Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 11:11 am BST

PDB ID	:	6PHF
Title	:	Pfs25 in complex with the human transmission blocking antibody 2587
Authors	:	McLeod, B.R.; Julien, J.P.
Deposited on	:	2019-06-25
Resolution	:	3.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references ①) were used in the production of this report:

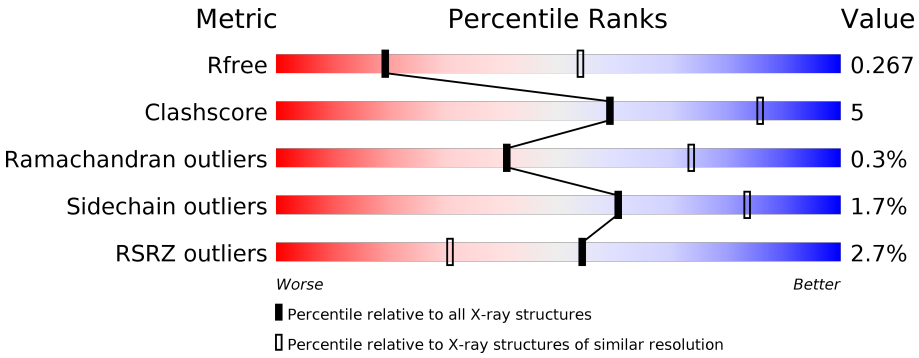
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 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 <=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	223	79% 17%
1	C	223	78% 19%
2	E	184	5% 79% 14% 8%
2	G	184	10% 77% 10% 13%
3	B	213	90% 7%
3	D	213	87% 10%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8766 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 2587 Antibody Fab, Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	215	Total	C	N	O	S	0	0	0
			1622	1022	274	319	7			
1	C	216	Total	C	N	O	S	0	0	0
			1630	1025	275	323	7			

- Molecule 2 is a protein called 25 kDa ookinete surface antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	161	Total	C	N	O	S	0	0	0
			1171	708	200	239	24			
2	E	170	Total	C	N	O	S	0	0	0
			1236	752	209	251	24			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	-2	GLU	-	expression tag	UNP P13829
G	-1	THR	-	expression tag	UNP P13829
G	0	GLY	-	expression tag	UNP P13829
G	91	GLN	ASN	conflict	UNP P13829
G	144	GLN	ASN	conflict	UNP P13829
G	166	GLN	ASN	conflict	UNP P13829
G	173	GLY	-	expression tag	UNP P13829
G	174	THR	-	expression tag	UNP P13829
G	175	LYS	-	expression tag	UNP P13829
G	176	HIS	-	expression tag	UNP P13829
G	177	HIS	-	expression tag	UNP P13829
G	178	HIS	-	expression tag	UNP P13829
G	179	HIS	-	expression tag	UNP P13829
G	180	HIS	-	expression tag	UNP P13829
G	181	HIS	-	expression tag	UNP P13829
E	-2	GLU	-	expression tag	UNP P13829

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	THR	-	expression tag	UNP P13829
E	0	GLY	-	expression tag	UNP P13829
E	91	GLN	ASN	conflict	UNP P13829
E	144	GLN	ASN	conflict	UNP P13829
E	166	GLN	ASN	conflict	UNP P13829
E	173	GLY	-	expression tag	UNP P13829
E	174	THR	-	expression tag	UNP P13829
E	175	LYS	-	expression tag	UNP P13829
E	176	HIS	-	expression tag	UNP P13829
E	177	HIS	-	expression tag	UNP P13829
E	178	HIS	-	expression tag	UNP P13829
E	179	HIS	-	expression tag	UNP P13829
E	180	HIS	-	expression tag	UNP P13829
E	181	HIS	-	expression tag	UNP P13829

- Molecule 3 is a protein called 2587 Antibody Fab, Light Chain.

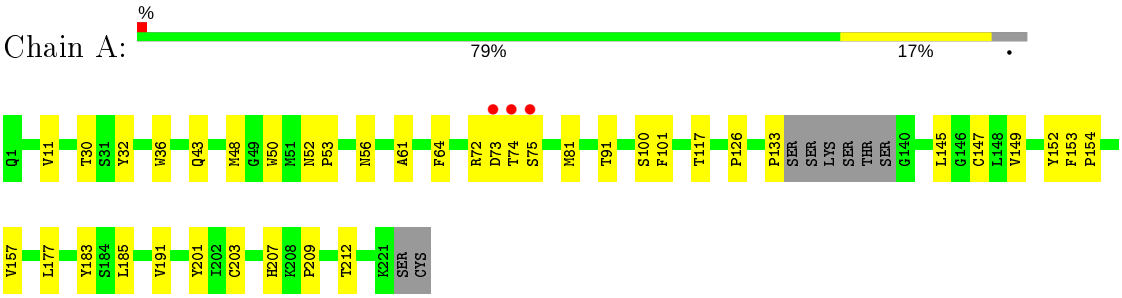
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	207	Total	C	N	O	S	0	0	0
			1553	974	259	315	5			
3	D	208	Total	C	N	O	S	0	0	0
			1554	974	259	316	5			

3 Residue-property plots

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.

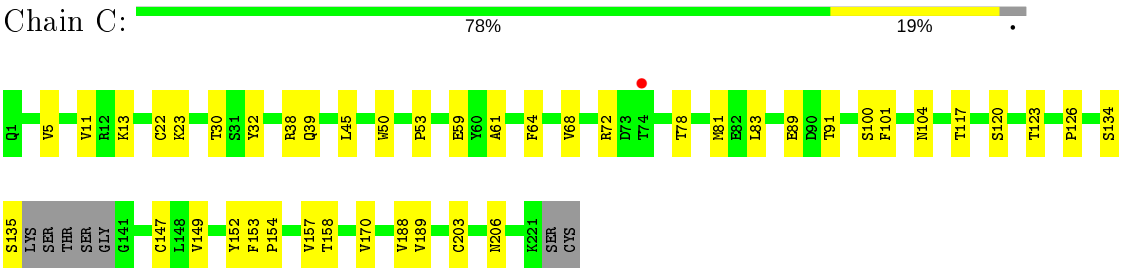
- Molecule 1: 2587 Antibody Fab, Heavy Chain

Chain A:



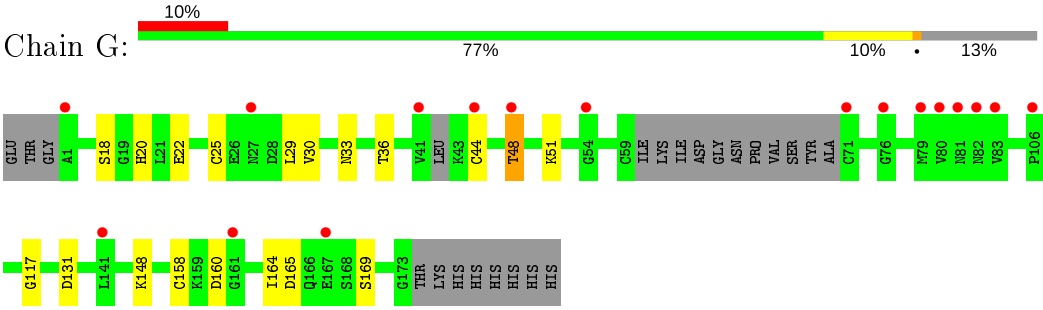
- Molecule 1: 2587 Antibody Fab, Heavy Chain

Chain C:



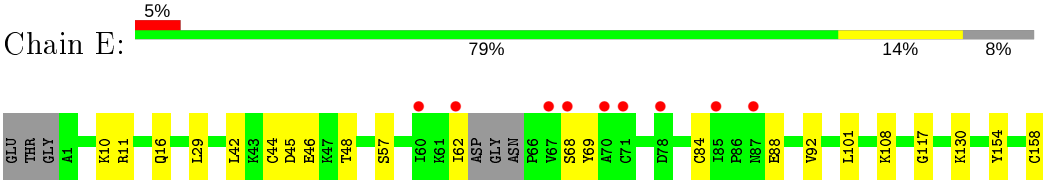
- Molecule 2: 25 kDa ookinete surface antigen

Chain G:



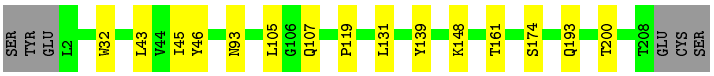
- Molecule 2: 25 kDa ookinete surface antigen

Chain E:



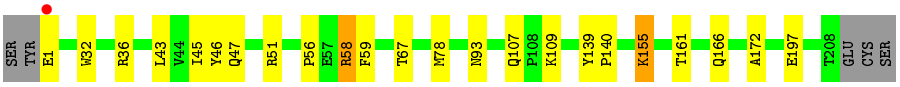
- Molecule 3: 2587 Antibody Fab, Light Chain

Chain B:



- Molecule 3: 2587 Antibody Fab, Light Chain

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.24Å 110.27Å 147.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.75 – 3.10 39.75 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.75-3.10) 90.8 (39.75-3.10)	Depositor EDS
R_{merge}	0.28	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.237 , 0.267 0.237 , 0.267	Depositor DCC
R_{free} test set	1291 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	50.1	Xtriage
Anisotropy	0.473	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 27.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	8766	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.35 % 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 4.7439e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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	0/1661	0.46	0/2263
1	C	0.25	0/1669	0.46	0/2275
2	E	0.24	0/1246	0.46	0/1678
2	G	0.24	0/1178	0.46	0/1584
3	B	0.25	0/1593	0.44	0/2176
3	D	0.25	0/1594	0.44	0/2179
All	All	0.25	0/8941	0.45	0/12155

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

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	1622	0	1571	23	0
1	C	1630	0	1573	21	0
2	E	1236	0	1175	13	0
2	G	1171	0	1110	10	0
3	B	1553	0	1499	8	0
3	D	1554	0	1493	13	0
All	All	8766	0	8421	80	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.

All (80) 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)
3:B:161:THR:HG22	3:B:174:SER:H	1.52	0.75
2:E:57:SER:OG	2:E:84:CYS:SG	2.49	0.70
1:A:133:PRO:HG3	1:A:145:LEU:HB3	1.74	0.69
2:E:101:LEU:HD21	2:E:108:LYS:HD3	1.75	0.69
2:E:11:ARG:HE	2:E:29:LEU:HD21	1.57	0.69
3:D:36:ARG:NH2	3:D:78:MET:O	2.28	0.67
2:G:18:SER:HB3	2:G:114:CYS:H	1.60	0.67
1:C:53:PRO:O	1:C:72:ARG:NH1	2.28	0.66
3:B:148:LYS:NZ	3:B:193:GLN:OE1	2.33	0.62
1:A:126:PRO:HB3	1:A:152:TYR:HB3	1.81	0.61
1:C:91:THR:HG23	1:C:117:THR:HA	1.83	0.60
3:D:32:TRP:HB2	3:D:45:ILE:HB	1.84	0.60
1:C:5:VAL:HG13	1:C:23:LYS:HB3	1.85	0.59
3:D:109:LYS:HE3	3:D:197:GLU:HG3	1.84	0.58
1:A:61:ALA:HB3	1:A:64:PHE:HD2	1.69	0.56
1:A:11:VAL:HB	1:A:154:PRO:HG3	1.87	0.56
1:C:126:PRO:HB3	1:C:152:TYR:HB3	1.86	0.56
1:A:91:THR:HG23	1:A:117:THR:HA	1.88	0.56
1:A:43:GLN:NE2	3:D:1:GLU:OE1	2.39	0.55
1:A:177:LEU:HD12	1:A:183:TYR:CZ	2.42	0.55
3:D:155:LYS:H	3:D:155:LYS:HD3	1.71	0.55
3:B:32:TRP:HB2	3:B:45:ILE:HB	1.89	0.54
1:A:52:ASN:ND2	2:G:131:ASP:OD2	2.40	0.53
1:A:101:PHE:HB3	2:G:117:GLY:HA2	1.89	0.53
2:E:16:GLN:HB3	2:E:154:TYR:HB2	1.89	0.53
1:A:43:GLN:OE1	1:A:43:GLN:N	2.42	0.52
1:A:73:ASP:OD2	1:A:75:SER:OG	2.21	0.52
1:A:177:LEU:HD22	3:D:67:THR:HG23	1.91	0.51
2:G:25:CYS:HB3	2:G:29:LEU:HB2	1.92	0.51
1:A:32:TYR:CD1	1:A:100:SER:HB2	2.46	0.51
3:B:119:PRO:HD3	3:B:131:LEU:HG	1.94	0.50
3:D:51:ARG:HD3	3:D:59:PHE:O	2.11	0.50
3:B:43:LEU:HD21	3:B:46:TYR:HB3	1.93	0.50
1:C:101:PHE:HB3	2:E:117:GLY:HA2	1.94	0.49
2:G:158:CYS:SG	2:G:164:ILE:HG23	2.52	0.49
2:E:88:GLU:HG2	2:E:108:LYS:O	2.13	0.48
1:A:72:ARG:HD2	1:A:74:THR:HG23	1.96	0.48
2:E:62:ILE:N	2:E:68:SER:OG	2.31	0.47
1:C:30:THR:HA	1:C:53:PRO:HG2	1.95	0.47
1:A:149:VAL:HG11	1:A:157:VAL:HG11	1.97	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:30:THR:HA	1:A:53:PRO:HG2	1.98	0.46
3:B:193:GLN:HE21	3:B:200:THR:HG21	1.80	0.46
1:C:13:LYS:NZ	1:C:120:SER:O	2.45	0.46
1:C:38:ARG:NH2	1:C:89:GLU:O	2.48	0.46
1:C:153:PHE:HA	1:C:154:PRO:HA	1.73	0.45
2:G:148:LYS:NZ	2:G:160:ASP:OD1	2.30	0.45
1:A:126:PRO:HD2	1:A:212:THR:HG21	1.99	0.45
1:C:59:GLU:OE2	2:E:130:LYS:NZ	2.47	0.45
3:D:56:PRO:HB2	3:D:58:ARG:HG3	1.99	0.44
1:C:134:SER:OG	1:C:135:SER:N	2.50	0.44
1:C:104:ASN:O	3:D:47:GLN:NE2	2.34	0.44
1:A:36:TRP:CE2	1:A:81:MET:HB2	2.53	0.44
1:C:39:GLN:HB2	1:C:45:LEU:HD12	1.98	0.44
1:C:149:VAL:HG11	1:C:157:VAL:HG11	2.00	0.43
2:G:165:ASP:O	2:G:169:SER:N	2.49	0.43
1:C:11:VAL:HG11	1:C:123:THR:HB	2.00	0.43
1:C:170:VAL:HG22	1:C:189:VAL:HG22	2.01	0.42
3:D:43:LEU:HD21	3:D:46:TYR:HB3	2.02	0.42
2:E:10:LYS:HE3	2:E:10:LYS:HB3	1.88	0.42
2:G:20:HIS:CD2	2:G:22:GLU:HG3	2.54	0.42
3:D:166:GLN:OE1	3:D:172:ALA:HB2	2.20	0.42
1:C:32:TYR:CD1	1:C:100:SER:HB2	2.55	0.42
3:D:107:GLN:HB2	3:D:139:TYR:CE2	2.55	0.42
3:B:105:LEU:HD13	3:D:140:PRO:HB2	2.02	0.42
1:A:152:TYR:OH	1:A:185:LEU:HD23	2.20	0.41
1:A:207:HIS:CD2	1:A:209:PRO:HD2	2.55	0.41
1:C:61:ALA:HB3	1:C:64:PHE:HD2	1.84	0.41
1:C:158:THR:HG23	1:C:206:ASN:HB3	2.03	0.41
2:E:45:ASP:H	2:E:48:THR:HG1	1.62	0.41
2:E:165:ASP:O	2:E:169:SER:N	2.47	0.41
2:E:46:GLU:HG2	2:E:69:TYR:OH	2.20	0.41
1:A:191:VAL:HG11	1:A:201:TYR:CE1	2.56	0.41
3:B:107:GLN:HB2	3:B:139:TYR:CE2	2.56	0.41
1:A:36:TRP:O	1:A:48:MET:HG2	2.22	0.41
1:C:22:CYS:O	1:C:78:THR:HA	2.21	0.41
1:A:153:PHE:HA	1:A:154:PRO:HA	1.80	0.40
2:G:48:THR:HG22	2:G:51:LYS:CB	2.51	0.40
2:E:158:CYS:HB3	2:E:162:PHE:HB2	2.02	0.40
2:G:33:ASN:OD1	2:G:36:THR:N	2.54	0.40
1:C:68:VAL:HG22	1:C:81:MET:HE2	2.04	0.40

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	211/223 (95%)	207 (98%)	4 (2%)	0	100	100
1	C	212/223 (95%)	207 (98%)	5 (2%)	0	100	100
2	E	166/184 (90%)	152 (92%)	13 (8%)	1 (1%)	25	59
2	G	155/184 (84%)	145 (94%)	10 (6%)	0	100	100
3	B	205/213 (96%)	199 (97%)	5 (2%)	1 (0%)	29	64
3	D	206/213 (97%)	199 (97%)	6 (3%)	1 (0%)	29	64
All	All	1155/1240 (93%)	1109 (96%)	43 (4%)	3 (0%)	41	73

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	42	LEU
3	B	93	ASN
3	D	93	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	181/191 (95%)	177 (98%)	4 (2%)	52	78
1	C	183/191 (96%)	178 (97%)	5 (3%)	44	74
2	E	144/167 (86%)	142 (99%)	2 (1%)	67	86
2	G	138/167 (83%)	135 (98%)	3 (2%)	52	78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	B	173/179 (97%)	173 (100%)	0	100 100
3	D	172/179 (96%)	169 (98%)	3 (2%)	60 83
All	All	991/1074 (92%)	974 (98%)	17 (2%)	60 83

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	TRP
1	A	56	ASN
1	A	147	CYS
1	A	203	CYS
2	G	30	VAL
2	G	44	CYS
2	G	48	THR
1	C	50	TRP
1	C	83	LEU
1	C	147	CYS
1	C	188	VAL
1	C	203	CYS
2	E	44	CYS
2	E	92	VAL
3	D	58	ARG
3	D	155	LYS
3	D	161	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
3	D	193	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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	215/223 (96%)	-0.09	3 (1%) 75 56	28, 48, 82, 120	0
1	C	216/223 (96%)	-0.10	1 (0%) 91 81	32, 54, 83, 127	0
2	E	170/184 (92%)	0.33	9 (5%) 26 12	36, 79, 134, 169	0
2	G	161/184 (87%)	0.66	18 (11%) 5 2	45, 95, 148, 173	0
3	B	207/213 (97%)	-0.19	0 100 100	33, 53, 76, 130	0
3	D	208/213 (97%)	-0.24	1 (0%) 91 81	29, 44, 65, 140	0
All	All	1177/1240 (94%)	0.03	32 (2%) 54 31	28, 55, 122, 173	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	74	THR	4.1
2	G	41	VAL	4.1
2	G	44	CYS	3.7
2	E	68	SER	3.4
2	E	70	ALA	3.4
2	E	67	VAL	3.3
2	E	62	ILE	3.2
2	G	1	ALA	3.0
2	G	71	CYS	2.9
2	G	109	THR	2.8
2	G	48	THR	2.7
2	G	83	VAL	2.6
2	E	71	CYS	2.6
2	G	76	GLY	2.6
2	G	141	LEU	2.6
2	E	60	ILE	2.4
1	A	75	SER	2.4
1	A	74	THR	2.3
2	G	106	PRO	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	G	54	GLY	2.3
2	G	27	ASN	2.3
2	G	167	GLU	2.3
2	G	81	ASN	2.3
1	A	73	ASP	2.2
2	G	79	MET	2.2
2	G	80	VAL	2.2
2	G	161	GLY	2.2
2	E	85	ILE	2.2
3	D	1	GLU	2.1
2	G	82	ASN	2.1
2	E	87	ASN	2.1
2	E	78	ASP	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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.