



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 8, 2024 – 06:34 am GMT

PDB ID : 5MSJ  
Title : Mouse PA28alpha  
Authors : Huber, E.M.; Groll, M.  
Deposited on : 2017-01-05  
Resolution : 3.50 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

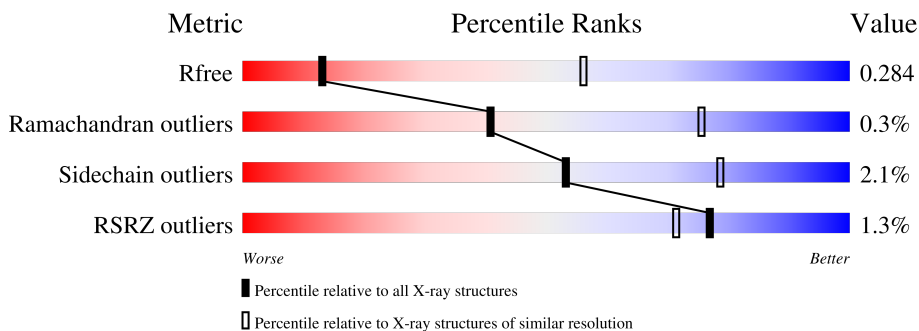
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 Å.

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







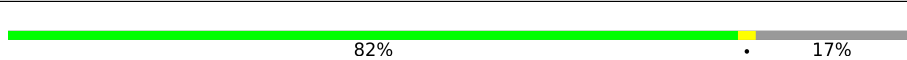
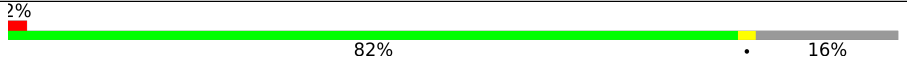
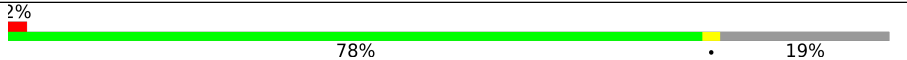
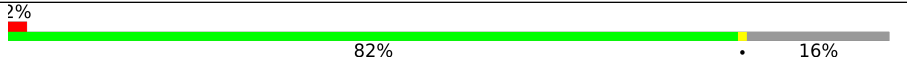
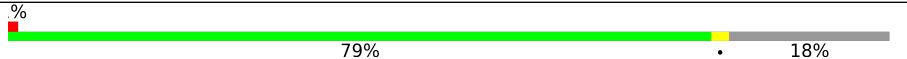
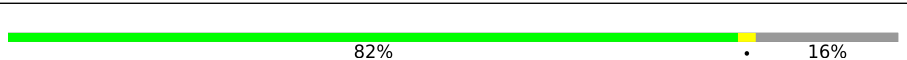
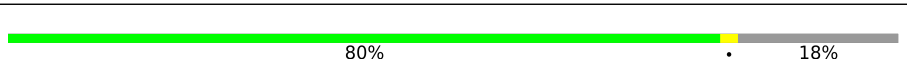
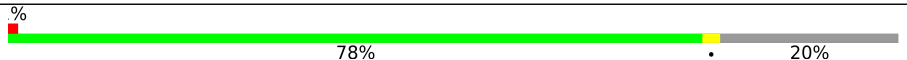


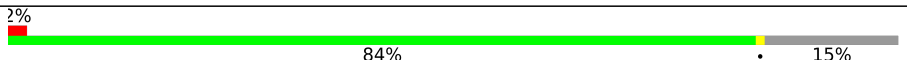
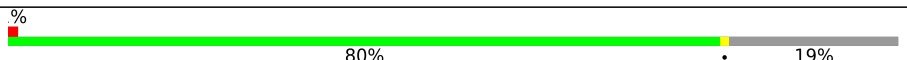
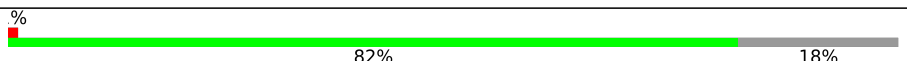
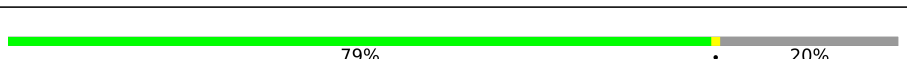
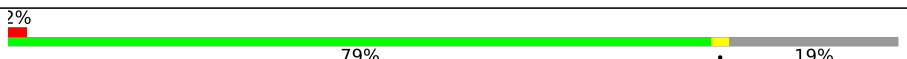

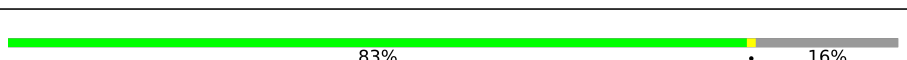
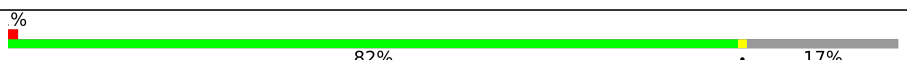
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1659 (3.60-3.40)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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

Mol	Chain	Length	Quality of chain
1	A	249	 81% 16%
1	B	249	 81% 18%
1	C	249	 2% 80% 18%
1	D	249	 80% 18%
1	E	249	 81% 16%
1	F	249	 79% 19%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	249	 81% 18%
1	H	249	 82% 16%
1	I	249	 78% 22%
1	J	249	 77% 19%
1	K	249	 82% 17%
1	L	249	 82% 16%
1	M	249	 78% 19%
1	N	249	 82% 16%
1	O	249	 79% 18%
1	P	249	 82% 16%
1	Q	249	 80% 18%
1	R	249	 78% 20%
1	S	249	 78% 20%
1	T	249	 80% 17%
1	U	249	 84% 15%
1	V	249	 80% 19%
1	W	249	 82% 18%
1	X	249	 79% 20%
1	Y	249	 79% 19%
1	Z	249	 81% 18%
1	a	249	 83% 16%
1	b	249	 82% 17%

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 46128 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 activator complex subunit 1.

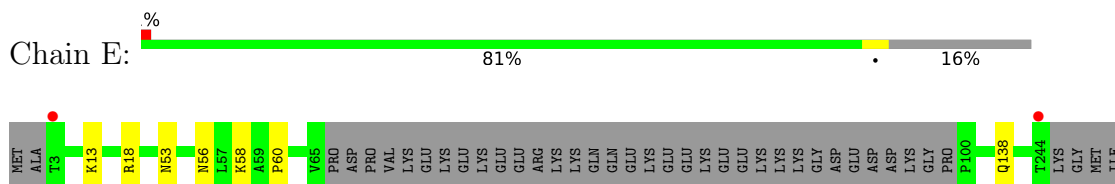
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	210	Total 1690	C 1081	N 288	O 315	S 6	0	0	0
1	B	204	Total 1646	C 1051	N 282	O 307	S 6	0	0	0
1	C	205	Total 1653	C 1056	N 283	O 308	S 6	0	0	0
1	D	203	Total 1635	C 1045	N 278	O 306	S 6	0	0	0
1	E	208	Total 1675	C 1070	N 286	O 313	S 6	0	0	0
1	F	201	Total 1619	C 1037	N 274	O 302	S 6	0	0	0
1	G	205	Total 1649	C 1054	N 280	O 309	S 6	0	0	0
1	H	208	Total 1676	C 1071	N 286	O 313	S 6	0	0	0
1	I	195	Total 1577	C 1006	N 269	O 297	S 5	0	0	0
1	J	202	Total 1635	C 1047	N 280	O 303	S 5	0	0	0
1	K	207	Total 1667	C 1066	N 285	O 310	S 6	0	0	0
1	L	209	Total 1682	C 1075	N 287	O 314	S 6	0	0	0
1	M	201	Total 1622	C 1038	N 276	O 302	S 6	0	0	0
1	N	208	Total 1675	C 1071	N 286	O 312	S 6	0	0	0
1	O	203	Total 1632	C 1044	N 276	O 306	S 6	0	0	0
1	P	208	Total 1675	C 1071	N 286	O 312	S 6	0	0	0

*Continued on next page...*

*Continued from previous page...*

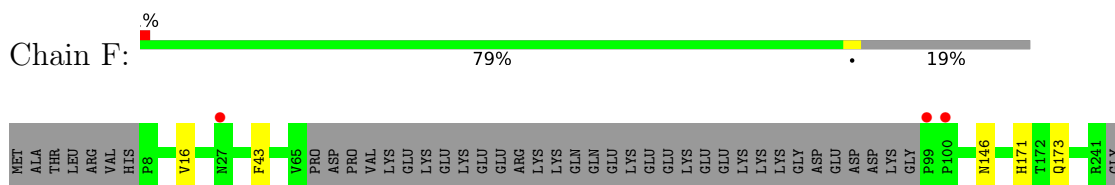
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Q	203	Total	C	N	O	S	0	0	0
			1635	1045	278	306	6			
1	R	200	Total	C	N	O	S	0	0	0
			1612	1032	273	301	6			
1	S	199	Total	C	N	O	S	0	0	0
			1602	1024	272	300	6			
1	T	206	Total	C	N	O	S	0	0	0
			1661	1062	284	309	6			
1	U	212	Total	C	N	O	S	0	0	0
			1711	1092	290	323	6			
1	V	202	Total	C	N	O	S	0	0	0
			1625	1039	275	305	6			
1	W	204	Total	C	N	O	S	0	0	0
			1647	1053	282	307	5			
1	X	200	Total	C	N	O	S	0	0	0
			1612	1031	273	303	5			
1	Y	201	Total	C	N	O	S	0	0	0
			1622	1037	276	304	5			
1	Z	204	Total	C	N	O	S	0	0	0
			1651	1057	282	307	5			
1	a	209	Total	C	N	O	S	0	0	0
			1683	1076	287	314	6			
1	b	206	Total	C	N	O	S	0	0	0
			1659	1062	284	307	6			



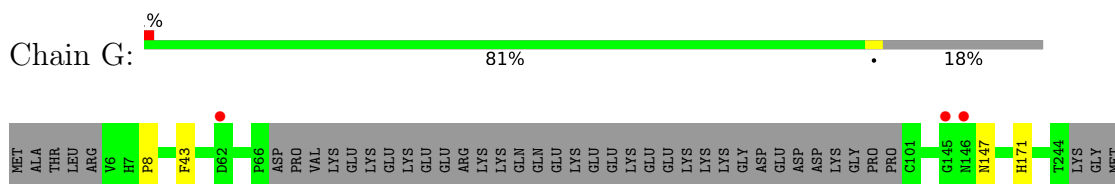


TYR

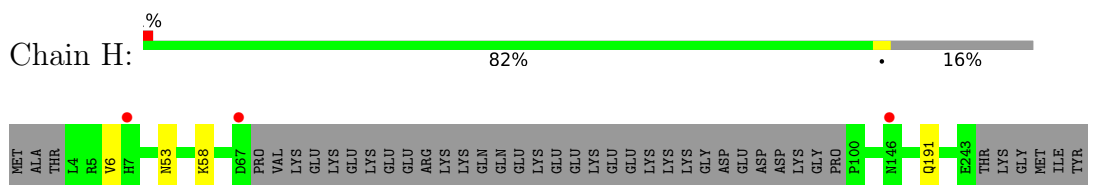
- Molecule 1: Proteasome activator complex subunit 1

GLU  
THR  
LYS  
GLY  
MET  
ILE  
TYR

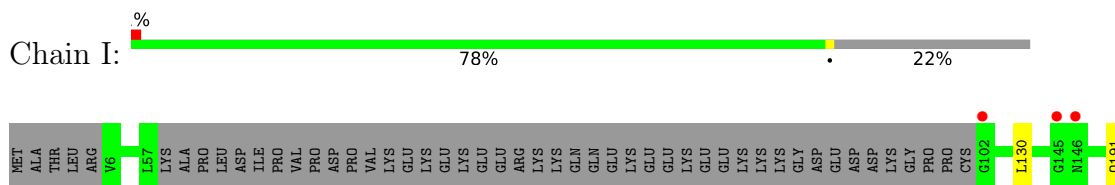
- Molecule 1: Proteasome activator complex subunit 1

ILE  
TYR

- Molecule 1: Proteasome activator complex subunit 1

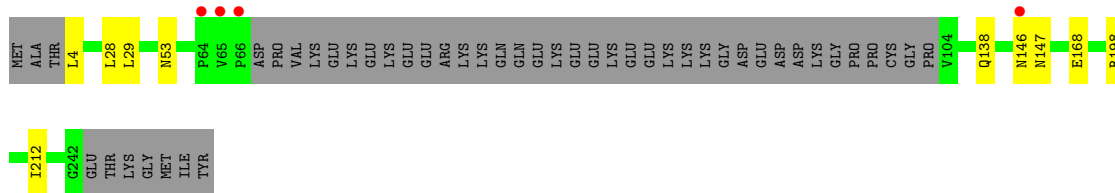


- Molecule 1: Proteasome activator complex subunit 1

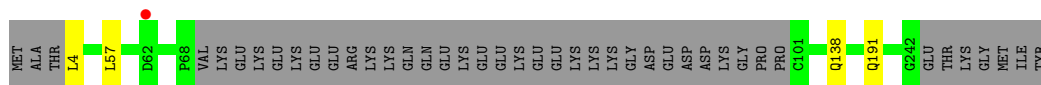
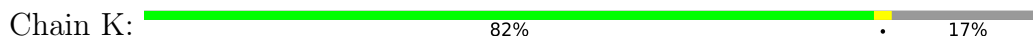
L244  
LYS  
GLY  
MET  
ILE  
TYR

- Molecule 1: Proteasome activator complex subunit 1

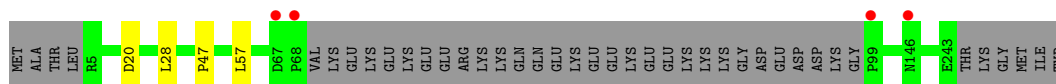
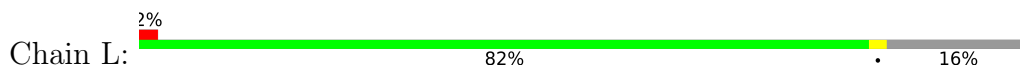




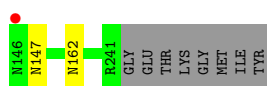
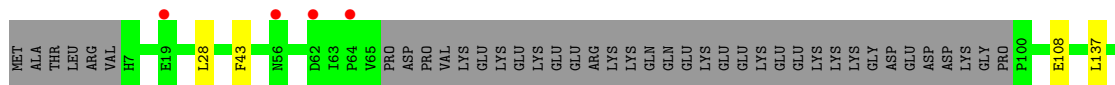
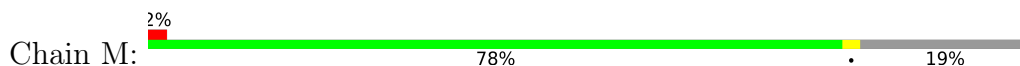
● Molecule 1: Proteasome activator complex subunit 1



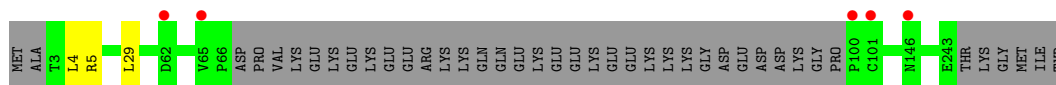
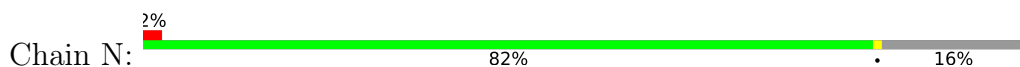
● Molecule 1: Proteasome activator complex subunit 1



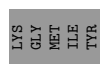
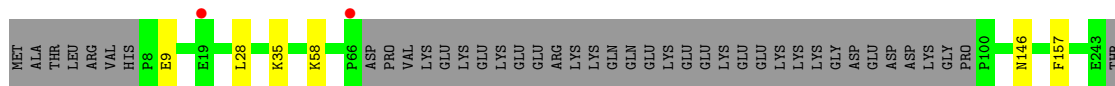
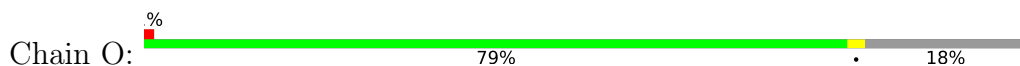
● Molecule 1: Proteasome activator complex subunit 1



● Molecule 1: Proteasome activator complex subunit 1



● Molecule 1: Proteasome activator complex subunit 1











## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.79Å 119.80Å 196.67Å 94.21° 98.52° 87.98°	Depositor
Resolution (Å)	29.65 – 3.50 29.65 – 3.50	Depositor EDS
% Data completeness (in resolution range)	93.1 (29.65-3.50) 93.2 (29.65-3.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 3.47Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.250 , 0.283 0.251 , 0.284	Depositor DCC
$R_{free}$ test set	4638 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.5	Xtrriage
Anisotropy	0.140	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 33.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	46128	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	100.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.50% 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	0.36	0/1722	0.52	0/2329
1	B	0.37	0/1675	0.52	0/2262
1	C	0.36	0/1683	0.52	0/2274
1	D	0.36	0/1665	0.51	0/2250
1	E	0.36	0/1705	0.53	0/2304
1	F	0.36	0/1649	0.53	0/2228
1	G	0.36	0/1679	0.53	0/2270
1	H	0.37	0/1707	0.52	0/2307
1	I	0.37	0/1604	0.51	0/2165
1	J	0.37	0/1664	0.55	0/2248
1	K	0.36	0/1698	0.52	0/2296
1	L	0.37	0/1715	0.53	0/2320
1	M	0.37	0/1652	0.50	0/2232
1	N	0.36	0/1706	0.53	0/2306
1	O	0.36	0/1662	0.51	0/2245
1	P	0.37	0/1706	0.53	0/2306
1	Q	0.36	0/1664	0.51	0/2248
1	R	0.36	0/1641	0.51	0/2217
1	S	0.36	0/1630	0.51	0/2200
1	T	0.36	0/1691	0.52	0/2285
1	U	0.36	0/1742	0.50	0/2352
1	V	0.37	0/1654	0.51	0/2234
1	W	0.36	0/1677	0.52	0/2266
1	X	0.36	0/1640	0.51	0/2214
1	Y	0.36	0/1651	0.49	0/2230
1	Z	0.36	0/1681	0.51	0/2271
1	a	0.36	0/1714	0.51	0/2318
1	b	0.36	0/1690	0.53	0/2284
All	All	0.36	0/46967	0.52	0/63461

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [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	206/249 (83%)	201 (98%)	4 (2%)	1 (0%)	29	68
1	B	200/249 (80%)	194 (97%)	6 (3%)	0	100	100
1	C	201/249 (81%)	197 (98%)	3 (2%)	1 (0%)	29	68
1	D	199/249 (80%)	195 (98%)	3 (2%)	1 (0%)	29	68
1	E	204/249 (82%)	199 (98%)	3 (2%)	2 (1%)	15	54
1	F	197/249 (79%)	190 (96%)	7 (4%)	0	100	100
1	G	201/249 (81%)	199 (99%)	1 (0%)	1 (0%)	29	68
1	H	204/249 (82%)	199 (98%)	3 (2%)	2 (1%)	15	54
1	I	191/249 (77%)	189 (99%)	2 (1%)	0	100	100
1	J	198/249 (80%)	194 (98%)	4 (2%)	0	100	100
1	K	203/249 (82%)	200 (98%)	3 (2%)	0	100	100
1	L	205/249 (82%)	200 (98%)	5 (2%)	0	100	100
1	M	197/249 (79%)	194 (98%)	3 (2%)	0	100	100
1	N	204/249 (82%)	201 (98%)	3 (2%)	0	100	100
1	O	199/249 (80%)	195 (98%)	3 (2%)	1 (0%)	29	68
1	P	204/249 (82%)	200 (98%)	3 (2%)	1 (0%)	29	68
1	Q	199/249 (80%)	195 (98%)	3 (2%)	1 (0%)	29	68
1	R	196/249 (79%)	192 (98%)	4 (2%)	0	100	100
1	S	195/249 (78%)	190 (97%)	5 (3%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	T	202/249 (81%)	198 (98%)	4 (2%)	0	100	100
1	U	208/249 (84%)	203 (98%)	4 (2%)	1 (0%)	29	68
1	V	198/249 (80%)	193 (98%)	5 (2%)	0	100	100
1	W	200/249 (80%)	196 (98%)	4 (2%)	0	100	100
1	X	196/249 (79%)	192 (98%)	2 (1%)	2 (1%)	15	54
1	Y	197/249 (79%)	192 (98%)	4 (2%)	1 (0%)	29	68
1	Z	200/249 (80%)	199 (100%)	1 (0%)	0	100	100
1	a	205/249 (82%)	196 (96%)	8 (4%)	1 (0%)	29	68
1	b	202/249 (81%)	198 (98%)	3 (2%)	1 (0%)	29	68
All	All	5611/6972 (80%)	5491 (98%)	103 (2%)	17 (0%)	41	75

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	X	146	ASN
1	D	58	LYS
1	H	58	LYS
1	Y	147	ASN
1	b	146	ASN

### 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	188/223 (84%)	181 (96%)	7 (4%)	34	65
1	B	182/223 (82%)	179 (98%)	3 (2%)	62	83
1	C	183/223 (82%)	179 (98%)	4 (2%)	52	78
1	D	181/223 (81%)	177 (98%)	4 (2%)	52	78
1	E	186/223 (83%)	181 (97%)	5 (3%)	44	73
1	F	180/223 (81%)	175 (97%)	5 (3%)	43	72
1	G	183/223 (82%)	180 (98%)	3 (2%)	62	83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	H	186/223 (83%)	184 (99%)	2 (1%)	73	88
1	I	174/223 (78%)	172 (99%)	2 (1%)	73	88
1	J	181/223 (81%)	171 (94%)	10 (6%)	21	54
1	K	185/223 (83%)	181 (98%)	4 (2%)	52	78
1	L	187/223 (84%)	183 (98%)	4 (2%)	53	79
1	M	180/223 (81%)	174 (97%)	6 (3%)	38	68
1	N	186/223 (83%)	183 (98%)	3 (2%)	62	83
1	O	181/223 (81%)	176 (97%)	5 (3%)	43	72
1	P	186/223 (83%)	183 (98%)	3 (2%)	62	83
1	Q	181/223 (81%)	177 (98%)	4 (2%)	52	78
1	R	179/223 (80%)	174 (97%)	5 (3%)	43	72
1	S	177/223 (79%)	172 (97%)	5 (3%)	43	72
1	T	184/223 (82%)	176 (96%)	8 (4%)	29	62
1	U	190/223 (85%)	189 (100%)	1 (0%)	88	94
1	V	180/223 (81%)	177 (98%)	3 (2%)	60	82
1	W	182/223 (82%)	182 (100%)	0	100	100
1	X	178/223 (80%)	177 (99%)	1 (1%)	86	94
1	Y	179/223 (80%)	176 (98%)	3 (2%)	60	82
1	Z	183/223 (82%)	180 (98%)	3 (2%)	62	83
1	a	187/223 (84%)	186 (100%)	1 (0%)	88	94
1	b	184/223 (82%)	182 (99%)	2 (1%)	73	88
All	All	5113/6244 (82%)	5007 (98%)	106 (2%)	53	79

5 of 106 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	147	ASN
1	Q	40	LEU
1	Y	146	ASN
1	N	4	LEU
1	O	146	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:



Mol	Chain	Res	Type
1	X	50	ASN
1	X	191	GLN
1	I	191	GLN
1	I	53	ASN
1	Y	105	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 [i](#)

There are no chain breaks in this entry.

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

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

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	210/249 (84%)	-0.33	3 (1%) 75 69	66, 89, 122, 152	0
1	B	204/249 (81%)	-0.26	2 (0%) 82 77	70, 95, 136, 141	0
1	C	205/249 (82%)	-0.21	4 (1%) 65 60	71, 97, 133, 149	0
1	D	203/249 (81%)	-0.21	3 (1%) 73 68	76, 100, 148, 160	0
1	E	208/249 (83%)	-0.26	2 (0%) 82 77	82, 105, 135, 148	0
1	F	201/249 (80%)	-0.22	3 (1%) 73 68	73, 95, 137, 152	0
1	G	205/249 (82%)	-0.34	3 (1%) 73 68	71, 88, 122, 138	0
1	H	208/249 (83%)	-0.26	3 (1%) 75 69	79, 100, 134, 162	0
1	I	195/249 (78%)	-0.20	3 (1%) 73 68	79, 105, 143, 158	0
1	J	202/249 (81%)	-0.23	4 (1%) 65 60	75, 98, 143, 154	0
1	K	207/249 (83%)	-0.27	1 (0%) 91 88	75, 101, 144, 157	0
1	L	209/249 (83%)	-0.25	4 (1%) 66 61	76, 106, 137, 165	0
1	M	201/249 (80%)	-0.13	5 (2%) 57 51	75, 107, 145, 157	0
1	N	208/249 (83%)	-0.21	5 (2%) 59 53	73, 98, 137, 147	0
1	O	203/249 (81%)	-0.29	2 (0%) 82 77	71, 93, 125, 147	0
1	P	208/249 (83%)	-0.28	1 (0%) 91 88	72, 95, 132, 144	0
1	Q	203/249 (81%)	-0.25	1 (0%) 91 88	76, 100, 129, 135	0
1	R	200/249 (80%)	-0.27	3 (1%) 73 68	81, 100, 140, 171	0
1	S	199/249 (79%)	-0.28	3 (1%) 73 68	69, 94, 137, 159	0
1	T	206/249 (82%)	-0.35	0 100 100	70, 93, 131, 139	0
1	U	212/249 (85%)	-0.30	6 (2%) 53 47	69, 92, 144, 174	0
1	V	202/249 (81%)	-0.25	3 (1%) 73 68	76, 95, 134, 161	0
1	W	204/249 (81%)	-0.28	2 (0%) 82 77	73, 95, 137, 150	0
1	X	200/249 (80%)	-0.31	1 (0%) 91 88	69, 94, 128, 136	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	Y	201/249 (80%)	-0.28	4 (1%) 65 60	67, 96, 136, 162	0
1	Z	204/249 (81%)	-0.29	2 (0%) 82 77	68, 95, 134, 159	0
1	a	209/249 (83%)	-0.28	1 (0%) 91 88	71, 98, 133, 155	0
1	b	206/249 (82%)	-0.32	2 (0%) 82 77	71, 97, 133, 148	0
All	All	5723/6972 (82%)	-0.27	76 (1%) 77 71	66, 97, 137, 174	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	100	PRO	4.2
1	O	66	PRO	3.9
1	V	66	PRO	3.8
1	I	146	ASN	3.7
1	U	68	PRO	3.7

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.