



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 10, 2024 – 02:17 am BST

PDB ID : 9FT1  
Title : Yeast 20S proteasome in complex with epoxyketone inhibitor 13  
Authors : Maurits, E.; Huber, E.M.; Dekker, P.M.; Wang, X.; Heinemeyer, W.; Florea, B.I.; Groll, M.; Overkleeft, H.S.  
Deposited on : 2024-06-23  
Resolution : 2.60 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : **FAILED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.37.1

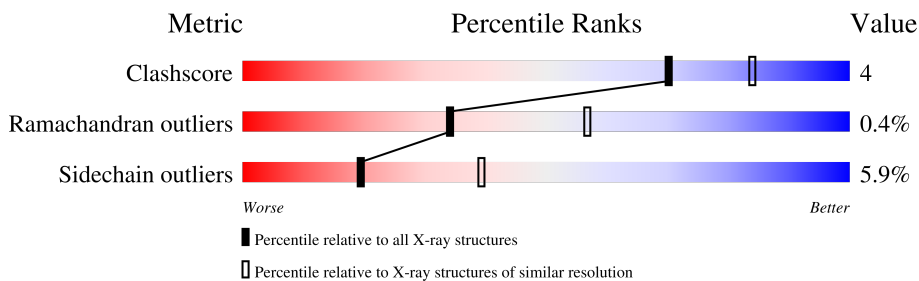
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)











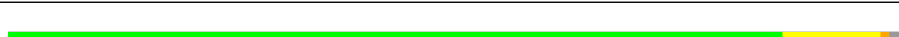


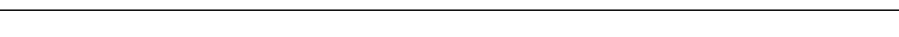
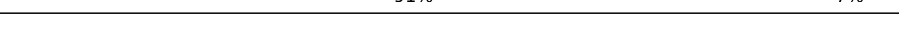
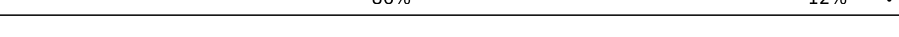

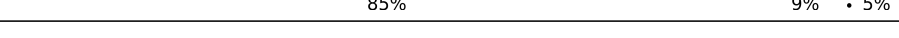

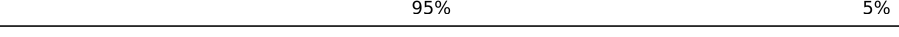
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria 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  $\leq 5\%$ .

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	250	
1	O	250	
2	B	258	
2	P	258	
3	C	254	
3	Q	254	
4	D	260	
4	R	260	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	234	 79% 17% ..
5	S	234	 80% 17% ..
6	F	288	 75% 8% • 16%
6	T	288	 75% 8% • 16%
7	G	252	 83% 12% ..
7	U	252	 83% 12% ..
8	H	231	 86% 10% •
8	V	231	 86% 10% •
9	I	205	 86% 13%
9	W	205	 85% 14%
10	J	198	 87% 11% ..
10	X	198	 87% 10% ..
11	K	211	 91% 7% •
11	Y	211	 91% 7% •
12	L	222	 86% 12% •
12	Z	222	 87% 12% •
13	M	246	 85% 9% • 5%
13	a	246	 89% 5% 5%
14	N	196	 95% 5%
14	b	196	 98% •

## 2 Entry composition

There are 18 unique types of molecules in this entry. The entry contains 49970 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Proteasome subunit alpha type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	Total 1915	C 1219	N 315	O 377	S 4	0	0	0
1	O	250	Total 1915	C 1219	N 315	O 377	S 4	0	0	0

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

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

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

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

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

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

- 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	221	Total	C	N	O	S	0	0	0
			1677	1057	292	321	7			
8	V	221	Total	C	N	O	S	0	0	0
			1677	1057	292	321	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-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	211	Total	C	N	O	S	0	0	0
			1637	1041	279	310	7			
11	Y	211	Total	C	N	O	S	0	0	0
			1637	1041	279	310	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			
12	Z	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	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).

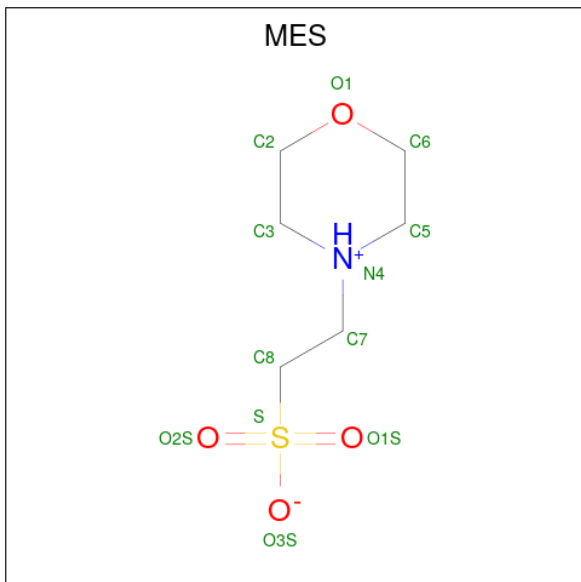
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	G	1	Total	Mg	0	0
			1	1		
15	H	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	I	1	Total	Mg	0	0
			1	1		
15	K	1	Total	Mg	0	0
			1	1		
15	N	1	Total	Mg	0	0
			1	1		
15	Y	1	Total	Mg	0	0
			1	1		
15	Z	1	Total	Mg	0	0
			1	1		

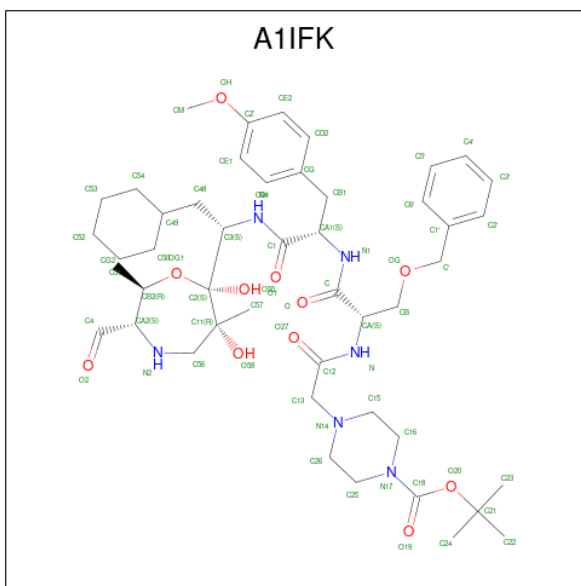
- Molecule 16 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	H	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
16	J	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
16	V	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
16	X	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
16	b	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 17 is tert-butyl 4-[2-[(2S)-1-[(2S)-1-[(1S)-2-cyclohexyl-1-[(2R,3S,6R,7S)-3-met

hanoyl-2,6-dimethyl-6,7-bis(oxidanyl)-1,4-oxazepan-7-yl]ethyl]amino]-3-(4-methoxyphenyl)-1-oxidanylidene-propan-2-yl]amino]-1-oxidanylidene-3-phenylmethoxy-propan-2-yl]amino]-2-oxidanylidene-ethyl]piperazine-1-carboxylate (three-letter code: A1IFK) (formula:  $C_{47}H_{70}N_6O_{11}$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
17	H	1	Total	C	N	O	0	0
			64	47	6	11		
17	K	1	Total	C	N	O	0	0
			64	47	6	11		
17	V	1	Total	C	N	O	0	0
			64	47	6	11		
17	Y	1	Total	C	N	O	0	0
			64	47	6	11		

- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	14	Total	O	0	0
			14	14		
18	B	14	Total	O	0	0
			14	14		
18	C	11	Total	O	0	0
			11	11		
18	D	11	Total	O	0	0
			11	11		
18	E	5	Total	O	0	0
			5	5		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	F	11	Total O 11 11	0	0
18	G	13	Total O 13 13	0	0
18	H	18	Total O 18 18	0	0
18	I	17	Total O 17 17	0	0
18	J	19	Total O 19 19	0	0
18	K	18	Total O 18 18	0	0
18	L	22	Total O 22 22	0	0
18	M	16	Total O 16 16	0	0
18	N	14	Total O 14 14	0	0
18	O	15	Total O 15 15	0	0
18	P	11	Total O 11 11	0	0
18	Q	8	Total O 8 8	0	0
18	R	7	Total O 7 7	0	0
18	S	5	Total O 5 5	0	0
18	T	7	Total O 7 7	0	0
18	U	20	Total O 20 20	0	0
18	V	10	Total O 10 10	0	0
18	W	16	Total O 16 16	0	0
18	X	15	Total O 15 15	0	0
18	Y	15	Total O 15 15	0	0
18	Z	14	Total O 14 14	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	a	15	Total	O	0	0
			15	15		
18	b	18	Total	O	0	0
			18	18		

### 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. 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. 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.

Note EDS failed to run properly.

- Molecule 1: Proteasome subunit alpha type-2

Chain A: 




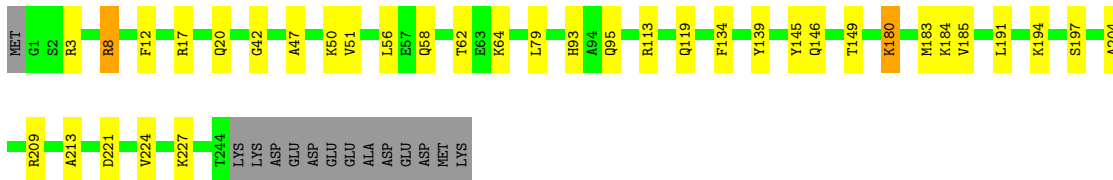
- Molecule 1: Proteasome subunit alpha type-2

Chain O: 




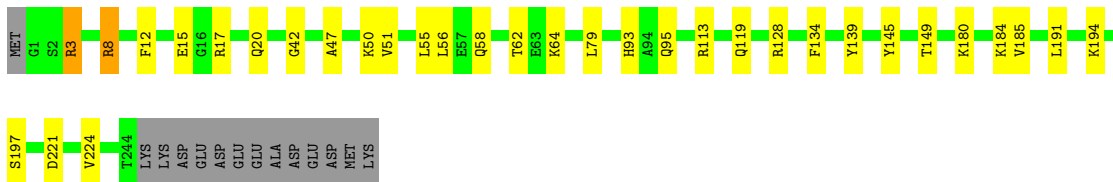
- Molecule 2: Proteasome subunit alpha type-3

Chain B: 




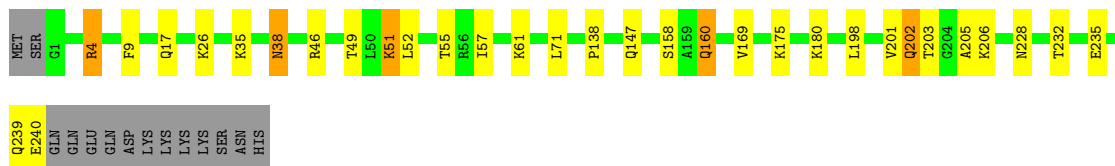
- Molecule 2: Proteasome subunit alpha type-3

Chain P: 



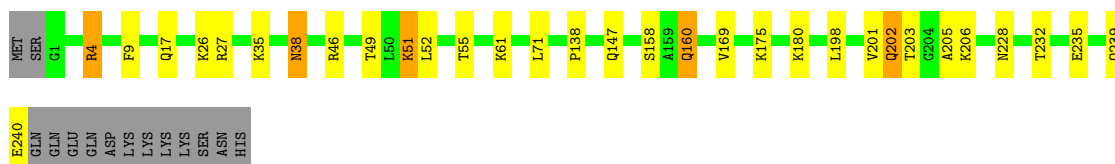
- Molecule 3: Proteasome subunit alpha type-4

Chain C: 



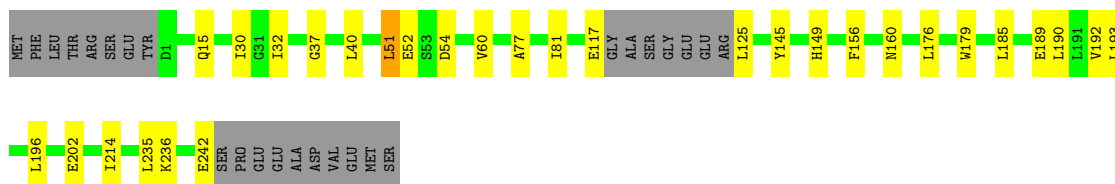
- Molecule 3: Proteasome subunit alpha type-4

Chain Q: 82% 11% 6%



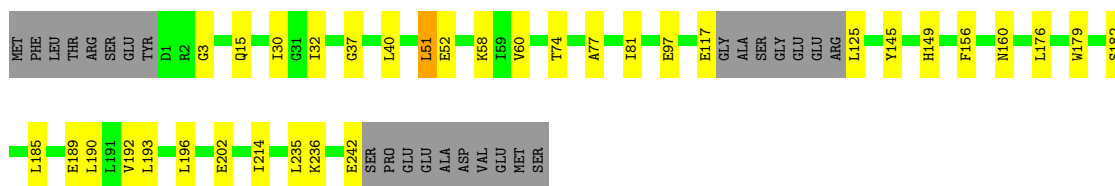
- Molecule 4: Proteasome subunit alpha type-5

Chain D: 79% 11% 10%



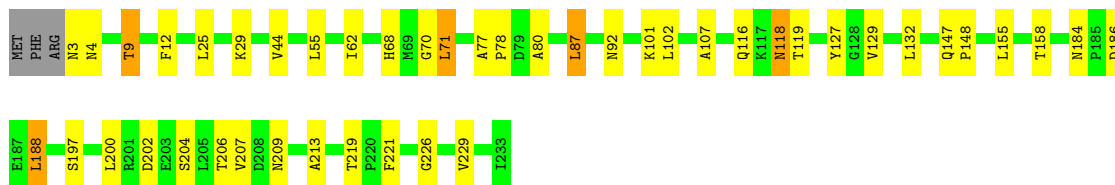
- Molecule 4: Proteasome subunit alpha type-5

Chain R: 77% 13% 10%



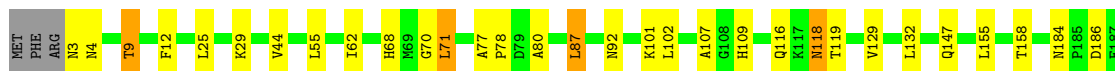
- Molecule 5: Proteasome subunit alpha type-6

Chain E: 79% 17% ..



- Molecule 5: Proteasome subunit alpha type-6

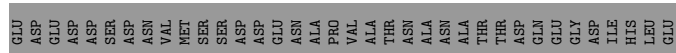
Chain S: 80% 17% ..



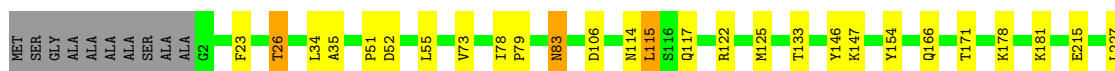
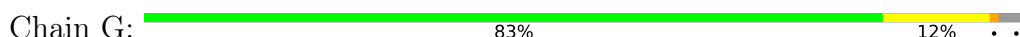
• Molecule 6: Probable proteasome subunit alpha type-7



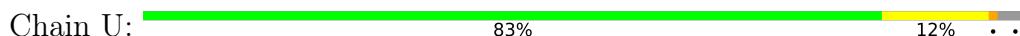
• Molecule 6: Probable proteasome subunit alpha type-7



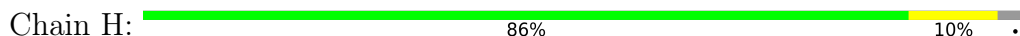
• Molecule 7: Proteasome subunit alpha type-1



• Molecule 7: Proteasome subunit alpha type-1



• Molecule 8: Proteasome subunit beta type-2





- Molecule 8: Proteasome subunit beta type-2

Chain V: 86% 10%



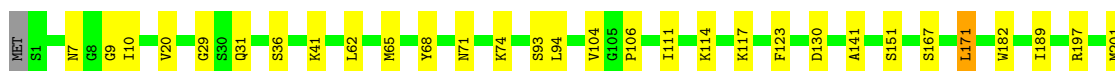
- Molecule 9: Proteasome subunit beta type-3

Chain I: 86% 13%



- Molecule 9: Proteasome subunit beta type-3

Chain W: 85% 14%



- Molecule 10: Proteasome subunit beta type-4

Chain J: 87% 11%



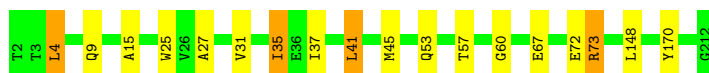
- Molecule 10: Proteasome subunit beta type-4

Chain X: 87% 10%




- Molecule 11: Proteasome subunit beta type-5

Chain K: 91% 7%




- Molecule 11: Proteasome subunit beta type-5

Chain Y:  91% 7%




- Molecule 12: Proteasome subunit beta type-6

Chain L:  86% 12%



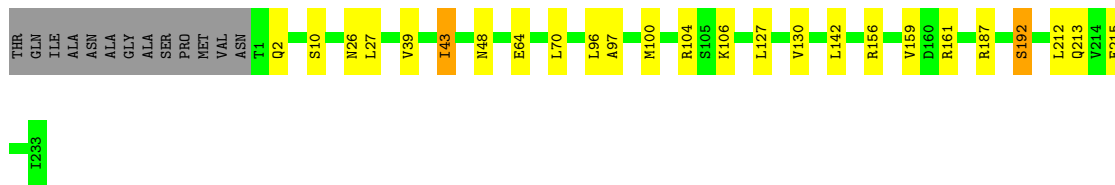
- Molecule 12: Proteasome subunit beta type-6

Chain Z:  87% 12%




- Molecule 13: Proteasome subunit beta type-7

Chain M:  85% 9% 5%



- Molecule 13: Proteasome subunit beta type-7

Chain a:  89% 5% 5%



- Molecule 14: Proteasome subunit beta type-1

Chain N:  95% 5%



- Molecule 14: Proteasome subunit beta type-1

Chain b:  98%



## 4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.00Å 300.67Å 144.54Å 90.00° 113.09° 90.00°	Depositor
Resolution (Å)	30.00 – 2.60	Depositor
% Data completeness (in resolution range)	97.4 (30.00-2.60)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.16 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.182 , 0.222	Depositor
Wilson B-factor (Å <sup>2</sup> )	37.4	Xtriage
Anisotropy	1.388	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	49970	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% 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: MES, A1IFK, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.67	0/1952	0.73	0/2642
1	O	0.67	0/1952	0.72	0/2642
2	B	0.66	0/1934	0.75	0/2618
2	P	0.66	0/1934	0.75	0/2618
3	C	0.67	0/1910	0.75	0/2586
3	Q	0.68	0/1910	0.76	0/2586
4	D	0.67	0/1837	0.74	0/2475
4	R	0.67	0/1837	0.74	0/2475
5	E	0.67	0/1800	0.74	0/2433
5	S	0.67	0/1800	0.74	0/2433
6	F	0.67	0/1932	0.72	0/2609
6	T	0.67	0/1932	0.73	0/2609
7	G	0.66	0/1945	0.75	0/2634
7	U	0.66	0/1945	0.75	0/2634
8	H	0.66	0/1708	0.75	0/2316
8	V	0.66	0/1708	0.75	0/2316
9	I	0.67	0/1611	0.75	0/2174
9	W	0.66	0/1611	0.74	0/2174
10	J	0.65	0/1589	0.74	0/2142
10	X	0.65	0/1589	0.74	0/2142
11	K	0.66	0/1674	0.76	0/2264
11	Y	0.66	0/1674	0.75	0/2264
12	L	0.66	0/1795	0.73	0/2420
12	Z	0.66	0/1795	0.73	0/2420
13	M	0.67	0/1855	0.76	0/2514
13	a	0.67	0/1855	0.76	0/2514
14	N	0.66	0/1541	0.73	0/2087
14	b	0.67	0/1541	0.73	0/2087
All	All	0.66	0/50166	0.74	0/67828

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

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	7	0
1	O	1915	0	1929	10	0
2	B	1904	0	1904	14	0
2	P	1904	0	1904	15	0
3	C	1881	0	1895	19	0
3	Q	1881	0	1895	22	0
4	D	1813	0	1797	10	0
4	R	1813	0	1797	13	0
5	E	1773	0	1775	21	0
5	S	1773	0	1775	21	0
6	F	1892	0	1883	10	0
6	T	1892	0	1883	11	0
7	G	1907	0	1901	15	0
7	U	1907	0	1901	17	0
8	H	1677	0	1678	10	0
8	V	1677	0	1678	10	0
9	I	1581	0	1574	18	0
9	W	1581	0	1574	21	0
10	J	1561	0	1569	9	0
10	X	1561	0	1569	10	0
11	K	1637	0	1585	15	0
11	Y	1637	0	1585	18	0
12	L	1757	0	1711	16	0
12	Z	1757	0	1711	15	0
13	M	1824	0	1832	9	0
13	a	1824	0	1832	0	0
14	N	1512	0	1481	3	0
14	b	1512	0	1481	0	0
15	G	1	0	0	0	0
15	H	1	0	0	0	0
15	I	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
15	K	1	0	0	0	0
15	N	1	0	0	0	0
15	Y	1	0	0	0	0
15	Z	1	0	0	0	0
16	H	12	0	13	0	0
16	J	12	0	13	0	0
16	V	12	0	13	0	0
16	X	12	0	13	1	0
16	b	12	0	13	0	0
17	H	64	0	0	6	0
17	K	64	0	0	2	0
17	V	64	0	0	5	0
17	Y	64	0	0	4	0
18	A	14	0	0	0	0
18	B	14	0	0	0	0
18	C	11	0	0	0	0
18	D	11	0	0	0	0
18	E	5	0	0	0	0
18	F	11	0	0	0	0
18	G	13	0	0	0	0
18	H	18	0	0	0	0
18	I	17	0	0	0	0
18	J	19	0	0	0	0
18	K	18	0	0	0	0
18	L	22	0	0	0	0
18	M	16	0	0	0	0
18	N	14	0	0	0	0
18	O	15	0	0	0	0
18	P	11	0	0	0	0
18	Q	8	0	0	1	0
18	R	7	0	0	0	0
18	S	5	0	0	0	0
18	T	7	0	0	0	0
18	U	20	0	0	0	0
18	V	10	0	0	0	0
18	W	16	0	0	0	0
18	X	15	0	0	0	0
18	Y	15	0	0	0	0
18	Z	14	0	0	0	0
18	a	15	0	0	0	0
18	b	18	0	0	0	0
All	All	49970	0	49093	321	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 321 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)
11:K:73:ARG:HH21	11:K:73:ARG:HG3	1.12	1.12
3:Q:160:GLN:HA	3:Q:160:GLN:HE21	1.40	0.85
3:C:160:GLN:HE21	3:C:160:GLN:HA	1.41	0.84
17:V:302:A1IFK:C5'	9:W:130:ASP:CB	2.62	0.77
7:G:23:PHE:O	7:G:26:THR:HB	1.85	0.75

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%)	240 (97%)	5 (2%)	3 (1%)	13	27
1	O	248/250 (99%)	240 (97%)	5 (2%)	3 (1%)	13	27
2	B	242/258 (94%)	234 (97%)	6 (2%)	2 (1%)	19	39
2	P	242/258 (94%)	234 (97%)	6 (2%)	2 (1%)	19	39
3	C	238/254 (94%)	225 (94%)	10 (4%)	3 (1%)	12	24
3	Q	238/254 (94%)	225 (94%)	10 (4%)	3 (1%)	12	24
4	D	231/260 (89%)	226 (98%)	5 (2%)	0	100	100
4	R	231/260 (89%)	227 (98%)	4 (2%)	0	100	100
5	E	229/234 (98%)	217 (95%)	12 (5%)	0	100	100
5	S	229/234 (98%)	216 (94%)	13 (6%)	0	100	100
6	F	241/288 (84%)	232 (96%)	8 (3%)	1 (0%)	34	57
6	T	241/288 (84%)	232 (96%)	8 (3%)	1 (0%)	34	57

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	G	239/252 (95%)	231 (97%)	7 (3%)	1 (0%)	34	57
7	U	239/252 (95%)	229 (96%)	9 (4%)	1 (0%)	34	57
8	H	219/231 (95%)	215 (98%)	4 (2%)	0	100	100
8	V	219/231 (95%)	215 (98%)	4 (2%)	0	100	100
9	I	202/205 (98%)	194 (96%)	8 (4%)	0	100	100
9	W	202/205 (98%)	193 (96%)	9 (4%)	0	100	100
10	J	193/198 (98%)	190 (98%)	1 (0%)	2 (1%)	15	32
10	X	193/198 (98%)	190 (98%)	1 (0%)	2 (1%)	15	32
11	K	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
11	Y	209/211 (99%)	202 (97%)	7 (3%)	0	100	100
12	L	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
12	Z	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
13	M	231/246 (94%)	223 (96%)	8 (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	6272/6610 (95%)	6061 (97%)	187 (3%)	24 (0%)	34	57

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	202	GLN
6	F	203	ASN
3	Q	202	GLN
6	T	203	ASN
3	C	239	GLN

### 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%)	201 (96%)	8 (4%)	33	59
1	O	209/209 (100%)	201 (96%)	8 (4%)	33	59
2	B	203/216 (94%)	188 (93%)	15 (7%)	13	28
2	P	203/216 (94%)	188 (93%)	15 (7%)	13	28
3	C	212/226 (94%)	196 (92%)	16 (8%)	13	27
3	Q	212/226 (94%)	196 (92%)	16 (8%)	13	27
4	D	194/215 (90%)	180 (93%)	14 (7%)	14	29
4	R	194/215 (90%)	180 (93%)	14 (7%)	14	29
5	E	190/193 (98%)	172 (90%)	18 (10%)	8	16
5	S	190/193 (98%)	172 (90%)	18 (10%)	8	16
6	F	201/239 (84%)	184 (92%)	17 (8%)	10	21
6	T	201/239 (84%)	185 (92%)	16 (8%)	12	24
7	G	206/210 (98%)	192 (93%)	14 (7%)	16	32
7	U	206/210 (98%)	192 (93%)	14 (7%)	16	32
8	H	180/189 (95%)	172 (96%)	8 (4%)	28	53
8	V	180/189 (95%)	171 (95%)	9 (5%)	24	47
9	I	172/173 (99%)	165 (96%)	7 (4%)	30	56
9	W	172/173 (99%)	165 (96%)	7 (4%)	30	56
10	J	173/175 (99%)	164 (95%)	9 (5%)	23	46
10	X	173/175 (99%)	166 (96%)	7 (4%)	31	57
11	K	168/168 (100%)	161 (96%)	7 (4%)	30	55
11	Y	168/168 (100%)	161 (96%)	7 (4%)	30	55
12	L	185/185 (100%)	176 (95%)	9 (5%)	25	48
12	Z	185/185 (100%)	177 (96%)	8 (4%)	29	54
13	M	199/208 (96%)	186 (94%)	13 (6%)	17	34
13	a	199/208 (96%)	186 (94%)	13 (6%)	17	34
14	N	162/162 (100%)	158 (98%)	4 (2%)	47	73
14	b	162/162 (100%)	158 (98%)	4 (2%)	47	73
All	All	5308/5536 (96%)	4993 (94%)	315 (6%)	19	39

5 of 315 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	T	47	GLU
11	Y	35	ILE
6	T	198	LEU
8	V	3	ILE
13	a	43	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 135 such sidechains are listed below:

Mol	Chain	Res	Type
11	Y	9	GLN
12	Z	1	GLN
13	a	194	ASN
10	J	191	GLN
10	J	147	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 7 are monoatomic - leaving 9 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	A1IFK	H	303	8	61,68,68	1.76	8 (13%)	74,96,96	1.62	10 (13%)
17	A1IFK	Y	302	11	61,68,68	2.06	10 (16%)	74,96,96	1.45	8 (10%)
16	MES	J	201	-	12,12,12	0.79	0	14,16,16	0.54	0
17	A1IFK	K	302	11	61,68,68	1.94	12 (19%)	74,96,96	1.41	7 (9%)
16	MES	V	301	-	12,12,12	0.73	0	14,16,16	0.40	0
16	MES	H	301	-	12,12,12	0.73	0	14,16,16	0.39	0
16	MES	b	201	-	12,12,12	0.82	0	14,16,16	0.93	1 (7%)
16	MES	X	201	-	12,12,12	0.74	0	14,16,16	0.36	0
17	A1IFK	V	302	8	61,68,68	1.75	8 (13%)	74,96,96	1.58	10 (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	A1IFK	H	303	8	-	2/49/99/99	0/4/5/5
17	A1IFK	Y	302	11	-	10/49/99/99	0/4/5/5
16	MES	J	201	-	-	5/6/14/14	0/1/1/1
17	A1IFK	K	302	11	-	10/49/99/99	0/4/5/5
16	MES	V	301	-	-	1/6/14/14	0/1/1/1
16	MES	H	301	-	-	2/6/14/14	0/1/1/1
16	MES	b	201	-	-	0/6/14/14	0/1/1/1
16	MES	X	201	-	-	2/6/14/14	0/1/1/1
17	A1IFK	V	302	8	-	3/49/99/99	0/4/5/5

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	Y	302	A1IFK	CB2-CA2	7.29	1.65	1.53
17	H	303	A1IFK	CB1-CG	-6.79	1.35	1.51
17	K	302	A1IFK	CB1-CG	-6.76	1.35	1.51
17	Y	302	A1IFK	CB1-CG	-6.66	1.35	1.51
17	K	302	A1IFK	O58-C11	-6.02	1.36	1.44

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	H	303	A1IFK	O20-C18-N17	6.31	119.30	111.01

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	Y	302	A1IFK	O20-C18-N17	5.67	118.45	111.01
17	V	302	A1IFK	O20-C18-N17	5.64	118.41	111.01
17	K	302	A1IFK	O20-C18-N17	5.54	118.28	111.01
17	Y	302	A1IFK	C25-N17-C16	4.87	122.00	112.62

There are no chirality outliers.

5 of 35 torsion outliers are listed below:

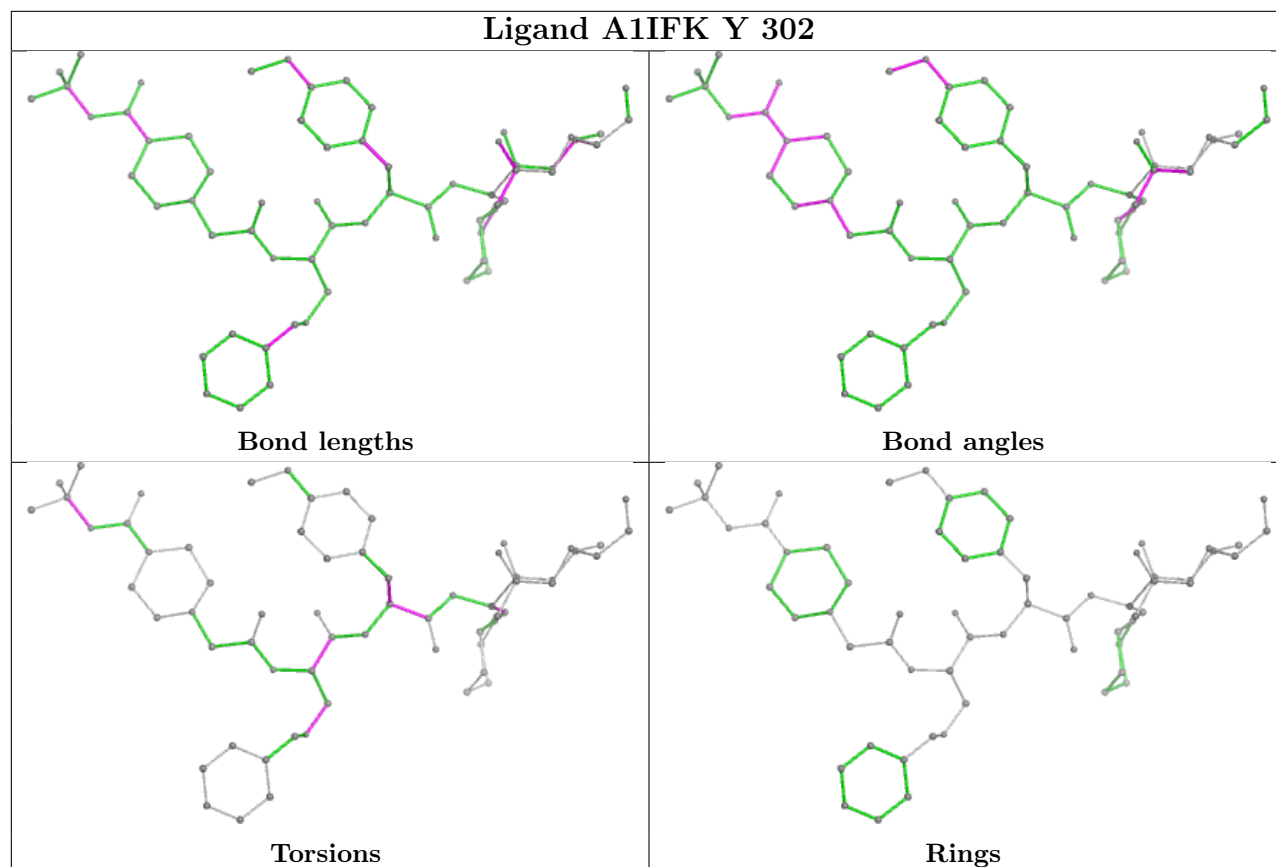
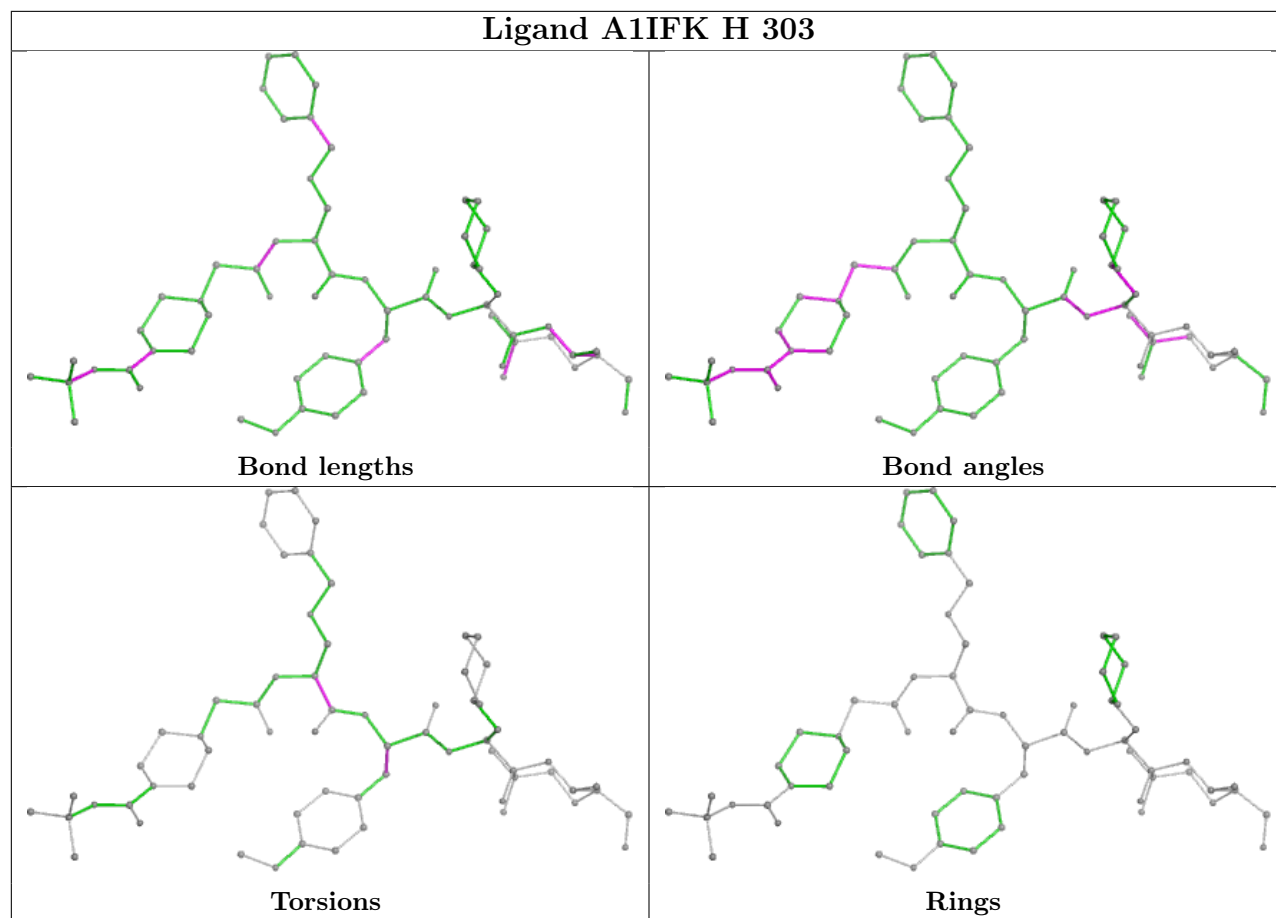
Mol	Chain	Res	Type	Atoms
16	J	201	MES	C8-C7-N4-C3
16	J	201	MES	C7-C8-S-O2S
16	J	201	MES	C7-C8-S-O3S
17	Y	302	A1IFK	CA-CB-OG-C'
17	K	302	A1IFK	C24-C21-O20-C18

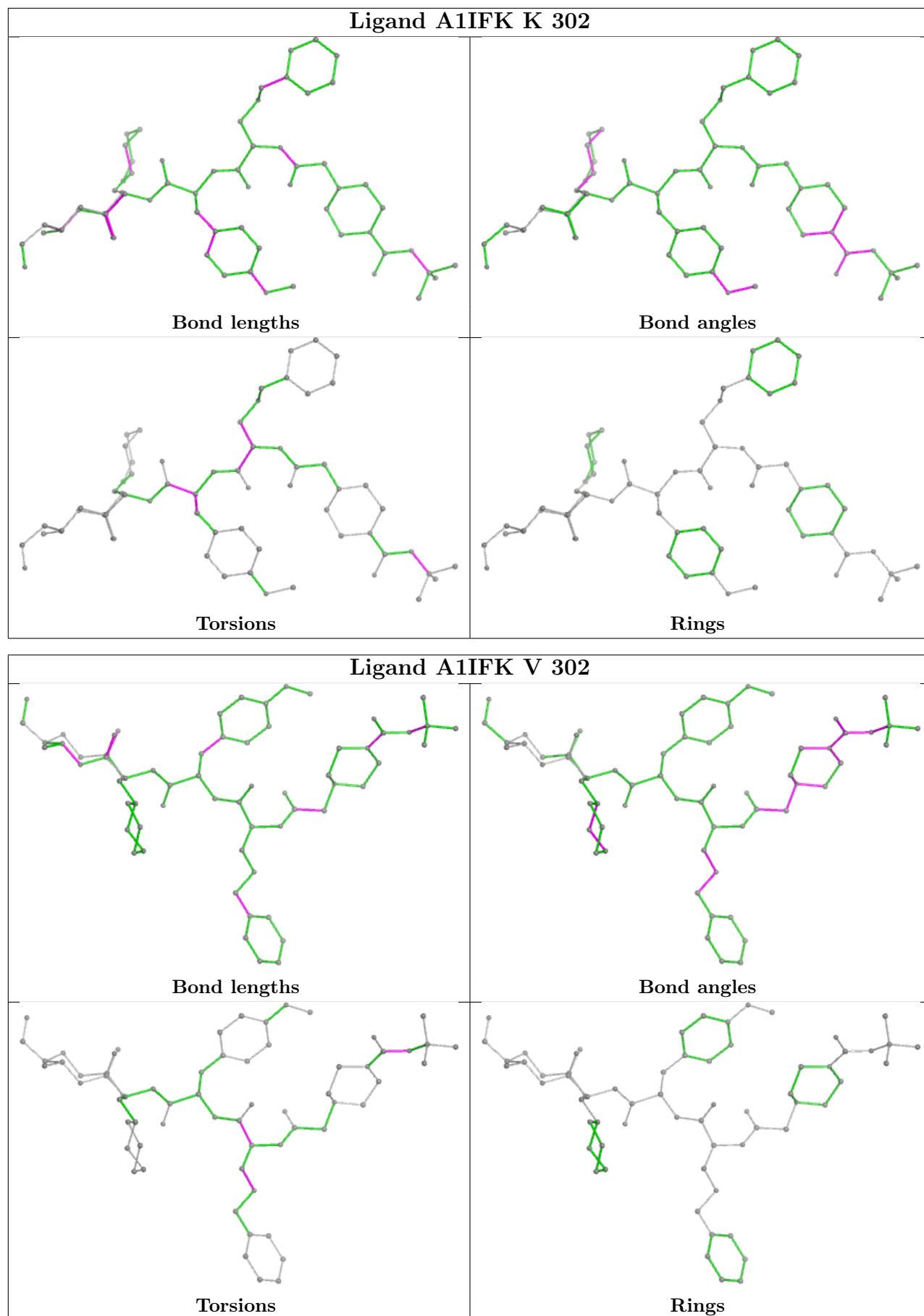
There are no ring outliers.

5 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
17	H	303	A1IFK	6	0
17	Y	302	A1IFK	4	0
17	K	302	A1IFK	2	0
16	X	201	MES	1	0
17	V	302	A1IFK	5	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 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

### 6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

### 6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

### 6.4 Ligands

EDS failed to run properly - this section is therefore empty.

### 6.5 Other polymers

EDS failed to run properly - this section is therefore empty.