



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 11:38 am BST

PDB ID : 2FHG  
Title : Crystal Structure of Mycobacterial Tuberculosis Proteasome  
Authors : Hu, G.; Lin, G.; Wang, M.; Dick, L.; Xu, R.M.; Nathan, C.; Li, H.  
Deposited on : 2005-12-23  
Resolution : 3.23 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	:	2.11
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)	:	2.11

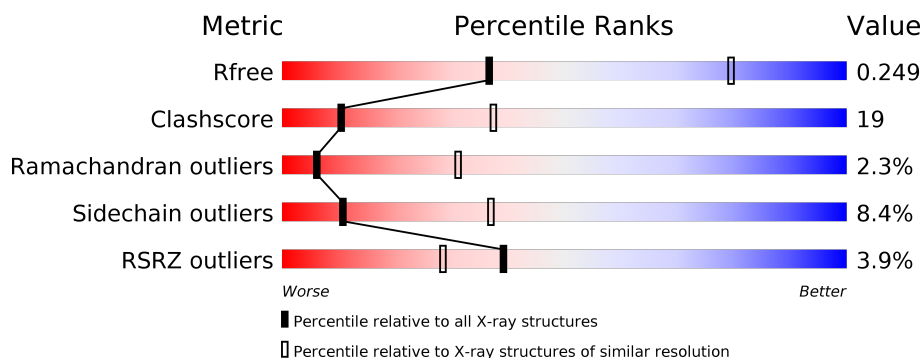
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.23 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	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)

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	1	250	<div> <div>8%</div> <div> <div></div> <div>53%</div> <div>31%</div> <div>•</div> <div>12%</div> </div> </div>
1	A	250	<div> <div>2%</div> <div> <div></div> <div>53%</div> <div>32%</div> <div>•</div> <div>12%</div> </div> </div>
1	B	250	<div> <div>2%</div> <div> <div></div> <div>54%</div> <div>30%</div> <div>•</div> <div>12%</div> </div> </div>
1	D	250	<div> <div>6%</div> <div> <div></div> <div>54%</div> <div>31%</div> <div>•</div> <div>12%</div> </div> </div>
1	F	250	<div> <div>4%</div> <div> <div></div> <div>54%</div> <div>30%</div> <div>•</div> <div>12%</div> </div> </div>
1	I	250	<div> <div>5%</div> <div> <div></div> <div>53%</div> <div>31%</div> <div>•</div> <div>12%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	250	
1	M	250	
1	O	250	
1	Q	250	
1	S	250	
1	U	250	
1	W	250	
1	Y	250	
2	2	240	
2	C	240	
2	E	240	
2	G	240	
2	H	240	
2	J	240	
2	L	240	
2	N	240	
2	P	240	
2	R	240	
2	T	240	
2	V	240	
2	X	240	
2	Z	240	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 46846 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 20S proteasome, alpha and beta subunits.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	B	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	D	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	F	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	I	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	K	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	M	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	O	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	Q	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	S	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	U	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	W	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	Y	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			
1	1	220	Total	C	N	O	S	0	0	0
			1692	1058	309	322	3			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	INITIATING METHIONINE	GB 13881852

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ASN	-	CLONING ARTIFACT	GB 13881852
A	1	SER	-	CLONING ARTIFACT	GB 13881852
B	-1	MET	-	INITIATING METHIONINE	GB 13881852
B	0	ASN	-	CLONING ARTIFACT	GB 13881852
B	1	SER	-	CLONING ARTIFACT	GB 13881852
D	-1	MET	-	INITIATING METHIONINE	GB 13881852
D	0	ASN	-	CLONING ARTIFACT	GB 13881852
D	1	SER	-	CLONING ARTIFACT	GB 13881852
F	-1	MET	-	INITIATING METHIONINE	GB 13881852
F	0	ASN	-	CLONING ARTIFACT	GB 13881852
F	1	SER	-	CLONING ARTIFACT	GB 13881852
I	-1	MET	-	INITIATING METHIONINE	GB 13881852
I	0	ASN	-	CLONING ARTIFACT	GB 13881852
I	1	SER	-	CLONING ARTIFACT	GB 13881852
K	-1	MET	-	INITIATING METHIONINE	GB 13881852
K	0	ASN	-	CLONING ARTIFACT	GB 13881852
K	1	SER	-	CLONING ARTIFACT	GB 13881852
M	-1	MET	-	INITIATING METHIONINE	GB 13881852
M	0	ASN	-	CLONING ARTIFACT	GB 13881852
M	1	SER	-	CLONING ARTIFACT	GB 13881852
O	-1	MET	-	INITIATING METHIONINE	GB 13881852
O	0	ASN	-	CLONING ARTIFACT	GB 13881852
O	1	SER	-	CLONING ARTIFACT	GB 13881852
Q	-1	MET	-	INITIATING METHIONINE	GB 13881852
Q	0	ASN	-	CLONING ARTIFACT	GB 13881852
Q	1	SER	-	CLONING ARTIFACT	GB 13881852
S	-1	MET	-	INITIATING METHIONINE	GB 13881852
S	0	ASN	-	CLONING ARTIFACT	GB 13881852
S	1	SER	-	CLONING ARTIFACT	GB 13881852
U	-1	MET	-	INITIATING METHIONINE	GB 13881852
U	0	ASN	-	CLONING ARTIFACT	GB 13881852
U	1	SER	-	CLONING ARTIFACT	GB 13881852
W	-1	MET	-	INITIATING METHIONINE	GB 13881852
W	0	ASN	-	CLONING ARTIFACT	GB 13881852
W	1	SER	-	CLONING ARTIFACT	GB 13881852
Y	-1	MET	-	INITIATING METHIONINE	GB 13881852
Y	0	ASN	-	CLONING ARTIFACT	GB 13881852
Y	1	SER	-	CLONING ARTIFACT	GB 13881852
1	-1	MET	-	INITIATING METHIONINE	GB 13881852
1	0	ASN	-	CLONING ARTIFACT	GB 13881852
1	1	SER	-	CLONING ARTIFACT	GB 13881852

- Molecule 2 is a protein called proteasome, beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	C	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	E	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	G	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	J	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	L	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	N	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	P	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	R	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	T	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	V	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	X	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	Z	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0
2	2	222	Total 1638	C 1027	N 282	O 324	S 5	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	535	HIS	-	EXPRESSION TAG	GB 13881852
H	536	HIS	-	EXPRESSION TAG	GB 13881852
H	537	HIS	-	EXPRESSION TAG	GB 13881852
H	538	HIS	-	EXPRESSION TAG	GB 13881852
H	539	HIS	-	EXPRESSION TAG	GB 13881852
H	540	HIS	-	EXPRESSION TAG	GB 13881852
C	535	HIS	-	EXPRESSION TAG	GB 13881852
C	536	HIS	-	EXPRESSION TAG	GB 13881852
C	537	HIS	-	EXPRESSION TAG	GB 13881852
C	538	HIS	-	EXPRESSION TAG	GB 13881852
C	539	HIS	-	EXPRESSION TAG	GB 13881852
C	540	HIS	-	EXPRESSION TAG	GB 13881852

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	535	HIS	-	EXPRESSION TAG	GB 13881852
E	536	HIS	-	EXPRESSION TAG	GB 13881852
E	537	HIS	-	EXPRESSION TAG	GB 13881852
E	538	HIS	-	EXPRESSION TAG	GB 13881852
E	539	HIS	-	EXPRESSION TAG	GB 13881852
E	540	HIS	-	EXPRESSION TAG	GB 13881852
G	535	HIS	-	EXPRESSION TAG	GB 13881852
G	536	HIS	-	EXPRESSION TAG	GB 13881852
G	537	HIS	-	EXPRESSION TAG	GB 13881852
G	538	HIS	-	EXPRESSION TAG	GB 13881852
G	539	HIS	-	EXPRESSION TAG	GB 13881852
G	540	HIS	-	EXPRESSION TAG	GB 13881852
J	535	HIS	-	EXPRESSION TAG	GB 13881852
J	536	HIS	-	EXPRESSION TAG	GB 13881852
J	537	HIS	-	EXPRESSION TAG	GB 13881852
J	538	HIS	-	EXPRESSION TAG	GB 13881852
J	539	HIS	-	EXPRESSION TAG	GB 13881852
J	540	HIS	-	EXPRESSION TAG	GB 13881852
L	535	HIS	-	EXPRESSION TAG	GB 13881852
L	536	HIS	-	EXPRESSION TAG	GB 13881852
L	537	HIS	-	EXPRESSION TAG	GB 13881852
L	538	HIS	-	EXPRESSION TAG	GB 13881852
L	539	HIS	-	EXPRESSION TAG	GB 13881852
L	540	HIS	-	EXPRESSION TAG	GB 13881852
N	535	HIS	-	EXPRESSION TAG	GB 13881852
N	536	HIS	-	EXPRESSION TAG	GB 13881852
N	537	HIS	-	EXPRESSION TAG	GB 13881852
N	538	HIS	-	EXPRESSION TAG	GB 13881852
N	539	HIS	-	EXPRESSION TAG	GB 13881852
N	540	HIS	-	EXPRESSION TAG	GB 13881852
P	535	HIS	-	EXPRESSION TAG	GB 13881852
P	536	HIS	-	EXPRESSION TAG	GB 13881852
P	537	HIS	-	EXPRESSION TAG	GB 13881852
P	538	HIS	-	EXPRESSION TAG	GB 13881852
P	539	HIS	-	EXPRESSION TAG	GB 13881852
P	540	HIS	-	EXPRESSION TAG	GB 13881852
R	535	HIS	-	EXPRESSION TAG	GB 13881852
R	536	HIS	-	EXPRESSION TAG	GB 13881852
R	537	HIS	-	EXPRESSION TAG	GB 13881852
R	538	HIS	-	EXPRESSION TAG	GB 13881852
R	539	HIS	-	EXPRESSION TAG	GB 13881852
R	540	HIS	-	EXPRESSION TAG	GB 13881852

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
T	535	HIS	-	EXPRESSION TAG	GB 13881852
T	536	HIS	-	EXPRESSION TAG	GB 13881852
T	537	HIS	-	EXPRESSION TAG	GB 13881852
T	538	HIS	-	EXPRESSION TAG	GB 13881852
T	539	HIS	-	EXPRESSION TAG	GB 13881852
T	540	HIS	-	EXPRESSION TAG	GB 13881852
V	535	HIS	-	EXPRESSION TAG	GB 13881852
V	536	HIS	-	EXPRESSION TAG	GB 13881852
V	537	HIS	-	EXPRESSION TAG	GB 13881852
V	538	HIS	-	EXPRESSION TAG	GB 13881852
V	539	HIS	-	EXPRESSION TAG	GB 13881852
V	540	HIS	-	EXPRESSION TAG	GB 13881852
X	535	HIS	-	EXPRESSION TAG	GB 13881852
X	536	HIS	-	EXPRESSION TAG	GB 13881852
X	537	HIS	-	EXPRESSION TAG	GB 13881852
X	538	HIS	-	EXPRESSION TAG	GB 13881852
X	539	HIS	-	EXPRESSION TAG	GB 13881852
X	540	HIS	-	EXPRESSION TAG	GB 13881852
Z	535	HIS	-	EXPRESSION TAG	GB 13881852
Z	536	HIS	-	EXPRESSION TAG	GB 13881852
Z	537	HIS	-	EXPRESSION TAG	GB 13881852
Z	538	HIS	-	EXPRESSION TAG	GB 13881852
Z	539	HIS	-	EXPRESSION TAG	GB 13881852
Z	540	HIS	-	EXPRESSION TAG	GB 13881852
2	535	HIS	-	EXPRESSION TAG	GB 13881852
2	536	HIS	-	EXPRESSION TAG	GB 13881852
2	537	HIS	-	EXPRESSION TAG	GB 13881852
2	538	HIS	-	EXPRESSION TAG	GB 13881852
2	539	HIS	-	EXPRESSION TAG	GB 13881852
2	540	HIS	-	EXPRESSION TAG	GB 13881852

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total O 2 2	0	0
3	H	1	Total O 1 1	0	0
3	B	7	Total O 7 7	0	0
3	C	6	Total O 6 6	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	11	Total 11	O 11	0	0
3	E	9	Total 9	O 9	0	0
3	F	10	Total 10	O 10	0	0
3	G	9	Total 9	O 9	0	0
3	I	7	Total 7	O 7	0	0
3	J	5	Total 5	O 5	0	0
3	K	15	Total 15	O 15	0	0
3	L	4	Total 4	O 4	0	0
3	M	11	Total 11	O 11	0	0
3	N	7	Total 7	O 7	0	0
3	O	8	Total 8	O 8	0	0
3	P	4	Total 4	O 4	0	0
3	Q	6	Total 6	O 6	0	0
3	R	5	Total 5	O 5	0	0
3	S	13	Total 13	O 13	0	0
3	T	6	Total 6	O 6	0	0
3	U	7	Total 7	O 7	0	0
3	V	10	Total 10	O 10	0	0
3	W	8	Total 8	O 8	0	0
3	X	4	Total 4	O 4	0	0
3	Y	12	Total 12	O 12	0	0

*Continued on next page...*

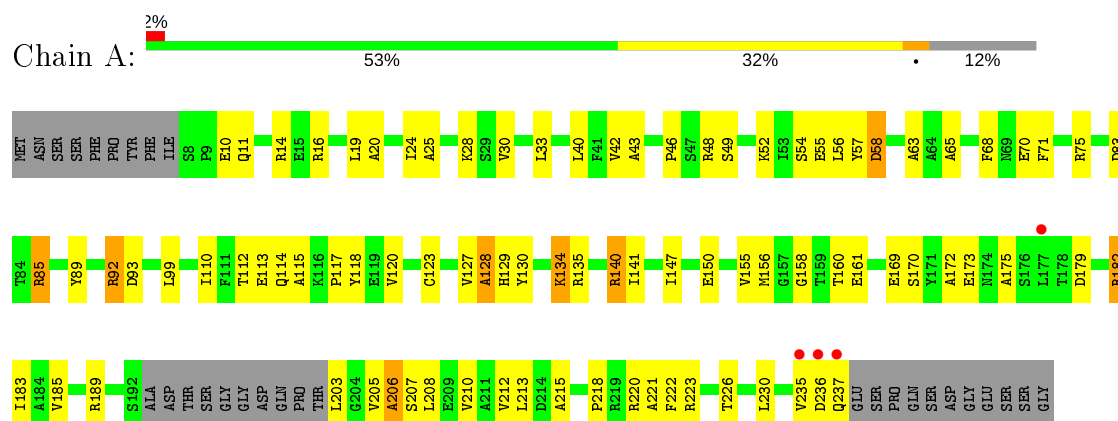
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	Z	7	Total 7	O 7	0	0
3	1	13	Total 13	O 13	0	0
3	2	19	Total 19	O 19	0	0

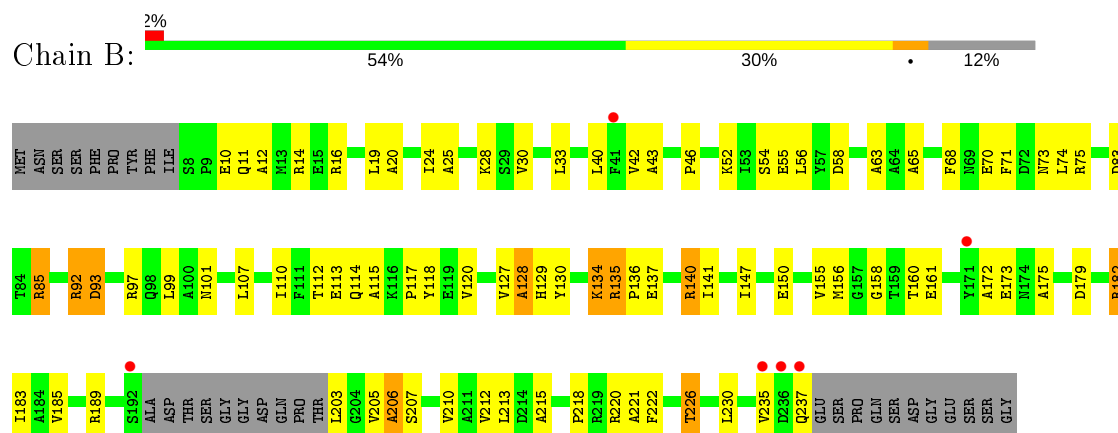
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

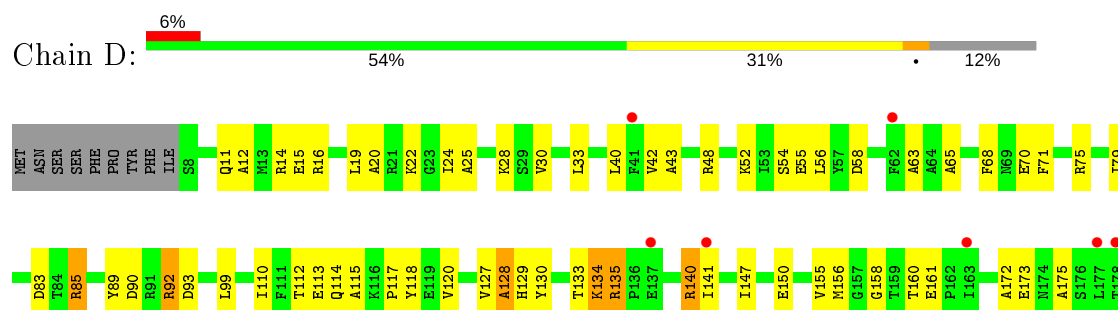
- Molecule 1: 20S proteasome, alpha and beta subunits

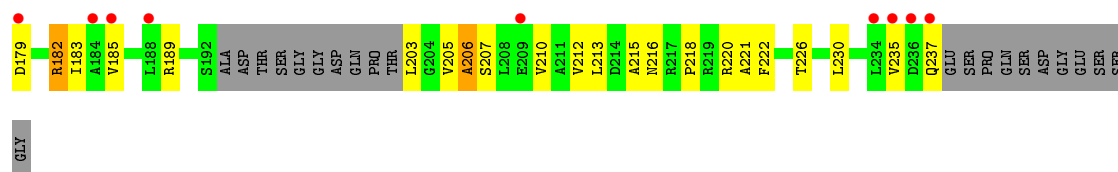


- Molecule 1: 20S proteasome, alpha and beta subunits

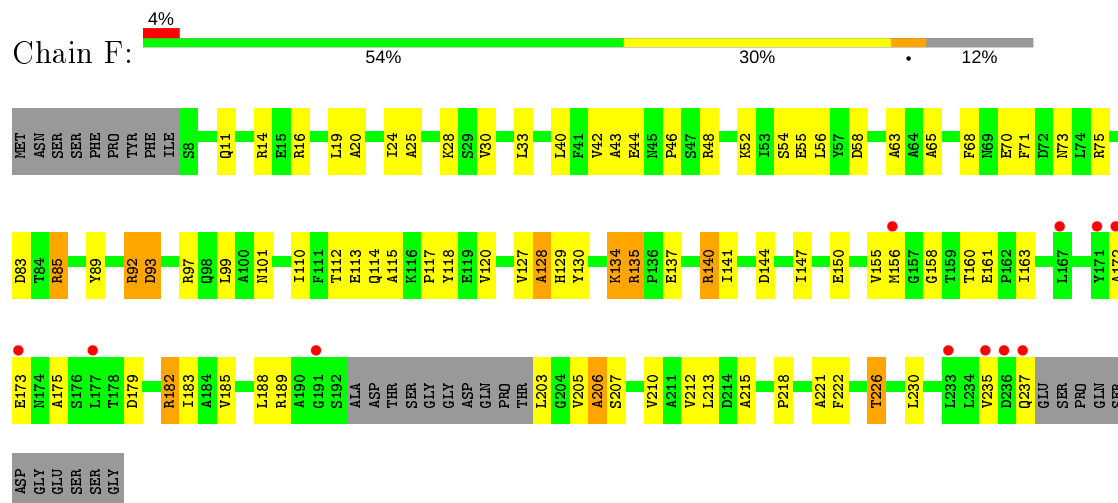


- Molecule 1: 20S proteasome, alpha and beta subunits

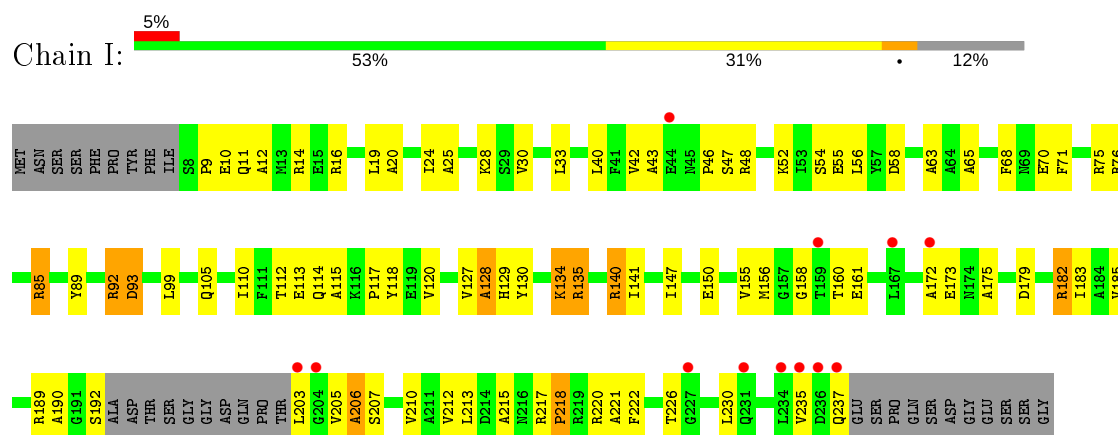




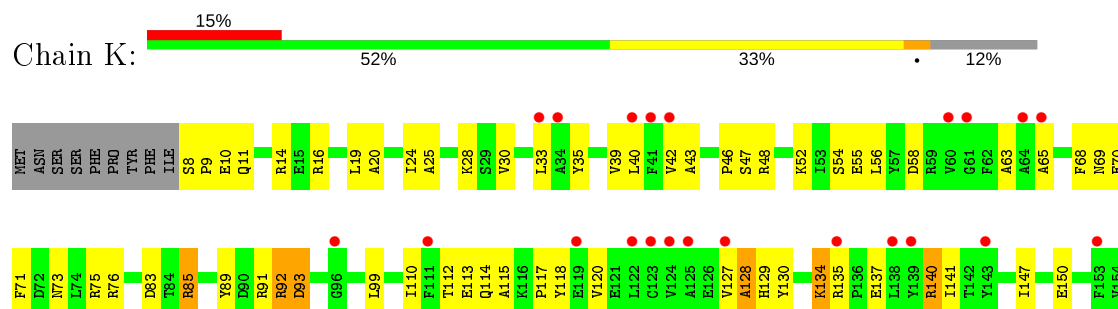
- Molecule 1: 20S proteasome, alpha and beta subunits

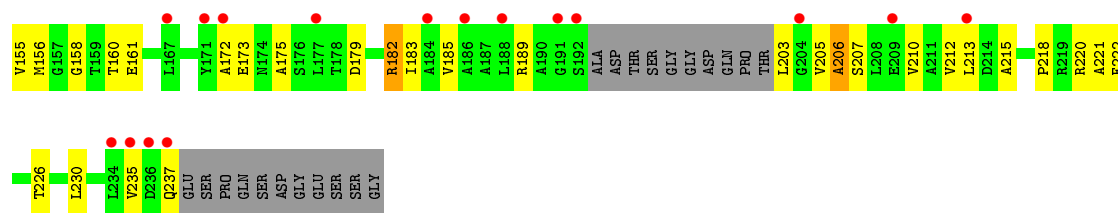


- Molecule 1: 20S proteasome, alpha and beta subunits

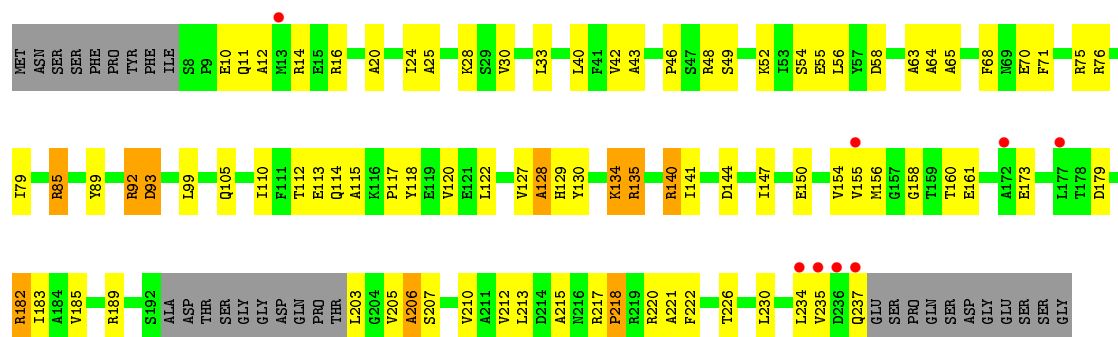


- Molecule 1: 20S proteasome, alpha and beta subunits

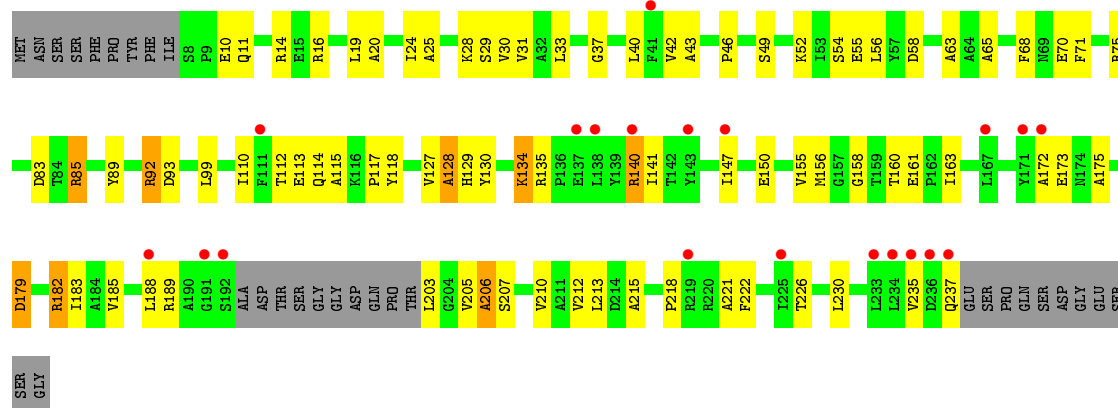




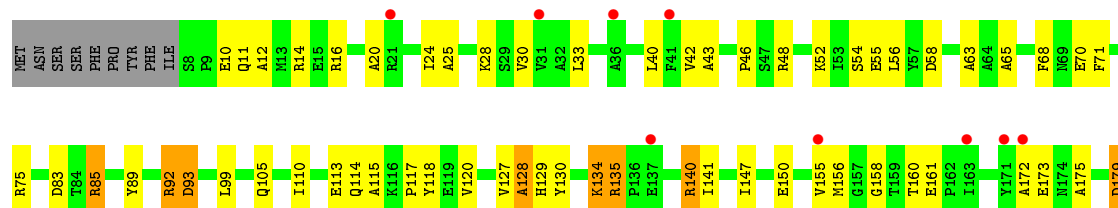
- Molecule 1: 20S proteasome, alpha and beta subunits

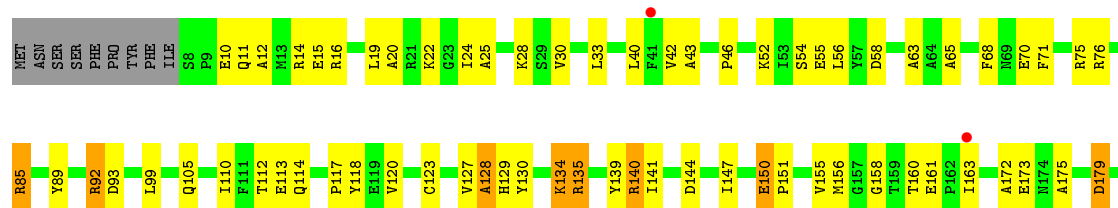
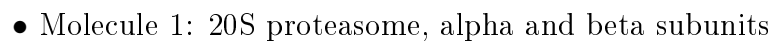
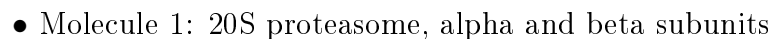


- Molecule 1: 20S proteasome, alpha and beta subunits



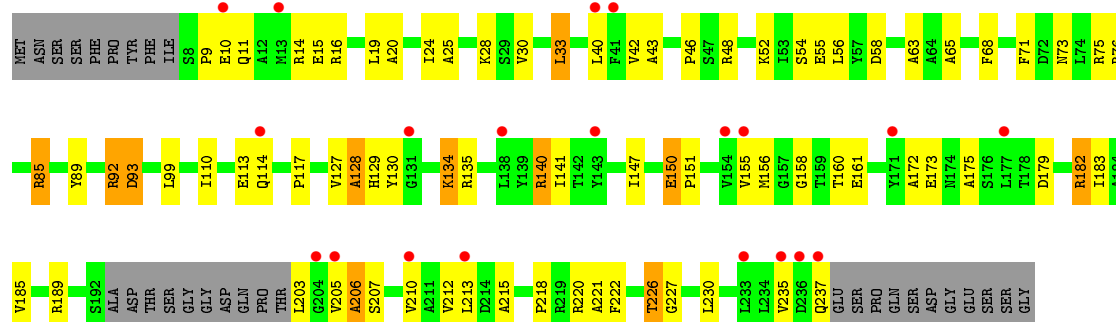
- Molecule 1: 20S proteasome, alpha and beta subunits



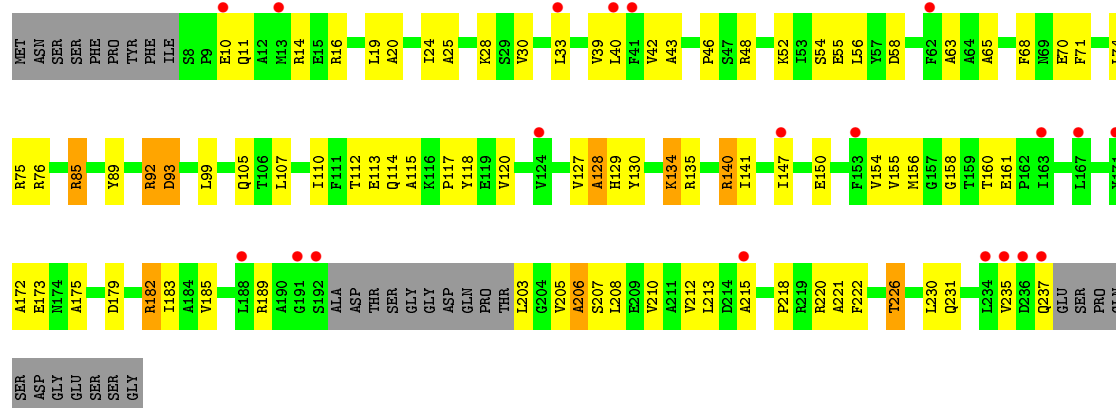




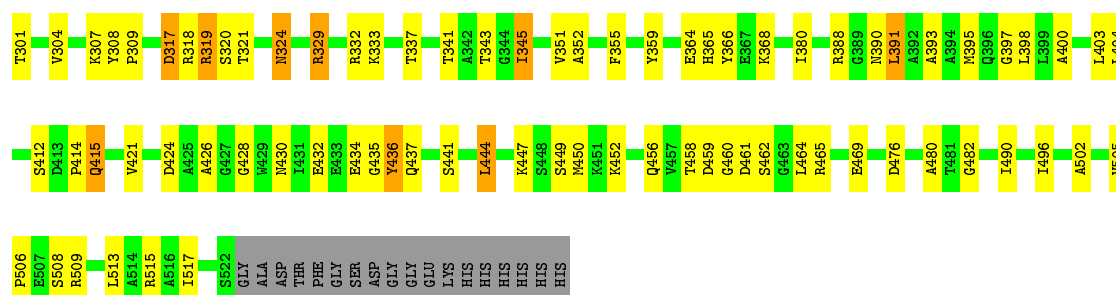
- Molecule 1: 20S proteasome, alpha and beta subunits



- Molecule 1: 20S proteasome, alpha and beta subunits

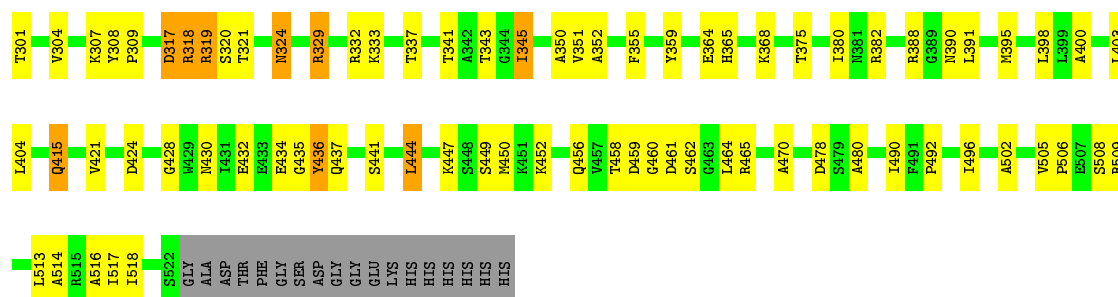


- Molecule 2: proteasome, beta subunit



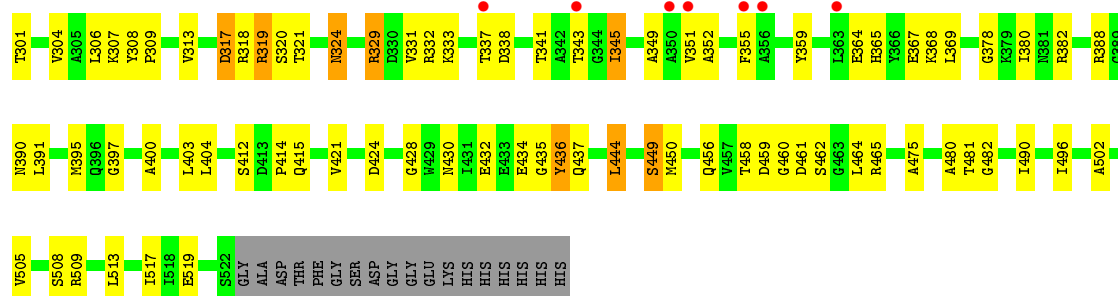
- Molecule 2: proteasome, beta subunit

Chain C: 



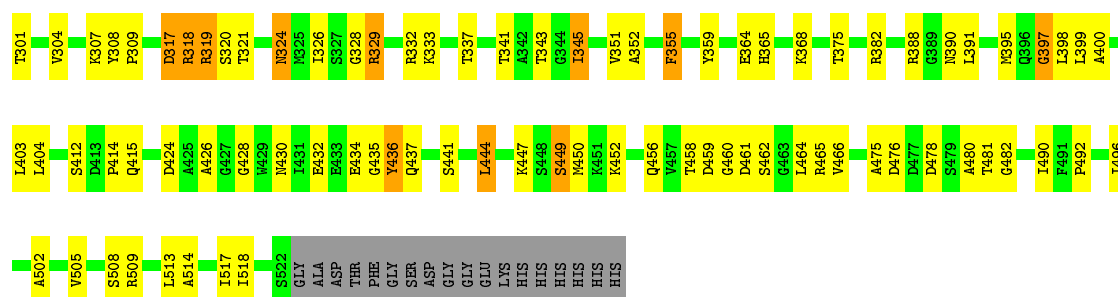
• Molecule 2: proteasome, beta subunit

Chain E: 



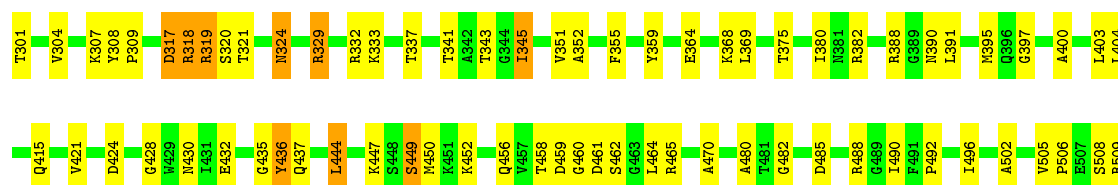
• Molecule 2: proteasome, beta subunit

Chain G: 

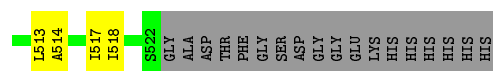


• Molecule 2: proteasome, beta subunit

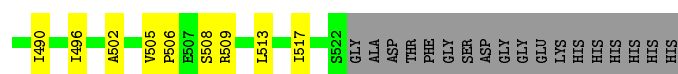
Chain J: 







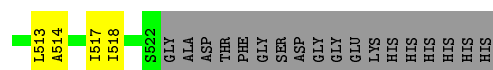
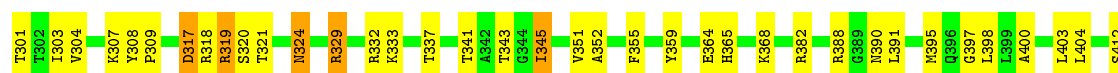
- Molecule 2: proteasome, beta subunit



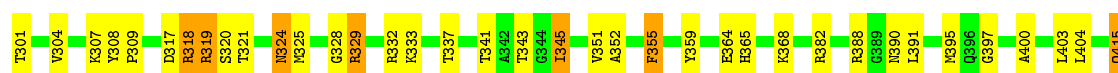
- Molecule 2: proteasome, beta subunit



- Molecule 2: proteasome, beta subunit



- Molecule 2: proteasome, beta subunit



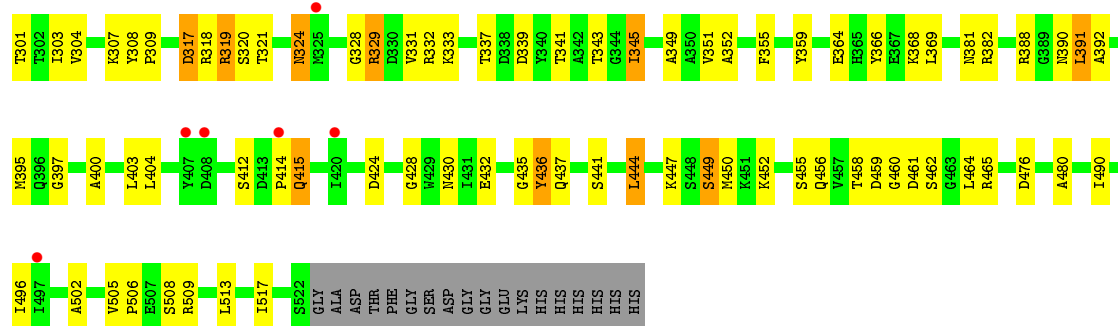
- Molecule 2: proteasome, beta subunit

- Molecule 2: proteasome, beta subunit

- Molecule 2: proteasome, beta subunit

- Molecule 2: proteasome, beta subunit

Chain Z:  59% 30% • 8%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	173.33Å 113.80Å 200.13Å 90.00° 113.93° 90.00°	Depositor
Resolution (Å)	50.00 – 3.23 49.63 – 3.23	Depositor EDS
% Data completeness (in resolution range)	79.6 (50.00-3.23) 79.6 (49.63-3.23)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 3.25Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.243 , 0.266 0.247 , 0.249	Depositor DCC
$R_{free}$ test set	2776 reflections (3.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	82.0	Xtriage
Anisotropy	0.332	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 102.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.046 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	46846	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	99.0	wwPDB-VP

Xtriage'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 [i](#)

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

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	1	0.43	0/1717	0.59	0/2320
1	A	0.54	1/1717 (0.1%)	0.63	0/2320
1	B	0.49	0/1717	0.62	0/2320
1	D	0.47	0/1717	0.61	0/2320
1	F	0.45	0/1717	0.60	0/2320
1	I	0.44	0/1717	0.60	0/2320
1	K	0.43	0/1717	0.59	0/2320
1	M	0.45	0/1717	0.60	0/2320
1	O	0.45	0/1717	0.60	0/2320
1	Q	0.44	0/1717	0.59	0/2320
1	S	0.43	0/1717	0.59	0/2320
1	U	0.43	0/1717	0.60	0/2320
1	W	0.47	0/1717	0.61	0/2320
1	Y	0.43	0/1717	0.59	0/2320
2	2	0.42	0/1662	0.61	0/2254
2	C	0.52	0/1662	0.65	0/2254
2	E	0.54	0/1662	0.66	0/2254
2	G	0.48	0/1662	0.65	0/2254
2	H	0.55	0/1662	0.66	0/2254
2	J	0.49	0/1662	0.63	0/2254
2	L	0.51	0/1662	0.65	0/2254
2	N	0.49	0/1662	0.64	0/2254
2	P	0.49	0/1662	0.64	0/2254
2	R	0.50	0/1662	0.65	0/2254
2	T	0.46	0/1662	0.63	0/2254
2	V	0.47	0/1662	0.65	0/2254
2	X	0.49	0/1662	0.64	0/2254
2	Z	0.48	0/1662	0.64	0/2254
All	All	0.47	1/47306 (0.0%)	0.62	0/64036

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	123	CYS	CB-SG	-5.45	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	1	1692	0	1688	75	0
1	A	1692	0	1688	77	4
1	B	1692	0	1688	75	0
1	D	1692	0	1688	72	4
1	F	1692	0	1688	80	0
1	I	1692	0	1688	72	2
1	K	1692	0	1688	84	0
1	M	1692	0	1688	88	0
1	O	1692	0	1688	77	0
1	Q	1692	0	1688	68	0
1	S	1692	0	1688	77	0
1	U	1692	0	1688	69	2
1	W	1692	0	1688	77	1
1	Y	1692	0	1688	72	0
2	2	1638	0	1630	67	1
2	C	1638	0	1630	58	0
2	E	1638	0	1630	60	1
2	G	1638	0	1630	67	3
2	H	1638	0	1630	62	0
2	J	1638	0	1630	56	2
2	L	1638	0	1630	68	0
2	N	1638	0	1630	62	1
2	P	1638	0	1630	61	1
2	R	1638	0	1630	54	0
2	T	1638	0	1630	60	0
2	V	1638	0	1630	64	2
2	X	1638	0	1630	59	3
2	Z	1638	0	1630	62	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	1	13	0	0	10	0
3	2	19	0	0	8	0
3	A	2	0	0	2	0
3	B	7	0	0	2	0
3	C	6	0	0	0	0
3	D	11	0	0	3	0
3	E	9	0	0	3	0
3	F	10	0	0	1	0
3	G	9	0	0	3	0
3	H	1	0	0	3	0
3	I	7	0	0	3	0
3	J	5	0	0	0	0
3	K	15	0	0	8	0
3	L	4	0	0	3	0
3	M	11	0	0	5	0
3	N	7	0	0	5	0
3	O	8	0	0	8	0
3	P	4	0	0	0	0
3	Q	6	0	0	2	0
3	R	5	0	0	2	0
3	S	13	0	0	3	0
3	T	6	0	0	4	0
3	U	7	0	0	1	0
3	V	10	0	0	4	0
3	W	8	0	0	3	0
3	X	4	0	0	1	0
3	Y	12	0	0	8	0
3	Z	7	0	0	4	0
All	All	46846	0	46452	1730	14

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 19.

The worst 5 of 1730 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:J:444:LEU:HD12	2:Z:444:LEU:HD12	1.27	1.15
2:L:444:LEU:HD12	2:P:444:LEU:HD12	1.30	1.13
2:J:452:LYS:HG3	3:Z:190:HOH:O	1.47	1.12
1:K:14:ARG:HG2	3:K:260:HOH:O	1.56	1.06
2:N:444:LEU:HD12	2:V:444:LEU:HD12	1.35	1.04

The worst 5 of 14 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)
2:G:465:ARG:NE	2:X:458:THR:CG2[2_554]	1.40	0.80
1:A:236:ASP:O	1:D:133:THR:OG1[2_655]	1.59	0.61
2:N:456:GLN:NE2	2:Z:415:GLN:OE1[1_565]	1.83	0.37
1:W:228:SER:O	2:2:414:PRO:CB[2_544]	1.97	0.23
2:J:456:GLN:NE2	2:V:415:GLN:OE1[1_545]	2.00	0.20

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	1	216/250 (86%)	188 (87%)	22 (10%)	6 (3%)	5	27
1	A	216/250 (86%)	192 (89%)	19 (9%)	5 (2%)	6	31
1	B	216/250 (86%)	189 (88%)	21 (10%)	6 (3%)	5	27
1	D	216/250 (86%)	193 (89%)	18 (8%)	5 (2%)	6	31
1	F	216/250 (86%)	190 (88%)	20 (9%)	6 (3%)	5	27
1	I	216/250 (86%)	189 (88%)	22 (10%)	5 (2%)	6	31
1	K	216/250 (86%)	191 (88%)	20 (9%)	5 (2%)	6	31
1	M	216/250 (86%)	187 (87%)	24 (11%)	5 (2%)	6	31
1	O	216/250 (86%)	188 (87%)	23 (11%)	5 (2%)	6	31
1	Q	216/250 (86%)	188 (87%)	23 (11%)	5 (2%)	6	31
1	S	216/250 (86%)	189 (88%)	22 (10%)	5 (2%)	6	31
1	U	216/250 (86%)	189 (88%)	22 (10%)	5 (2%)	6	31
1	W	216/250 (86%)	192 (89%)	19 (9%)	5 (2%)	6	31
1	Y	216/250 (86%)	187 (87%)	23 (11%)	6 (3%)	5	27
2	2	220/240 (92%)	195 (89%)	22 (10%)	3 (1%)	11	43
2	C	220/240 (92%)	195 (89%)	22 (10%)	3 (1%)	11	43

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	220/240 (92%)	196 (89%)	19 (9%)	5 (2%)	6	31
2	G	220/240 (92%)	196 (89%)	18 (8%)	6 (3%)	5	27
2	H	220/240 (92%)	193 (88%)	22 (10%)	5 (2%)	6	31
2	J	220/240 (92%)	194 (88%)	22 (10%)	4 (2%)	8	38
2	L	220/240 (92%)	194 (88%)	21 (10%)	5 (2%)	6	31
2	N	220/240 (92%)	198 (90%)	16 (7%)	6 (3%)	5	27
2	P	220/240 (92%)	195 (89%)	19 (9%)	6 (3%)	5	27
2	R	220/240 (92%)	193 (88%)	22 (10%)	5 (2%)	6	31
2	T	220/240 (92%)	195 (89%)	19 (9%)	6 (3%)	5	27
2	V	220/240 (92%)	194 (88%)	21 (10%)	5 (2%)	6	31
2	X	220/240 (92%)	193 (88%)	23 (10%)	4 (2%)	8	38
2	Z	220/240 (92%)	192 (87%)	24 (11%)	4 (2%)	8	38
All	All	6104/6860 (89%)	5375 (88%)	588 (10%)	141 (2%)	6	31

5 of 141 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	ALA
1	B	128	ALA
1	D	128	ALA
1	F	128	ALA
1	I	128	ALA

### 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	1	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	A	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	B	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	D	169/194 (87%)	156 (92%)	13 (8%)	13	41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	169/194 (87%)	155 (92%)	14 (8%)	11	38
1	I	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	K	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	M	169/194 (87%)	155 (92%)	14 (8%)	11	38
1	O	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	Q	169/194 (87%)	155 (92%)	14 (8%)	11	38
1	S	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	U	169/194 (87%)	156 (92%)	13 (8%)	13	41
1	W	169/194 (87%)	155 (92%)	14 (8%)	11	38
1	Y	169/194 (87%)	156 (92%)	13 (8%)	13	41
2	2	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	C	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	E	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	G	165/178 (93%)	151 (92%)	14 (8%)	10	36
2	H	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	J	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	L	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	N	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	P	165/178 (93%)	151 (92%)	14 (8%)	10	36
2	R	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	T	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	V	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	X	165/178 (93%)	150 (91%)	15 (9%)	9	32
2	Z	165/178 (93%)	149 (90%)	16 (10%)	8	30
All	All	4676/5208 (90%)	4281 (92%)	395 (8%)	11	37

5 of 395 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	150	GLU
2	P	444	LEU
2	Z	449	SER
2	N	319	ARG
1	O	113	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 154 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	98	GLN
2	P	415	GLN
1	1	51	GLN
1	M	165	ASN
1	O	73	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 carbohydrates 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 [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	1	220/250 (88%)	0.63	20 (9%) 9 7	100, 100, 100, 100	0
1	A	220/250 (88%)	0.35	4 (1%) 68 58	100, 100, 100, 100	0
1	B	220/250 (88%)	0.35	6 (2%) 54 42	100, 100, 100, 100	0
1	D	220/250 (88%)	0.60	16 (7%) 15 10	100, 100, 100, 100	0
1	F	220/250 (88%)	0.42	11 (5%) 28 18	100, 100, 100, 100	0
1	I	220/250 (88%)	0.55	12 (5%) 25 16	100, 100, 100, 100	0
1	K	220/250 (88%)	0.98	38 (17%) 1 1	100, 100, 100, 100	0
1	M	220/250 (88%)	0.43	8 (3%) 42 31	100, 100, 100, 100	0
1	O	220/250 (88%)	0.68	20 (9%) 9 7	100, 100, 100, 100	0
1	Q	220/250 (88%)	0.77	25 (11%) 5 4	100, 100, 100, 100	0
1	S	220/250 (88%)	0.64	19 (8%) 10 7	100, 100, 100, 100	0
1	U	220/250 (88%)	0.66	17 (7%) 13 8	100, 100, 100, 100	0
1	W	220/250 (88%)	0.42	9 (4%) 37 27	100, 100, 100, 100	0
1	Y	220/250 (88%)	0.65	20 (9%) 9 7	100, 100, 100, 100	0
2	2	222/240 (92%)	0.40	6 (2%) 54 42	100, 100, 100, 100	0
2	C	222/240 (92%)	0.21	0 100 100	100, 100, 100, 100	0
2	E	222/240 (92%)	0.22	7 (3%) 47 35	100, 100, 100, 100	0
2	G	222/240 (92%)	0.22	0 100 100	100, 100, 100, 100	0
2	H	222/240 (92%)	0.20	0 100 100	100, 100, 100, 100	0
2	J	222/240 (92%)	0.21	0 100 100	100, 100, 100, 100	0
2	L	222/240 (92%)	0.20	2 (0%) 84 78	100, 100, 100, 100	0
2	N	222/240 (92%)	0.14	0 100 100	100, 100, 100, 100	0
2	P	222/240 (92%)	0.22	0 100 100	100, 100, 100, 100	0
2	R	222/240 (92%)	0.27	0 100 100	100, 100, 100, 100	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å <sup>2</sup> )	Q<0.9
2	T	222/240 (92%)	0.22	0	100	100	100, 100, 100, 100	0
2	V	222/240 (92%)	0.24	2 (0%)	84	78	100, 100, 100, 100	0
2	X	222/240 (92%)	0.13	0	100	100	100, 100, 100, 100	0
2	Z	222/240 (92%)	0.24	0	100	100	100, 100, 100, 100	0
All	All	6188/6860 (90%)	0.40	242 (3%)	39	28	100, 100, 100, 100	0

The worst 5 of 242 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	236	ASP	17.0
1	K	235	VAL	15.5
1	S	236	ASP	14.4
1	K	236	ASP	12.1
1	Q	236	ASP	11.5

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

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.