wwPDB X-ray Structure Validation Summary Report (i)

Jan 11, 2017 – 07:41 PM EST

PDB ID	:	5TRS
Title	:	Structure of Mycobacterium tuberculosis proteasome in complex with N,C-capped dipeptide PKS2144
Authors	:	Hsu, H.-C.; Fan, H.; Singh, P.K.; Wang, R.; Sukenick, G.; Nathan, C.; Lin, G.; Li, H.
Deposited on	:	2016-10-27
Resolution	:	3.08 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*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

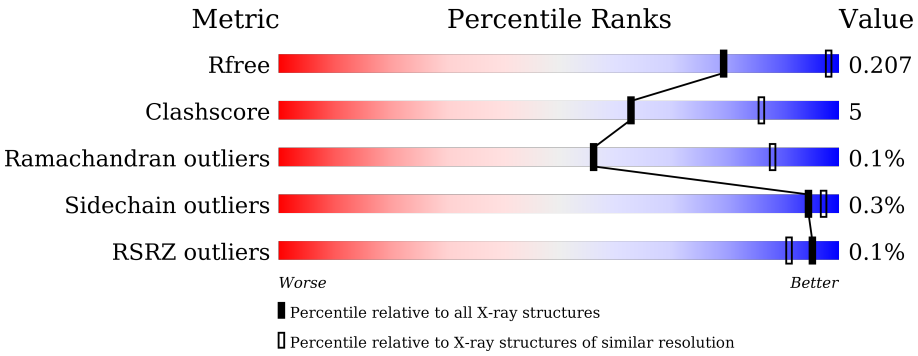
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.08 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1119 (3.12-3.04)
Clashscore	102246	1098 (3.10-3.06)
Ramachandran outliers	100387	1057 (3.10-3.06)
Sidechain outliers	100360	1057 (3.10-3.06)
RSRZ outliers	91569	1001 (3.10-3.06)

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 >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	240	79% 12% 9%
1	B	240	71% 18% 10%
1	C	240	72% 19% 9%
1	D	240	73% 17% 9%
1	E	240	76% 15% 10%
1	F	240	70% 18% • 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	240	74% 16% 10%
1	O	240	68% 23% 10%
1	P	240	73% 18% 9%
1	Q	240	72% 18% 10%
1	R	240	70% 20% 10%
1	S	240	79% 12% 9%
1	T	240	79% 12% 10%
1	U	240	74% 16% 10%
2	H	240	86% 7% 8%
2	I	240	86% 6% 8%
2	J	240	87% 6% 8%
2	K	240	83% 10% 7%
2	L	240	86% 7% 7%
2	M	240	83% 10% 8%
2	N	240	81% 12% 7%
2	V	240	83% 10% 7%
2	W	240	83% 10% 7%
2	X	240	82% 10% 8%
2	Y	240	80% 13% 7%
2	Z	240	84% 9% 8%
2	a	240	93% 7%
2	b	240	93% 7%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 47019 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Proteasome subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	219	Total	C	N	O	S	0	0	0
			1684	1055	307	318	4			
1	B	216	Total	C	N	O	S	0	0	0
			1668	1045	304	315	4			
1	C	218	Total	C	N	O	S	0	0	0
			1678	1050	306	318	4			
1	D	218	Total	C	N	O	S	0	0	0
			1677	1050	306	317	4			
1	E	217	Total	C	N	O	S	0	0	0
			1671	1047	305	315	4			
1	F	216	Total	C	N	O	S	0	0	0
			1663	1041	304	314	4			
1	G	216	Total	C	N	O	S	0	0	0
			1662	1040	304	314	4			
1	O	217	Total	C	N	O	S	0	0	0
			1671	1047	305	315	4			
1	P	219	Total	C	N	O	S	0	0	0
			1685	1054	307	320	4			
1	Q	216	Total	C	N	O	S	0	0	0
			1668	1045	304	315	4			
1	R	216	Total	C	N	O	S	0	0	0
			1663	1041	304	314	4			
1	S	218	Total	C	N	O	S	0	0	0
			1678	1050	306	318	4			
1	T	217	Total	C	N	O	S	0	0	0
			1671	1047	305	315	4			
1	U	216	Total	C	N	O	S	0	0	0
			1664	1043	304	313	4			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	MET	-	initiating methionine	UNP A5U4D5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	MET	-	initiating methionine	UNP A5U4D5
C	9	MET	-	initiating methionine	UNP A5U4D5
D	9	MET	-	initiating methionine	UNP A5U4D5
E	9	MET	-	initiating methionine	UNP A5U4D5
F	9	MET	-	initiating methionine	UNP A5U4D5
G	9	MET	-	initiating methionine	UNP A5U4D5
O	9	MET	-	initiating methionine	UNP A5U4D5
P	9	MET	-	initiating methionine	UNP A5U4D5
Q	9	MET	-	initiating methionine	UNP A5U4D5
R	9	MET	-	initiating methionine	UNP A5U4D5
S	9	MET	-	initiating methionine	UNP A5U4D5
T	9	MET	-	initiating methionine	UNP A5U4D5
U	9	MET	-	initiating methionine	UNP A5U4D5

- Molecule 2 is a protein called Proteasome subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	222	Total	C	N	O	S	0	0	0
			1638	1027	282	324	5			
2	I	222	Total	C	N	O	S	0	0	0
			1638	1027	282	324	5			
2	J	222	Total	C	N	O	S	0	0	0
			1638	1027	282	324	5			
2	K	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			
2	L	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			
2	M	222	Total	C	N	O	S	0	0	0
			1638	1027	282	324	5			
2	N	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			
2	V	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			
2	W	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			
2	X	222	Total	C	N	O	S	0	0	0
			1638	1027	282	324	5			
2	Y	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			
2	Z	222	Total	C	N	O	S	0	0	0
			1638	1027	282	324	5			
2	a	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	b	223	Total	C	N	O	S	0	0	0
			1642	1029	283	325	5			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	235	HIS	-	expression tag	UNP A5U4D6
H	236	HIS	-	expression tag	UNP A5U4D6
H	237	HIS	-	expression tag	UNP A5U4D6
H	238	HIS	-	expression tag	UNP A5U4D6
H	239	HIS	-	expression tag	UNP A5U4D6
H	240	HIS	-	expression tag	UNP A5U4D6
I	235	HIS	-	expression tag	UNP A5U4D6
I	236	HIS	-	expression tag	UNP A5U4D6
I	237	HIS	-	expression tag	UNP A5U4D6
I	238	HIS	-	expression tag	UNP A5U4D6
I	239	HIS	-	expression tag	UNP A5U4D6
I	240	HIS	-	expression tag	UNP A5U4D6
J	235	HIS	-	expression tag	UNP A5U4D6
J	236	HIS	-	expression tag	UNP A5U4D6
J	237	HIS	-	expression tag	UNP A5U4D6
J	238	HIS	-	expression tag	UNP A5U4D6
J	239	HIS	-	expression tag	UNP A5U4D6
J	240	HIS	-	expression tag	UNP A5U4D6
K	235	HIS	-	expression tag	UNP A5U4D6
K	236	HIS	-	expression tag	UNP A5U4D6
K	237	HIS	-	expression tag	UNP A5U4D6
K	238	HIS	-	expression tag	UNP A5U4D6
K	239	HIS	-	expression tag	UNP A5U4D6
K	240	HIS	-	expression tag	UNP A5U4D6
L	235	HIS	-	expression tag	UNP A5U4D6
L	236	HIS	-	expression tag	UNP A5U4D6
L	237	HIS	-	expression tag	UNP A5U4D6
L	238	HIS	-	expression tag	UNP A5U4D6
L	239	HIS	-	expression tag	UNP A5U4D6
L	240	HIS	-	expression tag	UNP A5U4D6
M	235	HIS	-	expression tag	UNP A5U4D6
M	236	HIS	-	expression tag	UNP A5U4D6
M	237	HIS	-	expression tag	UNP A5U4D6
M	238	HIS	-	expression tag	UNP A5U4D6
M	239	HIS	-	expression tag	UNP A5U4D6
M	240	HIS	-	expression tag	UNP A5U4D6

Continued on next page...

Continued from previous page...

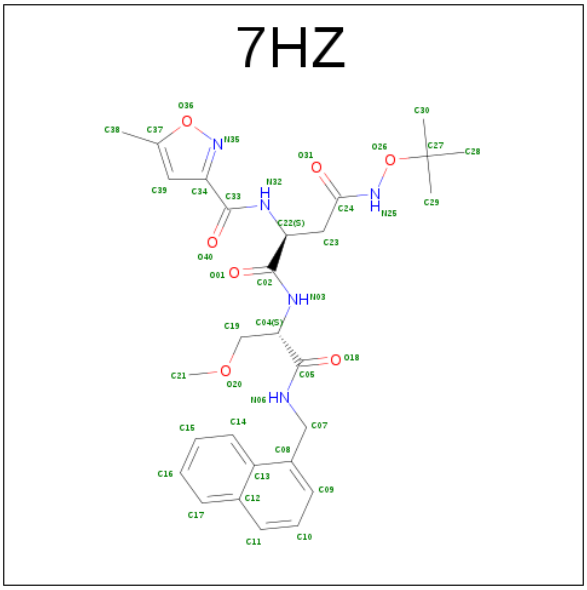
Chain	Residue	Modelled	Actual	Comment	Reference
N	235	HIS	-	expression tag	UNP A5U4D6
N	236	HIS	-	expression tag	UNP A5U4D6
N	237	HIS	-	expression tag	UNP A5U4D6
N	238	HIS	-	expression tag	UNP A5U4D6
N	239	HIS	-	expression tag	UNP A5U4D6
N	240	HIS	-	expression tag	UNP A5U4D6
V	235	HIS	-	expression tag	UNP A5U4D6
V	236	HIS	-	expression tag	UNP A5U4D6
V	237	HIS	-	expression tag	UNP A5U4D6
V	238	HIS	-	expression tag	UNP A5U4D6
V	239	HIS	-	expression tag	UNP A5U4D6
V	240	HIS	-	expression tag	UNP A5U4D6
W	235	HIS	-	expression tag	UNP A5U4D6
W	236	HIS	-	expression tag	UNP A5U4D6
W	237	HIS	-	expression tag	UNP A5U4D6
W	238	HIS	-	expression tag	UNP A5U4D6
W	239	HIS	-	expression tag	UNP A5U4D6
W	240	HIS	-	expression tag	UNP A5U4D6
X	235	HIS	-	expression tag	UNP A5U4D6
X	236	HIS	-	expression tag	UNP A5U4D6
X	237	HIS	-	expression tag	UNP A5U4D6
X	238	HIS	-	expression tag	UNP A5U4D6
X	239	HIS	-	expression tag	UNP A5U4D6
X	240	HIS	-	expression tag	UNP A5U4D6
Y	235	HIS	-	expression tag	UNP A5U4D6
Y	236	HIS	-	expression tag	UNP A5U4D6
Y	237	HIS	-	expression tag	UNP A5U4D6
Y	238	HIS	-	expression tag	UNP A5U4D6
Y	239	HIS	-	expression tag	UNP A5U4D6
Y	240	HIS	-	expression tag	UNP A5U4D6
Z	235	HIS	-	expression tag	UNP A5U4D6
Z	236	HIS	-	expression tag	UNP A5U4D6
Z	237	HIS	-	expression tag	UNP A5U4D6
Z	238	HIS	-	expression tag	UNP A5U4D6
Z	239	HIS	-	expression tag	UNP A5U4D6
Z	240	HIS	-	expression tag	UNP A5U4D6
a	235	HIS	-	expression tag	UNP A5U4D6
a	236	HIS	-	expression tag	UNP A5U4D6
a	237	HIS	-	expression tag	UNP A5U4D6
a	238	HIS	-	expression tag	UNP A5U4D6
a	239	HIS	-	expression tag	UNP A5U4D6
a	240	HIS	-	expression tag	UNP A5U4D6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
b	235	HIS	-	expression tag	UNP A5U4D6
b	236	HIS	-	expression tag	UNP A5U4D6
b	237	HIS	-	expression tag	UNP A5U4D6
b	238	HIS	-	expression tag	UNP A5U4D6
b	239	HIS	-	expression tag	UNP A5U4D6
b	240	HIS	-	expression tag	UNP A5U4D6

- Molecule 3 is N-tert-butoxy-N 2 -(5-methyl-1,2-oxazole-3-carbonyl)-L-asparaginyl-O-methyl -N-[(naphthalen-1-yl)methyl]-L-serinamide (three-letter code: 7HZ) (formula: $C_{28}H_{35}N_5O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	H	1	Total	C	N	O	0	0
			40	28	5	7		
3	I	1	Total	C	N	O	0	0
			40	28	5	7		
3	J	1	Total	C	N	O	0	0
			40	28	5	7		
3	K	1	Total	C	N	O	0	0
			40	28	5	7		
3	L	1	Total	C	N	O	0	0
			40	28	5	7		
3	M	1	Total	C	N	O	0	0
			40	28	5	7		
3	N	1	Total	C	N	O	0	0
			40	28	5	7		
3	V	1	Total	C	N	O	0	0
			40	28	5	7		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	W	1	Total	C	N	O	0	0
			40	28	5	7		
3	X	1	Total	C	N	O	0	0
			40	28	5	7		
3	Y	1	Total	C	N	O	0	0
			40	28	5	7		
3	Z	1	Total	C	N	O	0	0
			40	28	5	7		
3	a	1	Total	C	N	O	0	0
			40	28	5	7		
3	b	1	Total	C	N	O	0	0
			40	28	5	7		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	O	0	0
			2	2		
4	B	1	Total	O	0	0
			1	1		
4	D	2	Total	O	0	0
			2	2		
4	F	2	Total	O	0	0
			2	2		
4	G	1	Total	O	0	0
			1	1		
4	H	7	Total	O	0	0
			7	7		
4	I	5	Total	O	0	0
			5	5		
4	J	5	Total	O	0	0
			5	5		
4	K	6	Total	O	0	0
			6	6		
4	L	5	Total	O	0	0
			5	5		
4	M	7	Total	O	0	0
			7	7		
4	N	5	Total	O	0	0
			5	5		
4	O	1	Total	O	0	0
			1	1		

Continued on next page...

Continued from previous page...

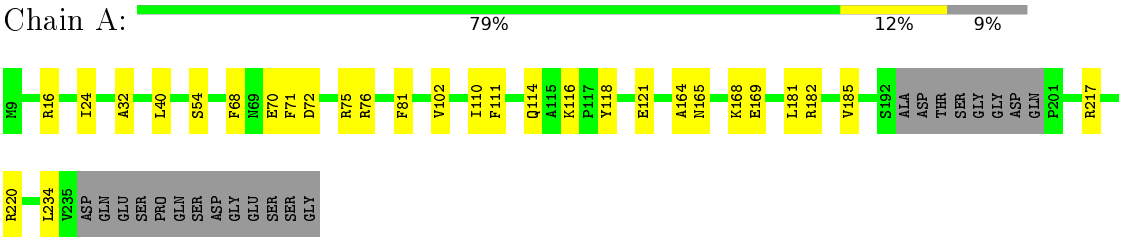
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	P	1	Total O 1 1	0	0
4	Q	1	Total O 1 1	0	0
4	R	1	Total O 1 1	0	0
4	S	2	Total O 2 2	0	0
4	T	1	Total O 1 1	0	0
4	U	1	Total O 1 1	0	0
4	V	4	Total O 4 4	0	0
4	W	3	Total O 3 3	0	0
4	X	5	Total O 5 5	0	0
4	Y	7	Total O 7 7	0	0
4	Z	5	Total O 5 5	0	0
4	a	6	Total O 6 6	0	0
4	b	6	Total O 6 6	0	0

3 Residue-property plots (i)

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

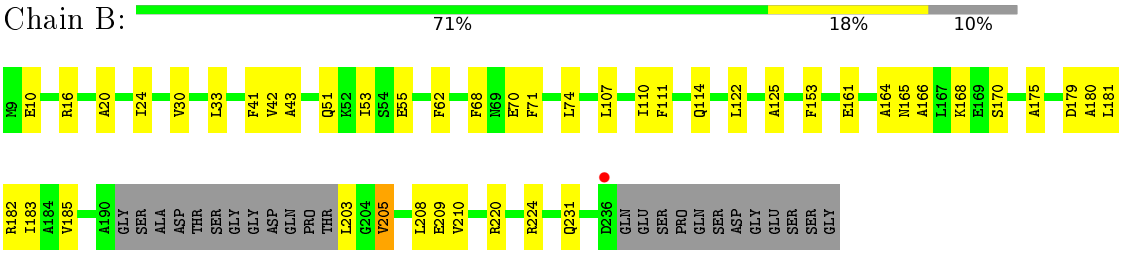
- Molecule 1: Proteasome subunit alpha

Chain A:



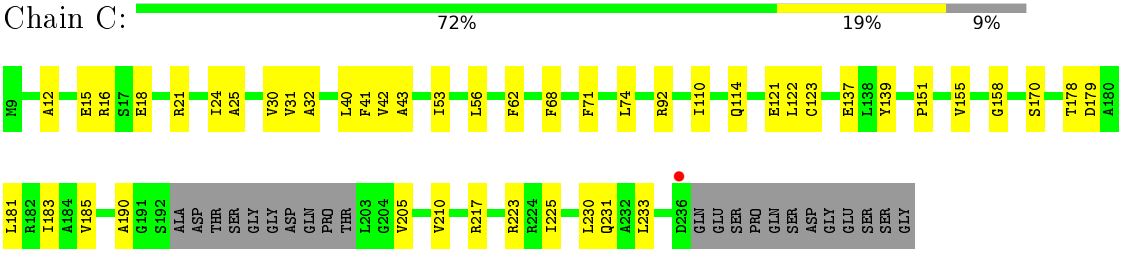
- Molecule 1: Proteasome subunit alpha

Chain B:



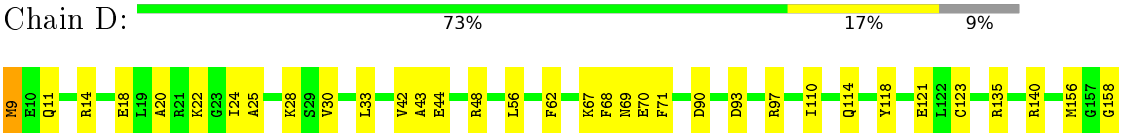
- Molecule 1: Proteasome subunit alpha

Chain C:



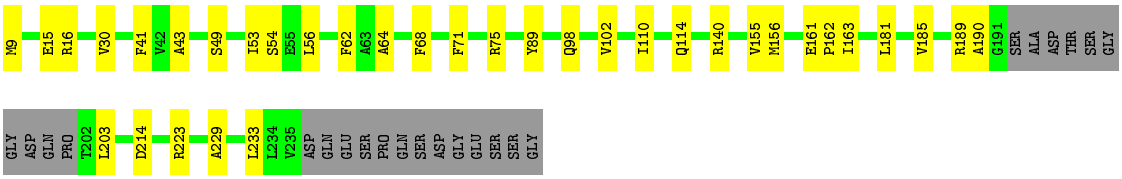
- Molecule 1: Proteasome subunit alpha

Chain D:



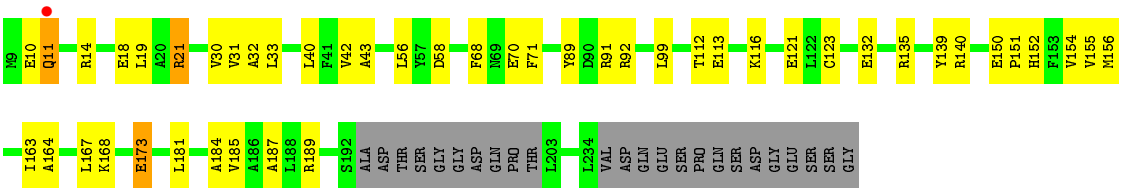
• Molecule 1: Proteasome subunit alpha

Chain E:



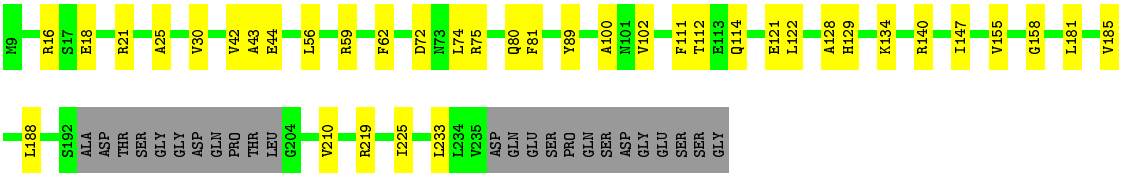
• Molecule 1: Proteasome subunit alpha

Chain F:



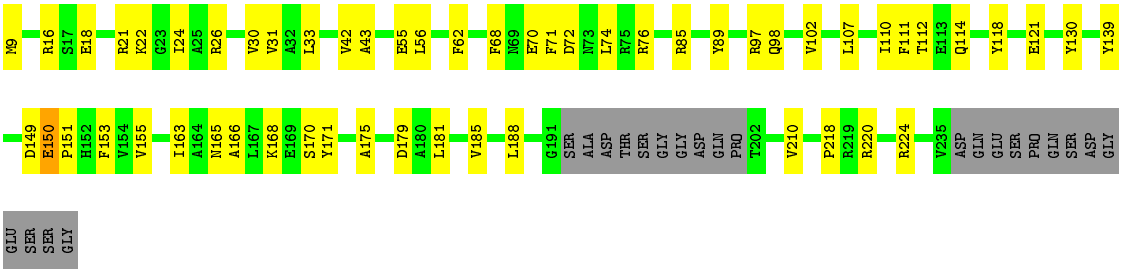
• Molecule 1: Proteasome subunit alpha

Chain G:



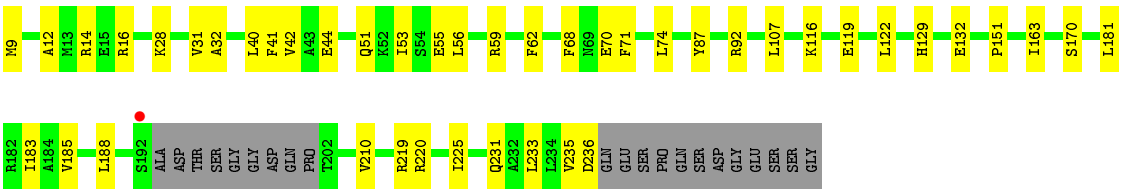
• Molecule 1: Proteasome subunit alpha

Chain O:



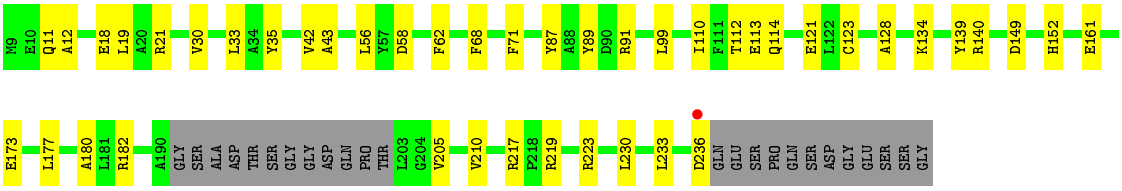
• Molecule 1: Proteasome subunit alpha

Chain P:



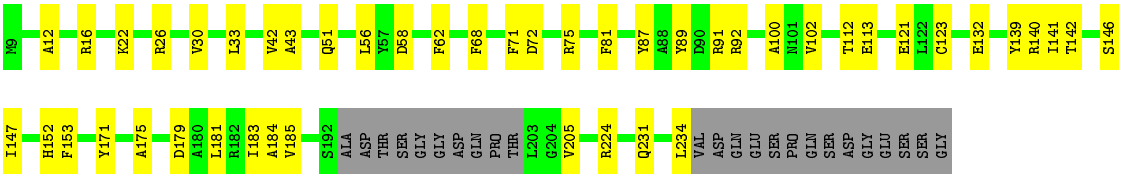
• Molecule 1: Proteasome subunit alpha

Chain Q:



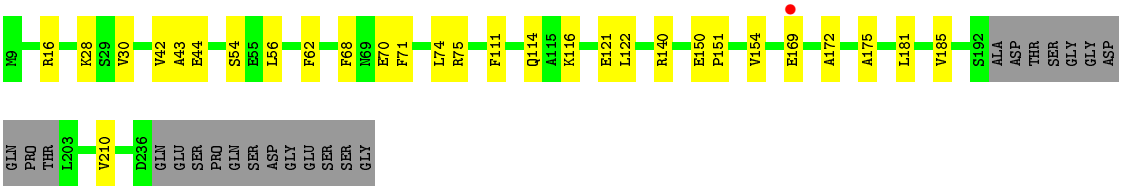
• Molecule 1: Proteasome subunit alpha

Chain R:



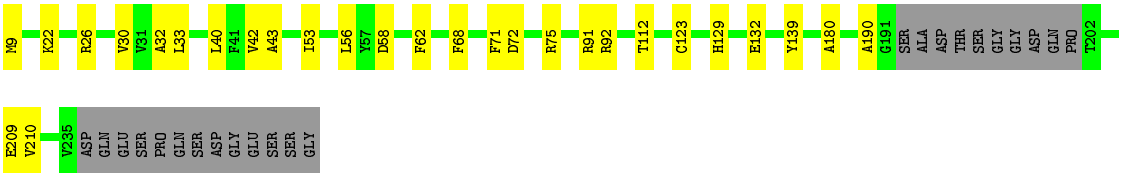
• Molecule 1: Proteasome subunit alpha

Chain S:



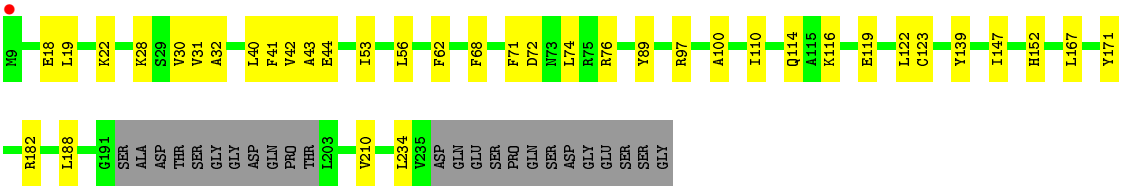
• Molecule 1: Proteasome subunit alpha

Chain T:



• Molecule 1: Proteasome subunit alpha

Chain U:

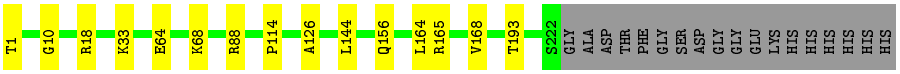


• Molecule 2: Proteasome subunit beta

Chain H:

• Molecule 2: Proteasome subunit beta

Chain I:

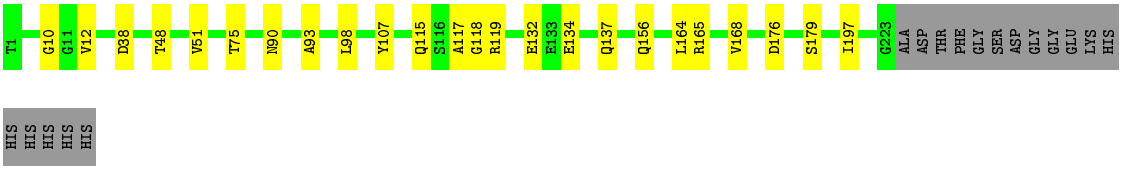


• Molecule 2: Proteasome subunit beta

Chain J:

• Molecule 2: Proteasome subunit beta

Chain K:

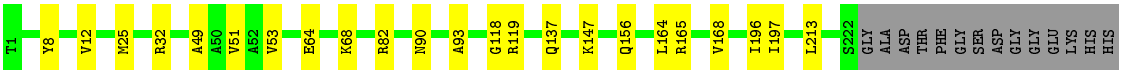


• Molecule 2: Proteasome subunit beta

Chain L:

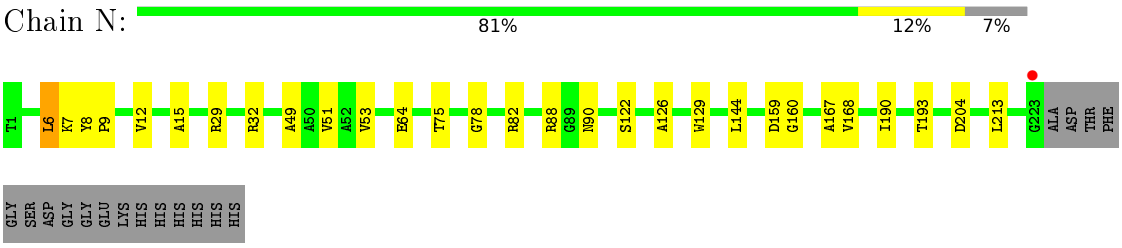
• Molecule 2: Proteasome subunit beta

Chain M:



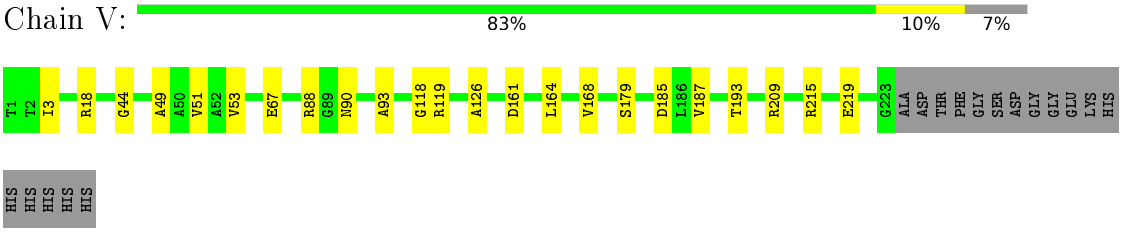
• Molecule 2: Proteasome subunit beta

Chain N:



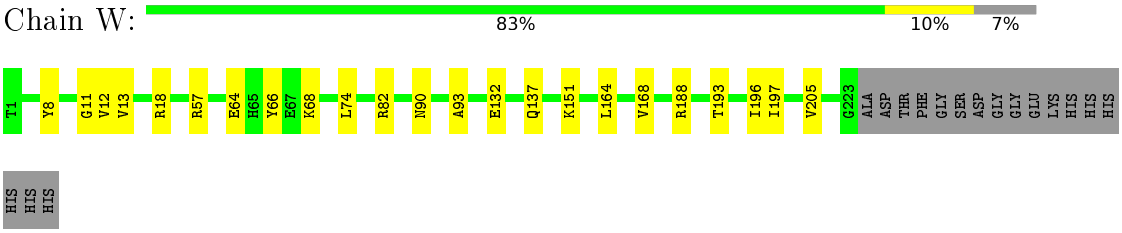
• Molecule 2: Proteasome subunit beta

Chain V:



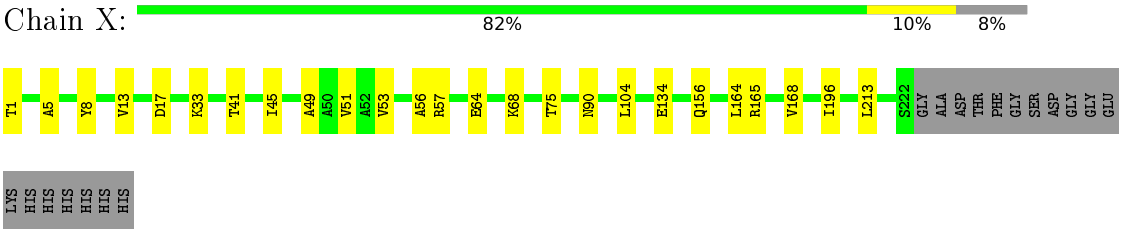
• Molecule 2: Proteasome subunit beta

Chain W:



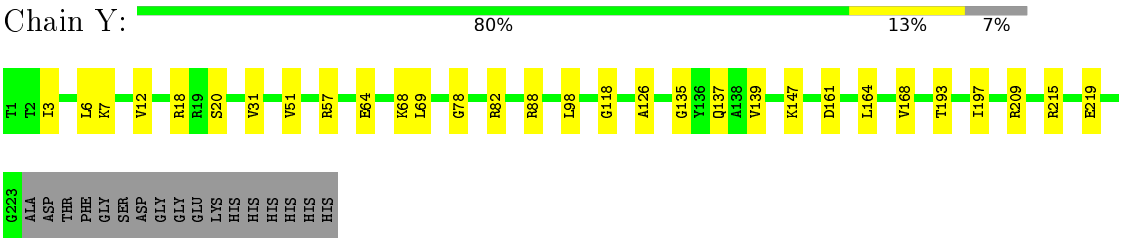
• Molecule 2: Proteasome subunit beta

Chain X:



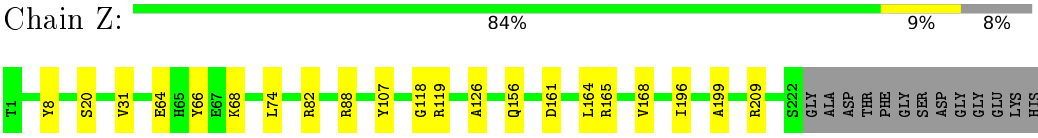
• Molecule 2: Proteasome subunit beta

Chain Y:



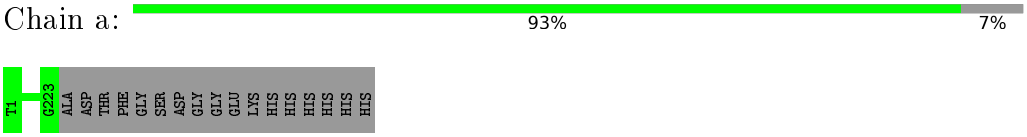
• Molecule 2: Proteasome subunit beta

Chain Z:



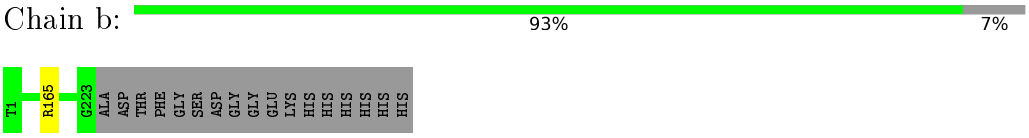
• Molecule 2: Proteasome subunit beta

Chain a:



• Molecule 2: Proteasome subunit beta

Chain b:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	120.51Å 198.30Å 165.94Å 90.00° 103.13° 90.00°	Depositor
Resolution (Å)	49.87 – 3.08 49.87 – 3.08	Depositor EDS
% Data completeness (in resolution range)	98.1 (49.87-3.08) 93.0 (49.87-3.08)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.84 (at 3.07Å)	Xtriage
Refinement program	Phenix	Depositor
R, R_{free}	0.181 , 0.211 0.180 , 0.207	Depositor DCC
R_{free} test set	6369 reflections (4.91%)	DCC
Wilson B-factor (Å^2)	66.7	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 21.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	47019	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.73% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 7HZ

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	0.37	0/1709	0.53	0/2308
1	B	0.34	0/1692	0.53	0/2285
1	C	0.34	0/1702	0.51	0/2298
1	D	0.34	0/1701	0.53	0/2297
1	E	0.34	0/1695	0.50	0/2289
1	F	0.35	0/1687	0.54	0/2277
1	G	0.36	0/1686	0.53	0/2276
1	O	0.35	0/1695	0.53	0/2289
1	P	0.34	0/1709	0.53	0/2308
1	Q	0.37	1/1692 (0.1%)	0.53	0/2285
1	R	0.37	0/1687	0.53	0/2277
1	S	0.37	0/1702	0.54	0/2298
1	T	0.36	0/1695	0.53	0/2289
1	U	0.36	0/1688	0.53	0/2279
2	H	0.37	0/1662	0.54	0/2254
2	I	0.37	0/1662	0.54	0/2254
2	J	0.35	0/1662	0.52	0/2254
2	K	0.38	0/1666	0.56	1/2259 (0.0%)
2	L	0.38	0/1666	0.55	0/2259
2	M	0.35	0/1662	0.54	0/2254
2	N	0.37	0/1666	0.55	0/2259
2	V	0.37	0/1666	0.55	0/2259
2	W	0.39	0/1666	0.55	0/2259
2	X	0.37	0/1662	0.55	0/2254
2	Y	0.37	0/1666	0.55	0/2259
2	Z	0.37	0/1662	0.55	0/2254
2	a	0.36	0/1666	0.53	0/2259
2	b	0.37	0/1666	0.54	0/2259
All	All	0.36	1/47040 (0.0%)	0.53	1/63651 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	Q	149	ASP	C-N	-5.34	1.21	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	K	38	ASP	CB-CG-OD1	5.14	122.93	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1684	0	1688	17	0
1	B	1668	0	1669	32	0
1	C	1678	0	1677	27	0
1	D	1677	0	1680	24	0
1	E	1671	0	1675	22	0
1	F	1663	0	1664	29	0
1	G	1662	0	1662	22	0
1	O	1671	0	1675	36	0
1	P	1685	0	1684	29	0
1	Q	1668	0	1668	27	0
1	R	1663	0	1664	30	0
1	S	1678	0	1677	14	0
1	T	1671	0	1675	15	0
1	U	1664	0	1668	23	0
2	H	1638	0	1633	11	0
2	I	1638	0	1633	8	0
2	J	1638	0	1633	9	0
2	K	1642	0	1636	14	0
2	L	1642	0	1636	9	0
2	M	1638	0	1633	14	0
2	N	1642	0	1636	20	0
2	V	1642	0	1636	13	0
2	W	1642	0	1636	17	0
2	X	1638	0	1633	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	Y	1642	0	1636	18	0
2	Z	1638	0	1633	12	0
2	a	1642	0	1636	0	0
2	b	1642	0	1636	0	0
3	H	40	0	0	0	0
3	I	40	0	0	0	0
3	J	40	0	0	0	0
3	K	40	0	0	0	0
3	L	40	0	0	0	0
3	M	40	0	0	0	0
3	N	40	0	0	0	0
3	V	40	0	0	0	0
3	W	40	0	0	0	0
3	X	40	0	0	0	0
3	Y	40	0	0	0	0
3	Z	40	0	0	0	0
3	a	40	0	0	0	0
3	b	40	0	0	0	0
4	A	2	0	0	2	0
4	B	1	0	0	0	0
4	D	2	0	0	1	0
4	F	2	0	0	0	0
4	G	1	0	0	0	0
4	H	7	0	0	0	0
4	I	5	0	0	0	0
4	J	5	0	0	0	0
4	K	6	0	0	0	0
4	L	5	0	0	0	0
4	M	7	0	0	0	0
4	N	5	0	0	0	0
4	O	1	0	0	0	0
4	P	1	0	0	0	0
4	Q	1	0	0	0	0
4	R	1	0	0	0	0
4	S	2	0	0	0	0
4	T	1	0	0	0	0
4	U	1	0	0	1	0
4	V	4	0	0	0	0
4	W	3	0	0	0	0
4	X	5	0	0	0	0
4	Y	7	0	0	0	0
4	Z	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	a	6	0	0	0	0
4	b	6	0	0	0	0
All	All	47019	0	46312	468	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 468 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)
1:D:11:GLN:HG3	1:D:14:ARG:HH12	1.37	0.90
1:E:140:ARG:NH1	1:E:155:VAL:O	2.10	0.83
2:N:159:ASP:OD1	2:N:160:GLY:N	2.15	0.80
1:S:16:ARG:NH2	1:S:114:GLN:O	2.16	0.78
2:X:156:GLN:OE1	2:X:165:ARG:NH2	2.18	0.77

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	215/240 (90%)	206 (96%)	9 (4%)	0	100	100
1	B	212/240 (88%)	203 (96%)	8 (4%)	1 (0%)	34	72
1	C	214/240 (89%)	205 (96%)	7 (3%)	2 (1%)	21	61
1	D	214/240 (89%)	203 (95%)	11 (5%)	0	100	100
1	E	213/240 (89%)	204 (96%)	9 (4%)	0	100	100
1	F	212/240 (88%)	204 (96%)	6 (3%)	2 (1%)	21	61
1	G	212/240 (88%)	201 (95%)	11 (5%)	0	100	100
1	O	213/240 (89%)	201 (94%)	12 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	P	215/240 (90%)	205 (95%)	9 (4%)	1 (0%)	34 72
1	Q	212/240 (88%)	205 (97%)	7 (3%)	0	100 100
1	R	212/240 (88%)	203 (96%)	9 (4%)	0	100 100
1	S	214/240 (89%)	203 (95%)	9 (4%)	2 (1%)	21 61
1	T	213/240 (89%)	204 (96%)	8 (4%)	1 (0%)	34 72
1	U	212/240 (88%)	205 (97%)	7 (3%)	0	100 100
2	H	220/240 (92%)	217 (99%)	3 (1%)	0	100 100
2	I	220/240 (92%)	215 (98%)	5 (2%)	0	100 100
2	J	220/240 (92%)	217 (99%)	3 (1%)	0	100 100
2	K	221/240 (92%)	216 (98%)	5 (2%)	0	100 100
2	L	221/240 (92%)	217 (98%)	4 (2%)	0	100 100
2	M	220/240 (92%)	216 (98%)	4 (2%)	0	100 100
2	N	221/240 (92%)	217 (98%)	4 (2%)	0	100 100
2	V	221/240 (92%)	217 (98%)	4 (2%)	0	100 100
2	W	221/240 (92%)	216 (98%)	5 (2%)	0	100 100
2	X	220/240 (92%)	216 (98%)	4 (2%)	0	100 100
2	Y	221/240 (92%)	217 (98%)	4 (2%)	0	100 100
2	Z	220/240 (92%)	217 (99%)	3 (1%)	0	100 100
2	a	221/240 (92%)	216 (98%)	5 (2%)	0	100 100
2	b	221/240 (92%)	217 (98%)	4 (2%)	0	100 100
All	All	6071/6720 (90%)	5883 (97%)	179 (3%)	9 (0%)	56 88

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	S	169	GLU
1	F	173	GLU
1	T	190	ALA
1	C	151	PRO
1	C	190	ALA

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	168/184 (91%)	168 (100%)	0	100	100
1	B	166/184 (90%)	166 (100%)	0	100	100
1	C	167/184 (91%)	167 (100%)	0	100	100
1	D	167/184 (91%)	166 (99%)	1 (1%)	90	96
1	E	166/184 (90%)	166 (100%)	0	100	100
1	F	165/184 (90%)	161 (98%)	4 (2%)	57	84
1	G	165/184 (90%)	164 (99%)	1 (1%)	90	96
1	O	166/184 (90%)	164 (99%)	2 (1%)	78	92
1	P	168/184 (91%)	168 (100%)	0	100	100
1	Q	166/184 (90%)	165 (99%)	1 (1%)	90	96
1	R	165/184 (90%)	165 (100%)	0	100	100
1	S	167/184 (91%)	167 (100%)	0	100	100
1	T	166/184 (90%)	166 (100%)	0	100	100
1	U	165/184 (90%)	164 (99%)	1 (1%)	90	96
2	H	165/178 (93%)	165 (100%)	0	100	100
2	I	165/178 (93%)	165 (100%)	0	100	100
2	J	165/178 (93%)	165 (100%)	0	100	100
2	K	165/178 (93%)	165 (100%)	0	100	100
2	L	165/178 (93%)	165 (100%)	0	100	100
2	M	165/178 (93%)	164 (99%)	1 (1%)	90	96
2	N	165/178 (93%)	164 (99%)	1 (1%)	90	96
2	V	165/178 (93%)	165 (100%)	0	100	100
2	W	165/178 (93%)	165 (100%)	0	100	100
2	X	165/178 (93%)	164 (99%)	1 (1%)	90	96
2	Y	165/178 (93%)	165 (100%)	0	100	100
2	Z	165/178 (93%)	165 (100%)	0	100	100
2	a	165/178 (93%)	165 (100%)	0	100	100
2	b	165/178 (93%)	164 (99%)	1 (1%)	90	96
All	All	4637/5068 (92%)	4623 (100%)	14 (0%)	94	97

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	32	ARG
2	N	6	LEU
1	U	97	ARG
1	G	80	GLN
1	Q	236	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	Q	152	HIS
2	W	137	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

14 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	7HZ	H	301	-	38,42,42	1.67	4 (10%)	43,58,58	2.00	12 (27%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	7HZ	I	301	-	38,42,42	1.57	4 (10%)	43,58,58	1.92	11 (25%)
3	7HZ	J	301	-	38,42,42	1.52	4 (10%)	43,58,58	1.91	11 (25%)
3	7HZ	K	301	-	38,42,42	1.64	4 (10%)	43,58,58	1.96	12 (27%)
3	7HZ	L	301	-	38,42,42	1.67	4 (10%)	43,58,58	1.88	11 (25%)
3	7HZ	M	301	-	38,42,42	1.53	5 (13%)	43,58,58	1.97	11 (25%)
3	7HZ	N	301	-	38,42,42	1.59	4 (10%)	43,58,58	1.97	12 (27%)
3	7HZ	V	301	-	38,42,42	1.49	4 (10%)	43,58,58	2.01	11 (25%)
3	7HZ	W	301	-	38,42,42	1.54	5 (13%)	43,58,58	1.91	11 (25%)
3	7HZ	X	301	-	38,42,42	1.55	4 (10%)	43,58,58	1.94	11 (25%)
3	7HZ	Y	301	-	38,42,42	1.62	5 (13%)	43,58,58	1.85	9 (20%)
3	7HZ	Z	301	-	38,42,42	1.62	5 (13%)	43,58,58	1.99	11 (25%)
3	7HZ	a	301	-	38,42,42	1.60	5 (13%)	43,58,58	1.83	10 (23%)
3	7HZ	b	301	-	38,42,42	1.55	4 (10%)	43,58,58	1.91	11 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	7HZ	H	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	I	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	J	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	K	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	L	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	M	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	N	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	V	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	W	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	X	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	Y	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	Z	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	a	301	-	-	0/34/38/38	0/2/3/3
3	7HZ	b	301	-	-	0/34/38/38	0/2/3/3

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	K	301	7HZ	O26-N25	-5.45	1.34	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	301	7HZ	O26-N25	-4.86	1.35	1.40
3	b	301	7HZ	O26-N25	-4.36	1.35	1.40
3	Y	301	7HZ	O26-N25	-4.10	1.36	1.40
3	M	301	7HZ	O26-N25	-3.79	1.36	1.40

The worst 5 of 154 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	Z	301	7HZ	O40-C33-N32	-4.19	114.86	122.45
3	V	301	7HZ	O40-C33-N32	-4.08	115.06	122.45
3	K	301	7HZ	O40-C33-N32	-4.04	115.13	122.45
3	N	301	7HZ	O40-C33-N32	-4.00	115.21	122.45
3	V	301	7HZ	O31-C24-N25	-3.99	118.74	123.57

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (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^{th} 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^2)	Q<0.9
1	A	219/240 (91%)	-0.47	0 100 100	41, 56, 84, 118	0
1	B	216/240 (90%)	-0.30	1 (0%) 91 82	42, 66, 105, 162	0
1	C	218/240 (90%)	-0.34	1 (0%) 91 82	42, 66, 103, 151	0
1	D	218/240 (90%)	-0.20	0 100 100	44, 71, 105, 121	0
1	E	217/240 (90%)	-0.35	0 100 100	40, 62, 90, 121	0
1	F	216/240 (90%)	-0.37	1 (0%) 91 82	41, 67, 101, 127	0
1	G	216/240 (90%)	-0.46	0 100 100	39, 57, 87, 135	0
1	O	217/240 (90%)	-0.26	0 100 100	44, 70, 110, 159	0
1	P	219/240 (91%)	-0.30	1 (0%) 91 82	43, 65, 97, 129	0
1	Q	216/240 (90%)	-0.43	1 (0%) 91 82	42, 61, 89, 130	0
1	R	216/240 (90%)	-0.37	0 100 100	40, 61, 91, 108	0
1	S	218/240 (90%)	-0.35	1 (0%) 91 82	40, 56, 90, 119	0
1	T	217/240 (90%)	-0.30	0 100 100	43, 64, 95, 117	0
1	U	216/240 (90%)	-0.45	1 (0%) 91 82	40, 59, 89, 118	0
2	H	222/240 (92%)	-0.44	0 100 100	40, 49, 74, 97	0
2	I	222/240 (92%)	-0.54	0 100 100	40, 47, 66, 83	0
2	J	222/240 (92%)	-0.57	0 100 100	41, 50, 74, 93	0
2	K	223/240 (92%)	-0.53	0 100 100	40, 49, 71, 88	0
2	L	223/240 (92%)	-0.55	0 100 100	39, 48, 69, 92	0
2	M	222/240 (92%)	-0.54	0 100 100	39, 50, 72, 109	0
2	N	223/240 (92%)	-0.48	1 (0%) 93 84	41, 54, 80, 130	0
2	V	223/240 (92%)	-0.50	0 100 100	39, 48, 67, 84	0
2	W	223/240 (92%)	-0.54	0 100 100	39, 49, 72, 98	0
2	X	222/240 (92%)	-0.58	0 100 100	40, 49, 72, 93	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
2	Y	223/240 (92%)	-0.53	0	100	100	41, 48, 72, 93	0
2	Z	222/240 (92%)	-0.56	0	100	100	39, 49, 73, 92	0
2	a	223/240 (92%)	-0.53	0	100	100	41, 52, 77, 106	0
2	b	223/240 (92%)	-0.48	0	100	100	40, 49, 76, 88	0
All	All	6155/6720 (91%)	-0.44	8 (0%)	95	91	39, 54, 91, 162	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	192	SER	3.0
1	C	236	ASP	2.6
1	F	11	GLN	2.2
1	S	169	GLU	2.1
2	N	223	GLY	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 carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	7HZ	Z	301	40/40	0.94	0.21	1.86	41,47,58,67	0
3	7HZ	N	301	40/40	0.94	0.23	1.46	48,51,70,82	0
3	7HZ	W	301	40/40	0.95	0.21	1.21	44,49,70,73	0
3	7HZ	H	301	40/40	0.95	0.21	1.19	43,49,74,76	0
3	7HZ	J	301	40/40	0.95	0.20	1.10	44,47,59,63	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	7HZ	Y	301	40/40	0.94	0.20	1.09	41,52,71,75	0
3	7HZ	M	301	40/40	0.94	0.20	1.03	44,51,69,76	0
3	7HZ	b	301	40/40	0.94	0.20	0.97	43,47,56,62	0
3	7HZ	a	301	40/40	0.95	0.20	0.86	45,51,61,62	0
3	7HZ	K	301	40/40	0.94	0.21	0.80	49,54,75,87	0
3	7HZ	I	301	40/40	0.94	0.18	0.72	41,52,66,91	0
3	7HZ	V	301	40/40	0.95	0.20	0.71	45,52,65,72	0
3	7HZ	L	301	40/40	0.94	0.19	0.48	40,48,62,75	0
3	7HZ	X	301	40/40	0.95	0.17	0.17	42,50,60,71	0

6.5 Other polymers (i)

There are no such residues in this entry.