



# wwPDB X-ray Structure Validation Summary Report

Jun 16, 2024 – 10:20 PM EDT

PDB ID : 5NIF  
Title : Yeast 20S proteasome in complex with Blm-pep activator  
Authors : Witkowska, J.; Grudnik, P.; Golik, P.; Dubin, G.; Jankowska, E.  
Deposited on : 2017-03-23  
Resolution : 3.00 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the  symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), 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

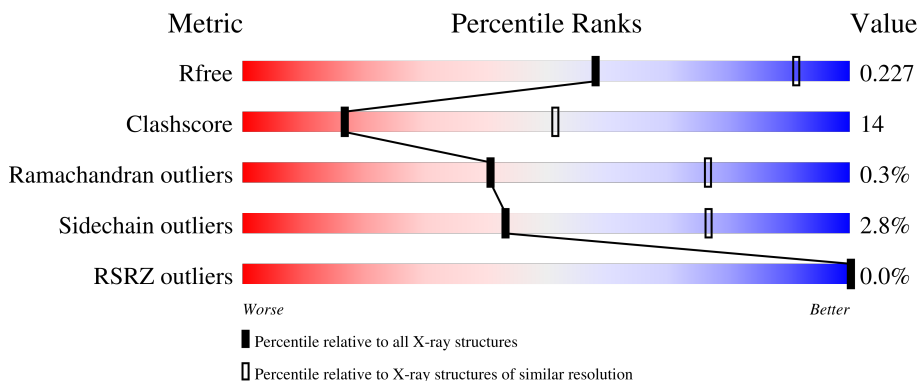
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	252	 67% 27% . .
1	O	252	 67% 27% . .
2	B	250	 73% 26% .
2	P	250	 70% 29% .
3	C	258	 60% 32% . 7%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	258	66% 28% 6%
4	D	254	66% 28% 5%
4	R	254	66% 28% 5%
5	E	260	64% 26% 9%
5	S	260	63% 27% 9%
6	F	234	75% 24% .
6	T	234	70% 27% ..
7	G	288	60% 23% . 15%
7	U	288	60% 23% . 15%
8	H	215	66% 24% . 9%
8	V	215	69% 21% 9%
9	I	261	61% 23% . 15%
9	W	261	61% 23% . 15%
10	J	205	72% 26% .
10	X	205	75% 24%
11	K	198	65% 32% ..
11	Y	198	69% 27% ..
12	L	287	55% 19% 26%
12	Z	287	55% 19% 26%
13	1	241	69% 22% . 8%
13	M	241	65% 27% 8%
14	2	266	64% 22% . 12%
14	N	266	65% 20% . 13%
15	3	14	14% 29% 57%
15	4	14	21% 21% 57%

## 2 Entry composition

There are 20 unique types of molecules in this entry. The entry contains 49177 atoms, of which 104 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-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	241	Total 1910	C 1217	N 317	O 368	S 8	0	2	0
1	O	241	Total 1862	C 1183	N 312	O 359	S 8	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	250	Total 1882	C 1200	N 306	O 372	S 4	0	0	0
2	P	250	Total 1904	C 1212	N 312	O 377	S 3	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	241	Total 1848	C 1167	N 305	O 373	S 3	0	0	0
3	Q	242	Total 1844	C 1163	N 307	O 371	S 3	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	241	Total 1821	C 1135	N 316	O 366	S 4	0	0	0
4	R	241	Total 1837	C 1147	N 321	O 365	S 4	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	236	Total	C	N	O	S	0	0	0
			1801	1133	295	366	7			
5	S	236	Total	C	N	O	S	0	0	0
			1800	1129	299	365	7			

- Molecule 6 is a protein called Proteasome subunit alpha type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	232	Total	C	N	O	S	0	0	0
			1756	1104	300	348	4			
6	T	232	Total	C	N	O	S	0	0	0
			1769	1110	307	348	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	244	Total	C	N	O	S	0	0	0
			1865	1184	324	353	4			
7	U	244	Total	C	N	O	S	0	0	0
			1882	1196	327	355	4			

- Molecule 8 is a protein called Proteasome subunit beta type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	196	Total	C	N	O	S	0	0	0
			1504	950	248	299	7			
8	V	196	Total	C	N	O	S	0	0	0
			1508	953	249	299	7			

- Molecule 9 is a protein called Proteasome subunit beta type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	222	Total	C	N	O	S	0	0	0
			1678	1059	292	320	7			
9	W	222	Total	C	N	O	S	0	0	0
			1680	1058	292	323	7			

- Molecule 10 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	204	Total	C	N	O	S	0	0	0
			1566	1000	253	305	8			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	X	204	Total	C	N	O	S	0	0	0
			1563	997	253	305	8			

- Molecule 11 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	196	Total	C	N	O	S	0	0	0
			1549	986	263	294	6			
11	Y	196	Total	C	N	O	S	0	0	0
			1565	995	265	299	6			

- Molecule 12 is a protein called Proteasome subunit beta type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	212	Total	C	N	O	S	0	0	0
			1634	1039	278	310	7			
12	Z	212	Total	C	N	O	S	0	0	0
			1626	1033	274	312	7			

- Molecule 13 is a protein called Proteasome subunit beta type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	222	Total	C	N	O	S	0	0	0
			1742	1104	299	335	4			
13	1	222	Total	C	N	O	S	0	0	0
			1740	1103	298	335	4			

- Molecule 14 is a protein called Proteasome subunit beta type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	232	Total	C	N	O	S	0	0	0
			1817	1150	311	349	7			
14	2	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			

- Molecule 15 is a protein called TRP-ARG-SER-TYR-TYR-ALA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	3	6	Total	C	N	O	0	0	0
			52	33	9	10			
15	4	6	Total	C	N	O	0	0	0
			52	33	9	10			

- Molecule 16 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	A	1	Total Cl 1 1	0	0
16	K	2	Total Cl 2 2	0	0
16	N	1	Total Cl 1 1	0	0
16	O	1	Total Cl 1 1	0	0
16	P	1	Total Cl 1 1	0	0
16	R	1	Total Cl 1 1	0	0
16	S	1	Total Cl 1 1	0	0
16	V	1	Total Cl 1 1	0	0
16	2	3	Total Cl 3 3	0	0

- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	C	1	Total Mg 1 1	0	0
17	D	1	Total Mg 1 1	0	0
17	H	1	Total Mg 1 1	0	0
17	I	1	Total Mg 1 1	0	0
17	J	1	Total Mg 1 1	0	0
17	N	2	Total Mg 2 2	0	0
17	Q	1	Total Mg 1 1	0	0

- Molecule 18 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
18	C	1	10	2	6	2	0	0
18	F	1	10	2	6	2	0	0
18	G	1	10	2	6	2	0	0
18	I	1	10	2	6	2	0	0
18	L	1	10	2	6	2	0	0
18	M	1	10	2	6	2	0	0
18	O	1	10	2	6	2	0	0
18	O	1	10	2	6	2	0	0
18	U	1	10	2	6	2	0	0
18	W	1	10	2	6	2	0	0
18	W	1	10	2	6	2	0	0
18	W	1	10	2	6	2	0	0
18	Y	1	10	2	6	2	0	0
18	Y	1	10	2	6	2	0	0

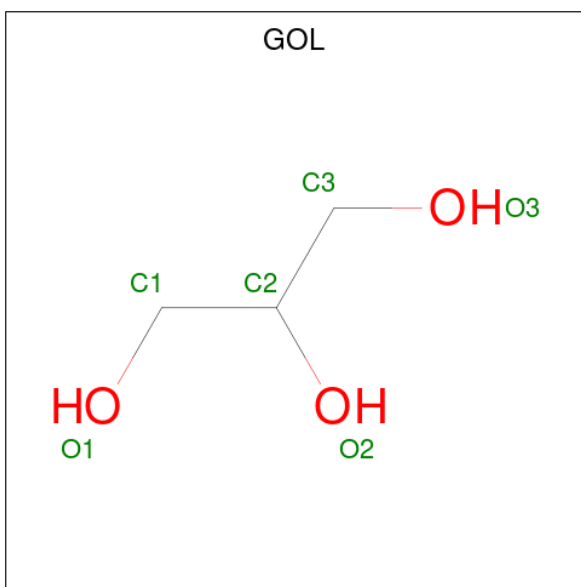
*Continued on next page...*



Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
18	1	1	Total	C	H	O	0	0
			10	2	6	2		
18	2	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 19 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
19	J	1	Total	C	O		0	0
			6	3	3			
19	L	1	Total	C	O		0	0
			6	3	3			
19	M	1	Total	C	H	O	0	0
			14	3	8	3		
19	Q	1	Total	C	O		0	0
			6	3	3			
19	V	1	Total	C	O		0	0
			6	3	3			
19	Z	1	Total	C	O		0	0
			6	3	3			
19	1	1	Total	C	O		0	0
			6	3	3			

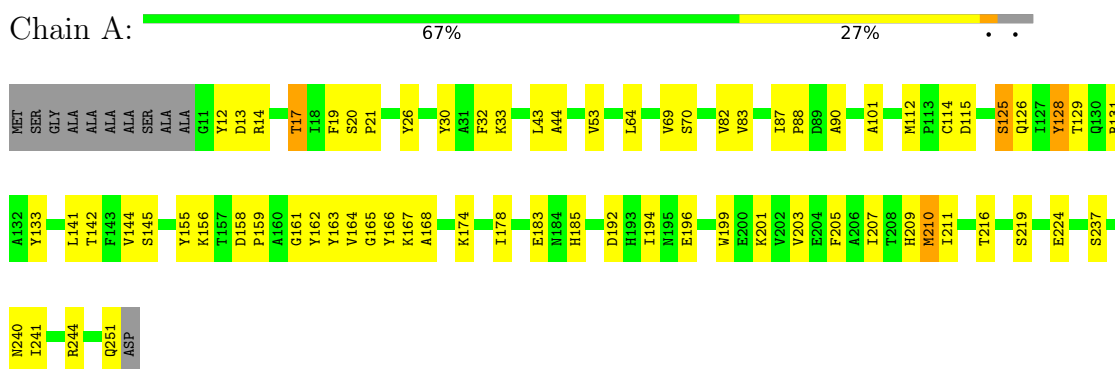
- Molecule 20 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
20	A	2	Total O 2 2	0	0
20	C	2	Total O 2 2	0	0
20	D	1	Total O 1 1	0	0
20	G	2	Total O 2 2	0	0
20	H	8	Total O 8 8	0	0
20	I	2	Total O 2 2	0	0
20	J	1	Total O 1 1	0	0
20	L	1	Total O 1 1	0	0
20	M	4	Total O 4 4	0	0
20	N	4	Total O 4 4	0	0
20	O	4	Total O 4 4	0	0
20	U	2	Total O 2 2	0	0
20	V	5	Total O 5 5	0	0
20	W	2	Total O 2 2	0	0
20	X	2	Total O 2 2	0	0
20	Z	6	Total O 6 6	0	0
20	1	9	Total O 9 9	0	0
20	2	9	Total O 9 9	0	0

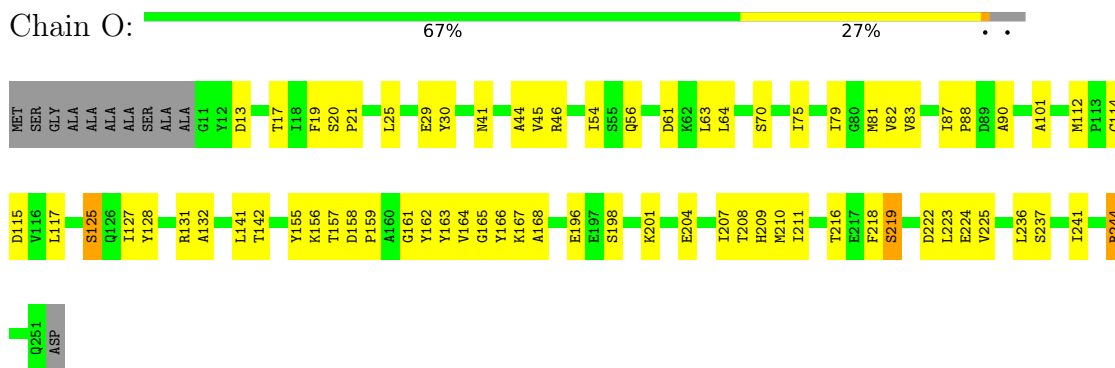
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry 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.

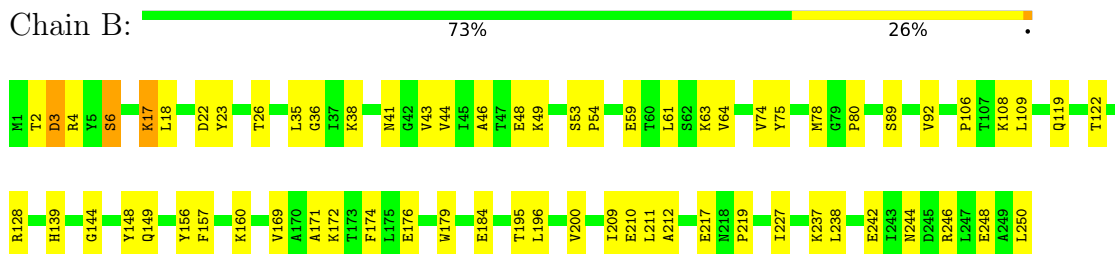
- Molecule 1: Proteasome subunit alpha type-1



- Molecule 1: Proteasome subunit alpha type-1



- Molecule 2: Proteasome subunit alpha type-2



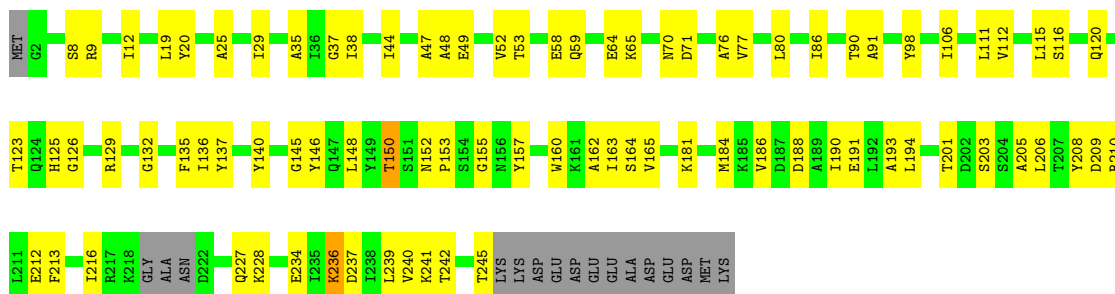
- Molecule 2: Proteasome subunit alpha type-2

Chain P:  70% 29%



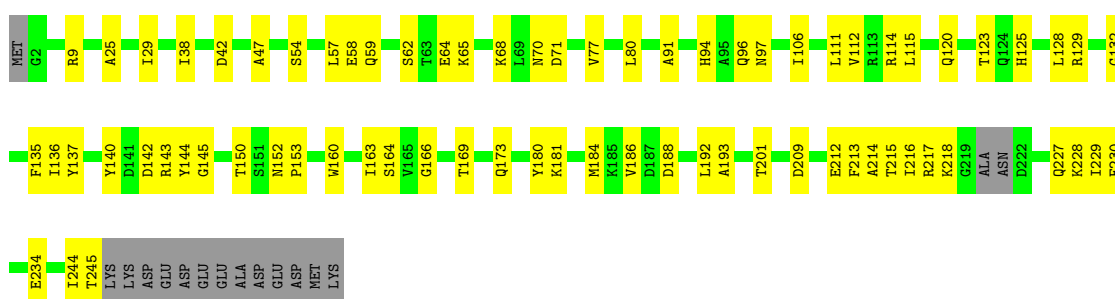
• Molecule 3: Proteasome subunit alpha type-3

Chain C:  60% 32% 7%



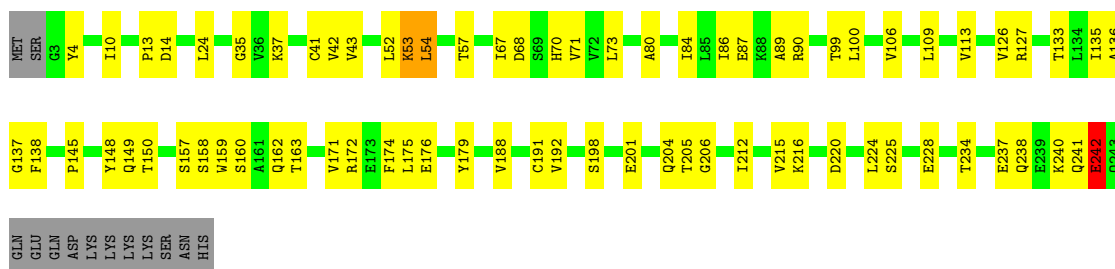
• Molecule 3: Proteasome subunit alpha type-3

Chain Q:  66% 28% 6%



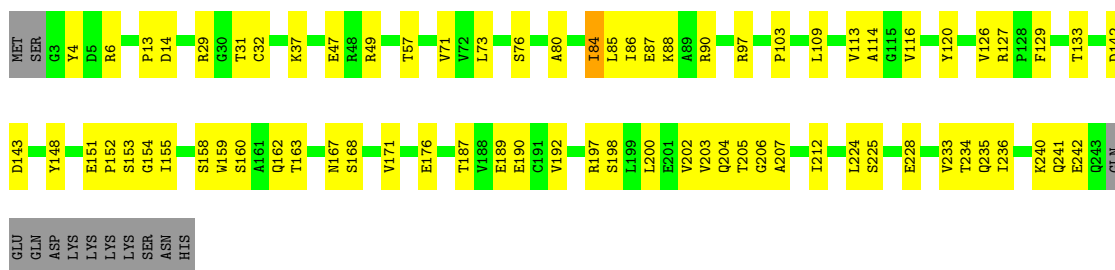
• Molecule 4: Proteasome subunit alpha type-4

Chain D:  66% 28% 5%



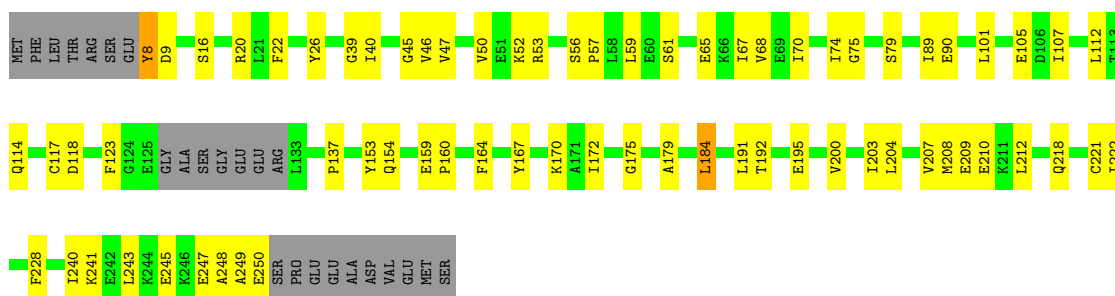
- Molecule 4: Proteasome subunit alpha type-4

Chain R: 66% 28% 5%



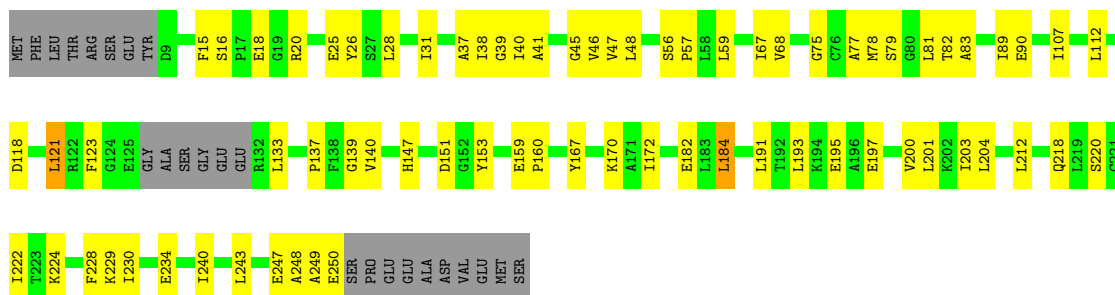
- Molecule 5: Proteasome subunit alpha type-5

Chain E: 64% 26% 9%



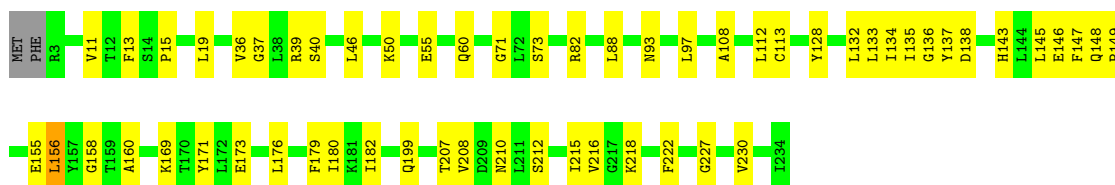
- Molecule 5: Proteasome subunit alpha type-5

Chain S: 63% 27% 9%



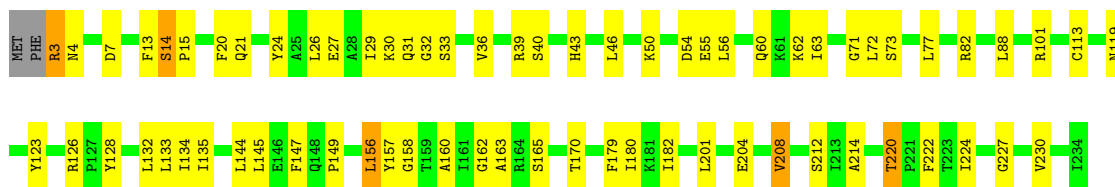
- Molecule 6: Proteasome subunit alpha type-6

Chain F: 75% 24% 1%



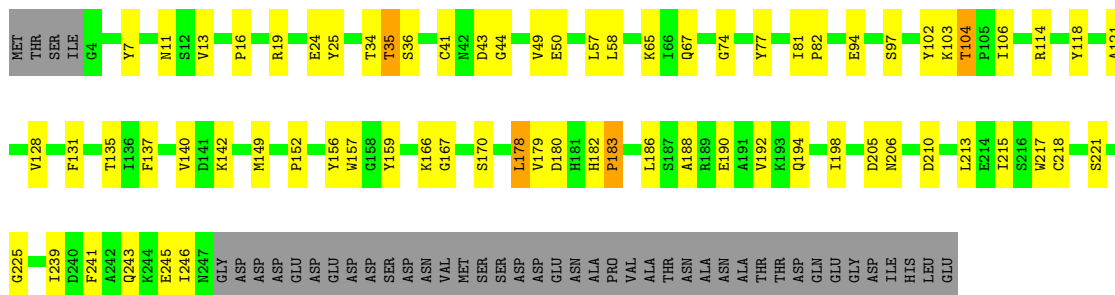
- Molecule 6: Proteasome subunit alpha type-6

Chain T:  70% 27% ..



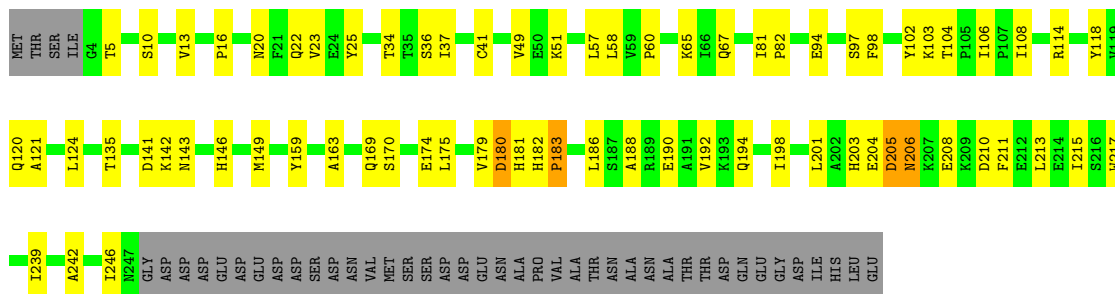
• Molecule 7: Probable proteasome subunit alpha type-7

Chain G:  60% 23% 15%



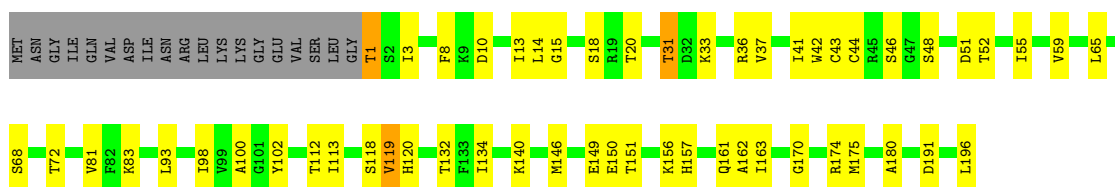
• Molecule 7: Probable proteasome subunit alpha type-7

Chain U:  60% 23% 15%



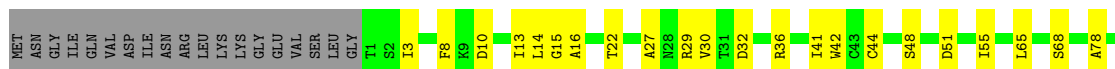
• Molecule 8: Proteasome subunit beta type-1

Chain H:  66% 24% 9%

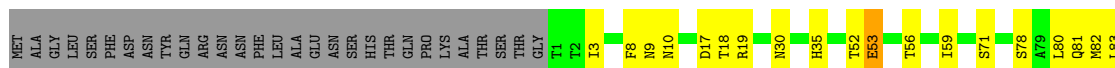


• Molecule 8: Proteasome subunit beta type-1

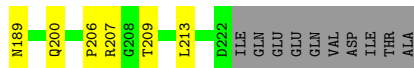
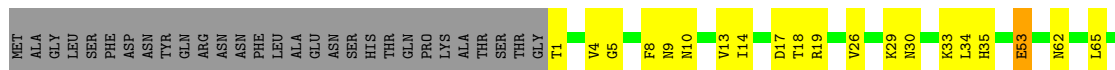
Chain V:  69% 21% 9%



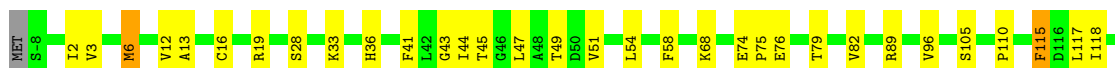
- Molecule 9: Proteasome subunit beta type-2



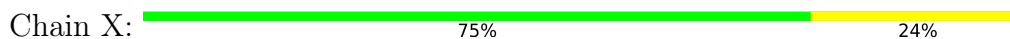
- Molecule 9: Proteasome subunit beta type-2



- Molecule 10: Proteasome subunit beta type-3

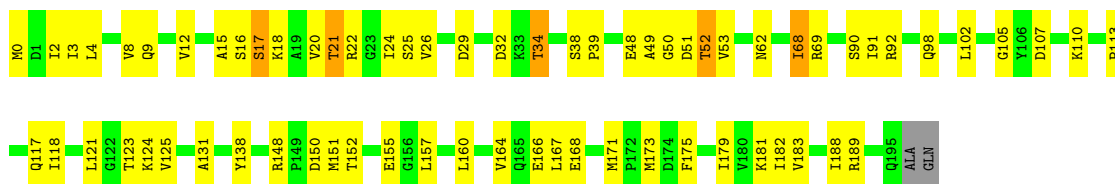


- Molecule 10: Proteasome subunit beta type-3

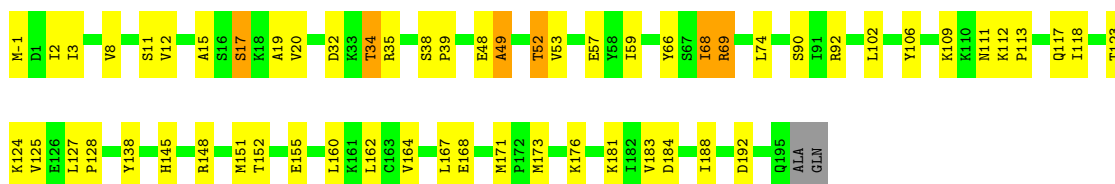




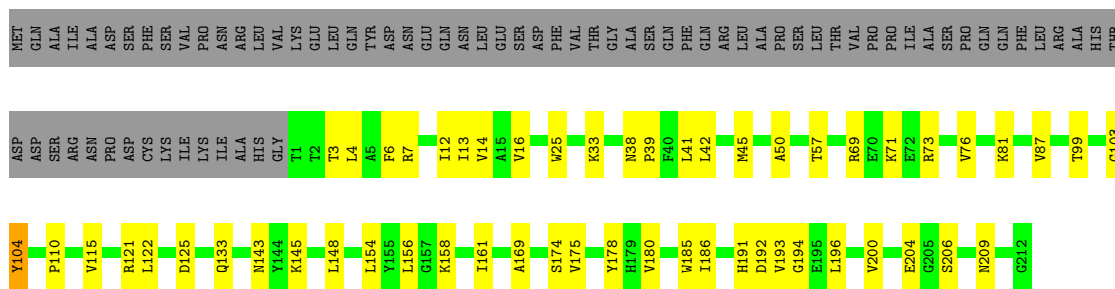
- Molecule 11: Proteasome subunit beta type-4



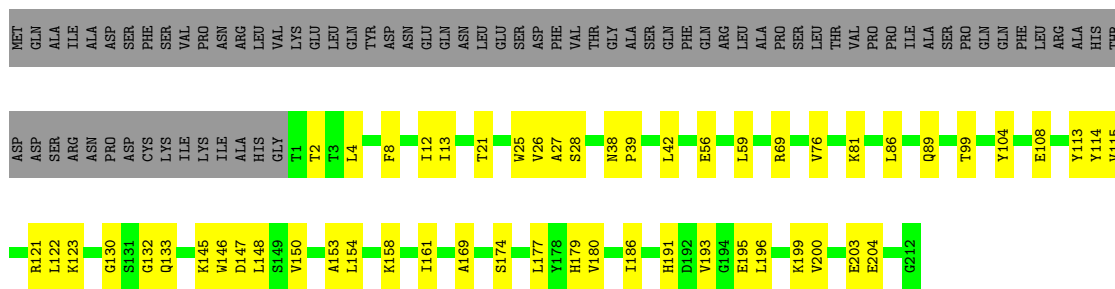
- Molecule 11: Proteasome subunit beta type-4



- Molecule 12: Proteasome subunit beta type-5



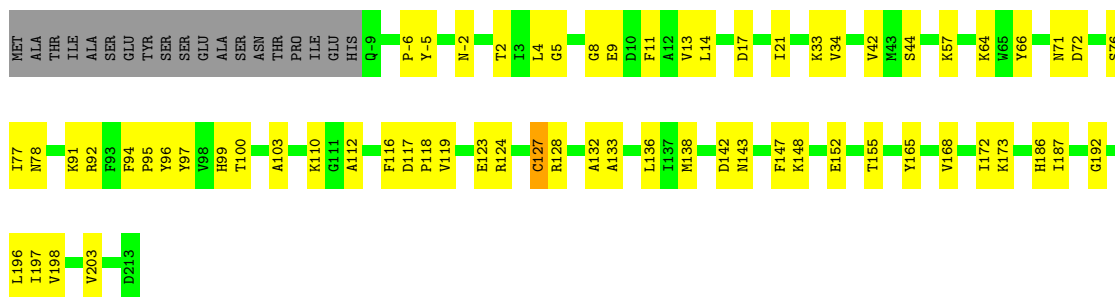
- Molecule 12: Proteasome subunit beta type-5



- Molecule 13: Proteasome subunit beta type-6

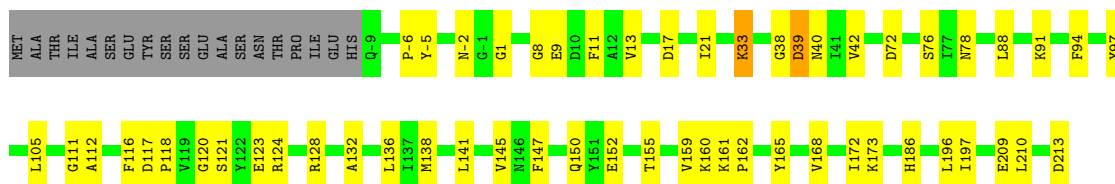


Chain M: 



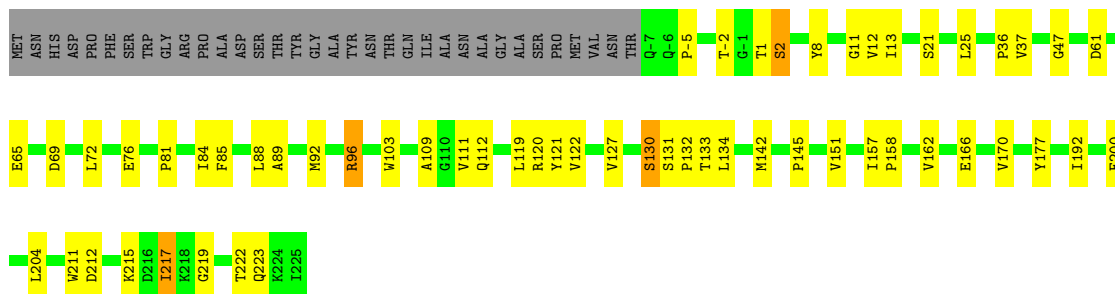
- Molecule 13: Proteasome subunit beta type-6

Chain 1: 



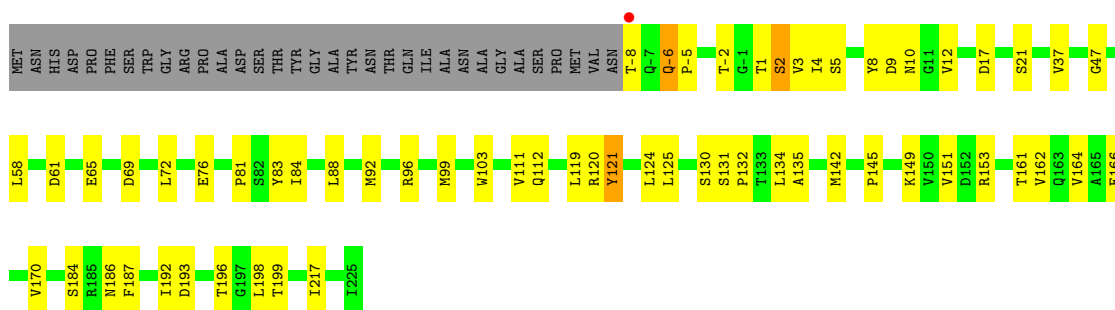
- Molecule 14: Proteasome subunit beta type-7

Chain N: 



- Molecule 14: Proteasome subunit beta type-7

Chain 2: 



- Molecule 15: TRP-ARG-SER-TYR-TYR-ALA

Chain 3:  14% 29% 57%

LYS	TYR	PHE	THR	GLY	SER	LYS	LEU	W9	R10	Y13	A14
-----	-----	-----	-----	-----	-----	-----	-----	----	-----	-----	-----

● Molecule 15: TRP-ARG-SER-TYR-TYR-ALA

Chain 4:  21% 21% 57%

LYS	TYR	PHE	THR	GLY	SER	LYS	LEU	W9	R10	Y13	A14
-----	-----	-----	-----	-----	-----	-----	-----	----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.04Å 302.01Å 143.84Å 90.00° 112.55° 90.00°	Depositor
Resolution (Å)	50.01 – 3.00 50.01 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.01-3.00) 100.0 (50.01-3.00)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, $R_{free}$	0.174 , 0.227 0.174 , 0.227	Depositor DCC
$R_{free}$ test set	10362 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.7	Xtrriage
Anisotropy	0.620	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	49177	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.34% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, CL, MG, EDO

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	0.47	0/1955	0.63	0/2652
1	O	0.47	0/1899	0.65	0/2581
2	B	0.46	0/1919	0.63	0/2605
2	P	0.47	0/1941	0.65	0/2631
3	C	0.44	0/1877	0.64	0/2550
3	Q	0.44	0/1873	0.65	0/2545
4	D	0.44	0/1850	0.61	0/2518
4	R	0.45	0/1866	0.63	0/2538
5	E	0.46	0/1826	0.63	0/2465
5	S	0.45	0/1824	0.63	1/2462 (0.0%)
6	F	0.44	0/1782	0.61	0/2412
6	T	0.45	0/1796	0.64	0/2431
7	G	0.52	0/1905	0.64	0/2581
7	U	0.46	0/1922	0.65	0/2599
8	H	0.47	0/1533	0.63	0/2078
8	V	0.47	0/1537	0.65	0/2082
9	I	0.53	2/1709 (0.1%)	0.67	2/2318 (0.1%)
9	W	0.51	2/1711 (0.1%)	0.68	2/2322 (0.1%)
10	J	0.46	0/1596	0.66	0/2159
10	X	0.46	0/1593	0.62	1/2156 (0.0%)
11	K	0.49	0/1577	0.66	0/2129
11	Y	0.48	0/1593	0.69	1/2148 (0.0%)
12	L	0.44	0/1671	0.61	0/2263
12	Z	0.45	0/1663	0.64	0/2255
13	1	0.46	0/1778	0.63	0/2403
13	M	0.48	1/1780 (0.1%)	0.62	0/2405
14	2	0.46	0/1855	0.65	0/2514
14	N	0.47	0/1848	0.71	0/2504
15	3	0.38	0/53	0.60	0/69
15	4	0.43	0/53	0.55	0/69
All	All	0.47	5/49785 (0.0%)	0.64	7/67444 (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
7	U	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	53	GLU	CB-CG	7.75	1.66	1.52
9	I	53	GLU	CG-CD	7.44	1.63	1.51
9	W	53	GLU	CG-CD	7.20	1.62	1.51
9	W	53	GLU	CB-CG	6.91	1.65	1.52
13	M	127	CYS	CB-SG	-6.15	1.71	1.82

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	Y	69	ARG	NE-CZ-NH1	-6.25	117.17	120.30
9	W	53	GLU	OE1-CD-OE2	-5.84	116.29	123.30
9	I	170	GLY	N-CA-C	-5.65	98.98	113.10
5	S	121	LEU	CA-CB-CG	-5.47	102.71	115.30
9	W	170	GLY	N-CA-C	-5.41	99.58	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	U	180	ASP	Peptide

## 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	1910	0	1883	62	1
1	O	1862	0	1804	64	0
2	B	1882	0	1863	59	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	P	1904	0	1897	64	1
3	C	1848	0	1796	69	0
3	Q	1844	0	1782	63	0
4	D	1821	0	1744	73	0
4	R	1837	0	1788	61	0
5	E	1801	0	1759	73	0
5	S	1800	0	1756	56	0
6	F	1756	0	1736	49	0
6	T	1769	0	1751	61	0
7	G	1865	0	1814	45	0
7	U	1882	0	1853	56	0
8	H	1504	0	1464	40	0
8	V	1508	0	1475	39	0
9	I	1678	0	1682	53	0
9	W	1680	0	1677	47	0
10	J	1566	0	1529	49	0
10	X	1563	0	1523	40	0
11	K	1549	0	1543	57	0
11	Y	1565	0	1571	55	0
12	L	1634	0	1573	46	0
12	Z	1626	0	1551	34	0
13	1	1740	0	1662	55	0
13	M	1742	0	1669	53	0
14	2	1824	0	1832	50	0
14	N	1817	0	1825	53	0
15	3	52	0	42	7	0
15	4	52	0	42	3	0
16	2	3	0	0	0	0
16	A	1	0	0	0	0
16	K	2	0	0	0	0
16	N	1	0	0	0	0
16	O	1	0	0	0	0
16	P	1	0	0	0	0
16	R	1	0	0	0	0
16	S	1	0	0	0	0
16	V	1	0	0	0	0
17	C	1	0	0	0	0
17	D	1	0	0	0	0
17	H	1	0	0	0	0
17	I	1	0	0	0	0
17	J	1	0	0	0	0
17	N	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	Q	1	0	0	0	0
18	1	4	6	6	0	0
18	2	4	6	6	0	0
18	C	4	6	6	2	0
18	F	4	6	6	1	0
18	G	4	6	6	1	0
18	I	4	6	6	0	0
18	L	4	6	6	0	0
18	M	4	6	6	0	0
18	O	8	12	12	2	0
18	U	4	6	6	1	0
18	W	12	18	18	0	0
18	Y	8	12	12	2	0
19	1	6	0	8	0	0
19	J	6	0	8	0	0
19	L	6	0	8	0	0
19	M	6	8	8	0	0
19	Q	6	0	8	1	0
19	V	6	0	8	2	0
19	Z	6	0	8	0	0
20	1	9	0	0	0	0
20	2	9	0	0	1	0
20	A	2	0	0	3	0
20	C	2	0	0	0	0
20	D	1	0	0	0	0
20	G	2	0	0	0	0
20	H	8	0	0	1	0
20	I	2	0	0	1	0
20	J	1	0	0	0	0
20	L	1	0	0	0	0
20	M	4	0	0	0	0
20	N	4	0	0	0	0
20	O	4	0	0	0	0
20	U	2	0	0	0	0
20	V	5	0	0	0	0
20	W	2	0	0	0	0
20	X	2	0	0	0	0
20	Z	6	0	0	0	0
All	All	49073	104	48038	1369	1

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 14.

The worst 5 of 1369 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)
3:C:201:THR:HG22	3:C:203:SER:H	1.09	1.10
3:C:123:THR:HG22	4:D:127:ARG:HH21	1.16	1.07
8:H:20:THR:HG23	8:H:31:THR:HG21	1.29	1.07
1:O:101:ALA:HA	1:O:112:MET:HE2	1.30	1.07
1:A:129:THR:HG22	2:B:128:ARG:HH21	1.21	1.05

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:185:HIS:ND1	2:P:201:GLU:OE2[2_546]	1.89	0.31

## 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	241/252 (96%)	232 (96%)	9 (4%)	0	100	100
1	O	239/252 (95%)	233 (98%)	6 (2%)	0	100	100
2	B	248/250 (99%)	231 (93%)	15 (6%)	2 (1%)	19	57
2	P	248/250 (99%)	233 (94%)	13 (5%)	2 (1%)	19	57
3	C	237/258 (92%)	222 (94%)	15 (6%)	0	100	100
3	Q	238/258 (92%)	225 (94%)	13 (6%)	0	100	100
4	D	239/254 (94%)	226 (95%)	11 (5%)	2 (1%)	19	57
4	R	239/254 (94%)	227 (95%)	10 (4%)	2 (1%)	19	57
5	E	232/260 (89%)	220 (95%)	12 (5%)	0	100	100
5	S	232/260 (89%)	218 (94%)	13 (6%)	1 (0%)	34	72
6	F	230/234 (98%)	213 (93%)	17 (7%)	0	100	100

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	T	230/234 (98%)	217 (94%)	13 (6%)	0	100	100
7	G	242/288 (84%)	226 (93%)	15 (6%)	1 (0%)	34	72
7	U	242/288 (84%)	228 (94%)	12 (5%)	2 (1%)	19	57
8	H	194/215 (90%)	183 (94%)	11 (6%)	0	100	100
8	V	194/215 (90%)	187 (96%)	7 (4%)	0	100	100
9	I	220/261 (84%)	209 (95%)	11 (5%)	0	100	100
9	W	220/261 (84%)	212 (96%)	8 (4%)	0	100	100
10	J	202/205 (98%)	195 (96%)	7 (4%)	0	100	100
10	X	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
11	K	194/198 (98%)	182 (94%)	10 (5%)	2 (1%)	15	53
11	Y	194/198 (98%)	184 (95%)	8 (4%)	2 (1%)	15	53
12	L	210/287 (73%)	201 (96%)	9 (4%)	0	100	100
12	Z	210/287 (73%)	204 (97%)	6 (3%)	0	100	100
13	1	220/241 (91%)	209 (95%)	9 (4%)	2 (1%)	17	55
13	M	220/241 (91%)	209 (95%)	10 (4%)	1 (0%)	29	68
14	2	231/266 (87%)	222 (96%)	9 (4%)	0	100	100
14	N	230/266 (86%)	222 (96%)	8 (4%)	0	100	100
15	3	4/14 (29%)	4 (100%)	0	0	100	100
15	4	4/14 (29%)	3 (75%)	1 (25%)	0	100	100
All	All	6286/6966 (90%)	5973 (95%)	294 (5%)	19 (0%)	41	76

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	53	SER
4	D	53	LYS
11	K	50	GLY
2	P	53	SER
4	R	240	LYS

### 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	205/210 (98%)	196 (96%)	9 (4%)	28	65
1	O	194/210 (92%)	188 (97%)	6 (3%)	40	75
2	B	201/209 (96%)	196 (98%)	5 (2%)	47	79
2	P	205/209 (98%)	200 (98%)	5 (2%)	49	79
3	C	193/216 (89%)	190 (98%)	3 (2%)	62	86
3	Q	190/216 (88%)	189 (100%)	1 (0%)	88	96
4	D	194/226 (86%)	189 (97%)	5 (3%)	46	78
4	R	199/226 (88%)	196 (98%)	3 (2%)	65	87
5	E	190/215 (88%)	185 (97%)	5 (3%)	46	78
5	S	189/215 (88%)	184 (97%)	5 (3%)	46	78
6	F	185/193 (96%)	181 (98%)	4 (2%)	52	81
6	T	187/193 (97%)	180 (96%)	7 (4%)	34	70
7	G	193/239 (81%)	183 (95%)	10 (5%)	23	59
7	U	197/239 (82%)	187 (95%)	10 (5%)	24	60
8	H	160/178 (90%)	154 (96%)	6 (4%)	33	69
8	V	161/178 (90%)	159 (99%)	2 (1%)	71	90
9	I	179/214 (84%)	175 (98%)	4 (2%)	52	81
9	W	180/214 (84%)	176 (98%)	4 (2%)	52	81
10	J	167/173 (96%)	163 (98%)	4 (2%)	49	79
10	X	167/173 (96%)	163 (98%)	4 (2%)	49	79
11	K	168/175 (96%)	161 (96%)	7 (4%)	30	66
11	Y	173/175 (99%)	167 (96%)	6 (4%)	36	71
12	L	166/235 (71%)	164 (99%)	2 (1%)	71	90
12	Z	165/235 (70%)	162 (98%)	3 (2%)	59	85
13	1	180/201 (90%)	178 (99%)	2 (1%)	73	90
13	M	181/201 (90%)	175 (97%)	6 (3%)	38	73
14	2	199/224 (89%)	190 (96%)	9 (4%)	27	64
14	N	198/224 (88%)	192 (97%)	6 (3%)	41	75
15	3	4/12 (33%)	4 (100%)	0	100	100
15	4	4/12 (33%)	4 (100%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	5174/5840 (89%)	5031 (97%)	143 (3%)	43 77

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	W	71	SER
10	X	115	PHE
12	Z	147	ASP
9	I	169	SER
9	I	71	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
9	I	30	ASN
9	I	189	ASN
9	I	194	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 43 ligands modelled in this entry, 20 are monoatomic - leaving 23 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
18	EDO	Y	202	-	3,3,3	0.42	0	2,2,2	0.76	0
18	EDO	W	302	-	3,3,3	0.62	0	2,2,2	0.19	0
18	EDO	2	304	-	3,3,3	0.60	0	2,2,2	0.34	0
19	GOL	1	301	-	5,5,5	0.36	0	5,5,5	0.97	0
19	GOL	V	202	-	5,5,5	0.56	0	5,5,5	0.91	0
18	EDO	M	302	-	3,3,3	0.65	0	2,2,2	0.07	0
18	EDO	O	303	-	3,3,3	0.78	0	2,2,2	0.06	0
18	EDO	O	302	-	3,3,3	0.63	0	2,2,2	0.39	0
18	EDO	U	301	-	3,3,3	0.32	0	2,2,2	0.72	0
19	GOL	J	201	-	5,5,5	0.36	0	5,5,5	0.35	0
18	EDO	1	302	-	3,3,3	0.66	0	2,2,2	0.13	0
19	GOL	L	301	-	5,5,5	0.38	0	5,5,5	0.28	0
18	EDO	W	301	-	3,3,3	0.95	0	2,2,2	0.44	0
18	EDO	I	302	-	3,3,3	0.53	0	2,2,2	0.22	0
18	EDO	C	302	-	3,3,3	0.56	0	2,2,2	0.50	0
19	GOL	M	301	-	5,5,5	0.52	0	5,5,5	0.45	0
18	EDO	L	302	-	3,3,3	0.66	0	2,2,2	0.14	0
19	GOL	Q	301	-	5,5,5	0.42	0	5,5,5	0.48	0
18	EDO	G	301	-	3,3,3	0.45	0	2,2,2	0.36	0
18	EDO	F	301	-	3,3,3	0.28	0	2,2,2	0.57	0
18	EDO	Y	201	-	3,3,3	0.65	0	2,2,2	0.15	0
19	GOL	Z	301	-	5,5,5	0.37	0	5,5,5	0.25	0
18	EDO	W	303	-	3,3,3	0.69	0	2,2,2	0.13	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
18	EDO	Y	202	-	-	0/1/1/1	-
18	EDO	W	302	-	-	0/1/1/1	-
18	EDO	2	304	-	-	0/1/1/1	-
19	GOL	1	301	-	-	2/4/4/4	-
19	GOL	V	202	-	-	3/4/4/4	-
18	EDO	M	302	-	-	0/1/1/1	-
18	EDO	O	303	-	-	1/1/1/1	-
18	EDO	O	302	-	-	0/1/1/1	-
18	EDO	U	301	-	-	0/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
19	GOL	J	201	-	-	1/4/4/4	-
18	EDO	I	302	-	-	0/1/1/1	-
19	GOL	L	301	-	-	2/4/4/4	-
18	EDO	W	301	-	-	1/1/1/1	-
18	EDO	I	302	-	-	0/1/1/1	-
18	EDO	C	302	-	-	0/1/1/1	-
19	GOL	M	301	-	-	2/4/4/4	-
18	EDO	L	302	-	-	0/1/1/1	-
19	GOL	Q	301	-	-	2/4/4/4	-
18	EDO	G	301	-	-	0/1/1/1	-
18	EDO	F	301	-	-	0/1/1/1	-
18	EDO	Y	201	-	-	0/1/1/1	-
19	GOL	Z	301	-	-	2/4/4/4	-
18	EDO	W	303	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
19	M	301	GOL	O1-C1-C2-O2
19	M	301	GOL	O1-C1-C2-C3
19	Q	301	GOL	C1-C2-C3-O3
19	Q	301	GOL	O2-C2-C3-O3
19	Z	301	GOL	O1-C1-C2-O2

There are no ring outliers.

8 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
19	V	202	GOL	2	0
18	O	302	EDO	2	0
18	U	301	EDO	1	0
18	C	302	EDO	2	0
19	Q	301	GOL	1	0
18	G	301	EDO	1	0
18	F	301	EDO	1	0
18	Y	201	EDO	2	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 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	241/252 (95%)	-0.78	0 100 100	22, 36, 56, 72	0
1	O	241/252 (95%)	-0.80	0 100 100	22, 34, 56, 74	0
2	B	250/250 (100%)	-0.65	0 100 100	25, 42, 71, 81	0
2	P	250/250 (100%)	-0.74	0 100 100	23, 36, 61, 82	0
3	C	241/258 (93%)	-0.65	0 100 100	26, 40, 70, 96	0
3	Q	242/258 (93%)	-0.73	0 100 100	24, 40, 67, 83	0
4	D	241/254 (94%)	-0.41	0 100 100	27, 49, 86, 100	1 (0%)
4	R	241/254 (94%)	-0.49	0 100 100	23, 44, 82, 95	1 (0%)
5	E	236/260 (90%)	-0.68	0 100 100	27, 44, 66, 83	0
5	S	236/260 (90%)	-0.75	0 100 100	26, 40, 61, 81	0
6	F	232/234 (99%)	-0.62	0 100 100	26, 42, 63, 73	0
6	T	232/234 (99%)	-0.64	0 100 100	27, 43, 64, 77	0
7	G	244/288 (84%)	-0.70	0 100 100	20, 36, 64, 81	0
7	U	244/288 (84%)	-0.70	0 100 100	24, 37, 65, 85	2 (0%)
8	H	196/215 (91%)	-0.85	0 100 100	22, 30, 46, 55	0
8	V	196/215 (91%)	-0.83	0 100 100	19, 29, 46, 54	0
9	I	222/261 (85%)	-0.80	0 100 100	22, 34, 52, 88	0
9	W	222/261 (85%)	-0.79	0 100 100	22, 33, 50, 91	0
10	J	204/205 (99%)	-0.85	0 100 100	22, 34, 50, 79	0
10	X	204/205 (99%)	-0.79	0 100 100	22, 35, 50, 73	0
11	K	196/198 (98%)	-0.77	0 100 100	22, 35, 50, 77	0
11	Y	196/198 (98%)	-0.82	0 100 100	24, 35, 51, 87	0
12	L	212/287 (73%)	-0.77	0 100 100	25, 36, 53, 68	0
12	Z	212/287 (73%)	-0.77	0 100 100	23, 35, 51, 65	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	1	222/241 (92%)	-0.76	0 100 100	24, 34, 53, 70	0
13	M	222/241 (92%)	-0.71	0 100 100	21, 35, 57, 74	0
14	2	233/266 (87%)	-0.82	1 (0%) 92 79	19, 32, 46, 59	0
14	N	232/266 (87%)	-0.86	0 100 100	19, 30, 44, 57	0
15	3	6/14 (42%)	-0.21	0 100 100	45, 57, 69, 80	0
15	4	6/14 (42%)	-0.14	0 100 100	49, 69, 70, 73	0
All	All	6352/6966 (91%)	-0.73	1 (0%) 100 100	19, 37, 63, 100	4 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
14	2	-8	THR	2.1

## 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 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
18	EDO	W	301	4/4	0.53	0.34	36,45,47,54	0
18	EDO	W	303	4/4	0.75	0.27	44,58,70,70	0
16	CL	2	301	1/1	0.83	0.20	60,60,60,60	0
17	MG	I	301	1/1	0.87	0.23	38,38,38,38	0
18	EDO	Y	202	4/4	0.87	0.24	47,57,64,75	0
19	GOL	Q	301	6/6	0.88	0.30	38,49,50,55	0
16	CL	A	301	1/1	0.89	0.17	53,53,53,53	0
16	CL	O	301	1/1	0.91	0.10	47,47,47,47	0

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
18	EDO	1	302	4/4	0.91	0.22	33,47,50,56	0
18	EDO	L	302	4/4	0.91	0.29	38,45,49,49	0
19	GOL	J	201	6/6	0.92	0.21	40,46,53,54	0
18	EDO	Y	201	4/4	0.92	0.19	32,54,69,69	0
19	GOL	V	202	6/6	0.92	0.26	31,40,48,50	0
16	CL	K	201	1/1	0.93	0.10	46,46,46,46	0
18	EDO	O	303	4/4	0.93	0.18	38,46,58,58	0
18	EDO	2	304	4/4	0.93	0.33	48,56,63,63	0
16	CL	V	201	1/1	0.93	0.08	44,44,44,44	0
17	MG	Q	302	1/1	0.93	0.12	29,29,29,29	0
18	EDO	C	302	4/4	0.93	0.31	41,48,57,66	0
19	GOL	Z	301	6/6	0.93	0.18	40,58,61,61	0
17	MG	N	302	1/1	0.94	0.18	33,33,33,33	0
17	MG	C	301	1/1	0.94	0.11	30,30,30,30	0
19	GOL	L	301	6/6	0.94	0.14	41,46,48,53	0
17	MG	H	201	1/1	0.95	0.10	20,20,20,20	0
18	EDO	G	301	4/4	0.95	0.22	41,49,55,65	0
18	EDO	I	302	4/4	0.95	0.20	34,49,67,67	0
16	CL	2	303	1/1	0.95	0.10	54,54,54,54	0
18	EDO	U	301	4/4	0.96	0.20	38,45,53,63	0
16	CL	P	301	1/1	0.96	0.08	41,41,41,41	0
16	CL	R	301	1/1	0.96	0.10	60,60,60,60	0
16	CL	K	202	1/1	0.96	0.07	40,40,40,40	0
19	GOL	1	301	6/6	0.96	0.17	34,44,48,49	0
17	MG	N	303	1/1	0.97	0.09	33,33,33,33	0
19	GOL	M	301	6/6	0.97	0.17	42,52,56,67	0
17	MG	D	301	1/1	0.97	0.07	31,31,31,31	0
18	EDO	M	302	4/4	0.97	0.33	31,37,42,44	0
18	EDO	W	302	4/4	0.97	0.13	38,45,49,49	0
18	EDO	O	302	4/4	0.97	0.24	33,43,51,51	0
16	CL	S	301	1/1	0.98	0.06	45,45,45,45	0
17	MG	J	202	1/1	0.98	0.09	32,32,32,32	0
18	EDO	F	301	4/4	0.98	0.17	34,39,42,44	0
16	CL	N	301	1/1	0.99	0.06	46,46,46,46	0
16	CL	2	302	1/1	0.99	0.10	47,47,47,47	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.