



# wwPDB X-ray Structure Validation Summary Report

Nov 28, 2021 – 02:14 pm GMT

PDB ID : 7B12  
Title : HUMAN IMMUNOPROTEASOME 20S PARTICLE IN COMPLEX WITH [2-(3-ethylphenyl)-1-[(2S)-3-phenyl-2-[(pyrazin-2-yl)formamido]propanamido]ethyl]boronic acid  
Authors : Musil, D.; Klein, M.; Crosignani, S.  
Deposited on : 2020-11-23  
Resolution : 2.43 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](mailto: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  symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

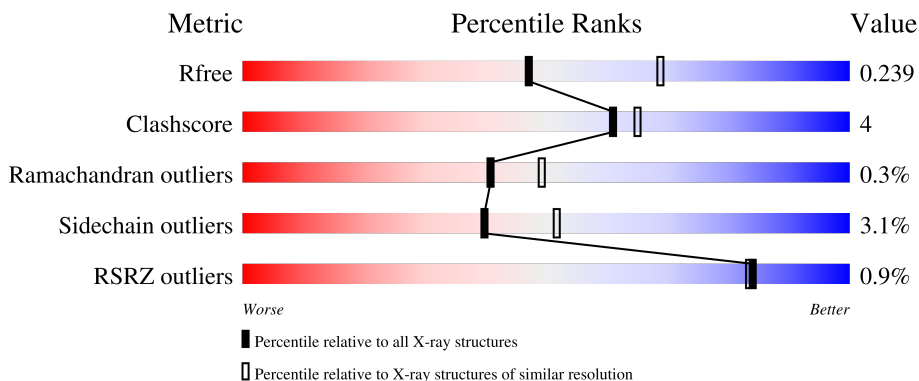
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.43 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.




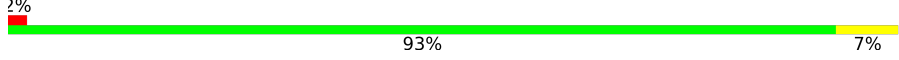
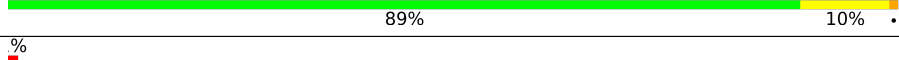
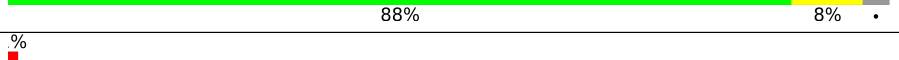
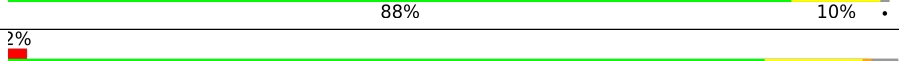
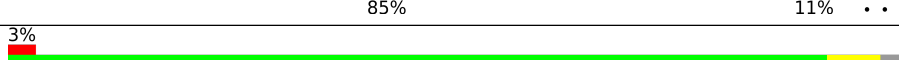
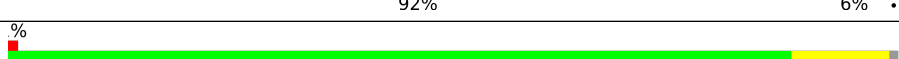
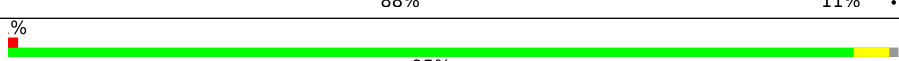
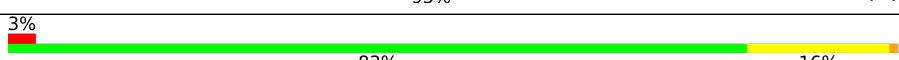
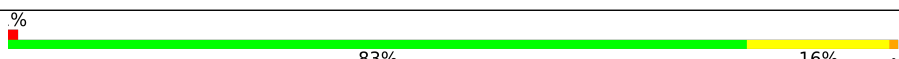
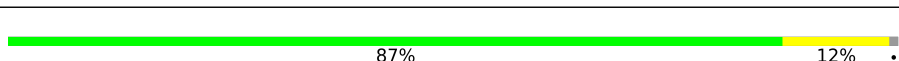
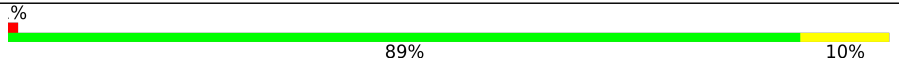

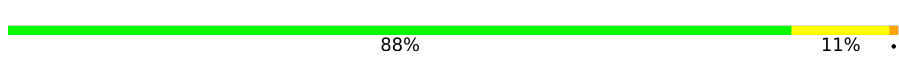



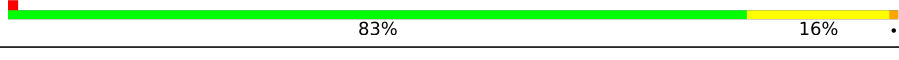
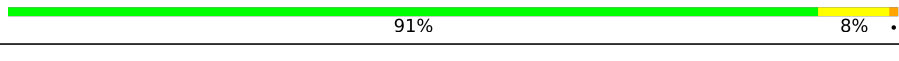
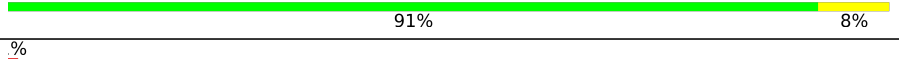
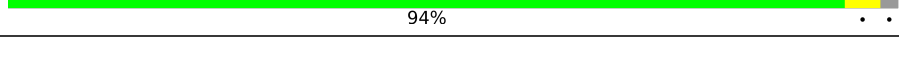
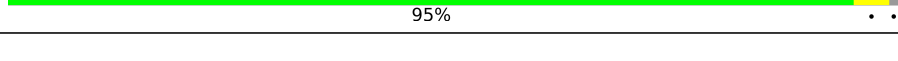

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	1	213	 88% 11% .
1	M	213	 89% 10% .
2	2	214	 88% 12%
2	N	214	 88% 12%
3	A	242	 2% 90% 9% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	O	242	 87% 12%
4	B	229	 93% 7%
4	P	229	 89% 10%
5	C	248	 88% 8%
5	Q	248	 88% 10%
6	D	235	 85% 11%
6	r	235	 92% 6%
7	E	232	 88% 11%
7	s	232	 95%
8	F	236	 83% 16%
8	T	236	 83% 16%
9	G	244	 87% 12%
9	U	244	 89% 10%
10	H	199	 90% 10%
10	V	199	 88% 11%
11	J	204	 83% 16%
11	X	204	 88% 11%
12	K	197	 85% 14%
12	Y	197	 83% 16%
13	L	203	 91% 8%
13	Z	203	 91% 8%
14	i	223	 94%
14	w	223	 95%

## 2 Entry composition [i](#)

There are 18 unique types of molecules in this entry. The entry contains 49338 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 beta type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	1	213	Total	C	N	O	S	54	0	0
			1654	1047	284	313	10			
1	M	213	Total	C	N	O	S	36	0	0
			1654	1047	284	313	10			

- Molecule 2 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	2	214	Total	C	N	O	S	20	0	0
			1673	1055	289	317	12			
2	N	214	Total	C	N	O	S	25	0	0
			1673	1055	289	317	12			

- Molecule 3 is a protein called Proteasome subunit alpha type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	240	Total	C	N	O	S	154	0	0
			1866	1185	309	359	13			
3	O	240	Total	C	N	O	S	114	0	0
			1866	1182	312	359	13			

- Molecule 4 is a protein called Proteasome subunit alpha type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	229	Total	C	N	O	S	123	0	0
			1787	1144	301	336	6			
4	P	229	Total	C	N	O	S	98	0	0
			1787	1144	301	336	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	241	Total	C	N	O	S	156	0	0
			1898	1202	325	361	10			
5	Q	245	Total	C	N	O	S	156	0	0
			1931	1220	332	369	10			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	227	Total	C	N	O	S	182	0	0
			1798	1132	318	343	5			
6	r	229	Total	C	N	O	S	239	0	0
			1816	1143	322	346	5			

- Molecule 7 is a protein called Proteasome subunit alpha type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	230	Total	C	N	O	S	117	0	0
			1758	1103	291	353	11			
7	s	230	Total	C	N	O	S	134	0	0
			1752	1100	290	351	11			

- Molecule 8 is a protein called Proteasome subunit alpha type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	236	Total	C	N	O	S	109	0	0
			1857	1162	334	350	11			
8	T	236	Total	C	N	O	S	129	0	0
			1857	1162	334	350	11			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	G	242	Total	C	N	O	S	149	0	0
			1893	1201	323	358	11			
9	U	243	Total	C	N	O	S	134	0	0
			1899	1204	324	360	11			

- Molecule 10 is a protein called Proteasome subunit beta type-9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	H	199	Total	C	N	O	S	26	0	0
			1493	939	254	291	9			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	V	199	1493	939	254	291	9	34	0	0

- Molecule 11 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	J	204	1591	1013	265	294	19	33	0	0
11	X	204	1591	1013	265	294	19	23	0	0

- Molecule 12 is a protein called Proteasome subunit beta type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	K	196	1571	1006	267	289	9	56	0	0
12	Y	197	1578	1011	268	290	9	51	0	0

- Molecule 13 is a protein called Proteasome subunit beta type-8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	L	203	1576	984	276	301	15	18	0	0
13	Z	203	1576	984	276	301	15	37	0	0

- Molecule 14 is a protein called Proteasome subunit beta type-10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
14	i	219	1609	1010	286	305	8	76	0	0
14	w	220	1616	1014	287	307	8	69	0	0

- Molecule 15 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



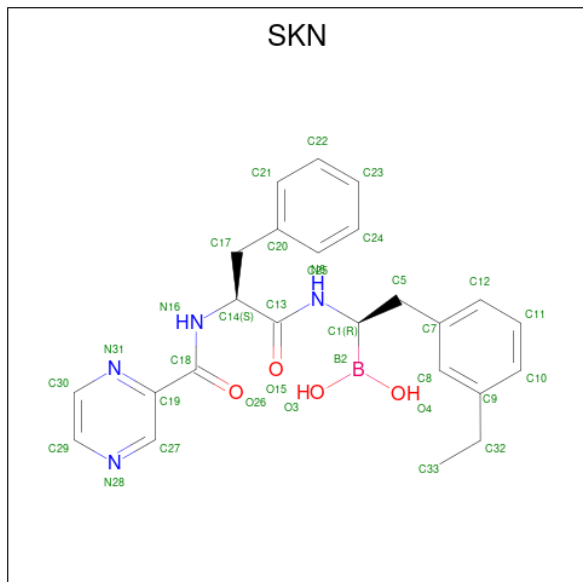
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	1	1	Total O S 5 4 1	0	0
15	1	1	Total O S 5 4 1	0	0
15	K	1	Total O S 5 4 1	0	0
15	L	1	Total O S 5 4 1	0	0
15	Y	1	Total O S 5 4 1	0	0

- Molecule 16 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	A	1	Total Na 1 1	0	0
16	L	1	Total Na 1 1	0	0
16	O	1	Total Na 1 1	0	0
16	Z	1	Total Na 1 1	0	0
16	i	1	Total Na 1 1	0	0
16	w	1	Total Na 1 1	0	0

- Molecule 17 is ((R)-2-(3-ethylphenyl)-1-((S)-3-phenyl-2-(pyrazine-2-carboxamido)prop

anamido)ethyl)boronic acid (three-letter code: SKN) (formula: C<sub>24</sub>H<sub>27</sub>BN<sub>4</sub>O<sub>4</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	B	C	N	O		
17	H	1	Total 33	B 1	C 24	N 4	O 4	0	0
17	L	1	Total 33	B 1	C 24	N 4	O 4	0	0
17	V	1	Total 33	B 1	C 24	N 4	O 4	0	0
17	Z	1	Total 33	B 1	C 24	N 4	O 4	0	0
17	i	1	Total 33	B 1	C 24	N 4	O 4	0	0
17	w	1	Total 33	B 1	C 24	N 4	O 4	0	0

- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	1	56	Total 56	O 56	0	0
18	2	60	Total 60	O 60	0	0
18	A	32	Total 32	O 32	0	0
18	B	26	Total 26	O 26	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	C	36	Total O 36 36	0	0
18	D	18	Total O 18 18	0	0
18	E	28	Total O 28 28	0	0
18	F	23	Total O 23 23	0	0
18	G	21	Total O 21 21	0	0
18	H	32	Total O 32 32	0	0
18	J	59	Total O 59 59	0	0
18	K	16	Total O 16 16	0	0
18	L	41	Total O 41 41	0	0
18	M	43	Total O 43 43	0	0
18	N	54	Total O 54 54	0	0
18	O	34	Total O 34 34	0	0
18	P	28	Total O 28 28	0	0
18	Q	33	Total O 33 33	0	0
18	T	38	Total O 38 38	0	0
18	U	33	Total O 33 33	0	0
18	V	42	Total O 42 42	0	0
18	X	47	Total O 47 47	0	0
18	Y	28	Total O 28 28	0	0
18	Z	38	Total O 38 38	0	0
18	i	35	Total O 35 35	0	0

*Continued on next page...*


*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
18	r	27	Total O 27 27	0	0
18	s	20	Total O 20 20	0	0
18	w	48	Total O 48 48	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 beta type-1

Chain 1:  88% 11%




- Molecule 1: Proteasome subunit beta type-1

Chain M:  89% 10%




- Molecule 2: Proteasome subunit beta type-4

Chain 2:  88% 12%

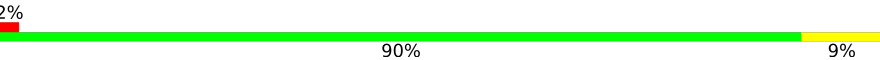


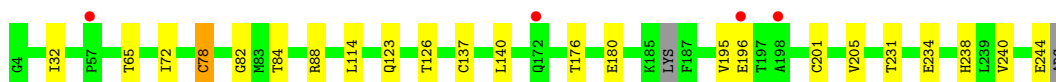
- Molecule 2: Proteasome subunit beta type-4

Chain N:  88% 12%




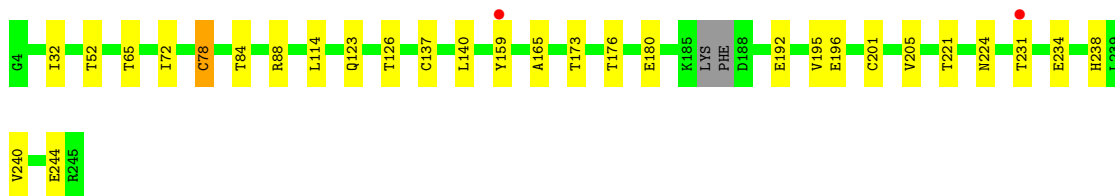
- Molecule 3: Proteasome subunit alpha type-6

Chain A:  90% 9% 2%

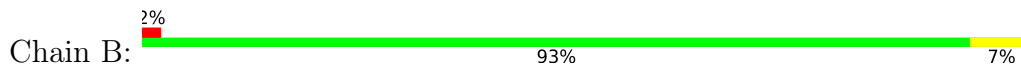


- Molecule 3: Proteasome subunit alpha type-6

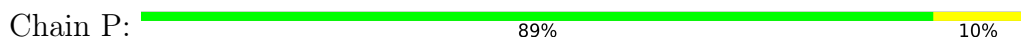
Chain O:  87% 12% 1%



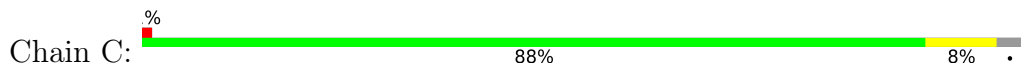
- Molecule 4: Proteasome subunit alpha type-2



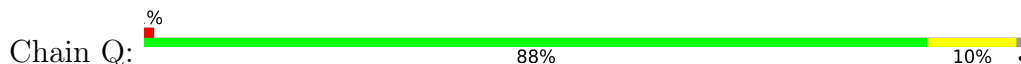
- Molecule 4: Proteasome subunit alpha type-2



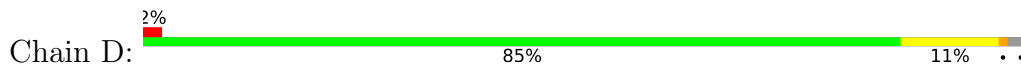
- Molecule 5: Proteasome subunit alpha type-4



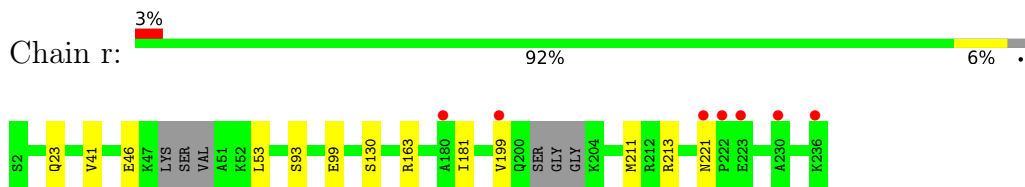
- Molecule 5: Proteasome subunit alpha type-4



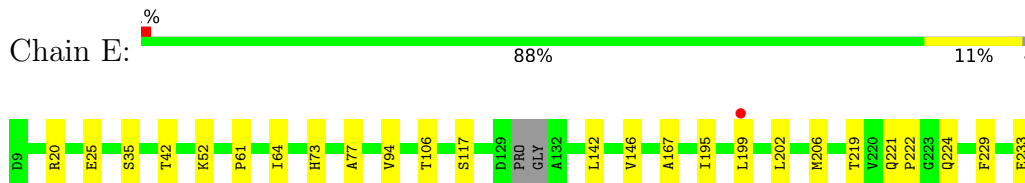
- Molecule 6: Proteasome subunit alpha type-7



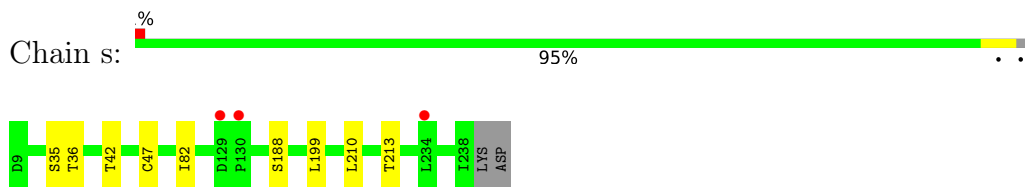
- Molecule 6: Proteasome subunit alpha type-7



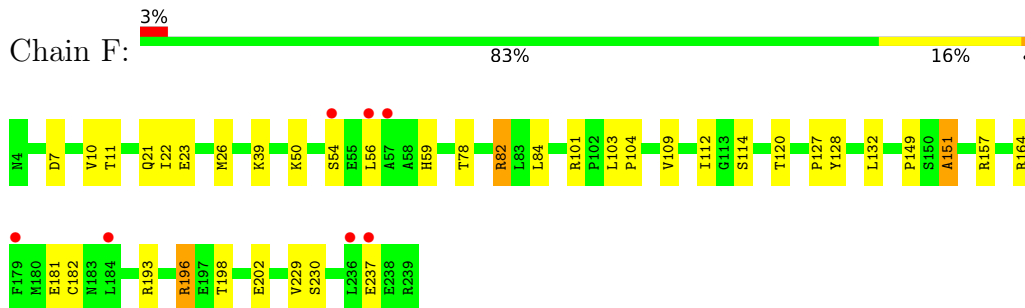
- Molecule 7: Proteasome subunit alpha type-5



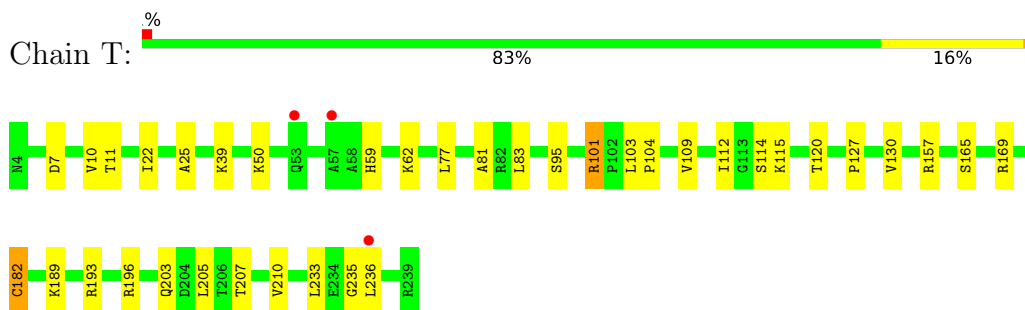
- Molecule 7: Proteasome subunit alpha type-5



- Molecule 8: Proteasome subunit alpha type-1



- Molecule 8: Proteasome subunit alpha type-1

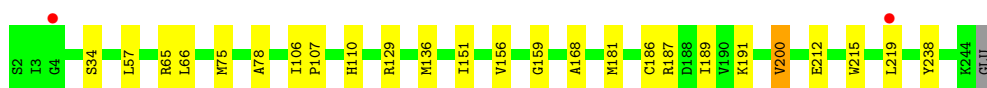
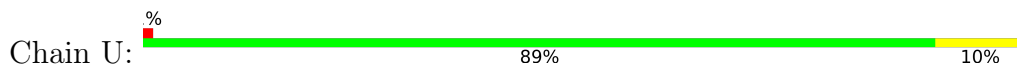


- Molecule 9: Proteasome subunit alpha type-3

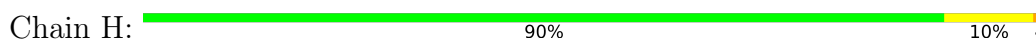




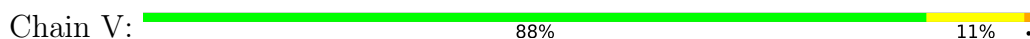
- Molecule 9: Proteasome subunit alpha type-3



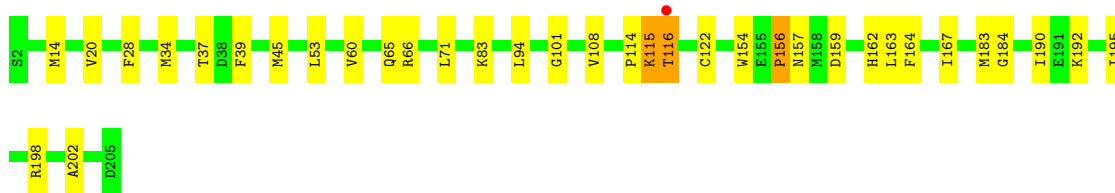
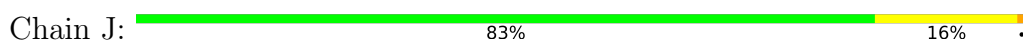
- Molecule 10: Proteasome subunit beta type-9



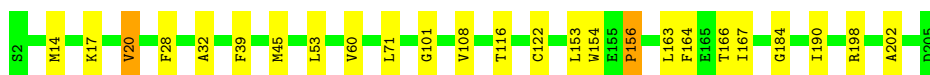
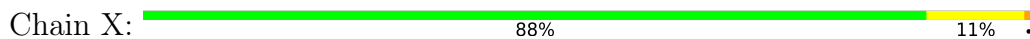
- Molecule 10: Proteasome subunit beta type-9



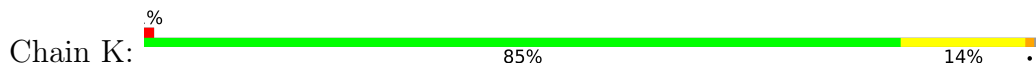
- Molecule 11: Proteasome subunit beta type-3

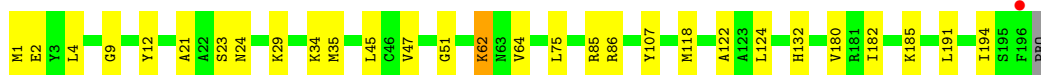


- Molecule 11: Proteasome subunit beta type-3

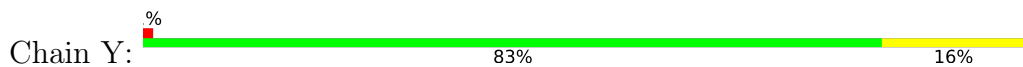


- Molecule 12: Proteasome subunit beta type-2

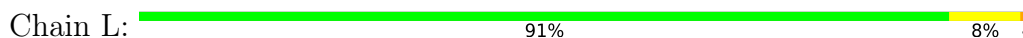




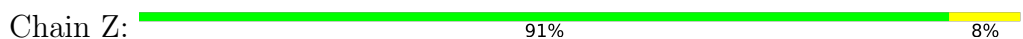
- Molecule 12: Proteasome subunit beta type-2



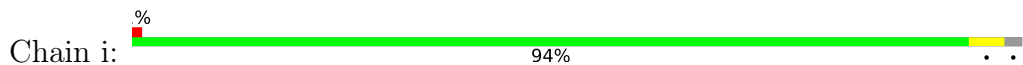
- Molecule 13: Proteasome subunit beta type-8



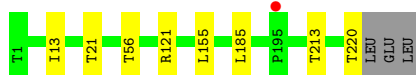
- Molecule 13: Proteasome subunit beta type-8



- Molecule 14: Proteasome subunit beta type-10



- Molecule 14: Proteasome subunit beta type-10



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.00Å 201.74Å 161.93Å 90.00° 106.83° 90.00°	Depositor
Resolution (Å)	107.89 – 2.43 107.89 – 2.43	Depositor EDS
% Data completeness (in resolution range)	93.4 (107.89-2.43) 64.8 (107.89-2.43)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 2.42Å)	Xtriage
Refinement program	BUSTER 2.11.7 (18-SEP-2020)	Depositor
R, $R_{free}$	0.180 , 0.225 0.196 , 0.239	Depositor DCC
$R_{free}$ test set	8735 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.3	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	49338	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.37% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: NA, SO4, SKN

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	1	0.50	0/1684	0.66	0/2268
1	M	0.49	0/1684	0.66	0/2268
2	2	0.46	0/1706	0.63	0/2309
2	N	0.46	0/1706	0.63	0/2309
3	A	0.39	0/1899	0.59	0/2567
3	O	0.41	0/1898	0.60	0/2565
4	B	0.43	0/1826	0.60	0/2474
4	P	0.45	0/1826	0.62	0/2474
5	C	0.42	0/1927	0.59	0/2595
5	Q	0.41	0/1960	0.59	0/2639
6	D	0.39	0/1822	0.54	0/2458
6	r	0.39	0/1840	0.54	0/2481
7	E	0.42	0/1783	0.55	0/2406
7	s	0.41	0/1779	0.57	0/2404
8	F	0.45	0/1891	0.66	0/2555
8	T	0.46	0/1891	0.65	0/2555
9	G	0.40	0/1928	0.57	0/2596
9	U	0.43	0/1934	0.59	0/2604
10	H	0.45	0/1521	0.66	0/2062
10	V	0.44	0/1521	0.65	1/2062 (0.0%)
11	J	0.46	0/1620	0.65	1/2184 (0.0%)
11	X	0.46	0/1620	0.65	0/2184
12	K	0.45	0/1603	0.65	0/2168
12	Y	0.44	0/1611	0.63	0/2180
13	L	0.46	0/1609	0.65	1/2170 (0.0%)
13	Z	0.43	0/1609	0.63	0/2170
14	i	0.46	0/1635	0.63	0/2222
14	w	0.45	0/1642	0.65	0/2232
All	All	0.44	0/48975	0.62	3/66161 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	V	104	ASP	CB-CA-C	-5.49	99.42	110.40
11	J	115	LYS	C-N-CA	5.39	135.16	121.70
13	L	115	ASP	N-CA-CB	-5.13	101.36	110.60

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	1	1654	0	1656	13	0
1	M	1654	0	1656	9	0
2	2	1673	0	1650	15	0
2	N	1673	0	1650	13	0
3	A	1866	0	1865	14	0
3	O	1866	0	1870	16	0
4	B	1787	0	1782	13	0
4	P	1787	0	1782	18	0
5	C	1898	0	1917	11	0
5	Q	1931	0	1945	12	0
6	D	1798	0	1820	15	0
6	r	1816	0	1841	0	0
7	E	1758	0	1742	19	0
7	s	1752	0	1736	0	0
8	F	1857	0	1845	24	0
8	T	1857	0	1845	20	0
9	G	1893	0	1882	19	0
9	U	1899	0	1887	15	0
10	H	1493	0	1457	9	0
10	V	1493	0	1457	11	0
11	J	1591	0	1609	19	0
11	X	1591	0	1609	11	0
12	K	1571	0	1573	19	0
12	Y	1578	0	1580	18	0
13	L	1576	0	1522	13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	Z	1576	0	1522	12	0
14	i	1609	0	1645	0	0
14	w	1616	0	1652	0	0
15	1	10	0	0	0	0
15	K	5	0	0	0	0
15	L	5	0	0	0	0
15	Y	5	0	0	0	0
16	A	1	0	0	0	0
16	L	1	0	0	0	0
16	O	1	0	0	0	0
16	Z	1	0	0	0	0
16	i	1	0	0	0	0
16	w	1	0	0	0	0
17	H	33	0	0	0	0
17	L	33	0	0	0	0
17	V	33	0	0	0	0
17	Z	33	0	0	0	0
17	i	33	0	0	0	0
17	w	33	0	0	0	0
18	1	56	0	0	1	0
18	2	60	0	0	1	0
18	A	32	0	0	0	0
18	B	26	0	0	0	0
18	C	36	0	0	0	0
18	D	18	0	0	0	0
18	E	28	0	0	0	0
18	F	23	0	0	1	0
18	G	21	0	0	0	0
18	H	32	0	0	0	0
18	J	59	0	0	1	0
18	K	16	0	0	0	0
18	L	41	0	0	0	0
18	M	43	0	0	0	0
18	N	54	0	0	0	0
18	O	34	0	0	0	0
18	P	28	0	0	0	0
18	Q	33	0	0	0	0
18	T	38	0	0	0	0
18	U	33	0	0	0	0
18	V	42	0	0	1	0
18	X	47	0	0	0	0
18	Y	28	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
18	Z	38	0	0	0	0
18	i	35	0	0	0	0
18	r	27	0	0	0	0
18	s	20	0	0	0	0
18	w	48	0	0	0	0
All	All	49338	0	47997	327	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 4.

The worst 5 of 327 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:U:34:SER:HB2	9:U:65:ARG:HH12	1.25	0.99
7:E:221:GLN:HG3	7:E:224:GLN:HB2	1.49	0.93
9:U:34:SER:HB2	9:U:65:ARG:NH1	1.90	0.86
1:1:145:LEU:HD21	1:1:182:ALA:HB2	1.59	0.84
1:M:145:LEU:HD21	1:M:182:ALA:HB2	1.60	0.83

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	1	211/213 (99%)	207 (98%)	4 (2%)	0	100	100
1	M	211/213 (99%)	206 (98%)	5 (2%)	0	100	100
2	2	212/214 (99%)	202 (95%)	10 (5%)	0	100	100
2	N	212/214 (99%)	202 (95%)	10 (5%)	0	100	100
3	A	236/242 (98%)	232 (98%)	4 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	O	236/242 (98%)	234 (99%)	2 (1%)	0	100	100
4	B	227/229 (99%)	222 (98%)	5 (2%)	0	100	100
4	P	227/229 (99%)	218 (96%)	9 (4%)	0	100	100
5	C	237/248 (96%)	233 (98%)	4 (2%)	0	100	100
5	Q	241/248 (97%)	233 (97%)	8 (3%)	0	100	100
6	D	221/235 (94%)	216 (98%)	3 (1%)	2 (1%)	17	20
6	r	223/235 (95%)	216 (97%)	5 (2%)	2 (1%)	17	20
7	E	226/232 (97%)	219 (97%)	7 (3%)	0	100	100
7	s	228/232 (98%)	220 (96%)	8 (4%)	0	100	100
8	F	234/236 (99%)	222 (95%)	10 (4%)	2 (1%)	17	20
8	T	234/236 (99%)	223 (95%)	10 (4%)	1 (0%)	34	41
9	G	240/244 (98%)	235 (98%)	5 (2%)	0	100	100
9	U	241/244 (99%)	232 (96%)	9 (4%)	0	100	100
10	H	197/199 (99%)	191 (97%)	6 (3%)	0	100	100
10	V	197/199 (99%)	191 (97%)	5 (2%)	1 (0%)	29	34
11	J	202/204 (99%)	191 (95%)	7 (4%)	4 (2%)	7	5
11	X	202/204 (99%)	196 (97%)	3 (2%)	3 (2%)	10	9
12	K	194/197 (98%)	186 (96%)	7 (4%)	1 (0%)	29	34
12	Y	195/197 (99%)	188 (96%)	6 (3%)	1 (0%)	29	34
13	L	201/203 (99%)	198 (98%)	3 (2%)	0	100	100
13	Z	201/203 (99%)	197 (98%)	4 (2%)	0	100	100
14	i	217/223 (97%)	209 (96%)	8 (4%)	0	100	100
14	w	218/223 (98%)	211 (97%)	7 (3%)	0	100	100
All	All	6121/6238 (98%)	5930 (97%)	174 (3%)	17 (0%)	41	49

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	D	213	ARG
6	r	213	ARG
11	J	156	PRO
11	X	156	PRO
6	D	163	ARG

### 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	1	178/178 (100%)	171 (96%)	7 (4%)	32	42
1	M	178/178 (100%)	167 (94%)	11 (6%)	18	23
2	2	177/177 (100%)	173 (98%)	4 (2%)	50	63
2	N	177/177 (100%)	170 (96%)	7 (4%)	31	41
3	A	204/206 (99%)	202 (99%)	2 (1%)	76	84
3	O	204/206 (99%)	201 (98%)	3 (2%)	65	76
4	B	188/188 (100%)	187 (100%)	1 (0%)	88	93
4	P	188/188 (100%)	181 (96%)	7 (4%)	34	45
5	C	202/208 (97%)	194 (96%)	8 (4%)	31	41
5	Q	205/208 (99%)	202 (98%)	3 (2%)	65	76
6	D	193/199 (97%)	186 (96%)	7 (4%)	35	46
6	r	195/199 (98%)	184 (94%)	11 (6%)	21	27
7	E	193/194 (100%)	191 (99%)	2 (1%)	76	84
7	s	192/194 (99%)	183 (95%)	9 (5%)	26	35
8	F	202/202 (100%)	193 (96%)	9 (4%)	27	36
8	T	202/202 (100%)	191 (95%)	11 (5%)	22	29
9	G	199/201 (99%)	194 (98%)	5 (2%)	47	60
9	U	200/201 (100%)	195 (98%)	5 (2%)	47	60
10	H	153/153 (100%)	146 (95%)	7 (5%)	27	35
10	V	153/153 (100%)	148 (97%)	5 (3%)	38	49
11	J	173/173 (100%)	169 (98%)	4 (2%)	50	63
11	X	173/173 (100%)	169 (98%)	4 (2%)	50	63
12	K	167/168 (99%)	165 (99%)	2 (1%)	71	81
12	Y	168/168 (100%)	164 (98%)	4 (2%)	49	61
13	L	165/165 (100%)	163 (99%)	2 (1%)	71	81
13	Z	165/165 (100%)	163 (99%)	2 (1%)	71	81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
14	i	171/175 (98%)	162 (95%)	9 (5%)	22	30
14	w	172/175 (98%)	164 (95%)	8 (5%)	26	35
All	All	5137/5174 (99%)	4978 (97%)	159 (3%)	40	52

5 of 159 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	X	122	CYS
7	s	36	THR
12	Y	185	LYS
14	i	213	THR
7	s	213	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
10	V	97	HIS
6	r	15	HIS
10	V	191	ASN
13	Z	53	GLN
12	K	174	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 17 ligands modelled in this entry, 6 are monoatomic - leaving 11 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

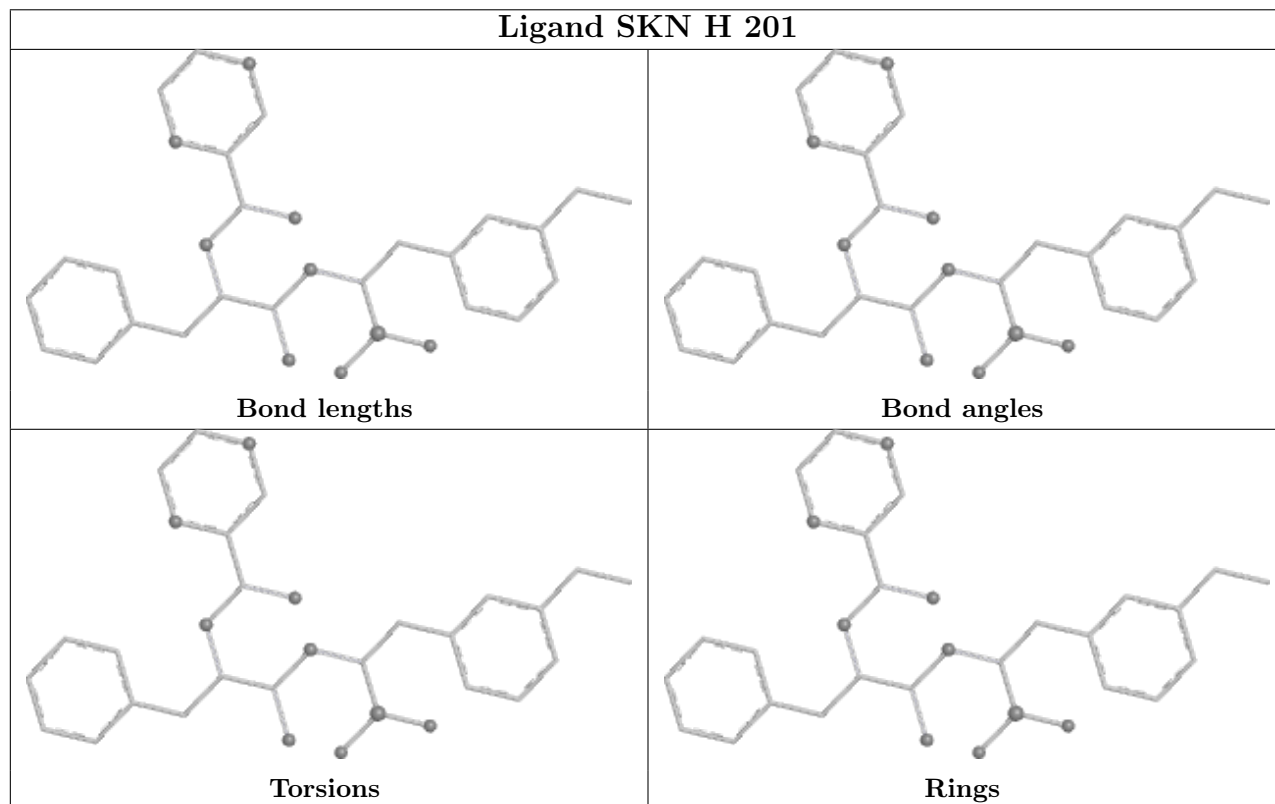
There are no chirality outliers.

There are no torsion outliers.

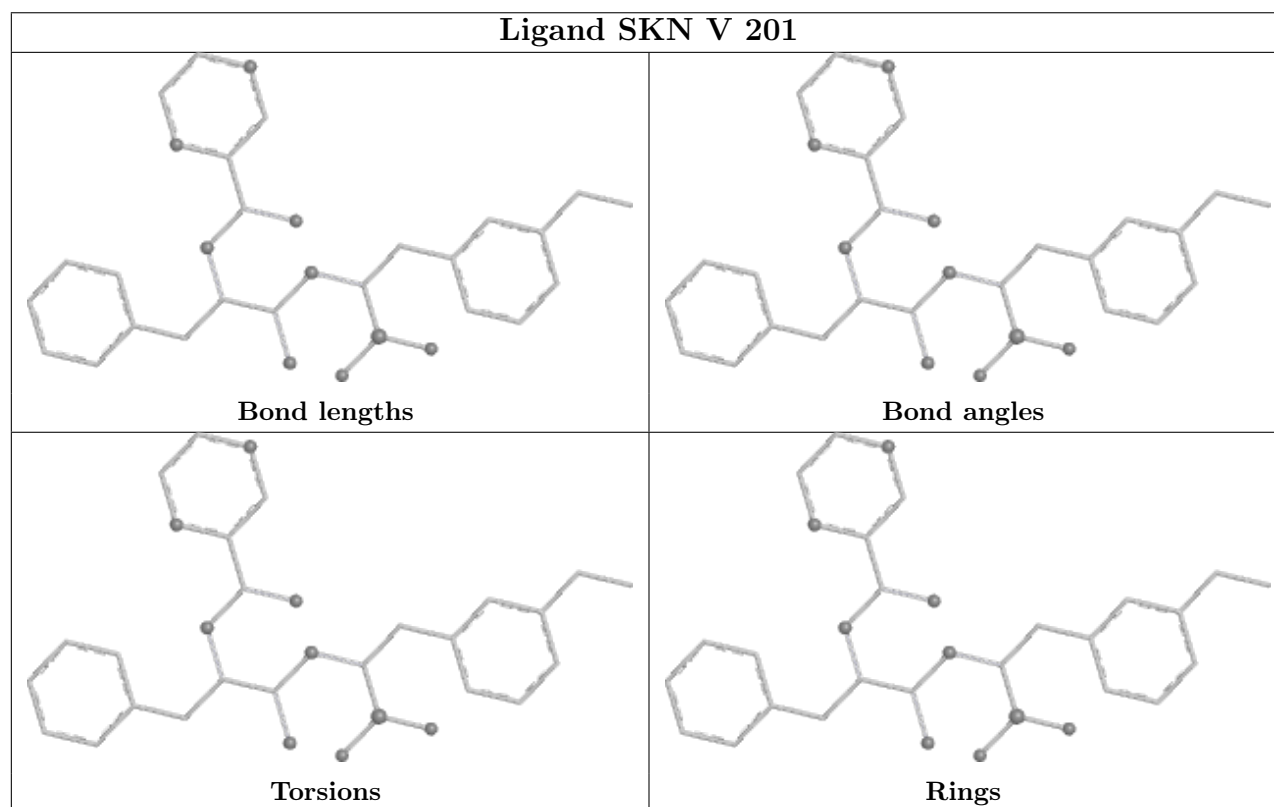
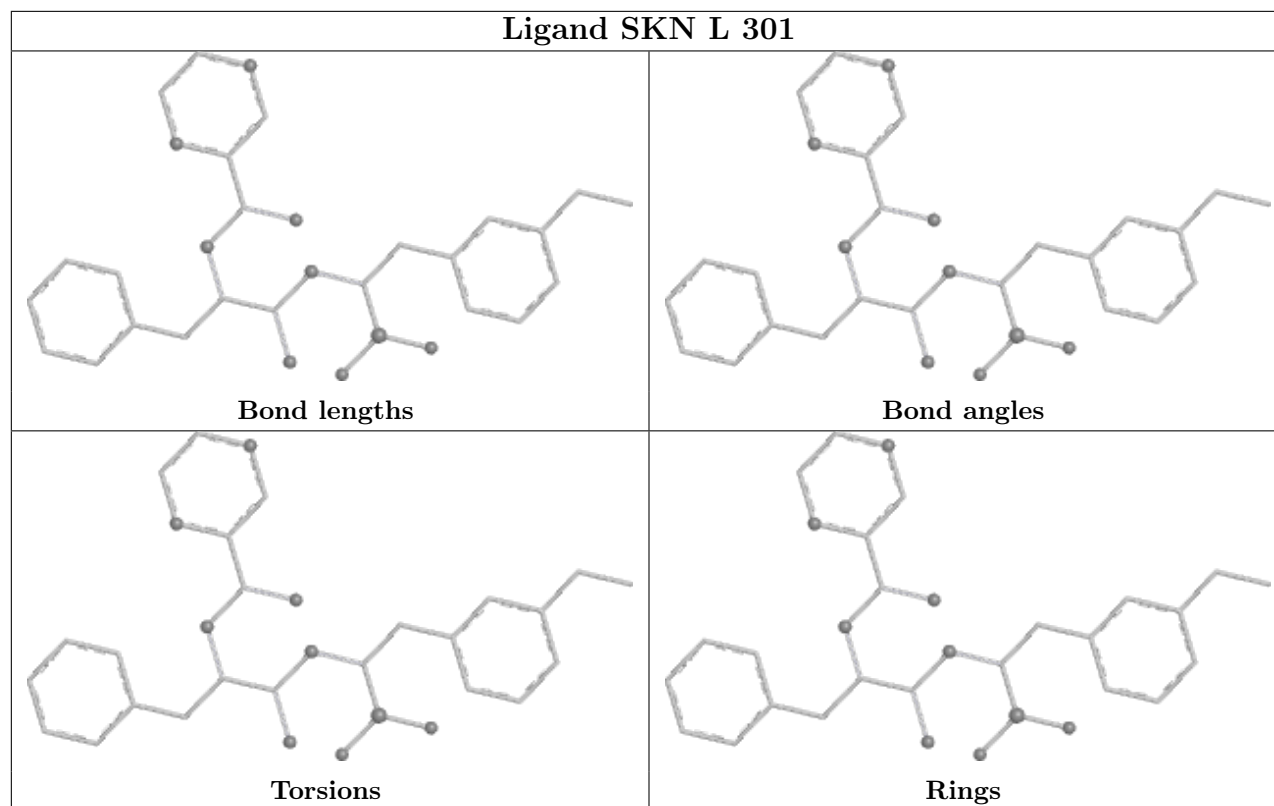
There are no ring outliers.

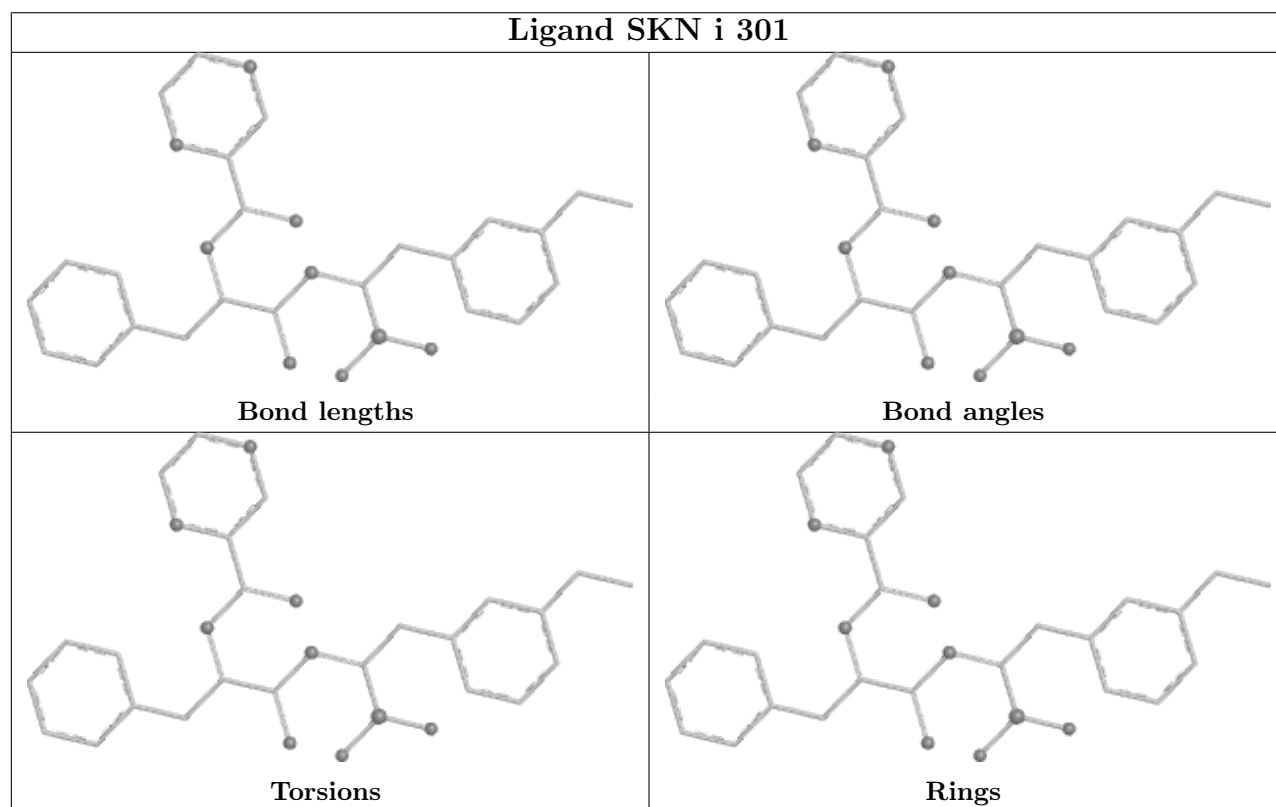
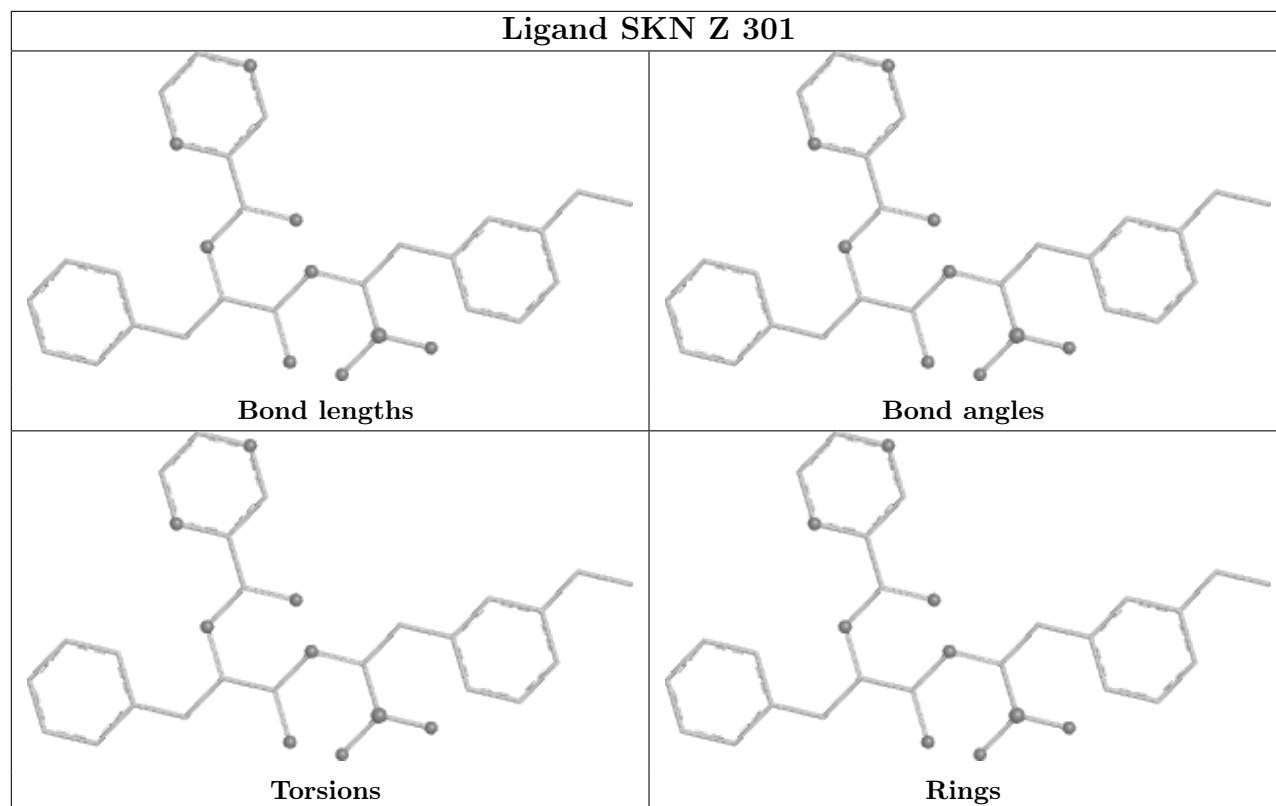
No monomer is involved in short contacts.

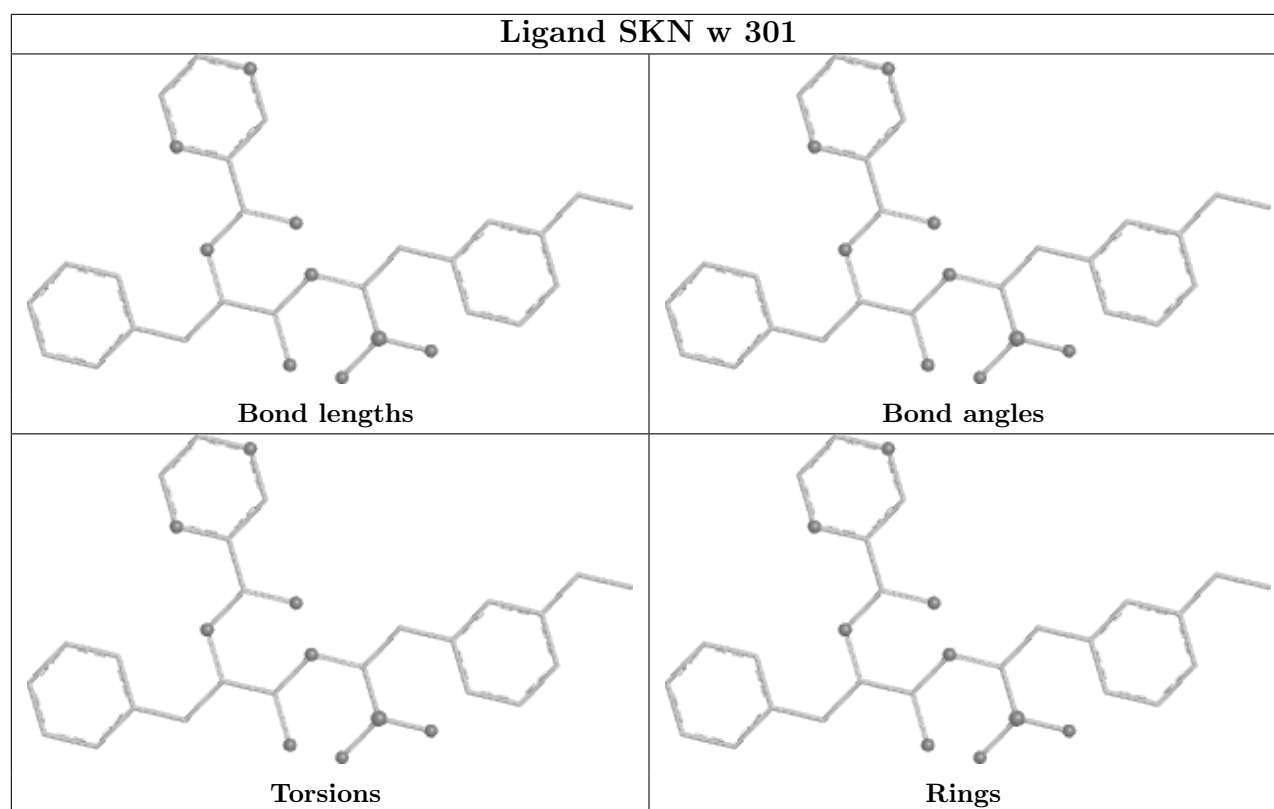
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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.











## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	1	213/213 (100%)	-0.10	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	32, 44, 65, 86	18 (8%)
1	M	213/213 (100%)	-0.08	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	32, 45, 66, 85	16 (7%)
2	2	214/214 (100%)	-0.14	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	27, 43, 65, 85	9 (4%)
2	N	214/214 (100%)	-0.18	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	33, 45, 62, 86	13 (6%)
3	A	240/242 (99%)	-0.04	4 (1%) <span style="border: 1px solid blue; padding: 2px;">70</span> <span style="border: 1px solid blue; padding: 2px;">66</span>	42, 65, 95, 114	58 (24%)
3	O	240/242 (99%)	-0.16	2 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">85</span>	32, 53, 74, 99	44 (18%)
4	B	229/229 (100%)	0.01	5 (2%) <span style="border: 1px solid blue; padding: 2px;">62</span> <span style="border: 1px solid blue; padding: 2px;">58</span>	40, 62, 85, 99	41 (17%)
4	P	229/229 (100%)	-0.10	1 (0%) <span style="border: 1px solid blue; padding: 2px;">92</span> <span style="border: 1px solid blue; padding: 2px;">92</span>	38, 53, 75, 84	33 (14%)
5	C	241/248 (97%)	0.01	3 (1%) <span style="border: 1px solid blue; padding: 2px;">79</span> <span style="border: 1px solid blue; padding: 2px;">77</span>	41, 61, 93, 119	55 (22%)
5	Q	245/248 (98%)	-0.03	2 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">85</span>	39, 57, 80, 106	54 (22%)
6	D	227/235 (96%)	-0.03	5 (2%) <span style="border: 1px solid blue; padding: 2px;">62</span> <span style="border: 1px solid blue; padding: 2px;">58</span>	45, 70, 104, 120	60 (26%)
6	r	229/235 (97%)	0.04	7 (3%) <span style="border: 1px solid blue; padding: 2px;">49</span> <span style="border: 1px solid blue; padding: 2px;">45</span>	40, 74, 126, 189	81 (35%)
7	E	230/232 (99%)	-0.06	2 (0%) <span style="border: 1px solid blue; padding: 2px;">84</span> <span style="border: 1px solid blue; padding: 2px;">83</span>	39, 69, 87, 107	45 (19%)
7	s	230/232 (99%)	-0.10	3 (1%) <span style="border: 1px solid blue; padding: 2px;">77</span> <span style="border: 1px solid blue; padding: 2px;">75</span>	37, 67, 92, 101	58 (25%)
8	F	236/236 (100%)	-0.03	7 (2%) <span style="border: 1px solid blue; padding: 2px;">50</span> <span style="border: 1px solid blue; padding: 2px;">46</span>	45, 63, 98, 116	42 (17%)
8	T	236/236 (100%)	-0.13	3 (1%) <span style="border: 1px solid blue; padding: 2px;">77</span> <span style="border: 1px solid blue; padding: 2px;">75</span>	38, 57, 90, 106	45 (19%)
9	G	242/244 (99%)	-0.08	1 (0%) <span style="border: 1px solid blue; padding: 2px;">92</span> <span style="border: 1px solid blue; padding: 2px;">92</span>	45, 68, 97, 110	53 (21%)
9	U	243/244 (99%)	-0.23	2 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">85</span>	36, 51, 76, 86	51 (20%)
10	H	199/199 (100%)	-0.17	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	32, 47, 63, 67	11 (5%)
10	V	199/199 (100%)	-0.13	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	31, 41, 59, 66	11 (5%)
11	J	204/204 (100%)	-0.11	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">91</span>	34, 47, 71, 83	11 (5%)
11	X	204/204 (100%)	-0.17	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	32, 46, 64, 76	8 (3%)
12	K	196/197 (99%)	-0.11	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">91</span>	35, 49, 65, 74	17 (8%)
12	Y	197/197 (100%)	-0.14	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">91</span>	34, 49, 65, 83	20 (10%)

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	L	203/203 (100%)	-0.12	0 100 100	34, 48, 66, 79	6 (2%)
13	Z	203/203 (100%)	-0.17	1 (0%) 91 91	30, 49, 67, 76	12 (5%)
14	i	219/223 (98%)	-0.08	2 (0%) 84 83	34, 49, 82, 105	25 (11%)
14	w	220/223 (98%)	-0.04	1 (0%) 91 91	33, 46, 79, 104	23 (10%)
All	All	6195/6238 (99%)	-0.09	54 (0%) 84 83	27, 53, 88, 189	920 (14%)

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
14	i	195	PRO	6.2
8	T	236	LEU	5.5
6	r	223	GLU	4.1
6	r	236	LYS	3.8
6	D	181	ILE	3.7

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
15	SO4	1	302	5/5	0.87	0.20	99,99,99,99	0
15	SO4	L	303	5/5	0.92	0.17	101,101,101,101	0
15	SO4	Y	201	5/5	0.92	0.17	104,104,104,104	0
15	SO4	1	301	5/5	0.93	0.15	77,77,78,78	0
16	NA	Z	302	1/1	0.93	0.15	47,47,47,47	0
16	NA	w	302	1/1	0.93	0.15	60,60,60,60	0
15	SO4	K	201	5/5	0.94	0.26	109,109,109,109	0

Continued on next page...

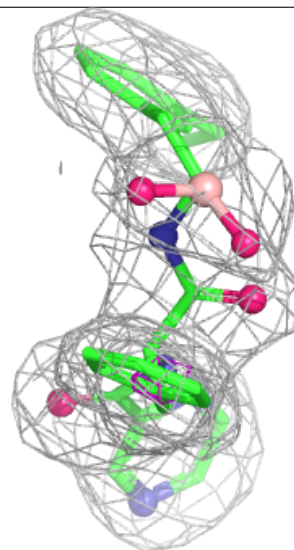
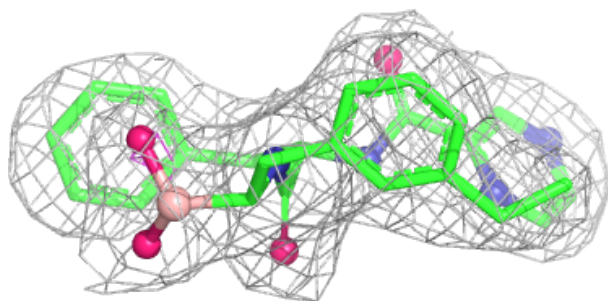
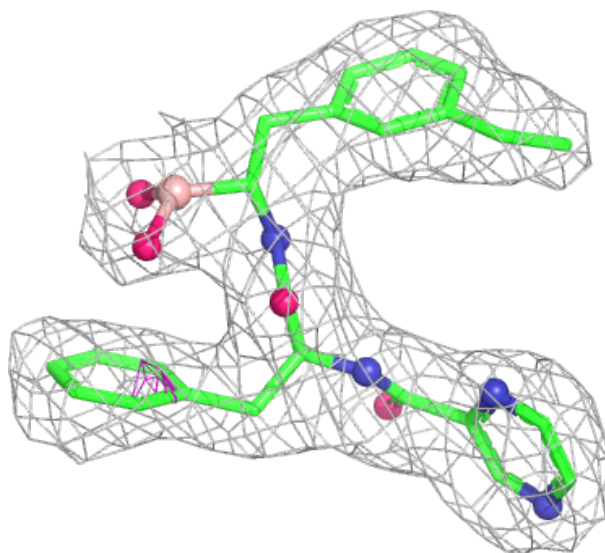
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
17	SKN	H	201	33/33	0.94	0.17	37,39,40,41	0
17	SKN	i	301	33/33	0.94	0.18	47,49,55,55	0
17	SKN	w	301	33/33	0.94	0.21	45,46,54,54	0
17	SKN	Z	301	33/33	0.96	0.17	34,40,44,44	0
16	NA	i	302	1/1	0.96	0.14	47,47,47,47	0
17	SKN	V	201	33/33	0.96	0.19	39,42,46,46	0
16	NA	A	301	1/1	0.97	0.07	42,42,42,42	0
17	SKN	L	301	33/33	0.97	0.15	40,44,46,46	0
16	NA	O	301	1/1	0.99	0.10	33,33,33,33	0
16	NA	L	302	1/1	0.99	0.10	30,30,30,30	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.

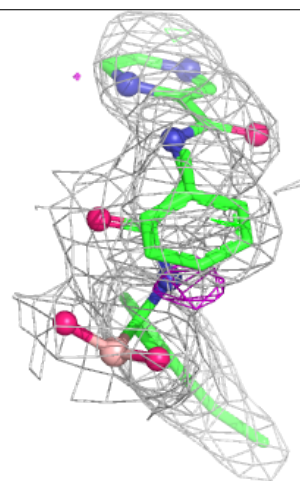
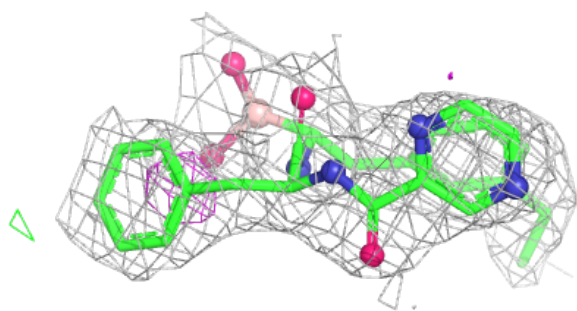
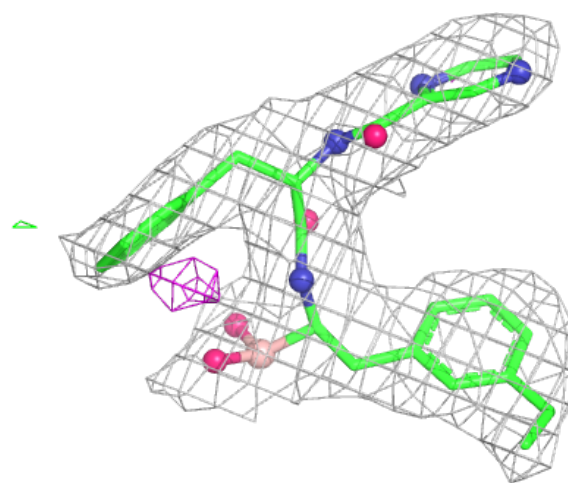
**Electron density around SKN H 201:**

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

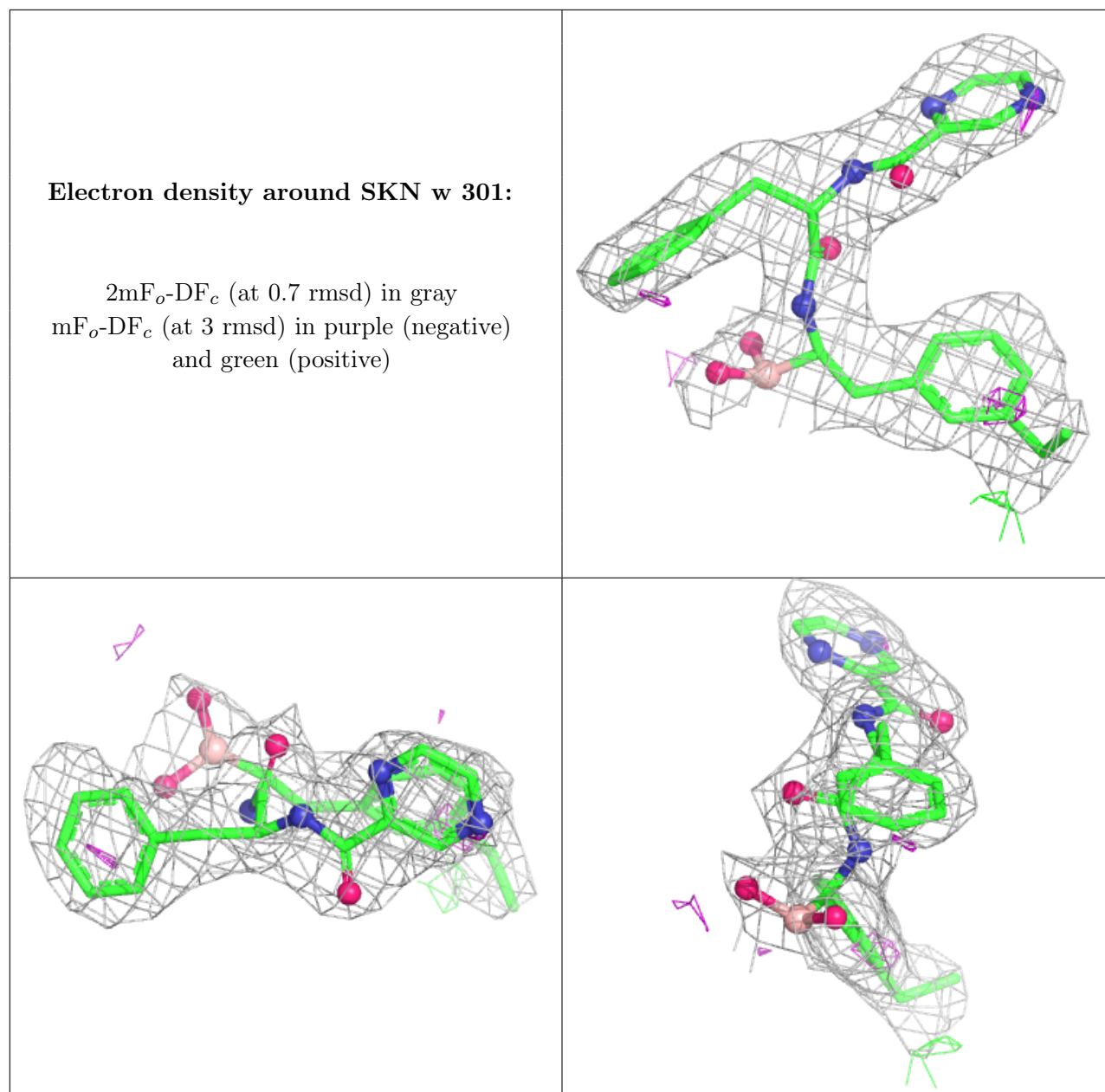


**Electron density around SKN i 301:**

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

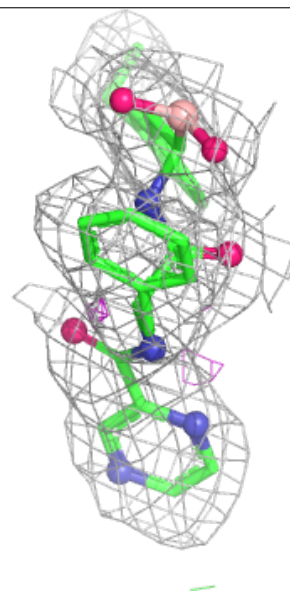
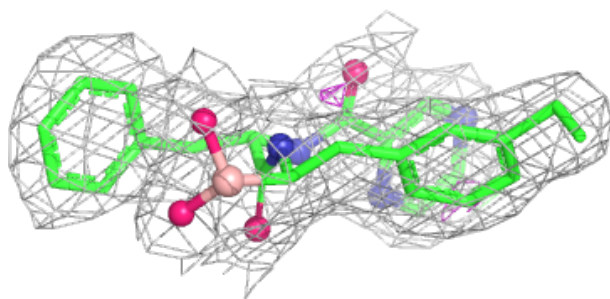
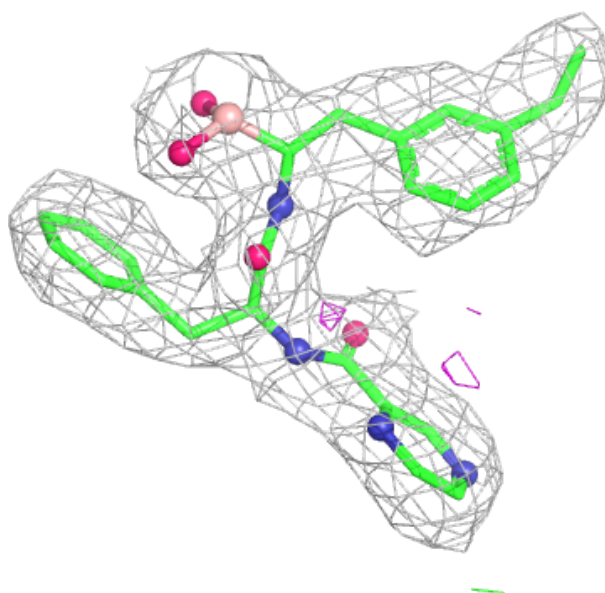


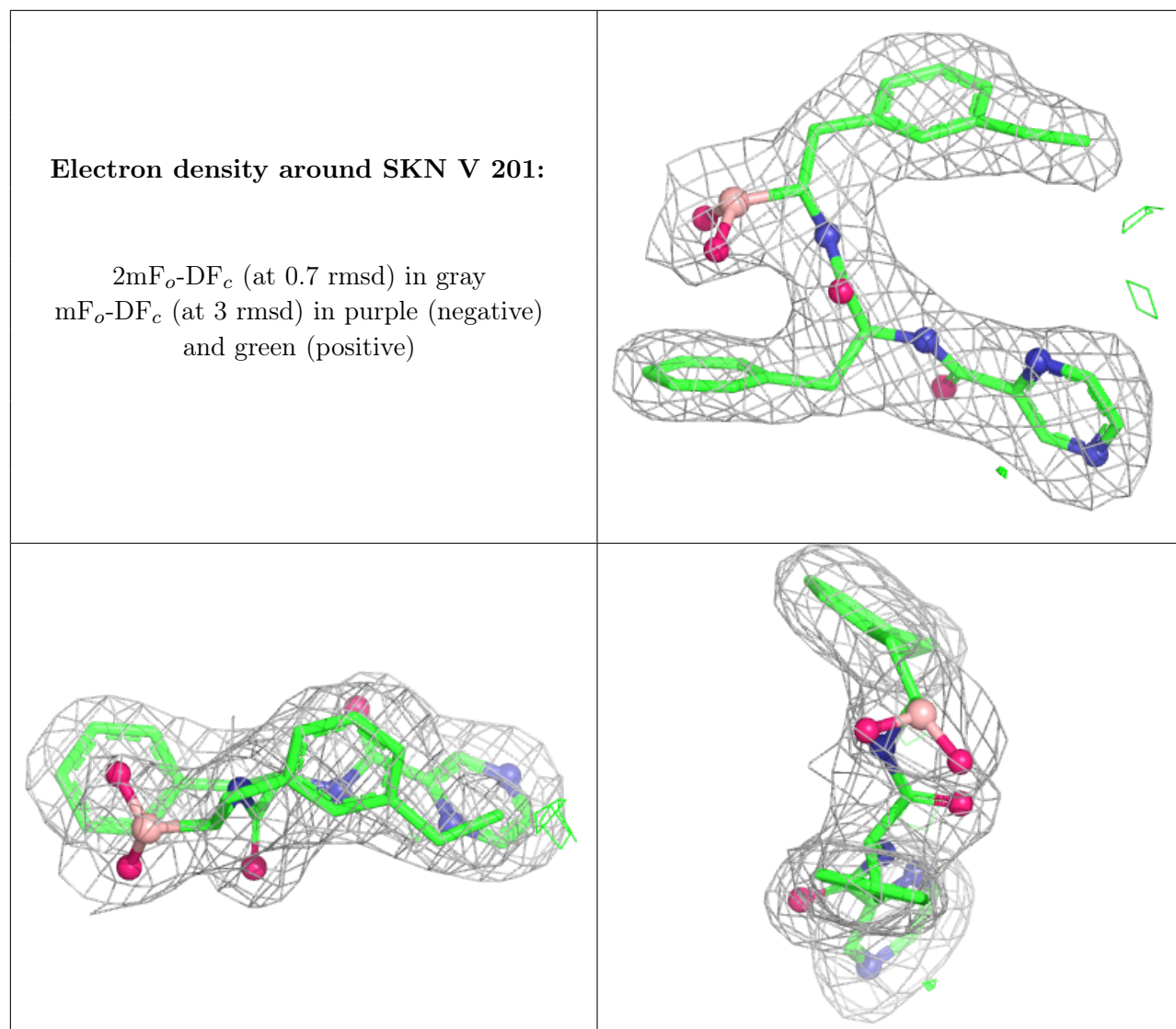


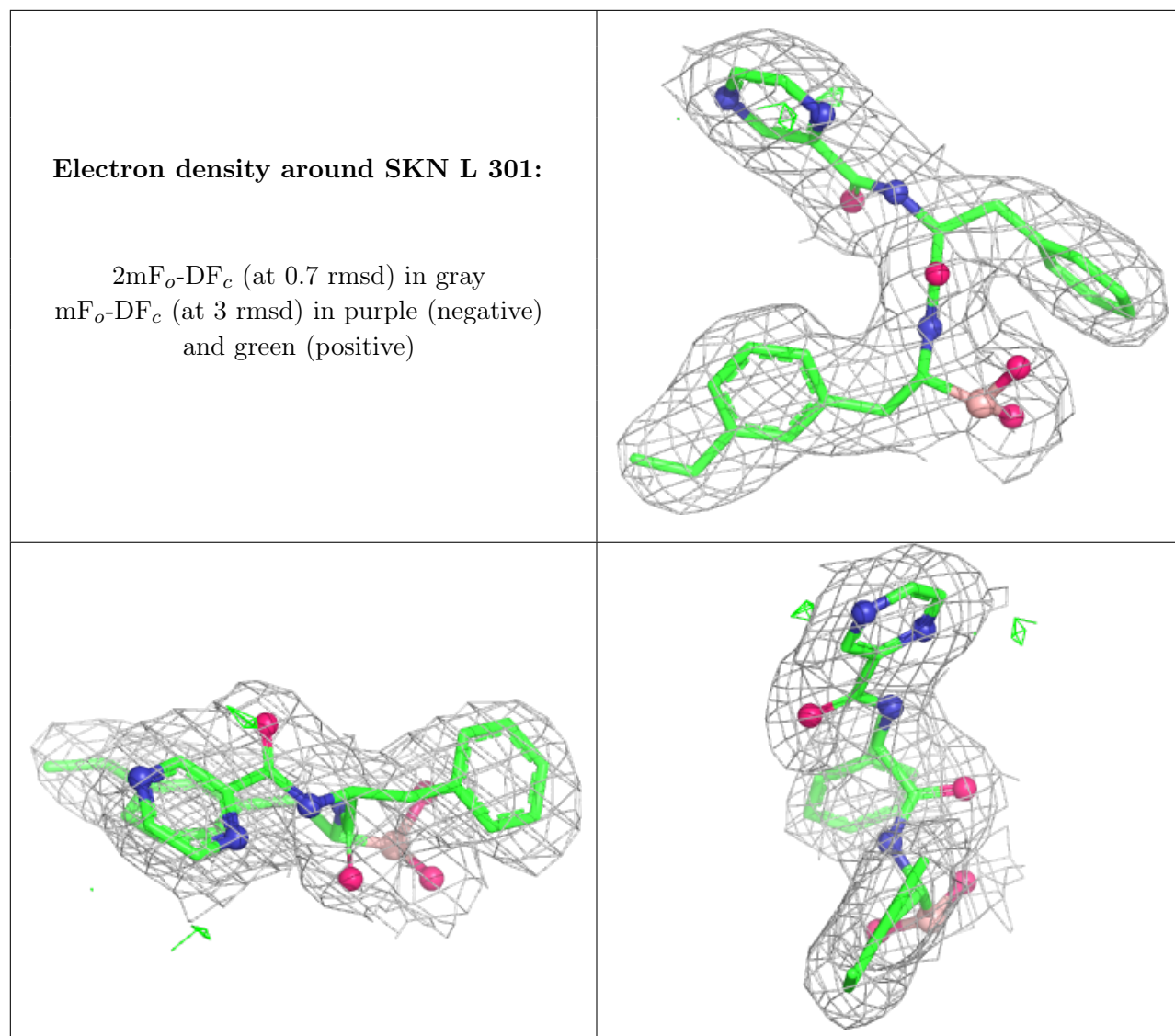


**Electron density around SKN Z 301:**

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







## 6.5 Other polymers [i](#)

There are no such residues in this entry.