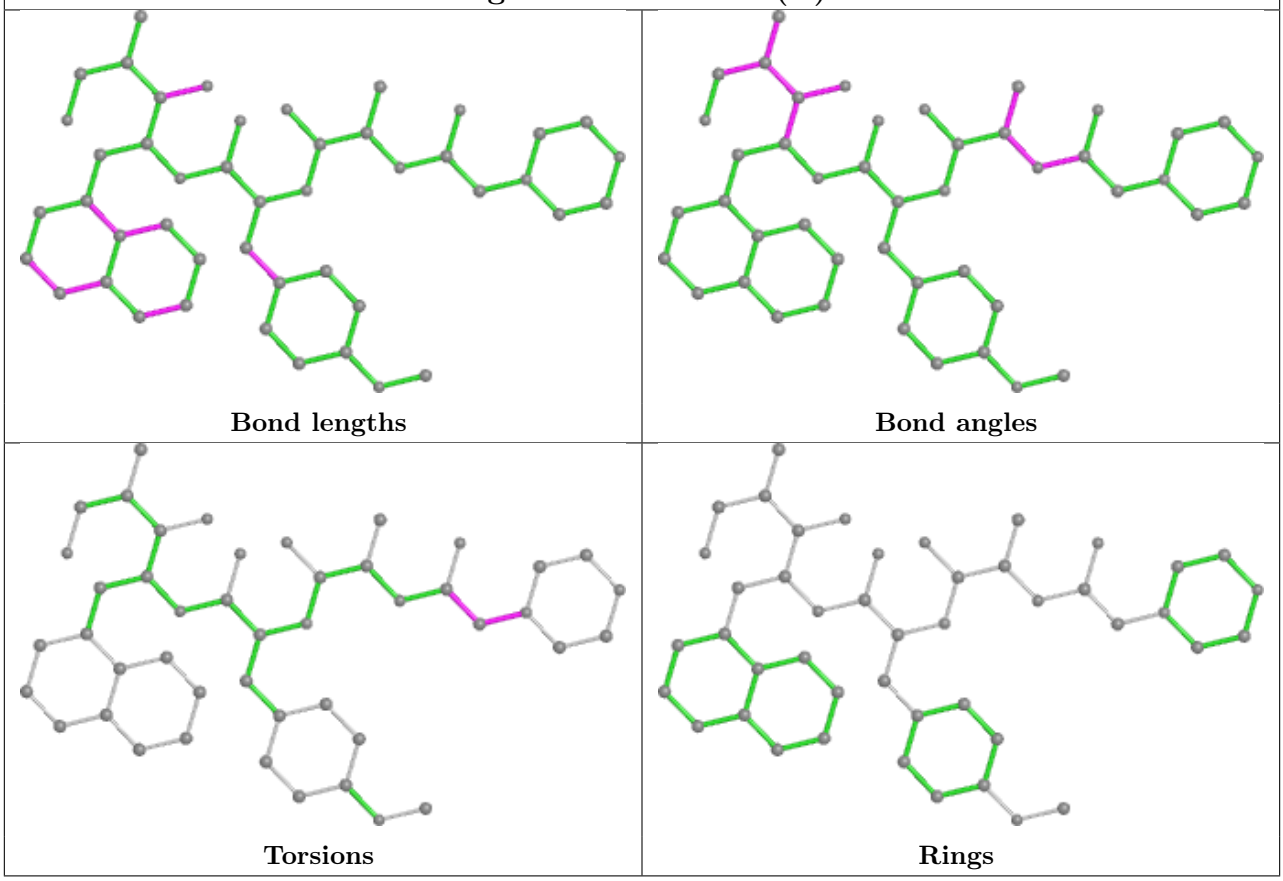
There are no ring outliers.

No monomer is involved in short contacts.

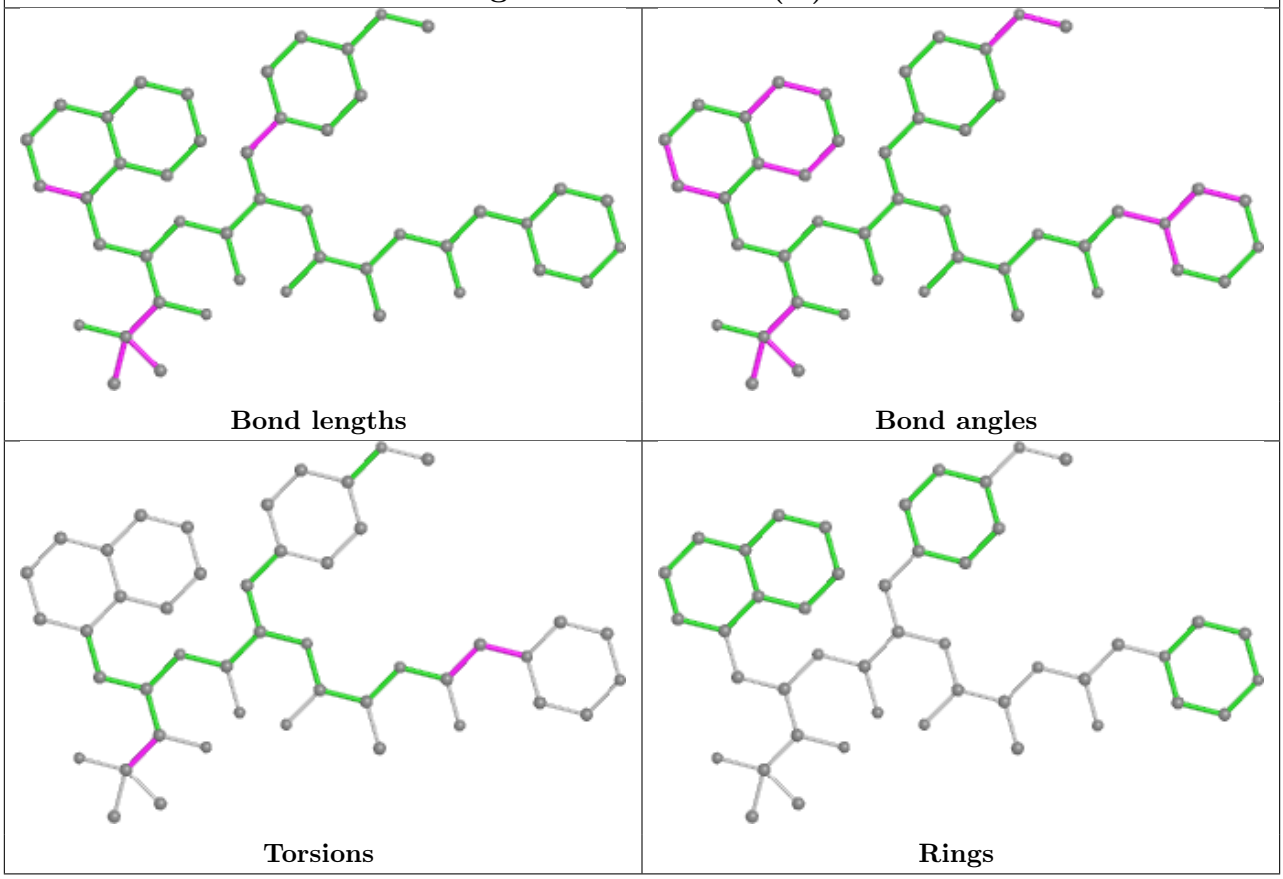
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 GW2 H 302 (B)	
 Bond lengths	 Bond angles
 Torsions	 Rings

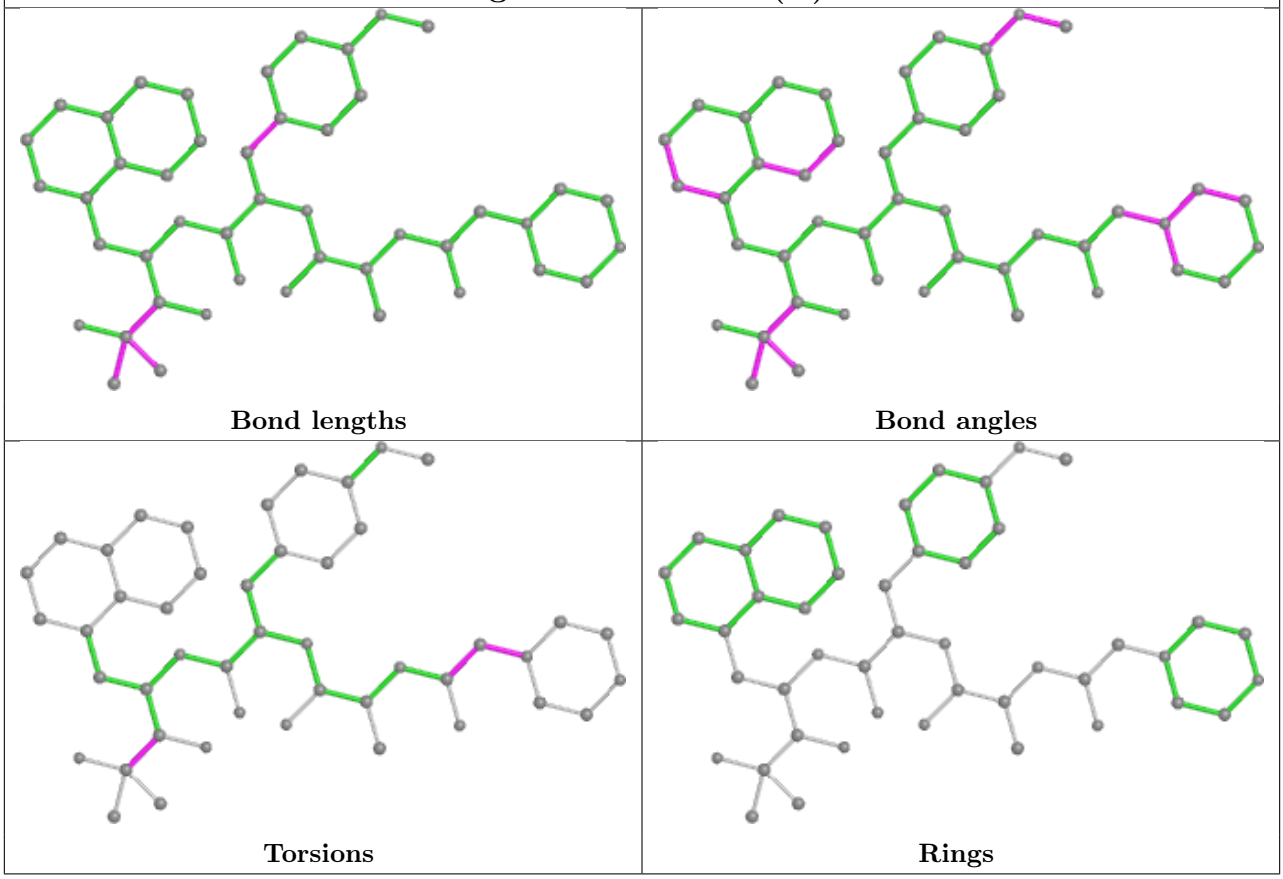
Ligand GW2 V 302 (B)



Ligand GVZ V 301 (A)



Ligand GVZ H 301 (A)



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%)	-0.39	8 (3%)	47	48	39, 52, 86, 126	0
1	O	250/250 (100%)	-0.31	6 (2%)	59	60	43, 59, 102, 131	0
2	B	244/258 (94%)	-0.27	6 (2%)	57	59	40, 57, 98, 149	0
2	P	244/258 (94%)	-0.18	11 (4%)	33	31	42, 62, 106, 147	0
3	C	240/254 (94%)	-0.11	14 (5%)	23	22	41, 63, 126, 147	0
3	Q	240/254 (94%)	0.10	20 (8%)	11	9	44, 72, 151, 177	0
4	D	235/260 (90%)	-0.38	5 (2%)	63	65	43, 63, 96, 137	0
4	R	235/260 (90%)	-0.28	6 (2%)	56	57	45, 66, 105, 136	0
5	E	231/234 (98%)	-0.18	7 (3%)	50	51	48, 67, 104, 149	0
5	S	231/234 (98%)	-0.07	8 (3%)	44	44	49, 74, 118, 161	0
6	F	243/288 (84%)	-0.34	5 (2%)	63	65	40, 60, 110, 136	0
6	T	243/288 (84%)	-0.25	10 (4%)	37	36	39, 69, 125, 156	0
7	G	241/252 (95%)	-0.43	5 (2%)	63	65	37, 55, 88, 134	0
7	U	241/252 (95%)	-0.34	5 (2%)	63	65	43, 59, 90, 138	0
8	H	226/232 (97%)	-0.45	7 (3%)	49	49	38, 52, 84, 139	0
8	V	226/232 (97%)	-0.37	5 (2%)	62	63	43, 55, 88, 152	0
9	I	204/205 (99%)	-0.66	1 (0%)	91	92	35, 48, 74, 97	0
9	W	204/205 (99%)	-0.65	3 (1%)	73	76	38, 50, 79, 103	0
10	J	195/198 (98%)	-0.49	3 (1%)	73	76	38, 52, 77, 127	0
10	X	195/198 (98%)	-0.49	3 (1%)	73	76	39, 54, 78, 142	0
11	K	212/212 (100%)	-0.54	0	100	100	36, 51, 72, 92	0
11	Y	212/212 (100%)	-0.55	1 (0%)	91	92	39, 51, 73, 94	0
12	L	222/222 (100%)	-0.51	1 (0%)	91	92	40, 55, 83, 121	0
12	Z	222/222 (100%)	-0.52	2 (0%)	84	85	39, 52, 82, 118	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%)	-0.57	3 (1%)	77	78	36, 54, 77, 94	0
13	a	233/246 (94%)	-0.54	1 (0%)	92	93	36, 53, 73, 89	0
14	N	196/196 (100%)	-0.58	2 (1%)	82	83	39, 49, 77, 103	0
14	b	196/196 (100%)	-0.52	2 (1%)	82	83	39, 50, 78, 109	0
All	All	6344/6614 (95%)	-0.38	150 (2%)	59	60	35, 57, 101, 177	0

The worst 5 of 150 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	49	THR	9.6
3	Q	50	LEU	7.3
3	Q	206	LYS	6.3
3	C	206	LYS	6.2
2	B	221	ASP	6.1

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
19	MES	V	303	12/12	0.89	0.29	82,87,96,98	0
18	GW2	H	302[B]	46/46	0.91	0.23	49,54,59,63	46
17	GVZ	H	301[A]	46/46	0.91	0.24	50,56,61,64	46
15	MG	Z	301	1/1	0.92	0.20	57,57,57,57	0
18	GW2	V	302[B]	46/46	0.92	0.24	49,55,59,62	46
17	GVZ	V	301[A]	46/46	0.92	0.25	50,57,61,64	46
19	MES	H	303	12/12	0.94	0.22	81,87,93,97	0

Continued on next page...

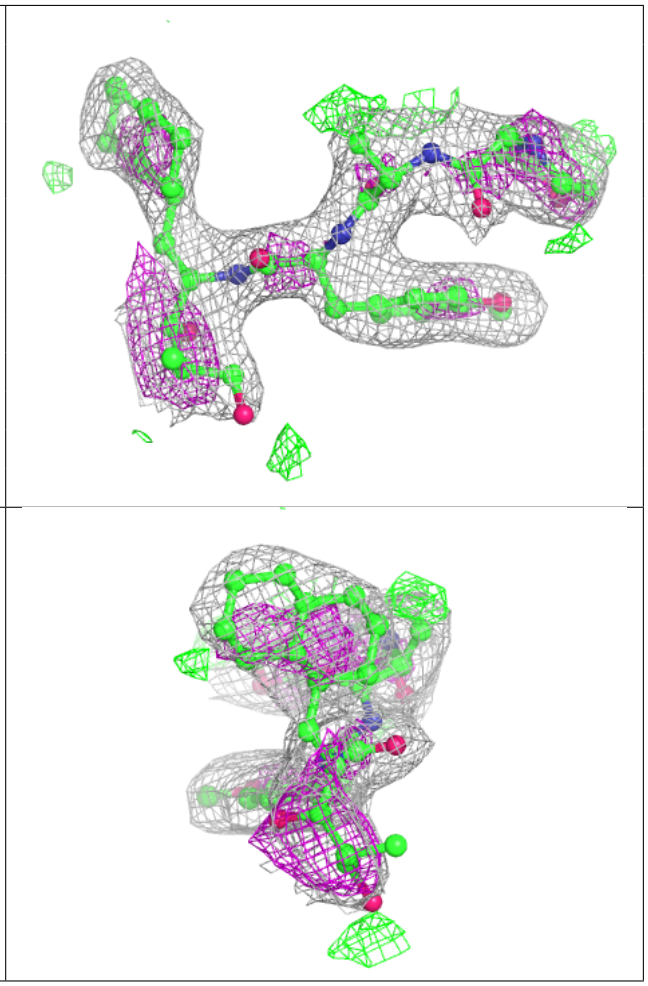
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
15	MG	I	302	1/1	0.95	0.07	57,57,57,57	0
15	MG	I	301	1/1	0.95	0.31	63,63,63,63	0
15	MG	W	301	1/1	0.96	0.26	60,60,60,60	0
15	MG	L	301	1/1	0.96	0.13	76,76,76,76	0
15	MG	N	201	1/1	0.98	0.10	52,52,52,52	0
15	MG	K	301	1/1	0.98	0.09	66,66,66,66	0
15	MG	G	301	1/1	0.98	0.06	49,49,49,49	0
16	CL	G	302	1/1	0.99	0.07	42,42,42,42	0
16	CL	U	301	1/1	0.99	0.18	44,44,44,44	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.

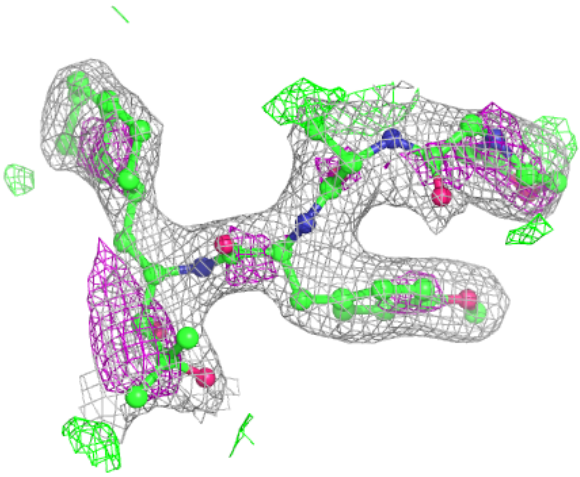
Electron density around GW2 H 302 (B):

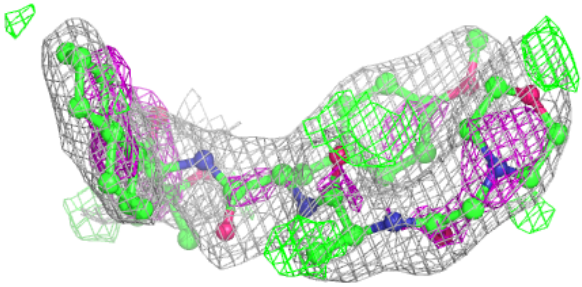
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

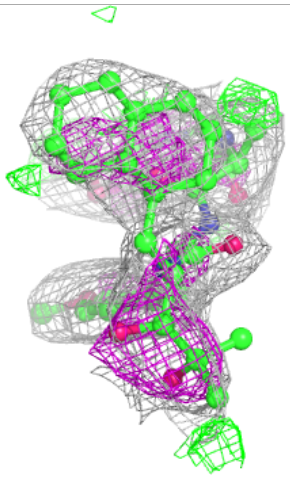


Electron density around GVZ H 301 (A):

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

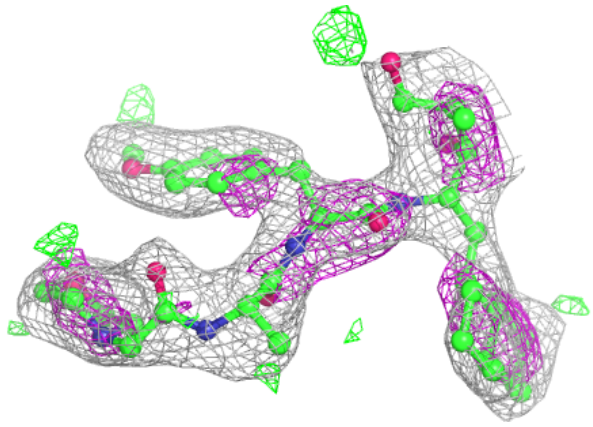


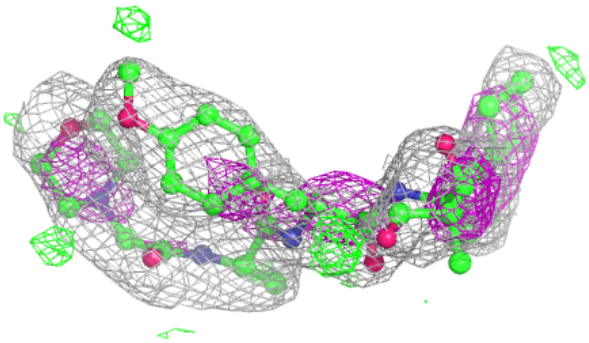


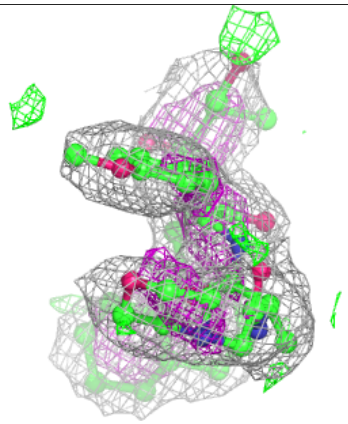


Electron density around GW2 V 302 (B):

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

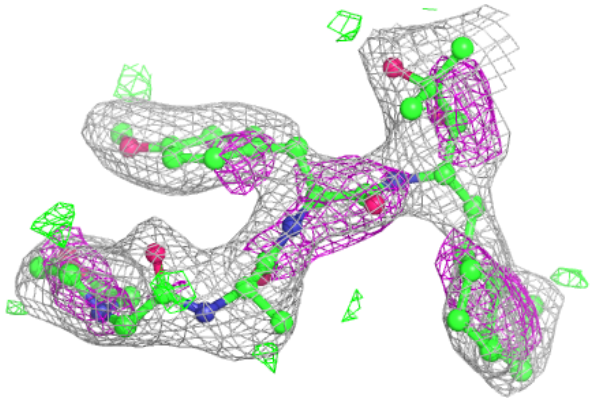


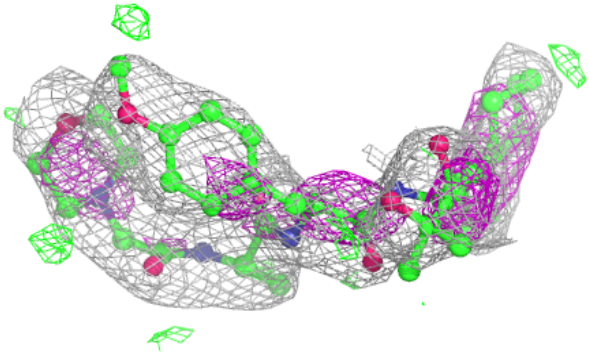


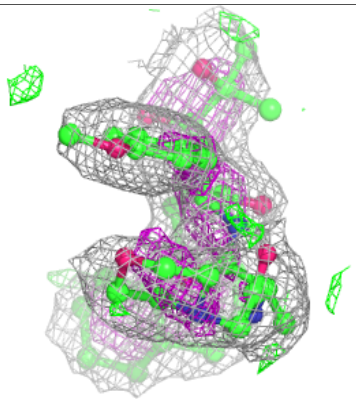


Electron density around GVZ V 301 (A):

$2mF_o$-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.