wwPDB X-ray Structure Validation Summary Report (i)

Jan 6, 2021 – 09:03 pm GMT

PDB ID	:	6XV8
Title	:	Crystal structure of Megabody Mb-Nb207-c7HopQ_G10
Authors	:	Steyaert, J.; Uchanski, T.; Fischer, B.
Deposited on	:	2020-01-21
Resolution	:	3.15 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 following versions of software and data (see references (1)) were used in the production of this report:

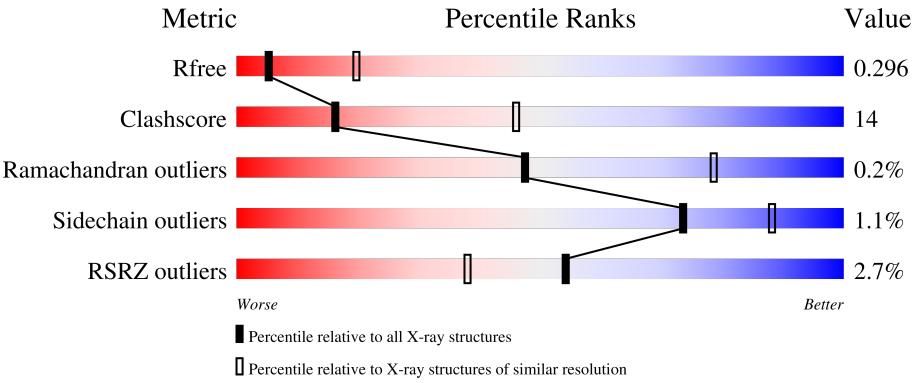
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.16
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

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 3.15 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	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)

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	521	2% 58% 21% • 21%
1	B	521	2% 57% 25% 18%
1	C	521	2% 60% 19% 21%
1	D	521	2% 51% 22% • 26%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12431 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 Outer membrane protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	412	Total	C	N	O	S	0	0	0
			3134	1942	552	627	13			
1	B	429	Total	C	N	O	S	0	0	0
			3251	2012	572	654	13			
1	C	412	Total	C	N	O	S	0	0	0
			3112	1924	549	626	13			
1	D	387	Total	C	N	O	S	0	0	0
			2934	1820	514	588	12			

There are 8 discrepancies between the modelled and reference sequences:

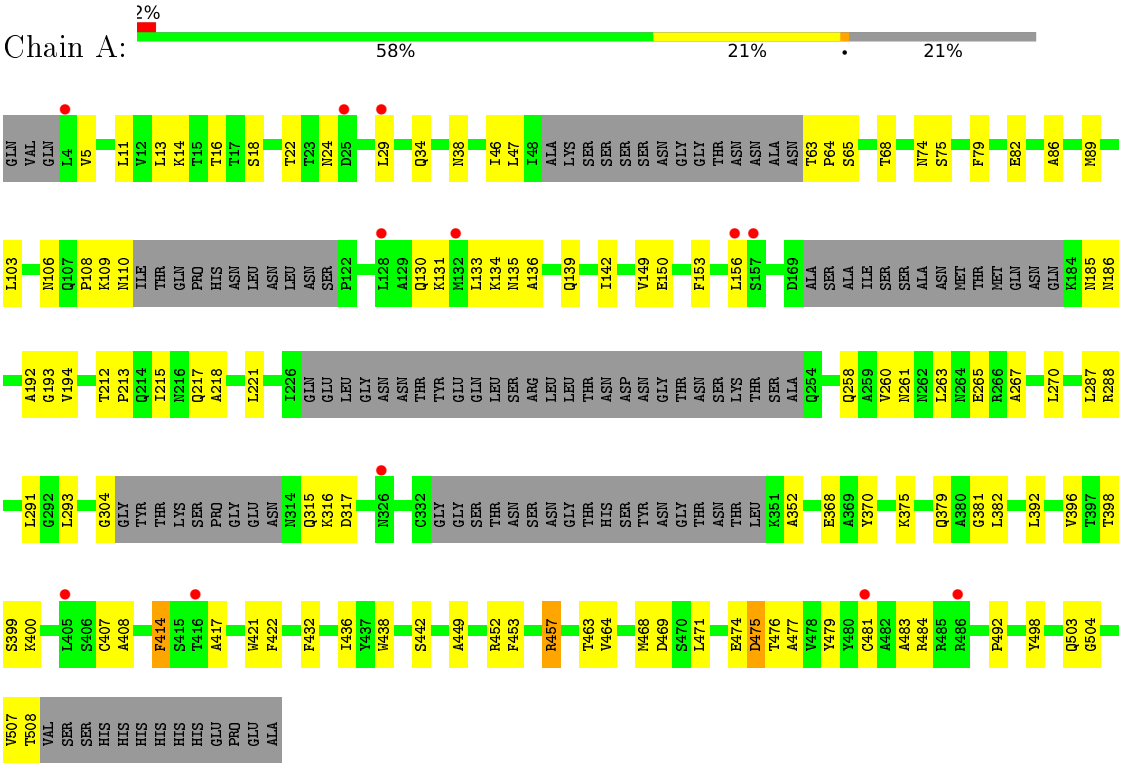
Chain	Residue	Modelled	Actual	Comment	Reference
A	13	LEU	-	linker	UNP B5Z8H1
A	401	SER	-	linker	UNP B5Z8H1
B	13	LEU	-	linker	UNP B5Z8H1
B	401	SER	-	linker	UNP B5Z8H1
C	13	LEU	-	linker	UNP B5Z8H1
C	401	SER	-	linker	UNP B5Z8H1
D	13	LEU	-	linker	UNP B5Z8H1
D	401	SER	-	linker	UNP B5Z8H1

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.

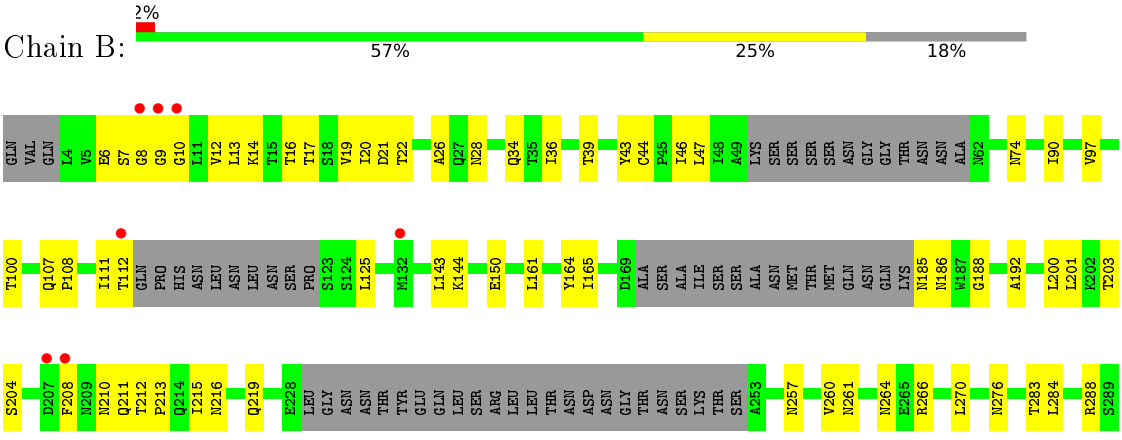
- Molecule 1: Outer membrane protein

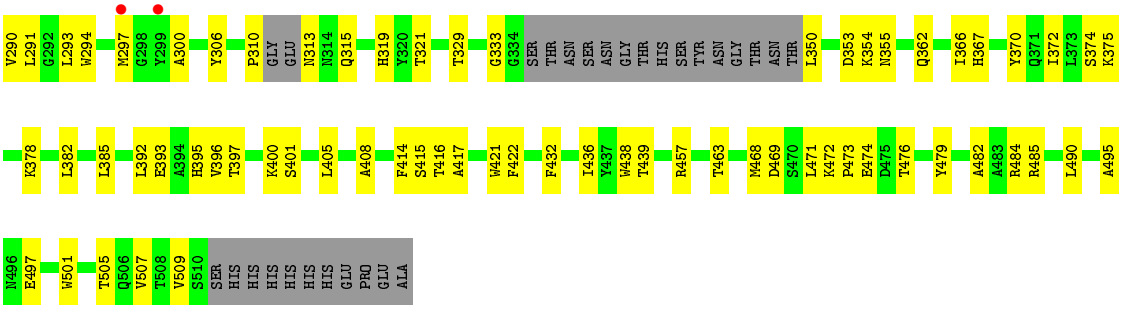
Chain A:



- Molecule 1: Outer membrane protein

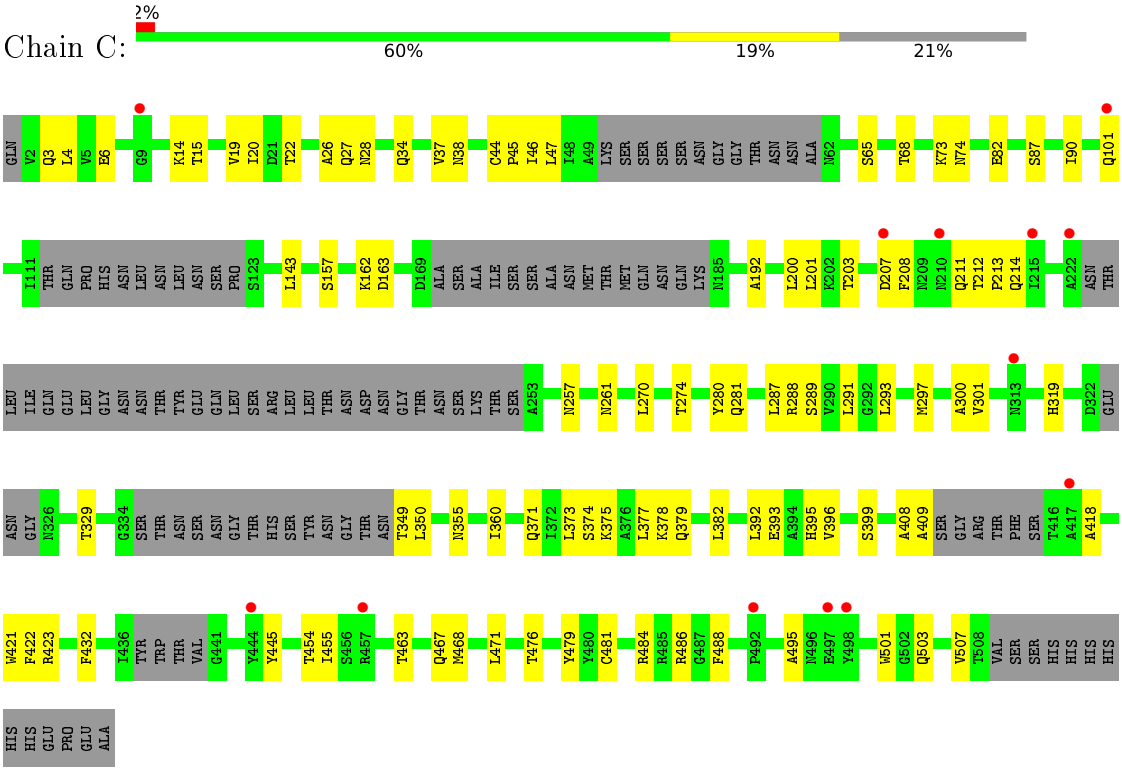
Chain B:





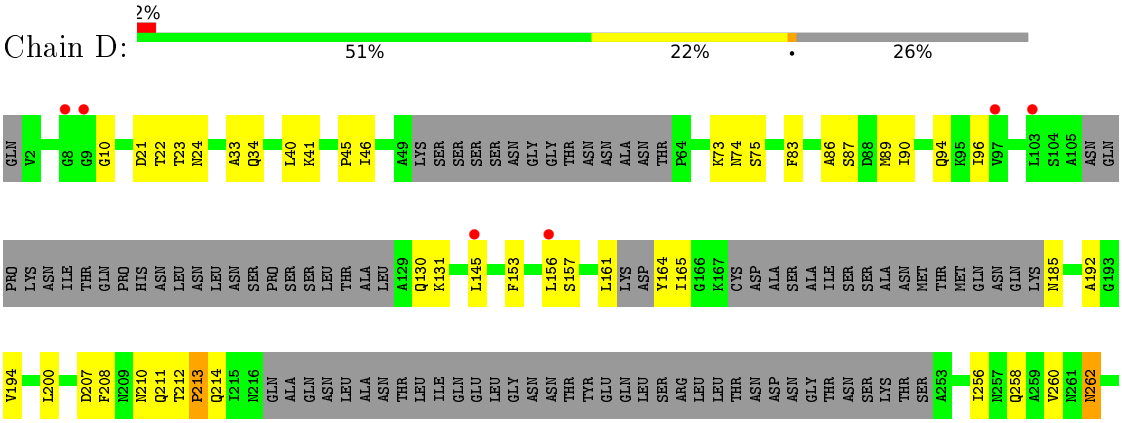
• Molecule 1: Outer membrane protein

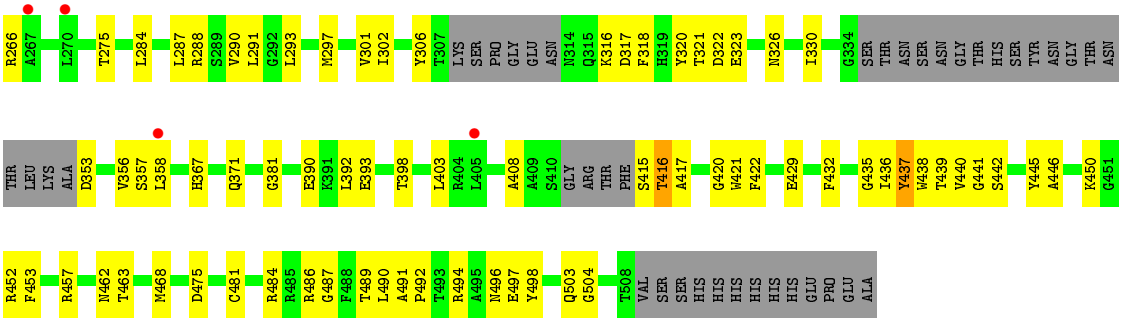
Chain C:



• Molecule 1: Outer membrane protein

Chain D:




R266 A267 L270 T275 L284 L287 R288 S289 V290 L291 G292 L293 M297 V301 I302 Y306 T307 LYS SER PRO GLY GLU ASN N314 Q315 K316 D317 F318 H319 Y320 T321 D322 E323 N326 I330 G334 SER THR ASN SER ASN GLY THR HIS SER TYR ASN GLY THR ASN
THR LEU LYS ALA D353 V356 S357 L358 H367 Q371 G381 E390 K391 L392 E393 T398 L403 R404 L405 A408 A409 S410 GLY ARG THR PHE S415 T416 A417 G420 W421 F422 E429 F432 G435 I436 Y437 W438 T439 V440 G441 S442 Y445 A446 K450 G451
R452 F453 R457 N462 T463 M468 D475 C481 R484 R485 R486 G487 F488 T489 L490 A491 P492 T493 R494 A495 N496 E497 Y498 Q503 G504 T508 VAL SER SER HIS HIS HIS HIS HIS HIS GLU PRO GLU ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	79.24Å 155.05Å 89.75Å 90.00° 91.73° 90.00°	Depositor
Resolution (Å)	79.21 – 3.15 79.21 – 3.15	Depositor EDS
% Data completeness (in resolution range)	96.0 (79.21-3.15) 96.1 (79.21-3.15)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.74 (at 3.13Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.255 , 0.297 0.255 , 0.296	Depositor DCC
R_{free} test set	1686 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å^2)	101.5	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 48.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.001 for h,-k,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	12431	wwPDB-VP
Average B, all atoms (Å^2)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% 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

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	0.45	0/3178	0.66	0/4298
1	B	0.51	0/3296	0.75	0/4460
1	C	0.54	0/3151	0.75	0/4258
1	D	0.43	0/2974	0.69	0/4019
All	All	0.48	0/12599	0.71	0/17035

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	3134	0	3069	80	0
1	B	3251	0	3177	116	0
1	C	3112	0	3046	73	0
1	D	2934	0	2856	96	0
All	All	12431	0	12148	354	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 14.

The worst 5 of 354 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)
1:B:211:GLN:HG3	1:B:266:ARG:NH2	1.68	1.07
1:D:256:ILE:O	1:D:260:VAL:HG23	1.62	0.98
1:B:108:PRO:HD2	1:B:264:ASN:OD1	1.75	0.87
1:B:7:SER:HB2	1:C:101:GLN:HE22	1.44	0.82
1:A:14:LYS:HG2	1:A:400:LYS:HB2	1.62	0.82

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	398/521 (76%)	378 (95%)	20 (5%)	0	100 100
1	B	415/521 (80%)	390 (94%)	24 (6%)	1 (0%)	47 78
1	C	394/521 (76%)	374 (95%)	20 (5%)	0	100 100
1	D	369/521 (71%)	350 (95%)	17 (5%)	2 (0%)	29 65
All	All	1576/2084 (76%)	1492 (95%)	81 (5%)	3 (0%)	47 78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	213	PRO
1	B	10	GLY
1	D	356	VAL

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	335/427 (78%)	331 (99%)	4 (1%)	71 87
1	B	347/427 (81%)	342 (99%)	5 (1%)	67 85
1	C	331/427 (78%)	329 (99%)	2 (1%)	86 94
1	D	310/427 (73%)	307 (99%)	3 (1%)	76 89
All	All	1323/1708 (78%)	1309 (99%)	14 (1%)	73 88

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	17	THR
1	B	293	LEU
1	D	262	ASN
1	B	16	THR
1	C	297	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	313	ASN
1	C	28	ASN
1	D	262	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

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	412/521 (79%)	0.12	12 (2%) 51 35	66, 123, 176, 241	0
1	B	429/521 (82%)	0.07	9 (2%) 63 49	57, 103, 144, 163	0
1	C	412/521 (79%)	0.16	13 (3%) 47 30	49, 99, 153, 182	0
1	D	387/521 (74%)	0.09	10 (2%) 56 40	69, 129, 181, 206	0
All	All	1640/2084 (78%)	0.11	44 (2%) 54 38	49, 112, 168, 241	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	10	GLY	4.7
1	C	9	GLY	4.5
1	D	145	LEU	4.3
1	D	9	GLY	4.2
1	D	8	GLY	3.6

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.