wwPDB X-ray Structure Validation Summary Report (i)

Oct 8, 2023 – 12:24 PM EDT

PDB ID	:	4R96
Title	:	Structure of a Llama Glama Fab 48A2 against human cMet
Authors	:	Klarenbeek, A.; El Mazouari, K.; Desmyter, A.; Blanchetot, C.; Hultberg, A.; Roovers, R.C.; Cambillau, C.; Spinelli, S.; Del-Favero, J.; Verrips, T.; de Haard, H.; Achour, I.
Deposited on	:	2014-09-03
Resolution	:	3.31 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 (i) 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 (1)) were used in the production of this report:

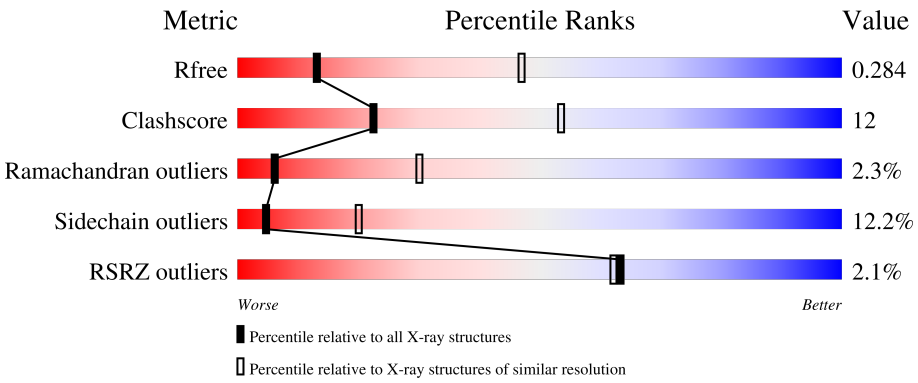
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.31 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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 <=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	221	.% 78% 19% • •
1	C	221	2% 76% 21% • •
1	E	221	4% 77% 19% •
1	L	221	.% 83% 14% •
2	B	219	% 64% 26% 6% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	219	2% 61% 26% 8%
2	F	219	4% 70% 19% 6%
2	H	219	% 64% 27% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13536 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 Llama glama Fab 48A2 against human cMet L chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	219	1703	1068	288	342	5	0	0	0
1	C	219	1703	1068	288	342	5	0	0	0
1	E	221	1714	1074	290	344	6	0	0	0
1	L	221	1718	1076	290	346	6	0	0	0

- Molecule 2 is a protein called Llama glama Fab 48A2 against human cMet H chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	213	1606	1015	263	322	6	0	0	0
2	D	213	1612	1018	266	322	6	0	0	0
2	F	213	1612	1018	266	322	6	0	0	0
2	H	217	1634	1030	270	328	6	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	39	39	39	0	0
3	C	16	16	16	0	0
3	E	28	28	28	0	0
3	L	29	29	29	0	0

Continued on next page...

Continued from previous page...

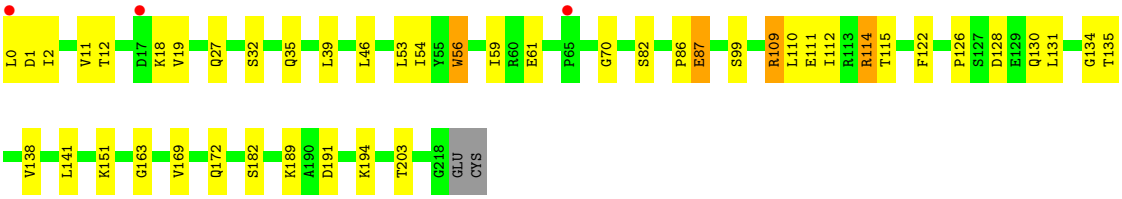
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	40	40	40	0	0
			Total	O		
3	D	25	25	25	0	0
			Total	O		
3	F	35	35	35	0	0
			Total	O		
3	H	22	22	22	0	0

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.

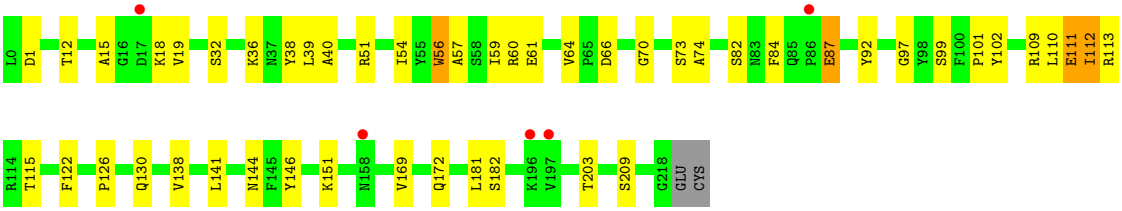
• Molecule 1: Llama glama Fab 48A2 against human cMet L chain

Chain A:



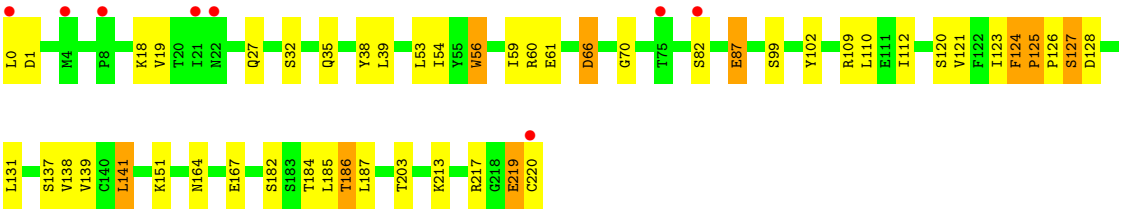
• Molecule 1: Llama glama Fab 48A2 against human cMet L chain

Chain C:



• Molecule 1: Llama glama Fab 48A2 against human cMet L chain

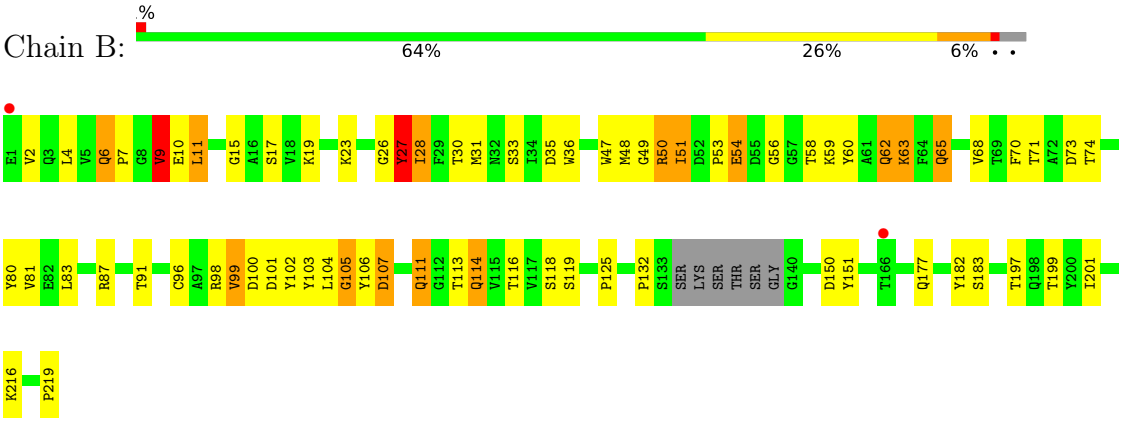
Chain E:



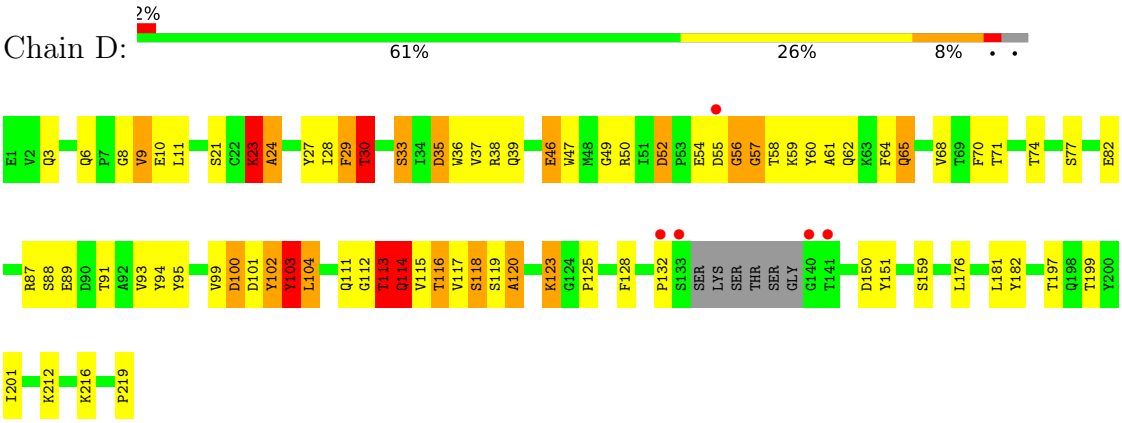
• Molecule 1: Llama glama Fab 48A2 against human cMet L chain

Chain L:

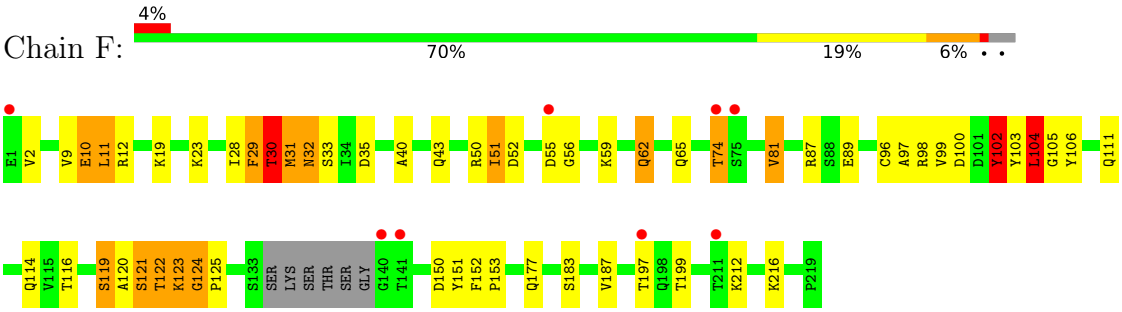
• Molecule 2: Llama glama Fab 48A2 against human cMet H chain



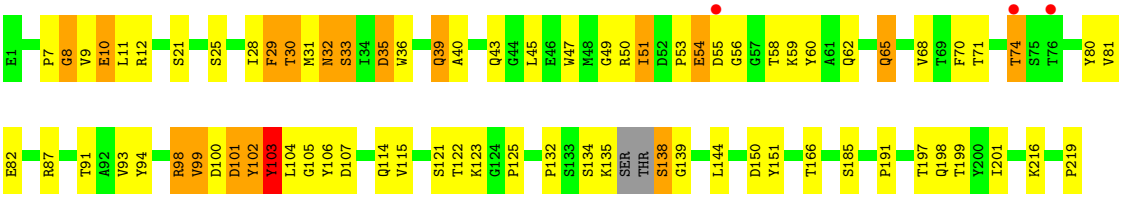
• Molecule 2: Llama glama Fab 48A2 against human cMet H chain



• Molecule 2: Llama glama Fab 48A2 against human cMet H chain



• Molecule 2: Llama glama Fab 48A2 against human cMet H chain


E1
P7
G8
V9
E10
L11
R12
S21
S25
I28
F29
T30
M31
N32
S33
I34
D35
W36
Q39
A40
Q43
G44
L45
E46
W47
M48
G49
R50
I51
D52
P53
E54
D55
G56
G57
T58
K59
Y60
A61
Q62
Q65
V68
T69
F70
T71
T74
S75
T76
Y80
V81
E82
R87
T91
A92
V93
Y94
R98
V99
D100
D101
Y102
Y103
L104
G105
Y106
D107
Q114
V115
S121
T122
K123
G124
P125
P132
S133
S134
K135
SER
THR
S138
G139
L144
D150
Y151
T166
S185
P191
T197
Q198
T199
Y200
I201
K216
P219

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	107.02Å 121.54Å 185.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.39 – 3.31 43.39 – 3.31	Depositor EDS
% Data completeness (in resolution range)	(Not available) (43.39-3.31) 97.6 (43.39-3.31)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.06 (at 3.32Å)	Xtriage
Refinement program	BUSTER 2.9.2	Depositor
R, R_{free}	0.217 , 0.261 0.241 , 0.284	Depositor DCC
R_{free} test set	3580 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å^2)	42.2	Xtriage
Anisotropy	0.891	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 64.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	13536	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% 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 (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	0.46	0/1741	0.68	0/2365
1	C	0.45	0/1741	0.68	0/2365
1	E	0.54	1/1752 (0.1%)	0.74	2/2380 (0.1%)
1	L	0.44	0/1756	0.69	0/2385
2	B	0.57	0/1645	0.85	1/2247 (0.0%)
2	D	0.50	0/1651	0.89	2/2254 (0.1%)
2	F	0.51	0/1651	0.84	2/2254 (0.1%)
2	H	0.50	0/1673	0.82	3/2283 (0.1%)
All	All	0.50	1/13610 (0.0%)	0.77	10/18533 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	125	PRO	N-CD	5.81	1.55	1.47

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	30	THR	C-N-CA	7.14	139.55	121.70
2	D	23	LYS	C-N-CA	7.10	139.45	121.70
1	E	125	PRO	C-N-CD	6.26	141.56	128.40
2	F	124	GLY	C-N-CD	6.24	141.50	128.40
2	H	103	TYR	CA-CB-CG	6.16	125.10	113.40

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	1703	0	1645	33	0
1	C	1703	0	1645	34	0
1	E	1714	0	1652	31	0
1	L	1718	0	1656	19	0
2	B	1606	0	1548	53	0
2	D	1612	0	1559	65	0
2	F	1612	0	1559	46	0
2	H	1634	0	1576	46	0
3	A	39	0	0	1	0
3	B	40	0	0	0	0
3	C	16	0	0	0	0
3	D	25	0	0	1	0
3	E	28	0	0	2	0
3	F	35	0	0	0	0
3	H	22	0	0	0	0
3	L	29	0	0	1	0
All	All	13536	0	12840	303	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 12.

The worst 5 of 303 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:F:30:THR:CB	2:F:31:MET:HB2	1.72	1.18
2:F:123:LYS:NZ	2:F:150:ASP:O	1.78	1.16
2:B:50:ARG:HH11	2:B:50:ARG:HG2	1.14	1.12
2:F:30:THR:HB	2:F:31:MET:CB	1.84	1.06
2:F:100:ASP:HB3	2:F:102:TYR:CE2	1.93	1.02

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	217/221 (98%)	201 (93%)	15 (7%)	1 (0%)	29	61
1	C	217/221 (98%)	205 (94%)	9 (4%)	3 (1%)	11	39
1	E	219/221 (99%)	203 (93%)	15 (7%)	1 (0%)	29	61
1	L	219/221 (99%)	204 (93%)	12 (6%)	3 (1%)	11	39
2	B	209/219 (95%)	190 (91%)	14 (7%)	5 (2%)	6	29
2	D	209/219 (95%)	177 (85%)	20 (10%)	12 (6%)	1	12
2	F	209/219 (95%)	186 (89%)	18 (9%)	5 (2%)	6	29
2	H	213/219 (97%)	185 (87%)	18 (8%)	10 (5%)	2	15
All	All	1712/1760 (97%)	1551 (91%)	121 (7%)	40 (2%)	6	30

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1	ASP
1	E	219	GLU
1	L	219	GLU
2	D	24	ALA
2	D	56	GLY

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	193/195 (99%)	182 (94%)	11 (6%)	20	52
1	C	193/195 (99%)	180 (93%)	13 (7%)	16	46
1	E	194/195 (100%)	171 (88%)	23 (12%)	5	21
1	L	195/195 (100%)	180 (92%)	15 (8%)	13	40
2	B	179/185 (97%)	151 (84%)	28 (16%)	2	12
2	D	180/185 (97%)	144 (80%)	36 (20%)	1	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	180/185 (97%)	153 (85%)	27 (15%)	3	13
2	H	182/185 (98%)	153 (84%)	29 (16%)	2	11
All	All	1496/1520 (98%)	1314 (88%)	182 (12%)	5	20

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	101	ASP
2	F	87	ARG
2	D	114	GLN
2	F	29	PHE
2	F	121	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	114	GLN
2	F	111	GLN
2	F	65	GLN
2	F	114	GLN
1	E	27	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	219/221 (99%)	0.34	3 (1%)	75	75	22, 46, 67, 82	3 (1%)
1	C	219/221 (99%)	0.39	5 (2%)	60	59	21, 42, 69, 85	3 (1%)
1	E	221/221 (100%)	0.43	8 (3%)	42	41	19, 46, 73, 125	3 (1%)
1	L	221/221 (100%)	0.12	3 (1%)	75	75	15, 38, 59, 120	3 (1%)
2	B	213/219 (97%)	0.19	2 (0%)	84	85	21, 44, 65, 79	0
2	D	213/219 (97%)	0.33	5 (2%)	60	59	21, 45, 68, 85	0
2	F	213/219 (97%)	0.44	8 (3%)	40	39	23, 48, 70, 84	0
2	H	217/219 (99%)	0.10	3 (1%)	75	75	17, 42, 66, 75	0
All	All	1736/1760 (98%)	0.29	37 (2%)	63	62	15, 44, 67, 125	12 (0%)

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	220	CYS	5.7
1	L	220	CYS	4.3
1	E	0	LEU	3.7
1	C	158	ASN	3.5
2	D	140	GLY	3.4

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.