



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 18, 2024 – 04:42 pm BST

PDB ID : 5CGH  
Title : Yeast 20S proteasome beta5-G48C mutant in complex with alpha-chloroacetamide 5  
Authors : Dubiella, C.; Groll, M.  
Deposited on : 2015-07-09  
Resolution : 2.50 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.37.1  
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) : 2.37.1

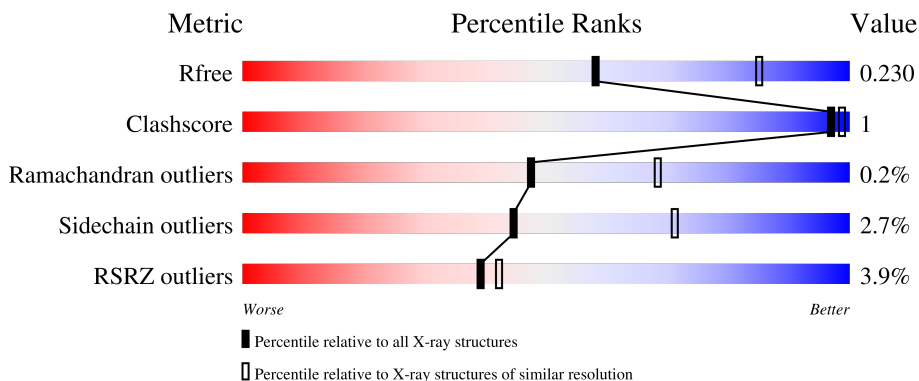
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	250	 3% 98% 5%
1	O	250	 4% 98% 5%
2	B	258	 7% 90% 5% 5%
2	P	258	 6% 90% 5% 5%
3	C	254	 9% 90% 5% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	11% 90% 6%
4	D	260	2% 87% 10%
4	R	260	3% 87% 10%
5	E	234	6% 95% ..
5	S	234	6% 95% ..
6	F	288	3% 81% 16%
6	T	288	4% 81% 16%
7	G	252	4% 90% 5%
7	U	252	4% 90% 6%
8	H	232	% 91% ..
8	V	232	2% 91% 5%
9	I	205	% 95% .
9	W	205	% 96% .
10	J	198	2% 93% ..
10	X	198	3% 93% ..
11	K	212	2% 97% .
11	Y	212	3% 96% .
12	L	222	3% 95% 5%
12	Z	222	4% 95% 5%
13	M	246	% 90% 5%
13	a	246	% 92% 5%
14	N	196	2% 98% .
14	b	196	2% 99% .
15	c	5	80% 20%
15	g	5	80% 20%

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
19	54L	c	101	-	-	-	X

## 2 Entry composition

There are 20 unique types of molecules in this entry. The entry contains 50544 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	244	Total	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
			Total	C	N	O	S			
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
			Total	C	N	O	S			
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
			Total	C	N	O	S			
5	E	231	Total 1773	C 1114	N 307	O 348	S 4	0	0	0
5	S	231	Total 1773	C 1114	N 307	O 348	S 4	0	0	0

- Molecule 6 is a protein called Probable proteasome subunit alpha type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	243	Total 1892	C 1203	N 329	O 356	S 4	0	0	0
6	T	243	Total 1892	C 1203	N 329	O 356	S 4	0	0	0

- Molecule 7 is a protein called Proteasome subunit alpha type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	241	Total 1907	C 1214	N 320	O 365	S 8	0	0	0
7	U	241	Total 1907	C 1214	N 320	O 365	S 8	0	0	0

- Molecule 8 is a protein called Proteasome subunit beta type-2.

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

- Molecule 9 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	204	Total 1581	C 1010	N 258	O 305	S 8	0	0	0
9	W	204	Total 1581	C 1010	N 258	O 305	S 8	0	0	0

- Molecule 10 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	195	Total 1561	C 992	N 264	O 299	S 6	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	X	195	1561	992	264	299	6	0	0	0

- Molecule 11 is a protein called Proteasome subunit beta type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	212	1646	1046	280	312	8	0	0	0
11	Y	212	1646	1046	280	312	8	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	48	CYS	GLY	conflict	UNP P30656
Y	48	CYS	GLY	conflict	UNP P30656

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

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

- Molecule 13 is a protein called Proteasome subunit beta type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	233	1824	1154	312	351	7	0	0	0
13	a	233	1824	1154	312	351	7	0	0	0

- Molecule 14 is a protein called Proteasome subunit beta type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
14	N	196	1512	955	250	300	7	0	0	0
14	b	196	1512	955	250	300	7	0	0	0

- Molecule 15 is a protein called carfilzomib-alpha-chloroacetamide 5.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	c	5	Total	C	N	O	0	0	0
			37	28	5	4			
15	g	5	Total	C	N	O	0	0	0
			37	28	5	4			

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

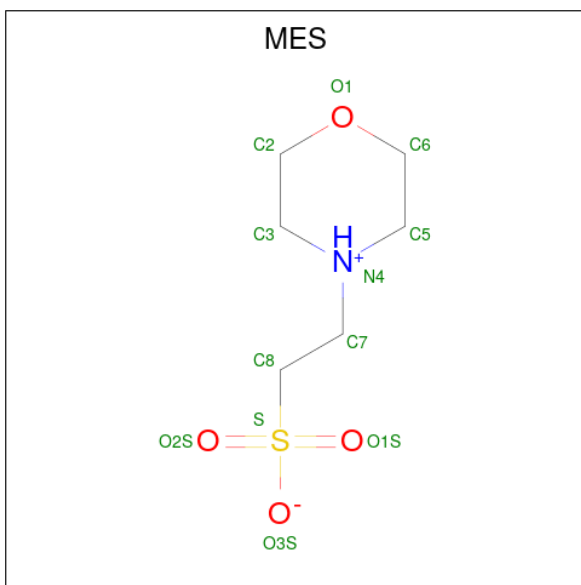
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	G	1	Total	Mg	0	0
			1	1		
16	I	1	Total	Mg	0	0
			1	1		
16	K	1	Total	Mg	0	0
			1	1		
16	N	1	Total	Mg	0	0
			1	1		
16	V	1	Total	Mg	0	0
			1	1		
16	Y	1	Total	Mg	0	0
			1	1		
16	Z	1	Total	Mg	0	0
			1	1		

- Molecule 17 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	G	1	Total	Cl	0	0
			1	1		
17	U	1	Total	Cl	0	0
			1	1		

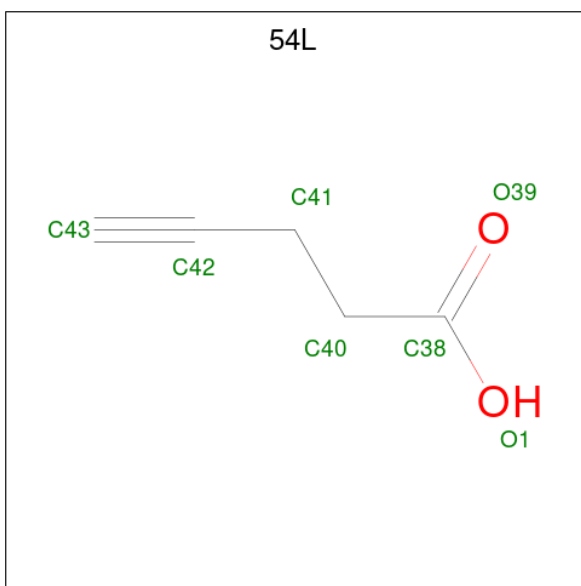
- Molecule 18 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
			Total	C	N	O			S	
18	K	1	Total	12	6	1	4	1	0	0
18	Y	1	Total	12	6	1	4	1	0	0

- Molecule 19 is pent-4-ynoic acid (three-letter code: 54L) (formula:  $C_5H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	O			
19	c	1	Total	6	5	1	0	0
19	g	1	Total	6	5	1	0	0

- Molecule 20 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
20	A	58	Total O 58 58	0	0
20	B	39	Total O 39 39	0	0
20	C	31	Total O 31 31	0	0
20	D	26	Total O 26 26	0	0
20	E	23	Total O 23 23	0	0
20	F	40	Total O 40 40	0	0
20	G	52	Total O 52 52	0	0
20	H	55	Total O 55 55	0	0
20	I	46	Total O 46 46	0	0
20	J	44	Total O 44 44	0	0
20	K	42	Total O 42 42	0	0
20	L	48	Total O 48 48	0	0
20	M	46	Total O 46 46	0	0
20	N	37	Total O 37 37	0	0
20	O	45	Total O 45 45	0	0
20	P	37	Total O 37 37	0	0
20	Q	16	Total O 16 16	0	0
20	R	24	Total O 24 24	0	0
20	S	14	Total O 14 14	0	0
20	T	31	Total O 31 31	0	0
20	U	55	Total O 55 55	0	0

*Continued on next page...*

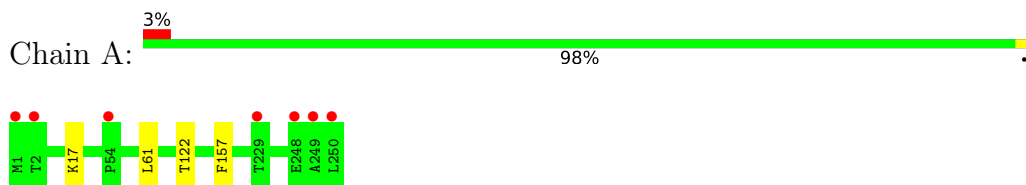
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
20	V	39	Total O 39 39	0	0
20	W	44	Total O 44 44	0	0
20	X	39	Total O 39 39	0	0
20	Y	33	Total O 33 33	0	0
20	Z	55	Total O 55 55	0	0
20	a	57	Total O 57 57	0	0
20	b	46	Total O 46 46	0	0
20	c	1	Total O 1 1	0	0
20	g	2	Total O 2 2	0	0

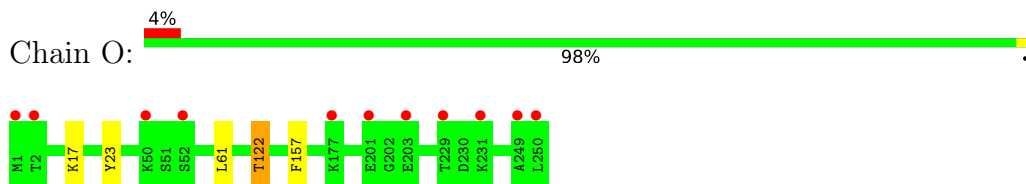
### 3 Residue-property plots

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

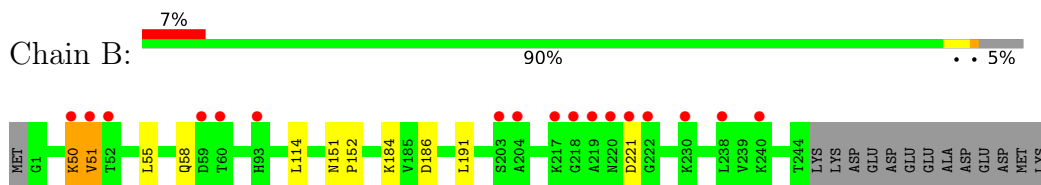
- Molecule 1: Proteasome subunit alpha type-2



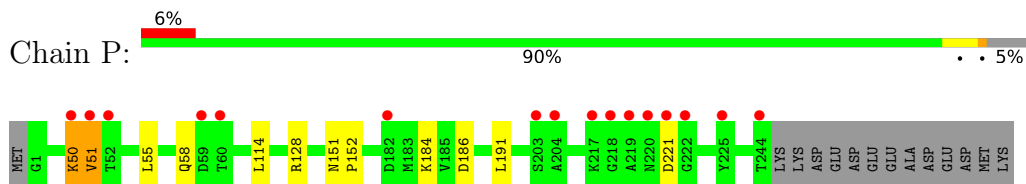
- Molecule 1: Proteasome subunit alpha type-2



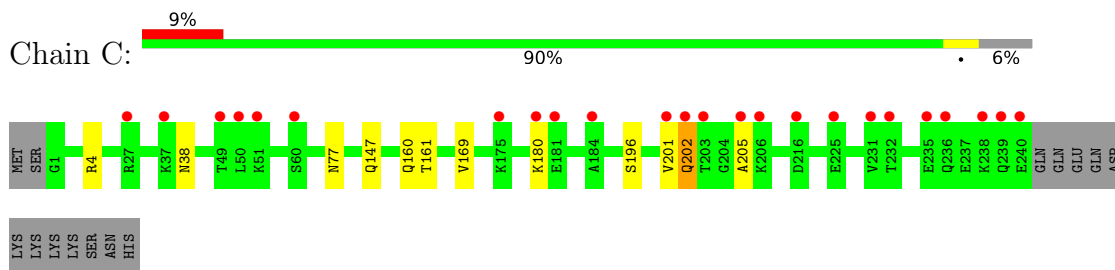
- Molecule 2: Proteasome subunit alpha type-3



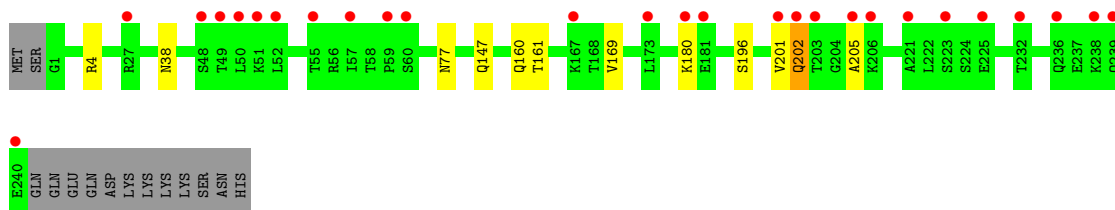
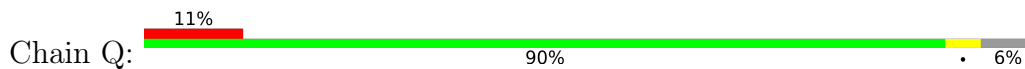
- Molecule 2: Proteasome subunit alpha type-3



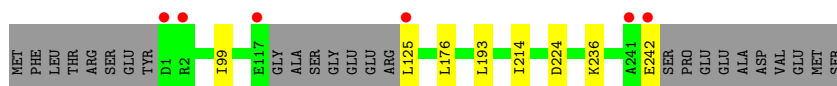
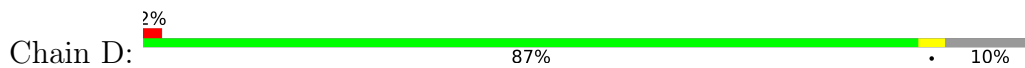
- Molecule 3: Proteasome subunit alpha type-4



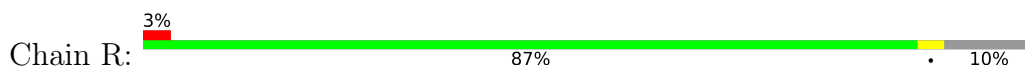
• Molecule 3: Proteasome subunit alpha type-4



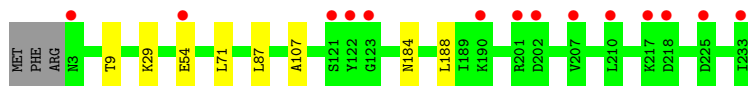
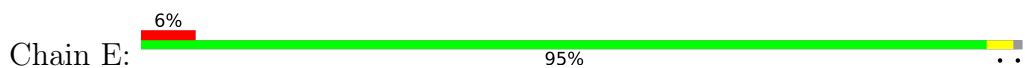
• Molecule 4: Proteasome subunit alpha type-5



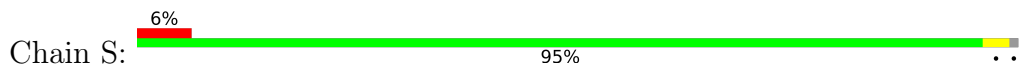
• Molecule 4: Proteasome subunit alpha type-5



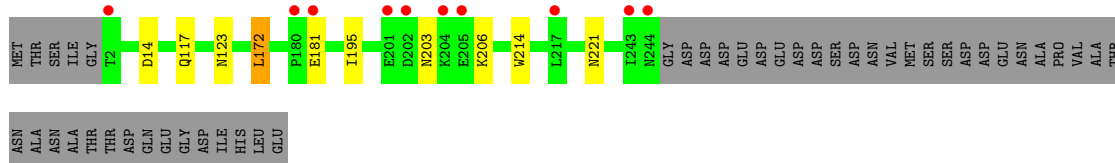
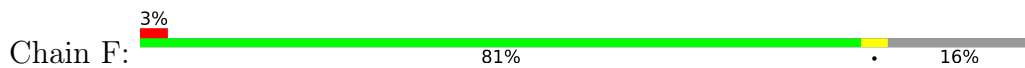
• Molecule 5: Proteasome subunit alpha type-6



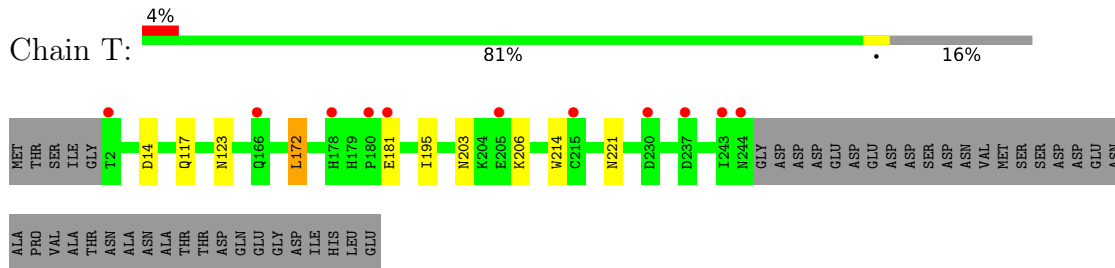
• Molecule 5: Proteasome subunit alpha type-6



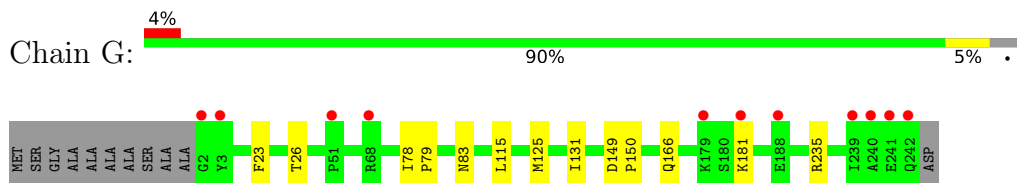
• Molecule 6: Probable proteasome subunit alpha type-7



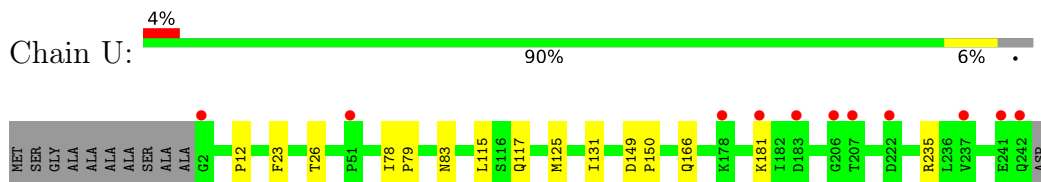
● Molecule 6: Probable proteasome subunit alpha type-7



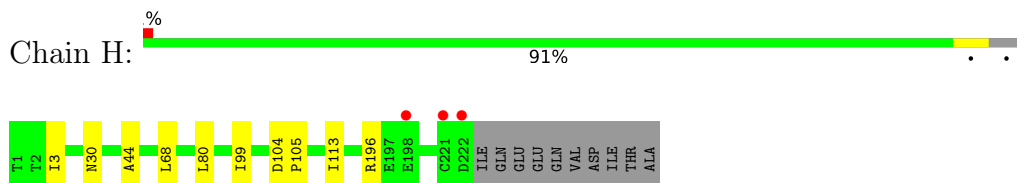
● Molecule 7: Proteasome subunit alpha type-1



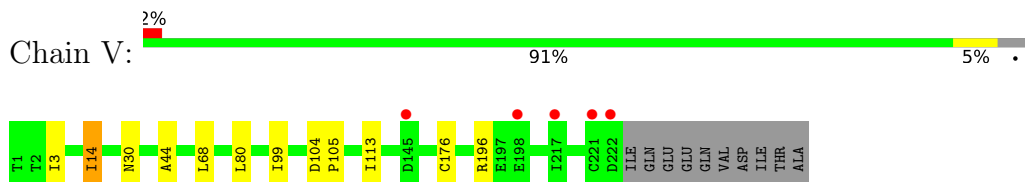
● Molecule 7: Proteasome subunit alpha type-1



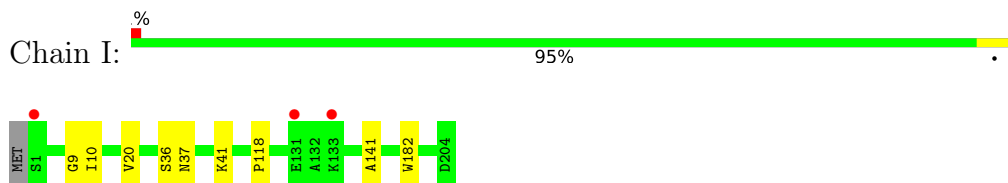
● Molecule 8: Proteasome subunit beta type-2



● Molecule 8: Proteasome subunit beta type-2



● Molecule 9: Proteasome subunit beta type-3

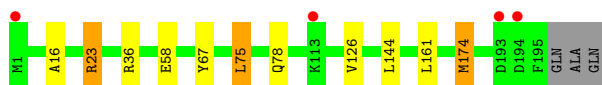
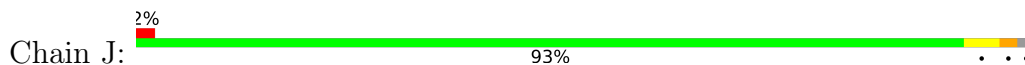


● Molecule 9: Proteasome subunit beta type-3

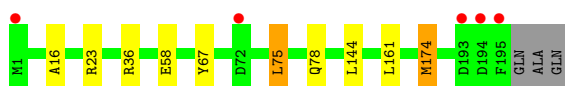
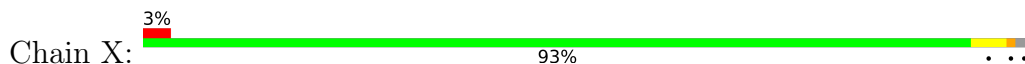




- Molecule 10: Proteasome subunit beta type-4



- Molecule 10: Proteasome subunit beta type-4



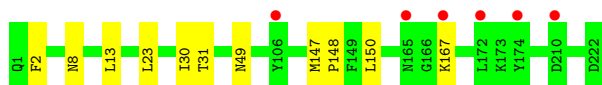
- Molecule 11: Proteasome subunit beta type-5



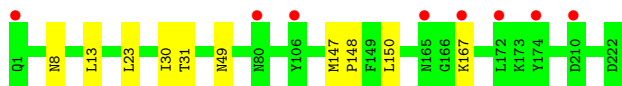
- Molecule 11: Proteasome subunit beta type-5



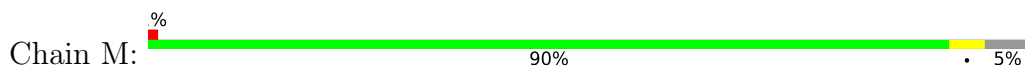
- Molecule 12: Proteasome subunit beta type-6



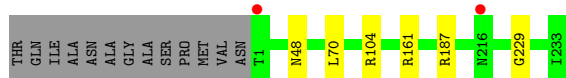
- Molecule 12: Proteasome subunit beta type-6



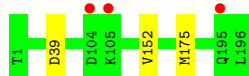
- Molecule 13: Proteasome subunit beta type-7



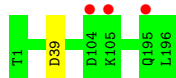
- Molecule 13: Proteasome subunit beta type-7



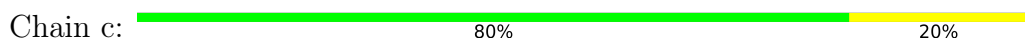
- Molecule 14: Proteasome subunit beta type-1



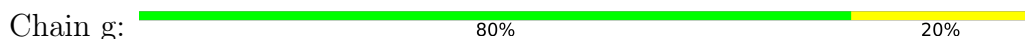
- Molecule 14: Proteasome subunit beta type-1



- Molecule 15: carfilzomib-alpha-chloroacetamide 5



- Molecule 15: carfilzomib-alpha-chloroacetamide 5





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.91Å 300.21Å 145.87Å 90.00° 112.90° 90.00°	Depositor
Resolution (Å)	15.00 – 2.50 15.00 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.1 (15.00-2.50) 96.1 (15.00-2.50)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.20 (at 2.51Å)	Xtrriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.207 , 0.227 0.210 , 0.230	Depositor DCC
$R_{free}$ test set	17551 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.4	Xtrriage
Anisotropy	0.047	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 36.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	50544	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.55% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, 54L, MES, DPN, DPP, R3W, MG, 02N

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	0.27	0/1952	0.45	0/2642
1	O	0.27	0/1952	0.45	0/2642
2	B	0.27	0/1934	0.47	0/2618
2	P	0.27	0/1934	0.47	0/2618
3	C	0.27	0/1910	0.49	0/2586
3	Q	0.27	0/1910	0.49	0/2586
4	D	0.26	0/1837	0.45	0/2475
4	R	0.26	0/1837	0.45	0/2475
5	E	0.27	0/1800	0.45	0/2433
5	S	0.26	0/1800	0.45	0/2433
6	F	0.27	0/1932	0.44	0/2609
6	T	0.27	0/1932	0.44	0/2609
7	G	0.27	0/1945	0.46	0/2634
7	U	0.27	0/1945	0.46	0/2634
8	H	0.27	0/1715	0.51	0/2326
8	V	0.27	0/1715	0.51	0/2326
9	I	0.27	0/1611	0.47	0/2174
9	W	0.27	0/1611	0.47	0/2174
10	J	0.26	0/1589	0.48	0/2142
10	X	0.26	0/1589	0.47	0/2142
11	K	0.26	0/1683	0.48	0/2277
11	Y	0.26	0/1683	0.48	0/2277
12	L	0.27	0/1795	0.46	0/2420
12	Z	0.26	0/1795	0.45	0/2420
13	M	0.28	0/1855	0.50	0/2514
13	a	0.27	0/1855	0.50	0/2514
14	N	0.25	0/1541	0.47	0/2087
14	b	0.25	0/1541	0.47	0/2087
15	c	0.79	0/7	1.28	0/8
15	g	0.81	0/7	1.24	0/8
All	All	0.27	0/50212	0.47	0/67890

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1915	0	1929	0	0
1	O	1915	0	1929	2	0
2	B	1904	0	1904	2	0
2	P	1904	0	1904	3	0
3	C	1881	0	1895	3	0
3	Q	1881	0	1895	3	0
4	D	1813	0	1797	0	0
4	R	1813	0	1797	0	0
5	E	1773	0	1775	1	0
5	S	1773	0	1775	1	0
6	F	1892	0	1883	1	0
6	T	1892	0	1883	1	0
7	G	1907	0	1901	4	0
7	U	1907	0	1901	5	0
8	H	1684	0	1688	7	0
8	V	1684	0	1688	7	0
9	I	1581	0	1574	4	0
9	W	1581	0	1574	3	0
10	J	1561	0	1569	6	0
10	X	1561	0	1569	4	0
11	K	1646	0	1596	2	0
11	Y	1646	0	1596	4	0
12	L	1757	0	1711	5	0
12	Z	1757	0	1711	4	0
13	M	1824	0	1832	4	0
13	a	1824	0	1832	0	0
14	N	1512	0	1481	1	0
14	b	1512	0	1481	0	0
15	c	37	0	32	0	0
15	g	37	0	32	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	G	1	0	0	0	0
16	I	1	0	0	0	0
16	K	1	0	0	0	0
16	N	1	0	0	0	0
16	V	1	0	0	0	0
16	Y	1	0	0	0	0
16	Z	1	0	0	0	0
17	G	1	0	0	0	0
17	U	1	0	0	0	0
18	K	12	0	13	0	0
18	Y	12	0	13	0	0
19	c	6	0	5	0	0
19	g	6	0	5	0	0
20	A	58	0	0	0	0
20	B	39	0	0	0	0
20	C	31	0	0	0	0
20	D	26	0	0	0	0
20	E	23	0	0	0	0
20	F	40	0	0	0	0
20	G	52	0	0	0	0
20	H	55	0	0	0	0
20	I	46	0	0	0	0
20	J	44	0	0	0	0
20	K	42	0	0	1	0
20	L	48	0	0	1	0
20	M	46	0	0	1	0
20	N	37	0	0	0	0
20	O	45	0	0	0	0
20	P	37	0	0	0	0
20	Q	16	0	0	0	0
20	R	24	0	0	0	0
20	S	14	0	0	0	0
20	T	31	0	0	0	0
20	U	55	0	0	0	0
20	V	39	0	0	0	0
20	W	44	0	0	0	0
20	X	39	0	0	0	0
20	Y	33	0	0	0	0
20	Z	55	0	0	0	0
20	a	57	0	0	0	0
20	b	46	0	0	0	0
20	c	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
20	g	2	0	0	0	0
All	All	50544	0	49170	73	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 1.

All (73) 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)
8:V:14:ILE:HD11	8:V:44:ALA:HB2	1.63	0.80
10:J:23:ARG:HD3	20:K:401:HOH:O	1.85	0.76
8:V:14:ILE:HD11	8:V:44:ALA:CB	2.21	0.70
8:V:14:ILE:CD1	8:V:44:ALA:HB2	2.20	0.70
11:K:5:ALA:HB3	11:K:100:MET:HE2	1.77	0.64
8:V:3:ILE:HG22	8:V:99:ILE:HD12	1.79	0.64
11:Y:5:ALA:HB3	11:Y:100:MET:HE2	1.80	0.62
14:N:152:VAL:HA	14:N:175:MET:HE1	1.85	0.57
8:H:3:ILE:HG22	8:H:99:ILE:HD12	1.88	0.54
3:C:201:VAL:O	3:C:202:GLN:CB	2.57	0.52
3:Q:201:VAL:O	3:Q:202:GLN:CB	2.58	0.51
10:J:67:TYR:CE1	10:J:75:LEU:HD13	2.47	0.50
12:Z:13:LEU:HD13	12:Z:150:LEU:HD21	1.94	0.50
12:L:13:LEU:HD13	12:L:150:LEU:HD21	1.94	0.49
10:X:67:TYR:CE1	10:X:75:LEU:HD13	2.47	0.49
7:G:23:PHE:O	7:G:26:THR:HB	2.13	0.49
7:U:23:PHE:O	7:U:26:THR:HB	2.13	0.48
10:J:16:ALA:HB2	10:J:161:LEU:HD21	1.96	0.47
11:K:35:ILE:HB	11:K:45:MET:CE	2.45	0.47
8:V:80:LEU:HD12	8:V:113:ILE:HD11	1.97	0.47
8:V:14:ILE:HG22	8:V:176:CYS:HB3	1.98	0.46
11:Y:35:ILE:HB	11:Y:45:MET:CE	2.45	0.46
8:H:3:ILE:HG21	8:H:44:ALA:CB	2.45	0.46
10:J:36:ARG:NH1	10:J:58:GLU:OE2	2.48	0.45
10:X:16:ALA:HB2	10:X:161:LEU:HD21	1.96	0.45
10:X:36:ARG:NH1	10:X:58:GLU:OE2	2.49	0.45
8:H:80:LEU:HD12	8:H:113:ILE:HD11	1.98	0.45
3:Q:161:THR:HG21	3:Q:169:VAL:HG13	1.99	0.45
1:O:23:TYR:CD1	7:U:12:PRO:HA	2.53	0.44
3:C:161:THR:HG21	3:C:169:VAL:HG13	1.99	0.44
7:G:78:ILE:N	7:G:79:PRO:CD	2.81	0.44
8:H:3:ILE:CG2	8:H:44:ALA:CB	2.95	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:V:104:ASP:HB2	8:V:105:PRO:HD2	2.00	0.43
11:Y:5:ALA:HB3	11:Y:100:MET:CE	2.47	0.43
3:C:169:VAL:HG23	3:C:196:SER:HB2	2.00	0.43
8:H:3:ILE:CG2	8:H:44:ALA:HB1	2.48	0.43
9:W:9:GLY:HA3	9:W:41:LYS:HE2	2.01	0.43
12:L:8:ASN:HA	12:L:30:ILE:O	2.19	0.43
8:H:104:ASP:HB2	8:H:105:PRO:HD2	1.99	0.43
7:U:78:ILE:N	7:U:79:PRO:CD	2.81	0.43
9:W:20:VAL:HG13	9:W:118:PRO:HB3	2.01	0.43
11:Y:100:MET:HE3	11:Y:127:PHE:HB2	2.00	0.43
12:Z:8:ASN:HA	12:Z:30:ILE:O	2.18	0.43
12:L:2:PHE:N	20:L:401:HOH:O	2.52	0.43
2:P:151:ASN:HB2	2:P:152:PRO:CD	2.49	0.43
9:I:9:GLY:HA3	9:I:41:LYS:HE2	2.01	0.42
12:L:13:LEU:CD1	12:L:150:LEU:HD21	2.49	0.42
9:I:36:SER:HB2	10:J:126:VAL:HG11	2.01	0.42
2:B:151:ASN:HB2	2:B:152:PRO:CD	2.49	0.42
9:I:20:VAL:HG13	9:I:118:PRO:HB3	2.01	0.42
3:Q:169:VAL:HG23	3:Q:196:SER:HB2	2.00	0.42
9:W:10:ILE:HG21	9:W:141:ALA:HB3	2.02	0.42
6:F:172:LEU:CD1	6:F:195:ILE:HD13	2.50	0.42
12:Z:13:LEU:CD1	12:Z:150:LEU:HD21	2.50	0.42
7:G:26:THR:HG21	7:G:131:ILE:HD12	2.02	0.41
7:U:26:THR:HG21	7:U:131:ILE:HD12	2.01	0.41
1:O:122:THR:HG22	2:P:128:ARG:HH21	1.85	0.41
13:M:96:LEU:O	13:M:100:MET:HG2	2.20	0.41
13:M:165:ILE:HB	13:M:166:PRO:HD3	2.03	0.41
13:M:187:ARG:NH2	20:M:301:HOH:O	2.47	0.41
6:T:172:LEU:CD1	6:T:195:ILE:HD13	2.50	0.41
12:Z:147:MET:N	12:Z:148:PRO:HD2	2.36	0.41
8:H:3:ILE:HG21	8:H:44:ALA:HB3	2.03	0.41
9:I:10:ILE:HG21	9:I:141:ALA:HB3	2.02	0.41
10:J:174:MET:HA	10:X:174:MET:HA	2.02	0.41
5:S:87:LEU:HD21	5:S:107:ALA:HB1	2.02	0.41
13:M:97:ALA:HA	13:M:130:VAL:HG21	2.04	0.40
2:P:50:LYS:O	2:P:51:VAL:C	2.60	0.40
12:L:147:MET:N	12:L:148:PRO:HD2	2.36	0.40
7:G:149:ASP:HB2	7:G:150:PRO:CD	2.52	0.40
2:B:50:LYS:O	2:B:51:VAL:C	2.60	0.40
5:E:87:LEU:HD21	5:E:107:ALA:HB1	2.03	0.40
7:U:149:ASP:HB2	7:U:150:PRO:CD	2.52	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	248/250 (99%)	241 (97%)	7 (3%)	0	100	100
1	O	248/250 (99%)	241 (97%)	7 (3%)	0	100	100
2	B	242/258 (94%)	233 (96%)	7 (3%)	2 (1%)	19	35
2	P	242/258 (94%)	233 (96%)	7 (3%)	2 (1%)	19	35
3	C	238/254 (94%)	232 (98%)	4 (2%)	2 (1%)	19	35
3	Q	238/254 (94%)	232 (98%)	4 (2%)	2 (1%)	19	35
4	D	231/260 (89%)	224 (97%)	7 (3%)	0	100	100
4	R	231/260 (89%)	224 (97%)	7 (3%)	0	100	100
5	E	229/234 (98%)	224 (98%)	5 (2%)	0	100	100
5	S	229/234 (98%)	224 (98%)	5 (2%)	0	100	100
6	F	241/288 (84%)	235 (98%)	6 (2%)	0	100	100
6	T	241/288 (84%)	235 (98%)	6 (2%)	0	100	100
7	G	239/252 (95%)	234 (98%)	5 (2%)	0	100	100
7	U	239/252 (95%)	234 (98%)	5 (2%)	0	100	100
8	H	220/232 (95%)	212 (96%)	8 (4%)	0	100	100
8	V	220/232 (95%)	213 (97%)	7 (3%)	0	100	100
9	I	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
9	W	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	J	193/198 (98%)	189 (98%)	4 (2%)	0	100	100
10	X	193/198 (98%)	189 (98%)	4 (2%)	0	100	100
11	K	210/212 (99%)	205 (98%)	5 (2%)	0	100	100
11	Y	210/212 (99%)	205 (98%)	5 (2%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
12	L	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
12	Z	220/222 (99%)	216 (98%)	4 (2%)	0	100	100
13	M	231/246 (94%)	222 (96%)	8 (4%)	1 (0%)	34	54
13	a	231/246 (94%)	222 (96%)	8 (4%)	1 (0%)	34	54
14	N	194/196 (99%)	188 (97%)	6 (3%)	0	100	100
14	b	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
15	c	1/5 (20%)	1 (100%)	0	0	100	100
15	g	1/5 (20%)	1 (100%)	0	0	100	100
All	All	6278/6624 (95%)	6102 (97%)	166 (3%)	10 (0%)	47	68

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
3	C	202	GLN
2	P	51	VAL
3	Q	202	GLN
2	B	221	ASP
2	P	221	ASP
3	C	205	ALA
3	Q	205	ALA
13	M	229	GLY
13	a	229	GLY

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	209/209 (100%)	205 (98%)	4 (2%)	57	80
1	O	209/209 (100%)	205 (98%)	4 (2%)	57	80
2	B	203/216 (94%)	196 (97%)	7 (3%)	37	63
2	P	203/216 (94%)	196 (97%)	7 (3%)	37	63

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	212/226 (94%)	206 (97%)	6 (3%)	43	70
3	Q	212/226 (94%)	206 (97%)	6 (3%)	43	70
4	D	194/215 (90%)	186 (96%)	8 (4%)	30	55
4	R	194/215 (90%)	186 (96%)	8 (4%)	30	55
5	E	190/193 (98%)	184 (97%)	6 (3%)	39	65
5	S	190/193 (98%)	184 (97%)	6 (3%)	39	65
6	F	201/239 (84%)	192 (96%)	9 (4%)	27	51
6	T	201/239 (84%)	192 (96%)	9 (4%)	27	51
7	G	206/210 (98%)	200 (97%)	6 (3%)	42	69
7	U	206/210 (98%)	199 (97%)	7 (3%)	37	63
8	H	181/190 (95%)	178 (98%)	3 (2%)	60	82
8	V	181/190 (95%)	177 (98%)	4 (2%)	52	77
9	I	172/173 (99%)	170 (99%)	2 (1%)	71	88
9	W	172/173 (99%)	170 (99%)	2 (1%)	71	88
10	J	173/175 (99%)	168 (97%)	5 (3%)	42	69
10	X	173/175 (99%)	168 (97%)	5 (3%)	42	69
11	K	170/170 (100%)	166 (98%)	4 (2%)	49	74
11	Y	170/170 (100%)	166 (98%)	4 (2%)	49	74
12	L	185/185 (100%)	181 (98%)	4 (2%)	52	77
12	Z	185/185 (100%)	181 (98%)	4 (2%)	52	77
13	M	199/208 (96%)	194 (98%)	5 (2%)	47	73
13	a	199/208 (96%)	194 (98%)	5 (2%)	47	73
14	N	162/162 (100%)	161 (99%)	1 (1%)	86	95
14	b	162/162 (100%)	161 (99%)	1 (1%)	86	95
15	c	1/1 (100%)	1 (100%)	0	100	100
15	g	1/1 (100%)	1 (100%)	0	100	100
All	All	5316/5544 (96%)	5174 (97%)	142 (3%)	44	71

All (142) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	LYS
1	A	61	LEU

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	122	THR
1	A	157	PHE
2	B	50	LYS
2	B	55	LEU
2	B	58	GLN
2	B	114	LEU
2	B	184	LYS
2	B	186	ASP
2	B	191	LEU
3	C	4	ARG
3	C	38	ASN
3	C	77	ASN
3	C	147	GLN
3	C	160	GLN
3	C	180	LYS
4	D	99	ILE
4	D	125	LEU
4	D	176	LEU
4	D	193	LEU
4	D	214	ILE
4	D	224	ASP
4	D	236	LYS
4	D	242	GLU
5	E	9	THR
5	E	29	LYS
5	E	54	GLU
5	E	71	LEU
5	E	184	ASN
5	E	188	LEU
6	F	14	ASP
6	F	117	GLN
6	F	123	ASN
6	F	172	LEU
6	F	181	GLU
6	F	203	ASN
6	F	206	LYS
6	F	214	TRP
6	F	221	ASN
7	G	83	ASN
7	G	115	LEU
7	G	125	MET
7	G	166	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
7	G	181	LYS
7	G	235	ARG
8	H	30	ASN
8	H	68	LEU
8	H	196	ARG
9	I	37	ASN
9	I	182	TRP
10	J	23	ARG
10	J	75	LEU
10	J	78	GLN
10	J	144	LEU
10	J	174	MET
11	K	4	LEU
11	K	9	GLN
11	K	35	ILE
11	K	106	ARG
12	L	23	LEU
12	L	31	THR
12	L	49	ASN
12	L	167	LYS
13	M	48	ASN
13	M	70	LEU
13	M	104	ARG
13	M	161	ARG
13	M	187	ARG
14	N	39	ASP
1	O	17	LYS
1	O	61	LEU
1	O	122	THR
1	O	157	PHE
2	P	50	LYS
2	P	55	LEU
2	P	58	GLN
2	P	114	LEU
2	P	184	LYS
2	P	186	ASP
2	P	191	LEU
3	Q	4	ARG
3	Q	38	ASN
3	Q	77	ASN
3	Q	147	GLN
3	Q	160	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
3	Q	180	LYS
4	R	99	ILE
4	R	125	LEU
4	R	176	LEU
4	R	193	LEU
4	R	214	ILE
4	R	224	ASP
4	R	236	LYS
4	R	242	GLU
5	S	9	THR
5	S	29	LYS
5	S	54	GLU
5	S	71	LEU
5	S	184	ASN
5	S	188	LEU
6	T	14	ASP
6	T	117	GLN
6	T	123	ASN
6	T	172	LEU
6	T	181	GLU
6	T	203	ASN
6	T	206	LYS
6	T	214	TRP
6	T	221	ASN
7	U	83	ASN
7	U	115	LEU
7	U	117	GLN
7	U	125	MET
7	U	166	GLN
7	U	181	LYS
7	U	235	ARG
8	V	14	ILE
8	V	30	ASN
8	V	68	LEU
8	V	196	ARG
9	W	37	ASN
9	W	182	TRP
10	X	23	ARG
10	X	75	LEU
10	X	78	GLN
10	X	144	LEU
10	X	174	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
11	Y	4	LEU
11	Y	9	GLN
11	Y	35	ILE
11	Y	106	ARG
12	Z	23	LEU
12	Z	31	THR
12	Z	49	ASN
12	Z	167	LYS
13	a	48	ASN
13	a	70	LEU
13	a	104	ARG
13	a	161	ARG
13	a	187	ARG
14	b	39	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (68) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	HIS
2	B	119	GLN
2	B	123	GLN
2	B	155	ASN
3	C	38	ASN
3	C	116	GLN
3	C	120	GLN
3	C	147	GLN
3	C	160	GLN
4	D	91	HIS
5	E	92	ASN
5	E	116	GLN
5	E	120	GLN
5	E	184	ASN
6	F	86	ASN
6	F	117	GLN
6	F	123	ASN
6	F	191	GLN
6	F	240	GLN
7	G	83	ASN
7	G	117	GLN
7	G	121	GLN
7	G	166	GLN
9	I	37	ASN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
9	I	203	GLN
10	J	55	GLN
11	K	85	ASN
11	K	176	ASN
12	L	3	ASN
12	L	70	ASN
12	L	79	HIS
12	L	95	HIS
12	L	158	ASN
13	M	48	ASN
13	M	102	GLN
1	O	94	HIS
2	P	20	GLN
2	P	119	GLN
2	P	123	GLN
3	Q	38	ASN
3	Q	116	GLN
3	Q	120	GLN
3	Q	147	GLN
3	Q	160	GLN
4	R	91	HIS
4	R	100	ASN
5	S	99	ASN
5	S	116	GLN
5	S	120	GLN
5	S	184	ASN
6	T	86	ASN
6	T	117	GLN
6	T	123	ASN
6	T	191	GLN
6	T	240	GLN
7	U	83	ASN
7	U	117	GLN
7	U	121	GLN
7	U	166	GLN
8	V	165	ASN
10	X	55	GLN
11	Y	85	ASN
11	Y	176	ASN
12	Z	3	ASN
12	Z	79	HIS
12	Z	158	ASN

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
13	a	48	ASN
13	a	102	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](#)

4 non-standard protein/DNA/RNA residues 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	DPP	g	2	19,15	3,5,6	2.13	1 (33%)	1,5,7	1.76	0
15	DPP	c	2	19,15	3,5,6	2.09	1 (33%)	1,5,7	1.72	0

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
15	DPP	g	2	19,15	-	0/2/4/6	-
15	DPP	c	2	19,15	-	0/2/4/6	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	g	2	DPP	CB-NG	-3.53	1.33	1.48
15	c	2	DPP	CB-NG	-3.46	1.34	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 9 are monoatomic - leaving 4 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
19	54L	c	101	15	5,5,6	1.30	1 (20%)	4,4,6	1.42	1 (25%)
18	MES	Y	302	-	12,12,12	2.31	1 (8%)	14,16,16	1.20	2 (14%)
18	MES	K	302	-	12,12,12	2.25	1 (8%)	14,16,16	1.19	2 (14%)
19	54L	g	101	15	5,5,6	1.27	1 (20%)	4,4,6	1.38	1 (25%)

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
19	54L	c	101	15	-	0/1/3/4	-
18	MES	Y	302	-	-	2/6/14/14	0/1/1/1
18	MES	K	302	-	-	1/6/14/14	0/1/1/1
19	54L	g	101	15	-	0/1/3/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	Y	302	MES	C8-S	-7.73	1.66	1.77

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	K	302	MES	C8-S	-7.48	1.66	1.77
19	c	101	54L	C41-C42	2.58	1.53	1.47
19	g	101	54L	C41-C42	2.12	1.52	1.47

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
18	K	302	MES	O3S-S-C8	3.07	110.73	105.77
18	Y	302	MES	O3S-S-C8	2.84	110.36	105.77
19	c	101	54L	C41-C40-C38	-2.47	103.61	111.93
19	g	101	54L	C41-C40-C38	-2.39	103.89	111.93
18	Y	302	MES	O1S-S-C8	2.31	109.69	106.92
18	K	302	MES	O1S-S-C8	2.18	109.55	106.92

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
18	Y	302	MES	C8-C7-N4-C5
18	K	302	MES	C8-C7-N4-C5
18	Y	302	MES	C8-C7-N4-C3

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	250/250 (100%)	-0.14	7 (2%) 53 56	33, 46, 83, 142	0
1	O	250/250 (100%)	-0.00	11 (4%) 34 37	37, 54, 97, 141	0
2	B	244/258 (94%)	0.12	17 (6%) 16 16	33, 52, 99, 147	0
2	P	244/258 (94%)	0.12	16 (6%) 18 19	36, 55, 100, 148	0
3	C	240/254 (94%)	0.28	24 (10%) 7 6	34, 56, 119, 159	0
3	Q	240/254 (94%)	0.46	27 (11%) 5 4	40, 67, 150, 178	0
4	D	235/260 (90%)	0.01	6 (2%) 56 59	38, 60, 91, 138	0
4	R	235/260 (90%)	0.08	7 (2%) 50 53	44, 62, 98, 142	0
5	E	231/234 (98%)	0.09	14 (6%) 21 22	39, 60, 94, 132	0
5	S	231/234 (98%)	0.20	14 (6%) 21 22	44, 67, 108, 140	0
6	F	243/288 (84%)	-0.06	10 (4%) 37 40	33, 54, 101, 132	0
6	T	243/288 (84%)	0.13	11 (4%) 33 36	34, 63, 119, 158	0
7	G	241/252 (95%)	-0.14	11 (4%) 32 34	31, 48, 88, 138	0
7	U	241/252 (95%)	-0.06	11 (4%) 32 34	28, 51, 85, 126	0
8	H	222/232 (95%)	-0.13	3 (1%) 75 77	33, 46, 77, 106	0
8	V	222/232 (95%)	-0.09	5 (2%) 60 63	36, 50, 80, 114	0
9	I	204/205 (99%)	-0.37	3 (1%) 73 75	26, 44, 74, 97	0
9	W	204/205 (99%)	-0.37	3 (1%) 73 75	28, 47, 74, 104	0
10	J	195/198 (98%)	-0.21	4 (2%) 63 66	28, 48, 76, 112	0
10	X	195/198 (98%)	-0.18	5 (2%) 56 59	32, 49, 76, 126	0
11	K	212/212 (100%)	-0.14	5 (2%) 59 62	32, 51, 86, 102	0
11	Y	212/212 (100%)	-0.11	6 (2%) 53 56	34, 51, 90, 115	0
12	L	222/222 (100%)	-0.22	6 (2%) 54 58	25, 49, 93, 128	0
12	Z	222/222 (100%)	-0.14	8 (3%) 42 46	28, 49, 89, 114	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%)	-0.27	3 (1%) 77 79	28, 47, 75, 83	0
13	a	233/246 (94%)	-0.28	2 (0%) 84 86	27, 47, 73, 82	0
14	N	196/196 (100%)	-0.27	3 (1%) 73 75	27, 45, 74, 97	0
14	b	196/196 (100%)	-0.34	3 (1%) 73 75	30, 44, 73, 98	0
15	c	1/5 (20%)	0.00	0 100 100	51, 51, 51, 51	0
15	g	1/5 (20%)	0.43	0 100 100	51, 51, 51, 51	0
All	All	6338/6624 (95%)	-0.06	245 (3%) 39 42	25, 52, 97, 178	0

All (245) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	10.2
1	O	1	MET	9.5
2	P	219	ALA	9.2
3	C	206	LYS	8.5
3	Q	50	LEU	7.7
3	Q	206	LYS	7.7
2	B	218	GLY	7.6
2	B	51	VAL	7.5
2	B	219	ALA	7.5
10	J	1	MET	7.4
2	P	218	GLY	7.3
2	P	51	VAL	7.2
9	I	1	SER	6.8
5	S	202	ASP	6.7
10	X	1	MET	6.4
3	Q	49	THR	6.1
9	W	1	SER	5.7
2	B	221	ASP	5.6
10	X	194	ASP	5.6
4	D	241	ALA	5.4
2	P	220	ASN	5.4
13	a	1	THR	5.3
3	C	49	THR	5.1
7	U	2	GLY	5.1
3	Q	48	SER	5.1
2	B	220	ASN	5.1
3	Q	239	GLN	4.9
2	P	59	ASP	4.9
5	E	202	ASP	4.8

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	C	205	ALA	4.8
3	Q	238	LYS	4.8
3	Q	202	GLN	4.7
2	P	221	ASP	4.7
8	V	221	CYS	4.6
7	U	242	GLN	4.6
6	T	243	ILE	4.5
3	C	50	LEU	4.5
13	M	1	THR	4.5
9	W	133	LYS	4.4
8	V	222	ASP	4.4
4	R	125	LEU	4.4
3	C	202	GLN	4.3
4	D	242	GLU	4.3
4	R	241	ALA	4.3
3	C	236	GLN	4.3
8	H	221	CYS	4.2
2	P	52	THR	4.2
1	O	50	LYS	4.2
2	B	59	ASP	4.2
7	G	3	TYR	4.2
5	E	207	VAL	4.2
6	F	205	GLU	4.1
5	S	210	LEU	4.1
11	Y	212	GLY	4.1
3	Q	225	GLU	4.1
6	F	244	ASN	4.0
9	I	133	LYS	4.0
2	P	222	GLY	4.0
1	O	201	GLU	3.9
5	E	123	GLY	3.9
3	Q	236	GLN	3.9
10	X	193	ASP	3.9
3	Q	51	LYS	3.9
1	A	2	THR	3.8
3	Q	240	GLU	3.8
2	P	182	ASP	3.8
6	F	2	THR	3.8
7	U	181	LYS	3.7
3	C	239	GLN	3.7
1	O	250	LEU	3.7
3	C	180	LYS	3.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
10	J	194	ASP	3.7
7	G	242	GLN	3.7
6	F	202	ASP	3.7
8	H	222	ASP	3.7
2	P	50	LYS	3.6
3	C	238	LYS	3.6
11	K	147	ASP	3.5
7	G	2	GLY	3.5
1	O	249	ALA	3.5
2	B	222	GLY	3.5
4	R	242	GLU	3.5
7	U	51	PRO	3.4
3	Q	203	THR	3.4
2	P	203	SER	3.4
1	O	231	LYS	3.3
5	E	122	TYR	3.3
6	T	2	THR	3.3
6	T	181	GLU	3.2
1	O	2	THR	3.2
3	C	240	GLU	3.2
5	E	233	ILE	3.1
14	b	105	LYS	3.1
11	Y	147	ASP	3.1
6	F	181	GLU	3.1
3	Q	205	ALA	3.1
12	Z	167	LYS	3.0
3	Q	57	ILE	3.0
7	G	239	ILE	3.0
6	T	230	ASP	3.0
2	B	60	THR	3.0
1	A	250	LEU	3.0
6	T	244	ASN	3.0
11	Y	106	ARG	3.0
2	B	203	SER	3.0
2	B	217	LYS	2.9
4	D	125	LEU	2.9
7	G	181	LYS	2.9
7	U	222	ASP	2.9
3	Q	180	LYS	2.9
1	O	52	SER	2.9
3	C	60	SER	2.9
3	Q	55	THR	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	C	235	GLU	2.8
2	B	50	LYS	2.8
3	C	225	GLU	2.8
7	G	179	LYS	2.8
13	M	47	ASP	2.8
3	C	175	LYS	2.7
13	a	216	ASN	2.7
6	T	205	GLU	2.7
6	T	166	GLN	2.7
5	E	201	ARG	2.7
1	A	249	ALA	2.7
3	Q	60	SER	2.7
12	Z	172	LEU	2.7
1	A	248	GLU	2.6
3	C	203	THR	2.6
5	E	3	ASN	2.6
12	L	167	LYS	2.6
8	H	198	GLU	2.6
6	T	215	CYS	2.6
7	G	240	ALA	2.6
6	F	243	ILE	2.6
5	S	194	GLU	2.6
5	S	165	GLN	2.6
11	Y	183	ASP	2.6
12	Z	210	ASP	2.6
5	S	171	LEU	2.5
5	S	173	ARG	2.5
1	A	54	PRO	2.5
4	R	201	GLU	2.5
14	b	104	ASP	2.5
10	X	195	PHE	2.5
2	P	244	THR	2.5
6	T	180	PRO	2.5
12	Z	165	ASN	2.5
3	Q	167	LYS	2.5
14	N	104	ASP	2.5
5	S	3	ASN	2.5
14	b	195	GLN	2.5
5	E	218	ASP	2.5
3	Q	59	PRO	2.5
3	Q	173	LEU	2.4
3	Q	223	SER	2.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
7	U	206	GLY	2.4
12	Z	106	TYR	2.4
13	M	233	ILE	2.4
1	O	177	LYS	2.4
3	C	216	ASP	2.4
12	L	210	ASP	2.4
9	W	192	ASP	2.3
12	Z	174	TYR	2.3
3	Q	232	THR	2.3
8	V	198	GLU	2.3
11	Y	182	GLU	2.3
11	K	182	GLU	2.3
10	J	113	LYS	2.3
3	C	184	ALA	2.3
3	Q	221	ALA	2.3
1	A	229	THR	2.3
2	P	60	THR	2.3
4	D	2	ARG	2.3
4	R	2	ARG	2.3
5	S	122	TYR	2.3
6	T	237	ASP	2.3
12	L	165	ASN	2.3
5	S	233	ILE	2.3
3	Q	52	LEU	2.3
2	B	93	HIS	2.2
12	L	172	LEU	2.2
12	Z	1	GLN	2.2
2	B	204	ALA	2.2
3	C	37	LYS	2.2
5	E	210	LEU	2.2
14	N	195	GLN	2.2
4	R	217	GLN	2.2
6	F	201	GLU	2.2
9	I	131	GLU	2.2
4	D	1	ASP	2.2
2	P	225	TYR	2.2
5	S	207	VAL	2.2
3	C	181	GLU	2.2
5	E	54	GLU	2.2
2	P	217	LYS	2.2
8	V	217	ILE	2.2
2	B	238	LEU	2.2

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
5	S	123	GLY	2.2
6	F	217	LEU	2.2
3	C	51	LYS	2.2
1	O	229	THR	2.2
5	S	54	GLU	2.2
2	B	230	LYS	2.1
5	E	190	LYS	2.1
14	N	105	LYS	2.1
5	E	225	ASP	2.1
7	G	51	PRO	2.1
6	T	178	HIS	2.1
1	O	203	GLU	2.1
4	D	117	GLU	2.1
3	C	27	ARG	2.1
6	F	180	PRO	2.1
3	Q	181	GLU	2.1
12	Z	80	ASN	2.1
3	C	201	VAL	2.1
3	C	231	VAL	2.1
3	C	232	THR	2.1
7	U	207	THR	2.1
12	L	174	TYR	2.1
5	S	225	ASP	2.1
10	X	72	ASP	2.1
5	S	187	GLU	2.1
10	J	193	ASP	2.1
7	G	188	GLU	2.1
2	B	240	LYS	2.1
3	Q	27	ARG	2.1
6	F	204	LYS	2.1
7	U	237	VAL	2.1
4	R	169	GLU	2.1
7	U	178	LYS	2.1
3	Q	201	VAL	2.0
5	E	121	SER	2.0
7	G	241	GLU	2.0
7	U	241	GLU	2.0
7	G	68	ARG	2.0
7	U	183	ASP	2.0
2	P	204	ALA	2.0
11	Y	151	GLU	2.0
11	K	106	ARG	2.0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
11	K	202	GLU	2.0
2	B	52	THR	2.0
5	E	217	LYS	2.0
11	K	212	GLY	2.0
12	L	106	TYR	2.0
8	V	145	ASP	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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	DPP	g	2	6/7	0.79	0.18	55,56,57,59	0
15	DPP	c	2	6/7	0.83	0.25	58,59,61,63	0
15	DPN	c	4	11/12	0.89	0.24	49,50,52,55	0
15	DPN	g	4	11/12	0.89	0.23	49,51,54,56	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
19	54L	c	101	6/7	0.77	0.47	60,66,69,71	0
19	54L	g	101	6/7	0.88	0.67	50,58,62,66	0
16	MG	G	301	1/1	0.89	0.07	43,43,43,43	0
16	MG	I	301	1/1	0.91	0.17	63,63,63,63	0
16	MG	Z	301	1/1	0.92	0.45	60,60,60,60	0
16	MG	V	301	1/1	0.93	0.19	58,58,58,58	0
18	MES	K	302	12/12	0.94	0.22	31,34,49,49	0
16	MG	K	301	1/1	0.95	0.07	52,52,52,52	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
16	MG	N	201	1/1	0.96	0.09	43,43,43,43	0
18	MES	Y	302	12/12	0.96	0.26	32,34,52,53	0
17	CL	G	302	1/1	0.98	0.07	39,39,39,39	0
16	MG	Y	301	1/1	0.99	0.05	47,47,47,47	0
17	CL	U	301	1/1	0.99	0.07	42,42,42,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.