



wwPDB X-ray Structure Validation Summary Report

Oct 9, 2023 – 01:57 PM EDT

PDB ID : 6WOZ
Title : Plasmodium vivax reticulocyte binding protein 2b (PvRBP2b) bound to human monoclonal antibody 251249
Authors : Chan, L.J.; Dietrich, M.H.; Tham, W.H.
Deposited on : 2020-04-26
Resolution : 2.90 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

<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 : 2.35.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

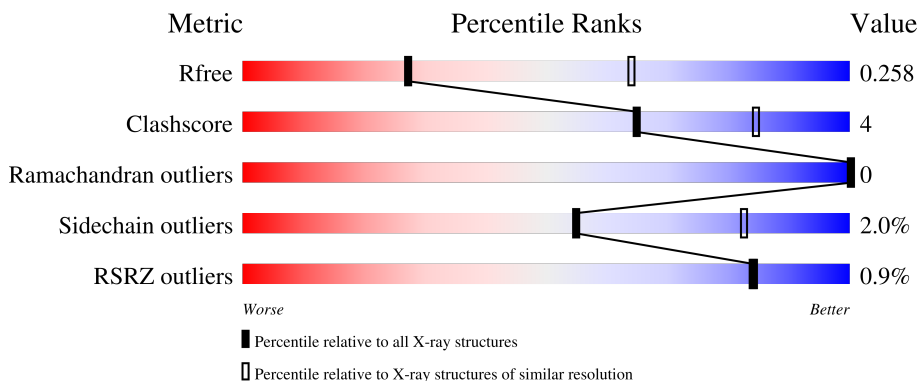
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 ≥ 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	307	 85% 11%
1	D	307	 84% 12%
1	G	307	 86% 11%
1	J	307	 85% 11%
2	B	241	 85% 9% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	241	 <p>% 83% 12% 5%</p>
2	H	241	 <p>2% 81% 13% 6%</p>
2	K	241	 <p>% 76% 18% 6%</p>
3	C	219	 <p>87% 11% .</p>
3	F	219	 <p>84% 13% .</p>
3	I	219	 <p>% 91% 6% .</p>
3	L	219	 <p>3% 84% 13% .</p>

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 22580 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 reticulocyte binding protein 2b.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	J	297	2354	1514	392	439	9	0	0	0
1	D	296	2390	1536	403	442	9	0	0	0
1	A	296	2396	1541	400	446	9	0	0	0
1	G	298	2407	1547	402	449	9	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	164	GLY	-	expression tag	UNP A5K736
J	165	ALA	-	expression tag	UNP A5K736
J	166	MET	-	expression tag	UNP A5K736
J	167	GLY	-	expression tag	UNP A5K736
J	168	SER	-	expression tag	UNP A5K736
D	164	GLY	-	expression tag	UNP A5K736
D	165	ALA	-	expression tag	UNP A5K736
D	166	MET	-	expression tag	UNP A5K736
D	167	GLY	-	expression tag	UNP A5K736
D	168	SER	-	expression tag	UNP A5K736
A	164	GLY	-	expression tag	UNP A5K736
A	165	ALA	-	expression tag	UNP A5K736
A	166	MET	-	expression tag	UNP A5K736
A	167	GLY	-	expression tag	UNP A5K736
A	168	SER	-	expression tag	UNP A5K736
G	164	GLY	-	expression tag	UNP A5K736
G	165	ALA	-	expression tag	UNP A5K736
G	166	MET	-	expression tag	UNP A5K736
G	167	GLY	-	expression tag	UNP A5K736
G	168	SER	-	expression tag	UNP A5K736

- Molecule 2 is a protein called 251249 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	E	229	Total 1674	C 1068	N 271	O 329	S 6	0	0	0
2	B	227	Total 1669	C 1065	N 269	O 329	S 6	0	0	0
2	H	227	Total 1668	C 1064	N 271	O 327	S 6	0	0	0
2	K	226	Total 1658	C 1057	N 267	O 328	S 6	0	0	0

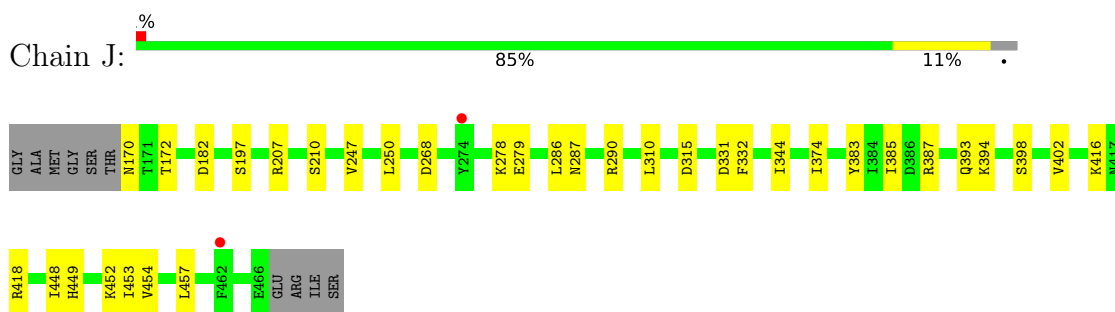
- Molecule 3 is a protein called 251249 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	F	213	Total 1593	C 995	N 271	O 323	S 4	0	0	0
3	C	213	Total 1603	C 1006	N 269	O 324	S 4	0	0	0
3	I	213	Total 1586	C 994	N 268	O 320	S 4	0	0	0
3	L	213	Total 1582	C 991	N 263	O 324	S 4	0	0	0

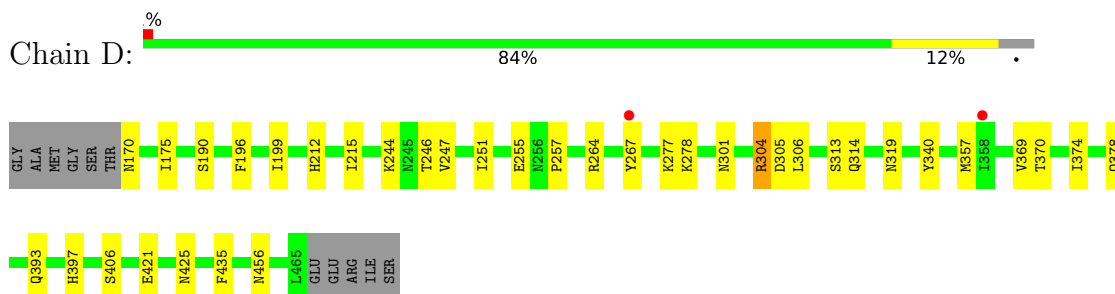
3 Residue-property plots [i](#)

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

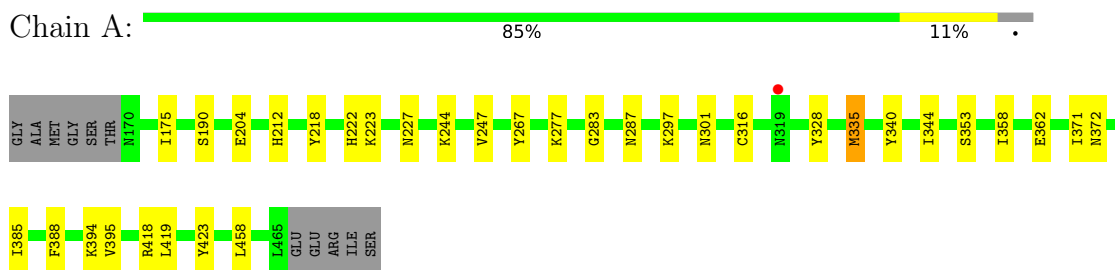
- Molecule 1: reticulocyte binding protein 2b



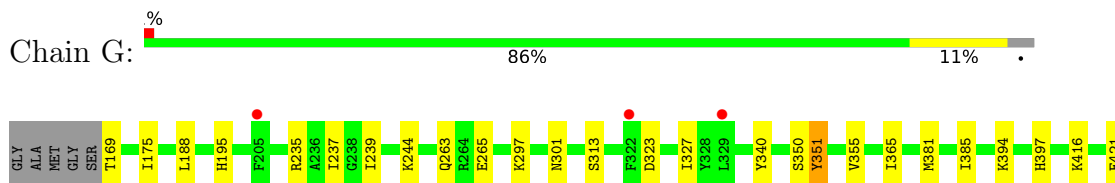
- Molecule 1: reticulocyte binding protein 2b



- Molecule 1: reticulocyte binding protein 2b

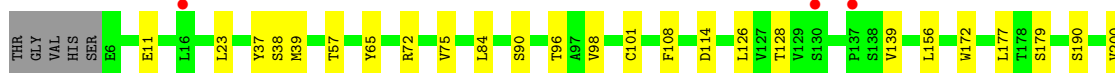
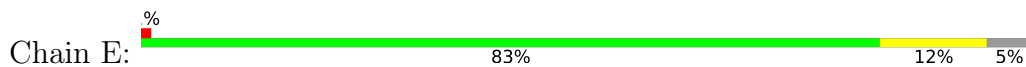


- Molecule 1: reticulocyte binding protein 2b

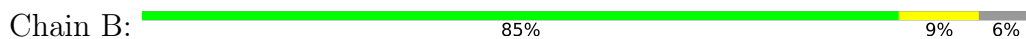




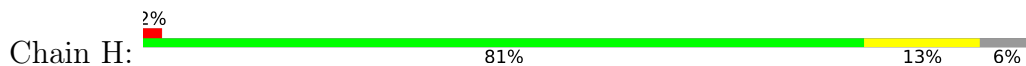
- Molecule 2: 251249 Fab heavy chain



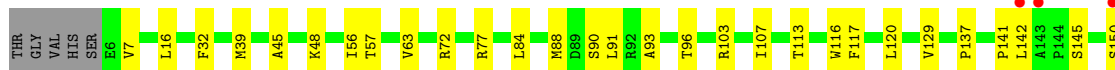
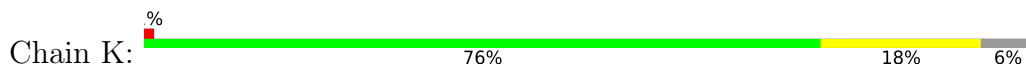
- Molecule 2: 251249 Fab heavy chain



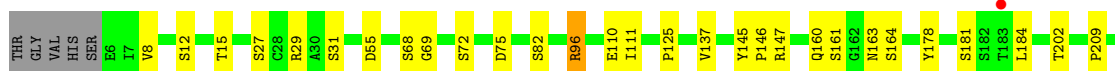
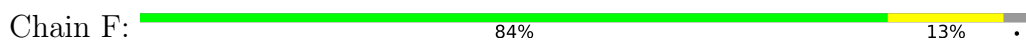
- Molecule 2: 251249 Fab heavy chain



- Molecule 2: 251249 Fab heavy chain




- Molecule 3: 251249 Fab light chain



E218
CYS

- Molecule 3: 251249 Fab light chain

Chain C:  87% 11%


THR	GLY	VAL	HIS	HIS	SER	E6	I7	S19	S27	Q32	Q42	L52	T61	S68	G69	S70	D75	T79	D87	Y91	P118	Q129	F144	K154	H194	A198	S208	N215	R216	G217	E218	CYS
-----	-----	-----	-----	-----	-----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	-----

- Molecule 3: 251249 Fab light chain

Chain I:  % 91% 6%

THR	GLY	VAL	HIS	SER	E6	S12	P13	Q42	L52	D65	E110	T134	V137	Q160	M163	S164	Q165	E166	S167	Y178	L184	T185	L186	S208	E218	CYS
-----	-----	-----	-----	-----	----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

- Molecule 3: 251249 Fab light chain

Chain L:  3% 84% 13%

THR	GLY	VAL	HIS	SER	E6	R23	A24	T25	L26	F37	D75	L78	L83	F88	R96	S97	V109	P118	S119	V120	T134	V137	F144	Y145	P146	R147	W153	S164	D175	S181	L184	S187	D190	Y191	H194	Y197	A198
-----	-----	-----	-----	-----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

V201	T202	S207	S208	P209	R216	G217	E218	CYS
------	------	------	------	------	------	------	------	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	99.57Å 163.55Å 121.76Å 90.00° 99.19° 90.00°	Depositor
Resolution (Å)	49.15 – 2.90 49.14 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.15-2.90) 99.9 (49.14-2.90)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.41 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.15_3459	Depositor
R, R_{free}	0.211 , 0.258 0.211 , 0.258	Depositor DCC
R_{free} test set	4253 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	64.4	Xtrriage
Anisotropy	0.165	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 33.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	22580	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.40 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.0486e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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	A	0.25	0/2443	0.35	0/3294
1	D	0.24	0/2438	0.36	0/3291
1	G	0.25	0/2456	0.36	0/3319
1	J	0.24	0/2402	0.35	0/3248
2	B	0.26	0/1711	0.46	0/2341
2	E	0.26	0/1716	0.45	0/2350
2	H	0.26	0/1710	0.46	0/2341
2	K	0.26	0/1700	0.45	0/2329
3	C	0.26	0/1638	0.45	0/2232
3	F	0.25	0/1628	0.45	0/2221
3	I	0.25	0/1621	0.46	0/2213
3	L	0.25	0/1617	0.44	0/2209
All	All	0.25	0/23080	0.42	0/31388

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](#)

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	2396	0	2314	18	0
1	D	2390	0	2282	22	0
1	G	2407	0	2295	19	0
1	J	2354	0	2203	22	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1669	0	1575	14	0
2	E	1674	0	1567	20	0
2	H	1668	0	1577	20	0
2	K	1658	0	1558	24	0
3	C	1603	0	1513	9	0
3	F	1593	0	1488	13	0
3	I	1586	0	1481	7	0
3	L	1582	0	1463	14	0
All	All	22580	0	21316	181	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 4.

The worst 5 of 181 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:K:96:THR:HG22	2:K:129:VAL:H	1.38	0.87
3:F:125:PRO:HD3	3:F:137:VAL:HG22	1.64	0.79
3:F:164:SER:HB2	3:F:184:LEU:HA	1.70	0.73
2:B:137:PRO:HB3	2:B:163:TYR:HB3	1.73	0.71
2:H:96:THR:HG23	2:H:128:THR:HA	1.75	0.69

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	294/307 (96%)	284 (97%)	10 (3%)	0	100 100
1	D	294/307 (96%)	288 (98%)	6 (2%)	0	100 100
1	G	296/307 (96%)	287 (97%)	9 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	J	295/307 (96%)	284 (96%)	11 (4%)	0	100	100
2	B	225/241 (93%)	218 (97%)	7 (3%)	0	100	100
2	E	227/241 (94%)	220 (97%)	7 (3%)	0	100	100
2	H	225/241 (93%)	217 (96%)	8 (4%)	0	100	100
2	K	224/241 (93%)	215 (96%)	9 (4%)	0	100	100
3	C	211/219 (96%)	198 (94%)	13 (6%)	0	100	100
3	F	211/219 (96%)	202 (96%)	9 (4%)	0	100	100
3	I	211/219 (96%)	204 (97%)	7 (3%)	0	100	100
3	L	211/219 (96%)	198 (94%)	13 (6%)	0	100	100
All	All	2924/3068 (95%)	2815 (96%)	109 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	253/286 (88%)	249 (98%)	4 (2%)	62	86
1	D	248/286 (87%)	243 (98%)	5 (2%)	55	82
1	G	252/286 (88%)	249 (99%)	3 (1%)	71	91
1	J	238/286 (83%)	235 (99%)	3 (1%)	69	90
2	B	180/207 (87%)	178 (99%)	2 (1%)	73	92
2	E	178/207 (86%)	174 (98%)	4 (2%)	52	81
2	H	180/207 (87%)	178 (99%)	2 (1%)	73	92
2	K	179/207 (86%)	176 (98%)	3 (2%)	60	86
3	C	174/189 (92%)	168 (97%)	6 (3%)	37	71
3	F	172/189 (91%)	166 (96%)	6 (4%)	36	70
3	I	170/189 (90%)	167 (98%)	3 (2%)	59	85
3	L	169/189 (89%)	161 (95%)	8 (5%)	26	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2393/2728 (88%)	2344 (98%)	49 (2%)	55 82

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	353	SER
2	K	145	SER
1	A	418	ARG
3	I	55	ASP
2	K	198	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	J	393	GLN
3	F	160	GLN
1	D	245	ASN
2	H	210	GLN
3	I	165	GLN

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, 95th 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	296/307 (96%)	0.06	1 (0%) 94 94	44, 60, 82, 96	0
1	D	296/307 (96%)	0.06	2 (0%) 87 87	44, 58, 82, 93	0
1	G	298/307 (97%)	0.11	3 (1%) 82 82	45, 61, 78, 95	0
1	J	297/307 (96%)	0.12	2 (0%) 87 87	45, 69, 89, 103	0
2	B	227/241 (94%)	0.10	0 100 100	39, 56, 74, 91	0
2	E	229/241 (95%)	0.20	3 (1%) 77 77	44, 61, 81, 92	0
2	H	227/241 (94%)	0.19	4 (1%) 68 67	43, 63, 85, 92	0
2	K	226/241 (93%)	0.13	3 (1%) 77 77	43, 61, 85, 106	0
3	C	213/219 (97%)	-0.02	0 100 100	41, 58, 71, 93	0
3	F	213/219 (97%)	0.03	1 (0%) 91 91	45, 56, 76, 82	0
3	I	213/219 (97%)	0.06	2 (0%) 84 84	42, 55, 82, 90	0
3	L	213/219 (97%)	0.17	6 (2%) 53 49	45, 63, 86, 99	0
All	All	2948/3068 (96%)	0.10	27 (0%) 84 84	39, 60, 83, 106	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	191	TYR	4.5
3	L	198	ALA	3.7
2	H	167	PRO	3.4
3	I	134	THR	3.1
3	L	197	TYR	2.9

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.