



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:14 PM GMT

PDB ID : 4R17  
Title : Ligand-induced aziridine-formation at subunit beta5 of the yeast 20S proteasome  
Authors : Dubiella, C.; Cui, H.; Gersch, M.; Brouwer, A.J.; Sieber, S.A.; Krueger, A.; Liskamp, R.; Groll, M.  
Deposited on : 2014-08-04  
Resolution : 2.10 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

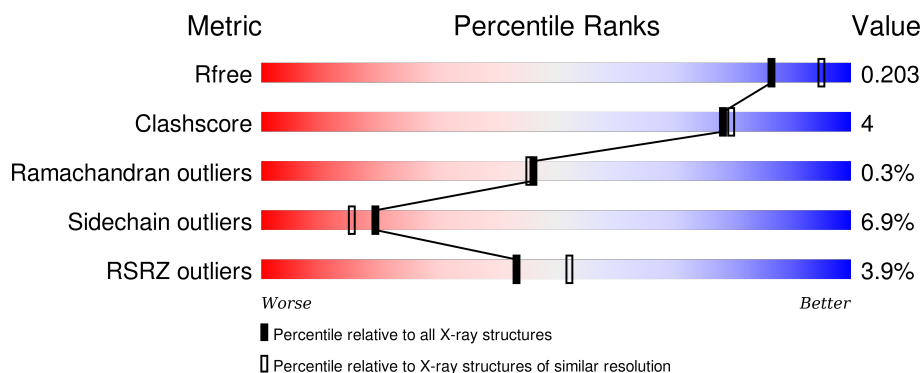
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	250	<div> <div>4%</div> <div>94%</div> <div>• •</div> </div>
1	O	250	<div> <div>4%</div> <div>94%</div> <div>5% •</div> </div>
2	B	258	<div> <div>5%</div> <div>79%</div> <div>14% • 5%</div> </div>
2	P	258	<div> <div>5%</div> <div>81%</div> <div>12% • 5%</div> </div>
3	C	254	<div> <div>8%</div> <div>79%</div> <div>11% 5% 6%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	
4	D	260	
4	R	260	
5	E	234	
5	S	234	
6	F	288	
6	T	288	
7	G	252	
7	U	252	
8	H	232	
8	V	232	
9	I	205	
9	W	205	
10	J	198	
10	X	198	
11	K	212	
11	Y	212	
12	L	222	
12	Z	222	
13	M	246	
13	a	246	
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	MG	I	302	-	-	-	X
15	MG	J	201	-	-	-	X
15	MG	K	303	-	-	-	X
16	3K4	K	301	-	-	-	X

## 2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 52831 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
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 subunit alpha type-3.

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 subunit alpha type-4.

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

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

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

- 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 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	226	Total	C	N	O	S	0	0	0
			1719	1082	298	332	7			
8	V	226	Total	C	N	O	S	0	0	0
			1719	1082	298	332	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).

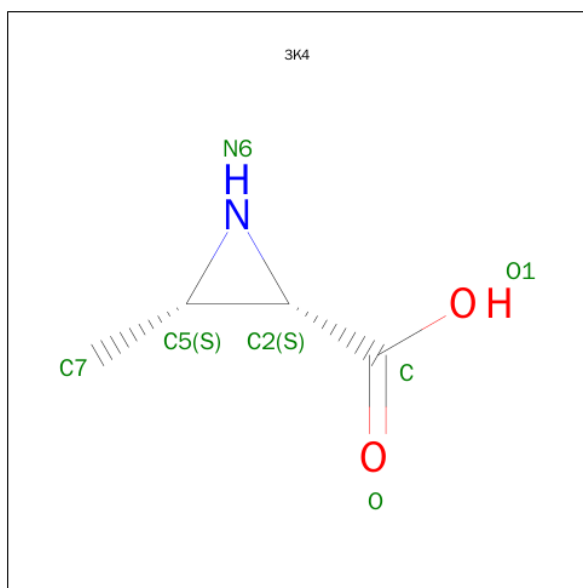
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	G	1	Total	Mg	0	0
			1	1		
15	J	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	K	2	Total	Mg	0	0
			2	2		
15	H	1	Total	Mg	0	0
			1	1		
15	I	2	Total	Mg	0	0
			2	2		
15	V	1	Total	Mg	0	0
			1	1		
15	Z	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		

- Molecule 16 is (2S,3S)-3-METHYLAZIRIDINE-2-CARBOXYLIC ACID (three-letter code: 3K4) (formula: C<sub>4</sub>H<sub>7</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
16	K	1	Total	C	N	O	0	0
			6	4	1	1		
16	Y	1	Total	C	N	O	0	0
			6	4	1	1		

- Molecule 17 is water.



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	167	Total 167	O 167	0	0
17	B	115	Total 115	O 115	0	0
17	C	105	Total 105	O 105	0	0
17	D	123	Total 123	O 123	0	0
17	E	69	Total 69	O 69	0	0
17	F	117	Total 117	O 117	0	0
17	G	158	Total 158	O 158	0	0
17	H	135	Total 135	O 135	0	0
17	I	140	Total 140	O 140	0	0
17	J	123	Total 123	O 123	0	0
17	K	152	Total 152	O 152	0	0
17	L	158	Total 158	O 158	0	0
17	M	168	Total 168	O 168	0	0
17	N	112	Total 112	O 112	0	0
17	O	118	Total 118	O 118	0	0
17	P	91	Total 91	O 91	0	0
17	Q	79	Total 79	O 79	0	0
17	R	107	Total 107	O 107	0	0
17	S	55	Total 55	O 55	0	0
17	T	94	Total 94	O 94	0	0
17	U	133	Total 133	O 133	0	0
17	V	111	Total 111	O 111	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	W	128	Total 128	O 128	0	0
17	X	118	Total 118	O 118	0	0
17	Y	148	Total 148	O 148	0	0
17	Z	166	Total 166	O 166	0	0
17	a	165	Total 165	O 165	0	0
17	b	117	Total 117	O 117	0	0

### 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 ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

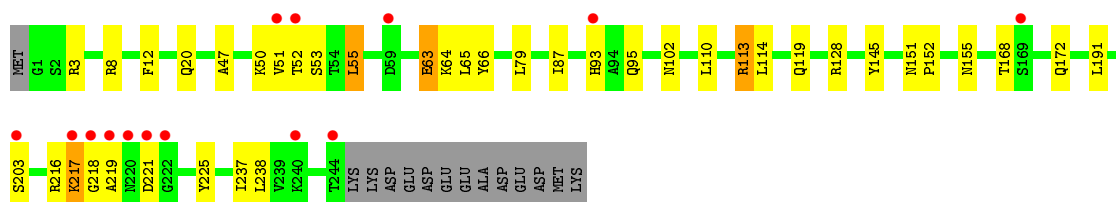
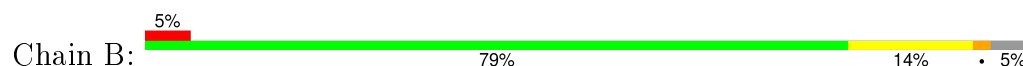
- Molecule 1: Proteasome subunit alpha type-2



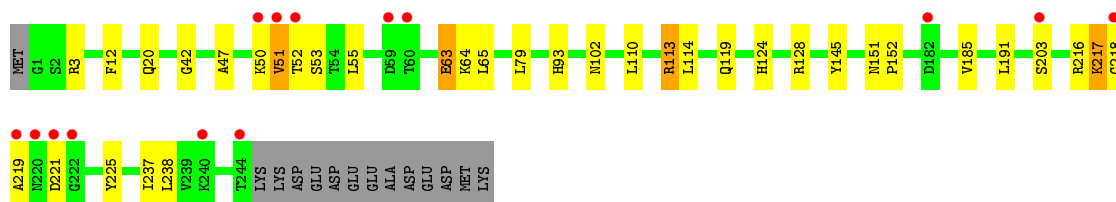
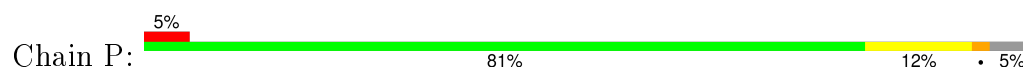
- Molecule 1: Proteasome subunit alpha type-2



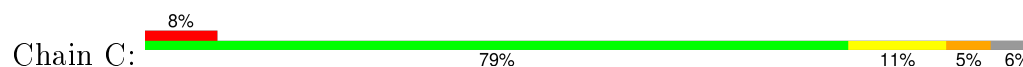
- Molecule 2: Proteasome subunit alpha type-3

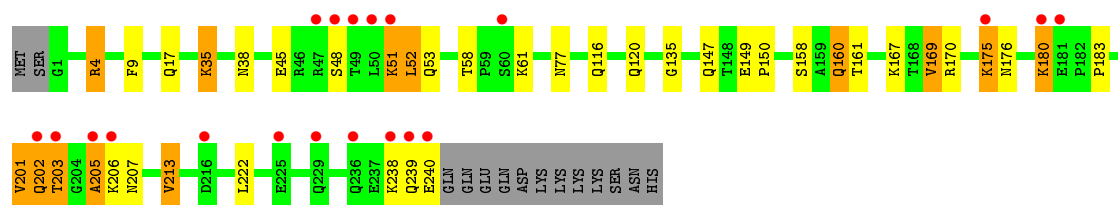


- Molecule 2: Proteasome subunit alpha type-3

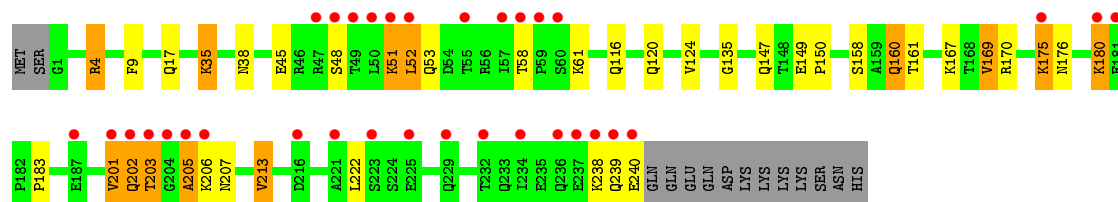
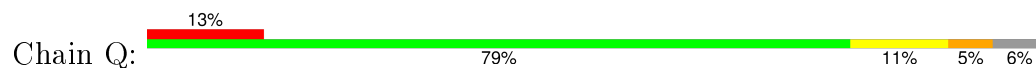


- Molecule 3: Proteasome subunit alpha type-4

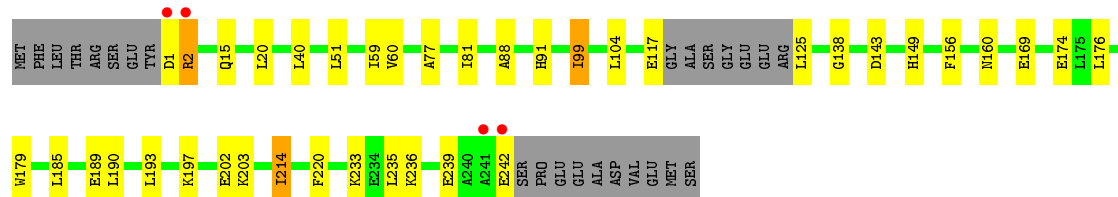
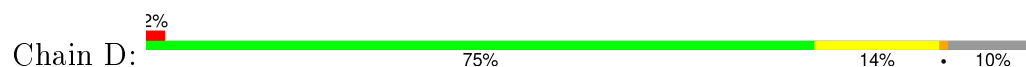




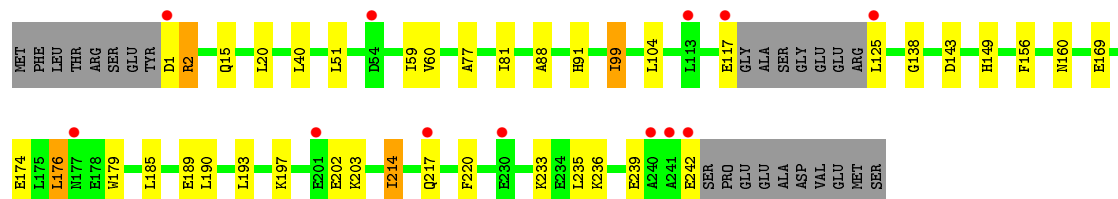
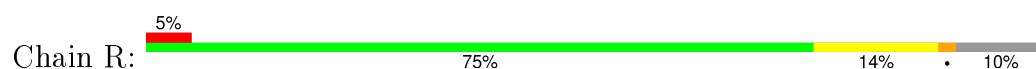
- Molecule 3: Proteasome subunit alpha type-4



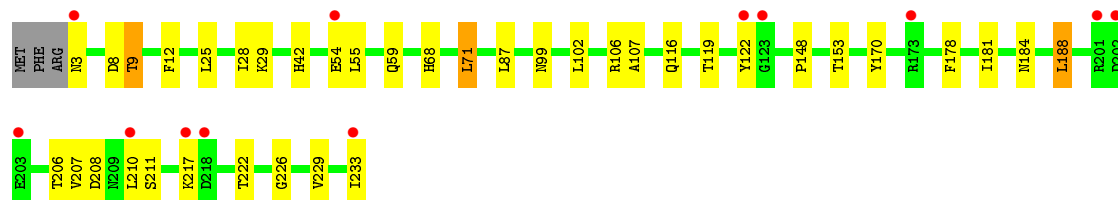
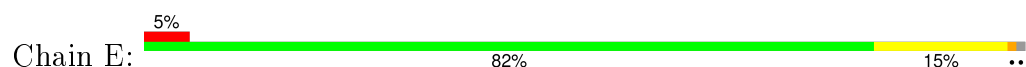
- Molecule 4: Proteasome subunit alpha type-5



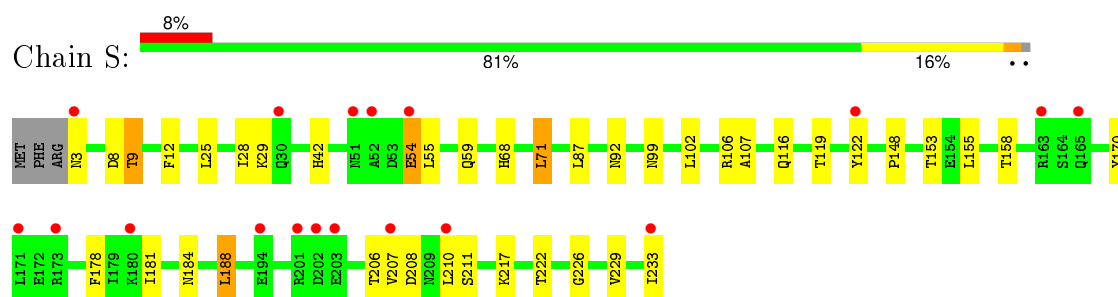
- Molecule 4: Proteasome subunit alpha type-5



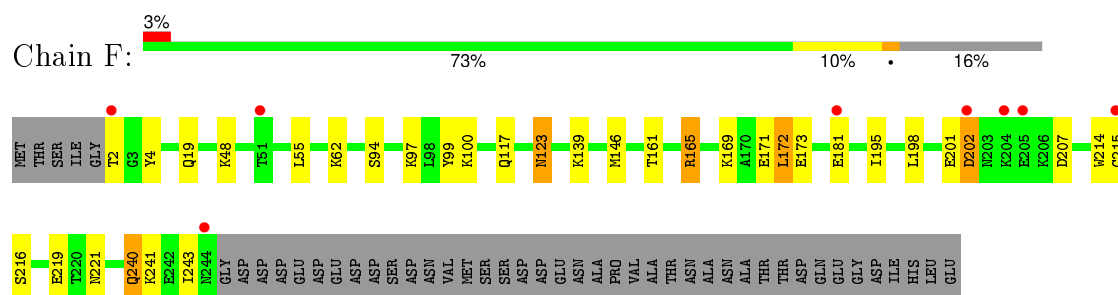
- Molecule 5: Proteasome subunit alpha type-6



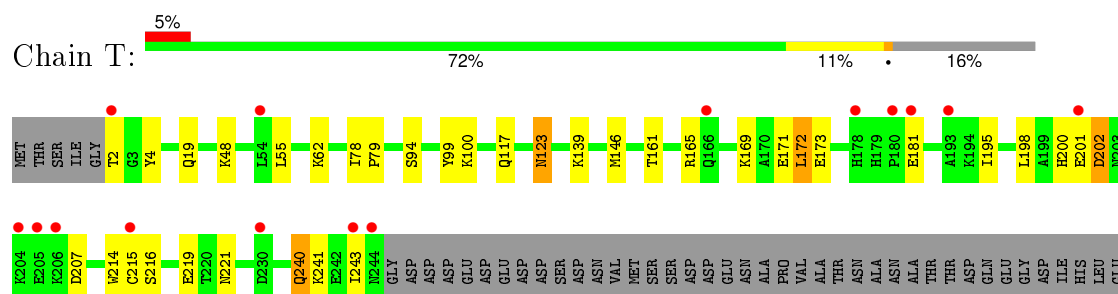
- Molecule 5: Proteasome subunit alpha type-6



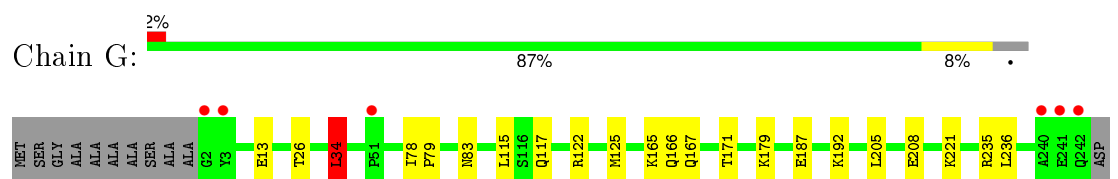
- Molecule 6: Proteasome subunit alpha type-7



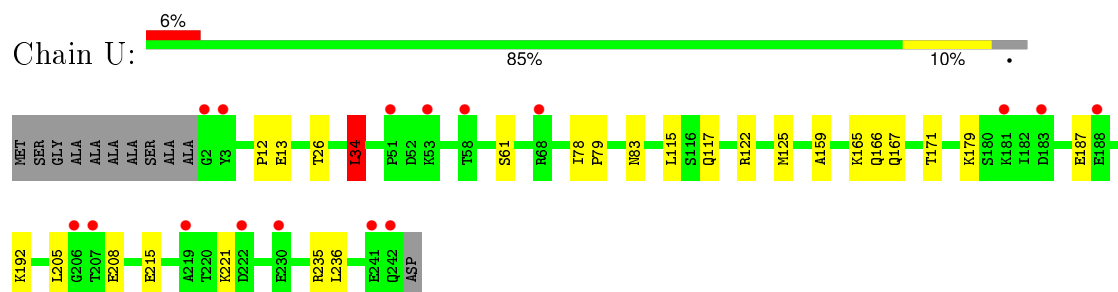
- Molecule 6: Proteasome subunit alpha type-7



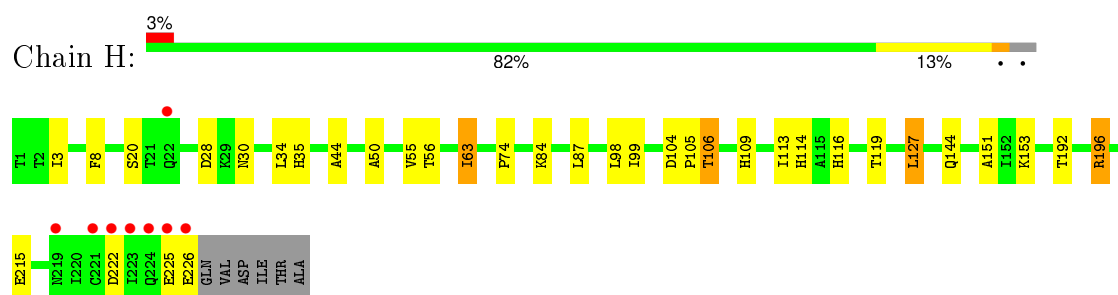
- Molecule 7: Proteasome subunit alpha type-1



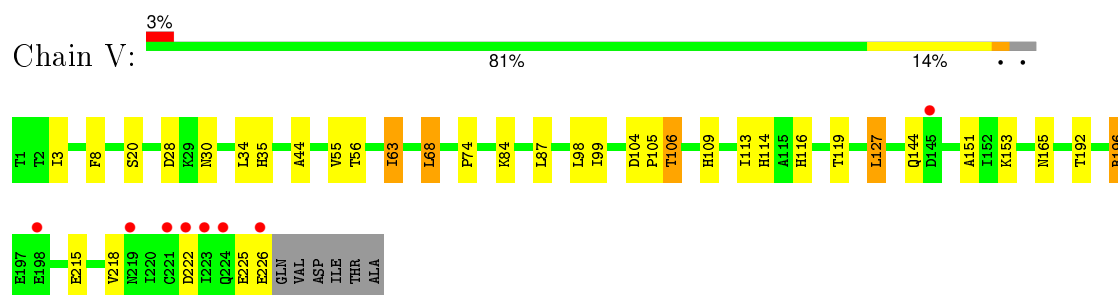
- Molecule 7: Proteasome subunit alpha type-1



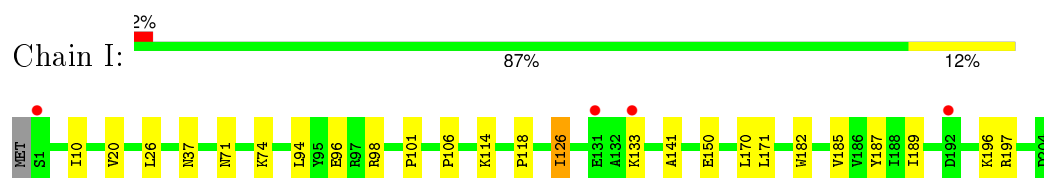
- Molecule 8: Proteasome subunit beta type-2



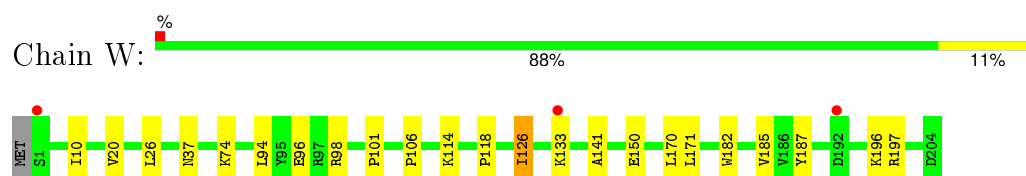
- Molecule 8: Proteasome subunit beta type-2



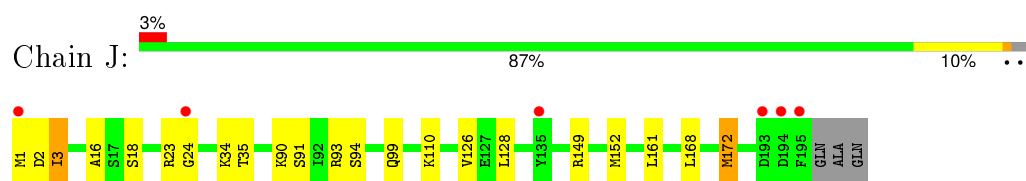
- Molecule 9: Proteasome subunit beta type-3



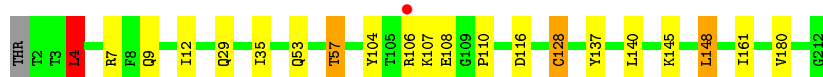
- Molecule 9: Proteasome subunit beta type-3



- Molecule 10: Proteasome subunit beta type-4



Chain K:  90% 8% .




- Molecule 11: Proteasome subunit beta type-5

Chain Y:  90% 8% .




- Molecule 12: Proteasome subunit beta type-6

Chain L:  89% 9% .




- Molecule 12: Proteasome subunit beta type-6

Chain Z:  89% 9% .




- Molecule 13: Proteasome subunit beta type-7

Chain M:  87% 6% • 5%



- Molecule 13: Proteasome subunit beta type-7

Chain a:  90% • • 5%

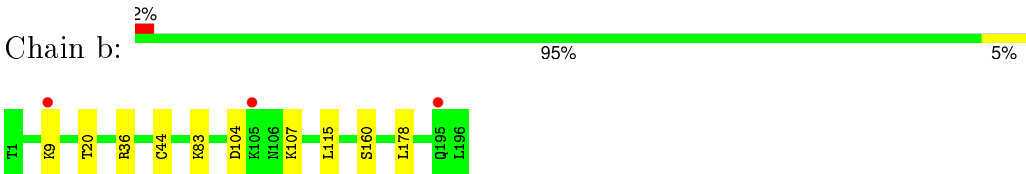


- Molecule 14: Proteasome subunit beta type-1

Chain N:  90% 8% .



- Molecule 14: Proteasome subunit beta type-1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.72Å 301.36Å 144.74Å 90.00° 112.81° 90.00°	Depositor
Resolution (Å)	15.00 – 2.10 15.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (15.00-2.10) 98.9 (15.00-2.10)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.04 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.185 , 0.198 0.191 , 0.203	Depositor DCC
$R_{free}$ test set	30359 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.1	Xtriage
Anisotropy	0.086	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 47.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 607186 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	52831	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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: 3K4, 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	0.29	0/1944	0.54	0/2632
1	O	0.28	0/1944	0.54	0/2632
2	B	0.30	0/1934	0.58	0/2618
2	P	0.31	0/1934	0.57	0/2618
3	C	0.29	0/1910	0.59	0/2586
3	Q	0.29	0/1910	0.59	0/2586
4	D	0.28	0/1837	0.57	0/2475
4	R	0.28	0/1837	0.57	0/2475
5	E	0.28	0/1800	0.55	1/2433 (0.0%)
5	S	0.28	0/1800	0.56	1/2433 (0.0%)
6	F	0.29	0/1932	0.54	0/2609
6	T	0.29	0/1932	0.54	0/2609
7	G	0.29	0/1945	0.56	1/2634 (0.0%)
7	U	0.29	0/1945	0.55	1/2634 (0.0%)
8	H	0.33	0/1750	0.55	0/2373
8	V	0.27	0/1750	0.56	1/2373 (0.0%)
9	I	0.28	0/1611	0.54	0/2174
9	W	0.28	0/1611	0.54	0/2174
10	J	0.27	0/1589	0.54	0/2142
10	X	0.27	0/1589	0.54	0/2142
11	K	0.27	0/1674	0.59	1/2264 (0.0%)
11	Y	0.27	0/1674	0.59	1/2264 (0.0%)
12	L	0.34	0/1795	0.55	0/2420
12	Z	0.28	0/1795	0.56	1/2420 (0.0%)
13	M	0.36	0/1855	0.66	3/2514 (0.1%)
13	a	0.34	0/1855	0.64	3/2514 (0.1%)
14	N	0.27	0/1541	0.54	0/2087
14	b	0.27	0/1541	0.54	0/2087
All	All	0.29	0/50234	0.57	14/67922 (0.0%)

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	M	104	ARG	NE-CZ-NH1	10.95	125.77	120.30
13	a	104	ARG	NE-CZ-NH1	8.21	124.41	120.30
13	a	104	ARG	NE-CZ-NH2	-8.11	116.25	120.30
13	M	104	ARG	NE-CZ-NH2	-7.98	116.31	120.30
11	Y	4	LEU	CA-CB-CG	5.83	128.70	115.30

There are no chirality outliers.

There are no planarity outliers.

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

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1907	0	1917	6	0
1	O	1907	0	1917	6	0
2	B	1904	0	1904	27	0
2	P	1904	0	1904	20	0
3	C	1881	0	1895	21	0
3	Q	1881	0	1895	20	0
4	D	1813	0	1797	14	0
4	R	1813	0	1797	15	0
5	E	1773	0	1775	13	0
5	S	1773	0	1775	17	0
6	F	1892	0	1883	20	0
6	T	1892	0	1883	21	0
7	G	1907	0	1901	7	0
7	U	1907	0	1901	10	0
8	H	1719	0	1719	25	0
8	V	1719	0	1719	27	0
9	I	1581	0	1574	15	0
9	W	1581	0	1574	13	0
10	J	1561	0	1569	14	0
10	X	1561	0	1569	13	0
11	K	1637	0	1585	9	0
11	Y	1637	0	1585	9	0
12	L	1757	0	1711	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	Z	1757	0	1711	10	0
13	M	1824	0	1832	8	0
13	a	1824	0	1832	0	0
14	N	1512	0	1481	11	0
14	b	1512	0	1481	0	0
15	G	1	0	0	0	0
15	H	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	K	2	0	0	0	0
15	N	1	0	0	0	0
15	V	1	0	0	0	0
15	Y	1	0	0	0	0
15	Z	1	0	0	0	0
16	K	6	0	5	0	0
16	Y	6	0	5	0	0
17	A	167	0	0	0	0
17	B	115	0	0	0	0
17	C	105	0	0	1	0
17	D	123	0	0	0	0
17	E	69	0	0	0	0
17	F	117	0	0	0	0
17	G	158	0	0	0	0
17	H	135	0	0	0	0
17	I	140	0	0	0	0
17	J	123	0	0	1	0
17	K	152	0	0	1	0
17	L	158	0	0	0	0
17	M	168	0	0	1	0
17	N	112	0	0	0	0
17	O	118	0	0	0	0
17	P	91	0	0	0	0
17	Q	79	0	0	1	0
17	R	107	0	0	0	0
17	S	55	0	0	0	0
17	T	94	0	0	0	0
17	U	133	0	0	0	0
17	V	111	0	0	0	0
17	W	128	0	0	0	0
17	X	118	0	0	0	0
17	Y	148	0	0	0	0
17	Z	166	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	a	165	0	0	0	0
17	b	117	0	0	0	0
All	All	52831	0	49096	336	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 336 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:52:THR:HG22	2:B:53:SER:H	1.22	1.03
2:B:52:THR:HG22	2:B:53:SER:N	1.80	0.97
14:N:35:THR:HG21	14:N:45:ARG:HE	1.36	0.90
2:P:52:THR:HG22	2:P:53:SER:N	1.88	0.89
2:P:52:THR:CG2	2:P:53:SER:N	2.39	0.86

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	247/250 (99%)	241 (98%)	5 (2%)	1 (0%)	39	37
1	O	247/250 (99%)	240 (97%)	6 (2%)	1 (0%)	39	37
2	B	242/258 (94%)	232 (96%)	9 (4%)	1 (0%)	39	37
2	P	242/258 (94%)	232 (96%)	9 (4%)	1 (0%)	39	37
3	C	238/254 (94%)	228 (96%)	6 (2%)	4 (2%)	11	5
3	Q	238/254 (94%)	228 (96%)	6 (2%)	4 (2%)	11	5
4	D	231/260 (89%)	230 (100%)	0	1 (0%)	39	37
4	R	231/260 (89%)	230 (100%)	0	1 (0%)	39	37

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	E	229/234 (98%)	218 (95%)	11 (5%)	0	100	100
5	S	229/234 (98%)	218 (95%)	11 (5%)	0	100	100
6	F	241/288 (84%)	234 (97%)	7 (3%)	0	100	100
6	T	241/288 (84%)	233 (97%)	8 (3%)	0	100	100
7	G	239/252 (95%)	238 (100%)	1 (0%)	0	100	100
7	U	239/252 (95%)	238 (100%)	1 (0%)	0	100	100
8	H	224/232 (97%)	218 (97%)	6 (3%)	0	100	100
8	V	224/232 (97%)	218 (97%)	6 (3%)	0	100	100
9	I	202/205 (98%)	197 (98%)	5 (2%)	0	100	100
9	W	202/205 (98%)	197 (98%)	5 (2%)	0	100	100
10	J	193/198 (98%)	187 (97%)	5 (3%)	1 (0%)	34	30
10	X	193/198 (98%)	187 (97%)	5 (3%)	1 (0%)	34	30
11	K	209/212 (99%)	206 (99%)	3 (1%)	0	100	100
11	Y	209/212 (99%)	206 (99%)	3 (1%)	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%)	220 (95%)	11 (5%)	0	100	100
13	a	231/246 (94%)	221 (96%)	10 (4%)	0	100	100
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
All	All	6280/6614 (95%)	6105 (97%)	159 (2%)	16 (0%)	46	45

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	202	GLN
3	C	205	ALA
3	Q	202	GLN
3	Q	205	ALA
2	B	51	VAL

### 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	208/209 (100%)	199 (96%)	9 (4%)	35	34
1	O	208/209 (100%)	199 (96%)	9 (4%)	35	34
2	B	203/216 (94%)	188 (93%)	15 (7%)	17	13
2	P	203/216 (94%)	187 (92%)	16 (8%)	15	11
3	C	212/226 (94%)	189 (89%)	23 (11%)	8	4
3	Q	212/226 (94%)	189 (89%)	23 (11%)	8	4
4	D	194/215 (90%)	175 (90%)	19 (10%)	10	6
4	R	194/215 (90%)	174 (90%)	20 (10%)	9	5
5	E	190/193 (98%)	171 (90%)	19 (10%)	9	6
5	S	190/193 (98%)	171 (90%)	19 (10%)	9	6
6	F	201/239 (84%)	185 (92%)	16 (8%)	15	11
6	T	201/239 (84%)	185 (92%)	16 (8%)	15	11
7	G	206/210 (98%)	192 (93%)	14 (7%)	20	16
7	U	206/210 (98%)	192 (93%)	14 (7%)	20	16
8	H	185/190 (97%)	176 (95%)	9 (5%)	31	28
8	V	185/190 (97%)	175 (95%)	10 (5%)	27	24
9	I	172/173 (99%)	165 (96%)	7 (4%)	37	36
9	W	172/173 (99%)	165 (96%)	7 (4%)	37	36
10	J	173/175 (99%)	165 (95%)	8 (5%)	33	31
10	X	173/175 (99%)	165 (95%)	8 (5%)	33	31
11	K	168/169 (99%)	159 (95%)	9 (5%)	27	24
11	Y	168/169 (99%)	159 (95%)	9 (5%)	27	24
12	L	185/185 (100%)	173 (94%)	12 (6%)	21	17
12	Z	185/185 (100%)	172 (93%)	13 (7%)	19	15
13	M	199/208 (96%)	189 (95%)	10 (5%)	30	27
13	a	199/208 (96%)	188 (94%)	11 (6%)	27	23
14	N	162/162 (100%)	152 (94%)	10 (6%)	23	19
14	b	162/162 (100%)	152 (94%)	10 (6%)	23	19
All	All	5316/5540 (96%)	4951 (93%)	365 (7%)	19	15

5 of 365 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	138	SER
2	P	238	LEU
12	Z	71	SER
13	M	212	LEU
1	O	201	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 139 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	48	ASN
3	Q	17	GLN
12	Z	79	HIS
13	M	108	ASN
2	P	20	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 11 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
16	3K4	K	301	11	4,6,7	4.75	3 (75%)	2,8,10	2.36	1 (50%)
16	3K4	Y	301	11	4,6,7	4.47	3 (75%)	2,8,10	2.14	1 (50%)

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
16	3K4	K	301	11	-	0/0/7/9	0/0/1/1
16	3K4	Y	301	11	-	0/0/7/9	0/0/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	K	301	3K4	C5-N6	-7.68	1.42	1.47
16	Y	301	3K4	C5-N6	-6.98	1.42	1.47
16	Y	301	3K4	C2-C5	-4.12	1.40	1.48
16	K	301	3K4	C2-C5	-4.07	1.40	1.48
16	K	301	3K4	C7-C5	-3.74	1.45	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	Y	301	3K4	C2-C5-N6	2.60	60.79	59.79
16	K	301	3K4	C2-C5-N6	2.95	60.93	59.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	249/250 (99%)	-0.19	10 (4%)	42 51	28, 38, 67, 111	0
1	O	249/250 (99%)	-0.03	11 (4%)	38 47	31, 44, 82, 118	0
2	B	244/258 (94%)	0.04	14 (5%)	27 35	26, 43, 80, 140	0
2	P	244/258 (94%)	0.10	14 (5%)	27 35	31, 46, 84, 131	0
3	C	240/254 (94%)	0.16	20 (8%)	14 19	27, 46, 102, 122	0
3	Q	240/254 (94%)	0.45	33 (13%)	4 5	31, 53, 120, 146	0
4	D	235/260 (90%)	-0.08	4 (1%)	73 78	31, 46, 75, 110	0
4	R	235/260 (90%)	-0.00	12 (5%)	32 40	32, 49, 79, 110	0
5	E	231/234 (98%)	0.02	12 (5%)	31 39	35, 48, 80, 126	0
5	S	231/234 (98%)	0.27	18 (7%)	16 22	35, 57, 90, 137	0
6	F	243/288 (84%)	-0.08	8 (3%)	50 59	29, 45, 85, 113	0
6	T	243/288 (84%)	0.02	15 (6%)	24 32	30, 50, 94, 118	0
7	G	241/252 (95%)	-0.20	6 (2%)	61 67	24, 38, 68, 115	0
7	U	241/252 (95%)	-0.06	16 (6%)	22 29	31, 43, 69, 102	0
8	H	226/232 (97%)	-0.20	8 (3%)	48 57	26, 38, 64, 116	0
8	V	226/232 (97%)	-0.12	8 (3%)	48 57	30, 42, 66, 124	0
9	I	204/205 (99%)	-0.42	4 (1%)	68 73	25, 35, 60, 82	0
9	W	204/205 (99%)	-0.42	3 (1%)	76 81	26, 36, 63, 92	0
10	J	195/198 (98%)	-0.30	6 (3%)	52 61	25, 38, 62, 129	0
10	X	195/198 (98%)	-0.32	4 (2%)	67 72	25, 39, 63, 125	0
11	K	211/212 (99%)	-0.40	1 (0%)	91 93	26, 36, 55, 71	0
11	Y	211/212 (99%)	-0.37	1 (0%)	91 93	29, 38, 57, 74	0
12	L	222/222 (100%)	-0.36	2 (0%)	85 88	27, 39, 61, 95	0
12	Z	222/222 (100%)	-0.39	3 (1%)	78 82	27, 38, 61, 96	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	233/246 (94%)	-0.45	4 (1%)	73 78	25, 39, 59, 77	0
13	a	233/246 (94%)	-0.42	4 (1%)	73 78	25, 37, 54, 72	0
14	N	196/196 (100%)	-0.42	2 (1%)	84 87	26, 36, 62, 88	0
14	b	196/196 (100%)	-0.34	3 (1%)	76 81	28, 37, 61, 95	0
All	All	6340/6614 (95%)	-0.15	246 (3%)	43 52	24, 42, 78, 146	0

The worst 5 of 246 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Q	50	LEU	11.1
2	B	218	GLY	9.0
2	B	219	ALA	8.9
2	P	51	VAL	7.8
1	O	249	ALA	7.6

## 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	MG	J	201	1/1	0.97	0.22	15.46	37,37,37,37	0
15	MG	I	302	1/1	0.89	0.37	13.96	59,59,59,59	0
15	MG	K	303	1/1	0.97	0.18	3.32	41,41,41,41	0
16	3K4	K	301	6/7	0.94	0.15	2.53	36,41,45,48	0
16	3K4	Y	301	6/7	0.95	0.13	1.81	41,47,50,54	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
15	MG	Y	302	1/1	0.97	0.08	-0.86	31,31,31,31	0
15	MG	G	301	1/1	0.98	0.04	-1.45	31,31,31,31	0
15	MG	V	301	1/1	0.98	0.04	-1.91	38,38,38,38	0
15	MG	K	302	1/1	0.99	0.05	-2.01	29,29,29,29	0
15	MG	I	301	1/1	0.99	0.03	-2.27	34,34,34,34	0
15	MG	Z	301	1/1	0.99	0.07	-2.44	40,40,40,40	0
15	MG	N	201	1/1	0.96	0.04	-2.46	40,40,40,40	0
15	MG	H	301	1/1	0.72	0.13	-	53,53,53,53	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.