wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2020 – 07:38 PM BST

PDB ID	:	5L66
Title	:	Yeast 20S proteasome with mouse beta5i (1-138) and mouse beta6 (97-111; 118-133) in complex with bortezomib
Authors	:	Groll, M.; Huber, E.M.
Deposited on	:	2016-05-28
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 following versions of software and data (see references (1)) were used in the production of this report:

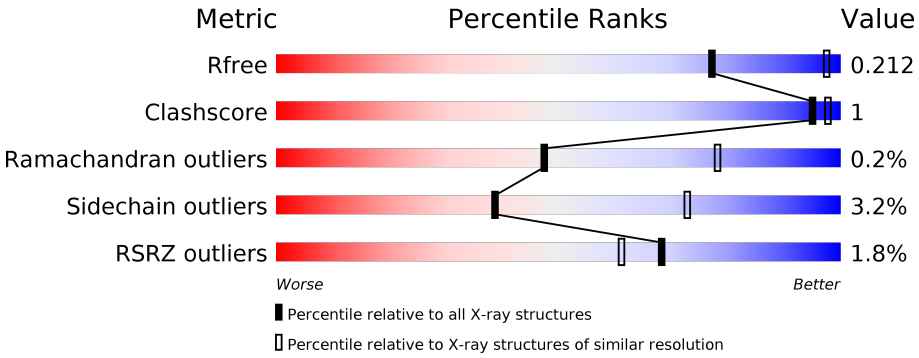
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.80 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.



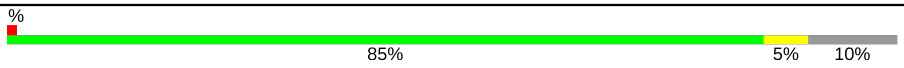
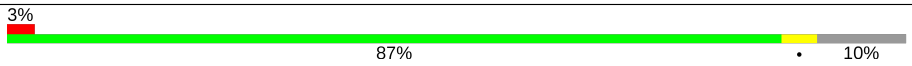
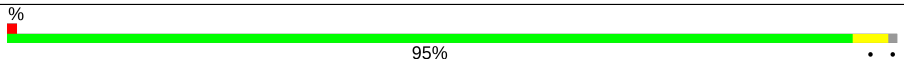
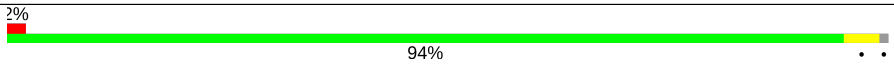
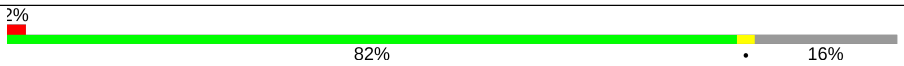
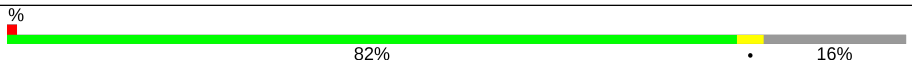
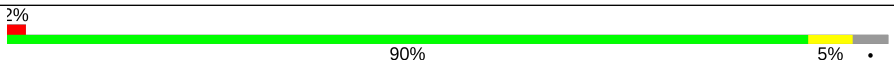
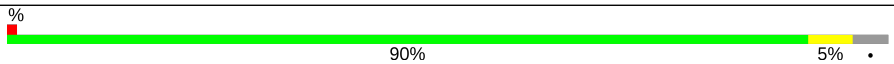
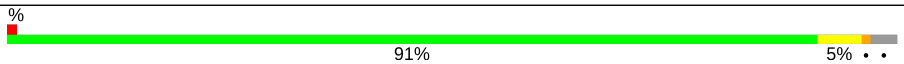
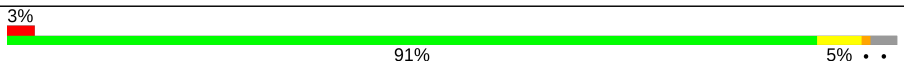
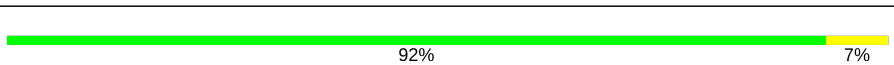
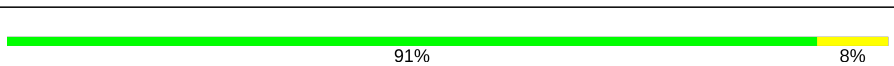
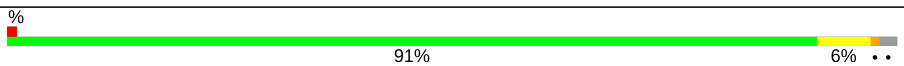
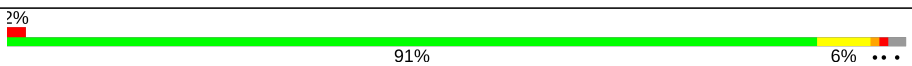
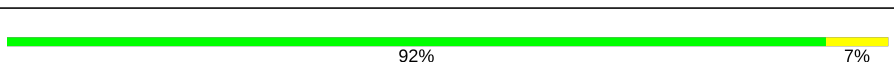
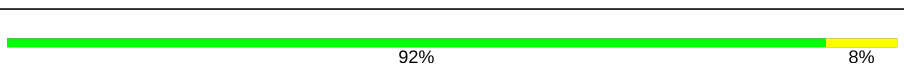
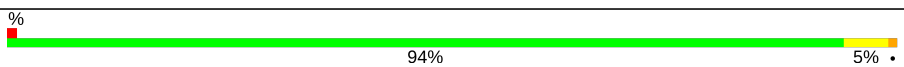
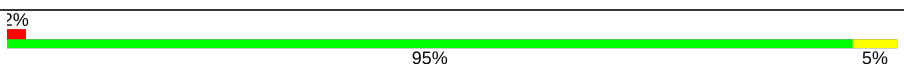
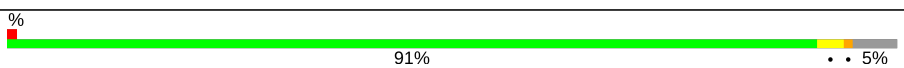
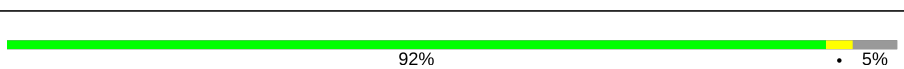
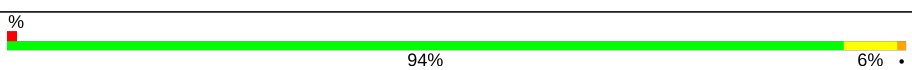
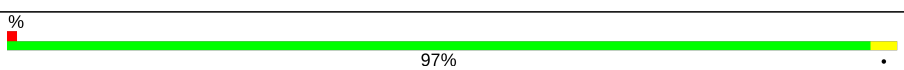
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 on 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% 98%
1	O	250	2% 97%
2	B	258	3% 89% 5% 5%
2	P	258	3% 89% 5% 5%
3	C	254	6% 87% 6% 6%
3	Q	254	6% 87% 6% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	260	
4	R	260	
5	E	234	
5	S	234	
6	F	288	
6	T	288	
7	G	252	
7	U	252	
8	H	232	
8	V	232	
9	I	205	
9	W	205	
10	J	198	
10	X	198	
11	K	211	
11	Y	211	
12	L	222	
12	Z	222	
13	M	246	
13	a	246	
14	N	196	
14	b	196	

2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 49840 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
1	A	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			
1	O	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			

- Molecule 2 is a protein called Proteasome subunit alpha type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			
2	P	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			

- Molecule 3 is a protein called Proteasome subunit alpha type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	240	Total	C	N	O	S	0	0	0
			1881	1176	329	372	4			
3	Q	240	Total	C	N	O	S	0	0	0
			1881	1176	329	372	4			

- Molecule 4 is a protein called Proteasome subunit alpha type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	235	Total	C	N	O	S	0	0	0
			1813	1136	304	366	7			
4	R	235	Total	C	N	O	S	0	0	0
			1813	1136	304	366	7			

- Molecule 5 is a protein called Proteasome subunit alpha type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	231	Total	C	N	O	S	0	0	0
			1773	1114	307	348	4			
5	S	231	Total	C	N	O	S	0	0	0
			1773	1114	307	348	4			

- Molecule 6 is a protein called Probable proteasome subunit alpha type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	243	Total	C	N	O	S	0	0	0
			1892	1203	329	356	4			
6	T	243	Total	C	N	O	S	0	0	0
			1892	1203	329	356	4			

- Molecule 7 is a protein called Proteasome subunit alpha type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	241	Total	C	N	O	S	0	0	0
			1907	1214	320	365	8			
7	U	241	Total	C	N	O	S	0	0	0
			1907	1214	320	365	8			

- Molecule 8 is a protein called Proteasome subunit beta type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	226	Total	C	N	O	S	0	0	0
			1719	1082	298	332	7			
8	V	226	Total	C	N	O	S	0	0	0
			1719	1082	298	332	7			

- Molecule 9 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			
9	W	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			

- Molecule 10 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	195	Total	C	N	O	S	0	0	0
			1561	992	264	299	6			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	195	Total	C	N	O	S	0	0	0
			1561	992	264	299	6			

- Molecule 11 is a protein called Proteasome subunit beta type-8,Proteasome subunit beta type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	211	Total	C	N	O	S	0	0	0
			1645	1039	281	313	12			
11	Y	211	Total	C	N	O	S	0	0	0
			1645	1039	281	313	12			

- Molecule 12 is a protein called Proteasome subunit beta type-6,Proteasome subunit beta type-1,Proteasome subunit beta type-6,Proteasome subunit beta type-1,Proteasome subunit beta type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	222	Total	C	N	O	S	0	0	0
			1764	1119	305	336	4			
12	Z	222	Total	C	N	O	S	0	1	0
			1773	1127	305	337	4			

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

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

- Molecule 14 is a protein called Proteasome subunit beta type-1.

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

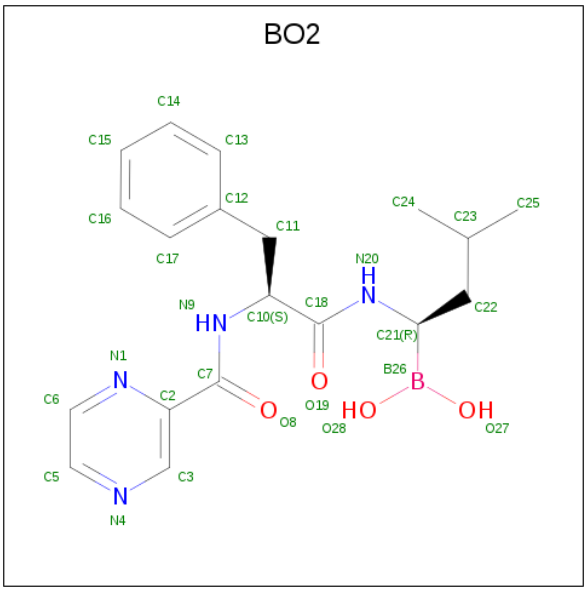
- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	G	1	Total Mg 1 1	0	0
15	J	1	Total Mg 1 1	0	0
15	K	1	Total Mg 1 1	0	0
15	I	1	Total Mg 1 1	0	0
15	Z	1	Total Mg 1 1	0	0
15	N	1	Total Mg 1 1	0	0
15	L	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	b	1	Total Cl 1 1	0	0
16	N	1	Total Cl 1 1	0	0
16	U	1	Total Cl 1 1	0	0

- Molecule 17 is N-[(1R)-1-(DIHYDROXYBORYL)-3-METHYLBUTYL]-N-(PYRAZIN-2-YLCARBONYL)-L-PHENYLALANINAMIDE (three-letter code: BO2) (formula: $C_{19}H_{25}BN_4O_4$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
17	H	1	Total	B	C	N	O	0	0
			28	1	19	4	4		
17	K	1	Total	B	C	N	O	0	0
			28	1	19	4	4		
17	N	1	Total	B	C	N	O	0	0
			28	1	19	4	4		
17	V	1	Total	B	C	N	O	0	0
			28	1	19	4	4		
17	Y	1	Total	B	C	N	O	0	0
			28	1	19	4	4		
17	b	1	Total	B	C	N	O	0	0
			28	1	19	4	4		

- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	8	Total	O	0	0
			8	8		
18	B	11	Total	O	0	0
			11	11		
18	C	4	Total	O	0	0
			4	4		
18	D	4	Total	O	0	0
			4	4		
18	E	6	Total	O	0	0
			6	6		
18	F	11	Total	O	0	0
			11	11		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	G	11	Total O 11 11	0	0
18	H	19	Total O 19 19	0	0
18	I	8	Total O 8 8	0	0
18	J	9	Total O 9 9	0	0
18	K	11	Total O 11 11	0	0
18	L	13	Total O 13 13	0	0
18	M	16	Total O 16 16	0	0
18	N	10	Total O 10 10	0	0
18	O	8	Total O 8 8	0	0
18	P	7	Total O 7 7	0	0
18	Q	7	Total O 7 7	0	0
18	R	9	Total O 9 9	0	0
18	S	6	Total O 6 6	0	0
18	T	8	Total O 8 8	0	0
18	U	12	Total O 12 12	0	0
18	V	15	Total O 15 15	0	0
18	W	8	Total O 8 8	0	0
18	X	8	Total O 8 8	0	0
18	Y	6	Total O 6 6	0	0
18	Z	10	Total O 10 10	0	0
18	a	16	Total O 16 16	0	0

Continued on next page...

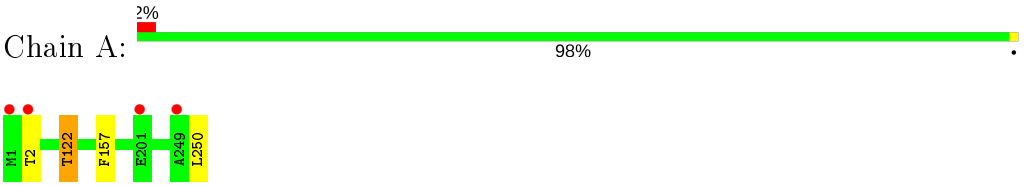
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	b	9	Total O 9 9	0	0

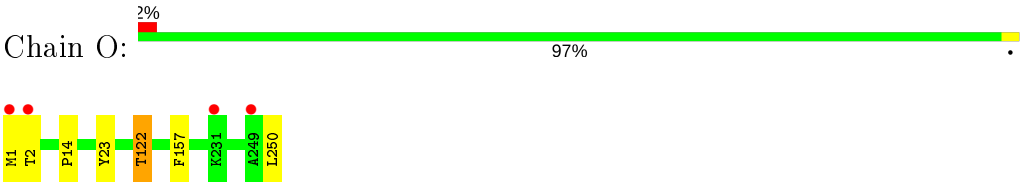
3 Residue-property plots

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.

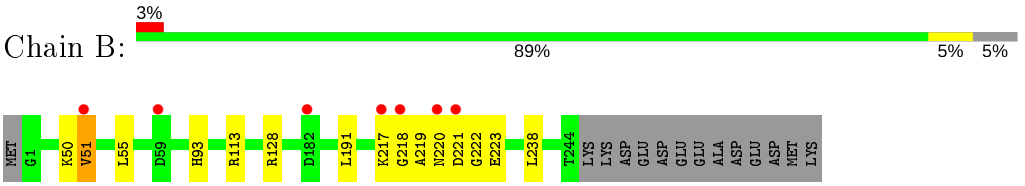
- Molecule 1: Proteasome subunit alpha type-2



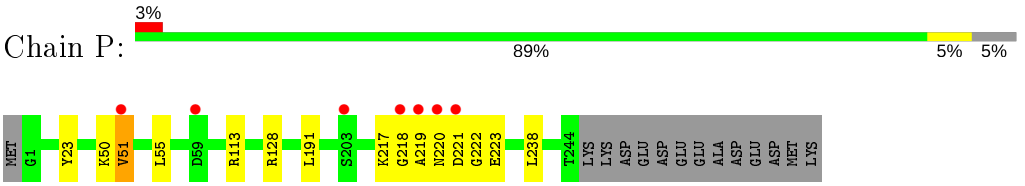
- Molecule 1: Proteasome subunit alpha type-2



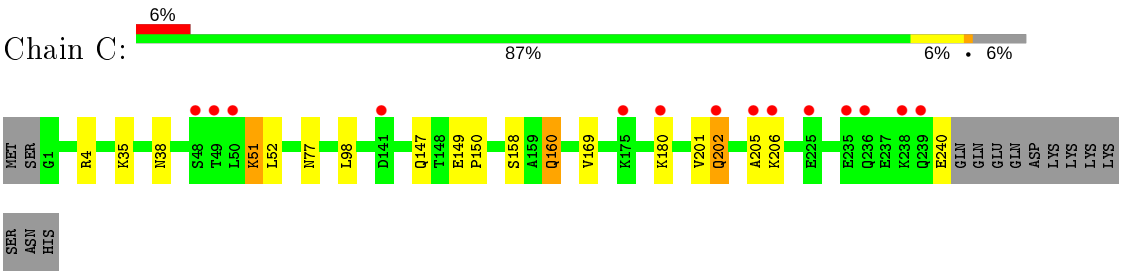
- Molecule 2: Proteasome subunit alpha type-3



- Molecule 2: Proteasome subunit alpha type-3

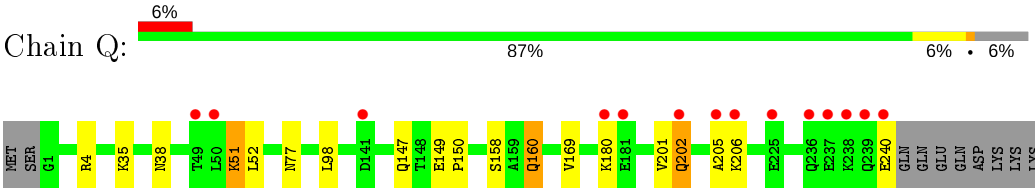


- Molecule 3: Proteasome subunit alpha type-4



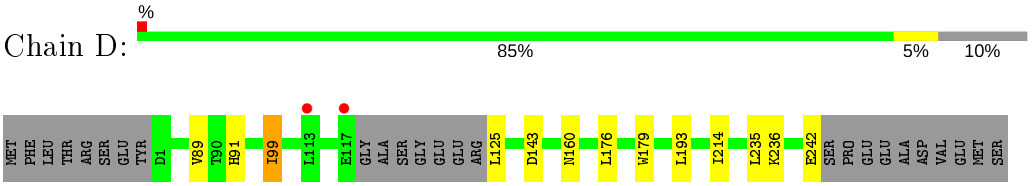
• Molecule 3: Proteasome subunit alpha type-4

Chain Q:



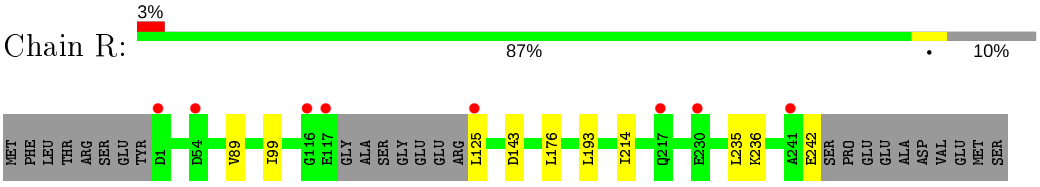
• Molecule 4: Proteasome subunit alpha type-5

Chain D:



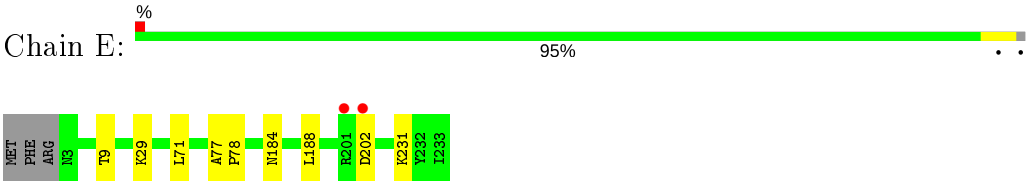
• Molecule 4: Proteasome subunit alpha type-5

Chain R:



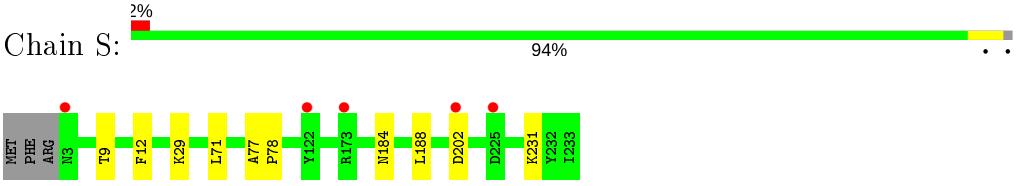
• Molecule 5: Proteasome subunit alpha type-6

Chain E:



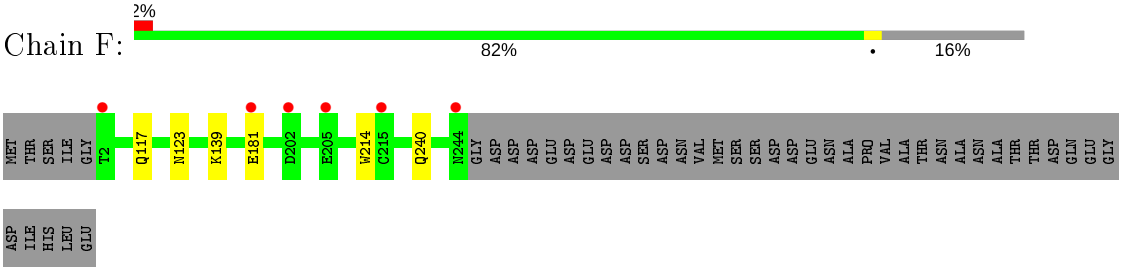
• Molecule 5: Proteasome subunit alpha type-6

Chain S:



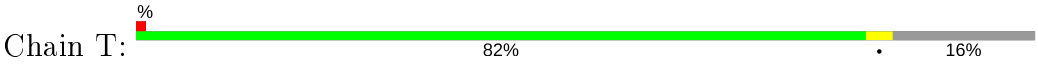
• Molecule 6: Probable proteasome subunit alpha type-7

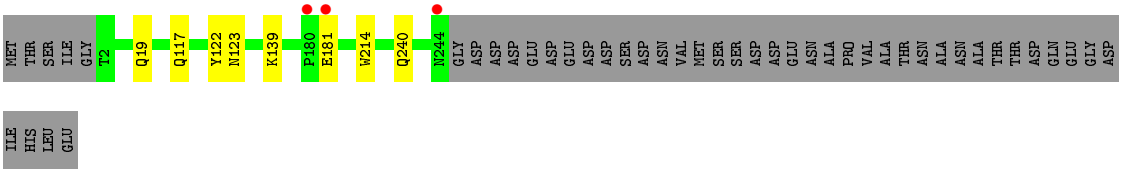
Chain F:



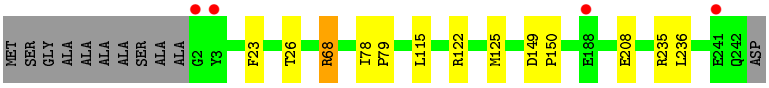
• Molecule 6: Probable proteasome subunit alpha type-7

Chain T:

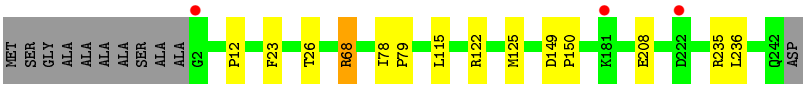




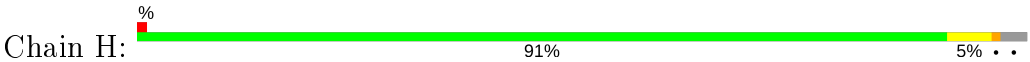
• Molecule 7: Proteasome subunit alpha type-1



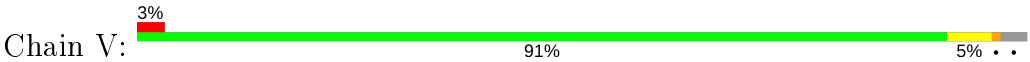
• Molecule 7: Proteasome subunit alpha type-1



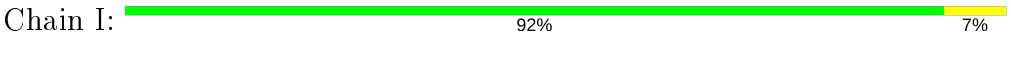
• Molecule 8: Proteasome subunit beta type-2

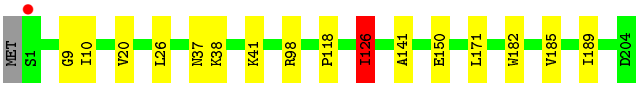


• Molecule 8: Proteasome subunit beta type-2

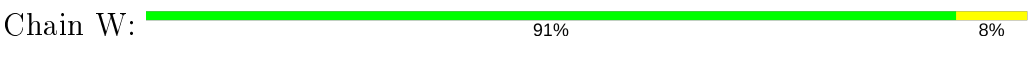


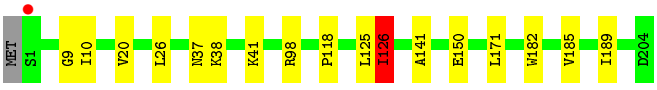
• Molecule 9: Proteasome subunit beta type-3





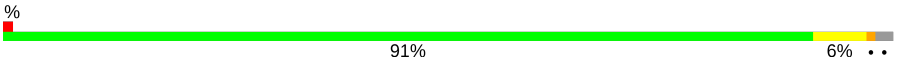
• Molecule 9: Proteasome subunit beta type-3

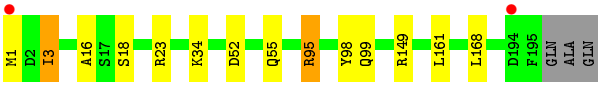




• Molecule 10: Proteasome subunit beta type-4

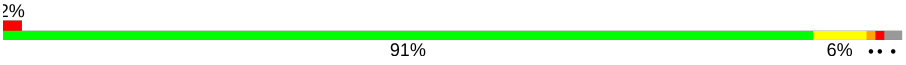
Chain J:

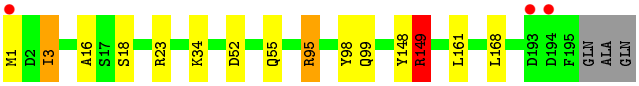




• Molecule 10: Proteasome subunit beta type-4

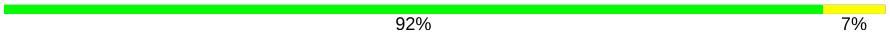
Chain X:





• Molecule 11: Proteasome subunit beta type-8,Proteasome subunit beta type-5

Chain K:

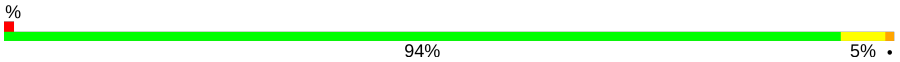


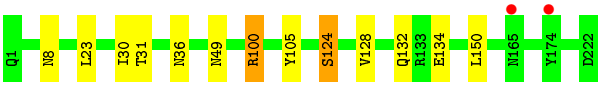
• Molecule 11: Proteasome subunit beta type-8,Proteasome subunit beta type-5

Chain Y:

• Molecule 12: Proteasome subunit beta type-6,Proteasome subunit beta type-1,Proteasome subunit beta type-6,Proteasome subunit beta type-1,Proteasome subunit beta type-6

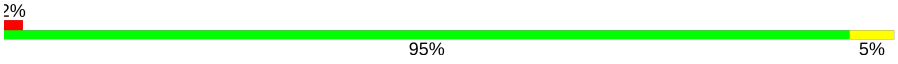
Chain L:





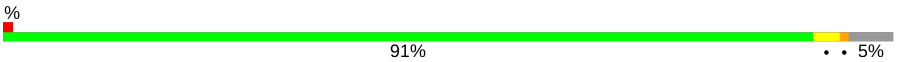
• Molecule 12: Proteasome subunit beta type-6,Proteasome subunit beta type-1,Proteasome subunit beta type-6,Proteasome subunit beta type-1,Proteasome subunit beta type-6

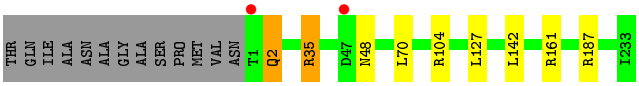
Chain Z:



• Molecule 13: Proteasome subunit beta type-7

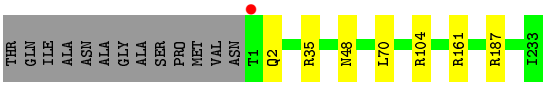
Chain M:





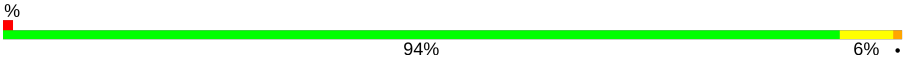
• Molecule 13: Proteasome subunit beta type-7

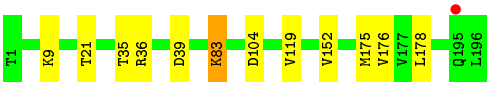
Chain a:



• Molecule 14: Proteasome subunit beta type-1

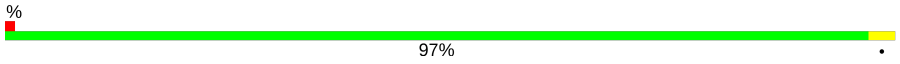
Chain N:

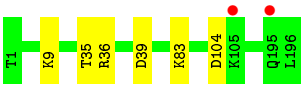




• Molecule 14: Proteasome subunit beta type-1

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	135.31Å 301.13Å 145.75Å 90.00° 112.92° 90.00°	Depositor
Resolution (Å)	15.00 – 2.80 15.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.3 (15.00-2.80) 98.3 (15.00-2.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.19 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.182 , 0.210 0.187 , 0.212	Depositor DCC
R_{free} test set	12828 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	60.7	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 38.9	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	49840	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.33% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BO2, 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	0.27	0/1952	0.46	0/2642
1	O	0.27	0/1952	0.46	0/2642
2	B	0.27	0/1934	0.49	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	0.27	0/1837	0.47	0/2475
4	R	0.27	0/1837	0.46	0/2475
5	E	0.27	0/1800	0.46	0/2433
5	S	0.27	0/1800	0.46	0/2433
6	F	0.27	0/1932	0.44	0/2609
6	T	0.27	0/1932	0.44	0/2609
7	G	0.27	0/1945	0.72	4/2634 (0.2%)
7	U	0.28	0/1945	0.74	4/2634 (0.2%)
8	H	0.25	0/1750	0.49	1/2373 (0.0%)
8	V	0.25	0/1750	0.49	1/2373 (0.0%)
9	I	0.27	0/1611	0.62	3/2174 (0.1%)
9	W	0.27	0/1611	0.64	2/2174 (0.1%)
10	J	0.27	0/1589	0.96	7/2142 (0.3%)
10	X	0.27	0/1589	1.01	7/2142 (0.3%)
11	K	0.26	0/1681	0.77	3/2268 (0.1%)
11	Y	0.27	0/1681	0.77	3/2268 (0.1%)
12	L	0.29	0/1802	0.75	3/2430 (0.1%)
12	Z	0.28	0/1815	0.79	3/2448 (0.1%)
13	M	0.46	2/1855 (0.1%)	0.76	5/2514 (0.2%)
13	a	0.46	2/1855 (0.1%)	0.77	5/2514 (0.2%)
14	N	0.25	0/1541	0.46	0/2087
14	b	0.25	0/1541	0.46	0/2087
All	All	0.29	4/50291 (0.0%)	0.62	51/67988 (0.1%)

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
7	G	0	1
7	U	0	1
10	J	0	1
10	X	0	2
All	All	0	5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	M	35	ARG	CZ-NH2	-12.34	1.17	1.33
13	a	35	ARG	CZ-NH2	-11.41	1.18	1.33
13	a	35	ARG	CZ-NH1	-11.29	1.18	1.33
13	M	35	ARG	CZ-NH1	-10.38	1.19	1.33

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	U	68	ARG	NE-CZ-NH2	-22.08	109.26	120.30
10	X	149	ARG	NE-CZ-NH2	-21.70	109.45	120.30
12	Z	100	ARG	NE-CZ-NH2	-21.19	109.70	120.30
10	X	95	ARG	NE-CZ-NH1	-20.77	109.92	120.30
12	Z	100	ARG	NE-CZ-NH1	20.10	130.35	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	G	68	ARG	Sidechain
10	J	95	ARG	Sidechain
7	U	68	ARG	Sidechain
10	X	149	ARG	Sidechain
10	X	95	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	1915	0	1929	1	0
1	O	1915	0	1929	4	0
2	B	1904	0	1904	5	0
2	P	1904	0	1904	6	0
3	C	1881	0	1895	6	0
3	Q	1881	0	1895	6	0
4	D	1813	0	1797	3	0
4	R	1813	0	1797	1	0
5	E	1773	0	1775	1	0
5	S	1773	0	1775	2	0
6	F	1892	0	1883	0	0
6	T	1892	0	1883	2	0
7	G	1907	0	1901	3	0
7	U	1907	0	1901	4	0
8	H	1719	0	1718	9	0
8	V	1719	0	1718	8	0
9	I	1581	0	1574	10	0
9	W	1581	0	1574	13	0
10	J	1561	0	1569	8	0
10	X	1561	0	1569	8	0
11	K	1645	0	1591	6	0
11	Y	1645	0	1591	6	0
12	L	1764	0	1716	4	0
12	Z	1773	0	1725	4	0
13	M	1824	0	1832	2	0
13	a	1824	0	1832	0	0
14	N	1512	0	1480	4	0
14	b	1512	0	1480	0	0
15	G	1	0	0	0	0
15	I	1	0	0	0	0
15	J	1	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	Z	1	0	0	0	0
16	G	1	0	0	0	0
16	N	1	0	0	0	0
16	U	1	0	0	0	0
16	b	1	0	0	0	0
17	H	28	0	25	0	0
17	K	28	0	25	0	0
17	N	28	0	25	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	V	28	0	25	0	0
17	Y	28	0	25	0	0
17	b	28	0	25	0	0
18	A	8	0	0	0	0
18	B	11	0	0	1	0
18	C	4	0	0	0	0
18	D	4	0	0	0	0
18	E	6	0	0	0	0
18	F	11	0	0	0	0
18	G	11	0	0	0	0
18	H	19	0	0	0	0
18	I	8	0	0	0	0
18	J	9	0	0	0	0
18	K	11	0	0	0	0
18	L	13	0	0	0	0
18	M	16	0	0	1	0
18	N	10	0	0	0	0
18	O	8	0	0	0	0
18	P	7	0	0	0	0
18	Q	7	0	0	0	0
18	R	9	0	0	0	0
18	S	6	0	0	0	0
18	T	8	0	0	0	0
18	U	12	0	0	0	0
18	V	15	0	0	0	0
18	W	8	0	0	0	0
18	X	8	0	0	0	0
18	Y	6	0	0	0	0
18	Z	10	0	0	0	0
18	a	16	0	0	0	0
18	b	9	0	0	0	0
All	All	49840	0	49287	114	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 114 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)
9:I:98:ARG:HD2	9:I:126:ILE:HD12	1.62	0.80
9:W:98:ARG:HD2	9:W:126:ILE:HD12	1.62	0.80

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:X:55:GLN:NE2	10:X:98:TYR:OH	2.19	0.75
10:J:55:GLN:NE2	10:J:98:TYR:OH	2.20	0.74
8:V:113:ILE:HG13	8:V:119:THR:HG22	1.70	0.73

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	66
1	O	248/250 (99%)	239 (96%)	8 (3%)	1 (0%)	34	66
2	B	242/258 (94%)	235 (97%)	3 (1%)	4 (2%)	9	29
2	P	242/258 (94%)	235 (97%)	3 (1%)	4 (2%)	9	29
3	C	238/254 (94%)	233 (98%)	3 (1%)	2 (1%)	19	49
3	Q	238/254 (94%)	233 (98%)	3 (1%)	2 (1%)	19	49
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%)	221 (96%)	8 (4%)	0	100	100
5	S	229/234 (98%)	223 (97%)	6 (3%)	0	100	100
6	F	241/288 (84%)	238 (99%)	3 (1%)	0	100	100
6	T	241/288 (84%)	238 (99%)	3 (1%)	0	100	100
7	G	239/252 (95%)	237 (99%)	2 (1%)	0	100	100
7	U	239/252 (95%)	237 (99%)	2 (1%)	0	100	100
8	H	224/232 (97%)	220 (98%)	4 (2%)	0	100	100
8	V	224/232 (97%)	220 (98%)	4 (2%)	0	100	100
9	I	202/205 (98%)	195 (96%)	7 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	W	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	J	193/198 (98%)	190 (98%)	3 (2%)	0	100	100
10	X	193/198 (98%)	190 (98%)	3 (2%)	0	100	100
11	K	209/211 (99%)	203 (97%)	6 (3%)	0	100	100
11	Y	209/211 (99%)	203 (97%)	6 (3%)	0	100	100
12	L	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
12	Z	221/222 (100%)	216 (98%)	5 (2%)	0	100	100
13	M	231/246 (94%)	222 (96%)	9 (4%)	0	100	100
13	a	231/246 (94%)	223 (96%)	8 (4%)	0	100	100
14	N	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
14	b	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
All	All	6283/6612 (95%)	6130 (98%)	139 (2%)	14 (0%)	47	78

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
3	C	202	GLN
2	P	51	VAL
3	Q	202	GLN
1	A	2	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	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%)	199 (98%)	4 (2%)	55	84
2	P	203/216 (94%)	199 (98%)	4 (2%)	55	84
3	C	212/226 (94%)	201 (95%)	11 (5%)	23	55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	Q	212/226 (94%)	201 (95%)	11 (5%)	23	55
4	D	194/215 (90%)	185 (95%)	9 (5%)	27	60
4	R	194/215 (90%)	185 (95%)	9 (5%)	27	60
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%)	195 (97%)	6 (3%)	41	75
6	T	201/239 (84%)	195 (97%)	6 (3%)	41	75
7	G	206/210 (98%)	200 (97%)	6 (3%)	42	76
7	U	206/210 (98%)	200 (97%)	6 (3%)	42	76
8	H	185/190 (97%)	179 (97%)	6 (3%)	39	73
8	V	185/190 (97%)	179 (97%)	6 (3%)	39	73
9	I	172/173 (99%)	168 (98%)	4 (2%)	50	82
9	W	172/173 (99%)	168 (98%)	4 (2%)	50	82
10	J	173/175 (99%)	170 (98%)	3 (2%)	60	87
10	X	173/175 (99%)	170 (98%)	3 (2%)	60	87
11	K	172/172 (100%)	166 (96%)	6 (4%)	36	70
11	Y	172/172 (100%)	165 (96%)	7 (4%)	30	64
12	L	186/186 (100%)	180 (97%)	6 (3%)	39	73
12	Z	187/186 (100%)	181 (97%)	6 (3%)	39	73
13	M	199/208 (96%)	192 (96%)	7 (4%)	36	70
13	a	199/208 (96%)	193 (97%)	6 (3%)	41	75
14	N	162/162 (100%)	156 (96%)	6 (4%)	34	68
14	b	162/162 (100%)	156 (96%)	6 (4%)	34	68
All	All	5329/5548 (96%)	5161 (97%)	168 (3%)	39	73

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

Mol	Chain	Res	Type
13	M	104	ARG
3	Q	77	ASN
12	Z	150	LEU
14	N	9	LYS
1	O	157	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 97 such sidechains are listed below:

Mol	Chain	Res	Type
12	L	79	HIS
2	P	123	GLN
12	Z	49	ASN
13	M	102	GLN
14	N	161	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 17 ligands modelled in this entry, 11 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	BO2	N	201	14	25,29,29	1.58	5 (20%)	32,38,38	1.29	4 (12%)
17	BO2	Y	301	11	25,29,29	1.59	5 (20%)	32,38,38	1.26	3 (9%)
17	BO2	b	201	14	25,29,29	1.58	5 (20%)	32,38,38	1.30	4 (12%)
17	BO2	K	301	11	25,29,29	1.60	5 (20%)	32,38,38	1.26	3 (9%)
17	BO2	H	301	8	25,29,29	1.62	5 (20%)	32,38,38	1.32	4 (12%)
17	BO2	V	301	8	25,29,29	1.61	5 (20%)	32,38,38	1.34	4 (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
17	BO2	N	201	14	-	4/22/28/28	0/2/2/2
17	BO2	Y	301	11	-	0/22/28/28	0/2/2/2
17	BO2	b	201	14	-	4/22/28/28	0/2/2/2
17	BO2	K	301	11	-	0/22/28/28	0/2/2/2
17	BO2	H	301	8	-	4/22/28/28	0/2/2/2
17	BO2	V	301	8	-	4/22/28/28	0/2/2/2

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	H	301	BO2	C2-C7	-4.42	1.39	1.50
17	V	301	BO2	C2-C7	-4.40	1.39	1.50
17	Y	301	BO2	C2-C7	-4.37	1.40	1.50
17	K	301	BO2	C2-C7	-4.31	1.40	1.50
17	N	201	BO2	C2-C7	-4.31	1.40	1.50

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
17	V	301	BO2	C21-C22-C23	-3.96	110.42	115.39
17	H	301	BO2	C21-C22-C23	-3.80	110.61	115.39
17	N	201	BO2	C21-C22-C23	-3.78	110.65	115.39
17	b	201	BO2	C21-C22-C23	-3.78	110.65	115.39
17	K	301	BO2	C6-N1-C2	3.56	121.54	116.93

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	H	301	BO2	C3-C2-C7-O8
17	H	301	BO2	C3-C2-C7-N9
17	V	301	BO2	C3-C2-C7-O8
17	V	301	BO2	C3-C2-C7-N9
17	H	301	BO2	N1-C2-C7-O8

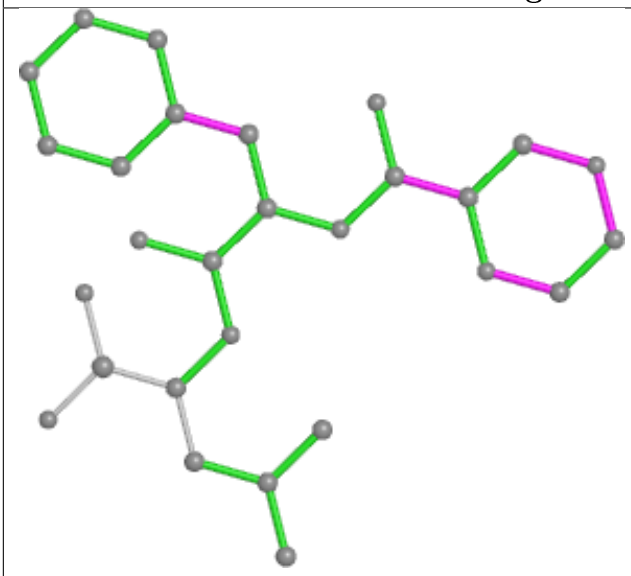
There are no ring outliers.

1 monomer is involved in 1 short contact:

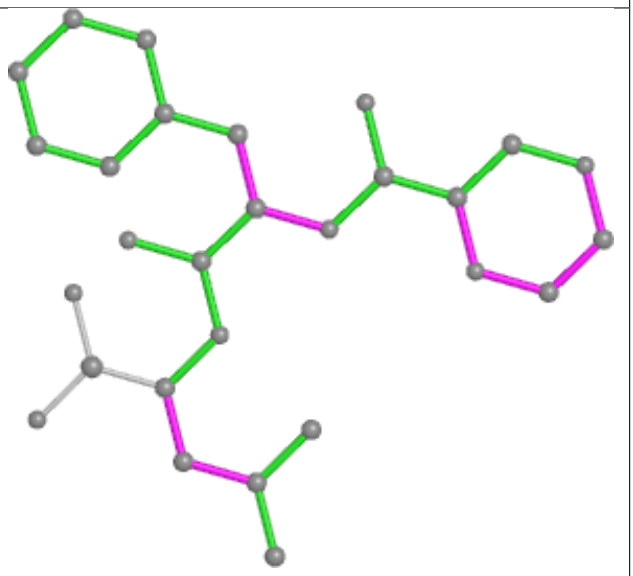
Mol	Chain	Res	Type	Clashes	Symm-Clashes
17	N	201	BO2	1	0

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.

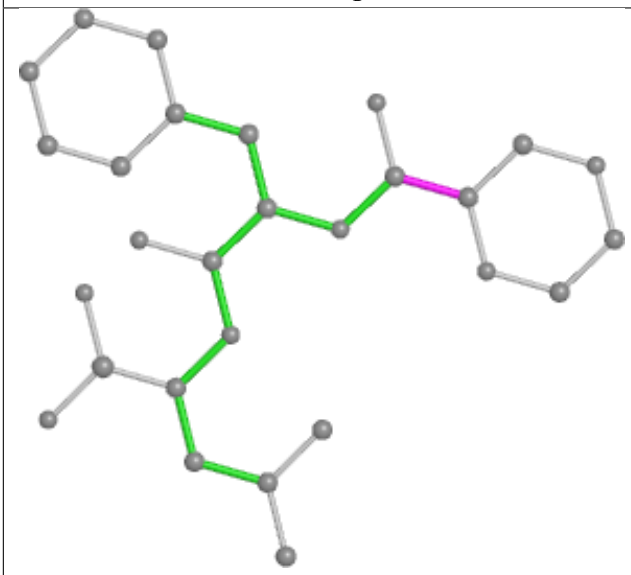
Ligand BO2 N 201



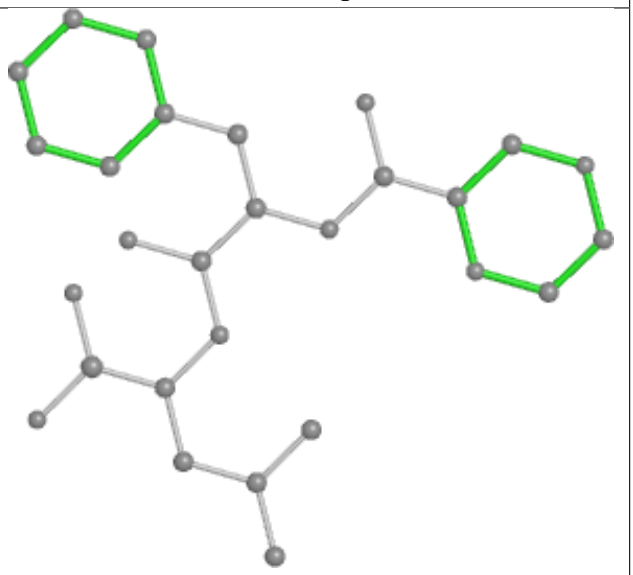
Bond lengths



Bond angles

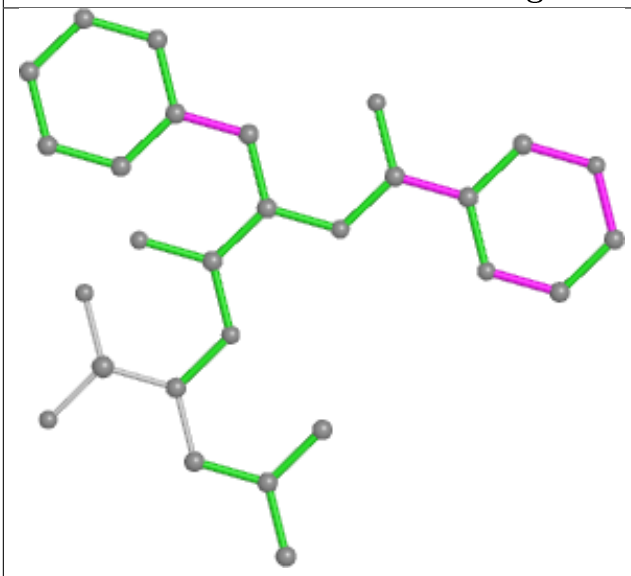


Torsions

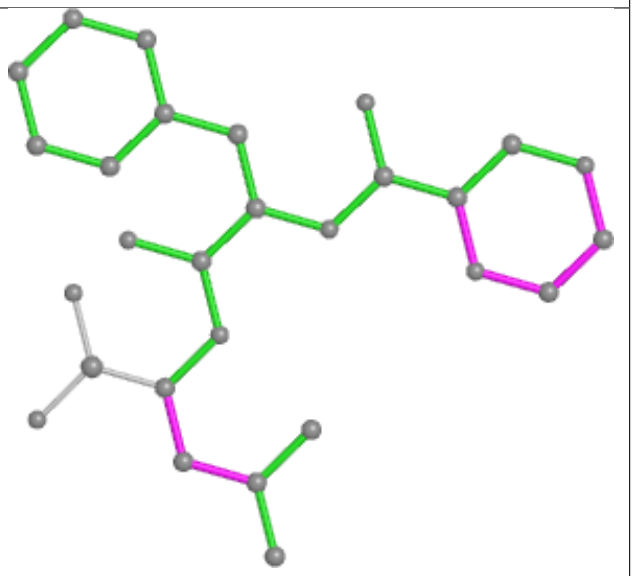


Rings

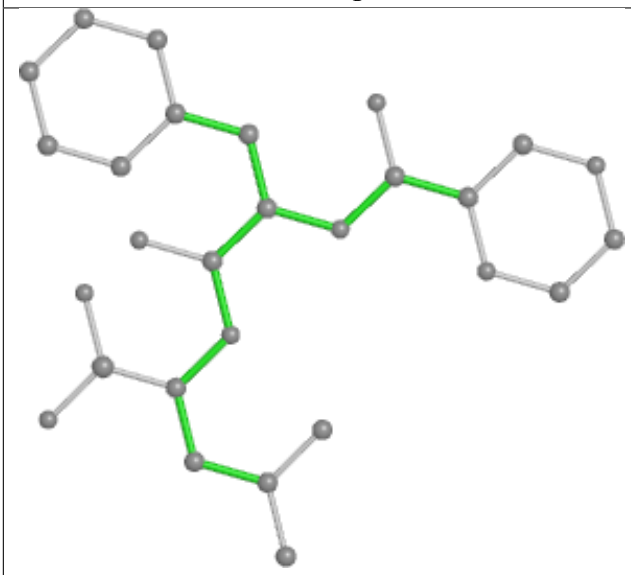
Ligand BO2 Y 301



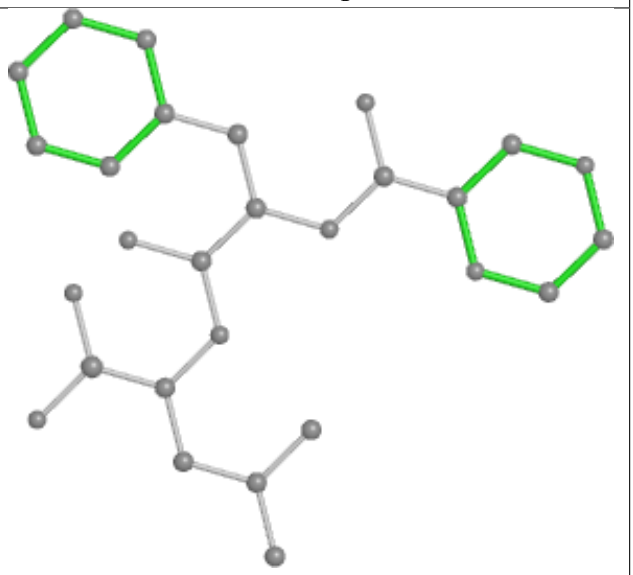
Bond lengths



Bond angles

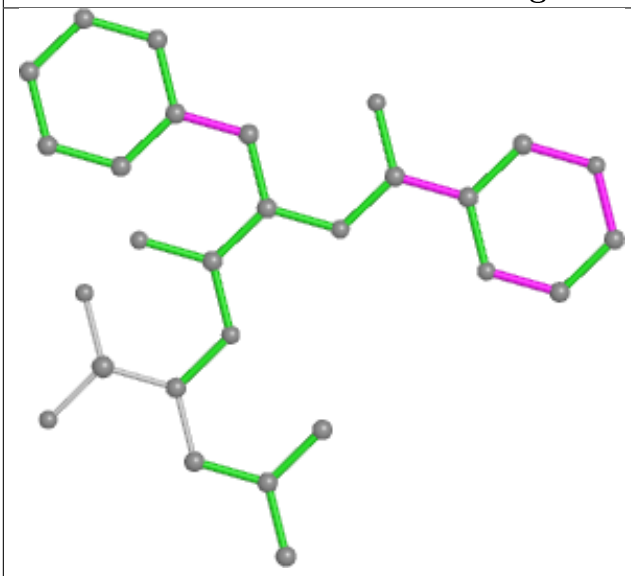


Torsions

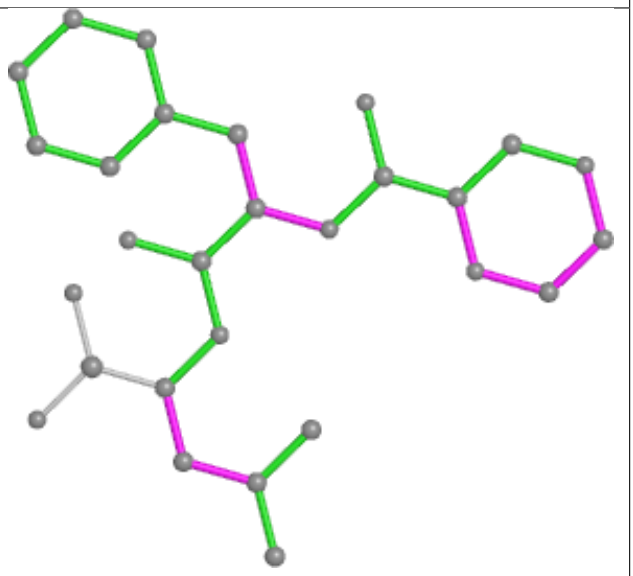


Rings

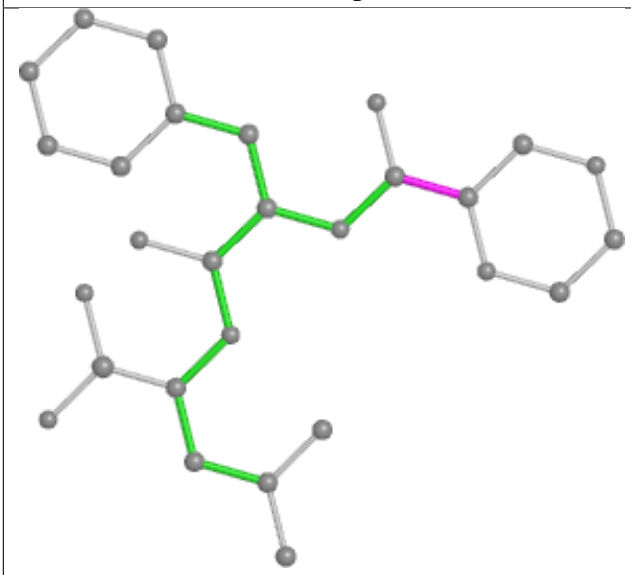
Ligand BO2 b 201



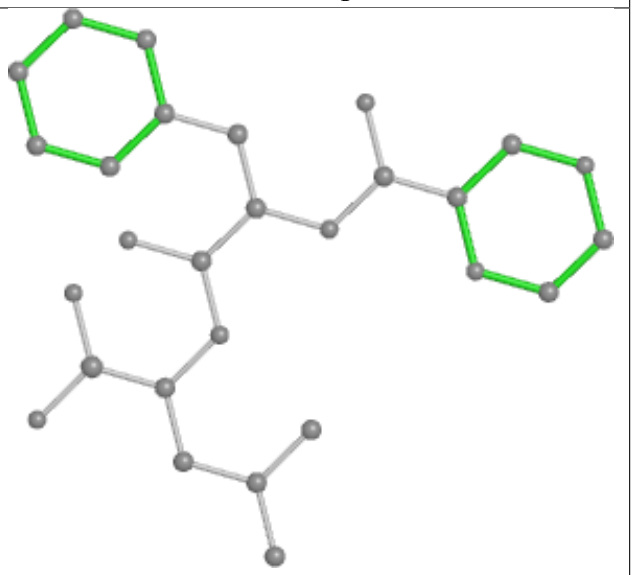
Bond lengths



Bond angles

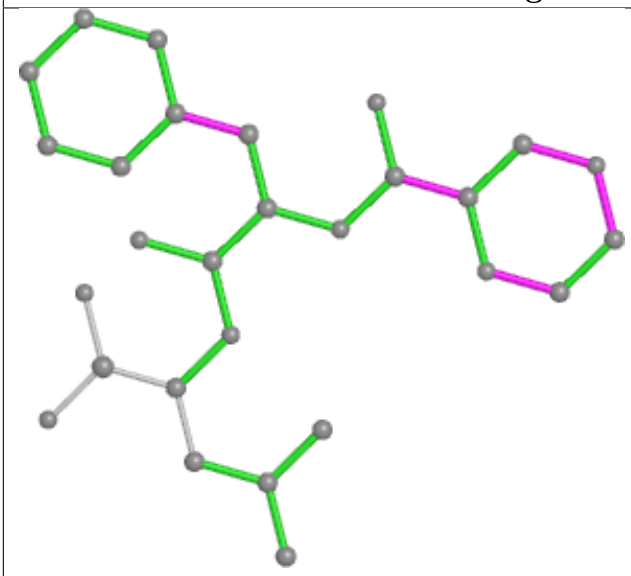


Torsions

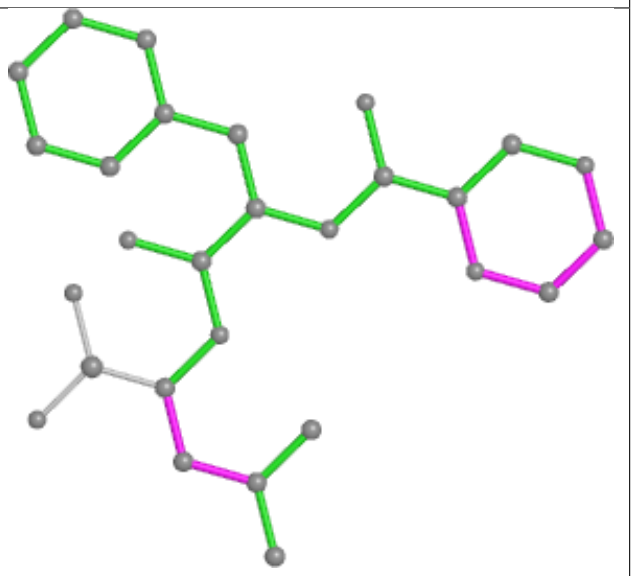


Rings

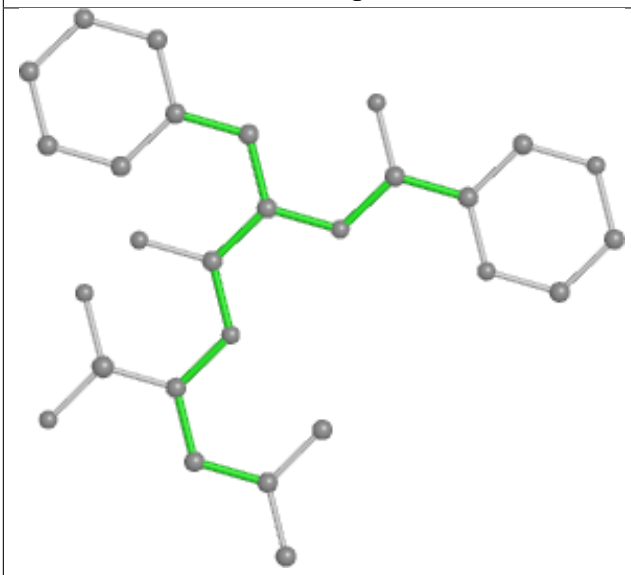
Ligand BO2 K 301



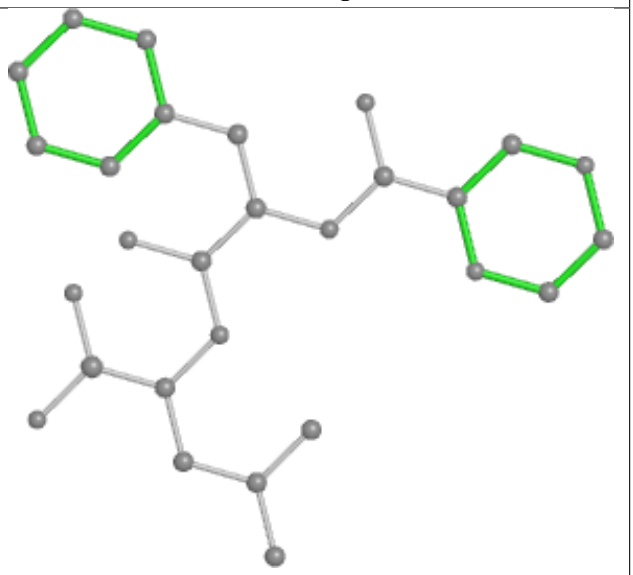
Bond lengths



Bond angles

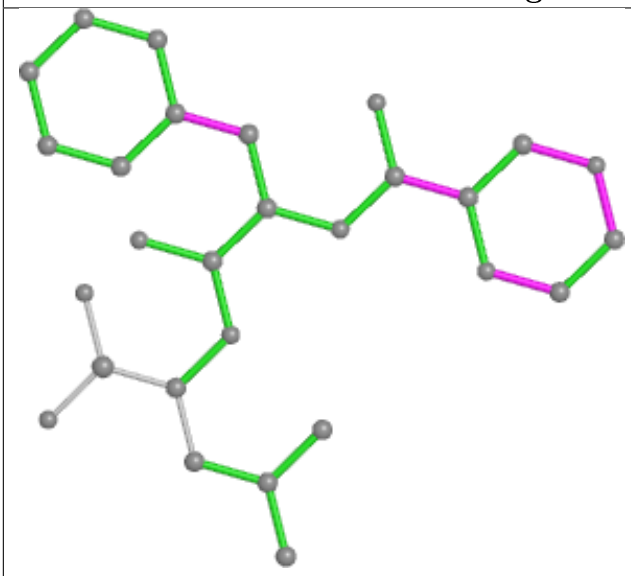


Torsions

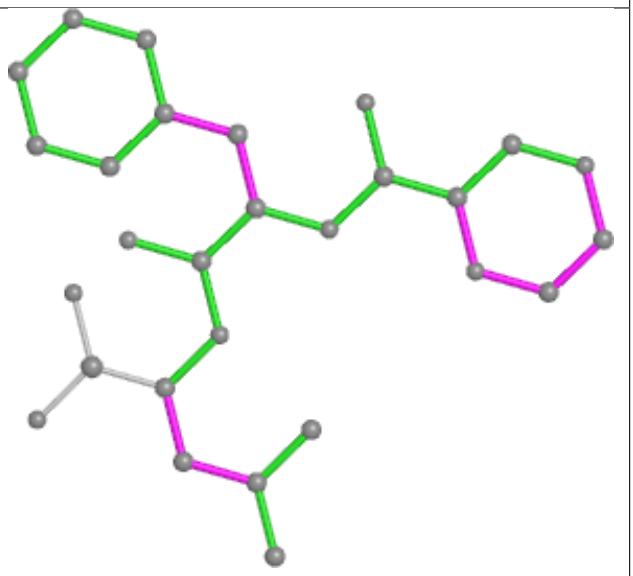


Rings

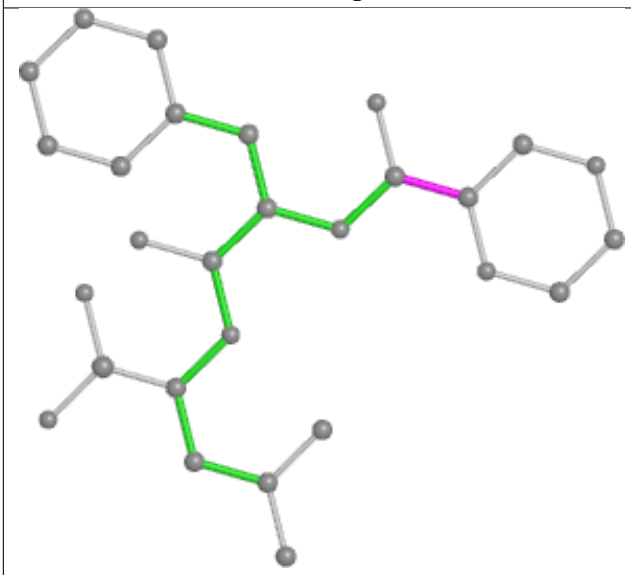
Ligand BO2 H 301



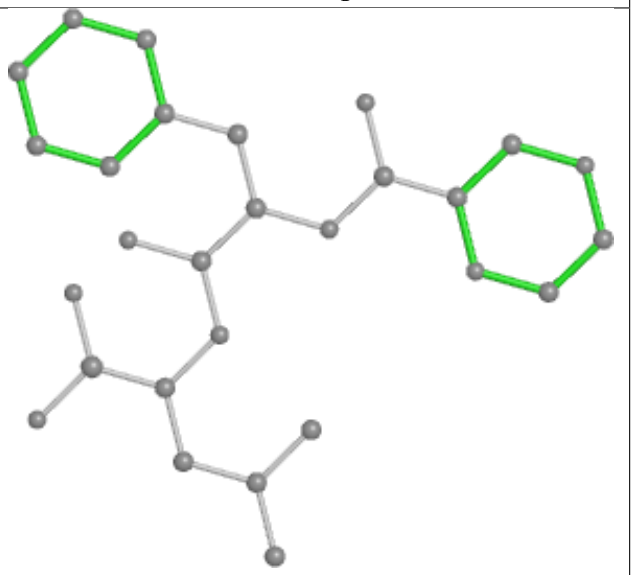
Bond lengths



Bond angles

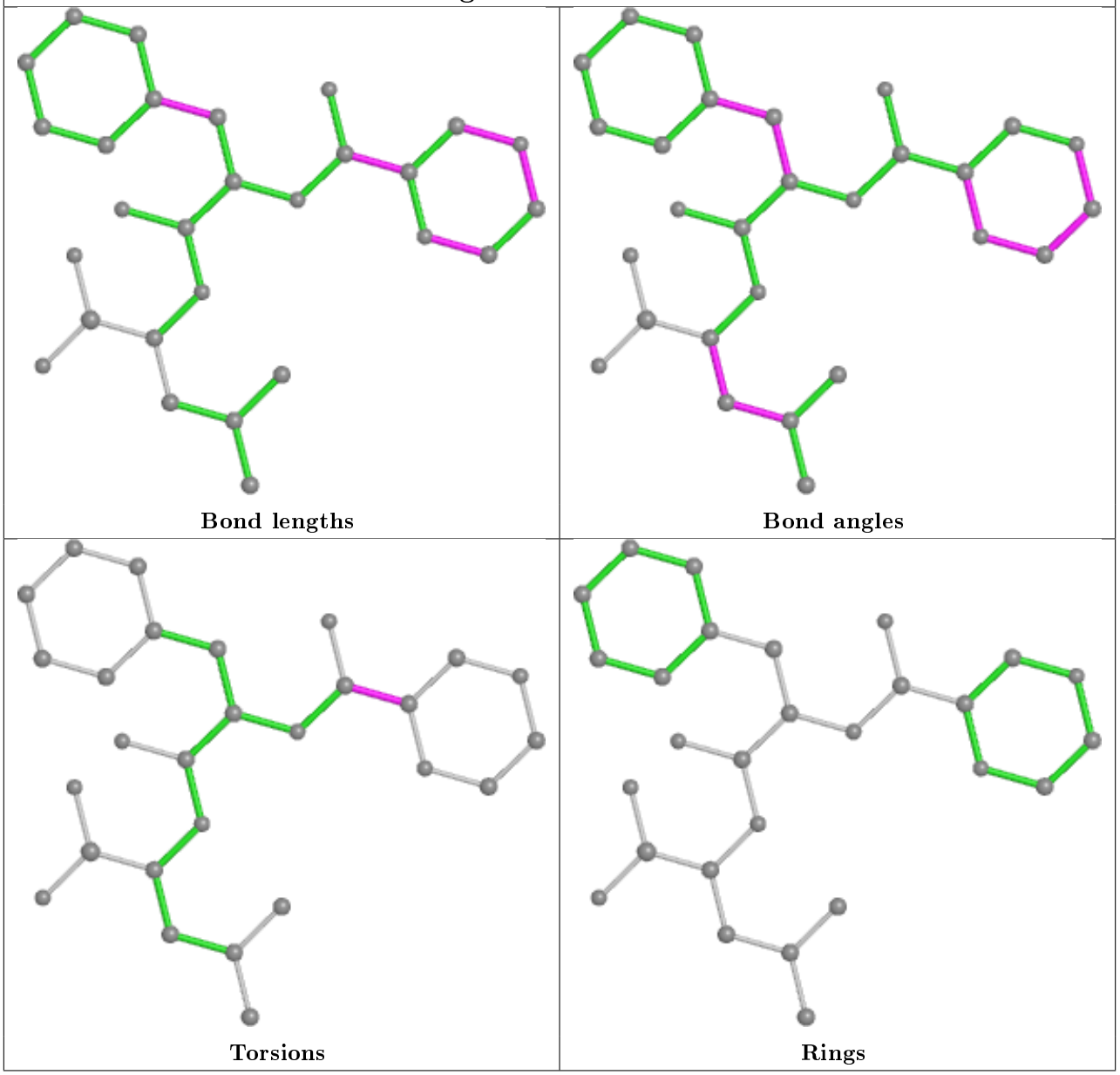


Torsions



Rings

Ligand BO2 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

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	250/250 (100%)	-0.45	4 (1%)	72	66	38, 55, 92, 128	0
1	O	250/250 (100%)	-0.45	4 (1%)	72	66	42, 59, 105, 138	0
2	B	244/258 (94%)	-0.36	7 (2%)	51	41	42, 62, 107, 161	0
2	P	244/258 (94%)	-0.38	7 (2%)	51	41	44, 63, 106, 155	0
3	C	240/254 (94%)	-0.11	14 (5%)	23	15	41, 68, 137, 175	0
3	Q	240/254 (94%)	-0.09	14 (5%)	23	15	32, 75, 152, 187	0
4	D	235/260 (90%)	-0.40	2 (0%)	84	80	48, 66, 97, 129	0
4	R	235/260 (90%)	-0.25	8 (3%)	45	35	55, 73, 113, 143	0
5	E	231/234 (98%)	-0.37	2 (0%)	84	80	47, 67, 107, 152	0
5	S	231/234 (98%)	-0.33	5 (2%)	62	52	49, 71, 112, 143	0
6	F	243/288 (84%)	-0.48	6 (2%)	57	47	40, 61, 108, 135	0
6	T	243/288 (84%)	-0.43	3 (1%)	79	73	41, 67, 125, 151	0
7	G	241/252 (95%)	-0.52	4 (1%)	70	63	39, 57, 96, 153	0
7	U	241/252 (95%)	-0.50	3 (1%)	79	73	42, 57, 92, 135	0
8	H	226/232 (97%)	-0.55	3 (1%)	77	72	35, 52, 91, 150	0
8	V	226/232 (97%)	-0.53	7 (3%)	49	39	38, 52, 87, 163	0
9	I	204/205 (99%)	-0.69	1 (0%)	91	88	37, 52, 80, 105	0
9	W	204/205 (99%)	-0.68	1 (0%)	91	88	39, 53, 80, 104	0
10	J	195/198 (98%)	-0.54	2 (1%)	82	77	39, 56, 82, 127	0
10	X	195/198 (98%)	-0.53	3 (1%)	73	68	41, 58, 83, 136	0
11	K	211/211 (100%)	-0.50	1 (0%)	91	88	44, 60, 90, 113	0
11	Y	211/211 (100%)	-0.52	0	100	100	45, 61, 90, 118	0
12	L	222/222 (100%)	-0.49	2 (0%)	84	80	42, 59, 100, 130	0
12	Z	222/222 (100%)	-0.47	4 (1%)	68	61	41, 60, 102, 136	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%)	-0.66	2 (0%) 84 80	38, 53, 77, 96	0
13	a	233/246 (94%)	-0.61	1 (0%) 92 91	39, 56, 79, 96	0
14	N	196/196 (100%)	-0.69	1 (0%) 91 88	35, 49, 79, 106	0
14	b	196/196 (100%)	-0.67	2 (1%) 82 77	37, 50, 82, 110	0
All	All	6342/6612 (95%)	-0.47	113 (1%) 68 61	32, 60, 104, 187	0

The worst 5 of 113 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
9	W	1	SER	5.5
3	C	238	LYS	4.9
2	B	220	ASN	4.8
3	Q	238	LYS	4.7
4	R	117	GLU	4.7

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
17	BO2	Y	301	28/28	0.90	0.28	43,52,66,69	0
17	BO2	b	201	28/28	0.90	0.21	37,48,60,61	0
17	BO2	N	201	28/28	0.90	0.21	33,48,59,61	0
15	MG	Z	301	1/1	0.91	0.17	66,66,66,66	0
17	BO2	K	301	28/28	0.93	0.21	38,47,61,61	0
17	BO2	V	301	28/28	0.93	0.20	44,48,68,70	0
17	BO2	H	301	28/28	0.94	0.18	42,48,64,67	0

Continued on next page...

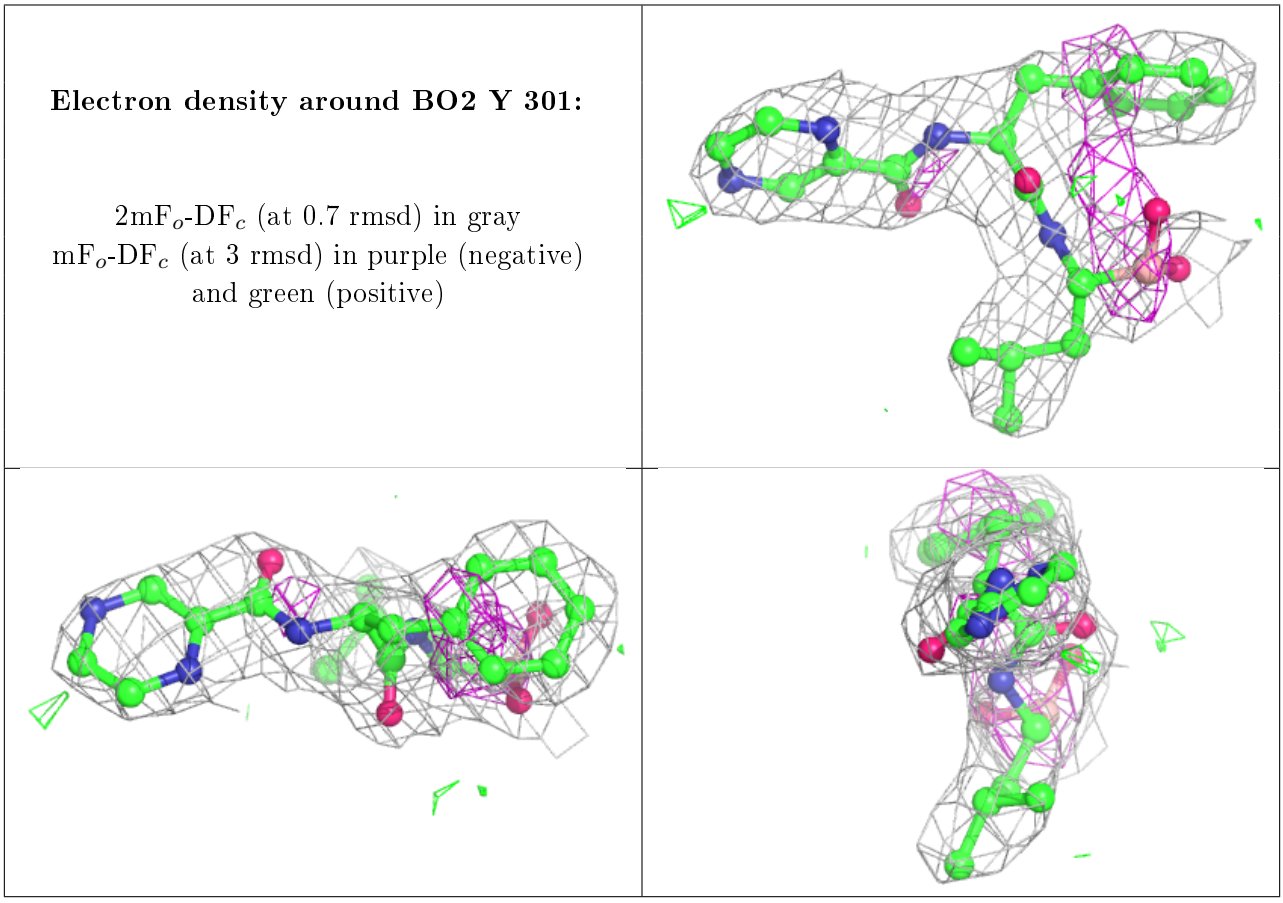
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
15	MG	J	201	1/1	0.96	0.09	50,50,50,50	0
16	CL	b	202	1/1	0.97	0.07	56,56,56,56	0
15	MG	N	202	1/1	0.97	0.10	49,49,49,49	0
16	CL	N	203	1/1	0.98	0.05	51,51,51,51	0
16	CL	G	302	1/1	0.98	0.17	41,41,41,41	0
15	MG	G	301	1/1	0.98	0.05	59,59,59,59	0
15	MG	I	301	1/1	0.98	0.07	55,55,55,55	0
15	MG	L	301	1/1	0.98	0.07	66,66,66,66	0
15	MG	K	302	1/1	0.99	0.07	56,56,56,56	0
16	CL	U	301	1/1	1.00	0.19	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.

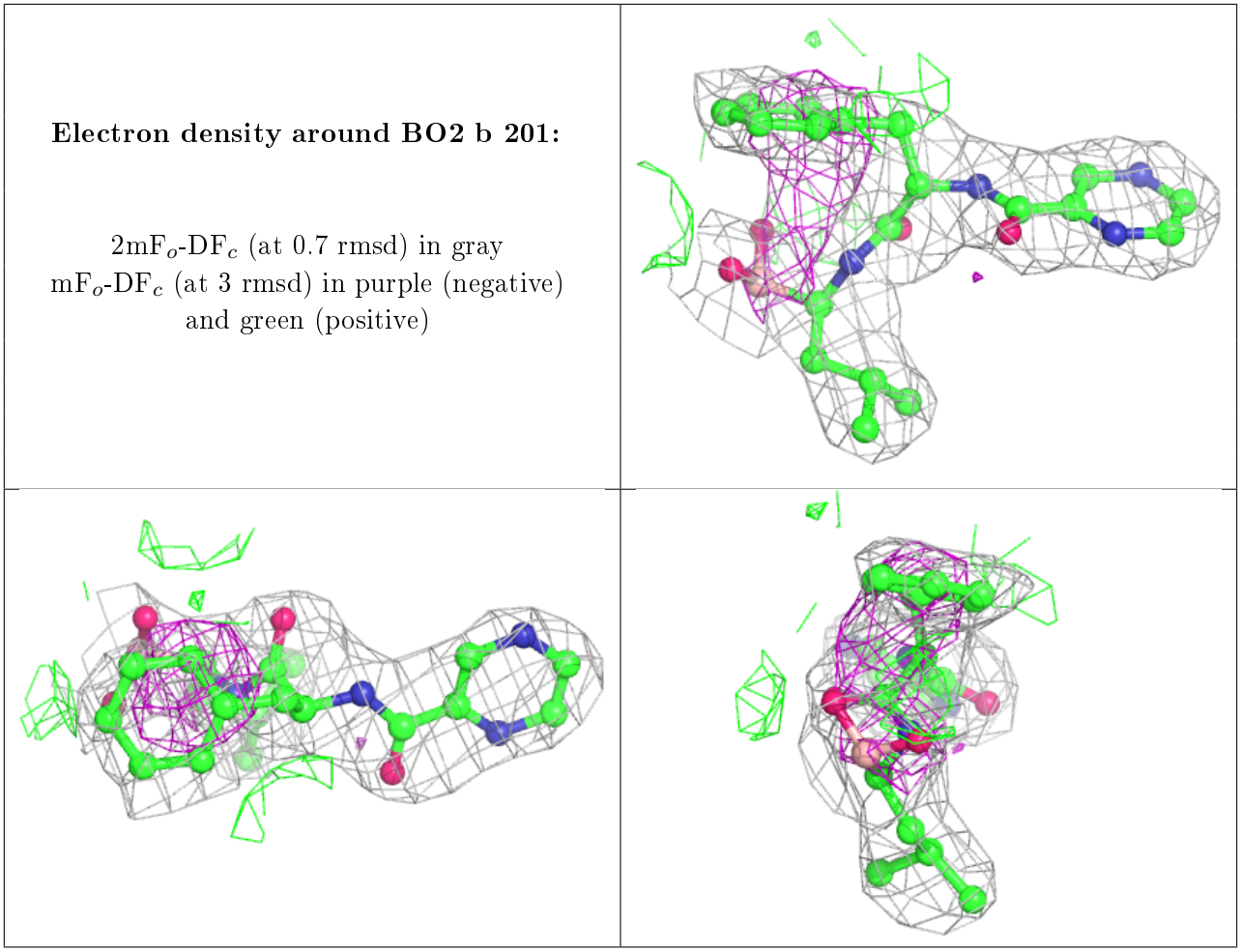
Electron density around BO2 Y 301:

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



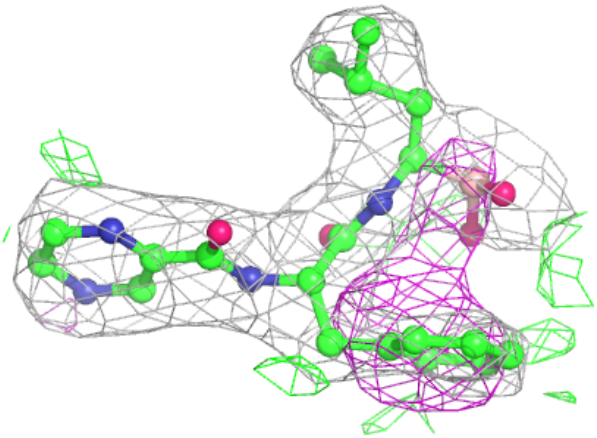
Electron density around BO2 b 201:

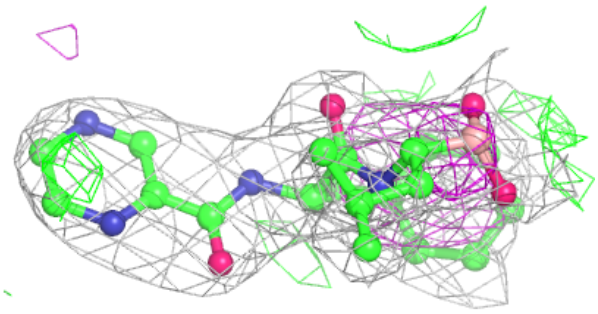
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

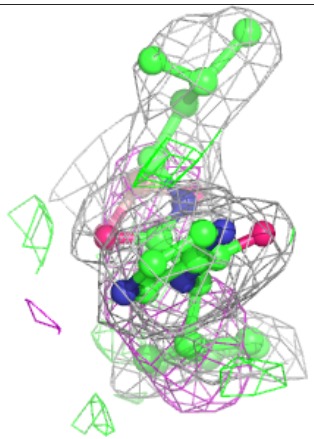


Electron density around BO2 N 201:

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

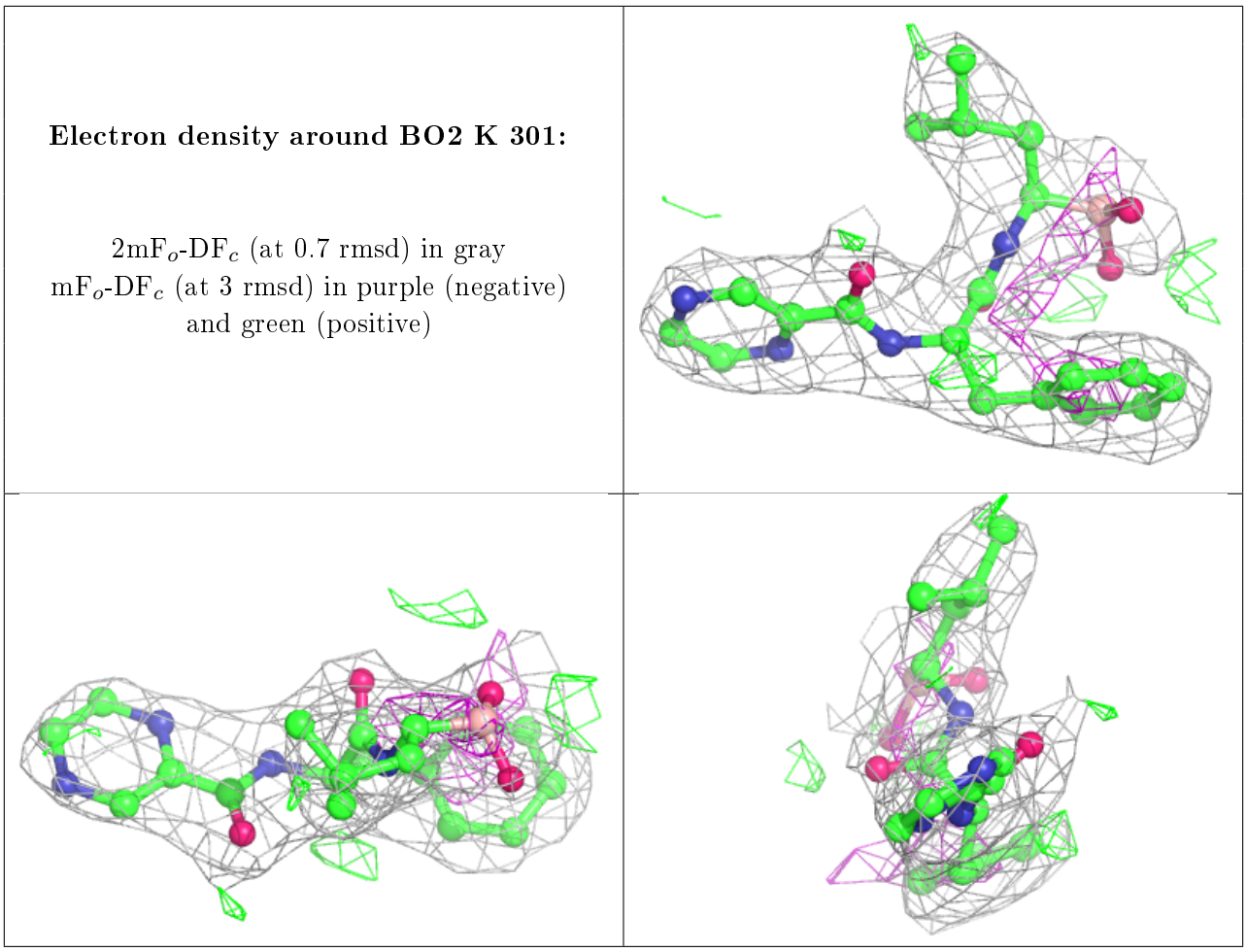






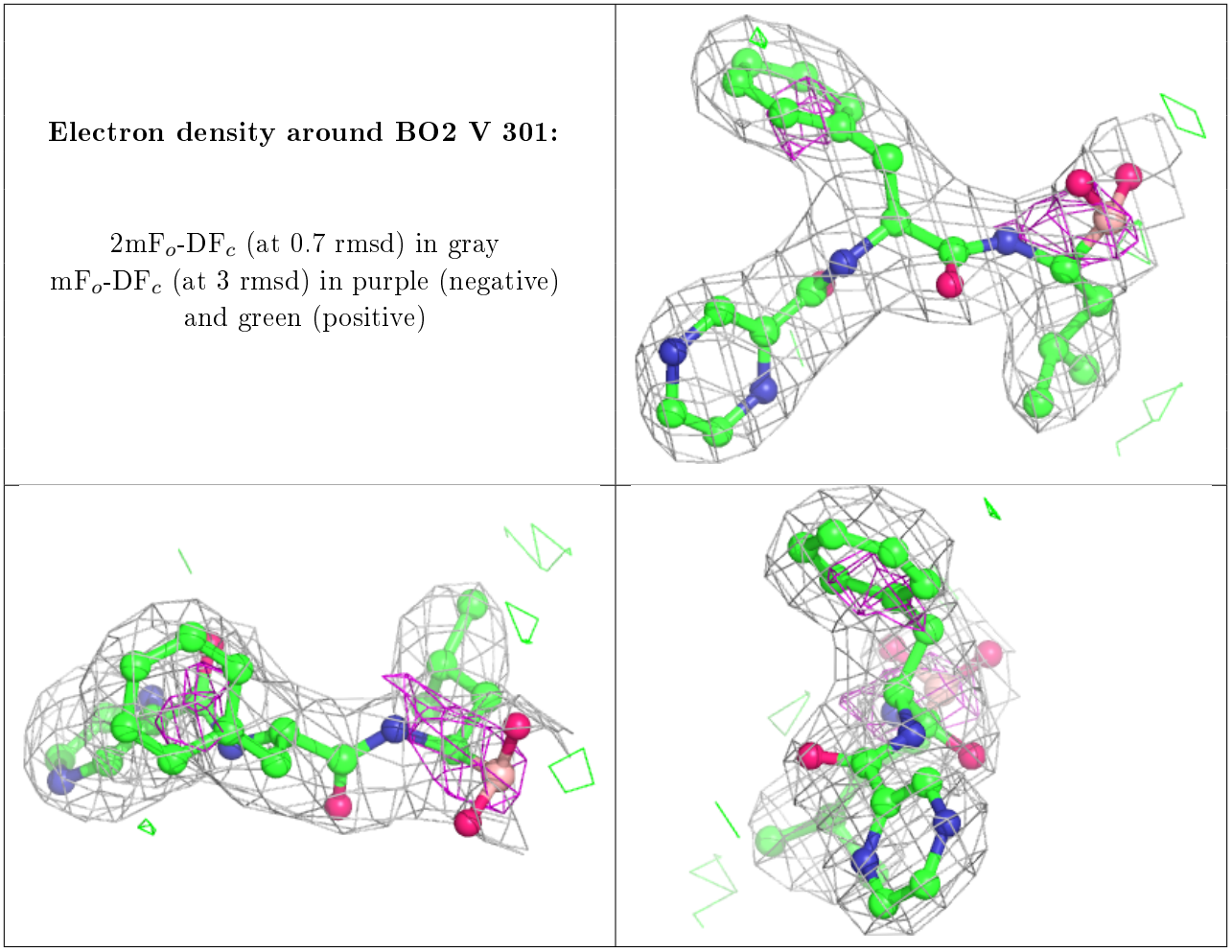
Electron density around BO2 K 301:

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



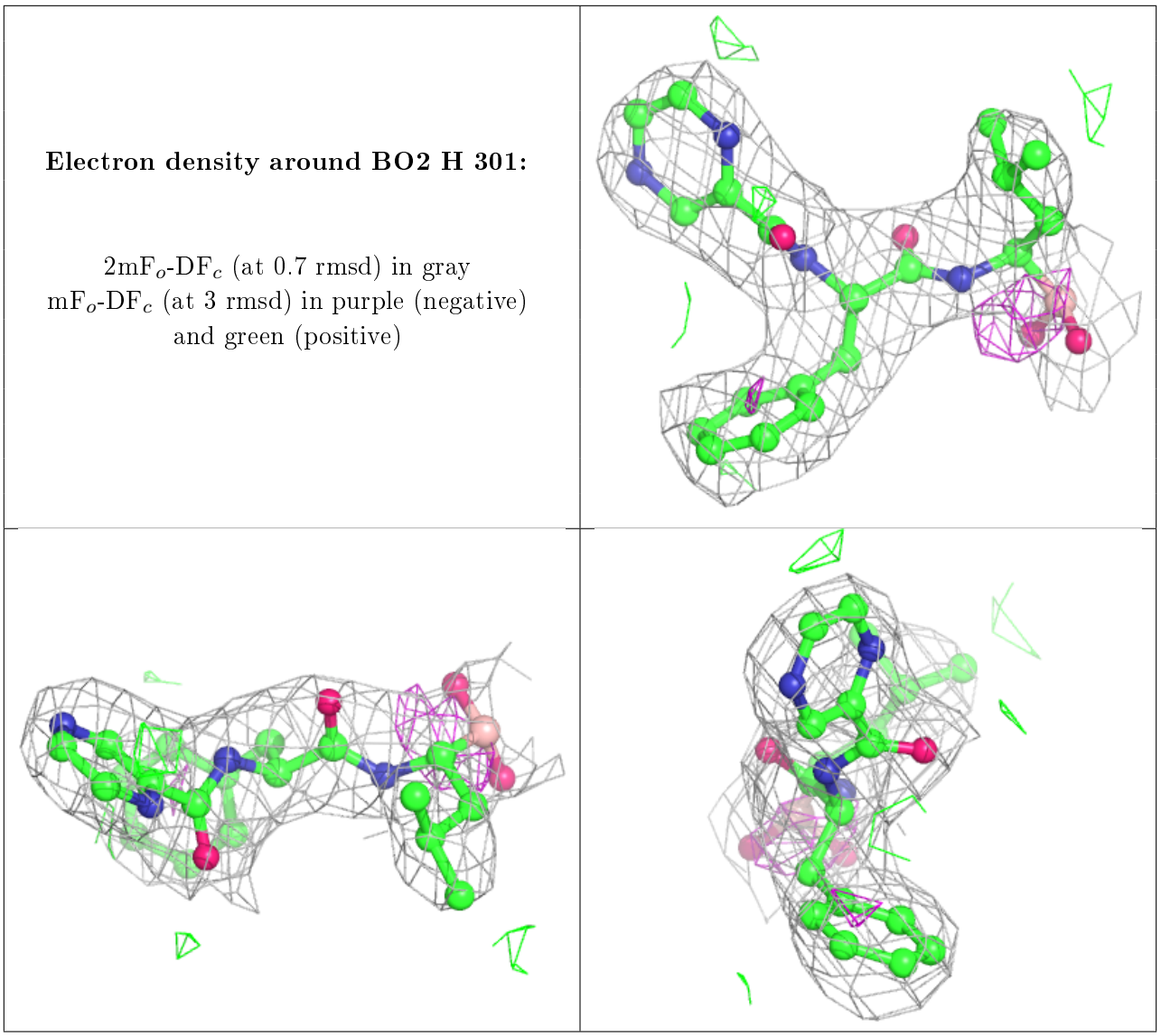
Electron density around BO2 V 301:

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



Electron density around BO2 H 301:

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



6.5 Other polymers

There are no such residues in this entry.