



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 02:35 PM BST

PDB ID : 1Z7Q  
Title : Crystal structure of the 20s proteasome from yeast in complex with the proteasome activator PA26 from Trypanosome brucei at 3.2 angstroms resolution  
Authors : Forster, A.; Whitby, F.G.; Hill, C.P.  
Deposited on : 2005-03-26  
Resolution : 3.22 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13
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)	:	2.13

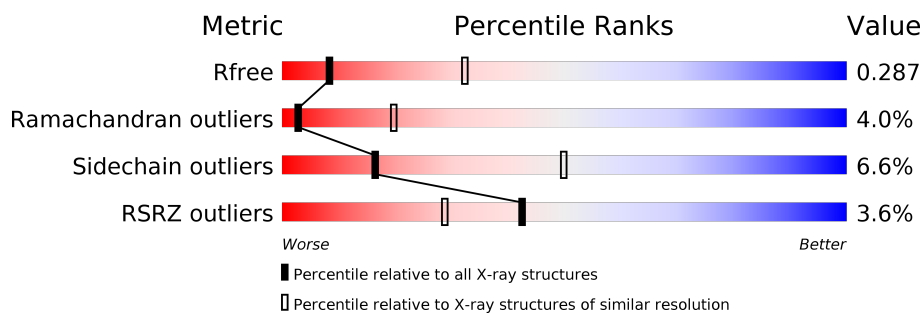
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.22 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	1335 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	252	<div> <div>4%</div> <div>83% 12% . .</div> </div>
1	O	252	<div> <div>4%</div> <div>83% 12% . .</div> </div>
2	B	250	<div> <div>3%</div> <div>86% 13% .</div> </div>
2	P	250	<div> <div>4%</div> <div>86% 12% .</div> </div>
3	C	258	<div> <div>2%</div> <div>81% 12% . 6%</div> </div>
3	Q	258	<div> <div>4%</div> <div>81% 12% . 6%</div> </div>
4	D	254	<div> <div>2%</div> <div>85% 10% 5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	R	254	
5	E	260	
5	S	260	
6	F	234	
6	T	234	
7	G	288	
7	U	288	
8	H	196	
8	V	196	
9	I	222	
9	W	222	
10	J	205	
10	X	205	
11	K	198	
11	Y	198	
12	L	212	
12	Z	212	
13	M	222	
13	a	222	
14	N	233	
14	b	233	
15	c	231	
15	d	231	
15	e	231	
15	f	231	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
15	g	231	
15	h	231	
15	i	231	
15	j	231	
15	k	231	
15	l	231	
15	m	231	
15	n	231	
15	o	231	
15	p	231	

## 2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 74222 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Proteasome component C7-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			
1	O	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			

- Molecule 2 is a protein called Proteasome component Y7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	249	Total	C	N	O	S	0	0	0
			1907	1214	314	376	3			
2	P	249	Total	C	N	O	S	0	0	0
			1907	1214	314	376	3			

- Molecule 3 is a protein called Proteasome component Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	243	Total	C	N	O	S	0	0	0
			1900	1199	320	378	3			
3	Q	243	Total	C	N	O	S	0	0	0
			1900	1199	320	378	3			

- Molecule 4 is a protein called Proteasome component PRE6.

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

- Molecule 5 is a protein called Proteasome component PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	245	Total	C	N	O	S	0	0	0
			1888	1179	317	385	7			
5	S	245	Total	C	N	O	S	0	0	0
			1888	1179	317	385	7			

- Molecule 6 is a protein called Proteasome component PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	234	Total	C	N	O	S	0	0	0
			1803	1134	313	351	5			
6	T	234	Total	C	N	O	S	0	0	0
			1803	1134	313	351	5			

- Molecule 7 is a protein called Proteasome component C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	243	Total	C	N	O	S	0	0	0
			1892	1203	329	356	4			
7	U	243	Total	C	N	O	S	0	0	0
			1892	1203	329	356	4			

- Molecule 8 is a protein called Proteasome component PRE3.

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

- Molecule 9 is a protein called Proteasome component PUP1.

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

- Molecule 10 is a protein called Proteasome component PUP3.

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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 11 is a protein called Proteasome component C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			
11	Y	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			

- Molecule 12 is a protein called Proteasome component PRE2.

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

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 13 is a protein called Potential proteasome component C5.

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

- Molecule 14 is a protein called Proteasome component PRE4.

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

- Molecule 15 is a protein called proteasome activator protein PA26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	c	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	d	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	e	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	f	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	g	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	h	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	i	230	Total 1760	C 1101	N 310	O 343	S 6	249	0	0
15	j	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	k	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	l	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	m	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	n	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	o	230	Total 1760	C 1101	N 310	O 343	S 6	219	0	0
15	p	230	Total 1760	C 1101	N 310	O 343	S 6	249	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
c	1049	VAL	THR	CONFLICT	UNP Q9U8G2
c	?	-	SER	DELETION	UNP Q9U8G2
c	1171	GLY	-	INSERTION	UNP Q9U8G2
d	1049	VAL	THR	CONFLICT	UNP Q9U8G2
d	?	-	SER	DELETION	UNP Q9U8G2
d	1171	GLY	-	INSERTION	UNP Q9U8G2
e	1049	VAL	THR	CONFLICT	UNP Q9U8G2
e	?	-	SER	DELETION	UNP Q9U8G2
e	1171	GLY	-	INSERTION	UNP Q9U8G2
f	1049	VAL	THR	CONFLICT	UNP Q9U8G2
f	?	-	SER	DELETION	UNP Q9U8G2
f	1171	GLY	-	INSERTION	UNP Q9U8G2

*Continued on next page...*

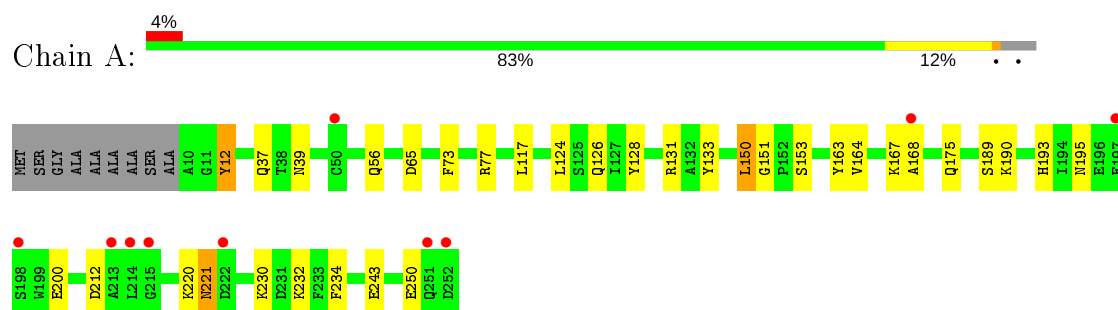
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
g	1049	VAL	THR	CONFLICT	UNP Q9U8G2
g	?	-	SER	DELETION	UNP Q9U8G2
g	1171	GLY	-	INSERTION	UNP Q9U8G2
h	1049	VAL	THR	CONFLICT	UNP Q9U8G2
h	?	-	SER	DELETION	UNP Q9U8G2
h	1171	GLY	-	INSERTION	UNP Q9U8G2
i	1049	VAL	THR	CONFLICT	UNP Q9U8G2
i	?	-	SER	DELETION	UNP Q9U8G2
i	1171	GLY	-	INSERTION	UNP Q9U8G2
j	1049	VAL	THR	CONFLICT	UNP Q9U8G2
j	?	-	SER	DELETION	UNP Q9U8G2
j	1171	GLY	-	INSERTION	UNP Q9U8G2
k	1049	VAL	THR	CONFLICT	UNP Q9U8G2
k	?	-	SER	DELETION	UNP Q9U8G2
k	1171	GLY	-	INSERTION	UNP Q9U8G2
l	1049	VAL	THR	CONFLICT	UNP Q9U8G2
l	?	-	SER	DELETION	UNP Q9U8G2
l	1171	GLY	-	INSERTION	UNP Q9U8G2
m	1049	VAL	THR	CONFLICT	UNP Q9U8G2
m	?	-	SER	DELETION	UNP Q9U8G2
m	1171	GLY	-	INSERTION	UNP Q9U8G2
n	1049	VAL	THR	CONFLICT	UNP Q9U8G2
n	?	-	SER	DELETION	UNP Q9U8G2
n	1171	GLY	-	INSERTION	UNP Q9U8G2
o	1049	VAL	THR	CONFLICT	UNP Q9U8G2
o	?	-	SER	DELETION	UNP Q9U8G2
o	1171	GLY	-	INSERTION	UNP Q9U8G2
p	1049	VAL	THR	CONFLICT	UNP Q9U8G2
p	?	-	SER	DELETION	UNP Q9U8G2
p	1171	GLY	-	INSERTION	UNP Q9U8G2

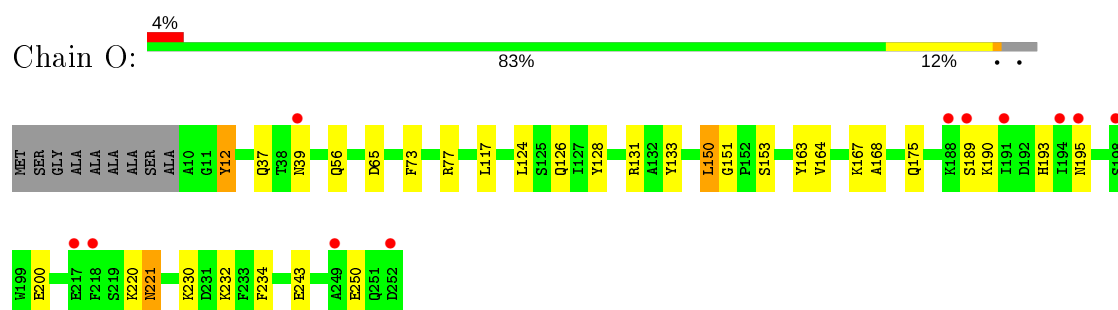
### 3 Residue-property plots [i](#)

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.

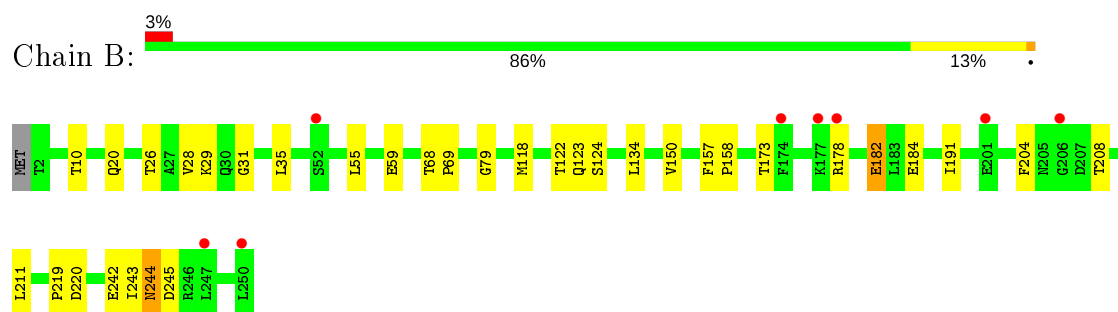
- Molecule 1: Proteasome component C7-alpha



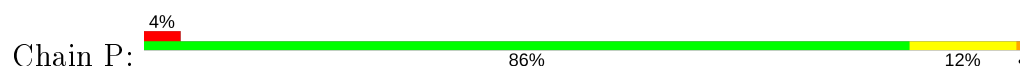
- Molecule 1: Proteasome component C7-alpha

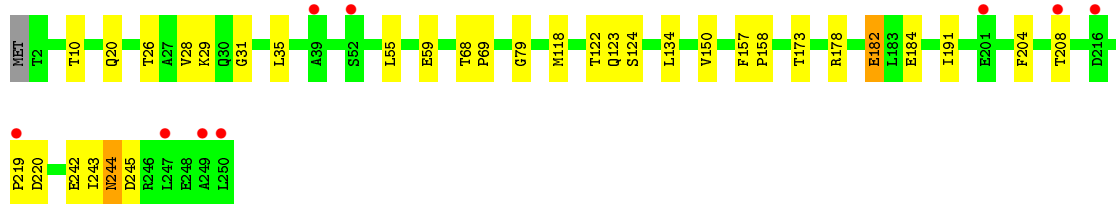


- Molecule 2: Proteasome component Y7

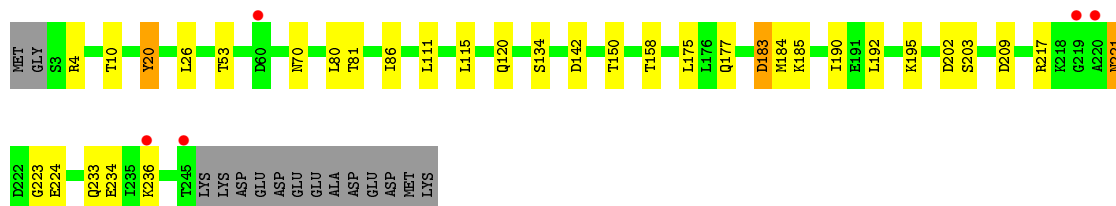
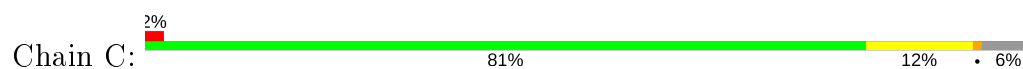


- Molecule 2: Proteasome component Y7

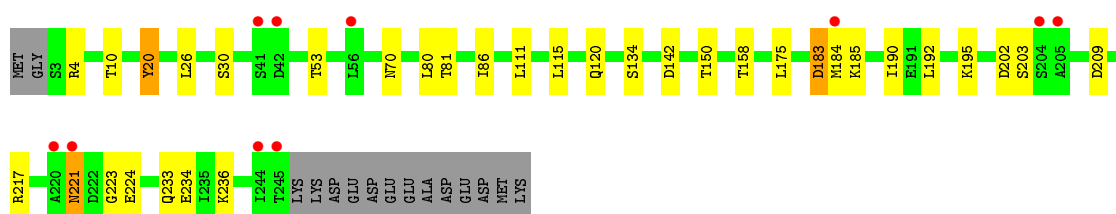
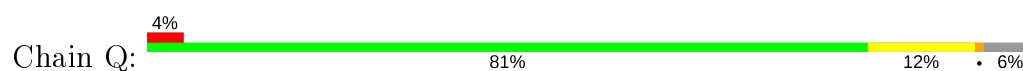




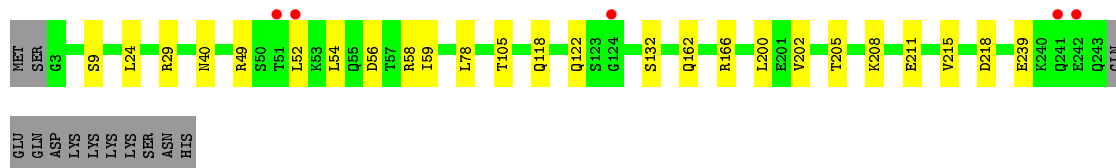
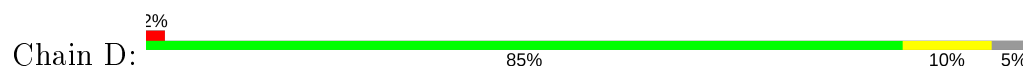
• Molecule 3: Proteasome component Y13



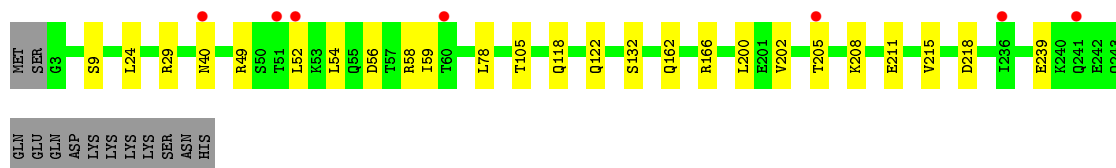
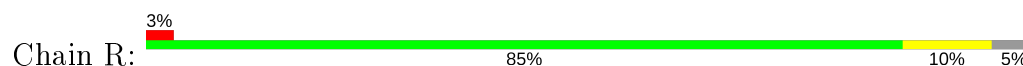
• Molecule 3: Proteasome component Y13



• Molecule 4: Proteasome component PRE6

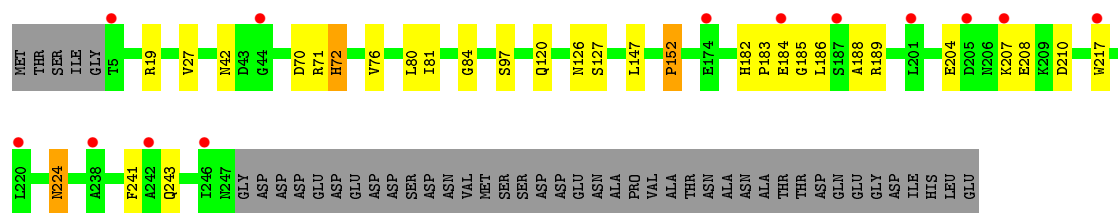


• Molecule 4: Proteasome component PRE6

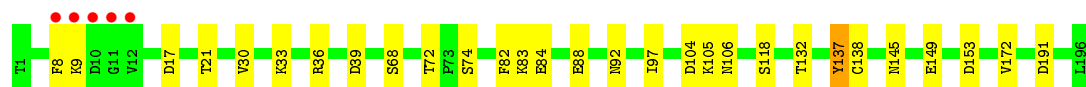
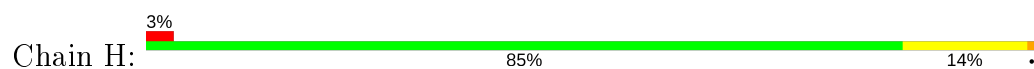


• Molecule 5: Proteasome component PUP2

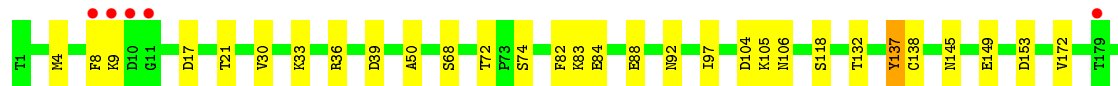
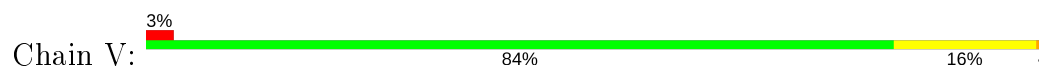




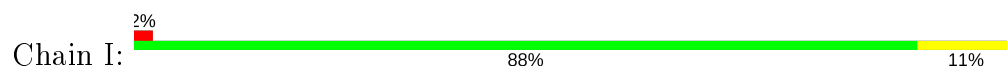
• Molecule 8: Proteasome component PRE3



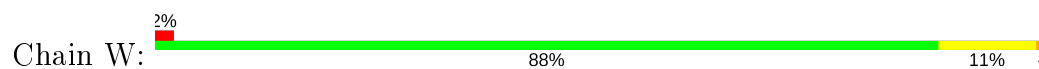
• Molecule 8: Proteasome component PRE3



• Molecule 9: Proteasome component PUP1



• Molecule 9: Proteasome component PUP1

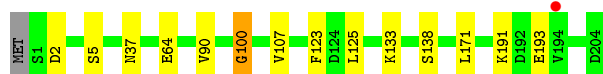


• Molecule 10: Proteasome component PUP3




• Molecule 10: Proteasome component PUP3

Chain X:  93% 6%




- Molecule 11: Proteasome component C11

Chain K:  3% 88% 12%



- Molecule 11: Proteasome component C11

Chain Y:  2% 89% 11%



- Molecule 12: Proteasome component PRE2

Chain L:  94% 6%



- Molecule 12: Proteasome component PRE2

Chain Z:  93% 6%




- Molecule 13: Potential proteasome component C5

Chain M:  89% 11%

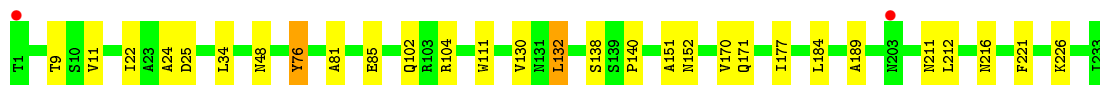
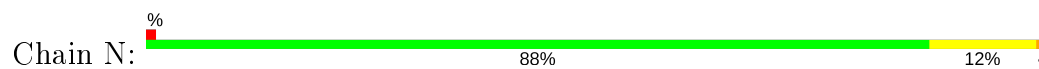


- Molecule 13: Potential proteasome component C5

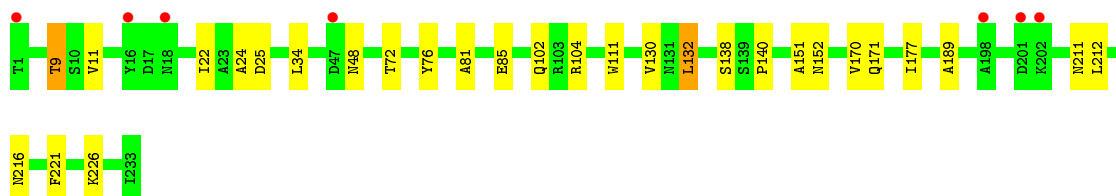
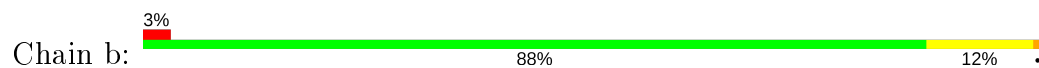
Chain a:  89% 11%



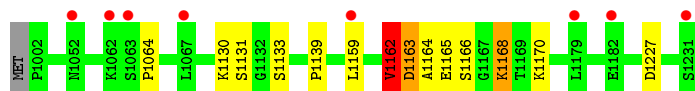
- Molecule 14: Proteasome component PRE4



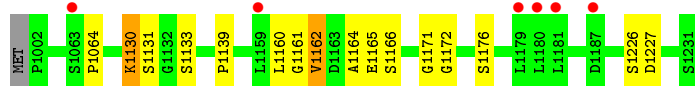
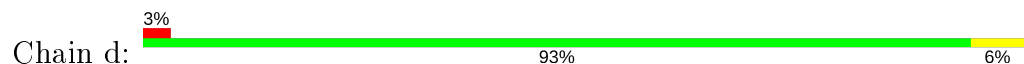
- Molecule 14: Proteasome component PRE4



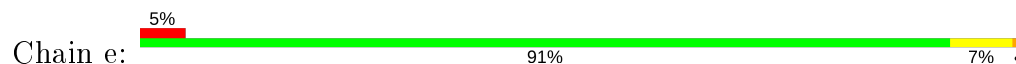
- Molecule 15: proteasome activator protein PA26



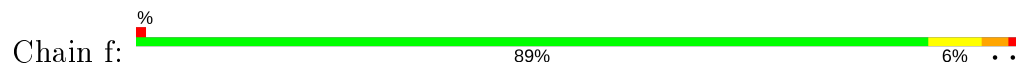
- Molecule 15: proteasome activator protein PA26



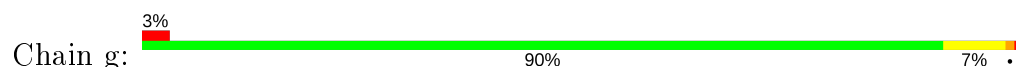
- Molecule 15: proteasome activator protein PA26

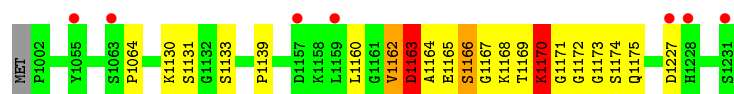


- Molecule 15: proteasome activator protein PA26

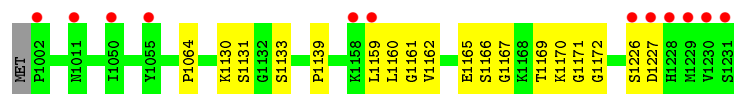


- Molecule 15: proteasome activator protein PA26





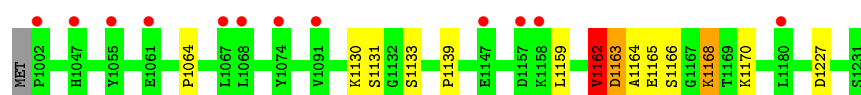
- Molecule 15: proteasome activator protein PA26



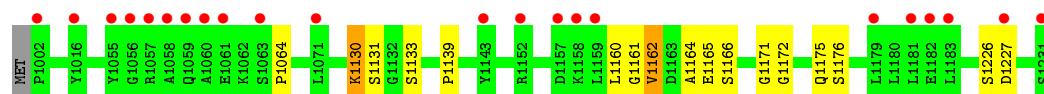
- Molecule 15: proteasome activator protein PA26



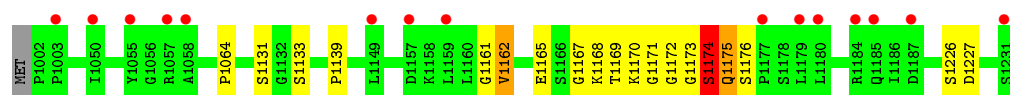
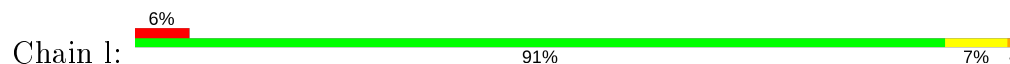
- Molecule 15: proteasome activator protein PA26



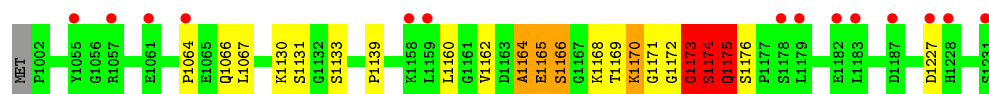
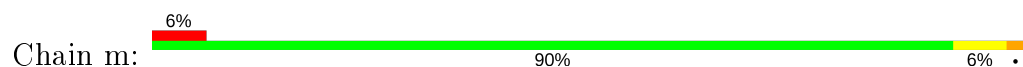
- Molecule 15: proteasome activator protein PA26



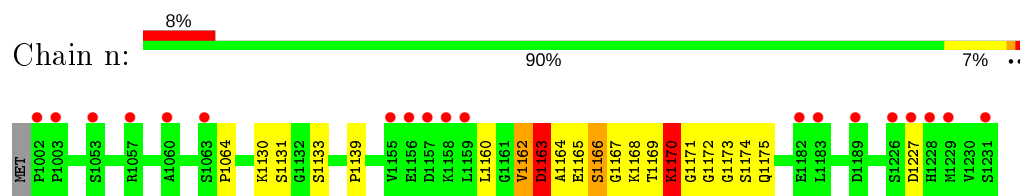
- Molecule 15: proteasome activator protein PA26



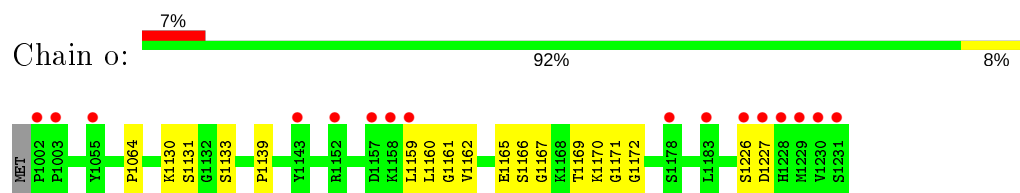
- Molecule 15: proteasome activator protein PA26



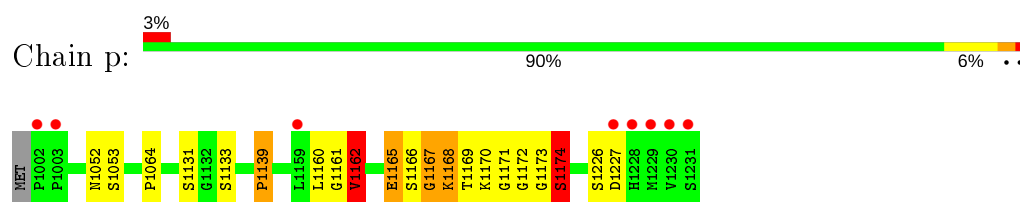
- Molecule 15: proteasome activator protein PA26



- Molecule 15: proteasome activator protein PA26



- Molecule 15: proteasome activator protein PA26



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	192.96 Å   232.13 Å   296.77 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	39.80 – 3.22 39.84 – 3.22	Depositor EDS
% Data completeness (in resolution range)	89.1 (39.80-3.22) 88.2 (39.84-3.22)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.19 (at 3.25 Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.263   ,   0.308 0.251   ,   0.287	Depositor DCC
$R_{free}$ test set	1264 reflections (0.67%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	71.3	Xtriage
Anisotropy	0.529	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 53.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	74222	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% 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 ⓘ

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	0.50	0/1959	0.70	1/2652 (0.0%)
1	O	0.49	0/1959	0.70	1/2652 (0.0%)
2	B	0.47	0/1944	0.73	0/2632
2	P	0.46	0/1944	0.72	0/2632
3	C	0.47	0/1930	0.69	0/2613
3	Q	0.47	0/1930	0.69	0/2613
4	D	0.47	0/1919	0.69	0/2598
4	R	0.45	0/1919	0.69	0/2598
5	E	0.49	0/1914	0.70	0/2579
5	S	0.47	0/1914	0.70	0/2579
6	F	0.48	0/1831	0.71	1/2473 (0.0%)
6	T	0.48	0/1831	0.70	0/2473
7	G	0.47	0/1932	0.68	0/2609
7	U	0.46	0/1932	0.67	0/2609
8	H	0.51	0/1541	0.74	0/2087
8	V	0.48	0/1541	0.73	0/2087
9	I	0.49	0/1716	0.71	0/2326
9	W	0.48	0/1716	0.70	0/2326
10	J	0.49	0/1611	0.76	2/2174 (0.1%)
10	X	0.50	0/1611	0.75	1/2174 (0.0%)
11	K	0.56	0/1613	0.74	0/2173
11	Y	0.53	0/1613	0.73	0/2173
12	L	0.54	0/1683	0.74	1/2277 (0.0%)
12	Z	0.54	0/1683	0.74	1/2277 (0.0%)
13	M	0.49	0/1795	0.73	0/2420
13	a	0.50	0/1795	0.73	0/2420
14	N	0.49	0/1855	0.75	0/2514
14	b	0.49	0/1855	0.75	0/2514
15	c	0.57	2/1786 (0.1%)	0.77	6/2415 (0.2%)
15	d	0.70	9/1786 (0.5%)	0.88	15/2415 (0.6%)
15	e	0.82	11/1786 (0.6%)	1.01	22/2415 (0.9%)
15	f	0.74	5/1786 (0.3%)	1.19	18/2415 (0.7%)
15	g	0.91	9/1786 (0.5%)	1.17	20/2415 (0.8%)
15	h	0.77	7/1786 (0.4%)	0.94	20/2415 (0.8%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
15	i	1.22	17/1786 (1.0%)	1.43	28/2415 (1.2%)
15	j	0.56	2/1786 (0.1%)	0.77	6/2415 (0.2%)
15	k	0.69	9/1786 (0.5%)	0.88	16/2415 (0.7%)
15	l	0.81	11/1786 (0.6%)	1.01	21/2415 (0.9%)
15	m	0.73	5/1786 (0.3%)	1.18	18/2415 (0.7%)
15	n	0.91	9/1786 (0.5%)	1.17	20/2415 (0.8%)
15	o	0.76	7/1786 (0.4%)	0.94	20/2415 (0.8%)
15	p	1.23	17/1786 (1.0%)	1.43	28/2415 (1.2%)
All	All	0.63	120/75490 (0.2%)	0.85	266/102064 (0.3%)

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
1	A	0	1
1	O	0	1
3	C	0	1
3	Q	0	1
14	N	0	1
15	e	0	1
15	f	0	4
15	g	0	1
15	i	0	4
15	l	0	1
15	m	0	4
15	n	0	1
15	p	0	4
All	All	0	25

The worst 5 of 120 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	i	1172	GLY	C-N	32.26	1.91	1.33
15	p	1172	GLY	C-N	32.26	1.91	1.33
15	g	1163	ASP	C-O	14.32	1.50	1.23
15	n	1163	ASP	C-O	14.30	1.50	1.23
15	i	1173	GLY	N-CA	13.87	1.66	1.46

The worst 5 of 266 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	n	1174	SER	N-CA-CB	-31.23	63.65	110.50
15	g	1174	SER	N-CA-CB	-31.22	63.68	110.50
15	p	1162	VAL	O-C-N	-27.33	78.97	122.70
15	i	1162	VAL	O-C-N	-27.26	79.09	122.70
15	f	1173	GLY	O-C-N	-27.22	79.15	122.70

There are no chirality outliers.

5 of 25 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	12	TYR	Mainchain
3	C	20	TYR	Sidechain
14	N	76	TYR	Sidechain
1	O	12	TYR	Mainchain
3	Q	20	TYR	Sidechain

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

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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%)	175 (73%)	53 (22%)	13 (5%)	2	13
1	O	241/252 (96%)	174 (72%)	54 (22%)	13 (5%)	2	13
2	B	247/250 (99%)	192 (78%)	39 (16%)	16 (6%)	1	9
2	P	247/250 (99%)	193 (78%)	38 (15%)	16 (6%)	1	9
3	C	241/258 (93%)	192 (80%)	35 (14%)	14 (6%)	1	12
3	Q	241/258 (93%)	192 (80%)	35 (14%)	14 (6%)	1	12
4	D	239/254 (94%)	203 (85%)	28 (12%)	8 (3%)	4	24

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	R	239/254 (94%)	203 (85%)	28 (12%)	8 (3%)	4	24
5	E	243/260 (94%)	204 (84%)	27 (11%)	12 (5%)	2	16
5	S	243/260 (94%)	201 (83%)	30 (12%)	12 (5%)	2	16
6	F	232/234 (99%)	192 (83%)	33 (14%)	7 (3%)	4	27
6	T	232/234 (99%)	191 (82%)	33 (14%)	8 (3%)	3	23
7	G	241/288 (84%)	188 (78%)	37 (15%)	16 (7%)	1	9
7	U	241/288 (84%)	188 (78%)	37 (15%)	16 (7%)	1	9
8	H	194/196 (99%)	143 (74%)	38 (20%)	13 (7%)	1	8
8	V	194/196 (99%)	142 (73%)	37 (19%)	15 (8%)	1	6
9	I	220/222 (99%)	173 (79%)	37 (17%)	10 (4%)	2	17
9	W	220/222 (99%)	174 (79%)	35 (16%)	11 (5%)	2	15
10	J	202/205 (98%)	173 (86%)	24 (12%)	5 (2%)	5	31
10	X	202/205 (98%)	174 (86%)	23 (11%)	5 (2%)	5	31
11	K	196/198 (99%)	171 (87%)	20 (10%)	5 (3%)	5	30
11	Y	196/198 (99%)	170 (87%)	21 (11%)	5 (3%)	5	30
12	L	210/212 (99%)	183 (87%)	24 (11%)	3 (1%)	11	45
12	Z	210/212 (99%)	185 (88%)	22 (10%)	3 (1%)	11	45
13	M	220/222 (99%)	193 (88%)	24 (11%)	3 (1%)	11	45
13	a	220/222 (99%)	192 (87%)	25 (11%)	3 (1%)	11	45
14	N	231/233 (99%)	176 (76%)	39 (17%)	16 (7%)	1	8
14	b	231/233 (99%)	176 (76%)	39 (17%)	16 (7%)	1	8
15	c	228/231 (99%)	207 (91%)	14 (6%)	7 (3%)	4	26
15	d	228/231 (99%)	206 (90%)	16 (7%)	6 (3%)	5	30
15	e	228/231 (99%)	206 (90%)	16 (7%)	6 (3%)	5	30
15	f	228/231 (99%)	200 (88%)	17 (8%)	11 (5%)	2	16
15	g	228/231 (99%)	202 (89%)	18 (8%)	8 (4%)	3	23
15	h	228/231 (99%)	209 (92%)	15 (7%)	4 (2%)	8	39
15	i	228/231 (99%)	207 (91%)	16 (7%)	5 (2%)	6	34
15	j	228/231 (99%)	207 (91%)	14 (6%)	7 (3%)	4	26
15	k	228/231 (99%)	206 (90%)	16 (7%)	6 (3%)	5	30
15	l	228/231 (99%)	206 (90%)	16 (7%)	6 (3%)	5	30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
15	m	228/231 (99%)	199 (87%)	18 (8%)	11 (5%)	2	16
15	n	228/231 (99%)	202 (89%)	18 (8%)	8 (4%)	3	23
15	o	228/231 (99%)	209 (92%)	15 (7%)	4 (2%)	8	39
15	p	228/231 (99%)	207 (91%)	16 (7%)	5 (2%)	6	34
All	All	9506/9802 (97%)	7986 (84%)	1140 (12%)	380 (4%)	3	20

5 of 380 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	150	LEU
1	A	168	ALA
1	A	221	ASN
1	A	232	LYS
2	B	123	GLN

### 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	207/210 (99%)	185 (89%)	22 (11%)	6	27
1	O	207/210 (99%)	186 (90%)	21 (10%)	7	29
2	B	208/209 (100%)	188 (90%)	20 (10%)	8	31
2	P	208/209 (100%)	189 (91%)	19 (9%)	9	33
3	C	203/216 (94%)	181 (89%)	22 (11%)	6	26
3	Q	203/216 (94%)	181 (89%)	22 (11%)	6	26
4	D	213/226 (94%)	196 (92%)	17 (8%)	12	42
4	R	213/226 (94%)	196 (92%)	17 (8%)	12	42
5	E	201/215 (94%)	188 (94%)	13 (6%)	17	50
5	S	201/215 (94%)	187 (93%)	14 (7%)	15	47
6	F	193/193 (100%)	174 (90%)	19 (10%)	8	30
6	T	193/193 (100%)	173 (90%)	20 (10%)	7	28

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	G	201/239 (84%)	182 (90%)	19 (10%)	8	31
7	U	201/239 (84%)	183 (91%)	18 (9%)	9	34
8	H	162/162 (100%)	145 (90%)	17 (10%)	7	27
8	V	162/162 (100%)	144 (89%)	18 (11%)	6	25
9	I	181/181 (100%)	164 (91%)	17 (9%)	8	32
9	W	181/181 (100%)	164 (91%)	17 (9%)	8	32
10	J	172/173 (99%)	163 (95%)	9 (5%)	23	58
10	X	172/173 (99%)	163 (95%)	9 (5%)	23	58
11	K	175/175 (100%)	157 (90%)	18 (10%)	7	28
11	Y	175/175 (100%)	158 (90%)	17 (10%)	8	31
12	L	169/169 (100%)	159 (94%)	10 (6%)	19	54
12	Z	169/169 (100%)	158 (94%)	11 (6%)	17	50
13	M	185/185 (100%)	164 (89%)	21 (11%)	5	24
13	a	185/185 (100%)	163 (88%)	22 (12%)	5	22
14	N	199/199 (100%)	185 (93%)	14 (7%)	15	47
14	b	199/199 (100%)	184 (92%)	15 (8%)	13	44
15	c	188/189 (100%)	182 (97%)	6 (3%)	39	70
15	d	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	e	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	f	188/189 (100%)	182 (97%)	6 (3%)	39	70
15	g	188/189 (100%)	184 (98%)	4 (2%)	53	79
15	h	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	i	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	j	188/189 (100%)	182 (97%)	6 (3%)	39	70
15	k	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	l	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	m	188/189 (100%)	182 (97%)	6 (3%)	39	70
15	n	188/189 (100%)	184 (98%)	4 (2%)	53	79
15	o	188/189 (100%)	186 (99%)	2 (1%)	73	88
15	p	188/189 (100%)	186 (99%)	2 (1%)	73	88
All	All	7970/8150 (98%)	7444 (93%)	526 (7%)	16	50

5 of 526 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
14	N	226	LYS
3	Q	190	ILE
15	e	1174	SER
1	O	126	GLN
2	P	134	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 259 such sidechains are listed below:

Mol	Chain	Res	Type
3	Q	70	ASN
6	T	152	ASN
15	l	1079	HIS
3	Q	97	ASN
4	R	209	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](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
15	p	2
15	i	2
15	e	1
15	h	1
15	o	1
15	l	1
15	f	1
15	m	1

The worst 5 of 10 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	i	1172:GLY	C	1173:GLY	N	1.91
1	p	1172:GLY	C	1173:GLY	N	1.91
1	i	1173:GLY	C	1174:SER	N	1.16
1	p	1173:GLY	C	1174:SER	N	1.16
1	f	1175:GLN	C	1176:SER	N	1.15

## 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	243/252 (96%)	0.27	10 (4%) 37 25	47, 82, 112, 124	0
1	O	243/252 (96%)	0.24	11 (4%) 33 21	47, 82, 112, 125	0
2	B	249/250 (99%)	0.14	8 (3%) 47 33	45, 67, 102, 123	0
2	P	249/250 (99%)	0.01	9 (3%) 42 29	46, 68, 102, 123	0
3	C	243/258 (94%)	-0.01	5 (2%) 63 50	32, 60, 108, 122	0
3	Q	243/258 (94%)	0.07	10 (4%) 37 25	32, 62, 108, 122	0
4	D	241/254 (94%)	-0.01	5 (2%) 63 50	34, 58, 105, 135	0
4	R	241/254 (94%)	-0.05	7 (2%) 51 37	35, 61, 104, 133	0
5	E	245/260 (94%)	0.10	15 (6%) 21 12	31, 55, 108, 133	0
5	S	245/260 (94%)	0.08	13 (5%) 26 15	33, 57, 108, 134	0
6	F	234/234 (100%)	-0.16	1 (0%) 92 89	34, 58, 78, 111	0
6	T	234/234 (100%)	-0.07	2 (0%) 84 76	36, 59, 79, 111	0
7	G	243/288 (84%)	0.14	4 (1%) 72 60	48, 69, 112, 121	0
7	U	243/288 (84%)	0.24	13 (5%) 26 15	50, 70, 112, 121	0
8	H	196/196 (100%)	0.01	5 (2%) 56 42	46, 66, 91, 98	0
8	V	196/196 (100%)	0.06	6 (3%) 49 34	46, 66, 91, 98	0
9	I	222/222 (100%)	-0.09	4 (1%) 68 56	40, 60, 84, 128	0
9	W	222/222 (100%)	-0.02	5 (2%) 60 48	39, 59, 84, 127	0
10	J	204/205 (99%)	-0.40	2 (0%) 82 73	28, 43, 63, 76	0
10	X	204/205 (99%)	-0.39	1 (0%) 91 86	27, 43, 62, 76	0
11	K	198/198 (100%)	-0.24	5 (2%) 57 44	23, 40, 61, 126	0
11	Y	198/198 (100%)	-0.35	3 (1%) 73 62	24, 41, 62, 126	0
12	L	212/212 (100%)	-0.40	0 100 100	17, 37, 55, 72	0
12	Z	212/212 (100%)	-0.50	0 100 100	18, 38, 54, 70	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	222/222 (100%)	-0.31	1 (0%) 91 86	24, 47, 65, 80	0
13	a	222/222 (100%)	-0.28	1 (0%) 91 86	24, 48, 66, 80	0
14	N	233/233 (100%)	-0.12	2 (0%) 84 76	34, 60, 87, 98	0
14	b	233/233 (100%)	0.02	7 (3%) 50 36	36, 61, 87, 98	0
15	c	195/231 (84%)	0.30	8 (4%) 37 25	46, 79, 110, 123	0
15	d	195/231 (84%)	0.15	6 (3%) 49 34	46, 75, 109, 125	0
15	e	195/231 (84%)	0.19	11 (5%) 24 14	40, 70, 104, 122	0
15	f	195/231 (84%)	-0.02	2 (1%) 82 73	38, 69, 103, 130	0
15	g	195/231 (84%)	0.12	7 (3%) 42 29	47, 75, 110, 159	0
15	h	195/231 (84%)	0.48	12 (6%) 20 12	55, 85, 119, 168	0
15	i	192/231 (83%)	0.55	16 (8%) 11 6	56, 86, 112, 159	1 (0%)
15	j	195/231 (84%)	0.49	12 (6%) 20 12	61, 96, 124, 134	0
15	k	195/231 (84%)	0.59	22 (11%) 5 3	66, 100, 130, 138	0
15	l	195/231 (84%)	0.46	15 (7%) 13 7	64, 98, 127, 136	0
15	m	195/231 (84%)	0.32	14 (7%) 15 9	56, 96, 129, 140	0
15	n	195/231 (84%)	0.37	19 (9%) 7 5	58, 92, 130, 169	0
15	o	195/231 (84%)	0.37	16 (8%) 11 7	64, 93, 122, 167	0
15	p	192/231 (83%)	0.41	8 (4%) 36 24	57, 93, 119, 165	1 (0%)
All	All	9094/9802 (92%)	0.06	323 (3%) 42 29	17, 66, 113, 169	2 (0%)

The worst 5 of 323 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
15	p	1231	SER	9.4
15	h	1227	ASP	8.5
3	Q	245	THR	8.5
15	h	1231	SER	7.0
3	Q	244	ILE	6.9

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

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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.