



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:37 AM GMT

PDB ID : 2B2X
Title : VLA1 RdeltaH I-domain complexed with a quadruple mutant of the AQC2 Fab
Authors : Clark, L.A.; Boriack-Sjodin, P.A.; Eldredge, J.; Fitch, C.; Friedman, B.; Hanf, K.J.; Jarpe, M.; Liparoto, S.F.; Li, Y.; Lugovskoy, A.
Deposited on : 2005-09-19
Resolution : 2.20 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

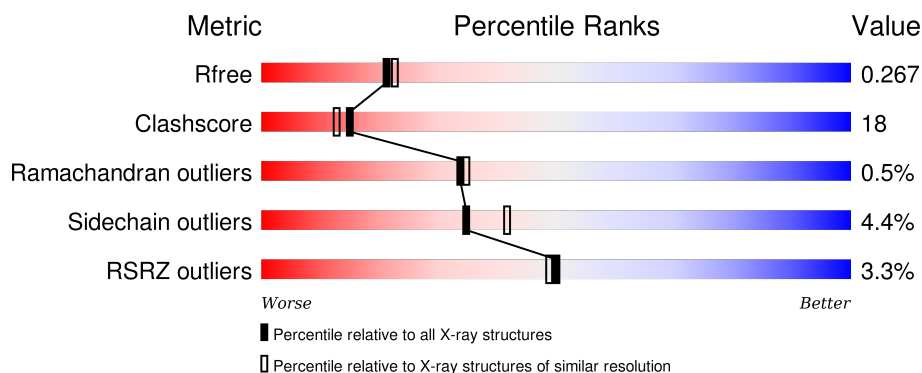
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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 $\leq 5\%$. 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	223	<div> <div>5%</div> <div> <div>48%</div> <div>33%</div> <div>•</div> <div>16%</div> </div> </div>
1	B	223	<div> <div>8%</div> <div> <div>48%</div> <div>29%</div> <div>•</div> <div>21%</div> </div> </div>
2	H	226	<div> <div>72%</div> <div>21%</div> <div>7%</div> </div>
2	I	226	<div> <div>72%</div> <div>18%</div> <div>•</div> <div>7%</div> </div>
3	L	213	<div> <div>2%</div> <div> <div>67%</div> <div>28%</div> <div>• •</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	M	213	 A horizontal bar chart showing the quality of chain M. The bar is divided into four segments: a small red segment at the beginning labeled '3%', a large green segment labeled '66%', a yellow segment labeled '31%', and a very small grey segment at the end labeled '..'. The segments are separated by thin white lines.

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9537 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 Integrin alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	188	Total	C	N	O	S	0	0	0
			1490	940	258	288	4			
1	B	176	Total	C	N	O	S	0	0	0
			1397	884	240	270	3			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	125	GLY	-	CLONING ARTIFACT	UNP P18614
A	126	SER	-	CLONING ARTIFACT	UNP P18614
A	217	VAL	GLY	ENGINEERED	UNP P18614
A	218	GLN	ARG	ENGINEERED	UNP P18614
A	219	ARG	GLN	ENGINEERED	UNP P18614
A	222	ARG	LEU	ENGINEERED	UNP P18614
A	341	LEU	-	CLONING ARTIFACT	UNP P18614
A	342	GLU	-	CLONING ARTIFACT	UNP P18614
A	343	ARG	-	CLONING ARTIFACT	UNP P18614
A	344	PRO	-	CLONING ARTIFACT	UNP P18614
A	345	HIS	-	CLONING ARTIFACT	UNP P18614
A	346	ARG	-	CLONING ARTIFACT	UNP P18614
A	347	ASP	-	CLONING ARTIFACT	UNP P18614
B	125	GLY	-	CLONING ARTIFACT	UNP P18614
B	126	SER	-	CLONING ARTIFACT	UNP P18614
B	217	VAL	GLY	ENGINEERED	UNP P18614
B	218	GLN	ARG	ENGINEERED	UNP P18614
B	219	ARG	GLN	ENGINEERED	UNP P18614
B	222	ARG	LEU	ENGINEERED	UNP P18614
B	341	LEU	-	CLONING ARTIFACT	UNP P18614
B	342	GLU	-	CLONING ARTIFACT	UNP P18614
B	343	ARG	-	CLONING ARTIFACT	UNP P18614
B	344	PRO	-	CLONING ARTIFACT	UNP P18614
B	345	HIS	-	CLONING ARTIFACT	UNP P18614
B	346	ARG	-	CLONING ARTIFACT	UNP P18614

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	347	ASP	-	CLONING ARTIFACT	UNP P18614

- Molecule 2 is a protein called Antibody AQC2 Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	210	Total	C	N	O	S	0	0	0
			1575	1000	260	308	7			
2	I	210	Total	C	N	O	S	0	0	0
			1575	1000	260	308	7			

- Molecule 3 is a protein called Antibody AQC2 Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	210	Total	C	N	O	S	0	0	0
			1636	1026	274	330	6			
3	M	210	Total	C	N	O	S	0	0	0
			1636	1026	274	330	6			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		
4	A	1	Total	Mg	0	0
			1	1		

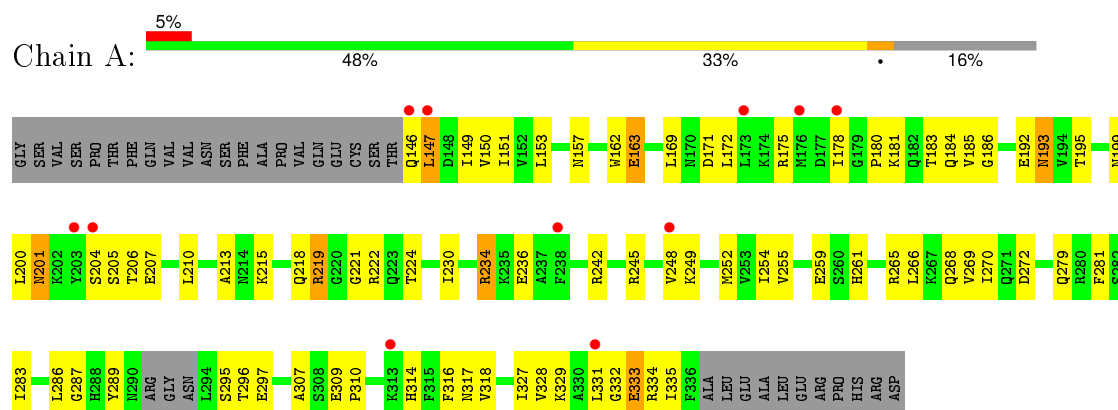
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	13	Total	O	0	0
			13	13		
5	B	5	Total	O	0	0
			5	5		
5	H	73	Total	O	0	0
			73	73		
5	I	77	Total	O	0	0
			77	77		
5	L	30	Total	O	0	0
			30	30		
5	M	28	Total	O	0	0
			28	28		

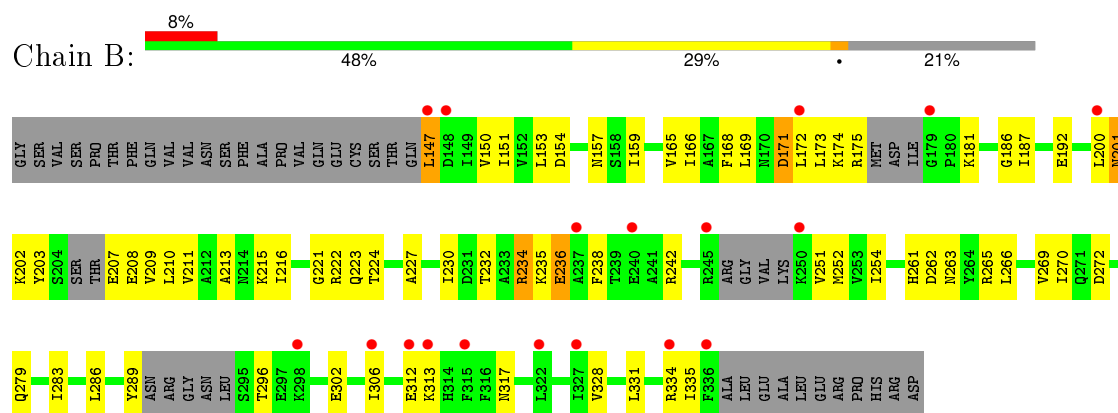
3 Residue-property plots [i](#)

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.

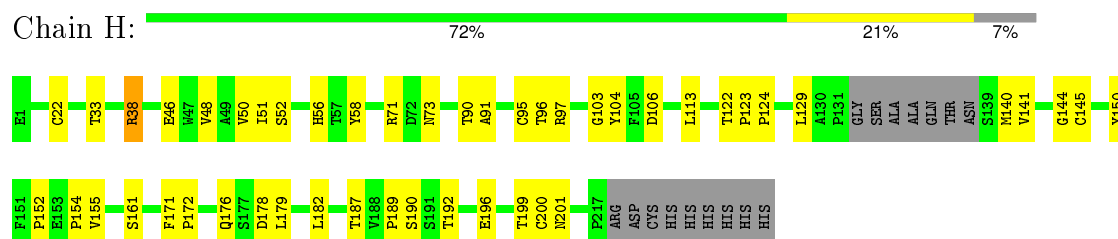
• Molecule 1: Integrin alpha-1



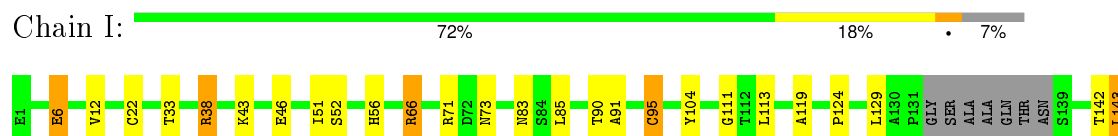
• Molecule 1: Integrin alpha-1



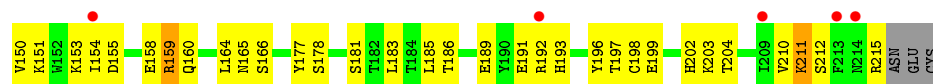
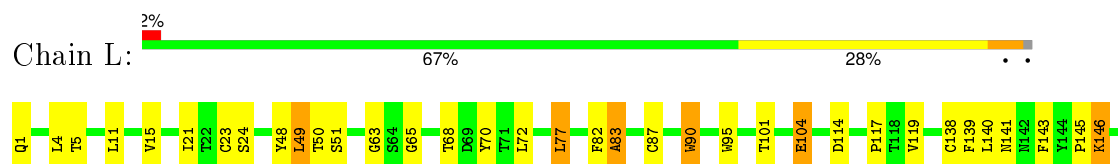
• Molecule 2: Antibody AQC2 Fab



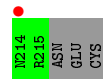
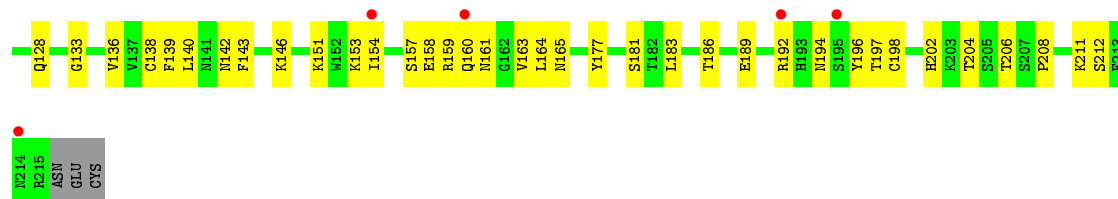
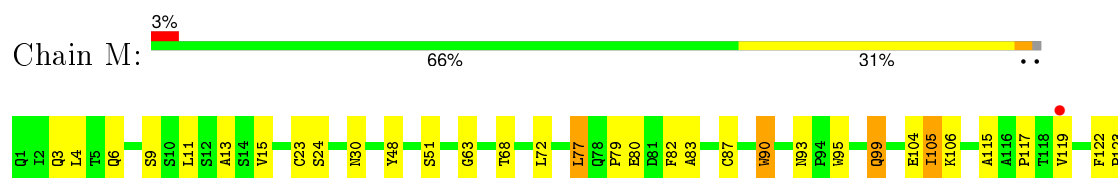
• Molecule 2: Antibody AQC2 Fab



• Molecule 3: Antibody AQC2 Fab



• Molecule 3: Antibody AQC2 Fab



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	106.12Å 43.68Å 153.88Å 90.00° 104.10° 90.00°	Depositor
Resolution (Å)	35.00 – 2.20 49.75 – 2.20	Depositor EDS
% Data completeness (in resolution range)	94.3 (35.00-2.20) 94.2 (49.75-2.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.04 (at 2.20Å)	Xtriage
Refinement program	CNX	Depositor
R, R_{free}	0.238 , 0.272 0.235 , 0.267	Depositor DCC
R_{free} test set	3380 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	38.6	Xtriage
Anisotropy	0.335	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 43.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	1 of 66552 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9537	wwPDB-VP
Average B, all atoms (Å ²)	46.0	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 36.80 % 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 4.7301e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

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	0.28	0/1510	0.56	0/2036
1	B	0.27	0/1414	0.55	0/1903
2	H	0.39	0/1616	0.71	0/2208
2	I	0.37	0/1616	0.70	0/2208
3	L	0.35	0/1680	0.63	0/2288
3	M	0.33	0/1680	0.61	0/2288
All	All	0.34	0/9516	0.63	0/12931

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	1490	0	1497	67	0
1	B	1397	0	1395	75	0
2	H	1575	0	1537	39	0
2	I	1575	0	1537	38	0
3	L	1636	0	1561	53	0
3	M	1636	0	1561	56	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
5	A	13	0	0	0	0
5	B	5	0	0	0	0
5	H	73	0	0	0	0
5	I	77	0	0	1	0
5	L	30	0	0	1	0
5	M	28	0	0	0	0
All	All	9537	0	9088	315	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 18.

The worst 5 of 315 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:L:191:GLU:HA	3:L:215:ARG:HH12	1.25	0.98
1:B:234:ARG:HH11	1:B:234:ARG:HB3	1.27	0.97
2:H:161:SER:H	2:H:201:ASN:HD21	1.00	0.95
2:I:161:SER:H	2:I:201:ASN:HD21	1.08	0.94
3:M:6:GLN:H	3:M:99:GLN:NE2	1.64	0.94

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	184/223 (82%)	164 (89%)	18 (10%)	2 (1%)	17	14
1	B	166/223 (74%)	147 (89%)	17 (10%)	2 (1%)	16	12
2	H	206/226 (91%)	201 (98%)	5 (2%)	0	100	100
2	I	206/226 (91%)	202 (98%)	4 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	L	208/213 (98%)	198 (95%)	9 (4%)	1 (0%)	34	35
3	M	208/213 (98%)	195 (94%)	12 (6%)	1 (0%)	34	35
All	All	1178/1324 (89%)	1107 (94%)	65 (6%)	6 (0%)	34	35

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	287	GLY
1	B	236	GLU
1	A	333	GLU
3	L	83	ALA
1	B	173	LEU

5.3.2 Protein sidechains ⓘ

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	164/194 (84%)	157 (96%)	7 (4%)	35	43
1	B	153/194 (79%)	148 (97%)	5 (3%)	45	56
2	H	177/190 (93%)	170 (96%)	7 (4%)	38	47
2	I	177/190 (93%)	166 (94%)	11 (6%)	23	25
3	L	187/190 (98%)	179 (96%)	8 (4%)	35	43
3	M	187/190 (98%)	179 (96%)	8 (4%)	35	43
All	All	1045/1148 (91%)	999 (96%)	46 (4%)	35	42

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	159	ARG
1	B	234	ARG
3	M	90	TRP
3	L	211	LYS
1	B	171	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	3	GLN
3	L	165	ASN
3	M	99	GLN
3	L	93	ASN
3	L	141	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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 ⓘ

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	188/223 (84%)	0.46	11 (5%) 26 25	28, 59, 77, 83	0
1	B	176/223 (78%)	0.73	18 (10%) 9 8	33, 66, 91, 96	0
2	H	210/226 (92%)	0.12	0 100 100	22, 31, 49, 58	0
2	I	210/226 (92%)	0.07	0 100 100	24, 34, 50, 61	0
3	L	210/213 (98%)	0.23	5 (2%) 62 61	22, 42, 74, 86	0
3	M	210/213 (98%)	0.22	6 (2%) 55 54	24, 41, 68, 86	0
All	All	1204/1324 (90%)	0.29	40 (3%) 50 49	22, 42, 78, 96	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	200	LEU	4.5
1	B	179	GLY	3.8
1	B	147	LEU	3.6
3	M	192	ARG	3.5
3	M	214	ASN	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 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(\AA^2)	Q<0.9
4	MG	B	401	1/1	0.88	0.16	1.63	54,54,54,54	0
4	MG	A	400	1/1	0.92	0.14	0.51	47,47,47,47	0

6.5 Other polymers [i](#)

There are no such residues in this entry.