



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 07:32 am BST

PDB ID : 4EU2  
Title : Crystal structure of 20s proteasome with novel inhibitor K-7174  
Authors : Kikuchi, J.; Shibayama, N.; Yamada, S.; Wada, T.; Nobuyoshi, M.; Izumi, T.; Akutsu, M.; Kano, Y.; Ohki, M.; Sugiyama, K.; Park, S.-Y.; Furukawa, Y.  
Deposited on : 2012-04-25  
Resolution : 2.51 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 : 2.11  
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) : 2.11

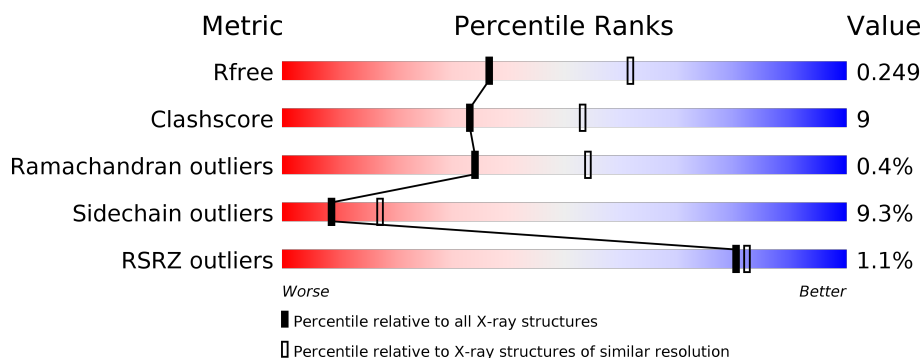
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.51 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.



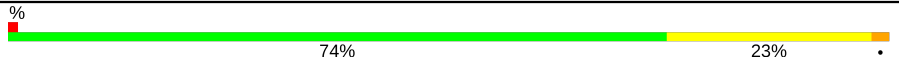



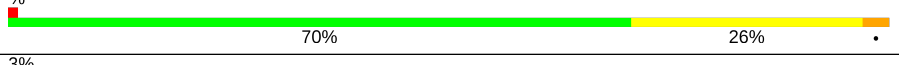
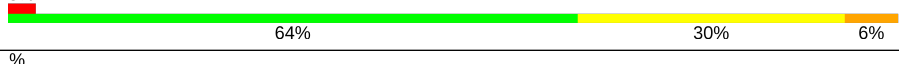
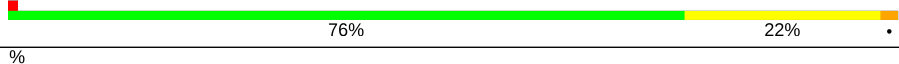

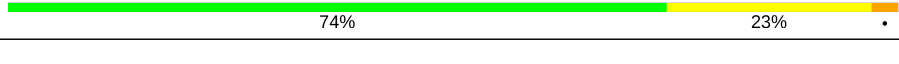


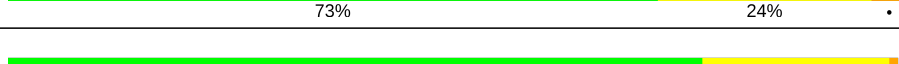

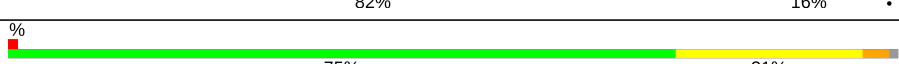

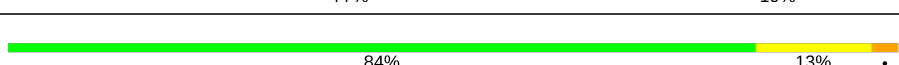
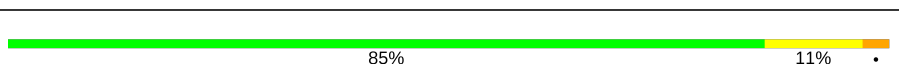
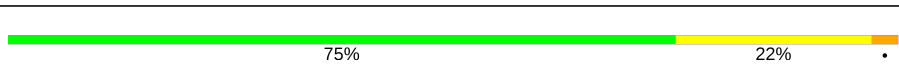
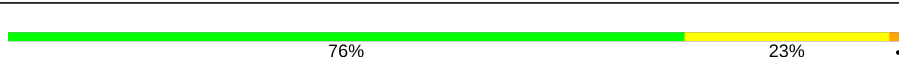



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	241	<div> <div>%</div> <div> <div></div> <div>75%</div> <div>22%</div> <div>.</div> </div> </div>
1	O	241	<div> <div>70%</div> <div>27%</div> <div>.</div> </div>
2	B	250	<div> <div>84%</div> <div>14%</div> <div>.</div> </div>
2	P	250	<div> <div>67%</div> <div>30%</div> <div>.</div> </div>
3	C	244	<div> <div>%</div> <div>76%</div> <div>23%</div> <div>.</div> </div>
3	Q	244	<div> <div>2%</div> <div>72%</div> <div>25%</div> <div>.</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	D	241	
4	R	241	
5	E	242	
5	S	242	
6	F	233	
6	T	233	
7	G	244	
7	U	244	
8	H	196	
8	V	196	
9	I	222	
9	W	222	
10	J	204	
10	X	204	
11	K	198	
11	Y	198	
12	L	212	
12	Z	212	
13	1	222	
13	M	222	
14	2	233	
14	N	233	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
15	WPI	H	300	-	-	-	X
15	WPI	I	301	-	X	-	X
15	WPI	L	301	-	-	-	X
15	WPI	W	301	-	-	-	X
15	WPI	Z	301	-	X	-	X

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 50701 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 component C7-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	S	0	0	0
			1903	1212	319	364	8			
1	O	241	Total	C	N	O	S	0	0	0
			1903	1212	319	364	8			

- Molecule 2 is a protein called Proteasome component Y7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	249	Total	C	N	O	S	0	0	0
			1906	1213	314	375	4			
2	P	249	Total	C	N	O	S	0	0	0
			1906	1213	314	375	4			

- Molecule 3 is a protein called Proteasome component Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	244	Total	C	N	O	S	0	0	0
			1905	1201	321	380	3			
3	Q	244	Total	C	N	O	S	0	0	0
			1905	1201	321	380	3			

- Molecule 4 is a protein called Proteasome component PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	241	Total	C	N	O	S	0	0	0
			1891	1181	331	375	4			
4	R	241	Total	C	N	O	S	0	0	0
			1891	1181	331	375	4			

- Molecule 5 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	242	Total	C	N	O	S	0	0	0
			1862	1162	314	379	7			
5	S	242	Total	C	N	O	S	0	0	0
			1862	1162	314	379	7			

- Molecule 6 is a protein called Proteasome component PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	232	Total	C	N	O	S	0	0	0
			1784	1120	311	349	4			
6	T	232	Total	C	N	O	S	0	0	0
			1784	1120	311	349	4			

- Molecule 7 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	244	Total	C	N	O	S	0	0	0
			1897	1205	330	358	4			
7	U	244	Total	C	N	O	S	0	0	0
			1897	1205	330	358	4			

- Molecule 8 is a protein called Proteasome component PRE3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			
8	V	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			

- Molecule 9 is a protein called Proteasome component PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	222	Total	C	N	O	S	0	0	0
			1685	1061	293	324	7			
9	W	222	Total	C	N	O	S	0	0	0
			1685	1061	293	324	7			

- Molecule 10 is a protein called Proteasome component PUP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			

- Molecule 11 is a protein called Proteasome component C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	196	Total	C	N	O	S	0	0	0
			1570	997	266	301	6			
11	Y	196	Total	C	N	O	S	0	0	0
			1570	997	266	301	6			

- Molecule 12 is a protein called Proteasome component PRE2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	212	Total	C	N	O	S	0	0	0
			1646	1045	282	312	7			
12	Z	212	Total	C	N	O	S	0	0	0
			1646	1045	282	312	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	33	ARG	LYS	CONFLICT	UNP P30656
Z	33	ARG	LYS	CONFLICT	UNP P30656

- Molecule 13 is a protein called Proteasome component C5.

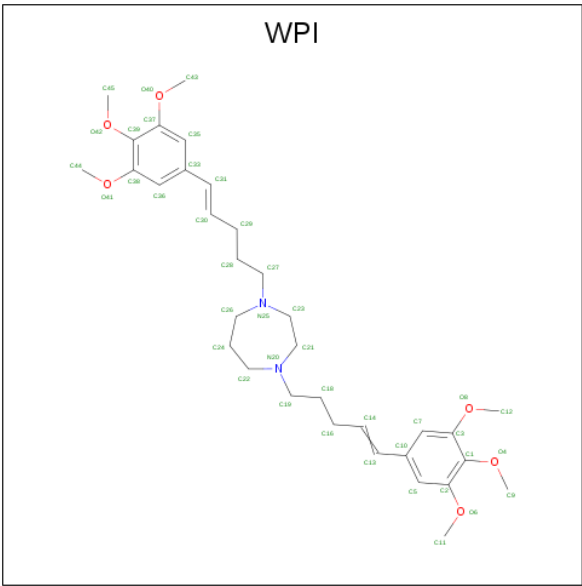
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			
13	1	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			

- Molecule 14 is a protein called Proteasome component PRE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			
14	2	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			

- Molecule 15 is 1,4-bis[(4E)-5-(3,4,5-trimethoxyphenyl)pent-4-en-1-yl]-1,4-diazepane

(three-letter code: WPI) (formula: C<sub>33</sub>H<sub>48</sub>N<sub>2</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	H	1	Total	C	N	O	0	0
			41	33	2	6		
15	I	1	Total	C	N	O	0	0
			41	33	2	6		
15	L	1	Total	C	N	O	0	0
			41	33	2	6		
15	V	1	Total	C	N	O	0	0
			41	33	2	6		
15	W	1	Total	C	N	O	0	0
			41	33	2	6		
15	Z	1	Total	C	N	O	0	0
			41	33	2	6		

- Molecule 16 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	47	Total	O	0	0
			47	47		
16	B	27	Total	O	0	0
			27	27		
16	C	34	Total	O	0	0
			34	34		
16	D	35	Total	O	0	0
			35	35		
16	E	30	Total	O	0	0
			30	30		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	F	18	Total 18	O 18	0	0
16	G	36	Total 36	O 36	0	0
16	H	34	Total 34	O 34	0	0
16	I	29	Total 29	O 29	0	0
16	J	47	Total 47	O 47	0	0
16	K	34	Total 34	O 34	0	0
16	L	46	Total 46	O 46	0	0
16	M	51	Total 51	O 51	0	0
16	N	44	Total 44	O 44	0	0
16	O	33	Total 33	O 33	0	0
16	P	32	Total 32	O 32	0	0
16	Q	33	Total 33	O 33	0	0
16	R	18	Total 18	O 18	0	0
16	S	23	Total 23	O 23	0	0
16	T	19	Total 19	O 19	0	0
16	U	38	Total 38	O 38	0	0
16	V	47	Total 47	O 47	0	0
16	W	36	Total 36	O 36	0	0
16	X	36	Total 36	O 36	0	0
16	Y	49	Total 49	O 49	0	0
16	Z	41	Total 41	O 41	0	0

*Continued on next page...*

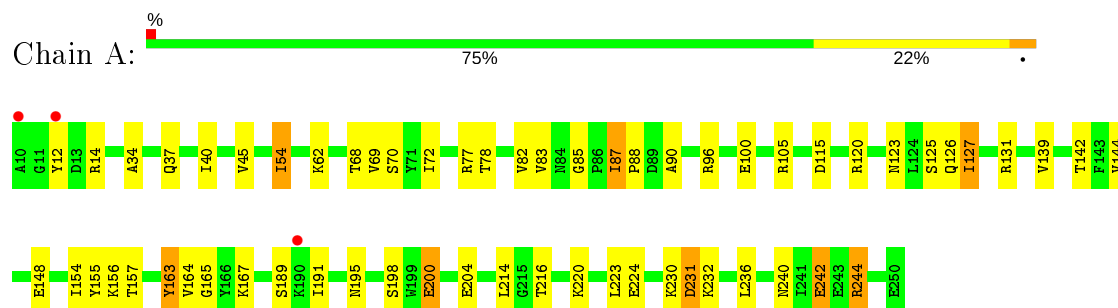
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	1	41	Total	O	0	0
			41	41		
16	2	51	Total	O	0	0
			51	51		

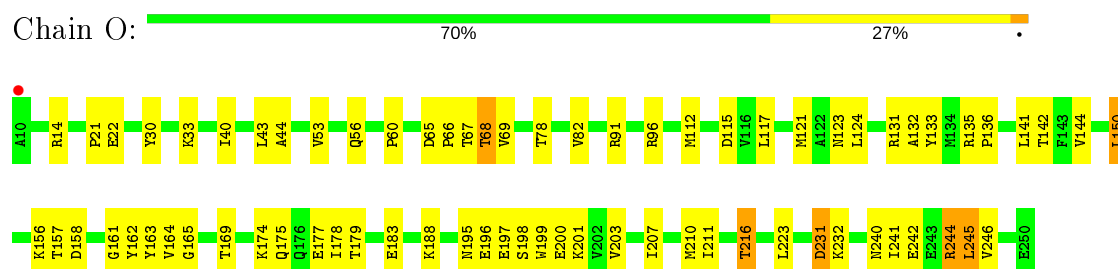
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

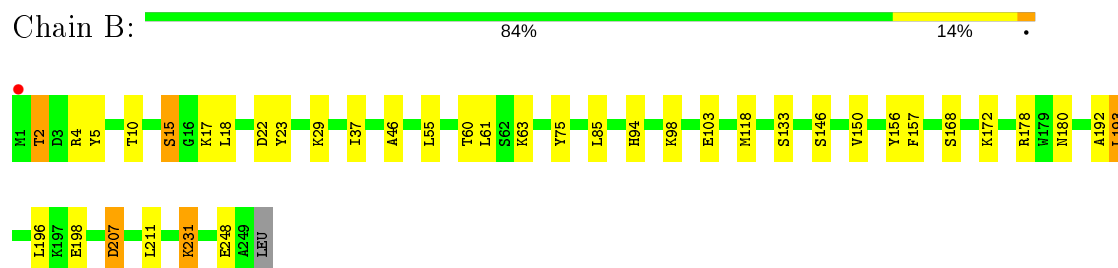
- Molecule 1: Proteasome component C7-alpha



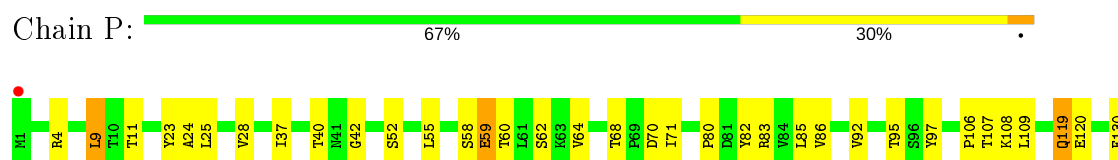
- Molecule 1: Proteasome component C7-alpha

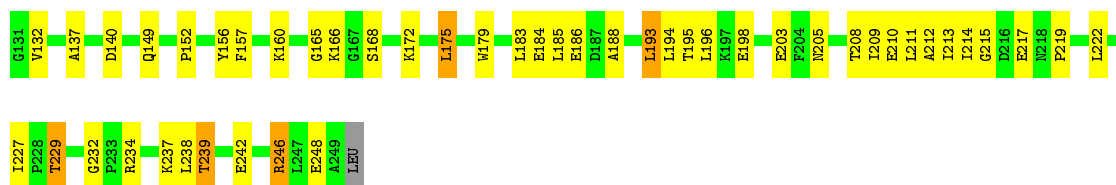


- Molecule 2: Proteasome component Y7

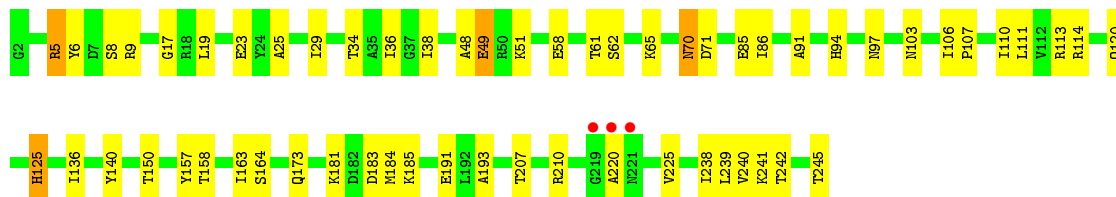
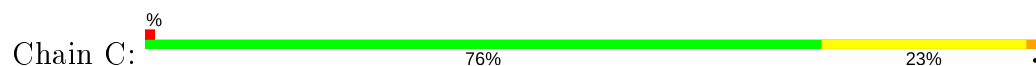


- Molecule 2: Proteasome component Y7

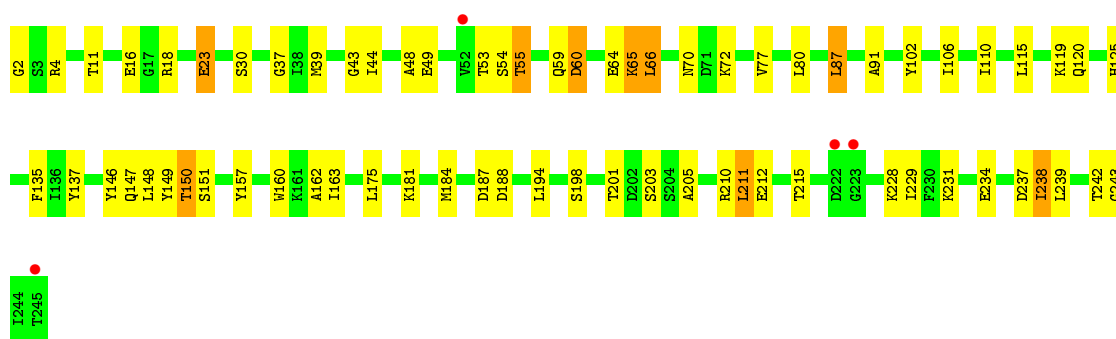




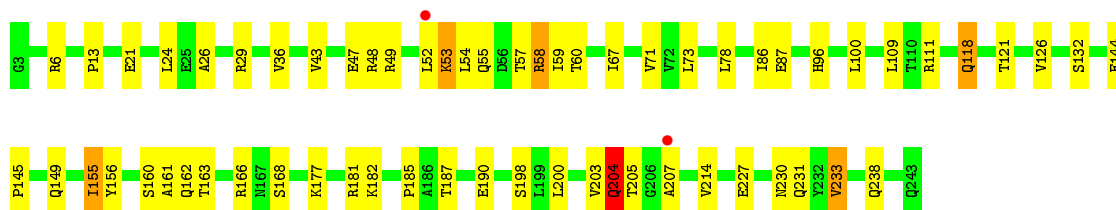
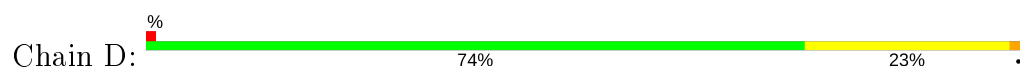
- Molecule 3: Proteasome component Y13



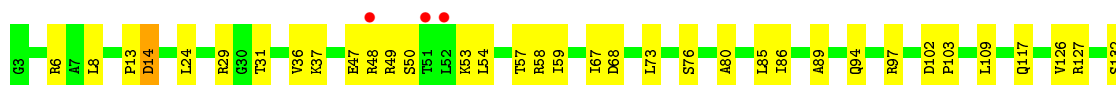
- Molecule 3: Proteasome component Y13

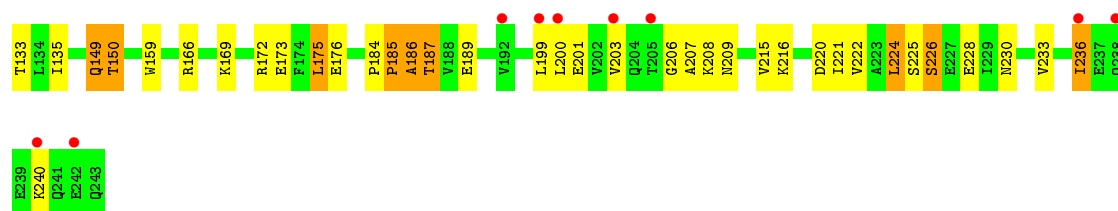


- Molecule 4: Proteasome component PRE6

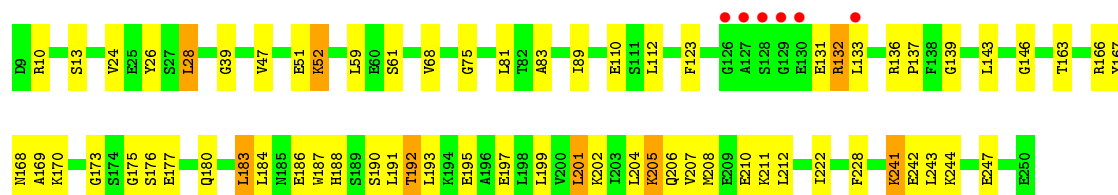
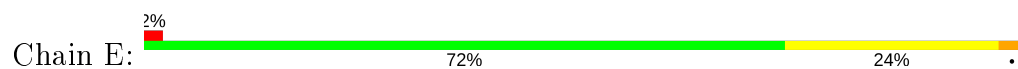


- Molecule 4: Proteasome component PRE6

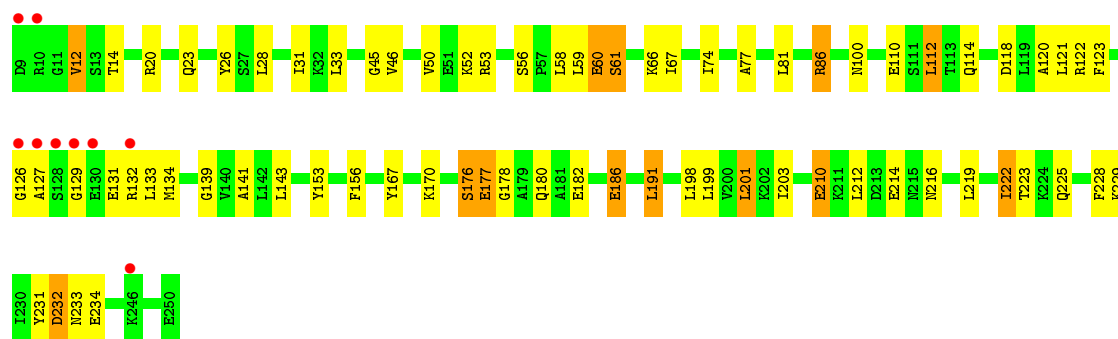




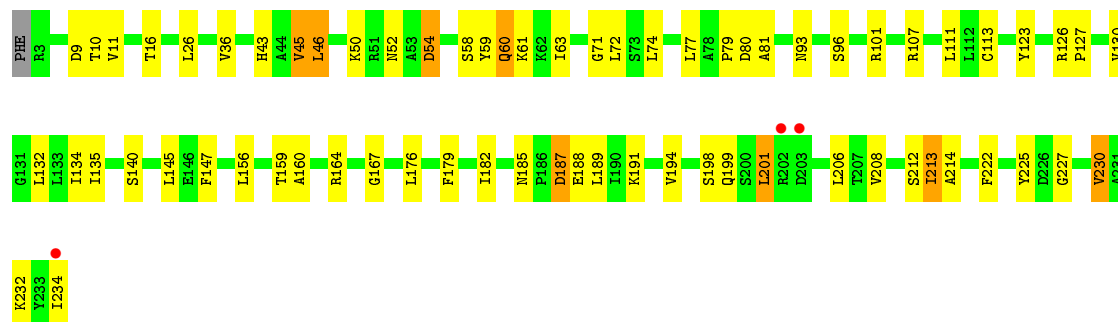
• Molecule 5: Proteasome component PUP2



• Molecule 5: Proteasome component PUP2

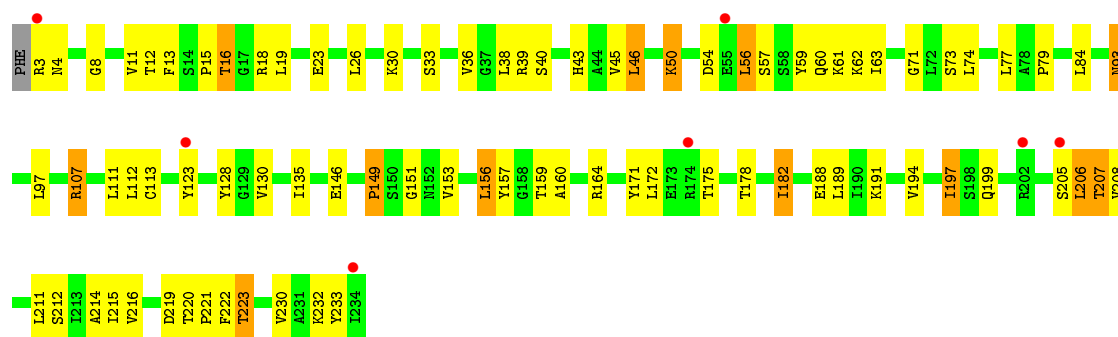


• Molecule 6: Proteasome component PRE5

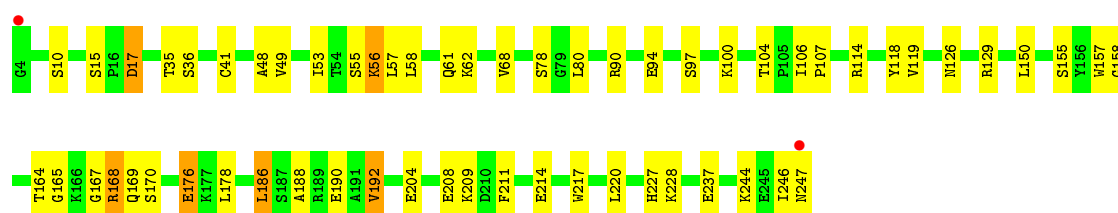
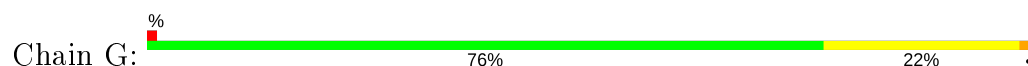


• Molecule 6: Proteasome component PRE5

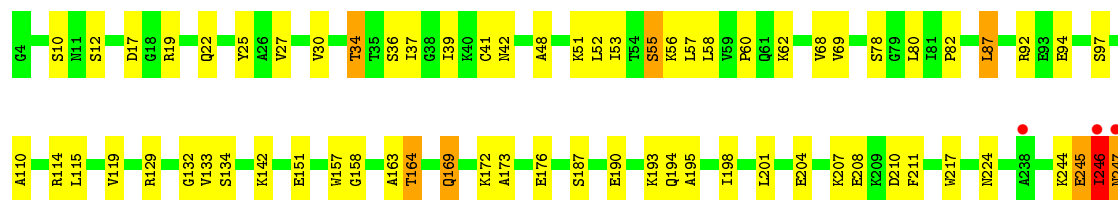




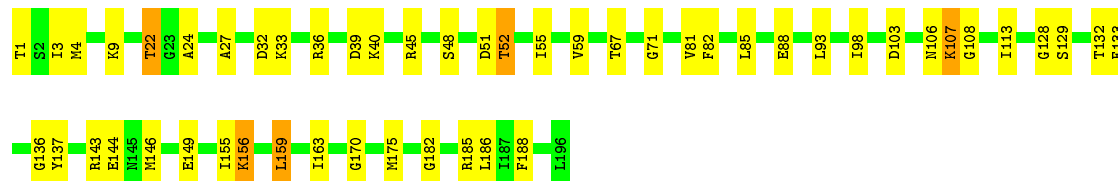
• Molecule 7: Proteasome component C1



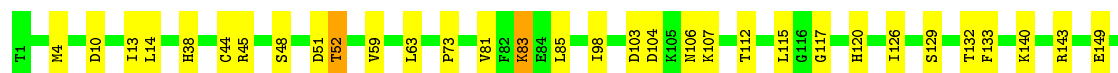
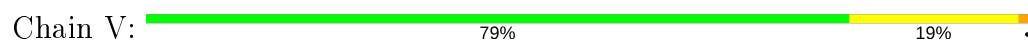
• Molecule 7: Proteasome component C1



• Molecule 8: Proteasome component PRE3

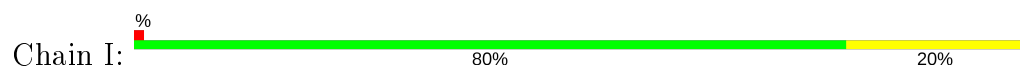


• Molecule 8: Proteasome component PRE3

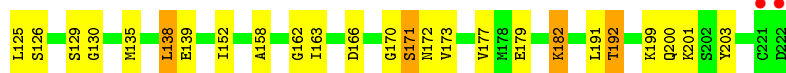
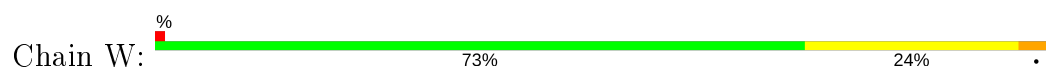




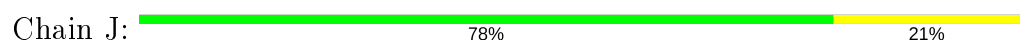
- Molecule 9: Proteasome component PUP1



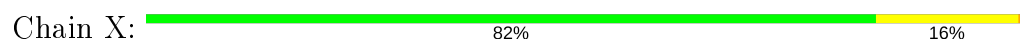
- Molecule 9: Proteasome component PUP1



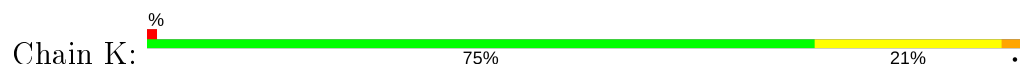
- Molecule 10: Proteasome component PUP3

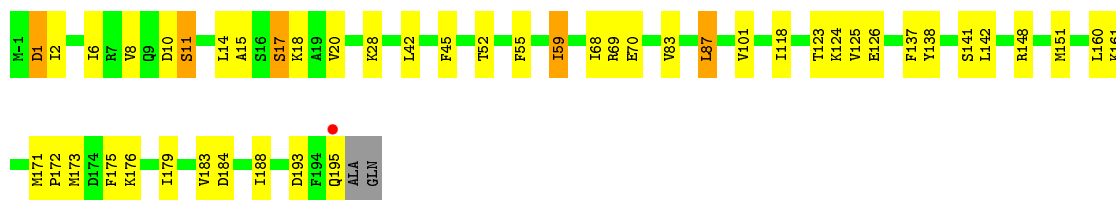


- Molecule 10: Proteasome component PUP3

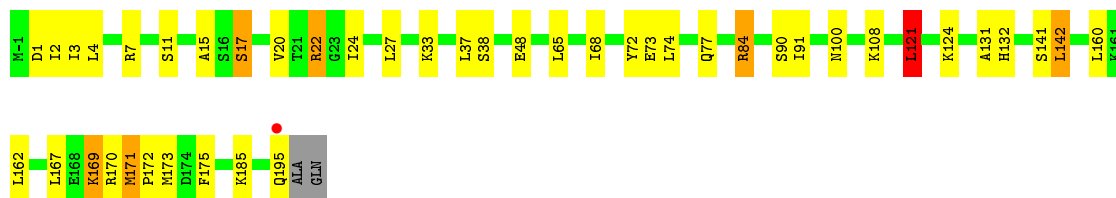
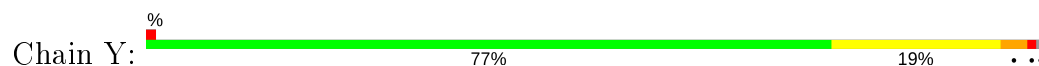


- Molecule 11: Proteasome component C11

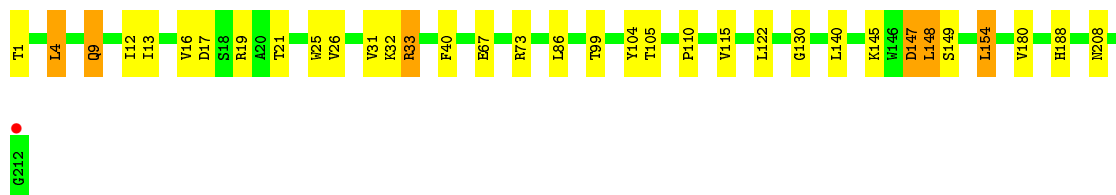
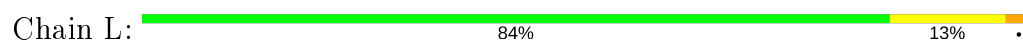




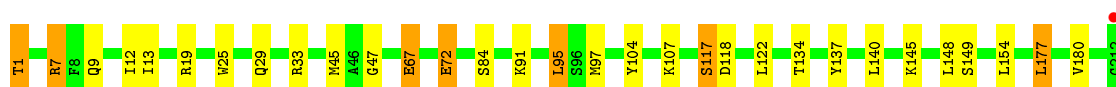
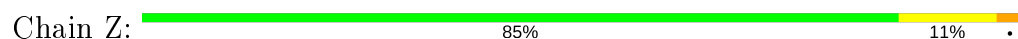
- Molecule 11: Proteasome component C11



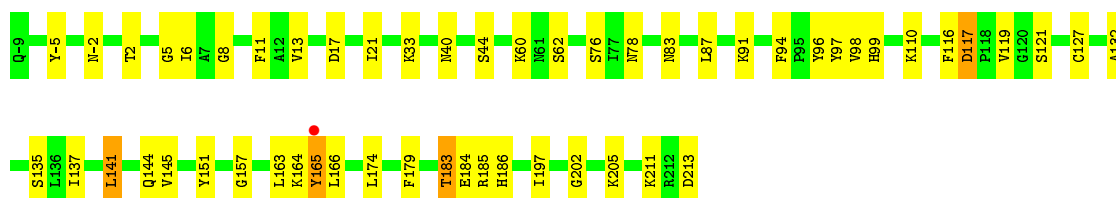
- Molecule 12: Proteasome component PRE2



- Molecule 12: Proteasome component PRE2

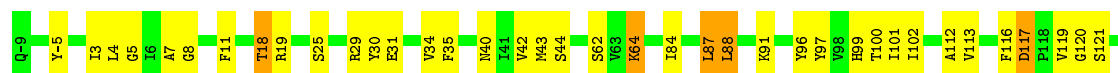


- Molecule 13: Proteasome component C5

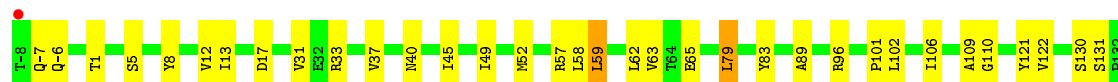
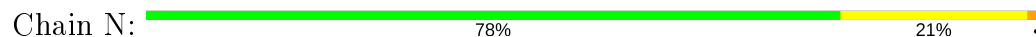


- Molecule 13: Proteasome component C5





• Molecule 14: Proteasome component PRE4



• Molecule 14: Proteasome component PRE4



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.26 Å   301.36 Å   143.96 Å 90.00°   112.86°   90.00°	Depositor
Resolution (Å)	45.70 – 2.51 47.42 – 2.51	Depositor EDS
% Data completeness (in resolution range)	87.6 (45.70-2.51) 87.6 (47.42-2.51)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.86 (at 2.51 Å)	Xtriage
Refinement program	PHENIX 1.7.3 _928	Depositor
R, $R_{free}$	0.202   ,   0.255 0.197   ,   0.249	Depositor DCC
$R_{free}$ test set	15842 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.4	Xtriage
Anisotropy	0.909	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 32.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	50701	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.24% 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: WPI

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	0.45	0/1941	0.62	0/2629
1	O	0.47	0/1941	0.59	0/2629
2	B	0.45	0/1943	0.59	0/2631
2	P	0.43	0/1943	0.57	0/2631
3	C	0.43	0/1935	0.58	0/2618
3	Q	0.44	0/1935	0.59	0/2618
4	D	0.41	0/1920	0.58	0/2598
4	R	0.42	0/1920	0.56	0/2598
5	E	0.45	0/1887	0.59	0/2541
5	S	0.41	0/1887	0.58	0/2541
6	F	0.40	0/1811	0.55	0/2447
6	T	0.41	0/1811	0.59	0/2447
7	G	0.42	0/1937	0.54	0/2614
7	U	0.43	0/1937	0.57	0/2614
8	H	0.45	0/1541	0.60	0/2087
8	V	0.46	0/1541	0.59	0/2087
9	I	0.45	0/1716	0.62	0/2326
9	W	0.43	0/1716	0.58	0/2326
10	J	0.44	0/1611	0.58	0/2174
10	X	0.47	0/1611	0.62	0/2174
11	K	0.44	0/1598	0.58	0/2154
11	Y	0.46	0/1598	0.63	1/2154 (0.0%)
12	L	0.46	0/1683	0.61	0/2277
12	Z	0.42	0/1683	0.58	0/2277
13	1	0.42	0/1795	0.60	0/2420
13	M	0.45	0/1795	0.61	0/2420
14	2	0.47	0/1855	0.62	0/2514
14	N	0.44	0/1855	0.61	0/2514
All	All	0.44	0/50346	0.59	1/68060 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	Y	121	LEU	CA-CB-CG	6.17	129.49	115.30

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	1903	0	1898	45	0
1	O	1903	0	1898	47	0
2	B	1906	0	1918	25	0
2	P	1906	0	1918	44	0
3	C	1905	0	1901	33	0
3	Q	1905	0	1901	43	0
4	D	1891	0	1900	38	0
4	R	1891	0	1900	44	0
5	E	1862	0	1836	37	0
5	S	1862	0	1836	42	0
6	F	1784	0	1788	39	0
6	T	1784	0	1788	44	0
7	G	1897	0	1886	29	0
7	U	1897	0	1886	45	0
8	H	1512	0	1481	40	0
8	V	1512	0	1481	27	0
9	I	1685	0	1688	24	0
9	W	1685	0	1688	35	0
10	J	1581	0	1574	21	0
10	X	1581	0	1574	18	0
11	K	1570	0	1577	27	0
11	Y	1570	0	1577	26	0
12	L	1646	0	1595	22	0
12	Z	1646	0	1595	21	0
13	1	1757	0	1711	36	0
13	M	1757	0	1711	27	0
14	2	1824	0	1832	34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	N	1824	0	1832	30	0
15	H	41	0	48	5	0
15	I	41	0	48	7	0
15	L	41	0	48	5	0
15	V	41	0	48	11	0
15	W	41	0	48	3	0
15	Z	41	0	48	9	0
16	1	41	0	0	1	0
16	2	51	0	0	2	0
16	A	47	0	0	0	0
16	B	27	0	0	0	0
16	C	34	0	0	0	0
16	D	35	0	0	0	0
16	E	30	0	0	0	0
16	F	18	0	0	0	0
16	G	36	0	0	0	0
16	H	34	0	0	1	0
16	I	29	0	0	0	0
16	J	47	0	0	0	0
16	K	34	0	0	1	0
16	L	46	0	0	0	0
16	M	51	0	0	0	0
16	N	44	0	0	1	0
16	O	33	0	0	1	0
16	P	32	0	0	0	0
16	Q	33	0	0	0	0
16	R	18	0	0	0	0
16	S	23	0	0	1	0
16	T	19	0	0	0	0
16	U	38	0	0	0	0
16	V	47	0	0	0	0
16	W	36	0	0	0	0
16	X	36	0	0	0	0
16	Y	49	0	0	1	0
16	Z	41	0	0	1	0
All	All	50701	0	49458	854	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 9.

The worst 5 of 854 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)
1:O:165:GLY:H	2:P:60:THR:HG21	1.14	1.09
7:U:245:GLU:O	7:U:246:ILE:HG22	1.56	1.04
7:U:246:ILE:O	7:U:246:ILE:HG13	1.60	0.99
8:V:45:ARG:HD2	8:V:52:THR:HG23	1.48	0.94
3:Q:70:ASN:HB3	3:Q:72:LYS:H	1.40	0.87

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	239/241 (99%)	229 (96%)	9 (4%)	1 (0%)	34	54
1	O	239/241 (99%)	230 (96%)	7 (3%)	2 (1%)	19	35
2	B	247/250 (99%)	230 (93%)	17 (7%)	0	100	100
2	P	247/250 (99%)	236 (96%)	9 (4%)	2 (1%)	19	35
3	C	242/244 (99%)	232 (96%)	8 (3%)	2 (1%)	19	35
3	Q	242/244 (99%)	231 (96%)	9 (4%)	2 (1%)	19	35
4	D	239/241 (99%)	224 (94%)	13 (5%)	2 (1%)	19	35
4	R	239/241 (99%)	225 (94%)	11 (5%)	3 (1%)	12	21
5	E	240/242 (99%)	228 (95%)	10 (4%)	2 (1%)	19	35
5	S	240/242 (99%)	222 (92%)	16 (7%)	2 (1%)	19	35
6	F	230/233 (99%)	217 (94%)	13 (6%)	0	100	100
6	T	230/233 (99%)	219 (95%)	10 (4%)	1 (0%)	34	54
7	G	242/244 (99%)	234 (97%)	8 (3%)	0	100	100
7	U	242/244 (99%)	229 (95%)	11 (4%)	2 (1%)	19	35
8	H	194/196 (99%)	186 (96%)	8 (4%)	0	100	100
8	V	194/196 (99%)	185 (95%)	9 (5%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	I	220/222 (99%)	212 (96%)	8 (4%)	0	100	100
9	W	220/222 (99%)	215 (98%)	4 (2%)	1 (0%)	29	48
10	J	202/204 (99%)	193 (96%)	8 (4%)	1 (0%)	29	48
10	X	202/204 (99%)	195 (96%)	7 (4%)	0	100	100
11	K	194/198 (98%)	185 (95%)	8 (4%)	1 (0%)	29	48
11	Y	194/198 (98%)	185 (95%)	9 (5%)	0	100	100
12	L	210/212 (99%)	203 (97%)	7 (3%)	0	100	100
12	Z	210/212 (99%)	200 (95%)	10 (5%)	0	100	100
13	1	220/222 (99%)	210 (96%)	10 (4%)	0	100	100
13	M	220/222 (99%)	208 (94%)	12 (6%)	0	100	100
14	2	231/233 (99%)	218 (94%)	12 (5%)	1 (0%)	34	54
14	N	231/233 (99%)	221 (96%)	10 (4%)	0	100	100
All	All	6300/6364 (99%)	6002 (95%)	273 (4%)	25 (0%)	34	54

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	131	GLU
5	E	132	ARG
7	U	246	ILE
4	R	185	PRO
4	R	186	ALA

### 5.3.2 Protein sidechains ⓘ

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	205/205 (100%)	192 (94%)	13 (6%)	18	34
1	O	205/205 (100%)	186 (91%)	19 (9%)	9	17
2	B	208/209 (100%)	194 (93%)	14 (7%)	16	31
2	P	208/209 (100%)	183 (88%)	25 (12%)	5	9

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	203/203 (100%)	188 (93%)	15 (7%)	13	27
3	Q	203/203 (100%)	183 (90%)	20 (10%)	8	15
4	D	213/213 (100%)	196 (92%)	17 (8%)	12	23
4	R	213/213 (100%)	189 (89%)	24 (11%)	6	11
5	E	198/198 (100%)	180 (91%)	18 (9%)	9	18
5	S	198/198 (100%)	176 (89%)	22 (11%)	6	11
6	F	191/192 (100%)	168 (88%)	23 (12%)	5	9
6	T	191/192 (100%)	156 (82%)	35 (18%)	1	3
7	G	201/201 (100%)	179 (89%)	22 (11%)	6	12
7	U	201/201 (100%)	172 (86%)	29 (14%)	3	6
8	H	162/162 (100%)	153 (94%)	9 (6%)	21	40
8	V	162/162 (100%)	148 (91%)	14 (9%)	10	20
9	I	181/181 (100%)	169 (93%)	12 (7%)	16	32
9	W	181/181 (100%)	162 (90%)	19 (10%)	7	13
10	J	172/172 (100%)	156 (91%)	16 (9%)	9	17
10	X	172/172 (100%)	156 (91%)	16 (9%)	9	17
11	K	174/175 (99%)	160 (92%)	14 (8%)	12	23
11	Y	174/175 (99%)	157 (90%)	17 (10%)	8	15
12	L	169/169 (100%)	157 (93%)	12 (7%)	14	28
12	Z	169/169 (100%)	155 (92%)	14 (8%)	11	22
13	1	185/185 (100%)	175 (95%)	10 (5%)	22	42
13	M	185/185 (100%)	168 (91%)	17 (9%)	9	18
14	2	199/199 (100%)	180 (90%)	19 (10%)	8	17
14	N	199/199 (100%)	187 (94%)	12 (6%)	19	37
All	All	5322/5328 (100%)	4825 (91%)	497 (9%)	9	17

5 of 497 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	O	169	THR
4	R	6	ARG
12	Z	84	SER
1	O	244	ARG
2	P	198	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	83	ASN
1	O	84	ASN
9	W	172	ASN
12	L	133	GLN
8	V	161	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

6 ligands are modelled in this entry.

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	$\# Z  > 2$
15	WPI	H	300	-	43,43,43	2.07	10 (23%)	51,55,55	3.22	25 (49%)
15	WPI	V	201	-	43,43,43	2.03	9 (20%)	51,55,55	3.12	22 (43%)
15	WPI	Z	301	-	43,43,43	2.09	11 (25%)	51,55,55	3.14	27 (52%)
15	WPI	W	301	-	43,43,43	2.01	9 (20%)	51,55,55	2.73	21 (41%)
15	WPI	I	301	-	43,43,43	2.01	12 (27%)	51,55,55	3.07	27 (52%)
15	WPI	L	301	-	43,43,43	2.13	12 (27%)	51,55,55	2.83	22 (43%)

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
15	WPI	H	300	-	-	18/28/39/39	0/3/3/3
15	WPI	V	201	-	-	12/28/39/39	0/3/3/3
15	WPI	Z	301	-	-	18/28/39/39	0/3/3/3
15	WPI	W	301	-	-	20/28/39/39	0/3/3/3
15	WPI	I	301	-	-	19/28/39/39	0/3/3/3
15	WPI	L	301	-	-	17/28/39/39	0/3/3/3

The worst 5 of 63 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	I	301	WPI	C31-C30	6.87	1.53	1.31
15	V	201	WPI	C31-C30	6.86	1.53	1.31
15	W	301	WPI	C31-C30	6.85	1.53	1.31
15	Z	301	WPI	C13-C14	6.77	1.53	1.31
15	Z	301	WPI	C31-C30	6.75	1.53	1.31

The worst 5 of 144 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	H	300	WPI	C29-C30-C31	-9.01	108.47	125.45
15	I	301	WPI	C16-C14-C13	-8.68	109.09	125.45
15	W	301	WPI	C16-C14-C13	-8.23	109.95	125.45
15	V	201	WPI	C16-C14-C13	-8.08	110.22	125.45
15	Z	301	WPI	C16-C14-C13	-7.65	111.03	125.45

There are no chirality outliers.

5 of 104 torsion outliers are listed below:

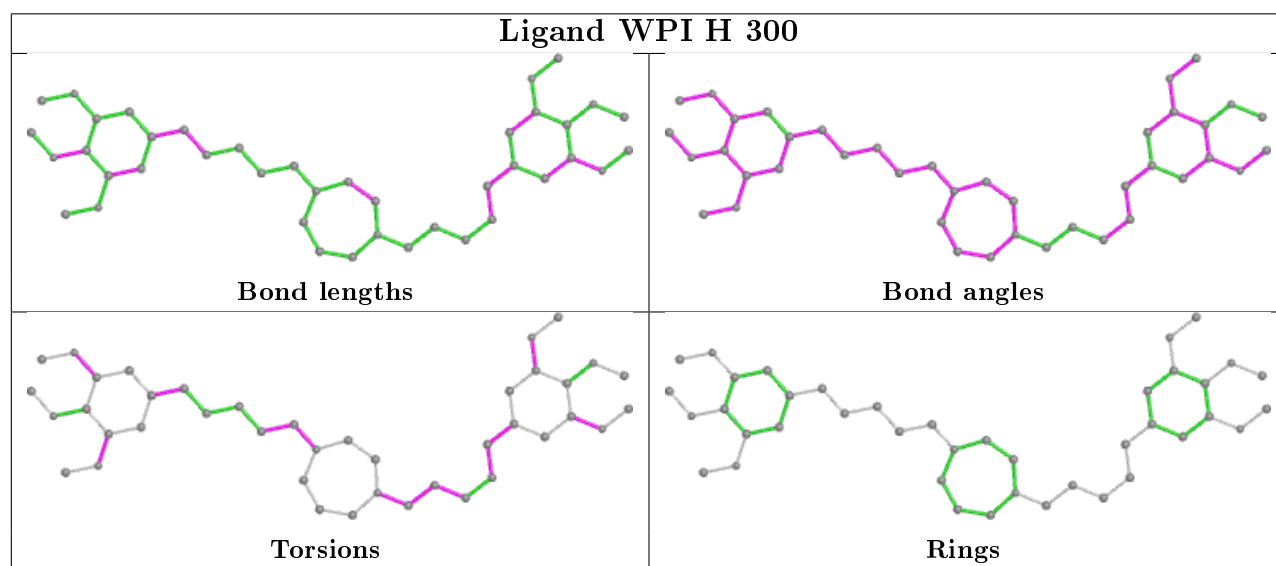
Mol	Chain	Res	Type	Atoms
15	H	300	WPI	C10-C13-C14-C16
15	H	300	WPI	C14-C16-C18-C19
15	H	300	WPI	C18-C19-N20-C22
15	H	300	WPI	C28-C27-N25-C26
15	V	201	WPI	C18-C19-N20-C22

There are no ring outliers.

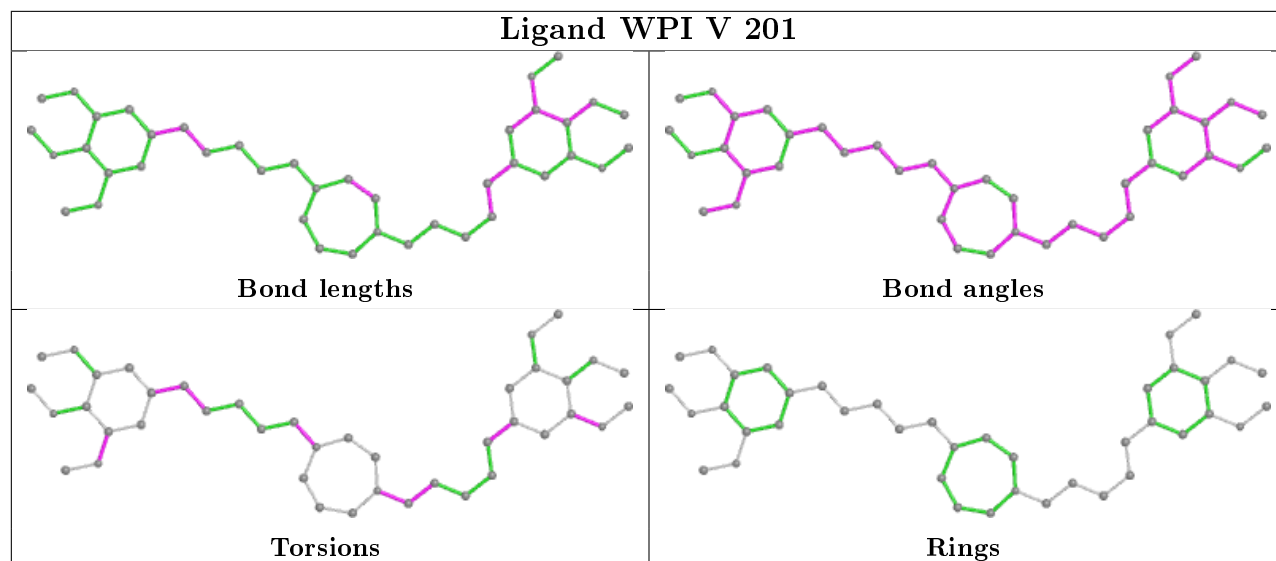
6 monomers are involved in 40 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	H	300	WPI	5	0
15	V	201	WPI	11	0
15	Z	301	WPI	9	0
15	W	301	WPI	3	0
15	I	301	WPI	7	0
15	L	301	WPI	5	0

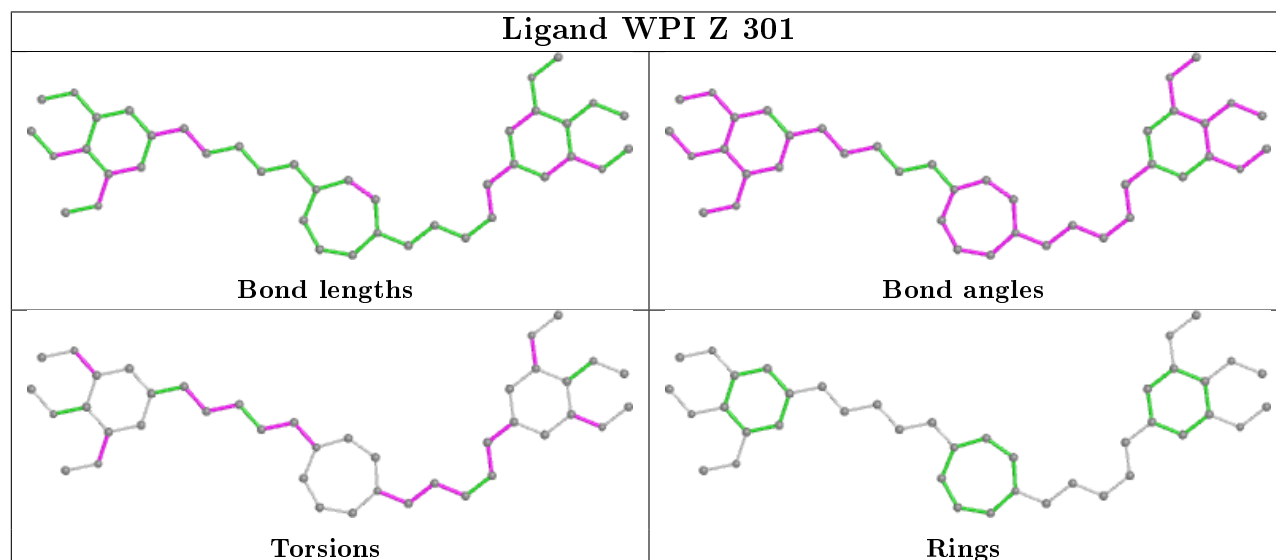
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



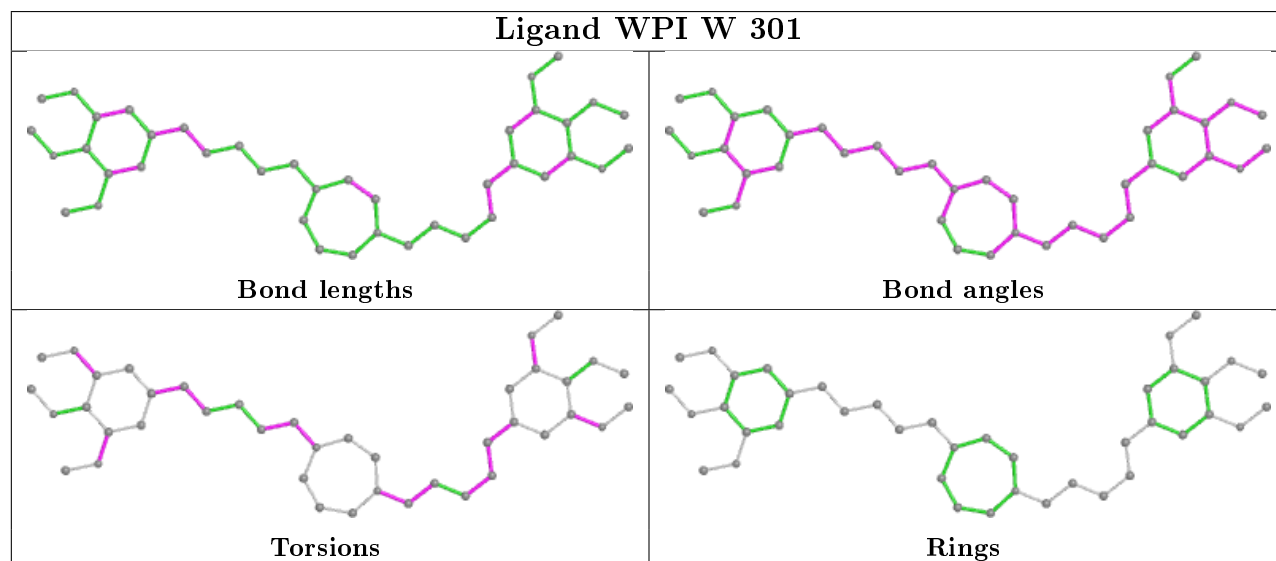
## Ligand WPI V 201

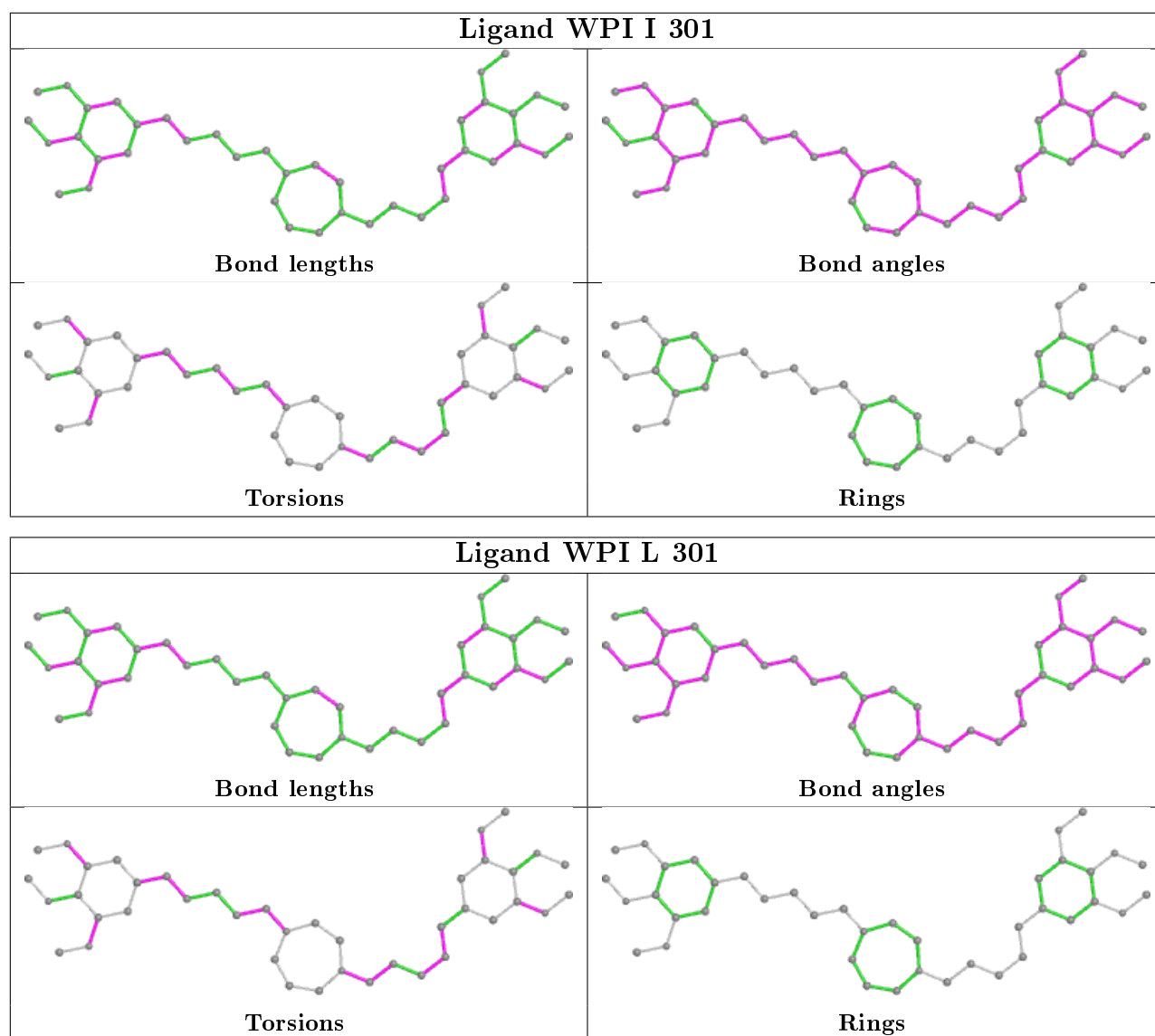


## Ligand WPI Z 301



## Ligand WPI W 301





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	241/241 (100%)	-0.48	3 (1%) 79 80	18, 32, 60, 80	0
1	O	241/241 (100%)	-0.50	1 (0%) 92 93	20, 36, 61, 88	0
2	B	249/250 (99%)	-0.49	1 (0%) 92 93	21, 34, 63, 85	0
2	P	249/250 (99%)	-0.43	1 (0%) 92 93	21, 39, 69, 84	0
3	C	244/244 (100%)	-0.42	3 (1%) 79 80	19, 37, 72, 128	0
3	Q	244/244 (100%)	-0.36	4 (1%) 72 74	18, 37, 71, 110	0
4	D	241/241 (100%)	-0.34	2 (0%) 86 87	19, 39, 86, 96	0
4	R	241/241 (100%)	-0.10	12 (4%) 28 30	22, 45, 93, 116	0
5	E	242/242 (100%)	-0.30	6 (2%) 57 61	20, 40, 69, 122	0
5	S	242/242 (100%)	-0.24	9 (3%) 41 45	24, 44, 78, 139	0
6	F	232/233 (99%)	-0.23	3 (1%) 77 79	25, 45, 70, 99	0
6	T	232/233 (99%)	-0.10	7 (3%) 50 53	25, 47, 75, 96	0
7	G	244/244 (100%)	-0.41	2 (0%) 86 87	22, 40, 71, 85	0
7	U	244/244 (100%)	-0.30	3 (1%) 79 80	22, 40, 72, 83	0
8	H	196/196 (100%)	-0.69	0 100 100	19, 27, 46, 64	0
8	V	196/196 (100%)	-0.60	0 100 100	19, 28, 48, 63	0
9	I	222/222 (100%)	-0.54	2 (0%) 84 86	18, 31, 52, 105	0
9	W	222/222 (100%)	-0.57	2 (0%) 84 86	21, 33, 51, 100	0
10	J	204/204 (100%)	-0.56	0 100 100	17, 31, 51, 71	0
10	X	204/204 (100%)	-0.51	1 (0%) 91 91	17, 30, 52, 73	0
11	K	196/198 (98%)	-0.69	1 (0%) 91 91	18, 31, 48, 83	0
11	Y	196/198 (98%)	-0.63	1 (0%) 91 91	17, 31, 48, 85	0
12	L	212/212 (100%)	-0.66	1 (0%) 91 91	18, 28, 44, 64	0
12	Z	212/212 (100%)	-0.61	1 (0%) 91 91	20, 31, 50, 69	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	1	222/222 (100%)	-0.56	1 (0%) 91 91	18, 32, 58, 74	0
13	M	222/222 (100%)	-0.65	1 (0%) 91 91	18, 31, 52, 71	0
14	2	233/233 (100%)	-0.65	1 (0%) 92 93	18, 29, 44, 59	0
14	N	233/233 (100%)	-0.62	1 (0%) 92 93	17, 30, 48, 58	0
All	All	6356/6364 (99%)	-0.46	70 (1%) 80 82	17, 35, 68, 139	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	127	ALA	9.1
5	S	129	GLY	8.7
5	S	128	SER	7.5
2	B	1	MET	6.8
7	U	247	ASN	6.3

## 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](#)

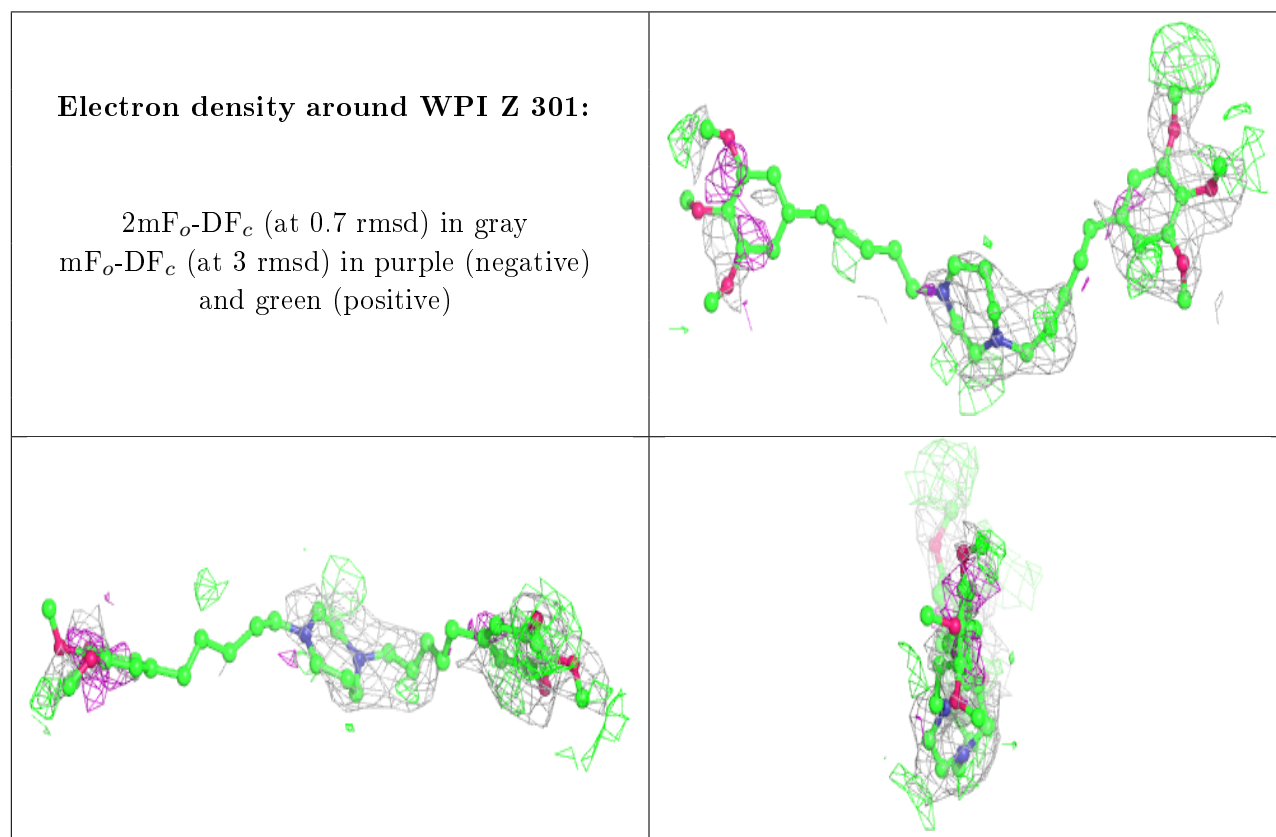
There are no carbohydrates in this entry.

## 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.

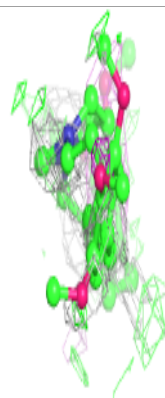
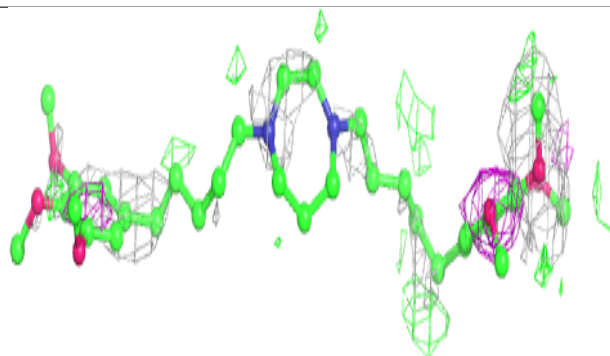
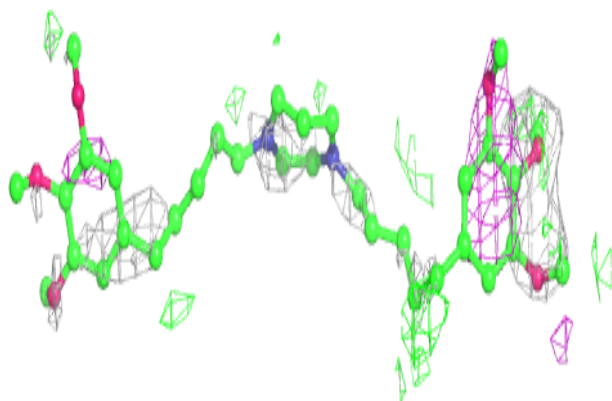
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
15	WPI	Z	301	41/41	0.61	0.50	36,76,123,131	0
15	WPI	W	301	41/41	0.62	0.63	46,99,128,133	0
15	WPI	I	301	41/41	0.62	0.53	46,94,124,133	0
15	WPI	L	301	41/41	0.65	0.41	38,67,111,118	0
15	WPI	H	300	41/41	0.69	0.58	40,78,102,107	0
15	WPI	V	201	41/41	0.77	0.39	38,67,81,90	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.

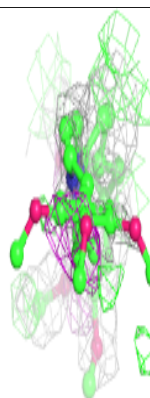
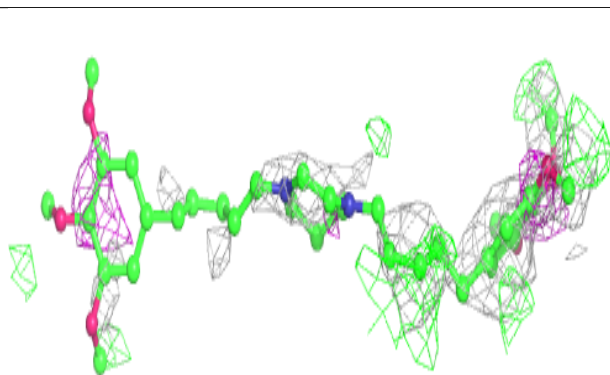
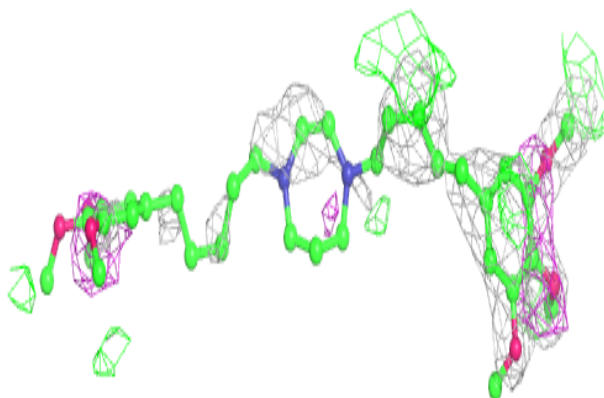


**Electron density around WPI W 301:**

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

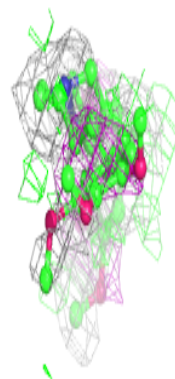
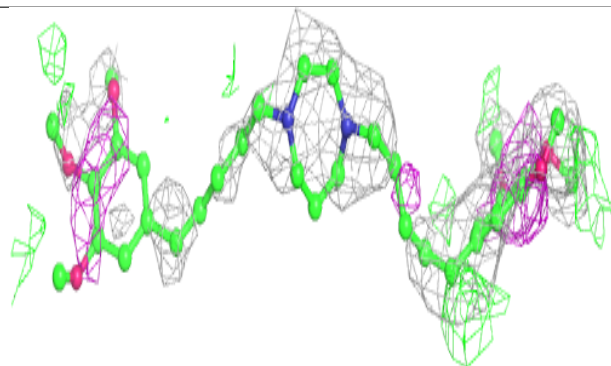
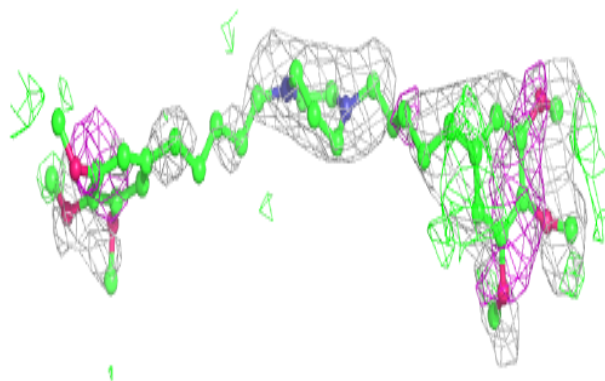
**Electron density around WPI I 301:**

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

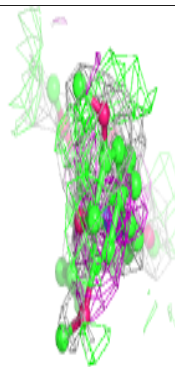
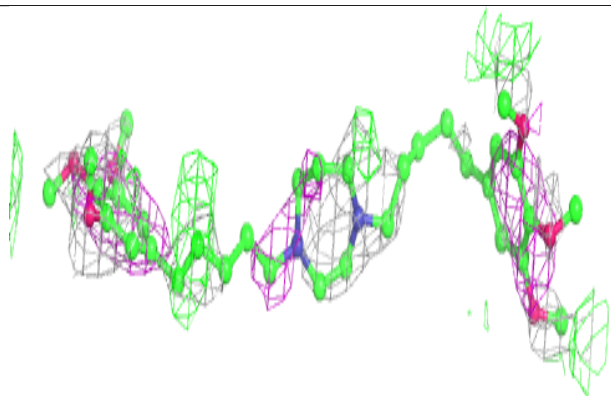
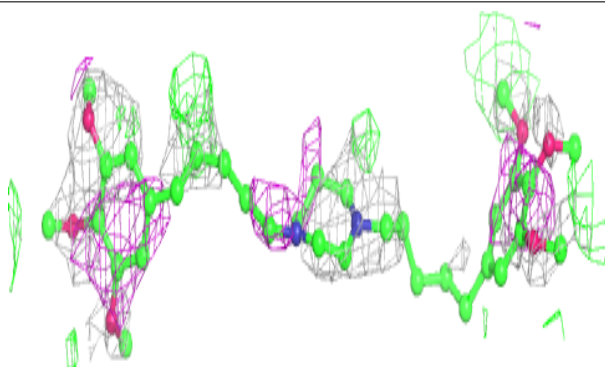


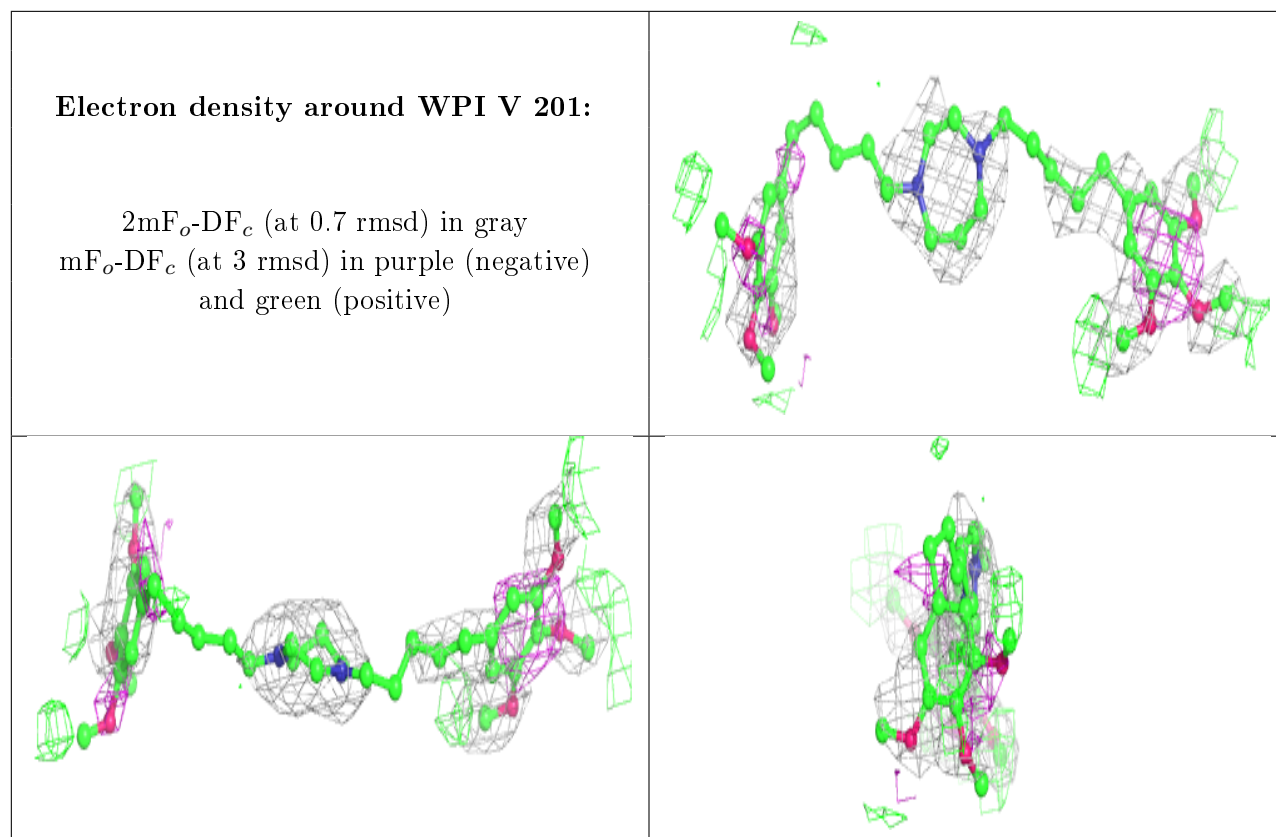
**Electron density around WPI L 301:**

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

**Electron density around WPI H 300:**

$2mF_o-DF_c$  (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.