



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 4, 2021 – 02:12 pm BST

PDB ID : 7AWE  
Title : HUMAN IMMUNOPROTEASOME 20S PARTICLE IN COMPLEX WITH [(1R)-2-(1-benzofuran-3-yl)-1-{{(1S,2R,4R)-7-oxabicyclo[2.2.1]heptan-2-yl}for mamido}ethyl]boronic acid  
Authors : Musil, D.; Klein, M.  
Deposited on : 2020-11-06  
Resolution : 2.29 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.1  
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) : 2.23.1

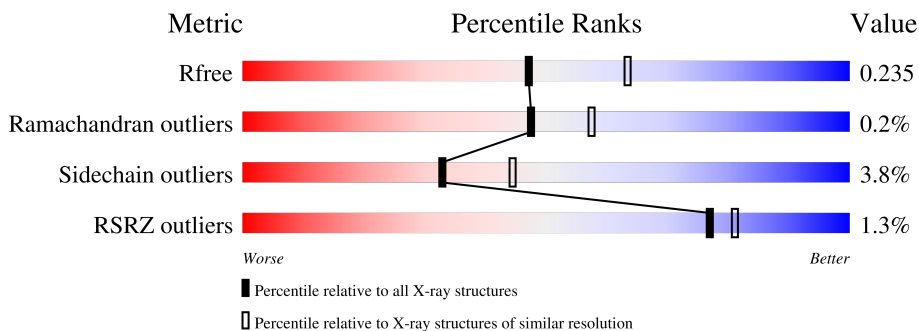
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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	6980 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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	A	242	 2% 97%
1	O	242	 2% 98%
2	B	229	 97%
2	P	229	 97%
3	C	256	 2% 91% 5%
3	Q	256	 2% 95%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	237	92%
4	R	237	4%
5	E	232	2%
5	S	232	3%
6	F	236	%
6	T	236	2%
7	G	242	3%
7	U	242	%
8	H	199	%
8	V	199	%
9	I	223	4%
9	W	223	4%
10	J	204	%
10	X	204	%
11	K	197	%
11	Y	197	%
12	L	203	%
12	Z	203	%
13	M	213	%
13	a	213	%
14	N	214	%
14	b	214	%

## 2 Entry composition [i](#)

There are 18 unique types of molecules in this entry. The entry contains 49839 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-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	240	1871	1189	310	359	13	31	0	0
1	O	242	1886	1197	315	361	13	30	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	229	1787	1144	301	336	6	34	0	0
2	P	229	1787	1144	301	336	6	32	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	243	1912	1210	327	365	10	38	0	0
3	Q	253	2003	1264	345	384	10	4	0	0

- Molecule 4 is a protein called Proteasome subunit alpha type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	227	1798	1132	318	343	5	48	0	0
4	R	234	1848	1160	327	356	5	54	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	230	Total 1758	C 1103	N 291	O 353	S 11	31	0	0
5	S	232	Total 1769	C 1110	N 293	O 355	S 11	58	0	0

- Molecule 6 is a protein called Proteasome subunit alpha type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	236	Total 1857	C 1162	N 334	O 350	S 11	38	0	0
6	T	235	Total 1846	C 1156	N 330	O 349	S 11	12	0	0

- Molecule 7 is a protein called Proteasome subunit alpha type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	242	Total 1894	C 1200	N 323	O 360	S 11	38	0	0
7	U	240	Total 1881	C 1193	N 321	O 356	S 11	25	0	0

- Molecule 8 is a protein called Proteasome subunit beta type-9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	199	Total 1493	C 939	N 254	O 291	S 9	3	0	0
8	V	199	Total 1493	C 939	N 254	O 291	S 9	7	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	223	Total 1641	C 1031	N 290	O 312	S 8	20	0	0
9	W	223	Total 1641	C 1031	N 290	O 312	S 8	16	0	0

- Molecule 10 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	204	Total 1591	C 1013	N 265	O 294	S 19	4	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	204	Total	C	N	O	S	7	0	0
			1591	1013	265	294	19			

- Molecule 11 is a protein called Proteasome subunit beta type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	196	Total	C	N	O	S	7	0	0
			1571	1006	267	289	9			
11	Y	197	Total	C	N	O	S	14	0	0
			1578	1011	268	290	9			

- Molecule 12 is a protein called Proteasome subunit beta type-8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	203	Total	C	N	O	S	4	0	0
			1576	984	276	301	15			
12	Z	203	Total	C	N	O	S	2	0	0
			1576	984	276	301	15			

- Molecule 13 is a protein called Proteasome subunit beta type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	213	Total	C	N	O	S	10	0	0
			1654	1047	284	313	10			
13	a	213	Total	C	N	O	S	16	0	0
			1654	1047	284	313	10			

- Molecule 14 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	214	Total	C	N	O	S	7	0	0
			1673	1055	289	317	12			
14	b	214	Total	C	N	O	S	3	0	0
			1673	1055	289	317	12			

- Molecule 15 is SODIUM ION (three-letter code: NA) (formula: Na).

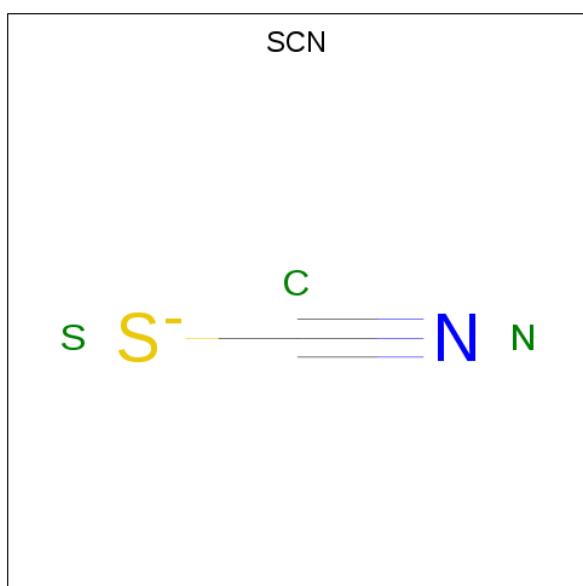
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	1	Total	Na	0	0
			1	1		
15	H	1	Total	Na	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	I	1	Total	Na	0	0
			1	1		
15	J	2	Total	Na	0	0
			2	2		
15	L	1	Total	Na	0	0
			1	1		
15	M	1	Total	Na	0	0
			1	1		
15	O	1	Total	Na	0	0
			1	1		
15	V	1	Total	Na	0	0
			1	1		
15	W	1	Total	Na	0	0
			1	1		
15	X	1	Total	Na	0	0
			1	1		
15	Z	1	Total	Na	0	0
			1	1		
15	a	1	Total	Na	0	0
			1	1		

- Molecule 16 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).



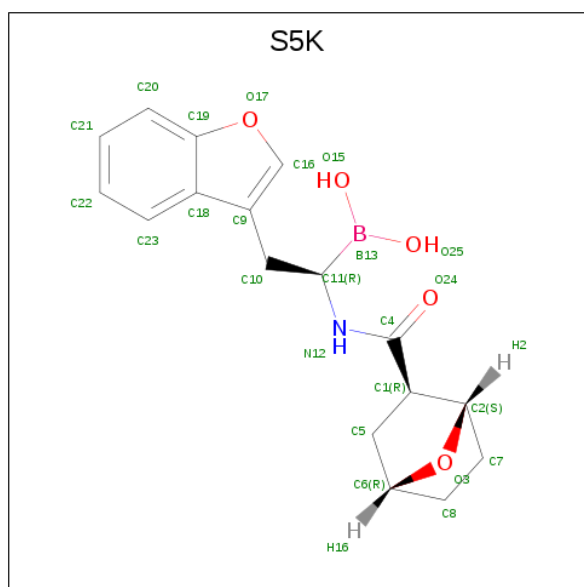
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
16	F	1	Total	C	N	S	0	0
			3	1	1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
16	K	1	Total	C	N	S	0	0
			3	1	1	1		
16	K	1	Total	C	N	S	0	0
			3	1	1	1		
16	M	1	Total	C	N	S	0	0
			3	1	1	1		
16	M	1	Total	C	N	S	0	0
			3	1	1	1		
16	M	1	Total	C	N	S	0	0
			3	1	1	1		
16	T	1	Total	C	N	S	0	0
			3	1	1	1		
16	U	1	Total	C	N	S	0	0
			3	1	1	1		
16	Z	1	Total	C	N	S	0	0
			3	1	1	1		
16	a	1	Total	C	N	S	0	0
			3	1	1	1		
16	a	1	Total	C	N	S	0	0
			3	1	1	1		

- Molecule 17 is [(1R)-2-(1-benzofuran-3-yl)-1-[[[(1S,2R,4R)-7-oxabicyclo[2.2.1]heptan-2-yl]formamido]ethyl]boronic acid (three-letter code: S5K) (formula: C<sub>17</sub>H<sub>20</sub>BNO<sub>5</sub>) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
17	H	1	Total	B	C	N	O	0	0
			24	1	17	1	5		
17	L	1	Total	B	C	N	O	0	0
			24	1	17	1	5		
17	V	1	Total	B	C	N	O	0	0
			24	1	17	1	5		
17	Z	1	Total	B	C	N	O	0	0
			24	1	17	1	5		

- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	44	Total	O	0	0
			44	44		
18	B	35	Total	O	0	0
			35	35		
18	C	60	Total	O	0	0
			60	60		
18	D	27	Total	O	0	0
			27	27		
18	E	38	Total	O	0	0
			38	38		
18	F	37	Total	O	0	0
			37	37		
18	G	34	Total	O	0	0
			34	34		
18	H	52	Total	O	0	0
			52	52		
18	I	52	Total	O	0	0
			52	52		
18	J	69	Total	O	0	0
			69	69		
18	K	48	Total	O	0	0
			48	48		
18	L	63	Total	O	0	0
			63	63		
18	M	80	Total	O	0	0
			80	80		
18	N	75	Total	O	0	0
			75	75		
18	O	35	Total	O	0	0
			35	35		
18	P	34	Total	O	0	0
			34	34		

*Continued on next page...*

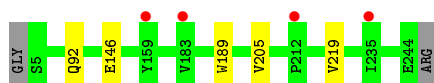
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
18	Q	40	Total 40	O 40	0	0
18	R	29	Total 29	O 29	0	0
18	S	22	Total 22	O 22	0	0
18	T	58	Total 58	O 58	0	0
18	U	43	Total 43	O 43	0	0
18	V	57	Total 57	O 57	0	0
18	W	52	Total 52	O 52	0	0
18	X	68	Total 68	O 68	0	0
18	Y	42	Total 42	O 42	0	0
18	Z	62	Total 62	O 62	0	0
18	a	67	Total 67	O 67	0	0
18	b	72	Total 72	O 72	0	0

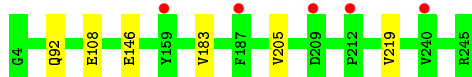
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Proteasome subunit alpha type-6



- Molecule 1: Proteasome subunit alpha type-6



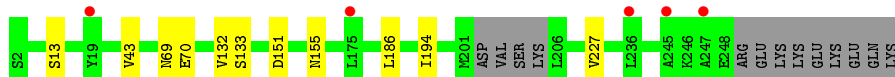
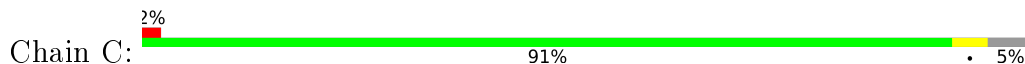
- Molecule 2: Proteasome subunit alpha type-2



- Molecule 2: Proteasome subunit alpha type-2



- Molecule 3: Proteasome subunit alpha type-4



- Molecule 3: Proteasome subunit alpha type-4



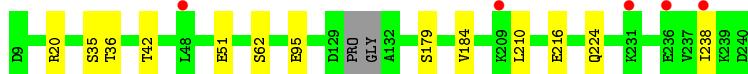
- Molecule 4: Proteasome subunit alpha type-7



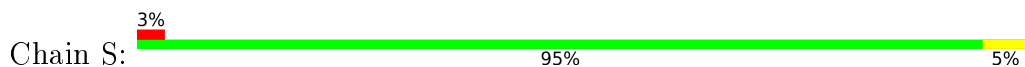
- Molecule 4: Proteasome subunit alpha type-7



- Molecule 5: Proteasome subunit alpha type-5



- Molecule 5: Proteasome subunit alpha type-5



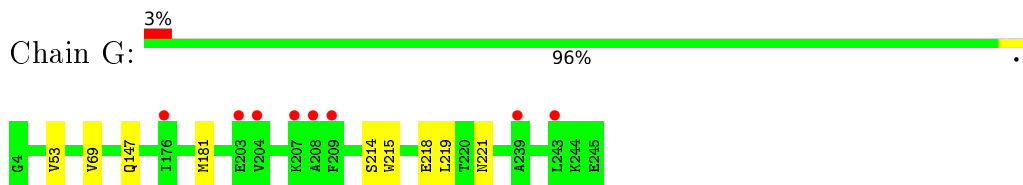
- Molecule 6: Proteasome subunit alpha type-1



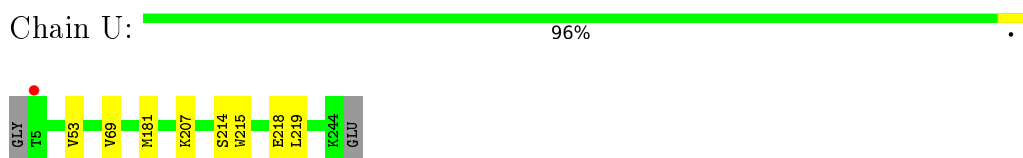
- Molecule 6: Proteasome subunit alpha type-1



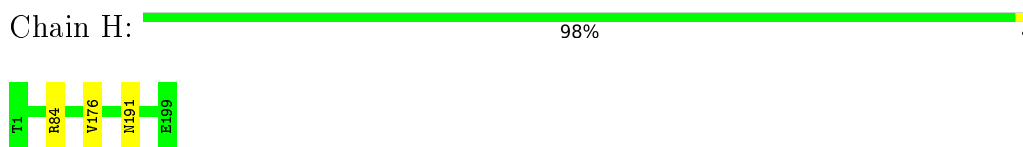
- Molecule 7: Proteasome subunit alpha type-3



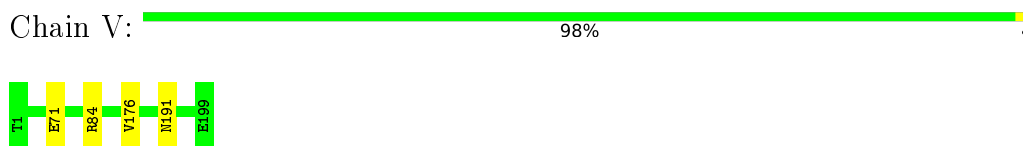
- Molecule 7: Proteasome subunit alpha type-3



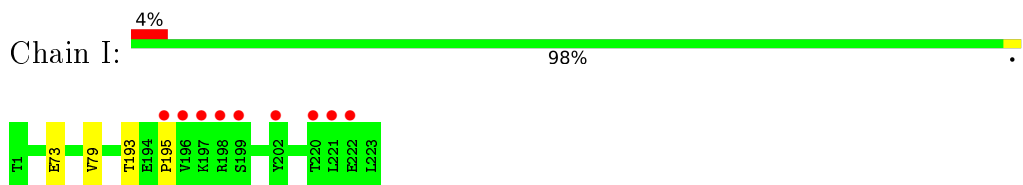
- Molecule 8: Proteasome subunit beta type-9



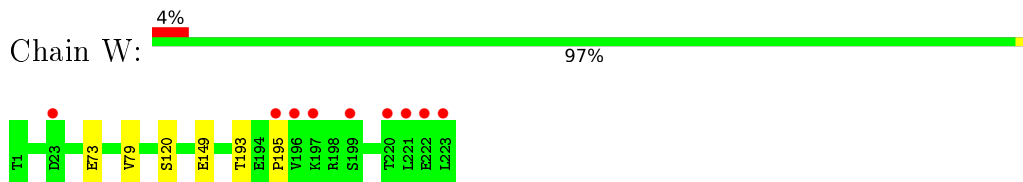
- Molecule 8: Proteasome subunit beta type-9



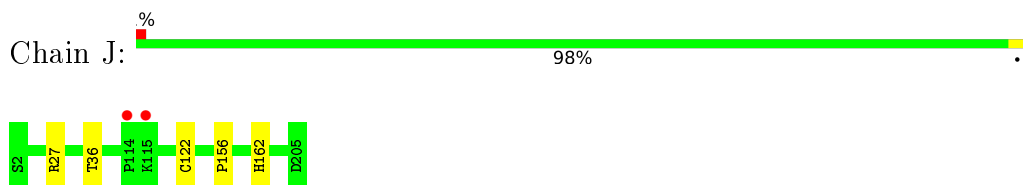
- Molecule 9: Proteasome subunit beta type-10



- Molecule 9: Proteasome subunit beta type-10



- Molecule 10: Proteasome subunit beta type-3



- Molecule 10: Proteasome subunit beta type-3

Chain X:  97%



- Molecule 11: Proteasome subunit beta type-2

Chain K:  98%



- Molecule 11: Proteasome subunit beta type-2

Chain Y:  97%



- Molecule 12: Proteasome subunit beta type-8

Chain L:  96%



- Molecule 12: Proteasome subunit beta type-8

Chain Z:  96%



- Molecule 13: Proteasome subunit beta type-1

Chain M:  96%



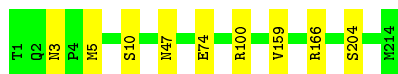
- Molecule 13: Proteasome subunit beta type-1

Chain a:  96%



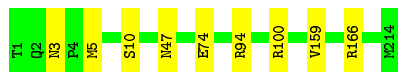
- Molecule 14: Proteasome subunit beta type-4

Chain N:  96%



- Molecule 14: Proteasome subunit beta type-4

Chain b:  96%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.47Å 206.05Å 162.81Å 90.00° 106.69° 90.00°	Depositor
Resolution (Å)	155.95 – 2.29 155.95 – 2.29	Depositor EDS
% Data completeness (in resolution range)	94.3 (155.95-2.29) 62.4 (155.95-2.29)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 2.29Å)	Xtrriage
Refinement program	BUSTER 2.11.7 (18-SEP-2020)	Depositor
R, $R_{free}$	0.178 , 0.214 0.203 , 0.235	Depositor DCC
$R_{free}$ test set	10251 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.7	Xtrriage
Anisotropy	0.051	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 53.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	49839	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.02% 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: S5K, NA, SCN

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	0.38	0/1905	0.60	0/2576
1	O	0.39	0/1920	0.59	0/2595
2	B	0.39	0/1826	0.61	0/2474
2	P	0.40	0/1826	0.59	0/2474
3	C	0.39	0/1941	0.58	0/2614
3	Q	0.39	0/2032	0.56	0/2731
4	D	0.37	0/1822	0.55	0/2458
4	R	0.36	0/1873	0.54	0/2526
5	E	0.40	0/1783	0.57	0/2406
5	S	0.40	0/1796	0.57	0/2426
6	F	0.39	0/1891	0.60	0/2555
6	T	0.41	0/1880	0.60	0/2541
7	G	0.38	0/1929	0.57	0/2597
7	U	0.39	0/1916	0.58	0/2580
8	H	0.42	0/1521	0.60	0/2062
8	V	0.42	0/1521	0.61	0/2062
9	I	0.43	0/1667	0.63	0/2266
9	W	0.43	0/1667	0.63	0/2266
10	J	0.47	0/1620	0.64	0/2184
10	X	0.47	0/1620	0.63	0/2184
11	K	0.44	0/1603	0.63	0/2168
11	Y	0.44	0/1611	0.62	0/2180
12	L	0.45	0/1609	0.65	0/2170
12	Z	0.45	0/1609	0.65	0/2170
13	M	0.48	0/1684	0.67	0/2268
13	a	0.48	0/1684	0.67	0/2268
14	N	0.41	0/1706	0.64	0/2309
14	b	0.42	0/1706	0.64	0/2309
All	All	0.42	0/49168	0.61	0/66419

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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	238/242 (98%)	237 (100%)	0	1 (0%)	34	40
1	O	240/242 (99%)	237 (99%)	3 (1%)	0	100	100
2	B	227/229 (99%)	217 (96%)	10 (4%)	0	100	100
2	P	227/229 (99%)	222 (98%)	5 (2%)	0	100	100
3	C	239/256 (93%)	235 (98%)	4 (2%)	0	100	100
3	Q	249/256 (97%)	244 (98%)	5 (2%)	0	100	100
4	D	221/237 (93%)	217 (98%)	4 (2%)	0	100	100
4	R	230/237 (97%)	226 (98%)	3 (1%)	1 (0%)	34	40
5	E	226/232 (97%)	219 (97%)	7 (3%)	0	100	100
5	S	230/232 (99%)	226 (98%)	4 (2%)	0	100	100
6	F	234/236 (99%)	228 (97%)	5 (2%)	1 (0%)	34	40
6	T	233/236 (99%)	227 (97%)	6 (3%)	0	100	100
7	G	240/242 (99%)	233 (97%)	7 (3%)	0	100	100
7	U	238/242 (98%)	231 (97%)	6 (2%)	1 (0%)	34	40
8	H	197/199 (99%)	191 (97%)	6 (3%)	0	100	100
8	V	197/199 (99%)	191 (97%)	6 (3%)	0	100	100
9	I	221/223 (99%)	214 (97%)	6 (3%)	1 (0%)	29	34

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	W	221/223 (99%)	214 (97%)	6 (3%)	1 (0%)	29	34
10	J	202/204 (99%)	195 (96%)	6 (3%)	1 (0%)	29	34
10	X	202/204 (99%)	196 (97%)	5 (2%)	1 (0%)	29	34
11	K	194/197 (98%)	188 (97%)	5 (3%)	1 (0%)	29	34
11	Y	195/197 (99%)	189 (97%)	5 (3%)	1 (0%)	29	34
12	L	201/203 (99%)	197 (98%)	4 (2%)	0	100	100
12	Z	201/203 (99%)	198 (98%)	3 (2%)	0	100	100
13	M	211/213 (99%)	208 (99%)	3 (1%)	0	100	100
13	a	211/213 (99%)	208 (99%)	3 (1%)	0	100	100
14	N	212/214 (99%)	203 (96%)	9 (4%)	0	100	100
14	b	212/214 (99%)	203 (96%)	9 (4%)	0	100	100
All	All	6149/6254 (98%)	5994 (98%)	145 (2%)	10 (0%)	47	57

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	U	207	LYS
1	A	189	TRP
11	K	174	ASN
4	R	213	ARG
11	Y	174	ASN

### 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	205/206 (100%)	201 (98%)	4 (2%)	55	70
1	O	206/206 (100%)	200 (97%)	6 (3%)	42	56
2	B	188/188 (100%)	180 (96%)	8 (4%)	29	38
2	P	188/188 (100%)	182 (97%)	6 (3%)	39	52
3	C	203/216 (94%)	192 (95%)	11 (5%)	22	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	Q	213/216 (99%)	204 (96%)	9 (4%)	30	39
4	D	193/201 (96%)	184 (95%)	9 (5%)	26	34
4	R	198/201 (98%)	191 (96%)	7 (4%)	36	48
5	E	193/194 (100%)	180 (93%)	13 (7%)	16	19
5	S	194/194 (100%)	183 (94%)	11 (6%)	20	26
6	F	202/202 (100%)	195 (96%)	7 (4%)	36	48
6	T	201/202 (100%)	195 (97%)	6 (3%)	41	54
7	G	199/199 (100%)	190 (96%)	9 (4%)	27	36
7	U	198/199 (100%)	191 (96%)	7 (4%)	36	48
8	H	153/153 (100%)	150 (98%)	3 (2%)	55	70
8	V	153/153 (100%)	149 (97%)	4 (3%)	46	60
9	I	175/175 (100%)	172 (98%)	3 (2%)	60	74
9	W	175/175 (100%)	170 (97%)	5 (3%)	42	56
10	J	173/173 (100%)	169 (98%)	4 (2%)	50	65
10	X	173/173 (100%)	168 (97%)	5 (3%)	42	56
11	K	167/168 (99%)	165 (99%)	2 (1%)	71	82
11	Y	168/168 (100%)	163 (97%)	5 (3%)	41	54
12	L	165/165 (100%)	156 (94%)	9 (6%)	21	27
12	Z	165/165 (100%)	156 (94%)	9 (6%)	21	27
13	M	178/178 (100%)	169 (95%)	9 (5%)	24	31
13	a	178/178 (100%)	169 (95%)	9 (5%)	24	31
14	N	177/177 (100%)	168 (95%)	9 (5%)	24	31
14	b	177/177 (100%)	168 (95%)	9 (5%)	24	31
All	All	5158/5190 (99%)	4960 (96%)	198 (4%)	33	44

5 of 198 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	Q	80	THR
6	T	166	GLN
3	Q	181	GLU
5	S	36	THR
7	U	218	GLU

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

Mol	Chain	Res	Type
6	T	175	HIS
13	a	131	GLN
6	T	185	ASN
12	Z	146	ASN
12	L	146	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 28 ligands modelled in this entry, 13 are monoatomic - leaving 15 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	$\# Z  > 2$
16	SCN	M	302	-	1,2,2	0.12	0	0,1,1	0.00	-
16	SCN	U	301	-	1,2,2	0.08	0	0,1,1	0.00	-
17	S5K	H	901	8	19,27,27	0.86	1 (5%)	23,39,39	1.48	4 (17%)
16	SCN	K	201	-	1,2,2	0.31	0	0,1,1	0.00	-
16	SCN	a	302	-	1,2,2	0.16	0	0,1,1	0.00	-
16	SCN	M	304	-	1,2,2	0.51	0	0,1,1	0.00	-
16	SCN	Z	903	-	1,2,2	0.25	0	0,1,1	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
16	SCN	T	301	-	1,2,2	0.42	0	0,1,1	0.00	-
16	SCN	F	301	-	1,2,2	0.67	0	0,1,1	0.00	-
17	S5K	Z	901	12	19,27,27	0.80	0	23,39,39	1.44	4 (17%)
16	SCN	M	303	-	1,2,2	0.43	0	0,1,1	0.00	-
16	SCN	a	303	-	1,2,2	0.57	0	0,1,1	0.00	-
16	SCN	K	202	-	1,2,2	0.63	0	0,1,1	0.00	-
17	S5K	L	901	12	19,27,27	0.81	1 (5%)	23,39,39	1.46	2 (8%)
17	S5K	V	901	8	19,27,27	0.81	1 (5%)	23,39,39	1.41	3 (13%)

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	S5K	L	901	12	-	6/10/33/33	0/5/4/4
17	S5K	Z	901	12	-	4/10/33/33	0/5/4/4
17	S5K	H	901	8	-	4/10/33/33	0/5/4/4
17	S5K	V	901	8	-	5/10/33/33	0/5/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	H	901	S5K	C18-C19	-2.25	1.38	1.43
17	L	901	S5K	C18-C19	-2.16	1.38	1.43
17	V	901	S5K	C18-C19	-2.08	1.39	1.43

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	L	901	S5K	C2-O3-C6	-5.13	96.89	108.92
17	V	901	S5K	C2-O3-C6	-4.90	97.43	108.92
17	H	901	S5K	C2-O3-C6	-4.65	98.01	108.92
17	Z	901	S5K	C2-O3-C6	-4.63	98.06	108.92
17	L	901	S5K	C23-C18-C19	-2.87	118.56	120.38

There are no chirality outliers.

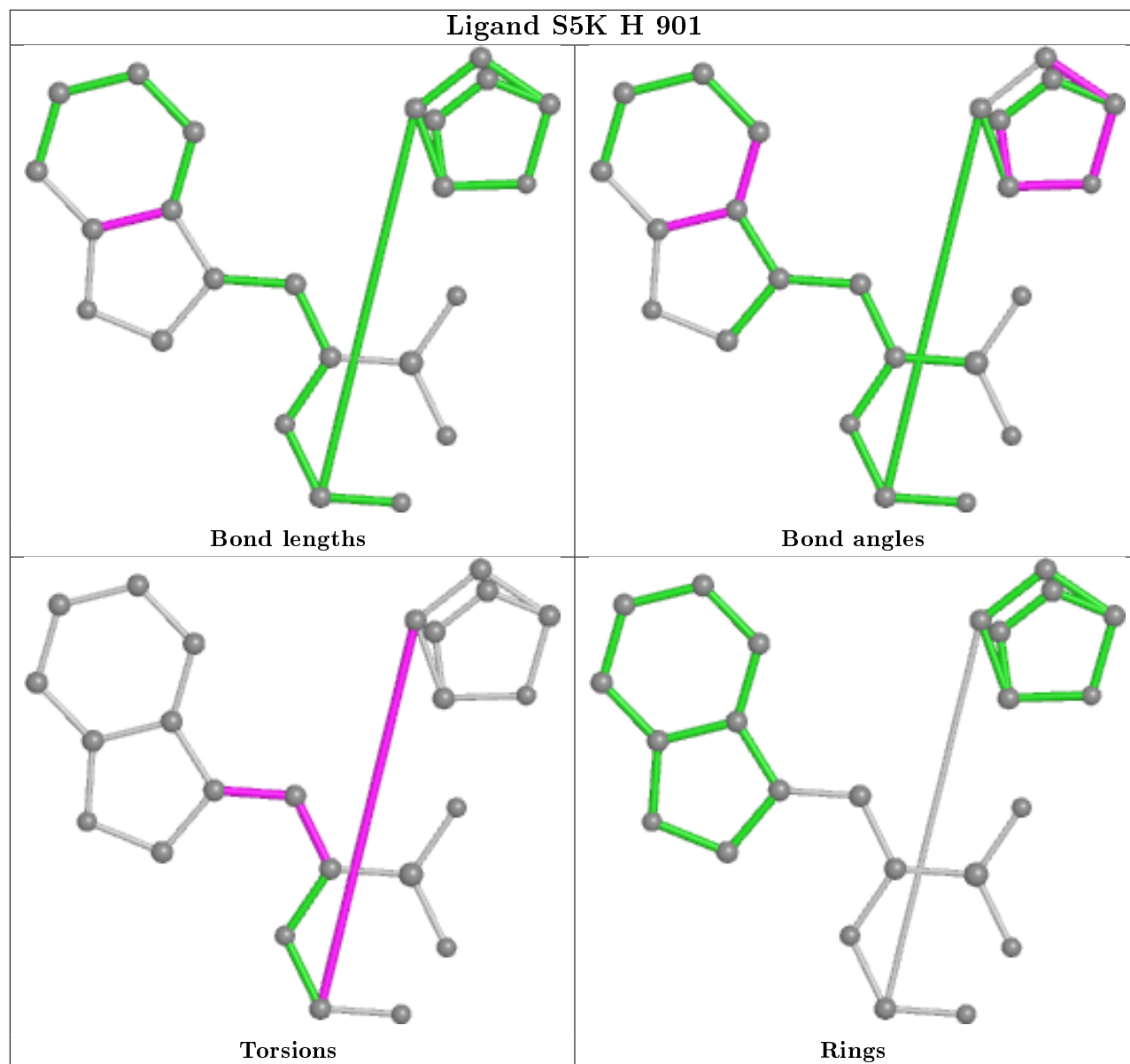
5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	H	901	S5K	C11-C10-C9-C16
17	H	901	S5K	C9-C10-C11-B13
17	L	901	S5K	C9-C10-C11-B13
17	V	901	S5K	C11-C10-C9-C16
17	V	901	S5K	C9-C10-C11-B13

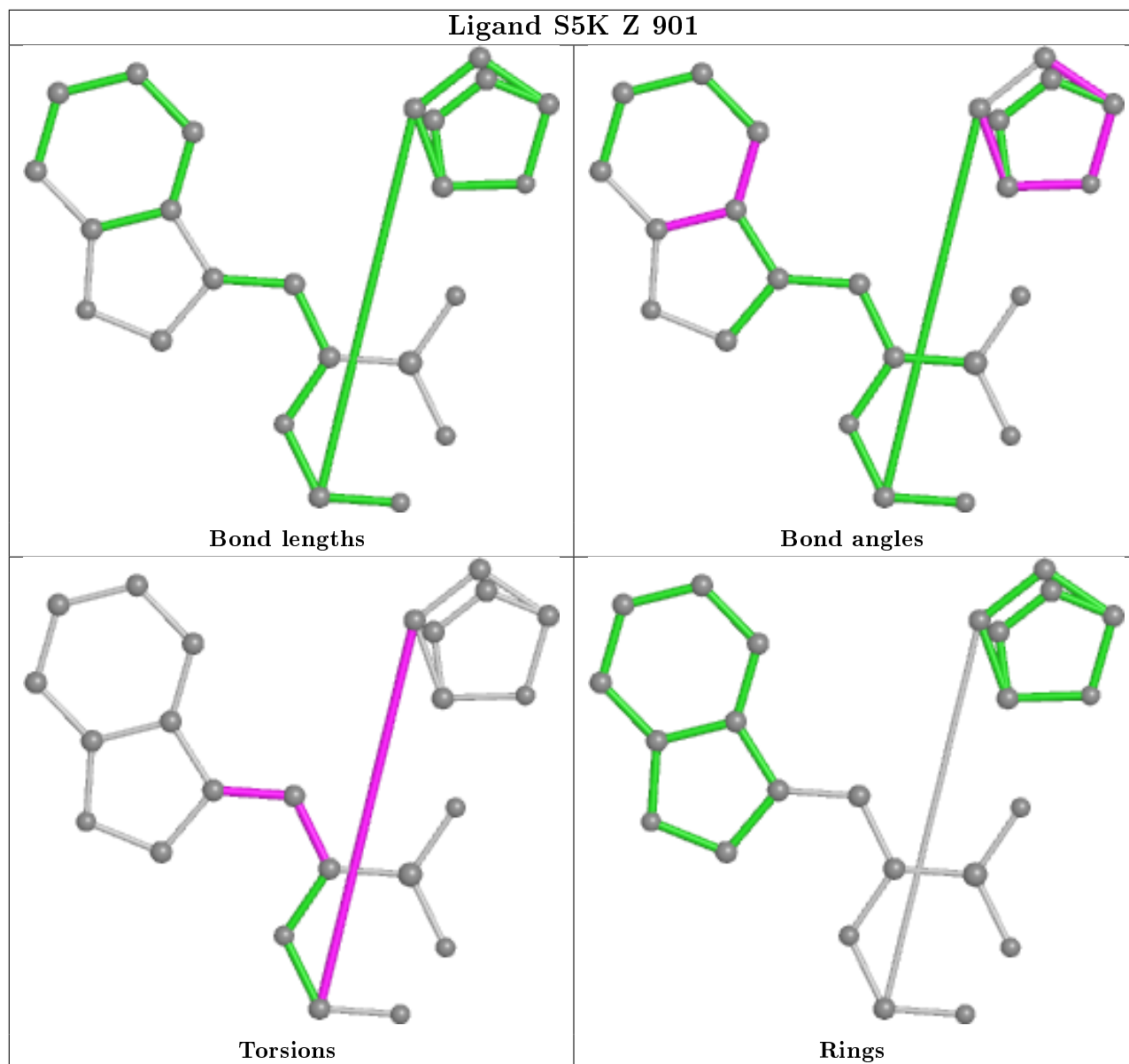
There are no ring outliers.

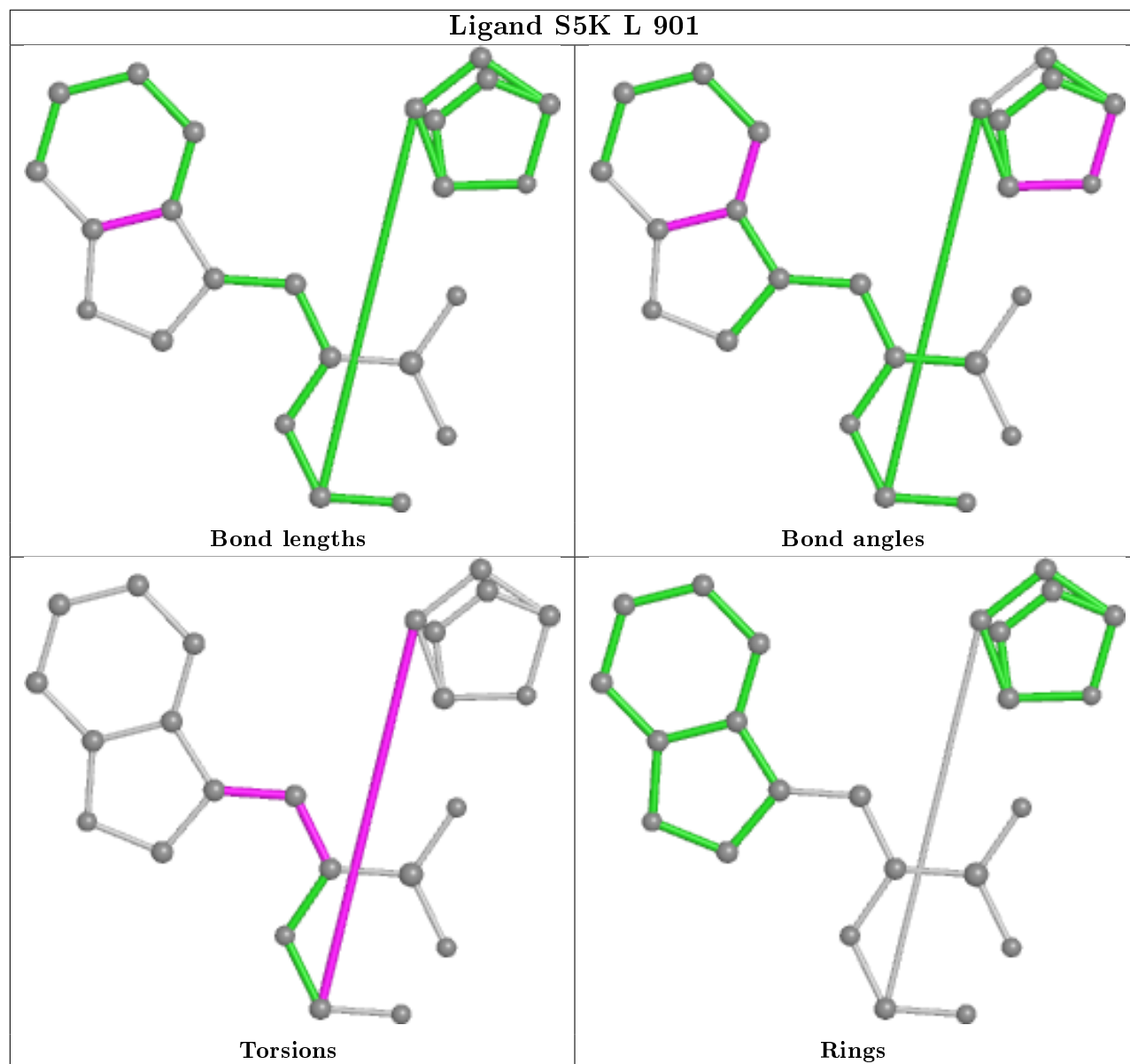
No monomer is involved in short contacts.

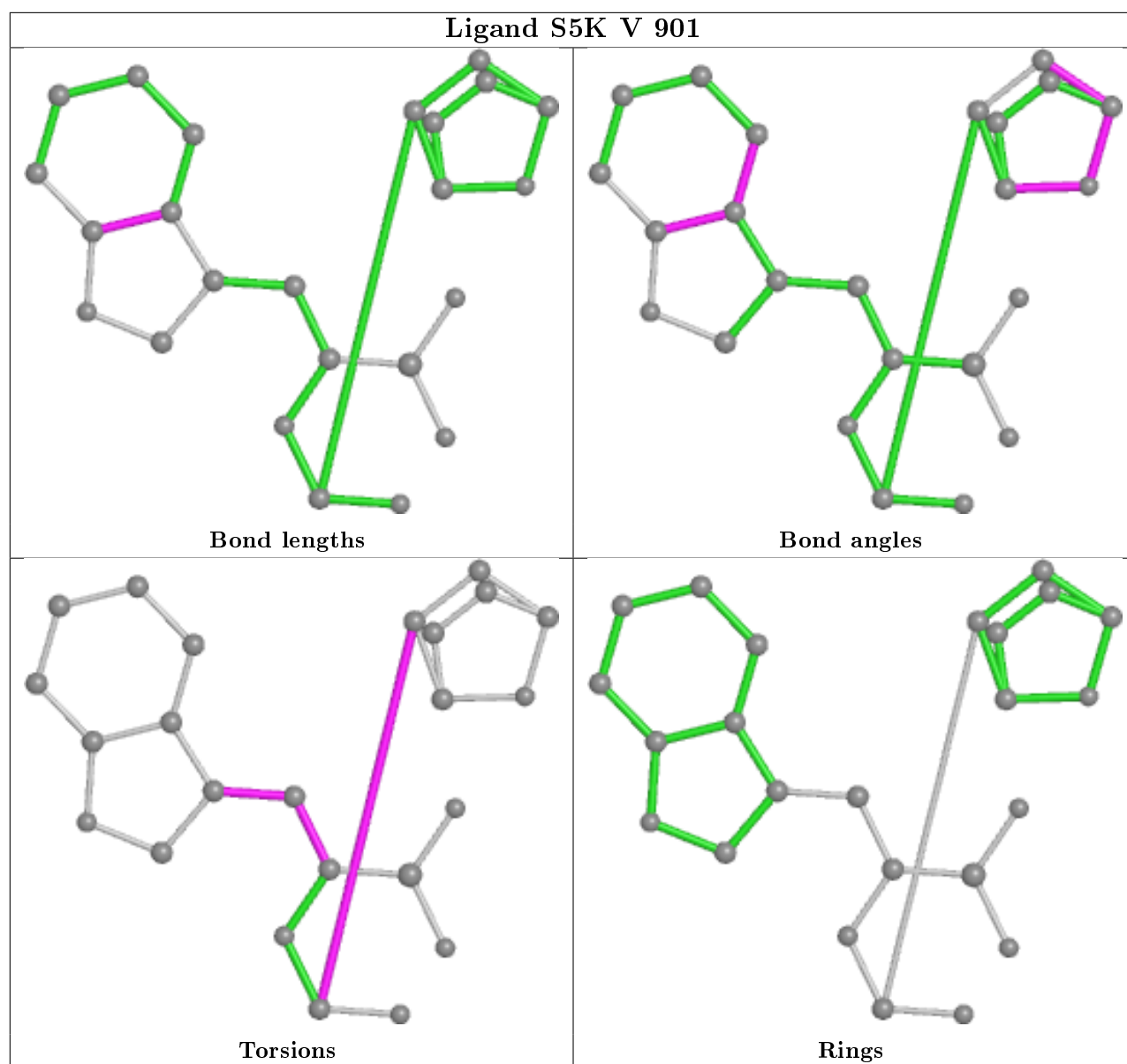
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	A	240/242 (99%)	0.17	4 (1%) 70 75	46, 68, 108, 129	9 (3%)
1	O	242/242 (100%)	0.09	5 (2%) 63 69	40, 62, 92, 121	9 (3%)
2	B	229/229 (100%)	0.04	0 100 100	43, 61, 90, 99	9 (3%)
2	P	229/229 (100%)	-0.10	0 100 100	43, 59, 85, 97	9 (3%)
3	C	243/256 (94%)	0.10	5 (2%) 63 69	40, 64, 95, 152	13 (5%)
3	Q	253/256 (98%)	0.19	5 (1%) 65 70	40, 65, 134, 192	2 (0%)
4	D	227/237 (95%)	0.16	1 (0%) 92 94	45, 71, 116, 154	13 (5%)
4	R	234/237 (98%)	0.19	9 (3%) 40 45	48, 71, 108, 133	16 (6%)
5	E	230/232 (99%)	0.22	5 (2%) 62 68	38, 69, 98, 129	10 (4%)
5	S	232/232 (100%)	0.19	7 (3%) 50 56	38, 71, 96, 122	17 (7%)
6	F	236/236 (100%)	0.18	2 (0%) 86 89	43, 64, 97, 131	11 (4%)
6	T	235/236 (99%)	0.09	4 (1%) 70 75	43, 61, 95, 119	4 (1%)
7	G	242/242 (100%)	0.32	8 (3%) 46 52	47, 69, 104, 138	12 (4%)
7	U	240/242 (99%)	0.14	1 (0%) 92 94	41, 61, 90, 104	8 (3%)
8	H	199/199 (100%)	-0.04	0 100 100	38, 50, 67, 73	1 (0%)
8	V	199/199 (100%)	-0.10	0 100 100	35, 47, 66, 76	3 (1%)
9	I	223/223 (100%)	0.09	9 (4%) 38 43	36, 52, 101, 125	6 (2%)
9	W	223/223 (100%)	0.09	9 (4%) 38 43	37, 50, 101, 120	4 (1%)
10	J	204/204 (100%)	-0.13	2 (0%) 82 86	32, 45, 66, 86	2 (0%)
10	X	204/204 (100%)	-0.17	0 100 100	34, 45, 65, 78	3 (1%)
11	K	196/197 (99%)	-0.03	0 100 100	36, 50, 64, 82	2 (1%)
11	Y	197/197 (100%)	-0.11	1 (0%) 91 93	36, 50, 70, 91	3 (1%)
12	L	203/203 (100%)	-0.16	0 100 100	32, 47, 67, 81	1 (0%)
12	Z	203/203 (100%)	-0.20	0 100 100	33, 48, 67, 78	1 (0%)

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	M	213/213 (100%)	-0.12	1 (0%) 91 93	30, 45, 66, 89	3 (1%)
13	a	213/213 (100%)	-0.15	0 100 100	30, 44, 66, 89	5 (2%)
14	N	214/214 (100%)	-0.15	0 100 100	31, 48, 68, 85	2 (0%)
14	b	214/214 (100%)	-0.15	0 100 100	30, 46, 68, 85	1 (0%)
All	All	6217/6254 (99%)	0.03	78 (1%) 77 81	30, 56, 95, 192	179 (2%)

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	254	LYS	5.8
9	W	220	THR	5.6
9	I	221	LEU	5.6
5	S	130	PRO	5.3
9	I	196	VAL	5.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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
17	S5K	V	901	24/24	0.78	0.37	62,70,72,72	0
17	S5K	H	901	24/24	0.81	0.32	61,71,74,74	0
16	SCN	M	303	3/3	0.86	0.18	60,60,61,62	0
16	SCN	Z	903	3/3	0.90	0.10	71,71,72,72	0
16	SCN	F	301	3/3	0.91	0.10	65,65,66,67	0
16	SCN	a	303	3/3	0.91	0.18	64,64,65,66	0
15	NA	X	301	1/1	0.92	0.09	43,43,43,43	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
15	NA	a	301	1/1	0.92	0.11	61,61,61,61	0
15	NA	J	301	1/1	0.92	0.07	44,44,44,44	0
15	NA	J	302	1/1	0.92	0.12	55,55,55,55	0
15	NA	M	301	1/1	0.93	0.14	51,51,51,51	0
16	SCN	K	202	3/3	0.93	0.15	69,69,69,70	0
16	SCN	a	302	3/3	0.93	0.18	69,69,69,69	0
16	SCN	T	301	3/3	0.95	0.09	78,78,79,79	0
15	NA	W	301	1/1	0.95	0.07	40,40,40,40	0
15	NA	H	902	1/1	0.95	0.16	51,51,51,51	0
15	NA	A	301	1/1	0.96	0.05	48,48,48,48	0
15	NA	Z	902	1/1	0.96	0.09	38,38,38,38	0
16	SCN	M	304	3/3	0.96	0.09	68,68,68,69	0
16	SCN	K	201	3/3	0.96	0.10	70,70,70,70	0
16	SCN	U	301	3/3	0.96	0.11	69,69,69,69	0
15	NA	V	902	1/1	0.97	0.12	52,52,52,52	0
15	NA	O	301	1/1	0.97	0.09	42,42,42,42	0
16	SCN	M	302	3/3	0.97	0.12	68,68,68,68	0
17	S5K	Z	901	24/24	0.97	0.13	30,35,38,38	0
17	S5K	L	901	24/24	0.98	0.13	31,34,36,37	0
15	NA	L	902	1/1	0.98	0.08	34,34,34,34	0
15	NA	I	301	1/1	0.98	0.08	39,39,39,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.