



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:24 PM GMT

PDB ID : 4HRC  
Title : Crystal structure of yeast 20S proteasome in complex with epoxyketone carma-  
phyacin analogue 3  
Authors : Trivella, D.B.B.; Stein, M.; Groll, M.  
Deposited on : 2012-10-27  
Resolution : 2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

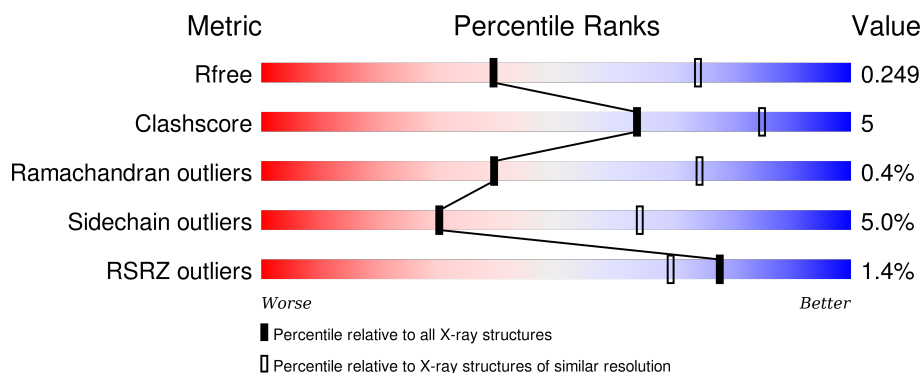
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	249	<div> <div>2%</div> <div>88% 12%</div> </div>
1	O	249	<div> <div>2%</div> <div>88% 12%</div> </div>
2	B	244	<div> <div>2%</div> <div>77% 18%</div> </div>
2	P	244	<div> <div>2%</div> <div>84% 15%</div> </div>
3	C	241	<div> <div>2%</div> <div>82% 16%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	Q	241	
4	D	242	
4	R	242	
5	E	233	
5	S	233	
6	F	244	
6	T	244	
7	G	243	
7	U	243	
8	H	222	
8	V	222	
9	I	204	
9	W	204	
10	J	198	
10	X	198	
11	K	212	
11	Y	212	
12	L	222	
12	Z	222	
13	M	233	
13	a	233	
14	N	196	
14	b	196	

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	OV2	N	301	-	-	-	X

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 50159 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 Y7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	249	Total	C	N	O	S	0	0	0
			1907	1214	314	376	3			
1	O	249	Total	C	N	O	S	0	0	0
			1907	1214	314	376	3			

- Molecule 2 is a protein called Proteasome component Y13.

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

- Molecule 3 is a protein called Proteasome component PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	241	Total	C	N	O	S	0	0	0
			1890	1181	331	374	4			
3	Q	241	Total	C	N	O	S	0	0	0
			1890	1181	331	374	4			

- Molecule 4 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	242	Total	C	N	O	S	0	0	0
			1861	1162	314	378	7			
4	R	242	Total	C	N	O	S	0	0	0
			1861	1162	314	378	7			

- Molecule 5 is a protein called Proteasome component PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	233	Total	C	N	O	S	0	0	0
			1795	1129	312	350	4			
5	S	233	Total	C	N	O	S	0	0	0
			1795	1129	312	350	4			

- Molecule 6 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	244	Total	C	N	O	S	0	0	0
			1896	1205	330	357	4			
6	T	244	Total	C	N	O	S	0	0	0
			1896	1205	330	357	4			

- Molecule 7 is a protein called Proteasome component C7-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			
7	U	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			

- Molecule 8 is a protein called Proteasome component PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	222	Total	C	N	O	S	0	0	0
			1684	1061	293	323	7			
8	V	222	Total	C	N	O	S	0	0	0
			1684	1061	293	323	7			

- Molecule 9 is a protein called Proteasome component PUP3.

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 component C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			

- Molecule 11 is a protein called Proteasome component PRE2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			
11	Y	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			

- Molecule 12 is a protein called Proteasome component C5.

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			

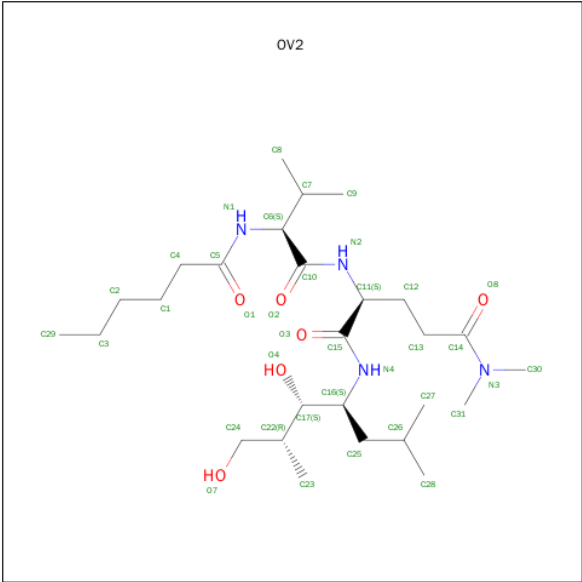
- Molecule 13 is a protein called Proteasome component PRE4.

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 component PRE3.

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 N-HEXANOYL-L-VALYL-N 1 -[(2R,3S,4S)-1,3-DIHYDROXY-2,6-DIMETHYLHEPTAN-4-YL]-N 5 ,N 5 -DIMETHYL-L-GLUTAMAMIDE (three-letter code: OV2) (formula: C<sub>27</sub>H<sub>52</sub>N<sub>4</sub>O<sub>6</sub>).





*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	G	18	Total O 18 18	0	0
16	H	12	Total O 12 12	0	0
16	I	19	Total O 19 19	0	0
16	J	19	Total O 19 19	0	0
16	K	13	Total O 13 13	0	0
16	L	22	Total O 22 22	0	0
16	M	22	Total O 22 22	0	0
16	N	16	Total O 16 16	0	0
16	O	19	Total O 19 19	0	0
16	P	4	Total O 4 4	0	0
16	Q	13	Total O 13 13	0	0
16	R	8	Total O 8 8	0	0
16	S	8	Total O 8 8	0	0
16	T	11	Total O 11 11	0	0
16	U	9	Total O 9 9	0	0
16	V	19	Total O 19 19	0	0
16	W	13	Total O 13 13	0	0
16	X	15	Total O 15 15	0	0
16	Y	15	Total O 15 15	0	0
16	Z	18	Total O 18 18	0	0
16	a	11	Total O 11 11	0	0

*Continued on next page...*

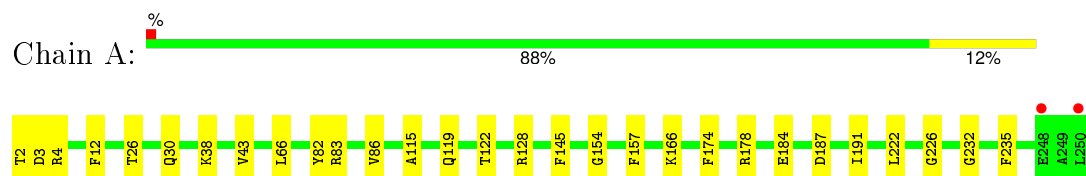
*Continued from previous page...*

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

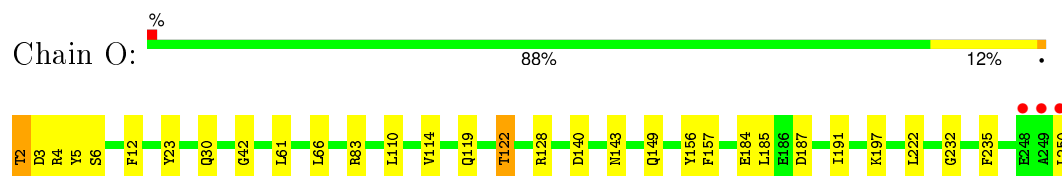
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

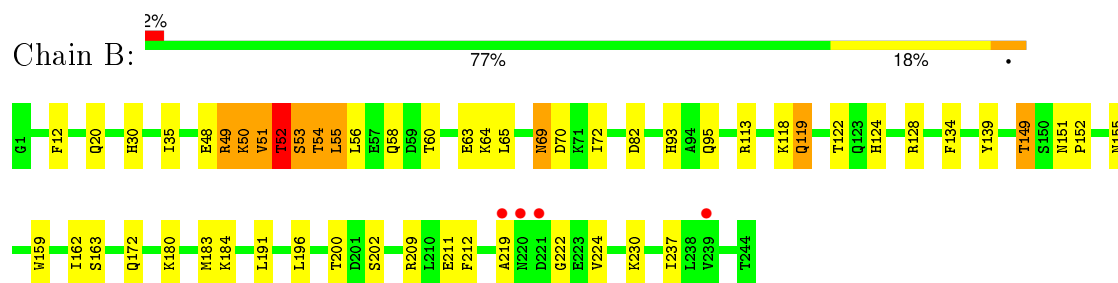
- Molecule 1: Proteasome component Y7



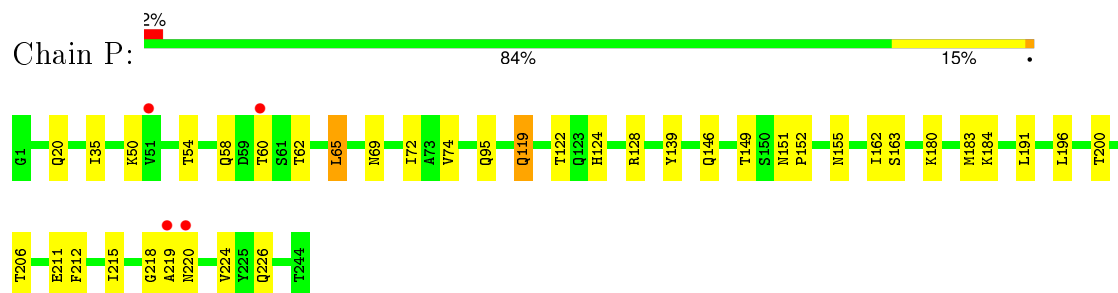
- Molecule 1: Proteasome component Y7



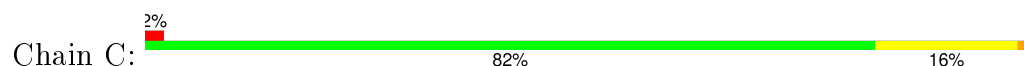
- Molecule 2: Proteasome component Y13

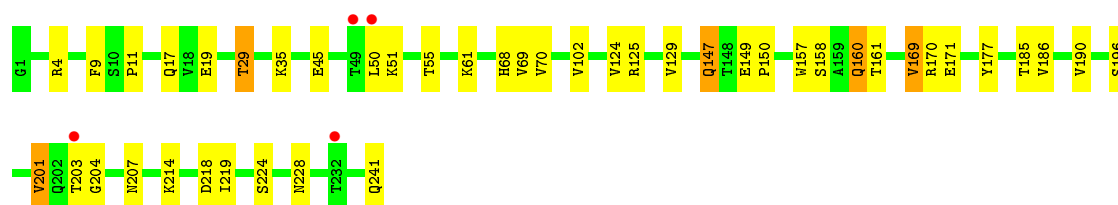


- Molecule 2: Proteasome component Y13

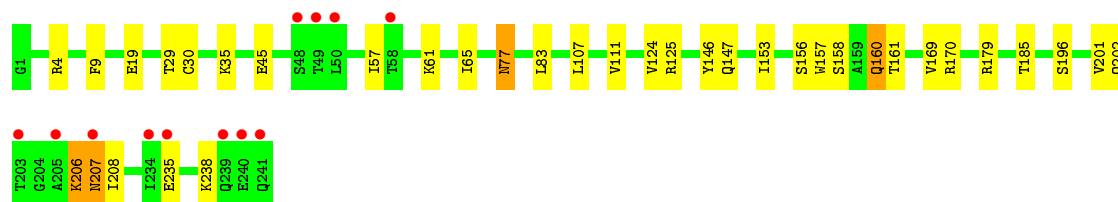
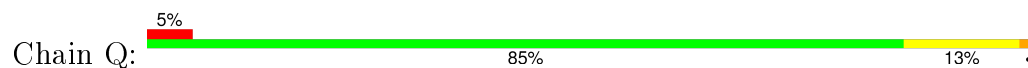


- Molecule 3: Proteasome component PRE6

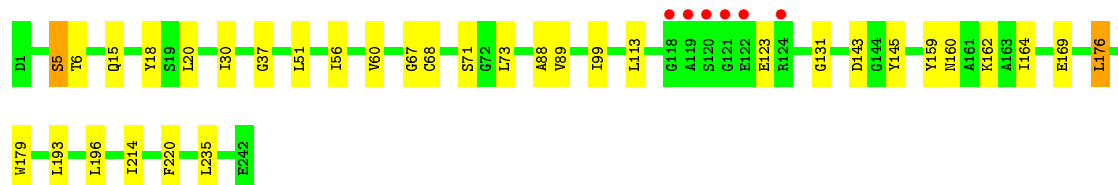
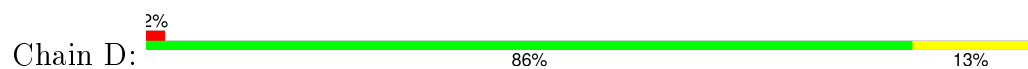




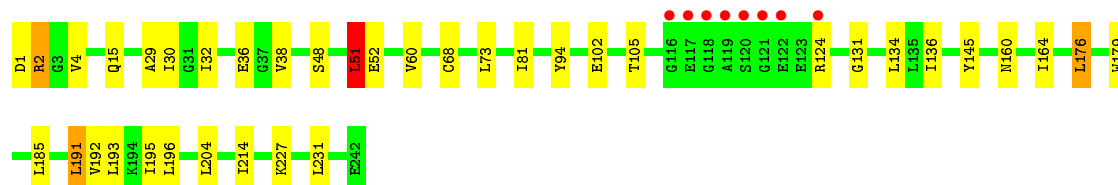
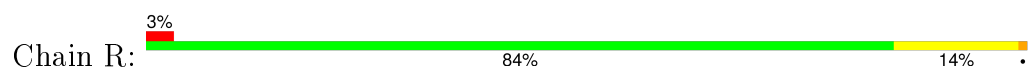
• Molecule 3: Proteasome component PRE6



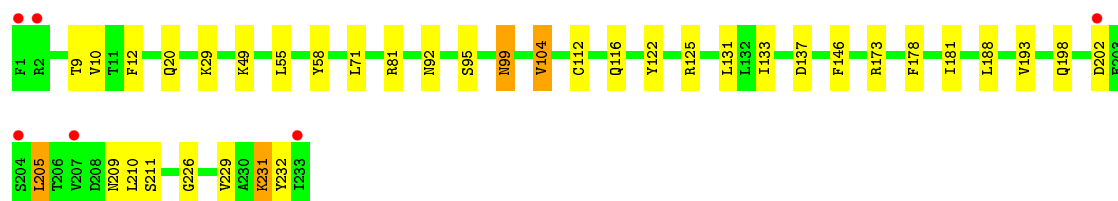
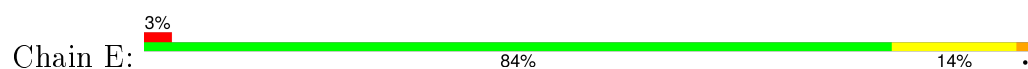
• Molecule 4: Proteasome component PUP2



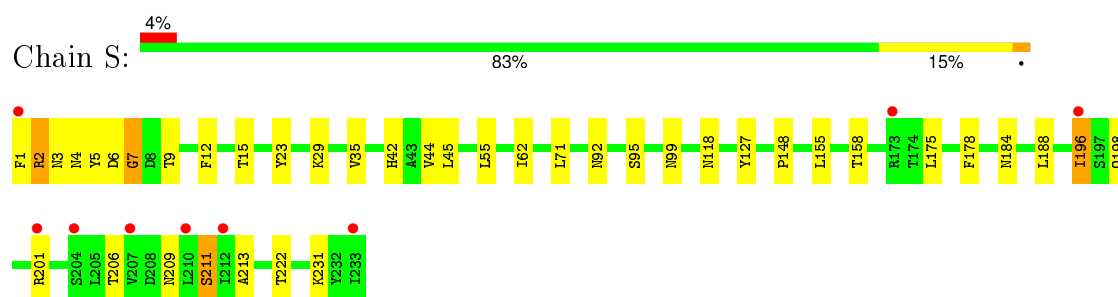
• Molecule 4: Proteasome component PUP2



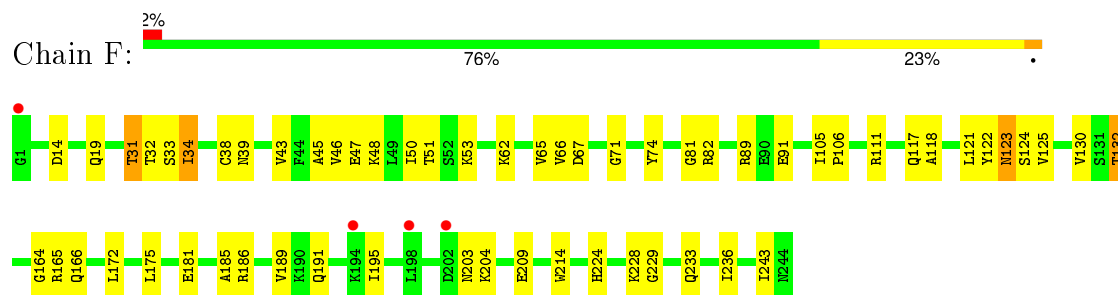
• Molecule 5: Proteasome component PRE5



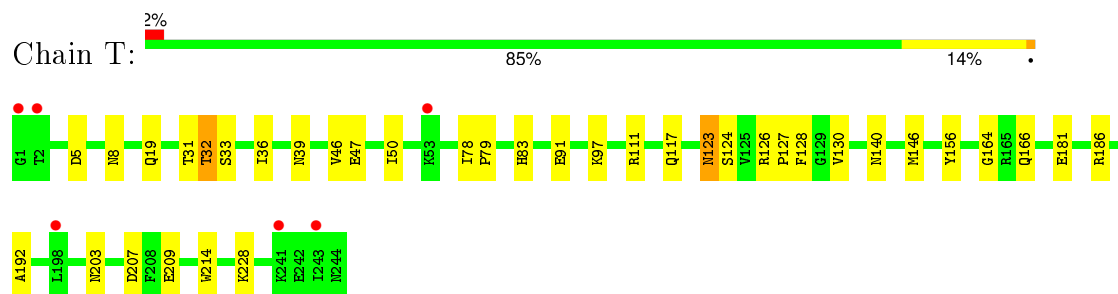
• Molecule 5: Proteasome component PRE5



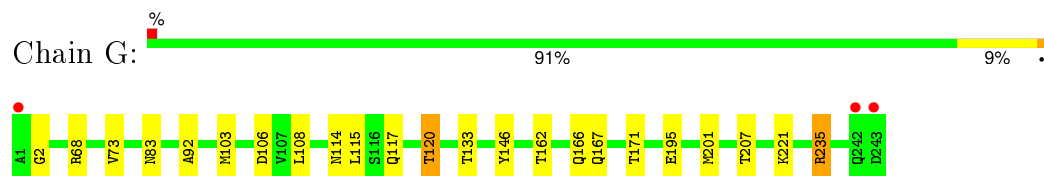
- Molecule 6: Proteasome component C1



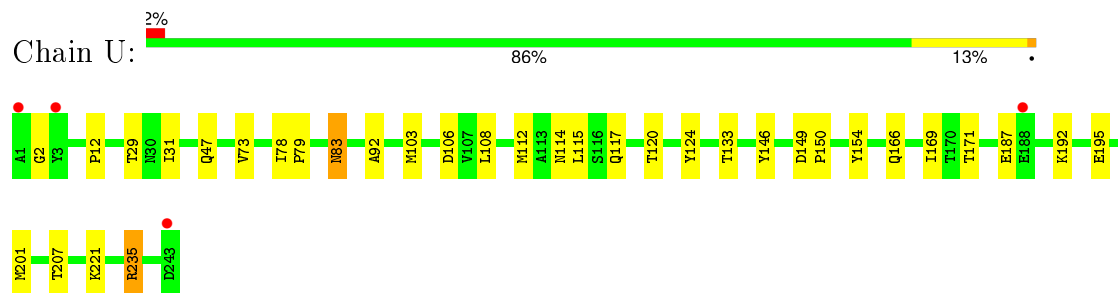
- Molecule 6: Proteasome component C1




- Molecule 7: Proteasome component C7-alpha

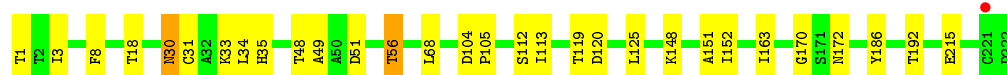


- Molecule 7: Proteasome component C7-alpha




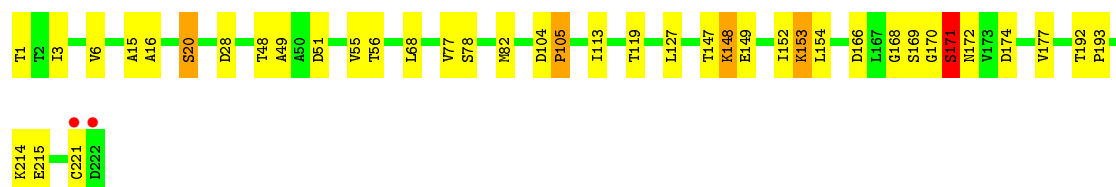
- Molecule 8: Proteasome component PUP1

Chain H:  86% 13%




- Molecule 8: Proteasome component PUP1

Chain V:  82% 16%



- Molecule 9: Proteasome component PUP3

Chain I:  87% 13%




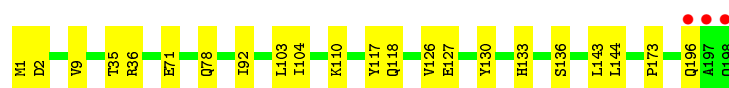
- Molecule 9: Proteasome component PUP3

Chain W:  86% 14%




- Molecule 10: Proteasome component C11

Chain J:  89% 11%




- Molecule 10: Proteasome component C11

Chain X:  90% 9%



- Molecule 11: Proteasome component PRE2

Chain K:  83% 15%





- Molecule 11: Proteasome component PRE2

Chain Y: 85% 14%



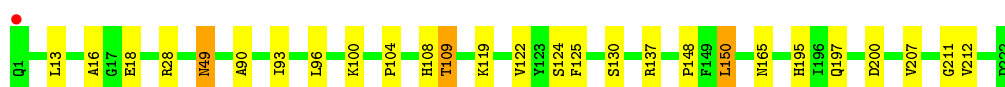
- Molecule 12: Proteasome component C5

Chain L: 87% 12%



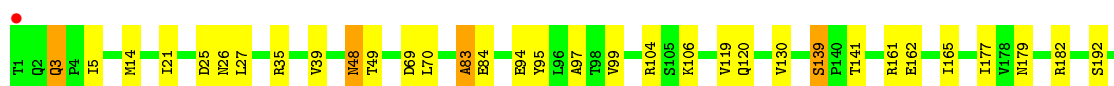
- Molecule 12: Proteasome component C5

Chain Z: 88% 11%



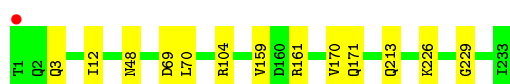
- Molecule 13: Proteasome component PRE4

Chain M: 84% 15%



- Molecule 13: Proteasome component PRE4

Chain a: 94% 6%



- Molecule 14: Proteasome component PRE3

Chain N: 86% 13%



- Molecule 14: Proteasome component PRE3

Chain b:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.35Å 300.88Å 144.25Å 90.00° 112.74° 90.00°	Depositor
Resolution (Å)	48.00 – 2.80 47.49 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.0 (48.00-2.80) 98.0 (47.49-2.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.67 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.210 , 0.243 0.215 , 0.249	Depositor DCC
$R_{free}$ test set	12670 reflections (5.23%)	DCC
Wilson B-factor (Å <sup>2</sup> )	63.7	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 44.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 255047 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	50159	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: OV2

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	0.34	0/1944	0.51	0/2632
1	O	0.36	0/1944	0.50	0/2632
2	B	0.36	0/1934	0.53	0/2618
2	P	0.33	0/1934	0.51	0/2618
3	C	0.35	0/1919	0.53	0/2598
3	Q	0.33	0/1919	0.52	0/2598
4	D	0.34	0/1886	0.54	0/2541
4	R	0.35	0/1886	0.54	1/2541 (0.0%)
5	E	0.34	0/1823	0.53	0/2463
5	S	0.40	0/1823	0.52	0/2463
6	F	0.36	0/1936	0.49	0/2614
6	T	0.34	0/1936	0.49	0/2614
7	G	0.34	0/1959	0.50	0/2652
7	U	0.34	0/1959	0.49	0/2652
8	H	0.33	0/1715	0.51	0/2326
8	V	0.37	0/1715	0.54	0/2326
9	I	0.36	0/1611	0.53	0/2174
9	W	0.37	0/1611	0.53	0/2174
10	J	0.33	0/1613	0.54	0/2173
10	X	0.34	0/1613	0.50	0/2173
11	K	0.35	0/1681	0.52	0/2274
11	Y	0.36	0/1681	0.52	0/2274
12	L	0.36	0/1795	0.53	0/2420
12	Z	0.36	0/1795	0.52	0/2420
13	M	0.34	0/1855	0.54	0/2514
13	a	0.34	0/1855	0.52	0/2514
14	N	0.33	0/1541	0.47	0/2087
14	b	0.32	0/1541	0.49	0/2087
All	All	0.35	0/50424	0.52	1/68172 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying

if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
8	V	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	R	51	LEU	CA-CB-CG	5.32	127.53	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	50	LYS	Peptide
8	V	1	THR	Peptide

## 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	1907	0	1917	20	0
1	O	1907	0	1917	24	0
2	B	1904	0	1904	47	0
2	P	1904	0	1904	19	0
3	C	1890	0	1903	28	0
3	Q	1890	0	1903	25	0
4	D	1861	0	1839	18	0
4	R	1861	0	1839	18	0
5	E	1795	0	1800	23	0
5	S	1795	0	1800	34	0
6	F	1896	0	1889	30	0
6	T	1896	0	1889	16	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	1921	0	1913	13	0
7	U	1921	0	1913	23	0
8	H	1684	0	1685	19	0
8	V	1684	0	1685	26	0
9	I	1581	0	1574	14	0
9	W	1581	0	1574	15	0
10	J	1585	0	1590	7	0
10	X	1585	0	1590	9	0
11	K	1644	0	1592	24	0
11	Y	1644	0	1592	25	0
12	L	1757	0	1711	19	0
12	Z	1757	0	1711	14	0
13	M	1824	0	1832	19	0
13	a	1824	0	1832	0	0
14	N	1512	0	1478	13	0
14	b	1512	0	1478	0	0
15	H	37	0	50	2	0
15	K	37	0	50	3	0
15	N	37	0	50	0	0
15	V	37	0	50	5	0
15	Y	37	0	50	1	0
15	b	37	0	50	0	0
16	A	10	0	0	0	0
16	B	20	0	0	1	0
16	C	13	0	0	0	0
16	D	17	0	0	1	0
16	E	18	0	0	0	0
16	F	15	0	0	1	0
16	G	18	0	0	0	0
16	H	12	0	0	0	0
16	I	19	0	0	1	0
16	J	19	0	0	0	0
16	K	13	0	0	0	0
16	L	22	0	0	0	0
16	M	22	0	0	2	0
16	N	16	0	0	2	0
16	O	19	0	0	1	0
16	P	4	0	0	0	0
16	Q	13	0	0	1	0
16	R	8	0	0	0	0
16	S	8	0	0	0	0
16	T	11	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	U	9	0	0	0	0
16	V	19	0	0	1	0
16	W	13	0	0	0	0
16	X	15	0	0	0	0
16	Y	15	0	0	0	0
16	Z	18	0	0	0	0
16	a	11	0	0	0	0
16	b	18	0	0	0	0
All	All	50159	0	49554	480	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 5.

The worst 5 of 480 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)
2:B:51:VAL:C	2:B:52:THR:HG23	1.59	1.20
11:Y:4:LEU:C	11:Y:4:LEU:HD23	1.61	1.14
5:S:2:ARG:HH11	5:S:2:ARG:HG3	0.95	1.04
5:S:2:ARG:HH11	5:S:2:ARG:CG	1.71	1.04
2:B:51:VAL:O	2:B:52:THR:HG23	1.56	1.04

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	247/249 (99%)	240 (97%)	5 (2%)	2 (1%)	24	58
1	O	247/249 (99%)	240 (97%)	7 (3%)	0	100	100
2	B	242/244 (99%)	228 (94%)	12 (5%)	2 (1%)	24	58

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	P	242/244 (99%)	228 (94%)	12 (5%)	2 (1%)	24	58
3	C	239/241 (99%)	231 (97%)	7 (3%)	1 (0%)	39	74
3	Q	239/241 (99%)	230 (96%)	7 (3%)	2 (1%)	24	58
4	D	240/242 (99%)	233 (97%)	6 (2%)	1 (0%)	39	74
4	R	240/242 (99%)	222 (92%)	17 (7%)	1 (0%)	39	74
5	E	231/233 (99%)	219 (95%)	12 (5%)	0	100	100
5	S	231/233 (99%)	217 (94%)	13 (6%)	1 (0%)	39	74
6	F	242/244 (99%)	229 (95%)	10 (4%)	3 (1%)	16	47
6	T	242/244 (99%)	230 (95%)	11 (4%)	1 (0%)	39	74
7	G	241/243 (99%)	231 (96%)	9 (4%)	1 (0%)	39	74
7	U	241/243 (99%)	236 (98%)	4 (2%)	1 (0%)	39	74
8	H	220/222 (99%)	213 (97%)	7 (3%)	0	100	100
8	V	220/222 (99%)	207 (94%)	9 (4%)	4 (2%)	11	34
9	I	202/204 (99%)	190 (94%)	11 (5%)	1 (0%)	34	69
9	W	202/204 (99%)	193 (96%)	9 (4%)	0	100	100
10	J	196/198 (99%)	184 (94%)	11 (6%)	1 (0%)	34	69
10	X	196/198 (99%)	186 (95%)	10 (5%)	0	100	100
11	K	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
11	Y	210/212 (99%)	207 (99%)	3 (1%)	0	100	100
12	L	220/222 (99%)	213 (97%)	7 (3%)	0	100	100
12	Z	220/222 (99%)	209 (95%)	11 (5%)	0	100	100
13	M	231/233 (99%)	219 (95%)	10 (4%)	2 (1%)	21	55
13	a	231/233 (99%)	225 (97%)	5 (2%)	1 (0%)	39	74
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
14	b	194/196 (99%)	182 (94%)	12 (6%)	0	100	100
All	All	6310/6366 (99%)	6037 (96%)	246 (4%)	27 (0%)	39	74

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	123	GLU
5	S	7	GLY
8	V	148	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	R	2	ARG
8	V	171	SER

### 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	208/208 (100%)	205 (99%)	3 (1%)	74	94
1	O	208/208 (100%)	198 (95%)	10 (5%)	31	66
2	B	203/203 (100%)	185 (91%)	18 (9%)	12	34
2	P	203/203 (100%)	189 (93%)	14 (7%)	19	48
3	C	213/213 (100%)	198 (93%)	15 (7%)	19	47
3	Q	213/213 (100%)	205 (96%)	8 (4%)	40	74
4	D	198/198 (100%)	190 (96%)	8 (4%)	38	73
4	R	198/198 (100%)	188 (95%)	10 (5%)	29	63
5	E	192/192 (100%)	177 (92%)	15 (8%)	16	41
5	S	192/192 (100%)	176 (92%)	16 (8%)	14	38
6	F	201/201 (100%)	183 (91%)	18 (9%)	12	34
6	T	201/201 (100%)	188 (94%)	13 (6%)	21	52
7	G	207/207 (100%)	197 (95%)	10 (5%)	31	66
7	U	207/207 (100%)	197 (95%)	10 (5%)	31	66
8	H	181/181 (100%)	175 (97%)	6 (3%)	45	79
8	V	181/181 (100%)	174 (96%)	7 (4%)	39	74
9	I	172/172 (100%)	165 (96%)	7 (4%)	37	72
9	W	172/172 (100%)	163 (95%)	9 (5%)	29	62
10	J	175/175 (100%)	164 (94%)	11 (6%)	22	53
10	X	175/175 (100%)	169 (97%)	6 (3%)	44	78
11	K	169/169 (100%)	160 (95%)	9 (5%)	28	61
11	Y	169/169 (100%)	164 (97%)	5 (3%)	48	82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
12	L	185/185 (100%)	181 (98%)	4 (2%)	60	89
12	Z	185/185 (100%)	179 (97%)	6 (3%)	46	80
13	M	199/199 (100%)	190 (96%)	9 (4%)	34	68
13	a	199/199 (100%)	187 (94%)	12 (6%)	24	56
14	N	162/162 (100%)	157 (97%)	5 (3%)	47	81
14	b	162/162 (100%)	157 (97%)	5 (3%)	47	81
All	All	5330/5330 (100%)	5061 (95%)	269 (5%)	30	64

5 of 269 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	49	ASN
2	P	60	THR
12	Z	109	THR
13	M	48	ASN
14	N	105	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 143 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	48	ASN
2	P	155	ASN
13	a	2	GLN
13	M	171	GLN
14	N	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	OV2	H	301	8	35,36,36	0.41	0	43,47,47	1.16	2 (4%)
15	OV2	K	301	11	35,36,36	0.40	0	43,47,47	1.32	2 (4%)
15	OV2	N	301	14	35,36,36	0.42	0	43,47,47	1.35	5 (11%)
15	OV2	V	301	8	35,36,36	0.37	0	43,47,47	0.86	2 (4%)
15	OV2	Y	301	11	35,36,36	0.39	0	43,47,47	1.40	3 (6%)
15	OV2	b	301	14	35,36,36	0.38	0	43,47,47	1.33	3 (6%)

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	OV2	H	301	8	-	4/52/52/52	0/0/0/0
15	OV2	K	301	11	-	4/52/52/52	0/0/0/0
15	OV2	N	301	14	-	0/52/52/52	0/0/0/0
15	OV2	V	301	8	-	0/52/52/52	0/0/0/0
15	OV2	Y	301	11	-	0/52/52/52	0/0/0/0
15	OV2	b	301	14	-	0/52/52/52	0/0/0/0

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	Y	301	OV2	C25-C16-N4	-7.09	100.61	110.15
15	K	301	OV2	C25-C16-N4	-5.77	102.39	110.15
15	H	301	OV2	C25-C16-N4	-5.72	102.46	110.15
15	N	301	OV2	C25-C16-N4	-5.60	102.62	110.15
15	b	301	OV2	C25-C16-N4	-5.03	103.39	110.15

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	K	301	OV2	O8-C14-N3-C31
15	K	301	OV2	O8-C14-N3-C30
15	K	301	OV2	C13-C14-N3-C31
15	K	301	OV2	C13-C14-N3-C30
15	H	301	OV2	C13-C14-N3-C31

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	H	301	OV2	2	0
15	K	301	OV2	3	0
15	V	301	OV2	5	0
15	Y	301	OV2	1	0

## 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	249/249 (100%)	-0.32	2 (0%) 87 81	54, 72, 100, 117	0
1	O	249/249 (100%)	-0.34	3 (1%) 81 73	56, 74, 104, 122	0
2	B	244/244 (100%)	-0.15	4 (1%) 74 66	51, 78, 119, 137	0
2	P	244/244 (100%)	-0.18	4 (1%) 74 66	55, 75, 107, 121	0
3	C	241/241 (100%)	-0.21	4 (1%) 73 63	48, 73, 115, 147	0
3	Q	241/241 (100%)	0.11	12 (4%) 32 21	62, 91, 132, 168	0
4	D	242/242 (100%)	-0.23	6 (2%) 61 48	54, 74, 101, 122	0
4	R	242/242 (100%)	-0.09	8 (3%) 50 38	63, 84, 116, 145	0
5	E	233/233 (100%)	-0.05	6 (2%) 59 47	63, 86, 122, 132	0
5	S	233/233 (100%)	-0.04	9 (3%) 43 31	53, 87, 119, 129	0
6	F	244/244 (100%)	-0.22	4 (1%) 74 66	60, 84, 116, 144	0
6	T	244/244 (100%)	-0.15	6 (2%) 61 48	59, 79, 111, 134	0
7	G	243/243 (100%)	-0.22	3 (1%) 81 73	58, 78, 105, 126	0
7	U	243/243 (100%)	-0.27	4 (1%) 74 66	57, 73, 100, 118	0
8	H	222/222 (100%)	-0.36	1 (0%) 91 88	49, 67, 88, 96	0
8	V	222/222 (100%)	-0.33	2 (0%) 85 79	50, 66, 90, 120	0
9	I	204/204 (100%)	-0.29	0 100 100	46, 65, 92, 99	0
9	W	204/204 (100%)	-0.37	1 (0%) 91 88	44, 60, 84, 91	0
10	J	198/198 (100%)	-0.29	3 (1%) 76 68	46, 66, 94, 119	0
10	X	198/198 (100%)	-0.40	2 (1%) 84 77	50, 64, 84, 104	0
11	K	212/212 (100%)	-0.41	0 100 100	43, 62, 85, 92	0
11	Y	212/212 (100%)	-0.44	1 (0%) 91 88	53, 67, 92, 103	0
12	L	222/222 (100%)	-0.44	1 (0%) 91 88	49, 65, 87, 91	0
12	Z	222/222 (100%)	-0.36	1 (0%) 91 88	48, 67, 91, 100	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	233/233 (100%)	-0.30	2 (0%)	85	79	53, 72, 99, 116	0
13	a	233/233 (100%)	-0.47	1 (0%)	93	90	44, 66, 87, 99	0
14	N	196/196 (100%)	-0.43	1 (0%)	91	88	53, 64, 87, 93	0
14	b	196/196 (100%)	-0.37	0	100	100	51, 63, 87, 93	0
All	All	6366/6366 (100%)	-0.27	91 (1%)	78	69	43, 72, 109, 168	0

The worst 5 of 91 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	R	121	GLY	9.1
4	D	121	GLY	7.6
6	T	1	GLY	6.5
2	P	220	ASN	6.2
10	J	197	ALA	5.8

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
15	OV2	N	301	37/37	0.94	0.20	2.07	54,59,65,66	0
15	OV2	V	301	37/37	0.94	0.21	1.65	54,58,61,61	0
15	OV2	H	301	37/37	0.93	0.20	1.13	55,60,62,63	0
15	OV2	K	301	37/37	0.95	0.19	0.98	51,53,60,61	0
15	OV2	b	301	37/37	0.93	0.19	0.88	53,57,61,61	0
15	OV2	Y	301	37/37	0.95	0.17	0.51	55,57,63,64	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.